EXISTENCE THEORY FOR NON-SEPARABLE MEAN FIELD GAMES IN SOBOLEV SPACES

DAVID M. AMBROSE

Abstract. The mean field games system is a coupled pair of nonlinear partial differential equations arising in differential game theory, as a limit as the number of agents tends to infinity. We prove existence and uniqueness theorems for time-dependent mean field games with Sobolev data. Many works in the literature assume additive separability of the Hamiltonian, as well as further structure such as convexity and monotonicity of the resulting components. Problems arising in practice, however, may not have this separable structure; we therefore consider the non-separable problem. For our existence and uniqueness results, we introduce new smallness constraints which simultaneously consider the size of the time horizon, the size of the data, and the strength of the coupling in the system.

1. Introduction

Mean field games have been introduced in the mathematics literature by Lasry and Lions as limits of problems from game theory, as the number of agents tends to infinity [13], [14], [15]. From a control theory perspective, mean field games were also introduced around the same time by Huang, Caines, and Malhame [11], [12]. The mean field games system of partial differential equations is the following coupled system for a value function, u, and a probability measure, m :

$$u_t + \Delta u + \mathcal{H}(t, x, m, Du) = 0, \tag{1}$$

$$m_t - \Delta m + \mathrm{div}\,(m\mathcal{H}_p(t, x, m, Du)) = 0, \tag{2}$$

with $x \in \mathbb{T}^d$ and $t \in [0, T]$, for some given $T > 0$. These equations are supplemented with boundary conditions, and these can be one of two types. The planning problem has as its boundary conditions

$$m(0, x) = m_0(x), \qquad u(T, x) = u_T(x). \tag{3}$$

An alternative is the payoff problem, which uses the following:

$$m(0, x) = m_0(x), \qquad u(T, x) = G(x, m(T, x)), \tag{4}$$

where G is known as the payoff function.

In the paper [2], the author initiated a program of proving existence results for strong solutions for time-dependent mean field games. Ideas from the work of Duchon and Robert on vortex sheets in incompressible flow [6] were used; in particular, Duchon and Robert developed a Duhamel formula for the vortex sheet which integrated both forward and backward in time, and found a contraction in function spaces based on the Wiener algebra, proving the existence of small spatially analytic solutions. The ideas of Duchon and Robert have been extended to finite time horizons and the spatially periodic setting by Milgrom and the author [17]. All of these features are thus also characteristics of the author's work [2] on mean field games. This was extended somewhat in [3], in which non-separable Hamiltonians were treated, and a result in the case of weak coupling, making use of the implicit function theorem, was also given.

Other authors have proved existence theorems for mean field games, focusing much attention on the case of separable mean field games. The assumption of separability is that the Hamiltonian, $\mathcal{H}$, separates additively as $\mathcal{H}(t,x,m,Du) = H(t,x,Du) + F(t,x,m)$. This H is then also known as the Hamiltonian, and the function F is known as the coupling (for if one were to take $F=0$, then the system decouples). The separability assumption, as well as further structural assumptions such as convexity of H and monotonicity of F, allow certain mathematical methods to be brought to bear on the problems (i.e., use of convex optimization and montonoicity methods, as well as techniques of optimal transportation). Poretta has proved in the separable case, using such techniques, the existence of weak solutions [19], [20], [21]. Results in this vein for strong solutions are by Gomes, Pimentel, and Sanchez-Morgado in the case of superquadratic and subquadratic Hamiltonians [8], [9], and by Gomes and Pimentel for the case of logarithmic coupling [7].

Although the separable case does have a number of sophisticated mathematical techniques available for existence theory, unfortunately, problems actually arising from game theory and economics do not tend to have this separable structure [18]. Therefore a study of existence theory not relying on this structure is essential. As mentioned above, the author has made one such study previously [3], and this work contained two different results, one with a smallness condition on the data, and the other considering a small parameter in front of the Hamiltonian. The author is only aware of one other existence proof for time-dependent non-separable mean field games in the literature, for a particular form of Hamiltonian related to modeling problems with congestion [10]; in this work, Gomes and Voskanyan made a smallness assumption on T, the length of the time horizon, and still do make structural assumptions such as monotonicity of part of the Hamiltonian. In the present work, we introduce a unified smallness condition which considers at once the size of the time horizon, the coupling parameter introduced by the author in [3], and in some cases, the size of the data. In addition to unifying the smallness constraints, a benefit of the present work as compared to [3] is the setting of more customary Sobolev spaces as opposed to the spaces based on the Wiener algebra used previously.

We also prove a uniqueness theorem, and as in the case of our existence theorem, a smallness condition must be satisfied. This smallness condition again considers at once a parameter which we describe as measuring the coupling in the system (unrelated to the concept of coupling in the separable case), the size of the time horizon, and in some cases, the size of the initial data. Such a smallness constraint is perhaps not just a feature of our proof, but may be more fundamental: Bardi and Fischer have recently given an example in mean field games of non-unique solutions, in the case of large time horizon [5]. While their setting may not be exactly the same as ours (we study the problem on the torus, and their construction uses the domain as the real line in a fundamental way), it is strongly suggestive that constraints such as those we impose are not in general avoidable. Bardi and Cirant have a related uniqueness theorem, for separable mean field games with some smallness constraints [4].

The plan of the paper is as follows. Immediately below, in Section 1.1, we give some elementary definitions and results on Sobolev spaces. In Section 2, we reformulate the problem slightly and introduce an approximating sequence for solutions. In Section 3, we prove our first main theorem (stated at the end of the section as Theorem 4), that under our smallness assumption, the approximating sequence converges to a solution of the mean field games system. We next treat uniqueness of solutions in Section 4, stating our second main theorem, Theorem 5, at the end of the section. We close with some discussion in Section 5.

1.1. **Function spaces and preliminaries.** We will make repeated use of Young's Inequality: for any $a\geq 0$, for any $b\geq 0$, and for any $\sigma>0$, we have

$$ab\leq\frac{a^2}{2\sigma}+\frac{\sigma b^2}{2}. \tag{5}$$

To be very definite, we say that we let $\mathbb{N}=\{0,1,2,\ldots\}$ be the natural numbers, including zero.

We now define our function spaces and norms. The d-dimensional torus is the set $[0,2\pi]^d$ with periodic boundary conditions. The Fourier transform of a function f may be denoted either as $\mathcal{F}f(k)$ or $\hat{f}(k)$, with Fourier variable $k\in\mathbb{Z}^d$. Of course, the Sobolev space $H^0(\mathbb{T}^d)$ is equal to $L^2(\mathbb{T}^d)$, with the same norm. We need multi-index notation for derivatives with respect to the x variables. We will use $\alpha\in\mathbb{N}^d$ for this purpose. Thus, given such an α, we will have $\partial^\alpha=\partial_{x_1}^{\alpha_1}\cdots\partial_{x_d}^{\alpha_d}$. The order of α is $|\alpha|=\displaystyle\sum_{\ell=1}^{d}\alpha_\ell$. For $s\in\mathbb{N}$, with $s>0$, the Sobolev space of order s is the set of functions

$$H^s(\mathbb{T}^d)=\left\{f\in L^2(\mathbb{T}^d):\|f\|_s<\infty\right\},$$

where the norm is defined by

$$\|f\|_s^2 = \sum_{0\leq|\alpha|\leq s} \|\partial^\alpha f\|_0^2.$$

Here, as is usual, the notation $\|\cdot\|_0 = \|\cdot\|_{L^2}$ is used. This definition is equivalent to any other usual definition of Sobolev spaces with index in the natural numbers. We need an elementary interpolation lemma, which we now state.

Lemma 1. *Let m and s be real numbers such that $0 < m < s$. There exists $c > 0$ such that for all $f \in H^s$,*

$$\|f\|_m \leq c\|f\|_s^{m/s}\|f\|_0^{1-m/s}.$$

We do not include a proof of Lemma 1; the proof can be found many places, one of which is [1]. We also need an elementary lemma about products in Sobolev spaces; this is part of Lemma 3.4 of [16], and the proof can be found there.

Lemma 2. *Let $m \in \mathbb{N}$. There exists $c > 0$ such that for all $f \in L^\infty \cap H^m$ and for all $g \in L^\infty \cap H^m$,*

$$\sum_{0\leq|\alpha|\leq m} \|\partial^\alpha(fg) - f\partial^\alpha g\|_{L^2} \leq c\left(|Df|_\infty\|D^{m-1}g\|_{L^2} + \|D^m f\|_{L^2}|g|_\infty\right).$$

2. Formulation and The Approximating Sequence

As we have said, we will track three effects in our existence theorem: the size of the time horizon (i.e., the magnitude of the value T), the size of the initial data, and the strength of the coupling between the two equations of the mean field games system. We now explain what we mean by this third item. We introduce a slight modification of the system (1), (2):

$$u_t + \Delta u + \varepsilon\mathcal{H}(t,x,m,Du) = 0,$$

$$m_t - \Delta m + \varepsilon\mathrm{div}(m\mathcal{H}_p(t,x,m,Du)) = 0,$$

for some $\varepsilon \in \mathbb{R}$. Obviously if $\varepsilon = 1$, this is exactly the system (1), (2). If instead $\varepsilon = 0$, then the system decouples – one could solve the linear heat equation for m and then the other linear heat equation for u. We may call the case of small values of ε the case of weak coupling of the system, and one of the theorems of [3] was in the case of weak coupling. We will perform our existence theory for the system with ε included as a parameter.

We let $\bar{m} = \dfrac{1}{\mathrm{vol}(\mathbb{T}^d)}$, which is the average value of m (since m is a probability distribution). It is convenient to introduce $\mu = m - \bar{m}$. We also subtract the mean from u, since inspection of the right-hand sides of the evolution equations indicates that the mean of u does not influence the evolution. We introduce a projection operator, P, which removes the mean of a periodic function (so, we could have said before that $\mu = Pm$), and denote $w = Pu$. Note that then $Dw = Du$.

We now introduce the system of equations satisfied by (w, μ), first giving notation for the Hamiltonian in terms of (w, μ) :

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

We then have the (w, μ) system:

$$w_t + \Delta w + \varepsilon P\Theta(t, x, \mu, Dw) = 0, \tag{6}$$

$$\mu_t - \Delta\mu + \varepsilon \text{div}(\mu\Theta_p(t, x, \mu, Dw)) + \varepsilon\bar{m}\text{div}(\Theta_p(t, x, \mu, Dw)) = 0. \tag{7}$$

Of course, we have initial data for μ :

$$\mu(0, x) = \mu_0(x) := m(0, x) - \bar{m}.$$

We will discuss the data for w soon below.

Remark 1. *We will be proving an existence theorem with a smallness condition. The reason for subtracting $\bar{m}$ from m to form μ and for replacing u with w is to clarify this smallness condition. Taking m arbitrarily small is not compatible with the fact that m should be a probability measure. Furthermore, since the mean of u is not relevant for the right-hand sides of the equation, requiring the mean of u to be small would be artificial. Instead, the smallness condition will include information about the initial size of μ, and about the size of the data for w; as far as m goes, then, we will be measuring how far m is from a uniform distribution.*

It is convenient to introduce a regularization operator, which will be useful as we construct solutions. Let $\delta > 0$ be given. We let $\mathbb{P}_\delta$ be the operator which projects onto Fourier modes with wavenumber at most $1/\delta$:

$$\mathbb{P}_\delta \mathcal{F} f(k) = \begin{cases} \mathcal{F}f(k), & |k| \leq 1/\delta, \\ 0, & |k| > 1/\delta. \end{cases}$$

We may use the convention $\mathbb{P}_0 = I$, where this signifies the identity operator.

We then set up an iterative approximation scheme, which will depend slightly on the choice of boundary conditions. In either case, we initialize in the same way, and we solve for μ in the same way. Define $\mu^0 = 0$ and $w^0 = 0$. Given (w^n, μ^n), we define μ^{n+1} to be the unique solution of the initial value problem for the following forced, linear heat equation:

$$\mu_t^{n+1} - \Delta\mu^{n+1} + \varepsilon \text{div}(\mu^n\Theta_p(t, x, \mu^n, Dw^n)) + \varepsilon\bar{m}\text{div}(\Theta_p(t, x, \mu^n, Dw^n)) = 0, \tag{8}$$

$$\mu^{n+1}(0, x) = \mathbb{P}_\delta\mu_0(x). \tag{9}$$

The problem which w^{n+1} solves depends on the choice of boundary conditions; consider first the case of the planning problem (3). Then, we define w^{n+1} to be the unique solution of the following forced, linear heat equation:

$$w_t^{n+1} + \Delta w^{n+1} + \varepsilon P\Theta(t, x, \mu^n, Dw^n) = 0, \tag{10}$$

$$w^{n+1}(T, x) = \mathbb{P}_\delta w_T(x), \tag{11}$$

where $w_T = Pu_T$. We may be completely explicit as to what these solutions (w^{n+1}, μ^{n+1}) are, by using Duhamel's formula; for w we have

$$w^{n+1}(t,\cdot) = e^{\Delta(T-t)}\mathbb{P}_\delta w_T - \varepsilon P \int_t^T e^{\Delta(s-t)}\Theta(s,\cdot,\mu^n(s,\cdot), Dw^n(s,\cdot))\ ds. \tag{12}$$

For μ, we instead integrate forward in time, finding

$$\begin{aligned}\mu^{n+1}(t,\cdot) = e^{\Delta t}\mathbb{P}_\delta \mu_0 + \varepsilon \int_0^t e^{\Delta(t-s)}\mathrm{div}\,(\mu^n(s,\cdot)\Theta_p(s,\cdot,\mu^n(s,\cdot), Dw^n(s,\cdot)))\ ds \\ + \varepsilon\bar{m}\int_0^t e^{\Delta(t-s)}\mathrm{div}\,(\Theta_p(s,\cdot,\mu^n(s,\cdot), Dw^n(s,\cdot)))\ ds.\end{aligned} \tag{13}$$

If we considered the payoff problem (4) instead of the planning problem, the change is simply that we replace (11) with the following:

$$w^{n+1}(T,x) = \mathbb{P}_\delta PG(x, m^n(T,x)).$$

Of course, there is the related change in the Duhamel formula as well:

$$\begin{aligned}w^{n+1}(t,\cdot) = e^{\Delta(T-t)}\mathbb{P}_\delta PG(\cdot, m^n(T,\cdot)) \\ - \varepsilon P \int_t^T e^{\Delta(s-t)}\Theta(s,\cdot,\mu^n(s,\cdot), Dw^n(s,\cdot))\ ds.\end{aligned}$$

Remark 2. *Because of the presence of the projection $\mathbb{P}_\delta$, the initial and terminal data for μ^n and w^n, respectively, for all n, is infinitely smooth. Furthermore, μ^n and w^n satisfy linear heat equations. It is trivial to show by induction, then, that for all n, the solutions given by* (12) *and* (13) *are infinitely smooth at each time $t \in [0,\infty)$, at least if the Hamiltonian is C^∞ (the regularity of the iterates is otherwise only limited by the regularity of $\mathcal{H}$). This fact helps to justify the estimates to be carried out in Section 3 below.*

3. Uniform Estimates and Existence of Solutions

Having defined a sequence of approximate solutions (w^n, μ^n) in Section 2, we will now work towards passing to the limit as n goes to infinity. In the present section, we will state assumptions on the Hamiltonian which will allow us to make estimates uniform in n. We focus on the planning problem; the payoff problem may be treated similarly, and would require just some assumption on the mapping properties of the payoff function.

We introduce now some further multi-index notation. We will use this for denoting derivatives of Θ. Consider $\Theta = \Theta(t, x_1, \ldots, x_d, q, p_1, \ldots, p_d)$. A multi-index β is an element of $\mathbb{N}^{2d+1}$; the first d positions correspond to the spatial variables $x_1, x_2, \ldots x_d$, the $(d+1)^{\mathrm{st}}$ position corresponds to the variable q (which is a placeholder for μ), and the final d positions correspond to the p variables. Derivatives with respect to such a multi-index β are denoted

∂^{β}, and the order of β is $|\beta| = \sum_{\ell=1}^{2d+1} \beta_{\ell}$, as is usual. We make the following assumption on $\mathcal{H}$:

(H1) The function $\mathcal{H}$ is such that there exists a non-decreasing function $\tilde{F} : [0,\infty) \to [0,\infty)$ such that for all $\beta \in \mathbb{N}^{2d+1}$ with $|\beta| \leq s+2$,

$$\left|\partial^{\beta}\Theta(\cdot,\cdot,\nu,Dy)\right|_{\infty} \leq \tilde{F}\left(|\nu|_{\infty} + |Dy|_{\infty}\right).$$

We use Sobolev embedding to replace $\tilde{F}$ with the closely related F, which is also a non-decreasing function and which instead satisfies

$$\left|\partial^{\beta}\Theta(\cdot,\cdot,\nu,Dy)\right|_{\infty} \leq F\left(\|\nu\|^2_{\lceil\frac{d+1}{2}\rceil} + \|Dy\|^2_{\lceil\frac{d+1}{2}\rceil}\right), \tag{14}$$

for all β as above. Based on this assumption, we may conclude a useful lemma.

Lemma 3. *Assume* **(H1)** *is satisfied. There exist constants $\bar{c} > 0$ and $\bar{C} > 0$ such that for multi-indices β (as defined at the beginning of this section) and for multi-indices α (as defined in Section 1.1),*

$$\sum_{|\beta|\leq 2}\left(\sum_{|\alpha|\leq s-1}\left\|\partial^{\alpha}\left((\partial^{\beta}\Theta)(t,\cdot,\mu,Dw)\right)\right\|_{L^2(\mathbb{T}^d)}\right)$$
$$\leq \bar{c}F\left(\|\mu\|^2_{\lceil\frac{d+1}{2}\rceil} + \|Dw\|^2_{\lceil\frac{d+1}{2}\rceil}\right)\left(1+\|\mu\|_{s-1}+\|w\|_s\right)^{s-1},$$

and furthermore,

$$\sum_{|\beta|\leq 2}\left\|(\partial^{\beta}\Theta)(t,\cdot,\mu,Du)\right\|_{s-1} \tag{15}$$
$$\leq \bar{C}F\left(\|\mu\|^2_{\lceil\frac{d+1}{2}\rceil} + \|Dw\|^2_{\lceil\frac{d+1}{2}\rceil}\right)\left(1+\|\mu\|_{s-1}+\|w\|_s\right)^{s-1}.$$

It is helpful to expand the divergence which appears in the evolution equation for μ^{n+1}. We have the following formula:

$$\operatorname{div}\left(\Theta_p(t,x,\mu^n,Dw^n)\right) = \sum_{i=1}^{d}\Theta_{x_ip_i}(t,x,\mu^n,Dw^n) \tag{16}$$
$$+\sum_{i=1}^{d}\left[\left(\Theta_{qp_i}(t,x,\mu^n,Dw^n)\right)(\partial_{x_i}\mu^n)\right]+\sum_{i=1}^{d}\sum_{j=1}^{d}\left[(\Theta_{p_ip_j}(t,x,\mu^n,Dw^n))(\partial^2_{x_ix_j}w^n)\right].$$

Remark 3. *There are clearly three kinds of terms on the right-hand side of (16). We will be making energy estimates for $(w^m,\mu^m) \in H^s \times H^{s-1}$, for all m. The first kind of term involves no derivatives of μ^n and first derivatives of w^n, and may be treated routinely in the estimates. The second kind of* *terms involve first derivatives on each of μ^n and w^n. The first derivatives on w^n cause no problems because of the choice of function space. The first derivatives on μ^n indicate that these are transport terms, which could typically be treated in the energy estimate by integration by parts. However, because of our iterative scheme, the necessary structure for integration by parts is not present. We will instead bound these terms using the available parabolic smoothing. The third kind of term on the right-hand side of* (16) *involves no derivatives on μ^n and second derivatives on w^n; these terms will also be bounded by taking advantage of parabolic smoothing.*

Similarly to the above, we apply ∂_{x_j} to (10):

$$\begin{aligned} \partial_{x_j} w_t^{n+1} &= -\Delta \partial_{x_j} w^{n+1} - \varepsilon \Theta_{x_j}(\cdot,\cdot,\mu^n, Dw^n) \\ &\quad - \varepsilon \left(\Theta_q(\cdot,\cdot,\mu^n, Dw^n)\right) \mu^n_{x_j} - \varepsilon \sum_{i=1}^{d} \left[\left(\Theta_{p_i}(\cdot,\cdot,\mu^n, Dw^n)\right) \partial^2_{x_i x_j} w^n \right]. \end{aligned} \tag{17}$$

Notice that we have dropped the operator P, since $\partial_{x_j} P = \partial_{x_j}$.

We provide some notation for certain norms which we will be useful for our estimates. For all $n \in \mathbb{N}$, we define M_n and N_n to be

$$M_n = \sup_{t \in [0,T]} \left(\|Dw^n\|_{s-1}^2 + \|\mu^n\|_{s-1}^2 \right),$$

$$N_n = \sum_{1 \le |\alpha| \le s} \int_0^T \|\partial^\alpha D w^n\|_0^2 \, d\tau + \sum_{0 \le |\alpha| \le s-1} \int_0^T \|\partial^\alpha D \mu^n\|_0^2 \, d\tau. \tag{18}$$

We will be proving an estimate for the solutions which is uniform in n. We will do so in stages; first, we will prove an estimate for (w^{n+1}, μ^{n+1}) in terms of (w^n, μ^n). Then we will proceed inductively, making an assumption about (w^n, μ^n), and showing that this implies the corresponding bound holds for (w^{n+1}, μ^{n+1}). This inductive step will use our smallness assumption (which remains to be stated).

Let α be a multi-index (as defined in Section 1.1) of order $|\alpha| = s-1$. We compute the time derivative of the square of the L^2-norm of $\partial^\alpha \mu$:

$$\begin{aligned}
(19)\quad & \frac{d}{dt}\frac{1}{2}\int_{\mathbb{T}^d}\left(\partial^\alpha \mu^{n+1}\right)^2\ dx \\
&= \int_{\mathbb{T}^d}\left(\partial^\alpha\mu^{n+1}\right)\left(\partial^\alpha\Delta\mu^{n+1}\right)\ dx - \varepsilon\int_{\mathbb{T}^d}\left(\partial^\alpha\mu^{n+1}\right)\partial^\alpha\left(D\mu^n\cdot\Theta_p(\cdot,x,\mu^n,Dw^n)\right)\ dx \\
&\quad -\varepsilon\int_{\mathbb{T}^d}\left(\partial^\alpha\mu^{n+1}\right)\partial^\alpha\left((\mu^n+\bar{m})\sum_{i=1}^d\Theta_{x_ip_i}(\cdot,x,\mu^n,Dw^n)\right)\ dx \\
&\quad -\varepsilon\int_{\mathbb{T}^d}\left(\partial^\alpha\mu^{n+1}\right)\partial^\alpha\left((\mu^n+\bar{m})\sum_{i=1}^d\left[\left(\Theta_{qp_i}(\cdot,x,\mu^n,Dw^n)\right)\left(\partial_{x_i}\mu^n\right)\right]\right)\ dx \\
&\quad -\varepsilon\int_{\mathbb{T}^d}\left(\partial^\alpha\mu^{n+1}\right)\partial^\alpha\left((\mu^n+\bar{m})\sum_{i=1}^d\sum_{j=1}^d\left[\left(\Theta_{p_ip_j}(\cdot,x,\mu^n,Dw^n)\right)\left(\partial^2_{x_ix_j}w^n\right)\right]\right)\ dx.
\end{aligned}$$

We integrate by parts in the first integral on the right-hand side, we move the resulting integral to the left-hand side, and we integrate (19) in time, over the interval $[0,t]$:

$$\begin{aligned}
(20)\quad & \\
& \frac{1}{2}\int_{\mathbb{T}^d}\left(\partial^\alpha\mu^{n+1}(t,x)\right)^2\ dx - \frac{1}{2}\int_{\mathbb{T}^d}\left(\partial^\alpha\mu^{n+1}(0,x)\right)^2\ dx + \int_0^t\int_{\mathbb{T}^d}\left|D\partial^\alpha\mu^{n+1}\right|^2\ dxd\tau \\
&= -\varepsilon\int_0^t\int_{\mathbb{T}^d}\left(\partial^\alpha\mu^{n+1}\right)\partial^\alpha\left(D\mu^n\cdot\Theta_p(\tau,x,\mu^n,Dw^n)\right)\ dxd\tau \\
&\quad -\varepsilon\int_0^t\int_{\mathbb{T}^d}\left(\partial^\alpha\mu^{n+1}\right)\partial^\alpha\left((\mu^n+\bar{m})\sum_{i=1}^d\Theta_{x_ip_i}(\tau,x,\mu^n,Dw^n)\right)\ dxd\tau \\
&\quad -\varepsilon\int_0^t\int_{\mathbb{T}^d}\left(\partial^\alpha\mu^{n+1}\right)\partial^\alpha\left((\mu^n+\bar{m})\sum_{i=1}^d\left[\left(\Theta_{qp_i}(\tau,x,\mu^n,Dw^n)\right)\left(\partial_{x_i}\mu^n\right)\right]\right)\ dxd\tau \\
&\quad -\varepsilon\int_0^t\int_{\mathbb{T}^d}\left(\partial^\alpha\mu^{n+1}\right)\partial^\alpha\left((\mu^n+\bar{m})\sum_{i=1}^d\sum_{j=1}^d\left[\left(\Theta_{p_ip_j}(\tau,x,\mu^n,Dw^n)\right)\left(\partial^2_{x_ix_j}w^n\right)\right]\right)\ dxd\tau \\
&= I + II + III + IV.
\end{aligned}$$

We first work to estimate I, and we begin by adding and subtracting:

$$\begin{aligned}
I &= -\varepsilon\int_0^t\int_{\mathbb{T}^d}(\partial^\alpha\mu^{n+1})(\partial^\alpha D\mu^n)\cdot\Theta_p(\tau,x,\mu^n,Dw^n)\ dxd\tau \\
&\quad +\varepsilon\int_0^t\int_{\mathbb{T}^d}(\partial^\alpha\mu^{n+1})\left[(\partial^\alpha D\mu^n)\cdot\Theta_p(\tau,x,\mu^n,Dw^n) - \partial^\alpha(D\mu^n\cdot\Theta_p(\tau,x,\mu^n,Dw^n))\right]\ dxd\tau \\
&= I_A + I_B.
\end{aligned}$$

We start with I_A, pulling the supremum of the Θ_p term outside the integrals:

$$I_A \leq \varepsilon \left(\sup_{t\in[0,T]} |\Theta_p(t,\cdot,\mu^n,Dw^n)|_\infty \right) \int_0^t \int_{\mathbb{T}^d} |\partial^\alpha \mu^{n+1}||\partial^\alpha D\mu^n| \ dxd\tau.$$

We use (14) to bound Θ_p in terms of M_n :

$$I_A \leq \varepsilon F(M_n) \int_0^t \int_{\mathbb{T}^d} |\partial^\alpha \mu^{n+1}||\partial^\alpha D\mu^n| \ dxd\tau.$$

Next, we continue by using (5) with positive parameter σ_1, which will be determined presently:

$$\begin{aligned} I_A &\leq \varepsilon F(M_n) \left(\frac{1}{2\sigma_1} \int_0^T \|\partial^\alpha \mu^{n+1}\|_0^2 \ d\tau + \frac{\sigma_1}{2} \int_0^T \|\partial^\alpha D\mu^n\|_0^2 \ d\tau \right) \\ &\leq \varepsilon F(M_n) \left(\frac{1}{2\sigma_1} \int_0^T \|\partial^\alpha \mu^{n+1}\|_0^2 \ d\tau + \frac{\sigma_1}{2} N_n \right). \end{aligned}$$

We let $\sigma_1 = 28T\varepsilon F(M_n)$, and this choice then yields the following:

$$\begin{aligned} I_A &\leq \frac{1}{56T} \int_0^T \|\partial^\alpha \mu^{n+1}\|_0^2 \ d\tau + 14\varepsilon^2 T(F(M_n))^2 N_n \\ &\leq \frac{1}{56} \left(\sup_{t\in[0,T]} \|\partial^\alpha \mu^{n+1}\|_0^2 \right) + 14\varepsilon^2 T(F(M_n))^2 N_n. \end{aligned} \tag{21}$$

We turn now to estimating I_B; we start by using (5) with parameter σ_2 :

$$\begin{aligned} I_B &\leq \varepsilon \int_0^t \int_{\mathbb{T}^d} |(\partial^\alpha \mu^{n+1})| \Big| (\partial^\alpha D\mu^n)\cdot\Theta_p(\tau,x,\mu^n,Dw^n) - \partial^\alpha(D\mu^n\cdot\Theta_p(\tau,x,\mu^n,Dw^n)) \Big| \ dxd\tau \\ &\leq \varepsilon \int_0^T \frac{1}{2\sigma_2} \|\partial^\alpha \mu^{n+1}\|_0^2 d\tau \\ &+ \frac{\varepsilon\sigma_2}{2} \int_0^T \Big\| (\partial^\alpha D\mu^n)\cdot\Theta_p(\tau,x,\mu^n,Dw^n) - \partial^\alpha(D\mu^n\cdot\Theta_p(\tau,x,\mu^n,Dw^n)) \Big\|_0^2 \ d\tau. \end{aligned}$$

We let $\sigma_2 = 28T\varepsilon$, and we continue:

$$\begin{aligned} I_B &\leq \frac{1}{56} \left(\sup_{t\in[0,T]} \|\partial^\alpha \mu^{n+1}\|_0^2 \right) \\ &+ 14\varepsilon^2 T \int_0^T \Big\| (\partial^\alpha D\mu^n)\cdot\Theta_p(\tau,\cdot,\mu^n,Dw^n) - \partial^\alpha(D\mu^n\cdot\Theta_p(\tau,\cdot,\mu^n,Dw^n)) \Big\|_0^2 \ d\tau. \end{aligned}$$

Next, we use Lemma 2 and Sobolev embedding as follows (we temporarily suppress the dependence of Θ_p on its arguments):

$$\begin{aligned} \|\partial^\alpha(D\mu^n\cdot\Theta_p) - (\partial^\alpha D\mu^n)\cdot\Theta_p\|_0^2 &\leq c\left(|D\Theta_p|_\infty \|D^{s-1}\mu^n\|_0 + \|D^{s-1}\Theta_p\|_0 |D\mu^n|_\infty \right)^2 \\ &\leq c\left(\|\Theta_p\|_{\lceil \frac{d+3}{2} \rceil} \|\mu^n\|_{s-1} + \|\Theta_p\|_{s-1} \|\mu^n\|_{\lceil \frac{d+3}{2} \rceil} \right)^2 \leq c\|\Theta_p\|_{s-1}^2 \|\mu^n\|_{s-1}^2. \end{aligned}$$

Here, we have used the condition $s \geq \left\lceil \frac{d+5}{2} \right\rceil$. By Lemma 3, we have

$$\|\Theta_p\|_{s-1} \leq cF(M_n)(1+M_n)^{(s-1)/2}.$$

Putting this information together, we complete our bound of I_B :

$$I_B \leq \frac{1}{56}\left(\sup_{t\in[0,T]} \|\partial^\alpha \mu^{n+1}\|_0^2\right) + c\varepsilon^2 T^2 (F(M_n))^2 M_n (1+M_n)^{s-1}. \tag{22}$$

Note that this constant c is independent of ε, T, n, and δ; instead it depends only upon s and d. The same will be true for any constants which we call c in the sequel.

We are ready to estimate the term II. We begin by using Young's inequality (5) with parameter $\sigma_3 = 28T\varepsilon$:

(23)

$$II \leq \varepsilon \int_0^T \int_{\mathbb{T}^d} \frac{|\partial^\alpha \mu^{n+1}|^2}{2\sigma_3} + \frac{\sigma_3}{2}\left|\partial^\alpha\left((\mu^n+\bar{m})\sum_{i=1}^d \Theta_{x_ip_i}(\tau,x,\mu^n,Dw^n)\right)\right|^2 dxd\tau$$

$$\leq \frac{1}{56}\left(\sup_{t\in[0,T]}\|\partial^\alpha\mu^{n+1}\|_0^2\right) + 14\varepsilon^2 T\int_0^T\left\|\partial^\alpha\left((\mu^n+\bar{m})\sum_{i=1}^d\Theta_{x_ip_i}(\tau,\cdot,\mu^n,Dw^n)\right)\right\|_0^2 d\tau.$$

We use the Sobolev algebra property and Lemma 3 to bound the integrand:

(24)

$$\left\|\partial^\alpha\left((\mu^n+\bar{m})\sum_{i=1}^d\Theta_{x_ip_i}(\tau,\cdot,\mu^n,Dw^n)\right)\right\|_0^2 \leq \left\|(\mu^n+\bar{m})\sum_{i=1}^d\Theta_{x_ip_i}(\tau,\cdot,\mu^n,Dw^n)\right\|_{s-1}^2$$

$$\leq c\|\mu^n+\bar{m}\|_{s-1}^2 (F(M_n))^2(1+M_n)^{s-1} \leq c(F(M_n))^2(1+M_n)^s.$$

Combining (24) with (23), we complete our bound for the term II :

$$II \leq \frac{1}{56}\left(\sup_{t\in[0,T]}\|\partial^\alpha\mu^{n+1}\|_0^2\right) + c\varepsilon^2T^2(F(M_n))^2(1+M_n)^s. \tag{25}$$

Before estimating III, we again add and subtract to isolate the leading-order term. We have $III = III_A + III_B$, with III_A given by

$$III_A = -\varepsilon\int_0^t\int_{\mathbb{T}^d}\left(\partial^\alpha\mu^{n+1}\right)(\mu^n+\bar{m})\sum_{i=1}^d\left[\left(\Theta_{qp_i}(\tau,x,\mu^n,Dw^n)\right)\left(\partial^\alpha\partial_{x_i}\mu^n\right)\right] dxd\tau,$$

and the remainder III_B given by

$$III_B = -\varepsilon\int_0^t\int_{\mathbb{T}^d}\left(\partial^\alpha\mu^{n+1}\right)\left\{\partial^\alpha\left((\mu^n+\bar{m})\sum_{i=1}^d\left[\left(\Theta_{qp_i}(\tau,x,\mu^n,Dw^n)\right)\left(\partial_{x_i}\mu^n\right)\right]\right)\right.$$

$$\left. - (\mu^n+\bar{m})\sum_{i=1}^d\left[\left(\Theta_{qp_i}(\tau,x,\mu^n,Dw^n)\right)\left(\partial^\alpha\partial_{x_i}\mu^n\right)\right]\right\} dxd\tau.$$

We begin estimating III_A by using the triangle inequality and pulling the lower-order terms through the integrals:

$$III_A \leq c\varepsilon(1+M_n)^{1/2}\left(\sup_{t\in[0,T],i\in\{1,\ldots,d\}}|\Theta_{qp_i}(t,\cdot,\mu^n,Dw^n)|_\infty\right) \times \sum_{i=1}^{d}\int_0^T\int_{\mathbb{T}^d}|\partial^\alpha\mu^{n+1}||\partial^\alpha\partial_{x_i}\mu^n|\ dxd\tau.$$

Next, we use (14), and we use Young's inequality (5) with parameter σ_4 :

$$III_A \leq c\varepsilon(1+M_n)^{1/2}F(M_n)\int_0^T \frac{1}{2\sigma_4}\|\partial^\alpha\mu^{n+1}\|_0^2 + \frac{\sigma_4}{2}\|\partial^\alpha D\mu^n\|_0^2\ d\tau.$$

We take $\sigma_4 = 28c\varepsilon T(1+M_n)^{1/2}F(M_n)$, and we complete our estimate of III_A :

$$III_A \leq \frac{1}{56}\left(\sup_{t\in[0,T]}\|\partial^\alpha\mu^{n+1}\|_0^2\right) + c\varepsilon^2T(1+M_n)(F(M_n))^2N_n. \tag{26}$$

We next estimate III_B. We begin by applying Young's inequality, with parameter $\sigma_5 > 0$:

$$III_B \leq \varepsilon\int_0^t\int_{\mathbb{T}^d}\frac{1}{2\sigma_5}|\partial^\alpha\mu^{n+1}|^2 + \frac{\sigma_5}{2}\left|\partial^\alpha\left((\mu^n+m)\sum_{i=1}^{d}(\Theta_{qp_i})(\partial_{x_i}\mu^n)\right) - (\mu^n+\bar{m})\sum_{i=1}^{d}(\Theta_{qp_i})(\partial^\alpha\partial_{x_i}\mu^n)\right|^2\ dxd\tau$$

$$= \frac{\varepsilon}{2\sigma_5}\int_0^t\|\partial^\alpha\mu^{n+1}\|_0^2\ d\tau + \frac{\varepsilon\sigma_5}{2}\int_0^t\left\|\partial^\alpha\left((\mu^n+\bar{m})\sum_{i=1}^{d}(\Theta_{qp_i})(\partial_{x_i}\mu^n)\right) - (\mu^n+\bar{m})\sum_{i=1}^{d}(\Theta_{qp_i})(\partial^\alpha\partial_{x_i}\mu^n)\right\|_0^2\ d\tau.$$

We proceed by using Lemma 2, as in our previous estimate for the term I_B; we find the following:

$$III_B \leq \frac{\varepsilon T}{2\sigma_5}\left(\sup_{t\in[0,T]}\|\partial^\alpha\mu^{n+1}\|_0^2\right) + cT\varepsilon\sigma_5(F(M_n))^2(1+M_n)^{s+1}.$$

We conclude the estimate of III_B by setting $\sigma_5 = 28T\varepsilon$:

$$III_B \leq \frac{1}{56}\left(\sup_{t\in[0,T]}\|\partial^\alpha\mu^{n+1}\|_0^2\right) + cT^2\varepsilon^2(F(M_n))^2(1+M_n)^{s+1}. \tag{27}$$

For the term IV, we again must separate the leading-order term by adding and subtracting. We write $IV = IV_A + IV_B$, with IV_A given by

$$IV_A = -\varepsilon \int_0^t \int_{\mathbb{T}^d} \left(\partial^\alpha \mu^{n+1}\right) (\mu^n + \bar{m}) \sum_{i=1}^d \sum_{j=1}^d \left[\left(\Theta_{p_i p_j}(\tau, x, \mu^n, Dw^n)\right) \left(\partial^\alpha \partial^2_{x_i x_j} w^n\right) \right] \, dx d\tau,$$

and with the remainder IV_B given by

$$IV_B =$$

$$-\varepsilon \int_0^t \int_{\mathbb{T}^d} \left(\partial^\alpha \mu^{n+1}\right) \left\{ \partial^\alpha \left((\mu^n + \bar{m}) \sum_{i=1}^d \sum_{j=1}^d \left[\left(\Theta_{p_i p_j}(\tau, x, \mu^n, Dw^n)\right) \left(\partial^2_{x_i x_j} w^n\right) \right] \right) \right.$$
$$\left. - (\mu^n + \bar{m}) \sum_{i=1}^d \sum_{j=1}^d \left[\left(\Theta_{p_i p_j}(\tau, x, \mu^n, Dw^n)\right) \left(\partial^\alpha \partial^2_{x_i x_j} w^n\right) \right] \right\} \, dx d\tau.$$

To estimate IV_A, we first pull $\mu^n + \bar{m}$ and $\Theta_{p_i p_j}$ through the integrals by taking supremums:

$$IV_A \leq c\varepsilon (1 + M_n)^{1/2} \left(\sup_{t \in [0,T]} \sup_{i,j} |\Theta_{p_i p_j}|_\infty \right) \sum_{i,j} \int_0^t \int_{\mathbb{T}^d} |\partial^\alpha \mu^{n+1}| |\partial^\alpha \partial^2_{x_i x_j} w^n| \, dx d\tau.$$

We estimate this by using (14), and we use Young's inequality with parameter $\sigma_6 > 0$:

$$IV_A \leq c\varepsilon F(M_n)(1 + M_n)^{1/2} \sum_{i,j} \int_0^T \frac{1}{2\sigma_6} \|\partial^\alpha \mu^{n+1}\|_0^2 + \frac{\sigma_6}{2} \|\partial^\alpha \partial^2_{x_i x_j} w^n\|_0^2 \, d\tau.$$

Taking $\sigma_6 = 28 c\varepsilon T (1 + M_n)^{1/2} F(M_n)$, and proceeding as we have previously, we arrive at our final bound for IV_A :

$$IV_A \leq \frac{1}{56} \left(\sup_{t \in [0,T]} \|\partial^\alpha \mu^{n+1}\|_0^2 \right) + c\varepsilon^2 T (1 + M_n)(F(M_n))^2 N_n. \tag{28}$$

We estimate IV_B just as we have estimated III_B, finding that

$$IV_B \leq \frac{1}{56} \left(\sup_{t \in [0,T]} \|\partial^\alpha \mu^{n+1}\|_0^2 \right) + cT^2 \varepsilon^2 (F(M_n))^2 (1 + M_n)^{s+1}. \tag{29}$$

To summarize our progress so far, we add (21), (22), (25), (26), (27), (28), and (29), and we make some elementary bounds, to find the following:

$$I + II + III + IV \tag{30}$$
$$\leq \frac{1}{8} \left(\sup_{t \in [0,T]} \|\partial^\alpha \mu^{n+1}\|_0^2 \right) + c\varepsilon^2 T (F(M_n))^2 \left((1+T)(1+N_n)(1+M_n)^{s+1} \right).$$

Continuing, we compute the time derivative of the square of the L^2-norm of $\partial^\alpha\partial_{x_j}w^{n+1}$, substituting from (17):

$$\begin{aligned}(31)\quad \frac{d}{dt}\frac{1}{2}\int_{\mathbb{T}^d}(\partial^\alpha\partial_{x_j}w^{n+1})^2\ dx &= -\int_{\mathbb{T}^d}(\partial^\alpha\partial_{x_j}w^{n+1})(\partial^\alpha\Delta\partial_{x_j}w^{n+1})\ dx \\ &-\varepsilon\int_{\mathbb{T}^d}(\partial^\alpha\partial_{x_j}w^{n+1})\partial^\alpha\left(\Theta_{x_j}(\cdot,x,\mu^n,Dw^n)\right)\ dx \\ &-\varepsilon\int_{\mathbb{T}^d}(\partial^\alpha\partial_{x_j}w^{n+1})\partial^\alpha\left(\left(\Theta_q(\cdot,x,\mu^n,Dw^n)\right)\mu^n_{x_j}\right)\ dx \\ &-\varepsilon\int_{\mathbb{T}^d}(\partial^\alpha\partial_{x_j}w^{n+1})\partial^\alpha\left(\sum_{i=1}^d\left[(\Theta_{p_i}(\cdot,x,\mu^n,Dw^n))\partial^2_{x_ix_j}w^n\right]\right)\ dx.\end{aligned}$$

We integrate by parts in the first integral on the right-hand side and we integrate (31) in time over the interval $[t,T]$; we also rearrange terms slightly, arriving at the following:

$$\begin{aligned}(32)\quad \frac{1}{2}\int_{\mathbb{T}^d}(\partial^\alpha\partial_{x_j}w^{n+1}(t,x))^2\ dx &- \frac{1}{2}\int_{\mathbb{T}^d}(\partial^\alpha\partial_{x_j}w^{n+1}(T,x))^2\ dx \\ &+\int_t^T\int_{\mathbb{T}^d}|D\partial^\alpha\partial_{x_j}w^{n+1}|^2\ dxd\tau \\ &=\varepsilon\int_t^T\int_{\mathbb{T}^d}(\partial^\alpha\partial_{x_j}w^{n+1})\partial^\alpha\left(\Theta_{x_j}(\cdot,x,\mu^n,Dw^n)\right)\ dxd\tau \\ &+\varepsilon\int_t^T\int_{\mathbb{T}^d}(\partial^\alpha\partial_{x_j}w^{n+1})\partial^\alpha\left(\left(\Theta_q(\cdot,x,\mu^n,Dw^n)\right)\mu^n_{x_j}\right)\ dxd\tau \\ &+\varepsilon\int_t^T\int_{\mathbb{T}^d}(\partial^\alpha\partial_{x_j}w^{n+1})\partial^\alpha\left(\sum_{i=1}^d\left[(\Theta_{p_i}(\cdot,x,\mu^n,Dw^n))\partial^2_{x_ix_j}w^n\right]\right)\ dxd\tau \\ &= V+VI+VII.\end{aligned}$$

The term V is straightforward to estimate; we begin with Young's inequality, with parameter $\sigma_7>0:$

$$V\le\varepsilon\int_t^T\int_{\mathbb{T}^d}\frac{1}{2\sigma_7}|\partial^\alpha\partial_{x_j}w^{n+1}|^2+\frac{\sigma_7}{2}|\partial^\alpha\Theta_{x_j}|^2\ dxd\tau.$$

We choose $\sigma_7=20\varepsilon T$, we use Lemma 3, and we estimate similarly to the previous terms to find the following:

$$(33)\qquad V\le\frac{1}{40}\left(\sup_{t\in[0,T]}\|\partial^\alpha\partial_{x_j}w^{n+1}\|_0^2\right)+c\varepsilon^2T^2(F(M_n))^2(1+M_n)^{s-1}.$$

Before estimating VI, we must add and subtract to isolate the leading-order term. We have $VI=VI_A+VI_B$, with VI_A given by

$$VI_A=\varepsilon\int_t^T\int_{\mathbb{T}^d}(\partial^\alpha\partial_{x_j}w^{n+1})(\Theta_q(\tau,x,\mu^n,Dw^n))(\partial^\alpha\partial_{x_j}\mu^n)\ dxd\tau,$$

and with the remainder VI_B given by

$$VI_B = \varepsilon \int_t^T \int_{\mathbb{T}^d} (\partial^\alpha \partial_{x_j} w^{n+1}) \Big\{ \partial^\alpha \left((\Theta_q(\tau, x, \mu^n, Dw^n)) \left(\partial_{x_j} \mu^n \right) \right) - (\Theta_q(\tau, x, \mu^n, Dw^n)) \left(\partial^\alpha \partial_{x_j} \mu^n \right) \Big\} \, dx d\tau.$$

We begin estimating VI_A by taking the supremum of Θ_q with respect to space and time, and pulling this through the integrals:

$$VI_A \leq \varepsilon \left(\sup_{t \in [0,T]} |\Theta_q(t, \cdot, \mu^n, Dw^n)|_\infty \right) \int_t^T \int_{\mathbb{T}^d} |\partial^\alpha \partial_{x_j} w^{n+1}| |\partial^\alpha \partial_{x_j} \mu^n| \, dx d\tau.$$

We then bound the Θ_q term by using (14) and by using Young's inequality with positive parameter σ_8 :

$$VI_A \leq \varepsilon F(M_n) \int_t^T \frac{1}{2\sigma_8} \|\partial^\alpha \partial_{x_j} w^{n+1}\|_0^2 + \frac{\sigma_8}{2} \|\partial^\alpha \partial_{x_j} \mu^n\|_0^2 \, d\tau.$$

We take $\sigma_8 = 20\varepsilon T F(M_n)$, and estimate as we have previously, finding the following:

$$VI_A \leq \frac{1}{40} \left(\sup_{t \in [0,T]} \|\partial^\alpha \partial_{x_j} w^{n+1}\|_0^2 \right) + c\varepsilon^2 T (F(M_n))^2 N_n. \tag{34}$$

We begin estimating VI_B with use of Young's inequality, with positive parameter σ_9 :

$$VI_B \leq \varepsilon \int_t^T \int_{\mathbb{T}^d} \frac{1}{2\sigma_9} |\partial^\alpha \partial_{x_j} w^{n+1}|^2 + \frac{\sigma_9}{2} \left| \partial^\alpha ((\Theta_q)(\partial_{x_j} \mu^n)) - (\Theta_q)(\partial^\alpha \partial_{x_j} \mu^n) \right|^2 \, dx d\tau.$$

We take $\sigma_9 = 20\varepsilon T$, and proceed as usual:

$$VI_B \leq \frac{1}{40} \left(\sup_{t \in [0,T]} \|\partial^\alpha \partial_{x_j} w^{n+1}\|_0^2 \right) + c\varepsilon^2 T^2 \left(\sup_{t \in [0,T]} \left\| \partial^\alpha ((\Theta_q)(\partial_{x_j} \mu^n)) - (\Theta_q)(\partial^\alpha \partial_{x_j} \mu^n) \right\|_0^2 \right).$$

Using Lemma 2 and Sobolev embedding, we bound this as follows:

$$VI_B \leq \frac{1}{40} \left(\sup_{t \in [0,T]} \|\partial^\alpha \partial_{x_j} w^{n+1}\|_0^2 \right) + c\varepsilon^2 T^2 \left(\sup_{t \in [0,T]} \|\Theta_q\|_{\lceil \frac{d+3}{2} \rceil}^2 \|\mu^n\|_{s-1}^2 + \|\Theta_q\|_{s-1}^2 \|\mu^n\|_{\lceil \frac{d+3}{2} \rceil}^2 \right).$$

Using Lemma 3, and the fact that s is sufficiently large ($s \geq \lceil \frac{d+5}{2} \rceil$ is needed here), we conclude our bound of VI_B :

$$VI_B \leq \frac{1}{40}\left(\sup_{t\in[0,T]} \|\partial^\alpha \partial_{x_j} w^{n+1}\|_0^2\right) + c\varepsilon^2 T^2 (F(M_n))^2(1+M_n)^s. \tag{35}$$

For the term VII, we must again add and subtract to isolate the leading-order contribution. We write $VII = VII_A + VII_B$, with VII_A given by

$$VII_A = \varepsilon \int_t^T \int_{\mathbb{T}^d} (\partial^\alpha \partial_{x_j} w^{n+1}) \sum_{i=1}^d \left[(\Theta_{p_i}(\tau, x, \mu^n, Dw^n)) \left(\partial^\alpha \partial^2_{x_i x_j} w^n\right)\right] \, dx d\tau,$$

and with the remainder VII_B given by

$$\begin{aligned} VII_B = \varepsilon \int_t^T \int_{\mathbb{T}^d} (\partial^\alpha \partial_{x_j} w^{n+1}) \Bigg\{ & \partial_\alpha \sum_{i=1}^d \left[(\Theta_{p_i}(\tau, x, \mu^n, Dw^n))\, \partial^2_{x_i x_j} w^n \right] \\ & - \sum_{i=1}^d \left[(\Theta_{p_i}(\tau, x, \mu^n, Dw^n)) \left(\partial^\alpha \partial^2_{x_i x_j} w^n\right)\right] \Bigg\} \, dx d\tau. \end{aligned}$$

To begin to estimate VII_A, we pull Θ_{p_i} through the integrals (after taking its supremum):

$$\begin{aligned} VII_A \leq \sum_{i=1}^d \varepsilon & \left(\sup_{t\in[0,T]} |\Theta_{p_i}(t, \cdot, \mu^n, Dw^n)|_\infty \right) \\ & \times \int_t^T \int_{\mathbb{T}^d} |\partial^\alpha \partial_{x_j} w^{n+1}| |\partial^\alpha \partial^2_{x_i x_j} w^n| \, dx d\tau. \end{aligned}$$

We estimate Θ_{p_i} by using (14), and we use Young's inequality with positive parameter σ_{10} :

$$VII_A \leq \sum_{i=1}^d \varepsilon F(M_n) \int_t^T \int_{\mathbb{T}^d} \frac{1}{2\sigma_{10}} |\partial^\alpha \partial_{x_j} w^{n+1}|^2 + \frac{\sigma_{10}}{2} |\partial^\alpha \partial^2_{x_i x_j} w^n| \, dx d\tau.$$

We choose the value $\sigma_{10} = 20 d\varepsilon T F(M_n)$, and thus find the following bound:

$$VII_A \leq \frac{1}{40}\left(\sup_{t\in[0,T]} \|\partial^\alpha \partial_{x_j} w^{n+1}\|_0^2\right) + c\varepsilon^2 (F(M_n))^2 T N_n. \tag{36}$$

We now estimate the final term, VII_B. We interchange the summation and the integrals, and we use Young's inequality with positive parameter

σ_{11} :

$$
\begin{aligned}
VII_B \leq \varepsilon \sum_{i=1}^{d} \int_t^T \int_{\mathbb{T}^d} & \frac{1}{2\sigma_{11}} |\partial^\alpha \partial_{x_j} w^{n+1}|^2 \\
& + \frac{\sigma_{11}}{2} |\partial^\alpha((\Theta_{p_i})(\partial^2_{x_i x_j} w^n)) - (\Theta_{p_i})(\partial^\alpha \partial^2_{x_i x_j} w^n)|^2 \, dx d\tau \\
& \leq \frac{\varepsilon T d}{2\sigma_{11}} \left(\sup_{t\in[0,T]} \|\partial^\alpha \partial_{x_j} w^{n+1}\|_0^2 \right) \\
+ \frac{\sigma_{11}\varepsilon T}{2} \sum_{i=1}^{d} & \left(\sup_{t\in[0,T]} \|\partial^\alpha((\Theta_{p_i})(\partial^2_{x_i x_j} w^n)) - (\Theta_{p_i})(\partial^\alpha \partial^2_{x_i x_j} w^n)\|_0^2 \right).
\end{aligned}
$$

We thus choose $\sigma_{11} = 20\varepsilon T d$, and we use Lemma 2 and Lemma 3 as we have previously:

$$
VII_B \leq \frac{1}{40} \left(\sup_{t\in[0,T]} \|\partial^\alpha \partial_{x_j} w^{n+1}\|_0^2 \right) + c\varepsilon^2 T^2 (F(M_n))^2 (1 + M_n)^s. \tag{37}
$$

We are now in a position to add (33), (34), (35), (36), and (37); we then make some elementary estimates, finding the following:

$$
\begin{aligned}
&V + VI + VII \\
&\leq \frac{1}{8} \left(\sup_{t\in[0,T]} \|\partial^\alpha \partial_{x_j} w^{n+1}\|_0^2 \right) + c\varepsilon^2 T (F(M_n))^2 \left((1+T)(1+N_n)(1+M_n)^{s+1} \right).
\end{aligned} \tag{38}
$$

We return to (20), considering (30). We isolate the first term on the left-hand side of (20), finding the following bound:

$$
\begin{aligned}
\frac{1}{2}\|\partial^\alpha \mu^{n+1}(t,\cdot)\|_0^2 \leq \frac{1}{2}\|\partial^\alpha \mu^{n+1}(0,\cdot)\|_0^2 + \frac{1}{8} & \left(\sup_{t\in[0,T]} \|\partial^\alpha \mu^{n+1}\|_0^2 \right) \\
& + c\varepsilon^2 T (F(M_n))^2 \left((1+T)(1+N_n)(1+M_n)^{s+1} \right).
\end{aligned}
$$

Taking the supremum with respect to t (which does not change the right-hand side) and rearranging, we have

$$
\begin{aligned}
\frac{3}{8} \left(\sup_{t\in[0,T]} \|\partial^\alpha \mu^{n+1}\|_0^2 \right) & \leq \frac{1}{2}\|\partial^\alpha \mu^{n+1}(0,\cdot)\|_0^2 \\
& + c\varepsilon^2 T (F(M_n))^2 \left((1+T)(1+N_n)(1+M_n)^{s+1} \right).
\end{aligned} \tag{39}
$$

We next isolate the time integral on the left-hand side of (20), again using this with (30). As before, we find

$$\int_0^t \|D\partial^\alpha \mu^{n+1}\|_0^2\,d\tau \le \frac{1}{2}\|\partial^\alpha\mu^{n+1}(0,\cdot)\|_0^2 + \frac{1}{8}\left(\sup_{t\in[0,T]}\|\partial^\alpha\mu^{n+1}\|_0^2\right) + c\varepsilon^2 T(F(M_n))^2\Bigg((1+T)(1+N_n)(1+M_n)^{s+1}\Bigg).$$

We again take the supremum in time, and this again does not affect the right-hand side. We add the result to (39), and rearrange to find the following:

$$\frac{1}{4}\left(\sup_{t\in[0,T]}\|\partial^\alpha\mu^{n+1}\|_0^2\right) + \int_0^T \|D\partial^\alpha\mu^{n+1}\|_0^2\,d\tau \le \|\partial^\alpha\mu^{n+1}(0,\cdot)\|_0^2 + c\varepsilon^2T(F(M_n))^2\Bigg((1+T)(1+N_n)(1+M_n)^{s+1}\Bigg). \tag{40}$$

We perform the same manipulations regarding (32) and (38), and add the results to (40). These considerations imply the following:

$$\begin{aligned}\frac{1}{4}&\left(\sup_{t\in[0,T]}\|\partial^\alpha\mu^{n+1}\|_0^2 + \sup_{t\in[0,T]}\|\partial^\alpha\partial_{x_j}w^{n+1}\|_0^2\right)\\ &+\int_0^T \|D\partial^\alpha\mu^{n+1}(\tau,\cdot)\|_0^2\,d\tau + \int_0^T\|D\partial^\alpha\partial_{x_j}w^{n+1}(\tau,\cdot)\|_0^2\,d\tau\\ &\le \|\partial^\alpha\mu^{n+1}(0,\cdot)\|_0^2 + \|\partial^\alpha\partial_{x_j}w^{n+1}(T,\cdot)\|_0^2\\ &\quad + c\varepsilon^2T(F(M_n))^2\Bigg((1+T)(1+N_n)(1+M_n)^{s+1}\Bigg).\end{aligned} \tag{41}$$

We sum (41) over multi-indices α such that $0\le|\alpha|\le s-1$ and also over natural numbers j such that $1\le j\le n$, and we multiply by 4; this results in the following:

$$M_{n+1}+4N_{n+1} \le 4\|\mu^{n+1}(0,\cdot)\|_{s-1}^2 + 4\|Dw^{n+1}(T,\cdot)\|_{s-1}^2 + c\varepsilon^2T(F(M_n))^2\Bigg((1+T)(1+N_n)(1+M_n)^{s+1}\Bigg).$$

We substitute the boundary conditions (9) and (11):

$$M_{n+1}+4N_{n+1} \le 4\|\mathbb{P}_\delta\mu_0\|_{s-1}^2 + 4\|D\mathbb{P}_\delta w_T\|_{s-1}^2 + c\varepsilon^2T(F(M_n))^2\Bigg((1+T)(1+N_n)(1+M_n)^{s+1}\Bigg). \tag{42}$$

We are now ready to make our inductive hypothesis. Let $\mathcal{S}\in\mathbb{R}$ satisfy

$$4\|\mu_0\|_{s-1}^2 + 4\|Dw_T\|_{s-1}^2 \le \mathcal{S}.$$

Note that because of the definition of the smoothing operator $\mathbb{P}_\delta$ and Plancherel's theorem, an immediate consequence is

$$4\|\mathbb{P}_\delta\mu_0\|_{s-1}^2 + 4\|D\mathbb{P}_\delta w_T\|_{s-1}^2 \leq 4\|\mu_0\|_{s-1}^2 + 4\|Dw_T\|_{s-1}^2 \leq \mathcal{S}, \qquad \forall\delta > 0.$$

We make our smallness assumption:

(H2) The function F and the constants c, ε, T, and $\mathcal{S}$ satisfy

$$c\varepsilon^2 T(F(2\mathcal{S}))^2\left((1+T)(1+2\mathcal{S})^{s+2}\right) \leq \mathcal{S}.$$

Remark 4. *In* [3]*, we gave two existence theorems for non-separable mean field games with data in the Wiener algebra, with each of these theorems having a different smallness constraint. Here, we treat up to three different sources of smallness in a unified constraint. Clearly, either by taking ε sufficiently small for fixed T and $\mathcal{S}$, or by instead taking T sufficiently small for fixed ε and $\mathcal{S}$, we may satisfy* **(H2)***. The third source of smallness depends on the form of the Hamiltonian; if, for instance $\mathcal{H}(t,x,m,Du) = m|Du|^4$, then the function F could go to zero with $\mathcal{S}$, and by taking $\mathcal{S}$ sufficiently small, with fixed values of ε and T, the condition* (**H2**) *would be satisfied. For other choices of the Hamiltonian, however, it may not be the case that F vanishes as $\mathcal{S}$ vanishes. In summary, this unified condition treats the size of the time horizon, the strength of the coupling in the model, and in some cases, the size of the data.*

Our inductive hypothesis is that

$$M_n + 4N_n \leq 2\mathcal{S}. \tag{43}$$

It is trivial to see that when $n = 0$, since $\mu^0 = w^0 = 0$, that $M_0 + 4N_0 \leq 2\mathcal{S}$. We assume the inductive hypothesis for some $n \in \mathbb{N}$. Then, we see that $M_n \leq 2\mathcal{S}$ and $N_n \leq 2\mathcal{S}$ as well. Since the function F is monotonic, we have $F(M_n) \leq F(2\mathcal{S})$. Combining the inductive hypothesis (43) with **(H2)** and the bound (42), we conclude $M_{n+1} + 4N_{n+1} \leq 2\mathcal{S}$. Thus, we have proved that for all $n \in \mathbb{N}$, (43) holds.

The estimate (43), together with the definition of M_n, implies that the sequence (μ^n, w^n) is bounded in the space $C([0,T]; H^{s-1} \times H^s)$, uniformly with respect to n. Our specification of s is sufficiently large so that inspection of (8), (10) shows that μ_t^n and w_t^n are uniformly bounded. From this, we are able to conclude that (μ^n, w^n) forms an equicontinuous family, with compact domain $[0,T] \times \mathbb{T}^d$. Applying the Arzela-Ascoli theorem, we find that a subsequence converges uniformly to a limit $(\mu, w) \in (C([0,T] \times \mathbb{T}^d))^2$. This implies, since the domain is compact, that the convergence also holds in $C([0,T]; H^0 \times H^0)$. Applying Lemma 1, using the uniform bound in $H^{s+1} \times H^s$, we also find convergence in the space $C([0,T]; H^{s'-1} \times H^{s'})$, for any $s' \in [0, s)$.

The uniform bound (43) also implies that the sequence w^n is uniformly bounded in $L^2([0,T]; H^{s+1})$ and the sequence μ^n is uniformly bounded in

$L^2([0,T];H^s)$. These are Hilbert spaces and therefore bounded sequences have subsequences with weak limits, with the weak limits obeying the same bounds. So, we may take $w\in L^2([0,T];H^{s+1})$ and $\mu\in L^2([0,T];H^s)$, with the bound

$$\int_0^T \|w(t,\cdot)\|_{H^{s+1}}^2+\|\mu(t,\cdot)\|_{H^s}\ dt\le c\mathcal{S}, \tag{44}$$

with c being an absolute constant related to the definitions of the norms.

Integrating (8), (10) in time, and using the boundary conditions (9) and (11), we see that μ^{n+1} and w^{n+1} satisfy the equations

$$\mu^{n+1}(t,\cdot)=\mathbb{P}_\delta\mu_0+\int_0^t\left[\Delta\mu^{n+1}(\tau,\cdot)-\varepsilon\mathrm{div}\left((\bar{m}+\mu^n(\tau,\cdot))\,\Theta_p(\tau,\cdot,\mu^n,Dw^n)\right)\right]\ d\tau, \tag{45}$$

$$w^{n+1}(t,\cdot)=\mathbb{P}_\delta w_T+\int_t^T\left[\Delta w^{n+1}(\tau,\cdot)+\varepsilon\Theta(\tau,\cdot,\mu^n,Dw^n)\right]\ d\tau. \tag{46}$$

We have established sufficient regularity of the solution (μ,w) to pass to the limit as $n\to\infty$ and as $\delta\to 0$ in (45) and (46); taking these limits, we have

$$\mu(t,\cdot)=\mu_0+\int_0^t\left[\Delta\mu(\tau,\cdot)-\varepsilon\mathrm{div}\left((\bar{m}+\mu(\tau,\cdot))\,\Theta_p(\tau,\cdot,\mu,Dw)\right)\right]\ d\tau, \tag{47}$$

$$w(t,\cdot)=w_T+\int_t^T\left[\Delta w(\tau,\cdot)+\varepsilon\Theta(\tau,\cdot,\mu,Dw)\right]\ d\tau. \tag{48}$$

Differentiating (47) and (48) with respect to time, we find that (μ,w) satisfy (6) and (7), with the boundary values μ_0 and w_T.

To recover m from μ, one simply needs to add $\bar{m}$. To recover u from w and μ, one simply integrates (1) with respect to t, since the right-hand side is determined in terms of w and m.

We summarize what we have proved so far in our first main theorem.

Theorem 4. *Let $T>0$ and $\varepsilon>0$ be given. Let $s\ge\left\lceil\frac{d+5}{2}\right\rceil$ and let $\mu_0\in H^{s-1}(\mathbb{T}^d)$ be such that $\bar{m}+\mu_0$ is a probability measure. Let $u_T\in H^s(\mathbb{T}^d)$ be given. Assume that the conditions* **(H1)** *and* **(H2)** *are satisfied. Then there exists $\mu\in L^\infty([0,T];H^{s-1})\cap L^2([0,T];H^s)$ and there exists $u\in L^\infty([0,T];H^s)\cap L^2([0,T];H^{s+1})$ such that $\bar{m}+\mu$ is a probability measure for all $t\in[0,T]$, and such that $(u,\bar{m}+\mu)$ satisfies* (1)*,* (2)*,* (3)*. Furthermore, for all $s'\in[0,s)$, we have $\mu\in C([0,T];H^{s'-1})$ and $u\in C([0,T];H^{s'})$.*

4. Uniqueness

We consider two solutions, (w^1,μ^1) and (w^2,μ^2) in $H^s\times H^{s-1}$, for $s>2+\frac{d}{2}$, with the norm of these solutions bounded in these spaces by some $K>0$. We define $E(t)=E_\mu(t)+E_w(t)$, with

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

$$E_w(t) = \frac{1}{2}\sum_{i=1}^{d}\int_{\mathbb{T}^d} \left(\partial_{x_i} w^1(t,x) - \partial_{x_i} w^2(t,x)\right)^2\ dx.$$

Thus, we are measuring the difference of Dw in L^2 and the difference of μ also in L^2.

We must have a Lipschitz property for the Hamiltonian for our uniqueness argument. We make the following assumption:

(H3) For all multi-indices β (as described in the beginning of Section 3) with $0 \leq |\beta| \leq 2$, for any (p^i, q^i) in a bounded subset of $\mathbb{R}^{d+1}$, there exists a constant $c > 0$ such that

$$\left|\partial^\beta \Theta(t,x,p^1,q^1) - \partial^\beta \Theta(t,x,p^2,q^2)\right| \leq c\left(|p^1 - p^2| + \sum_{i=1}^{d} |q_i^1 - q_i^2|\right).$$

To estimate the growth of the difference of the two solutions, we take the time derivative of E, starting with E_μ. To begin, we have simply

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

We substitute for μ_t^1 and μ_t^2 from (7), and we do some preliminary adding and subtracting. This leads us to the expression

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

We apply the divergence operators on the right-hand side, making the expansion

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

where we now list each of these terms:

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

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

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

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

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

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

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

$$V_8 = -\varepsilon \int_{\mathbb{T}^d} (\mu^1 - \mu^2)(\mu^2) \sum_{i=1}^d \sum_{j=1}^d \left[\Theta_{p_i p_j}(t,x,\mu^1,Dw^1)\frac{\partial^2 w^1}{\partial x_i \partial x_j} \right.$$
$$\left. - \Theta_{p_i p_j}(t,x,\mu^2,Dw^2)\frac{\partial^2 w^1}{\partial x_i \partial x_j}\right]\ dx,$$

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

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

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

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

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

and finally,

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

We integrate V_1 by parts:

$$V_1 = -\int_{\mathbb{T}^d} \left|\nabla\left(\mu^1 - \mu^2\right)\right|^2\ dx. \tag{49}$$

We also integrate each of V_2, V_7, and V_{12} by parts:

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

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

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

These terms, and also V_3, are then bounded in terms of the energy, using the bound on the solutions. We distinguish between two kinds of bounds, though; there exists a nondecreasing function $\mathcal{G}_1$ which may be taken so as to converge to zero if K vanishes, such that

$$V_7 \leq \varepsilon\mathcal{G}_1(K)E_\mu. \tag{50}$$

The fact that $\mathcal{G}_1$ can be taken to vanish with K is because of the presence of the linear factor μ^2 in V_7. Also note that the regularity requirement $s>2+\frac{d}{2}$ allowed us here to estimate μ^2 in L^∞; the requirement comes into play in the same way several times throughout the rest of the argument. On the other hand, we have a nondecreasing function $\mathcal{G}_2$ such that

$$V_2+V_3+V_{12} \leq \varepsilon\mathcal{G}_2(K)E_\mu. \tag{51}$$

For most of the remaining terms, we estimate them using the Lipschitz properties of Θ_p and its derivatives; these terms satisfy

$$V_4+V_5+V_6+V_8+V_{11}+V_{13} \leq \varepsilon\mathcal{G}_1(K)(E_\mu+E_\mu^{1/2}E_w^{1/2}), \tag{52}$$

where $\mathcal{G}_1(K)$ is as before, and where its vanishing property is again because of the presence of linear factors such as μ^2 in the terms. Another term relies on the Lipschitz estimate for Θ_p, but does not have such a linear factor of the unknowns present; for this, we again have the existence of $\mathcal{G}_2(K)$ such that

$$V_{10} \leq \varepsilon\mathcal{G}_2(K)(E_\mu+E_\mu^{1/2}E_w^{1/2}). \tag{53}$$

This leaves two more terms to deal with, V_9 and V_{14}. We will use Young's inequality for these, and later bound them by a contribution from E_w. For V_9, we begin by bounding $\Theta_{p_ip_j}$ and μ^2 with $\mathcal{G}_1(K)$:

$$V_9 \leq \varepsilon\mathcal{G}_1(K)\sum_{i=1}^{d}\sum_{j=1}^{d}\int_{\mathbb{T}^d}(\mu^1-\mu^2)\frac{\partial^2(w^1-w^2)}{\partial_{x_i}\partial_{x_j}}\ dx.$$

We then apply Young's inequality, with parameter $4\varepsilon\mathcal{G}_1(K)$:

$$\begin{aligned} V_9 &\leq \varepsilon^2\mathcal{G}_1(K)\int_{\mathbb{T}^d}(\mu^1-\mu^2)^2\ dx + \frac{1}{8}\sum_{j=1}^{d}\int_{\mathbb{T}^d}(\partial_{x_j}(Dw^1-Dw^2))^2\ dx \\ &\leq \varepsilon^2\mathcal{G}_1(K)E_\mu + \frac{1}{8}\sum_{j=1}^{d}\int_{\mathbb{T}^d}(\partial_{x_j}(Dw^1-Dw^2))^2\ dx. \end{aligned}$$

The remaining term, V_{14}, is entirely similar, except that we use $\mathcal{G}_2$ instead of $\mathcal{G}_1$:

$$V_{14} \leq \varepsilon^2 \mathcal{G}_2(K) E_\mu + \frac{1}{8} \sum_{j=1}^{d} \int_{\mathbb{T}^d} (\partial_{x_j}(Dw^1 - Dw^2))^2 \, dx.$$

Adding these results for V_9 and V_{14}, we have

$$V_9 + V_{14} \leq \varepsilon(\mathcal{G}_1(K) + \mathcal{G}_2(K)) E_\mu + \frac{1}{4} \sum_{j=1}^{d} \int_{\mathbb{T}^d} (\partial_{x_j}(Dw^1 - Dw^2))^2 \, dx. \tag{54}$$

Summarizing the terms we have estimated so far, leaving V_1 out for the moment, by adding the bounds (50), (51), (52), (53), and (54), we have concluded the following:

$$\sum_{\ell=2}^{14} V_\ell \leq \varepsilon(\mathcal{G}_1(K)+\mathcal{G}_2(K))(E_w+E_\mu)+\frac{1}{4} \sum_{j=1}^{d} \int_{\mathbb{T}^d} (\partial_{x_j}^2(Dw^1-Dw^2))^2 \, dx. \tag{55}$$

We turn our attention to E_w, and we write $E_w = \sum_{j=1}^{d} E_w^j$, 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 then add and subtract to make the following decomposition:

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

with

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

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

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

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

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

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

We integrate W_1^j by parts:

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

We may also integrate W_6^j by parts, to find the following:

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

With $\mathcal{G}_1$ and $\mathcal{G}_2$ as before, we may estimate many of the terms forthwith:

$$W_2^j + W_6^j \leq \varepsilon\mathcal{G}_2(K)(E_w + E_w^{1/2}E_\mu^{1/2}), \tag{56}$$

$$W_3^j + W_5^j \leq \varepsilon\mathcal{G}_1(K)(E_w + E_w^{1/2}E_\mu^{1/2}). \tag{57}$$

We are left with one term to deal with more carefully, W_4^j. We first bound $\Theta_p(t,x,\mu^2,Dw^2)$ in L^∞ by $\mathcal{G}_2(K)$, finding

$$W_4^j \leq \varepsilon\mathcal{G}_2(K)\int_{\mathbb{T}^d}(\partial_{x_j}w^1 - \partial_{x_j}w^2)\partial_{x_j}(\mu^1-\mu^2)\ dx.$$

We next use Young's inequality, with $2\varepsilon\mathcal{G}_2(K)$ as the parameter, finding

$$W_4^j \leq \varepsilon^2\mathcal{G}_2(K)E_w + \frac{1}{4}\int_{\mathbb{T}^d}(\partial_{x_j}\mu^1 - \partial_{x_j}\mu^2)^2\ dx.$$

Adding contributions from (56) and (57) to this, we have

$$\sum_{\ell=2}^{6}\sum_{j=1}^{d}W_\ell^j \leq \varepsilon(\mathcal{G}_1(K)+\mathcal{G}_2(K))(E_w+E_\mu) + \frac{1}{4}\int_{\mathbb{T}^d}|D\mu^1 - D\mu^2|^2\ dx. \tag{58}$$

We are now ready to integrate with respect to time. Integrating $\frac{dE_\mu}{dt}$ over the interval $[0,t]$, we find

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

We then use the bound (55), finding

$$\begin{aligned} E_\mu(t) \leq E_\mu(0) &+ \varepsilon T(\mathcal{G}_1(K)+\mathcal{G}_2(K))(E_\mu + E_w) \\ &+ \int_0^t\left[V_1 + \frac{1}{4}\sum_{j=1}^{d}\int_{\mathbb{T}^d}(\partial_{x_j}(Dw^1 - Dw^2))^2\ dx\right]\ d\tau. \end{aligned}$$

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

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

We use the estimate (58), then, as follows:

$$E_w(t) \le E_w(T) + \varepsilon T(\mathcal{G}_1(K) + \mathcal{G}_2(K))(E_w + E_\mu) - \sum_{j=1}^{d} \int_t^T W_1^j \ d\tau \\ + \frac{1}{4}\int_t^T \int_{\mathbb{T}^d} |D\mu^1 - D\mu^2|^2 \ dxd\tau.$$

We apply the definitions of V_1 and W_1^j, and summarize what we have found thus far:

$$\begin{aligned} E_w(t) + E_\mu(t) + \int_0^t \|D\mu^1 - D\mu^2\|_0^2 \ d\tau + \int_t^T \|D^2w^1 - D^2w^2\|_0^2 \ d\tau \\ \le E_w(T) + E_\mu(0) + \varepsilon T(\mathcal{G}_1(K) + \mathcal{G}_2(K))(E_w(t) + E_\mu(t)) \\ + \frac{1}{4}\int_t^T \|D\mu^1 - D\mu^2\|_0^2 \ d\tau + \frac{1}{4}\int_0^t \|D^2w^1 - D^2w^2\|_0^2 \ d\tau. \end{aligned} \tag{59}$$

On the right-hand side, we bound the integrals with integrals over the entire time interval:

$$\begin{aligned} E_w(t) + E_\mu(t) + \int_0^t \|D\mu^1 - D\mu^2\|_0^2 \ d\tau + \int_t^T \|D^2w^1 - D^2w^2\|_0^2 \ d\tau \\ \le E_w(T) + E_\mu(0) + \varepsilon T(\mathcal{G}_1(K) + \mathcal{G}_2(K))(E_w(t) + E_\mu(t)) \\ + \frac{1}{4}\int_0^T \|D\mu^1 - D\mu^2\|_0^2 \ d\tau + \frac{1}{4}\int_0^T \|D^2w^1 - D^2w^2\|_0^2 \ d\tau. \end{aligned} \tag{60}$$

Isolating the first integral on the left-hand side of (60), we conclude from this that

$$\begin{aligned} \int_0^t \|D\mu^1 - D\mu^2\|_0^2 \ d\tau \le E_w(T) + E_\mu(0) + \varepsilon T(\mathcal{G}_1(K) + \mathcal{G}_2(K))(E_w(t) + E_\mu(t)) \\ + \frac{1}{4}\int_0^T \|D\mu^1 - D\mu^2\|_0^2 \ d\tau + \frac{1}{4}\int_0^T \|D^2w^1 - D^2w^2\|_0^2 \ d\tau. \end{aligned} \tag{61}$$

We take the supremum over time in both sides of (61), finding

$$\begin{aligned} \int_0^T \|D\mu^1 - D\mu^2\|_0^2 \ d\tau \le E_w(T) + E_\mu(0) \\ + \varepsilon T(\mathcal{G}_1(K) + \mathcal{G}_2(K))\left(\sup_{t\in[0,T]} (E_w(t) + E_\mu(t))\right) \\ + \frac{1}{4}\int_0^T \|D\mu^1 - D\mu^2\|_0^2 \ d\tau + \frac{1}{4}\int_0^T \|D^2w^1 - D^2w^2\|_0^2 \ d\tau. \end{aligned} \tag{62}$$

We similarly treat the other integral on the left-hand side of (60), finding

$$\begin{aligned}(63)\qquad \int_0^T \|D^2w^1 - D^2w^2\|_0^2\,d\tau &\leq E_w(T) + E_\mu(0)\\ &+ \varepsilon T(\mathcal{G}_1(K)+\mathcal{G}_2(K))\left(\sup_{t\in[0,T]}(E_w(t)+E_\mu(t))\right)\\ &+ \frac{1}{4}\int_0^T \|D\mu^1 - D\mu^2\|_0^2\,d\tau + \frac{1}{4}\int_0^T \|D^2w^1 - D^2w^2\|_0^2\,d\tau.\end{aligned}$$

We add (62) and (63) and rearrange the integrals, finding

$$\begin{aligned}&\int_0^T \|D\mu^1 - D\mu^2\|_0^2\,d\tau + \int_0^T \|D^2w^1 - D^2w^2\|_0^2\,d\tau\\ &\leq 2E_w(T) + 2E_\mu(0) + \varepsilon T(\mathcal{G}_1(K)+\mathcal{G}_2(K))\left(\sup_{t\in[0,T]}(E_w(t)+E_\mu(t))\right).\end{aligned}$$

Using this with (60), we have

$$\begin{aligned}(64)\qquad &\sup_{t\in[0,T]}(E_w(t)+E_\mu(t))\\ &\leq \frac{3}{2}E_w(T) + \frac{3}{2}E_\mu(0) + \varepsilon T(\mathcal{G}_1(K)+\mathcal{G}_2(K))\left(\sup_{t\in[0,T]}(E_w(t)+E_\mu(t))\right).\end{aligned}$$

If $\varepsilon T(\mathcal{G}_1(K)+\mathcal{G}_2(K)) < 1$, then we have established (local) uniqueness of solutions. For then, if $w^1(T,\cdot) = w^2(T,\cdot)$ and if $\mu^1(0,\cdot) = \mu^2(0,\cdot)$, then $E_w(T) = E_\mu(0) = 0$, and (64) implies

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

Thus $w^1 = w^2$ and $\mu^1 = \mu^2$. We have proved the following, our second main theorem:

Theorem 5. *Let $(u^1, \bar{m}+\mu^1)$ and $(u^2, \bar{m}+\mu^2)$ be two solutions of* (1), (2), (3)*, with the same data:*

$$m^1(0,\cdot) = m^2(0,\cdot), \qquad u^1(T,\cdot) = u^2(T,\cdot).$$

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 $s > 2 + \frac{d}{2}$. Assume **(H3)** *holds, and let $\mathcal{G}_1$ and $\mathcal{G}_2$ be as above. Assume $\varepsilon T(\mathcal{G}_1(K)+\mathcal{G}_2(K)) < 1$. Then $(u^1,\mu^1) = (u^2,\mu^2)$.*

Remark 5. *As we did in Remark 4 following the statement of our smallness condition for use in Theorem 4, we remark now on the smallness condition for Theorem 5. Clearly it may be satisfied by taking either ε or T sufficiently small, and we remarked on this similarly in Remark 4 for existence. It is also possible, for some Hamiltonians, that uniqueness follows in the case* *of small data. The function $\mathcal{G}_1(K)$ does go to zero as K goes to zero, but whether $\mathcal{G}_2(K)$ does as well depends on the Hamiltonian. For example, for the Hamiltonian $\mathcal{H} = m|Du|^2$, we have $\mathcal{H}_p = 2mDu$, and the Lipschitz constant for $\mathcal{H}_p$ then does not become small when the solution is small. For other choices of Hamiltonian, however, the Lipschitz constant may be small when the solution is, and for such Hamiltonians, one would gain from this theorem uniqueness of small solutions for larger values of ε and T.*

5. Discussion

Having proved our main existence and uniqueness theorems, we now close with some remarks.

We have carried out the above analysis for the planning problem, i.e. with boundary conditions (3). It is straightforward instead to prove both existence and uniqueness for the payoff problem instead, i.e. with boundary conditions (4). All that is required is a suitable hypothesis about the mapping properties of the payoff function G, but other than making such an assumption, there would be very little changed in the proofs.

Some other works have considered mean field games with congestion, in which the Hamiltonian has the particular form with a power of m in the denominator. The solutions proven to exist in [10], for example, are shown to always satisfy $m > 0$, then. Our analysis could be extended to such Hamiltonians when the data satisfies $m_0 > 0$. In this case, smallness constraints can again be used to keep m positive and to control the nonlinear evolution.

Future work to be done includes studying questions regarding regularity and uniqueness. For questions of regularity, this includes both loweing regularity requirements on the data to get results such as those in the present work, and also inferring still higher regularity of solutions such as those we have proved to exist. The uniqueness question also certainly deserves further attention; while we have presented a uniqueness theorem with a smallness constraint, as have Bardi and Cirant [4], as we have discussed in the introduction, Bardi and Fischer have given an example of non-uniquness [5]. Understanding when solutions are and are not unique and characterizing the mutiple possible solutions is an important problem to be studied.

Acknowledgments

The author gratefully acknowledges support from the National Science Foundation through grant DMS-1515849. The author is also grateful to the Institute for Computational and Experimental Research in Mathematics (ICERM) at Brown University, for hosting him while this paper was substantially completed. The author also is grateful to the Institute for Pure and Applied Mathematics (IPAM) at University of California, Los Angeles for hosting him while the article was finished.

References

[1] D.M. Ambrose. Well-posedness of vortex sheets with surface tension. *SIAM J. Math. Anal.*, 35(1):211–244 (electronic), 2003.

[2] D.M. Ambrose. Small strong solutions for time-dependent mean field games with local coupling. *C.R. Acad. Sci. Paris*, 354:589–594, 2016.

[3] D.M. Ambrose. Strong solutions for time-dependent mean field games with non-separable Hamiltonians. *J. Math. Pures Appl. (9)*, 113:141–154, 2018.

[4] M. Bardi and M. Cirant. Uniqueness of solutions in mean field games with several populations and Neumann conditions. 2017. Preprint. arXiv:1709.02158.

[5] M. Bardi and M. Fischer. On non-uniqueness and uniqueness of solutions in finite-horizon mean field games. 2017. Preprint. arXiv:1707.00628.

[6] J. Duchon and R. Robert. Global vortex sheet solutions of Euler equations in the plane. *J. Differential Equations*, 73(2):215–224, 1988.

[7] D.A. Gomes and E. Pimentel. Time-dependent mean-field games with logarithmic nonlinearities. *SIAM J. Math. Anal.*, 47(5):3798–3812, 2015.

[8] D.A. Gomes, E. Pimentel, and H. Sánchez-Morgado. Time-dependent mean-field games in the superquadratic case. *ESAIM Control Optim. Calc. Var.*, 22(2):562–580, 2016.

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

[10] D.A. Gomes and V.K. Voskanyan. Short-time existence of solutions for mean-field games with congestion. *J. Lond. Math. Soc. (2)*, 92(3):778–799, 2015.

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

[12] M. Huang, R.P. Malhamé, and P.E. Caines. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Commun. Inf. Syst.*, 6(3):221–251, 2006.

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

[14] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen II: Horizin fini et contrôle optimal. *C.R. Acad. Sci. Paris*, 343:679–684, 2006.

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

[16] A.J. Majda and A.L. Bertozzi. *Vorticity and incompressible flow*, volume 27 of *Cambridge Texts in Applied Mathematics*. Cambridge University Press, Cambridge, 2002.

[17] T. Milgrom and D.M. Ambrose. Temporal boundary value problems in interfacial fluid dynamics. *Applicable Analysis*, 92:922–948, 2013.

[18] B. Moll. Mean field games in economics, 2012. http://www.princeton.edu/∼moll/EconMFG.pdf. Accessed on April 4, 2016.

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

[20] A. Porretta. On the planning problem for the mean field games system. *Dyn. Games Appl.*, 4(2):231–256, 2014.

[21] A. Porretta. Weak solutions to Fokker-Planck equations and mean field games. *Arch. Ration. Mech. Anal.*, 216(1):1–62, 2015.

Drexel University, Department of Mathematics, Philadelphia, PA, USA
E-mail address: `dma68@drexel.edu`