Global Well-Posedness of Master Equations for Deterministic Displacement Convex Potential Mean Field Games

WILFRID GANGBO
University of California, Los Angeles

ALPÁR R. MÉSZÁROS
Durham University

Abstract

This manuscript constructs global in time solutions to *master equations* for potential mean field games. The study concerns a class of Lagrangians and initial data functions that are *displacement convex*, and so this property may be in dichotomy with the so-called Lasry–Lions monotonicity, widely considered in the literature. We construct solutions to both the scalar and vectorial master equations in potential mean field games, when the underlying space is the whole space $\mathbb{R}^d$, and so it is not compact. © 2022 The Authors. *Communications on Pure and Applied Mathematics* published by Wiley Periodicals LLC.

Introduction

In this manuscript, we study a Hamilton-Jacobi equation on $\mathscr{P}_2(\mathbb{R}^d)$, the set of Borel probability measures on $\mathbb{R}^d$ of finite second moments. This allows us to make inferences on the *master equation* in mean field games, introduced by P.-L.

Communications on Pure and Applied Mathematics, Vol. LXXV, 2685–2801 (2022)

Lions in [37]. Our study relies on a special notion of convexity, the so-called *displacement convexity*, which is natural for functions $\mathscr{V} : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$. It differs from the classical notion of convexity on the set of measures, which corresponds to the so-called *Lasry–Lions monotonicity* condition, central in most prior works aiming to study global-in-time solutions to the master equation. A comparison between the classical notion of convexity and displacement convexity can already be made by considering ways of interpolating Dirac masses. Given two Dirac masses δ_{q_0} and δ_{q_1}, the paths

$$[0,1] \ni t \mapsto \sigma_t := (1-t)\delta_{q_0} + t\delta_{q_1}, \quad [0,1] \ni t \mapsto \sigma_t^* := \delta_{(1-t)q_0+tq_1},$$

provide two distinct interpolations, these two elements of $\mathscr{P}_2(\mathbb{R}^d)$. The function $\mathscr{V}$ is called convex in the classical sense if it is convex along classical interpolation, which in particular implies $t \mapsto \mathscr{V}(\sigma_t)$ is a convex function on $[0,1]$. The function is called *displacement convex* [40] if its restriction to any W_2-geodesics is convex, which in particular means $t \mapsto \mathscr{V}(\sigma_t^*)$ is a convex function on $[0,1]$.

A blatant example which shows that convexity and displacement convexity cannot be the same is when

$$2\mathscr{V}(\mu) = \int_{\mathbb{R}^{2d}} |q-q'|^2 \mu(dq)\mu(dq'), \quad \mu \in \mathscr{P}_2(\mathbb{R}^d).$$

In this case, it has long been known that $\mathscr{V}$ is concave in the classical sense while $\mathscr{V}$ is obviously displacement convex. However, for the purpose of our study, we need to come up with a richer class of examples consistent with our analysis. For instance, let us consider two functions $\phi, \phi_1 \in C^2(\mathbb{R}^d)$ with bounded second derivatives and such that ϕ_1 is even and define

$$2\mathscr{V}(\mu) := \int_{\mathbb{R}^d} \big(2\phi(q) + (\phi_1 * \mu)(q)\big)\mu(dq), \quad \mu \in \mathscr{P}_2(\mathbb{R}^d).$$

Let us recall that (see Lemma B.2) the function $\mathscr{V}$ is convex in the classical sense if and only if $\widehat{\phi}_1$— the Fourier transform of ϕ_1—is nonnegative, independently of whether or not additional requirements are imposed on ϕ. Suppose for instance that ϕ is 2λ-convex for some $\lambda > 0$. If ϕ_1 is λ_1-convex for some $2\lambda_1 \in (-\lambda, \lambda)$, then $\mathscr{V}$ is displacement convex. As discussed in Sections 4.3 and B.1, we can choose ϕ_1 such that $\widehat{\phi}_1$ changes sign, so that $\mathscr{V}$ fails to be convex in the classical sense.

The theory of well-posedness of the master equation in mean field games is well developed on the set of probability measures [14] (for a probabilistic approach to study such equations, we refer the reader to [17]) under the Lasry-Lions monotonicity condition [13, 35, 36, 38] for games where the individual and/or common noises are essential mechanisms governing the games. In the same setting of monotone data, global solutions were also constructed in [19], where the authors can handle even degenerate diffusions in the equations. In the same context, [42] improves the regularity restrictions on the data, which need to be still monotone, and propose a notion of weak solutions for the master equation. When the monotonicity condition fails (even in the presence of the noise), only short time existence results for the scalar master equation were achieved (in the deterministic case, we refer to [10, 31, 39]; in the presence of noise, we refer to [17, 19]). For classical mean field games systems, the smallness of the time horizon sometimes can be replaced by a smallness condition on the data (see, for instance, [3, 4]). Via a "lifting procedure", it is possible to study master equations on a Hilbert space of square integrable random variables. The main benefit of this process is to instead use the more familiar Fréchet derivatives on flat spaces and bypass the differential calculus on the space of probability measures, which is a curved infinite-dimensional manifold. Such analyses were carried out for a special class of mechanical Lagrangians and for potential games, either in a deterministic setting [9] or in the presence of individual noise in [7, 8]. Furthermore, the authors needed to impose higher than second-order Fréchet differentiability on the data functions. It turns out (see below) that this may sometimes be a too severe restriction. Therefore, from this point of view the Hilbert space approach has a serious drawback.

This manuscript constructs global solutions to potential mean field games master equations, where the widely used Lasry–Lions monotonicity condition is replaced by *displacement convexity*, a concept which appeared in optimal transport theory in the early 1990s. The use of displacement convexity in mean field control problems and mean field games goes back to [15], where the authors study control problems of McKean–Vlasov type. In the case of mean field game systems with common noise, we refer the reader to [1, 2], where their so-called *weak monotonicity condition*, assumed on the data, is equivalent to displacement convexity in the potential game case. As mentioned before, in [7, 8], this condition is used in the Hilbertian setting. In the study of master equations arising in control problems of McKean-Vlasov type (in the presence of individual noise), [19] seems to be the first work in the literature that imposed displacement convexity on their data to obtain well-posedness of a master equation in the spirit of [15].

In potential mean field games, one considers smooth enough real-valued functions $\mathscr{U}_0, \mathscr{F}$ defined on $\mathscr{P}_2(\mathbb{R}^d)$. We assume that there are smooth real-valued functions u_0, f defined on $\mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d)$ that are related to $\mathscr{U}_0, \mathscr{F}$ in the sense that the Wasserstein gradient of $\mathscr{U}_0$ at $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ equals the finite-dimensional gradient $D_q u_0(\cdot, \mu)$ and the Wasserstein gradient of $\mathscr{F}$ at $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ equals the finite-dimensional gradient $D_q f(\cdot, \mu)$. Given a Hamiltonian $H \in C^3(\mathbb{R}^{2d})$ and a time horizon $T > 0$, the master equation consists in finding a real-valued function u defined on $[0, T) \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d)$, a solution to the nonlocal equation

$$\begin{cases} \partial_t u + H(q, D_q u) + \mathscr{N}_\mu\big[D_q u(t, \cdot, \mu), \quad \nabla_w u(t, q, \mu)(\cdot)\big] = f(x, \mu), \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{in } (0, T) \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d), \\ u(0, \cdot, \cdot) = u_0, \qquad\qquad\qquad\qquad\qquad\quad \text{in } \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d). \end{cases}$$

Here, $\mathscr{N}_\mu : L^2(\mu) \times L^2(\mu) \to \mathbb{R}$ is the nonlocal operator defined as

$$\mathscr{N}_\mu[\eta, \theta] := \int_{\mathbb{R}^d} D_p H(c, \eta(c)) \cdot \theta(c)\mu(dc). \tag{0.1}$$

Let $L(q, \cdot)$ be the Legendre transform of $H(q, \cdot)$ and assume L is strictly convex, and both functions have bounded second-order derivatives. Under the assumption that $\mathscr{U}_0$ and $\mathscr{F}$ are displacement convex (convex along the Wasserstein geodesics), we construct classical solutions and weak solutions to the master equation, depending on the regularity properties imposed on the data. Following [32], the starting point of our study relies on the point of view that the differential structure on $(\mathscr{P}_2(\mathbb{R}^d), W_2)$ is inherited from the differential structure on the flat space $\mathbb{H} := L^2((0,1)^d; \mathbb{R}^d)$, and the former space can be viewed as the quotient space of the latter. The functions $\mathscr{U}_0, \mathscr{F}$ are lifted to obtain functions $\widetilde{\mathscr{U}_0}, \widetilde{\mathscr{F}}$ defined on the Hilbert space $\mathbb{H}$, with the property that they are rearrangement invariant. What we mean by rearrangement invariant is that $\widetilde{\mathscr{U}_0}(x) = \widetilde{\mathscr{U}_0}(y)$ whenever the push forward of Lebesgue measure restricted to $(0,1)^d$ by $x, y \in \mathbb{H}$ coincide. In this case, we sometimes say that x and y have the same law. The Hamiltonian H is used to define on the cotangent bundle $\mathbb{H}^2$, another Hamiltonian denoted

$$\widetilde{\mathscr{H}}(x, b) := \int_{(0,1)^d} H(x(\omega), b(\omega))d\omega - \widetilde{\mathscr{F}}(x).$$

The corresponding Lagrangian $\widetilde{\mathscr{L}}$ is on $\mathbb{H}^2$, the tangent bundle, and is

$$\widetilde{\mathscr{L}}(x, a) := \int_{(0,1)^d} L(x(\omega), a(\omega))d\omega + \widetilde{\mathscr{F}}(x).$$

Both the Lagrangian and the Hamiltonian are invariant under the action of the group of bijections of $(0,1)^d$ onto $(0,1)^d$, which preserve the Lebesgue measure. We are interested in regularity properties of $\widetilde{\mathscr{U}} : (0,\infty) \times \mathbb{H} \to \mathbb{R}$, solutions to the Hamilton–Jacobi equation

$$\begin{cases} \partial_t \widetilde{\mathscr{U}} + \widetilde{\mathscr{H}}\big(\cdot, \nabla_x \widetilde{\mathscr{U}}\big) = 0 & \text{in } (0,\infty) \times \mathbb{H}, \\ \widetilde{\mathscr{U}}(0, \cdot) = \widetilde{\mathscr{U}_0} & \text{on } \mathbb{H}. \end{cases}$$

The characteristics of this infinite-dimensional PDE and the smoothness properties of $\widetilde{\mathscr{U}}$ will play an essential role in the application of our study to mean field games. They allow us to obtain an explicit representation formula of the solution to the master equation for arbitrarily large times. Similar observations were made also by P.-L. Lions during a recorded seminar talk [37]. This lecture seems to suggest that is was not clear at all how far the displacement convexity assumptions on the data could be used to advance the study of the global-in-time well-posedness of master equations.

Under appropriate growth and convexity conditions on the data, the classical theory of Hamilton–Jacobi equations on Hilbert spaces ensures that $\widetilde{\mathscr{U}}(t, \cdot)$ is of class $C^{1,1}_{\text{loc}}(\mathbb{H})$. Our Hamiltonian and Lagrangian being rearrangement invariant, by the uniqueness theory of Hamilton–Jacobi equation, $\widetilde{\mathscr{U}}(t,\cdot)$ is rearrangement invariant. This allows us to define a function $\mathscr{U}(t,\cdot)$ on $\mathscr{P}_2(\mathbb{R}^d)$ such that $\mathscr{U}(t,\mu) = \widetilde{\mathscr{U}}(t,x)$ whenever $x \in \mathbb{H}$ has μ as its law. At the same time, $\mathscr{U}$ will be the unique classical solution to the corresponding Hamilton–Jacobi equation set on $\mathscr{P}_2(\mathbb{R}^d)$.

By Lemma 3.11, a function $\mathscr{V} : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is of class $C^{1,1}_{\text{loc}}$ on the Wasserstein space if and only if its lift $\widetilde{\mathscr{V}} : \mathbb{H} \to \mathbb{R}$ is of class $C^{1,1}_{\text{loc}}$ on the Hilbert space. Since the Hilbert space theory ensures that $\widetilde{\mathscr{U}}(t,\cdot)$ is of class $C^{1,1}_{\text{loc}}$ on the Hilbert space, we obtain as a by-product that $\mathscr{U}(t,\cdot)$ is of class $C^{1,1}_{\text{loc}}$ on the Wasserstein space. This is how far one could push the Hilbert approach in terms of regularity theory if one would like to make useful inference in mean field games. Indeed, imposing that a rearrangement invariant function $\widetilde{\mathscr{V}} : \mathbb{H} \to \mathbb{R}$ is of class C^2 (twice Fréchet differentiable) is too stringent for the purpose of mean field games. For instance, if $\phi \in C^\infty_c(\mathbb{R}^d)$, unless $\phi \equiv 0$, the function $\widetilde{\mathscr{V}}$ defined on $\mathbb{H}$ by

$$\widetilde{\mathscr{V}}(x) := \int_{(0,1)^d} \phi(x(\omega))d\omega,$$

does not belong to $C^2(\mathbb{H})$ (cf. Proposition A.4). The reader should compare this to another subtlety in [11, sec. 2]. Similar conclusions can be drawn on other functionals with a local representation such as

$$\mathbb{H} \ni x \mapsto \widetilde{\mathscr{V}}(x) := \int_{(0,1)^{nd}} \phi(x(\omega_1),\dots,x(\omega_n))d\omega_1\cdots d\omega_n,$$

when $\phi \in C^3(\mathbb{R}^{nd})$ is symmetric and has bounded second- and third-order derivatives (cf. Proposition A.2). Pursuing a deeper analysis, we assume $\alpha \in (0,1]$, $\widetilde{\mathscr{V}} \in C^{2,\alpha}_{\text{loc}}\big(\mathbb{H}\big)$ is rearrangement invariant so that it is the lift of a function $\mathscr{V} : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$. We show in Lemma A.1 that if (A.1) holds for all $h, h_* \in \mathbb{H}$, then $D_q\big(\nabla_w \mathscr{V}(\mu)\big)$ is a constant function on $\text{spt}(\mu)$.

A final argument to support the fact that we need a new concept of higher-order derivatives on the set of probability measures is the following: When $k \geq 3$, making assumptions on k-order differentials of Hamiltonians $\widetilde{\mathscr{H}} : \mathbb{H}^2 \to \mathbb{R}$ and treating them as continuous multilinear forms on Cartesian products of $\mathbb{H}^2$ is too restrictive for a theory in mean field games. Indeed, frequently used Hamiltonians in mean field games theory are of the form

$$\widetilde{\mathscr{H}}(x,b) = \widetilde{\mathscr{H}}_H(x,b) - \widetilde{\mathscr{F}}(x), \quad \widetilde{\mathscr{H}}_H(x,b) \equiv \int_{(0,1)^d} H(x(\omega),b(\omega))d\omega,$$

where $H \in C^3(\mathbb{R}^{2d})$ is such that D^2H is bounded. Let $\alpha \in (0,1]$. Even if $C^{2,\alpha}_{\text{loc}}(\mathbb{H}^2)$ is an infinite-dimensional space, its intersection with the set of functions that have a local representation is contained in a finite-dimensional space. For instance,

$$\dim\Big(C^{2,\alpha}_{\mathrm{loc}}(\mathbb{H}^2) \cap \big\{\widetilde{\mathscr{H}}_H : H \in C^{2,\alpha}_{\mathrm{loc}}(\mathbb{R}^{2d}),\ D^2 H \text{ is bounded}\big\}\Big) < \infty. \tag{0.2}$$

In this manuscript, to write a meaningful master equation, we are interested in functions $\mathscr{V} : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ that satisfy higher regularity properties than being of $C^{1,1}_{\mathrm{loc}}$. We assume at least that their lifts $\widetilde{\mathscr{V}} : \mathbb{H} \to \mathbb{R}$ are such that $\nabla\mathscr{V}$ is Gâteaux differentiable with bounded second-order differential in a sense to be made precise. Due to the rearrangement invariance property of $\widetilde{\mathscr{V}}$, $\nabla^2\widetilde{\mathscr{V}}$ must have a special form. Given $x \in \mathbb{H}$, there exist matrix-valued maps

$$A^*_{12} \in L^\infty((0,1)^d; \mathbb{R}^{d\times d}), \quad A^*_{22} \in L^\infty((0,1)^{2d}; \mathbb{R}^{d\times d})$$

such that A^*_{12} is symmetric almost everywhere, $A^*_{22}(\omega, o) = A^*_{22}(o,\omega)^\top$ almost everywhere and the operator $\mathbb{H} \ni \zeta \mapsto \nabla^2\widetilde{\mathscr{V}}(x)\zeta$ can be written as

$$\big(\nabla^2\widetilde{\mathscr{V}}(x)\zeta\big)(\omega) = A^*_{12}(\omega)\zeta(\omega) + \int_{(0,1)^d} A^*_{22}(\omega, o)\zeta(o)\,do. \tag{0.3}$$

In fact, as observed in [11] (cf. also [14, 16, 17, 19, 21]), there exists a matrix field A_{12} defined on $R(x)$, the range of x, and a matrix field A_{22} defined on $R(x)\times R(x)$ such that the following factorization holds:

$$A^*_{12}(\omega) = A_{12}\big(x(\omega)\big), \quad A^*_{22}(\omega, o) = A_{22}\big(x(\omega), x(o)\big).$$

We argue in Remark 3.14 that A_{12} can be interpreted as $D_q\big(\nabla_w\mathscr{V}(\mu)(q)\big)$ and indicate the relation between A_{22} and the Wasserstein gradient of $\nabla_w\mathscr{V}$.

When $\mathscr{B} \subseteq \mathscr{P}_2(\mathbb{R}^d)$ is an open set, we introduce vector spaces of functions $C^{2,\alpha,w}(\mathscr{B})$ as substitutes for the spaces $C^{2,\alpha}(\mathbb{H})$. These spaces are such that whenever $\mathscr{V} \in C^{2,\alpha,w}(\mathscr{B})$, its restrictions

$$\mathbb{R}^{md} \ni (q_1, \dots, q_m) \mapsto \mathscr{V}\left(\frac{1}{m}\sum_{i=1}^m \delta_{q_i}\right)$$

belong to $C^{2,\alpha}_{\mathrm{loc}}(\mathbb{R}^{md})$. The precise definition of this space can be found in Definition 3.13. At least we require that if $\mathscr{V} \in C^{2,\alpha,w}(\mathscr{B})$, since the second-order Gâteaux differential of its lift $\widetilde{\mathscr{V}}$ exists, it must satisfy the property

$$\begin{aligned}\Big|\nabla\widetilde{\mathscr{V}}(y)(\omega) - \nabla\widetilde{\mathscr{V}}(x)(\omega) - \nabla^2\widetilde{\mathscr{V}}(x)(\omega)\big((y(\omega) - x(\omega)\big)\Big| \\ \le C\big(|y(\omega) - x(\omega)|^\alpha + \|x-y\|^\alpha\big)\end{aligned} \tag{0.4}$$

whenever $x, y \in \mathbb{H}$, x pushes $\mathscr{L}^d_{(0,1)^d}$ forward to μ, y pushes $\mathscr{L}^d_{(0,1)^d}$ forward to ν, and $\|x - y\| = W_2(\mu,\nu)$. In fact, spaces of type $C^{2,1}(\mathscr{P}_2(\mathbb{M}))$ have already been considered in the framework of mean field models in [11], based on a construction very similar to ours in Definition 3.13.

A discretization approach (which consists in restricting our study to the subsets of $\mathscr{P}_2(\mathbb{R}^d)$ that are averages of Dirac masses) greatly facilitates the task to show (0.3), with $\widetilde{\mathscr{V}}$ replaced by the solution to the Hamilton–Jacobi equation we constructed on the Hilbert space. This helps us show that

$$A_{12} \in L^\infty(R(x); \mathbb{R}^{d\times d}), \quad A_{22} \in L^\infty(R(x) \times R(x); \mathbb{R}^{d\times d})$$

and for $\varphi \in C_c^\infty(\mathbb{R}^d)$ and $h := D\varphi \circ x$,

$$D^2\widetilde{\mathscr{V}}(x)(h,h) = \int_{(0,1)^d} A_{12}(x(\omega))h(\omega)\cdot h(\omega)d\omega + \int_{(0,1)^{2d}} A_{22}(x(\omega_1), x(\omega_2))h(\omega_1)\cdot h(\omega_2)d\omega_1 d\omega_2.$$

This allows us to make inference beyond an estimate such as

$$\sup_{x,h\in\mathbb{H}} \big\{|D^2\widetilde{\mathscr{U}}(t,x)(h,h)| : \|h\| \leq 1,\ \|x\| \leq r\big\} < +\infty \quad \forall r > 0.$$

Unlike studies of the master equation in compact settings such as the periodic setting $\mathbb{R}^d/\mathbb{Z}^d$, the fact that the range of $\widetilde{\mathscr{U}}$ is certainly unbounded is a source of additional complications in our study,

When $\nabla\widetilde{\mathscr{H}}$ is Lipschitz, the characteristics of the Hamilton–Jacobi equation are the Hamiltonian flow $\Sigma = (\Sigma^1, \Sigma^2) : [0,\infty) \times \mathbb{H}^2 \to \mathbb{H}^2$, uniquely defined by the solution of

$$\begin{cases} \dot\Sigma^1(t,\cdot) = \ \ \nabla_b\widetilde{\mathscr{H}}\big(\Sigma(t,\cdot)\big), & \text{in } (0,\infty)\times\mathbb{H}^2, \\ \dot\Sigma^2(t,\cdot) = -\nabla_x\widetilde{\mathscr{H}}\big(\Sigma(t,\cdot)\big) & \text{in } (0,\infty)\times\mathbb{H}^2, \\ \Sigma(0,\cdot) = \mathrm{id}_{\mathbb{H}^2}. \end{cases} \tag{0.5}$$

The vector field $\nabla^\perp\widetilde{\mathscr{H}}$ is the velocity in Eulerian coordinates for the trajectory Σ on the cotangent bundle $\mathbb{H}^2$. We denote as

$$(\tilde\xi, \tilde\eta) : [0,\infty) \times \mathbb{H} \to \mathbb{H}^2$$

the restriction of Σ to the graph of $\nabla\widetilde{\mathscr{U}_0}$, i.e.,

$$(\tilde\xi, \tilde\eta) := \Sigma\big(\cdot, \cdot, \nabla\widetilde{\mathscr{U}_0}\big). \tag{0.6}$$

When $\widetilde{\mathscr{L}}$ and $\widetilde{\mathscr{U}_0}$ are convex, under appropriate standard conditions on $\widetilde{\mathscr{L}}$ and $\widetilde{\mathscr{H}}$, differentiability properties of $\widetilde{\mathscr{U}}$ are obtained by standard methods. A strict convexity property of $\widetilde{\mathscr{L}}$ ensures that for any fixed $t \geq 0$, $\tilde\xi(t,\cdot)$ is a bijection of $\mathbb{H}$ onto $\mathbb{H}$. The trajectories

$$[0,t] \ni s \mapsto \tilde S_s^t[x] := \tilde\xi\big(s, \tilde\xi^{-1}(t,x)\big) \in \mathbb{H}$$

are useful to write the representation formula

$$\widetilde{\mathscr{U}}(t,x) = \widetilde{\mathscr{U}_0}(\tilde S_0^t[x]) + \int_0^t \widetilde{\mathscr{L}}\big(\tilde S_s^t[x], \partial_s \tilde S_s^t[x]\big)ds.$$

The identity

$$\nabla\widetilde{\mathscr{U}}(t,\cdot) = \tilde\eta\big(t, \tilde S_0^t\big) \tag{0.7}$$

suggests that the smoothness properties of $\widetilde{\mathscr{U}}$ rest on the smoothness properties of $\widetilde{S}_0^t$ and $\widetilde{\eta}$. While strict convexity of $\widetilde{\mathscr{L}}$ is sufficient to get that the restriction of $\widetilde{\xi}(t,\cdot)^{-1}$ to appropriate finite-dimensional spaces is continuously differentiable, it becomes much harder to show that $\widetilde{\xi}(t,\cdot)^{-1}$ is continuous on the whole space $\mathbb{H}$ unless appropriate convexity properties are imposed on the data.

Let us consider the vector field

$$B(t,\cdot) := \nabla_b \mathscr{H}\big(\cdot, \widetilde{\eta}(t, \widetilde{S}_0^t)\big),$$

which helps to study the second-order derivatives of $\widetilde{\mathscr{U}}$ and which represents the velocity of the flow $\widetilde{\xi}$ in physical space, since $\dot{\widetilde{\xi}} = B(s, \widetilde{\xi})$. When $\widetilde{\mathscr{U}}(t,\cdot)$ is twice differentiable then $\nabla^2 \widetilde{\mathscr{U}}(t,x), \nabla B(t,x) : \mathbb{H}^2 \to \mathbb{R}$ are bilinear forms which satisfy the relation

$$\nabla B(t,x)(h,a) = \nabla^2 \widetilde{\mathscr{U}}(t,x)\Big(a, D^2_{pp} H\big(x, \nabla \widetilde{\mathscr{U}}(t,x)\big)h\Big) + \int_{(0,1)^d} \big(D^2_{qp} H(x, \nabla \widetilde{\mathscr{U}}(t,x))a\big)\cdot h\, d\omega, \quad (\forall h, a \in \mathbb{H}).$$

Summary of our main results

Coming back to the description of our main results, after having provided the $C^{1,1}_{\mathrm{loc}}$ regularity for the viscosity solutions $\mathscr{U}$ to the corresponding Hamilton–Jacobi equations on $\mathscr{P}_2(\mathbb{R}^d)$, we completely abandon the setting of the Hilbert space and via the mentioned discretization approach we show that $\mathscr{U}(t,\cdot)$ is actually of class $C^{2,1,w}_{\mathrm{loc}}$. We note that our approach seems to be novel and, although similar in flavor, it is completely different from the ones developed in [31, 39]. It relies on fine quantitative derivative estimates with respect to $m \in \mathbb{N}$ on the Hamiltonian flow for m-particles, then these in turn translate to higher-regularity estimates on $\mathscr{U}$ by carefully differentiating the identity (0.7), written for the restriction of $\mathscr{U}$ to the set of averages of Dirac masses. Let us emphasize that this finite-dimensional projection of the value function solves the corresponding optimization problem but driven by the finite-dimensional projections of the cost coefficients (see Remark 1.4); this is in fact what allows for a preliminary analysis of the optimal trajectories of the mean field control problem when restricting initial states of the population to uniform finite distributions. A key point is then to obtain regularity estimates that are independent of the cardinality of those finite distributions. This is one crucial step where the convexity structure plays a key role. This idea is in fact the heart of our analysis and works only for deterministic mean field games; the approach in this manuscript is entirely different from the existing ones to tackle mean field games master equations: most of them consist in working directly at the level of PDE system of mean field games.

Having $\mathscr{U}(t,\cdot) \in C^{2,1,w}_{\mathrm{loc}}(\mathscr{P}_2(\mathbb{R}^d))$ allows us to obtain *weak solutions* (see in Theorem 4.4) $\mathscr{V} : [0,T] \times \mathscr{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ to the so-called *vectorial master* *equation*,

$$(0.8)\qquad \begin{cases} \partial_t \mathscr{V} + D_q H(q, \mathscr{V}(t,\mu,q)) + D_q \mathscr{V}(t,\mu,q) \nabla_p H(q, \mathscr{V}(t,\mu,q)) \\ + \mathscr{N}_\mu\big[\mathscr{V}, \nabla_w^\top \mathscr{V}\big](t,\mu,q) = \nabla_w \mathscr{F}(\mu)(q), \\ \mathscr{V}(0,\mu,\cdot) = \nabla_w \mathscr{U}_0(\mu)(\cdot), \end{cases}$$

where for $\mathscr{V} : \mathscr{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ we define

$$\mathscr{N}_\mu\big[\mathscr{V}, \nabla_w^\top \mathscr{V}\big](t,\mu,q) := \int_{\mathbb{R}^d} \nabla_w^\top \mathscr{V}(t,\mu,q)(b) D_p H\big(b, \mathscr{V}(t,\mu,b)\big)\mu(db).$$

This equation can be seen as a vectorial conservation law on $(0,T) \times \mathscr{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$ and be derived formally by taking the Wasserstein gradient of the Hamilton-Jacobi equation satisfied by $\mathscr{U}$. Such a method is possible in the setting of the Hilbert space as well (provided one has the sufficient regularity to justify the differentiation), and this is done for instance in [7,9] for short time and special Hamiltonians. Let us emphasize that there is a subtlety in this derivation and in particular at a first glance the vectorial master equation in the setting of $\mathscr{P}_2(\mathbb{R}^d)$ is satisfied pointwise only on $(0,T) \times \bigcup_{\mu \in \mathscr{P}_2(\mathbb{R}^d)} \{\mu\} \times \mathrm{spt}(\mu)$. Therefore, we refer to such a solution as a *weak solution*. Thus, additional effort is needed to extend the vectorial master equation to $(0,T) \times \mathscr{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$, and actually, this is possible through the solution to the scalar master equation. One cannot observe this phenomenon in the setting of $\mathbb{H}$, because $\nabla \widetilde{\mathscr{U}}(t,x)$, as an element of $\mathbb{H}$, does not carry explicitly the dependence on the range of $x \in \mathbb{H}$.

Let us stress that even though there is a deep connection between the vectorial and scalar master equations, while formally speaking the former one is the Wasserstein gradient of a Hamilton–Jacobi equation, additional effort is needed to justify the well-posedness of the latter one. In particular, this is not a simple consequence of the well-posedness of the vectorial equation at all. In the same time, while the vectorial master equation might have physical relevance as a vectorial conservation law, in the theory of mean field games the scalar master equation is the one that has profound significance. One of the reasons for this is that this equation deeply carries the features of m-player differential games. In particular, as we can see this in [14], it provides an important tool to prove the convergence of Nash equilibria of m-player differential games to the mean field games system as $m \to +\infty$. At the same time, typically it provides quantified rates on propagation of chaos. Therefore, such equations are very natural, and they were successfully used in the literature in the context of mean field limits of a large particle system (see, for instance, in [20, 41]).

The candidate for the solution of the scalar master equation is constructed as follows. Given $t \in [0, T]$, $q \in \mathbb{R}^d$, and $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, we define

$$u(t, q, \mu) := \inf_{\gamma} \Big\{ u_0(\gamma_0, \sigma_0^t[\mu]) + \int_0^t \big(L(\gamma_s, \dot{\gamma}_s) + f(\gamma_s, \sigma_s^t[\mu]) \big) ds, \quad \gamma \in W^{1,2}([0,t], \mathbb{R}^d), \gamma_t = q \Big\}, \tag{0.9}$$

where the curve $(\sigma_s^t[\mu])_{s\in[0,t]}$ is the projection onto $\mathscr{P}_2(\mathbb{R}^d)$ of the Hamiltonian flow. We underline the important fact that the previous formula defines $u(t, \cdot, \mu)$ for every $q \in \mathbb{R}^d$ (and not just for $q \in \operatorname{spt}(\mu)$).

After obtaining the sufficient regularity of the mapping $\mu \mapsto \sigma_s^t[\mu]$ (using also the fact that $\mathscr{U}(t, \cdot) \in C^{2,1,w}_{\mathrm{loc}}(\mathscr{P}_2(\mathbb{R}^d))$), we show that u is of class $C^{1,1}_{\mathrm{loc}}([0, T] \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d))$ (see Lemma 4.13). The connection between u and $\mathscr{U}$ is that

$$D_q u(t, \cdot, \mu) = \nabla_w \mathscr{U}(t, \mu)(\cdot) \quad \text{on } \operatorname{spt}(\mu). \tag{0.10}$$

This is an important remark, since it means that $D_q u(t, \cdot, \mu)$ provides the natural Lipschitz-continuous extension for $\nabla_w \mathscr{U}(t, \mu)(\cdot)$ to $\mathbb{R}^d$. By these arguments we can prove Theorem 4.19, the main theorem of this manuscript, which states that under our standing assumptions u defined in (0.9) is the unique classical solution to the scalar master equation which is of class $C^{1,1}_{\mathrm{loc}}([0, T] \times \mathbb{R}^d \times \mathscr{P}_2(\mathbb{R}^d))$.

Theorem 4.19 has several implications. First, the obtained regularity of u and (0.10) allow us to deduce that $D_q u$ is a solution to the vectorial master equation and (0.8) is satisfied for all $(t, \mu) \in (0, T) \times \mathscr{P}_2(\mathbb{R}^d)$ and for $\mathscr{L}^d$-a.e. $q \in \mathbb{R}^d$. Second, since the scalar master equation, and in particular our definition (0.9) possesses the features of m-player differential games, we could easily deduce that $u(t, \cdot, \cdot)$, when restricted to $\bigcup_{q\in\mathbb{R}^{md}} \mu_q^{(m)} \times \operatorname{spt}(\mu_q^{(m)})$, provides approximate solutions to a system of Hamilton–Jacobi equations, characterizing the Nash equilibria of the associated m-player differential game (such a construction would be similar to the ones in [14, 24, 25], so we omit the details on this). At the same time, the regularity of u would allow us to deduce the local convergence of Nash equilibria as $m \to +\infty$, provided we know that the m-player Nash system of Hamilton–Jacobi equations has a smooth enough classical solution. In such a fortunate scenario, the proof of this result, even in the deterministic setting, would follow similar ideas as the ones in [14], [24, 25]. However, let us emphasize that the well-posedness question of systems of Hamilton–Jacobi equations in the deterministic setting is not a settled issue in the literature. It's worth mentioning the recent work [28], which studies this convergence question in the deterministic setting in a suitable weak sense, without relying on the well-posedness of either the Nash system or the master equation.

The structure of the rest of the paper is the following. In Section 1 we provide the first part of our standing assumptions, and we present the discretization approach and show a direct argument that provides $C^{1,1}_{\text{loc}}$ regularity for solutions to a class of Hamilton-Jacobi equations set on Hilbert spaces.

Section 2 contains the important quantitative estimates with respect to m on the Hamiltonian flows of m-particle systems and the corresponding derivative estimates of the solutions to Hamilton-Jacobi equations set on $\mathbb{R}^{md}$.

In Section 3 we compare notions of convexity and regularity for functions defined on $\mathscr{P}_2(\mathbb{R}^d)$, their lifts defined on $\mathbb{H}$, and their restrictions to discrete measures. Here we also show how can we deduce regularity estimates for functions on $\mathscr{P}_2(\mathbb{R}^d)$ from precise quantitative derivative estimates on their restrictions to discrete measures.

Section 4 is the core of the manuscript, where we investigate the well-posedness of both vectorial and scalar master equations. Additional assumptions need to be imposed to establish the well-posedness of the scalar master equation. These are listed in this section.

In Section 5 we have collected an important implication of the scalar master equation. We use scalar master equations to improve the notion of weak solution for the vectorial equations.

To facilitate the reading of the main text, our manuscript has several appendices. In Appendix A we demonstrate the limitations of the Hilbert space approach, when studying or assuming $C^{2,\alpha}$ type regularity on rearrangement invariant functionals having local representations.

In Appendix B we emphasize how our setting by imposing displacement convexity of the data can replace the more standard, so-called Lasry–Lions monotonicity assumptions imposed typically in the mean field games literature. Here we provide examples of functionals which produce nonmonotone coupling functions in the Lasry–Lions sense and an example of a Hamilton–Jacobi equation on $\mathscr{P}_2(\mathbb{R}^d)$, for which the data have this standard monotonicity condition, yet its classical solution ceases to exist after finite time.

In Appendix C we have collected some standard results on Hamiltonian flows on Hilbert spaces, and we explain how the regularity of these flows can be used to show regularity of solutions to a Hamilton–Jacobi equations.

1 Preliminaries

We start this section with some well-known definitions in the Hilbert setting as well as in the Wasserstein space. We denote by $\Omega := (0,1)^d \subset \mathbb{R}^d$ the unit cube and as $\mathscr{L}^d_\Omega$ the Lebesgue measure restricted to Ω. We sometimes refer to any Borel map of Ω to $\mathbb{M}$ as a random variable. We shall work on the Hilbert space

$$\mathbb{H} := L^2(\Omega; \mathbb{R}^d),$$

the set of square-integrable Borel vector fields with respect to Ω.

Since it is more convenient to write $\mathbb{M}^m$ instead of $(\mathbb{R}^d)^m$, we shall write $\mathbb{M}$ in place of $\mathbb{R}^d$. Letters x, y are typically used for elements of $\mathbb{H}$, while elements of

$\mathbb{M}$ are typically denoted by q, p, v. Sometimes, we also use the notation $\mathbb{R}_+ := [0, +\infty)$.

Given two topological spaces $\mathbb{S}_1$ and $\mathbb{S}_2$, a Borel measure μ on $\mathbb{S}_1$, and a Borel map $X : \mathbb{S}_1 \to \mathbb{S}_2$, $X_\sharp\mu$ is the measure on $\mathbb{S}_2$ defined as $X_\sharp\mu(B) = \mu(X^{-1}(B))$ for $B \subset \mathbb{S}_2$.

The canonical projections $\pi^1, \pi^2 : \mathbb{M} \times \mathbb{M} \to \mathbb{M}$ are defined as

$$\pi^1(q_1, q_2) = q_1, \quad \pi^2(q_1, q_2) = q_2 \quad \forall q_1, q_2 \in \mathbb{M}.$$

Given $\mu_0, \mu_1 \in \mathscr{P}_2(\mathbb{M})$, we denote as $\Gamma(\mu_0, \mu_1)$ the set of Borel probability measures γ on $\mathbb{M} \times \mathbb{M}$ such that $\pi^1_\sharp\gamma = \mu_0$ and $\pi^2_\sharp\gamma = \mu_1$. We denote as $\Gamma_o(\mu_0, \mu_1)$ the set of $\gamma \in \Gamma(\mu_0, \mu_1)$ such that

$$W_2^2(\mu_0, \mu_1) = \int_{\mathbb{R}^{2d}} |q_1 - q_2|^2 \gamma(dq_1, dq_2).$$

The law of $x \in \mathbb{H}$ is the Borel probability measure $\sharp(x) := x_\sharp \mathscr{L}^d_\Omega$. The map $\sharp$ maps $\mathbb{H}$ onto $\mathscr{P}_2(\mathbb{M})$, the set of Borel probability measure on $\mathbb{M}$ of finite second moments. One basic result in measure theory is that as Ω has no atoms, any Borel probability measure on $\mathbb{R}^d$ is the law of a Borel map $z : \Omega \to \mathbb{R}^d$.

If $\mu \in \mathscr{P}_2(\mathbb{M})$, the set of Borel vector fields $\xi : \mathbb{M} \to \mathbb{M}$ that are square integrable is denoted by $L^2(\mu)$. The tangent space to $\mathscr{P}_2(\mathbb{M})$ at μ denoted by $T_\mu\mathscr{P}_2(\mathbb{M})$ is the closure of $\nabla C_c^\infty(\mathbb{M})$ in $L^2(\mu)$.

If $\widetilde{\mathscr{U}} : \mathbb{H} \to \mathbb{R}$ is differentiable at $x \in \mathbb{H}$, we use the notations $\nabla\widetilde{\mathscr{U}}(x)$ or $\nabla_x\widetilde{\mathscr{U}}(x)$ to denote its Fréchet derivative at x (as an element of $\mathbb{H}$). If $\widetilde{\mathscr{U}}$ is twice differentiable at x, we use the notations $\nabla^2\widetilde{\mathscr{U}}(x)$ or $\nabla^2_{xx}\widetilde{\mathscr{U}}(x)$ to denote its Hessian (as a bi-linear form on $\mathbb{H} \times \mathbb{H}$). If $u : \mathbb{M} \to \mathbb{R}$ is differentiable at $q \in \mathbb{M}$, we use the notation $Du(q)$ or $D_q u(q)$ to denote its gradient at q. If it is twice differentiable at q, we use the notation $D^2u(q)$ or $D^2_{qq}u(q)$ to denote its Hessian matrix at q.

For $r > 0$, we define $\mathscr{B}_r$ to be the closed ball in $(\mathscr{P}_2(\mathbb{M}), W_2)$, centered at δ_0 and of radius r. $\mathbb{B}_r(0)$ stands for the closed ball in $\mathbb{H}$ centered at 0 and of radius r.

For any integer $m > 1$ we fix $(\Omega_i^m)_{i=1}^m$ to be a partition of Ω into Borel sets of the same volume. Given

$$q := (q_1, \ldots, q_m), \ p := (p_1, \ldots, p_m) \in \mathbb{M}^m,$$

we set

$$M^q := \sum_{i=1}^m q_i \chi_{\Omega_i^m}, \quad M^{mp} := \sum_{i=1}^m (mp_i)\chi_{\Omega_i^m} \equiv mM^p, \tag{1.1}$$
$$\mu_q^{(m)} := \frac{1}{m}\sum_{i=1}^m \delta_{q_i}.$$

We set

$$\mathbb{B}_r^m := \left\{ q \in \mathbb{M}^m : m^{-1} \sum_{j=1}^m |q_j|^2 \le r^2 \right\}, \quad \mathscr{P}_2^{(m)}(\mathbb{M}) := \left\{ \frac{1}{m} \sum_{i=1}^m \delta_{q_i} : q \in \mathbb{M}^m \right\}.$$

1.1 Assumptions

In this manuscript $N \ge 1$ is an integer, $m_*, \lambda_0 \in \mathbb{R}$, and $\kappa_0, \lambda_1, \kappa_3 > 0$. We shall denote by $\bar{\kappa}$ a generic constant depending on $m_*, \kappa_0, r_2, \kappa_3 > 0$. Let $-\infty < s < t < \infty$, and let $m > 1$ be an integer.

When $\mathbb{S}$ is a metric space, we denote by $AC_2(s,t;\mathbb{S})$ the set of $S : [s,t] \to \mathbb{S}$, which are 2-absolutely continuous. When $\tau \in [s,t]$, when convenient, we write S_τ in place of $S(\tau)$. We are imposing the following standing assumptions throughout the paper.

Suppose

$$\text{(H1)} \qquad \widetilde{\mathscr{F}}, \widetilde{\mathscr{U}_0} \in C^{1,1}(\mathbb{H}), \quad \widetilde{\mathscr{F}} \ge 0, \quad \widetilde{\mathscr{U}_0} \ge m_*,$$

and are rearrangement invariant in the sense that if $x, y \in \mathbb{H}$ have the same law, then $\widetilde{\mathscr{F}}(x) = \widetilde{\mathscr{F}}(y)$ and $\widetilde{\mathscr{U}_0}(x) = \widetilde{\mathscr{U}_0}(y)$. Note that (H1) implies in particular that there exists $\kappa_0 > 0$ such that and

$$\text{(1.2)} \qquad \nabla \widetilde{\mathscr{F}}, \nabla \widetilde{\mathscr{U}_0} \quad \text{are } \kappa_0\text{-Lipschitz-continuous.}$$

We assume

$$\text{(H2)} \qquad \widetilde{\mathscr{U}_0} \text{ is convex.}$$

Let

$$\text{(H3)} \qquad H, L \in C^{N+1}(\mathbb{M} \times \mathbb{R}^d), \quad L \ge 0,$$

such that $L(q,\cdot)$ and $H(q,\cdot)$ are Legendre transforms of each other for any $q \in \mathbb{M}$. We assume

$$\text{(H4)} \qquad D^2_{vv} L \ge \kappa_3 I_d, \quad D^2_{pp} H > 0,$$

and

$$\text{(H5)} \qquad DH, DL \text{ are } \kappa_0\text{-Lipschitz-continuous.}$$

We further assume

$$\text{(H6)} \qquad \lambda_1 |v|^2 + \lambda_0 \le L(q,v).$$

We set

$$\widetilde{\mathscr{L}}(x,a) = \int_\Omega L\big(x(\omega), a(\omega)\big) d\omega + \widetilde{\mathscr{F}}(x),$$
$$\widetilde{\mathscr{H}}(x,b) = \int_\Omega H\big(x(\omega), b(\omega)\big) d\omega - \widetilde{\mathscr{F}}(x),$$

for $x, a, b \in \mathbb{H}$ and assume

$$\text{(H7)} \qquad \widetilde{\mathscr{L}} \text{ is jointly strictly convex in both variables.}$$

Observe that a sufficient condtion for (H7) to be satisfied is to assume existence of a constant $\kappa_1 > 0$ such that $\widetilde{\mathscr{F}}$ is κ_1-convex and that there exists $\kappa_2 > 0$ such that

$$D^2 L(\overline{q}, \overline{v}) \begin{pmatrix} q \\ v \end{pmatrix} \cdot \begin{pmatrix} q \\ v \end{pmatrix} \geq \kappa_2 |v|^2 \quad \forall q, \overline{q}, v, \overline{v} \in \mathbb{R}^d. \tag{1.3}$$

In this case, the strict convexity of $\widetilde{\mathscr{L}}$ would follow from the fact that

$$\frac{d^2}{dt^2} \widetilde{\mathscr{L}}(\overline{x} + tx, \overline{a} + ta)\Big|_{t=0} \geq \kappa_1 \|x\|^2 + \kappa_2 \|a\|^2 \quad \forall x, a, \overline{x}, \overline{a} \in \mathbb{H}. \tag{1.4}$$

The regularity assumptions (H1) and (H3) will be important to derive regularity estimates on the classical solution $\widetilde{\mathscr{U}}$ to the corresponding Hamilton–Jacobi equation. At first glance these are sufficient to obtain well-known semiconcavity and Lipschitz estimates on this solution. The convexity of $\widetilde{\mathscr{L}}$ in (H7) and of $\widetilde{\mathscr{U}}_0$ in (H2) will then imply that $\widetilde{\mathscr{U}}(t, \cdot)$ (as a value function in an optimal control problem) is convex. Together with the previous properties this will lead to the $C^{1,1}$ regularity on $\widetilde{\mathscr{U}}(t, \cdot)$. To be able to achieve higher regularity estimates on $\widetilde{\mathscr{U}}(t, \cdot)$ that will be necessary to derive the corresponding master equations, additional assumptions will be introduced in Section 4. The combination of (H1) and (H5) ensures that the underlying Hamiltonian flow is globally well-posed. We combine (H6) and (H7) to obtain existence and uniqueness of solutions to the optimal control problems associated to $\widetilde{\mathscr{U}}(t, \cdot)$. Finally, the strict convexity assumptions in (H3) will help us to deduce the invertibility of the Hamiltonian flow and by this linking it to the optimal curve in the definition of $\widetilde{\mathscr{U}}(t, \cdot)$.

For any $S \in AC_2(s, t; \mathbb{H})$ we set

$$\widetilde{\mathscr{A}}_s^t(S) := \int_s^t \widetilde{\mathscr{L}}(S, \dot{S}) d\tau.$$

When $x, y \in \mathbb{H}$ we set

$$\widetilde{C}_s^t(x, y) := \inf_S \Big\{ \widetilde{\mathscr{A}}_s^t(S) : S(0) = x, S(t) = y,\ S \in AC_2(s, t; \mathbb{H}) \Big\}$$

and define for $t > 0$,

$$\widetilde{\mathscr{U}}(t, y) = \inf_{z \in \mathbb{H}} \Big\{ \widetilde{C}_0^t(z, y) + \widetilde{\mathscr{U}}_0(z) \Big\}. \tag{1.5}$$

We denote as $AC_2(0, t; \mathbb{H}_y)$ the set of $S \in AC_2(0, t; \mathbb{H})$ such that $\widetilde{\mathscr{A}}_0^t(S) < \infty$ and $S(t) = y$. Strict convexity of $\widetilde{\mathscr{A}}_s^t$ is ensured by (H7).

Remark 1.1. The following holds.

(i) Using (H5), we obtain that $|H|$ and $|L|$ are bounded above by quadratic forms.

(ii) Note that by (H1) and (H6),

$$\widetilde{\mathscr{A}}_0^t(S) \geq \lambda_1 \int_0^t \|\dot{S}\|^2 d\tau + \lambda_0 t + m_*.$$

This ensures a precompactness property to the sublevel sets of $\widetilde{\mathscr{A}}_0^t$ when they are contained in $AC_2(0, t; \mathbb{H}_y)$ for some $y \in \mathbb{H}$.

(iii) The functions DL, DH, $\nabla \widetilde{\mathscr{U}}_0$, and $\nabla \widetilde{\mathscr{F}}$ being Lipschitz, there is a constant $\bar{\kappa}$ such that

$$|DL(q,v)| \leq \bar{\kappa}(|v| + |q| + 1), \ |DH(q,p)| \leq \bar{\kappa}(|p| + |q| + 1),$$
$$\text{and } \|\nabla \widetilde{\mathscr{U}}_0(x)\| + \|\nabla \widetilde{\mathscr{F}}(x)\| \leq \bar{\kappa}(\|x\| + 1).$$

The assumptions imposed on H and $\widetilde{\mathscr{F}}$ ensure $\nabla \widetilde{\mathscr{H}} : \mathbb{H}^2 \to \mathbb{R}$ is Lipschitz, and so there exists a unique Hamiltonian flow $\Sigma : \mathbb{R} \times \mathbb{H}^2 \to \mathbb{H}^2$ on the phase space, a solution to the initial value problem (0.5). By Remark 1.1(iii) there exists a constant $\tilde{\kappa} > \bar{\kappa}$ depending only on $\bar{\kappa}$ such that

$$\|\Sigma(t,x,b)\| + 1 \leq \big(\|(x,b)\| + 1\big)e^{\tilde{\kappa}t} \tag{1.6}$$

for any $t > 0$ and $x, b \in \mathbb{H}$. The restriction of Σ to the graph of $\nabla \widetilde{\mathscr{U}}_0$ is the flow map denoted by $(\tilde{\xi}, \tilde{\eta})$ (defined in (0.6)) on the spatial space, with values in the cotangent bundle. We combine (1.2) and (1.6) to find $c_5 > 0$ depending only on κ_0 and $\|\nabla \widetilde{\mathscr{U}}_0(0)\|$ such that

$$\|(\tilde{\xi}, \tilde{\eta})\| + 1 \leq c_5(\|x\| + 1)e^{\tilde{\kappa}t}. \tag{1.7}$$

We discuss some more classical properties of the Hamiltonian flow in the setting of Hilbert spaces in Appendix C.

1.2 Discretization

Fix a natural number $m > 1$. For $q, v, p \in \mathbb{M}^m$ we define

$$L^{(m)}(q,v) := \int_\Omega L(M^q, M^v)d\omega = \frac{1}{m}\sum_{i=1}^m L(q_i, v_i), \quad F^{(m)}(q) := \widetilde{\mathscr{F}}(M^q),$$

and

$$H^{(m)}(q,p) := \int_\Omega H(M^q, M^{mp})d\omega = \frac{1}{m}\sum_{i=1}^m H(q_i, mp_i).$$

Then we set

$$\mathscr{L}^m(q,v) := L^{(m)}(q,v) + F^{(m)}(q), \quad \mathscr{H}^m(q,p) := H^{(m)}(q,p) - F^{(m)}(q),$$
$$U^{(m)}(t,q) := \widetilde{\mathscr{U}}(t, M^q).$$

One checks that for each $j \in \{1, \dots, m\}$, $\nabla \widetilde{\mathscr{U}}(t, M^q)$ is constant on Ω_j^m and the following useful identities (see, for instance, [16, 31]) hold:

$$D_{q_j} U^{(m)}(t, q_1, \dots, q_m) = \frac{1}{m}\nabla \widetilde{\mathscr{U}}(t, M^q)|_{\Omega_j^m}. \tag{1.8}$$

Note this means in particular,

$$\nabla \widetilde{\mathscr{U}_0} : \{M^q : q \in \mathbb{M}^m\} \to \{M^q : q \in \mathbb{M}^m\}. \tag{1.9}$$

We infer

$$\nabla \widetilde{\mathscr{U}}(t, M^q) = m \sum_{j=1}^{m} \chi_{\Omega_j^m} D_{q_j} U^{(m)}(t, q). \tag{1.10}$$

Observe

$$\begin{aligned} D_{q_j} \mathscr{L}^m(q, v) &= \frac{1}{m} \nabla_x \widetilde{\mathscr{L}}(M^q, M^v)|_{\Omega_j^m}, \\ D_{v_j} \mathscr{L}^m(q, v) &= \frac{1}{m} \nabla_a \widetilde{\mathscr{L}}(M^q, M^v)|_{\Omega_j^m}, \end{aligned} \tag{1.11}$$

and so

$$\begin{aligned} \nabla_x \widetilde{\mathscr{L}}(M^q, M^v) &= m \sum_{j=1}^{m} \chi_{\Omega_j^m} D_{q_j} \mathscr{L}^m(q, v), \\ \nabla_a \widetilde{\mathscr{L}}(M^q, M^v) &= m \sum_{j=1}^{m} \chi_{\Omega_j^m} D_{v_j} \mathscr{L}^m(q, v). \end{aligned} \tag{1.12}$$

Similarly,

$$\begin{aligned} D_{q_j} \mathscr{H}^m(q, p) &= \frac{1}{m} \nabla_x \widetilde{\mathscr{H}}(M^q, M^{mp})|^m_{\Omega_j}, \\ D_{p_j} \mathscr{H}^m(q, p) &= \nabla_b \widetilde{\mathscr{H}}(M^q, M^{mp})|_{\Omega_j^m}. \end{aligned} \tag{1.13}$$

Note that the fact that the coefficient in front of $\nabla_b \widetilde{\mathscr{H}}(M^q, M^{mp})$ is not divided by m is not a misprint. However, we have

$$D_{q_j} \mathscr{H}^m\big(q, D_q U^{(m)}(t, q)\big) = \frac{1}{m} \nabla_x \widetilde{\mathscr{H}}\big(M^q, \nabla \widetilde{\mathscr{U}}(t, M^q)\big)|_{\Omega_j^m}, \tag{1.14}$$

and so

$$\frac{1}{m} \nabla_x \widetilde{\mathscr{H}}\big(M^q, \nabla \widetilde{\mathscr{U}}(t, M^q)\big) = \sum_{j=1}^{m} D_{q_j} \mathscr{H}^m\big(q, D_q U^{(m)}(t, q)\big) \chi_{\Omega_j^m}. \tag{1.15}$$

For any natural number m denote by $(\Sigma_1^m, \Sigma_2^m) : \mathbb{R} \times \mathbb{M}^{2m} \to \mathbb{M}^{2m}$ the Hamiltonian flow for $\mathscr{H}^m$.

For $x \in \mathbb{H}$ such that $\sharp(x) = \mu_q^{(m)}$ (i.e., $x = M^q$), we consider the spatially discretized flows

$$\xi_i^m(s, q) := \widetilde{\xi}_s[x]|_{\Omega_i^m}, \quad \eta_i^m(s, q) = \frac{1}{m} \widetilde{\eta}_s[x]|_{\Omega_i^m}. \tag{1.16}$$

Using the notation $(\xi^m, \eta^m) = (\xi_1^m, \dots, \xi_m^m, \eta_1^m, \dots, \eta_m^m)$, these flows are uniquely defined to satisfy

$$\begin{cases} \dot{\xi}_i^m(s,q) = D_{p_i}\mathscr{H}^m\big(\xi_i^m(s,q), \eta_i^m(s,q)\big), & (s,q) \in (0,\infty) \times \mathbb{M}^m, \\ \dot{\eta}_i^m(s,q) = -D_{q_i}\mathscr{H}^m\big(\xi_i^m(s,q), \eta_i^m(s,q)\big), & (s,q) \in (0,\infty) \times \mathbb{M}^m, \\ \big(\xi^m(0,q), \eta^m(0,q)\big) = \big(q, D_q U_0^{(m)}(q)\big), & q \in \mathbb{M}^m. \end{cases} \tag{1.17}$$

1.3 Direct arguments for $C^{1,1}_{\mathrm{loc}}$–regularity in Hilbert setting

Throughout this subsection, we apply (H1)–(H7). We rely on the theory of existence of solutions to Hamilton–Jacobi equations on Hilbert spaces developed in [22, 23]. The function $\widetilde{U}$ defined in (1.5) is the unique viscosity solution to

$$\begin{cases} \partial_t \widetilde{\mathscr{U}} + \widetilde{\mathscr{H}}\big(x, \nabla \widetilde{\mathscr{U}}\big) = 0 & \text{in } (0,\infty) \times \mathbb{H}, \\ \widetilde{\mathscr{U}}(0,\cdot) = \widetilde{\mathscr{U}}_0 & \text{on } \mathbb{H}. \end{cases} \tag{1.18}$$

In this subsection, basic analytical tools are used to verify that $\widetilde{\mathscr{U}}$ is of class $C^{1,1}_{\mathrm{loc}}$. We refer the reader to [33] for the proof of the following proposition.

PROPOSITION 1.2. *There exists $e_1 \in C(\mathbb{R}_+, \mathbb{R}_+)$ monotone nondecreasing such that the following hold for $T > 0$ and $r > 0$:*

(i) *$\widetilde{\mathscr{U}}$ is $e_1\big(r(T+1)\big)$-Lipschitz on $[0,T] \times \mathbb{B}_r(0)$.*

(ii) *$\widetilde{\mathscr{U}}(t,\cdot)$ is $e_1\big(r(t+1)\big)$-semiconcave on $\mathbb{B}_r(0)$ for $t \in [0,T]$.*

PROPOSITION 1.3. *There is an increasing function $e_1 \in C(\mathbb{R}_+, \mathbb{R}_+)$ such that if $t > 0$ then*

(i) *$\widetilde{\mathscr{U}}(t,\cdot)$ is rearrangement invariant.*

(ii) *$\widetilde{\mathscr{U}}(t,\cdot)$ is convex, and so it is differentiable and $\nabla\widetilde{\mathscr{U}}(t,\cdot)$ is $e_1\big(r(t+1)\big)$-Lipschitz on $\mathbb{B}_r(0)$.*

PROOF. (i) The invariance property imposed on $\widetilde{\mathscr{U}}_0$ and $\widetilde{\mathscr{F}}$ implies $\widetilde{\mathscr{L}}$ satisfies the invariance property

$$\widetilde{\mathscr{L}}(x,a) = \widetilde{\mathscr{L}}(x \circ E, a \circ E)$$

for $x, a \in \mathbb{H}$, $E : \Omega \to \Omega$ such that E preserves Lebesgue measure. Since $\widetilde{\mathscr{L}}$ is further continuous, we conclude that $\widetilde{\mathscr{U}}(t,\cdot)$ is rearrangement invariant for $t \geq 0$ (cf. [32]).

(ii) The convexity of $\widetilde{\mathscr{A}}_0^t$ on $AC_2(0,t;\mathbb{H})$ and (H2) yields the convexity of $\widetilde{\mathscr{U}}(t,\cdot)$ on $\mathbb{H}$. This, together with Proposition 1.2 (ii) completes the proof. □

Remark 1.4. Let $q \in \mathbb{M}^m$. Note $\sigma \mapsto \int_0^t \mathscr{L}^m(\sigma,\dot{\sigma})d\tau + U_0^{(m)}(\sigma(0))$ is strictly convex on $AC_2\big(0,t;q;\mathbb{R}^{md}\big)$, and the set of paths $\sigma \in AC_2\big(0,t;\mathbb{R}^{md}\big)$ is such that $\sigma(t) = q$. Since $\mathscr{L}^m$ is of class C^2 and satisfies the assumptions in Section 1.1, standard results of the calculus of variations ensure that $\int_0^t \mathscr{L}^m(\sigma,\dot{\sigma})d\tau +$

$U_0^{(m)}(\sigma(0))$ admits a unique minimizer σ^m on $AC_2\big(0,t;q;\mathbb{M}^m\big)$. The minimizer is completely characterized by the Euler–Lagrange equations

$$\begin{aligned}\frac{d}{d\tau}\big(D_v\mathscr{L}^m(\sigma^m,\dot\sigma^m)\big) &= D_q\mathscr{L}^m(\sigma^m,\dot\sigma^m), \quad \sigma^m(t)=q,\\ D_qU_0^{(m)}(\sigma^m(0)) &= D_q\mathscr{L}^m(\sigma^m(0),\dot\sigma^m(0)).\end{aligned}\tag{1.19}$$

Define

$$U^m(t,q) := \int_0^t \mathscr{L}^m(\sigma^m,\dot\sigma^m)d\tau + U_0^{(m)}(\sigma^m(0)).$$

It is well-known that U^m is the unique continuous viscosity solution to

$$\partial_t U^m + \mathscr{H}^m\big(q, D_qU^m\big) = 0 \quad \text{on } (0,\infty)\times\mathbb{M}^m, \qquad U^m(0,\cdot) = U_0^{(m)}.\tag{1.20}$$

Setting $S := M^{\sigma^m}$, we have $\dot S = M^{\dot\sigma^m}$. We use (1.10) at $t=0$, and then use (1.12) and (1.19) to obtain

$$\frac{d}{d\tau}\big(\nabla_a\widetilde{\mathscr{L}}(S,\dot S)\big) = \nabla_x\widetilde{\mathscr{L}}(S,\dot S), \quad \nabla\widetilde{\mathscr{U}}_0(S(0)) = \nabla_a\widetilde{\mathscr{L}}(S(0),\dot S(0)).$$

This means S is a critical point of $\widetilde{\mathscr{A}}_0^t$ over $AC_2(0,t;\mathbb{H}_y)$ if we set $y := M^q$. Since $\widetilde{\mathscr{A}}_0^t$ is convex over $AC_2(0,t;\mathbb{H}_y)$, we conclude that S is a minimizer of $\widetilde{\mathscr{A}}_0^t$ over $AC_2(0,t;\mathbb{H}_y)$. Thus,

$$U^m(t,q) = \widetilde{\mathscr{A}}_0^t(S) = \widetilde{\mathscr{U}}(t,M^q) = U^{(m)}(t,q).\tag{1.21}$$

Consequently, $U^{(m)}$ is the unique viscosity solution to (1.20). We emphasize that the observation (1.21) is crucial in our consideration and in fact represents the heart of our analysis. This is a feature of the deterministic setting, and so this approach might not be applicable to stochastic Hamiltonian systems.

The proof of the following proposition will be provided in Appendix C.3.

PROPOSITION 1.5. *There exists $e_0 : [0,\infty)\to[0,\infty)$, monotone nondecreasing, such that the following hold*:

(i) *If $0\le t_1<t_2\le T$, then*

$$\widetilde{\mathscr{U}}(t_2,y) - \widetilde{\mathscr{U}}(t_1,y) = -\int_{t_1}^{t_2}\widetilde{\mathscr{H}}\big(y,\nabla\widetilde{\mathscr{U}}(\tau,y)\big)d\tau \quad \forall y\in\mathbb{H}.$$

(ii) *$\widetilde{\mathscr{U}}$ is continuously differentiable on $(0,\infty)\times\mathbb{H}$, and $\partial_t\widetilde{\mathscr{U}}$, $\nabla\widetilde{\mathscr{U}}$ are Lipschitz on $[0,T]\times\mathbb{B}_r(0)$.*

(iii) *For any $y\in\mathbb{H}$, there exists a unique $S\in AC_2(0,t;\mathbb{H}_y)$ such that*

$$\widetilde{\mathscr{U}}(t,y) = \widetilde{\mathscr{A}}_0^t(S) + \widetilde{\mathscr{U}}_0(S(0)).$$

(iv) *Let S be as in* (iii) *and set $P := \nabla_a\widetilde{\mathscr{L}}(S,\dot S)$. Then $S,P\in C^2([0,t];\mathbb{H})$,*

$$\begin{aligned}\dot S = \nabla_b\widetilde{\mathscr{H}}(S,P), \quad \dot P &= \nabla_x\widetilde{\mathscr{L}}(S,\dot S) = -\nabla_x\widetilde{\mathscr{H}}(S,P),\\ \nabla\widetilde{\mathscr{U}}(\cdot,S) &= \nabla_a\widetilde{\mathscr{L}}(S,\dot S) \quad \text{on } [0,t].\end{aligned}\tag{1.22}$$

In particular,

$$\nabla\widetilde{\mathscr{U}}_0(S(0)) = \nabla_a\widetilde{\mathscr{L}}(S(0),\dot S(0)). \tag{1.23}$$

(v) *We have*

$$\begin{gathered}\widetilde{C}_0^t(S(0),y),\quad \|\dot S(\tau)\| \le e_0\big((t+1)\|y\|\big),\\ \|S(\tau)\| \le \|y\| + te_0\big((t+1)\|y\|\big)\quad \forall\tau\in[0,t].\end{gathered} \tag{1.24}$$

Remark 1.6. (i) We denote the unique S that appears in Proposition 1.5(iii) as

$$\widetilde{S}_s^t[y](\omega) := S(s,\omega),\quad 0\le s\le t,\ \ \omega\in\Omega.$$

It is uniquely characterized by the equation

$$\widetilde{\mathscr{U}}(t,y) = \int_0^t \widetilde{\mathscr{L}}\big(\widetilde{S}_s^t[y],\partial_s\widetilde{S}_s^t[y]\big)ds + \widetilde{\mathscr{U}}_0\big(\widetilde{S}_0^t[y]\big),\quad S_t^t[y] = y. \tag{1.25}$$

Defining

$$\widetilde{P}_s^t[y] = \nabla_a\widetilde{\mathscr{L}}\big(\widetilde{S}_s^t[y],\partial_s\widetilde{S}_s^t[y]\big),$$

we have

$$\begin{cases}\partial_s\widetilde{S}_s^t[y] = \nabla_b\widetilde{\mathscr{H}}\big(\widetilde{S}_s^t[y],\widetilde{P}_s^t[y]\big) & \text{for } (s,y)\in(0,t)\times\mathbb{H},\\ \partial_s\widetilde{P}_s^t[y] = -\nabla_x\widetilde{\mathscr{H}}\big(\widetilde{S}_s^t[y],\widetilde{P}_s^t[y]\big) & \text{for } (s,y)\in(0,t)\times\mathbb{H},\\ \big(\widetilde{S}_t^t[y],\widetilde{P}_0^t[y]\big) = \big(y,\nabla\widetilde{\mathscr{U}}_0(y)\big) & \text{for } y\in\mathbb{H}.\end{cases} \tag{1.26}$$

(ii) For any natural number m and $q\in\mathbb{M}^m$, we have

$$\widetilde{S}_s^t[M^q] = M^{\sigma_s^{t,m}[q]}, \tag{1.27}$$

where $(\sigma_s^{t,m}[q])_{s\in(0,t)}$ is the optimizer discussed in Remark 1.4. Let us emphasize that only in the case of deterministic Hamiltonian systems like ours, (1.27) provides the characteristics not only for the viscosity solutions of the Hamilton–Jacobi equation on $\mathbb{H}$ but also for the one on $\mathbb{M}^m$.

(iii) When the conditions in Remark 1.6 are satisfied, we define the vector field

$$B(t,\cdot) := \nabla_b\widetilde{\mathscr{H}}\big(\cdot,\widetilde{\eta}(t,\widetilde{S}_0^t)\big). \tag{1.28}$$

which will turn out to be the velocity in Eulerian coordinates for the trajectory $\widetilde{\xi}$.

2 Regularity Estimates for HJEs and Hamiltonian Systems for Systems of *m* Particles

In this section, we assume that (H3)–(H6) hold. Let $u_0\in C^N(\mathbb{M})$ be a convex function with bounded second derivatives. Let $F\in C^N(\mathbb{M})$ and L be such that the corresponding Lagrangian action, as in (H7), is strictly convex. We fix $T>0$. We shall show that classical solutions to Hamilton–Jacobi equations set on $\mathbb{M}^m$ possess higher derivative estimates that we precisely quantify in terms of m. As we will see in the next sections, when $m\to+\infty$, these estimates will provide the necessary regularity estimates on $\mathscr{U}$, the solution to the corresponding Hamilton–Jacobi equation set on $\mathscr{P}_2(\mathbb{M})$.

2.1 One-Particle Hamiltonian Flow

We study the regularity of viscosity solutions $u : [0,T] \times \mathbb{M} \to \mathbb{R}$ of Cauchy problems of the form

$$\begin{cases} \partial_t u + H(q, \nabla u) - F(q) = 0, & (0,T) \times \mathbb{M}, \\ u(0,\cdot) = u_0, & \mathbb{M}. \end{cases} \tag{2.1}$$

Given $t \in (0,T]$, we consider the Hamiltonian system

$$\begin{cases} \dot S(s,q) = D_p H(S(s,q), P(s,q)), & s \in (0,t),\ q \in \mathbb{M}, \\ \dot P(s,q) = -D_q H(S(s,q), P(s,q)) + D_q F(Q(s,q)), & s \in (0,t),\ q \in \mathbb{M}, \\ S(t,q) = q,\ P(0,q) = Du_0(S(0,q)), & q \in \mathbb{M}, \end{cases} \tag{2.2}$$

Such a flow has been considered in greater generality in Remark 1.6. Recall S is the unique optimizer in

$$u(t,x) := \inf\left\{ u_0(\gamma(0)) + \int_0^t L(\gamma(s), \dot\gamma(s)) + F(\gamma(s))ds :\ \gamma(t) = x \right\}. \tag{2.3}$$

Similarly, we shall use the flow

$$\begin{cases} \dot\xi(s,z) = D_p H(\xi(s,z), \eta(s,z)), & s \in (0,t),\ z \in \mathbb{M}, \\ \dot\eta(s,z) = -D_q H(\xi(s,z), \eta(s,z)) + \nabla_q F(\xi(s,z)), & s \in (0,t),\ z \in \mathbb{M}, \\ \xi(0,z) = z,\ \eta(0,z) = Du_0(z), & z \in \mathbb{M}, \end{cases} \tag{2.4}$$

denoted as $(\tilde\xi, \tilde\eta)$ in (0.5) when our Hilbert space reduces to $\mathbb{M}$.

LEMMA 2.1. *Let* $t \in [0,T]$.

(1) *The map* $\xi_t : \mathbb{M} \to \mathbb{M}$ *is a homeomorphism* $S_s := \xi_s \circ \xi_t^{-1}$ *and* $P_s := \eta_s \circ \xi_t^{-1}$. *We have* $\xi_t, \eta_t \in C^{N-1}(\mathbb{M})$.

(2) *If we further assume* $N \geq 2$, *then* $u \in C^{1,1}_{\mathrm{loc}}([0,T] \times \mathbb{M})$ *is a classical solution to* (2.1) *and* $z \mapsto \xi(t,z)$ *is a* C^{N-1} *diffeomorphism from* $\mathbb{M}$ *onto itself.*

PROOF. (1) The existence and smooth dependence on the data of the solution of (2.2) is classical, Proposition C.2 ensures $\xi_t : \mathbb{M} \to \mathbb{M}$ is a homeomorphism and $S(s,\cdot) := \xi_s \circ \xi_t^{-1},\ P(s,\cdot) := \eta_s \circ \xi_t^{-1}$.

(2) By Proposition 1.5, $u \in C^{1,1}_{\mathrm{loc}}([0,T] \times \mathbb{M})$ and is a classical solution to (2.1). Let us show that $z \mapsto \xi(t,z)$ is a global C^{N-1} diffeomorphism. Recall that by Proposition C.2, ξ is a solution to

$$\begin{cases} \dot\xi(s,z) = D_p H(\xi(s,z), Du(s, \xi(s,z))), & s \in (0,t), \\ \xi(0,z) = z, \end{cases}$$

from where one has

$$\begin{cases} \partial_s D_z \xi(s,z) = A(s,z) D_z \xi(s,z), & s \in (0,t), \\ D_z \xi(0,z) = I_d. \end{cases}$$

Here we used the notation

$$A(s,z) := D^2_{xp}H(\xi(s,z), Du(s,\xi(s,z))) + D^2_{pp}H(\xi(s,z), Du(s,\xi(s,z)))D^2u(s,\xi(s,z)).$$

Since $A(s,z)$ is locally uniformly bounded, we have that for $s > 0$ small enough $D_z\xi(s,z)$ is invertible. Therefore, Jacobi's formula yields

$$\det(D_z\xi(s,z)) = \exp\left(\int_0^s \operatorname{tr}(A(\tau,z))d\tau\right).$$

Since $A(\tau,\cdot) \in L^\infty_{\mathrm{loc}}(\mathbb{M})$, uniformly with respect to $\tau \in [0,t]$, we have that $\det(D_z\xi(s,z)) > 0$ for all $z \in \mathbb{M}$, uniformly with respect to s. Therefore, $D_z\xi(s,z)$ is invertible for any $z \in \mathbb{M}$ and for any $s \in [0,t]$. Thus, by the fact that $\xi(t,\cdot) \in C^{N-1}(\mathbb{M})$ and that $\xi(t,\cdot)$ is bijective, we conclude that $z \mapsto \xi(t,z)$ is a global C^{N-1} diffeomorphism of $\mathbb{M}$ onto itself. □

2.2 *m*-particles Hamiltonian flow

Throughout this subsection, we assume to be given a positive monotone nondecreasing function $C_0 : (0,\infty) \to (0,\infty)$. Furthermore, we impose that in the assumption (H3) $N \geq 2$ and $F^{(m)}, U_0^{(m)} \in C^3(\mathbb{M}^m)$.

As in Section 1.2 we define

$$U_0^{(m)}(q) := \mathscr{U}_0\left(\frac{1}{m}\sum_{i=1}^m \delta_{q_i}\right), \quad F^{(m)}(q) := \mathscr{F}\left(\frac{1}{m}\sum_{i=1}^m \delta_{q_i}\right) \quad \forall q \in \mathbb{M}^m.$$

We also assume we are given that $U_0^{(m)}, F^{(m)} : \mathbb{M}^m \to \mathbb{R}$ satisfy Property 2.2(2) with $C = C_0(r)$. We also consider viscosity solutions $U^{(m)} : [0,T] \times \mathbb{M}^m \to \mathbb{R}$ of the Hamilton–Jacobi equation

$$\begin{cases} \partial_t U^{(m)}(t,q) + H^{(m)}(q, D_q U^{(m)}(t,q)) - F^{(m)}(q) = 0 & \text{on } (0,T)\times\mathbb{M}^m, \\ U^{(m)}(0,\cdot) = U_0^{(m)} & \text{on } \mathbb{M}^m. \end{cases} \tag{2.5}$$

By Remark 1.4

$$U^{(m)}(t,q) \equiv \widetilde{\mathscr{U}}(t, M^q) \quad \forall (t,q) \in [0,\infty)\times\mathbb{M}^m.$$

Given $t \in (0,T)$ we consider the m particles flows $S^{t,m}, P^{t,m} : \mathbb{M}^m \to \mathbb{M}^m$. In other words,

$$\begin{cases} \dot S_i^{t,m}(s,q) = D_pH(S_i^{t,m}(s,q), mP_i^{t,m}(s,q)), & (s,q) \in (0,t)\times\mathbb{M}^m, \\ \dot P_i^{t,m}(s,q) = -\frac{1}{m}D_qH(S_i^{t,m}(s,q), mP_i^{t,m}(s,q)) & \\ \qquad\qquad + D_{q_i}F^{(m)}(S^{t,m}(s,q)), & (s,q) \in (0,t)\times\mathbb{M}^m, \\ S_i^{t,m}(t,q) = q_i, \ P_i^{t,m}(0,q) = D_{q_i}U_0^{(m)}(S^{t,m}(0,q)) & q \in \mathbb{M}^m. \end{cases} \tag{2.6}$$

This is analogous to the flow $(S^{t,m}, P^{t,m})$ in Remark 1.6 where we have not displayed the m- and t-dependence to simplify the notation. We also consider the

m-particle flows $\xi^m, \eta^m : [0,\infty) \times \mathbb{M}^m \to \mathbb{M}^m$, similar to (2.4) (which also correspond to the discretized flow (1.17)). They are defined as

$$\begin{cases} \dot{\xi}_i^m(s,z) = D_p H(\xi_i^m(s,z), m\eta_i^m(s,z)), & s \in (0,t), \\ \dot{\eta}_i^m(s,z) = -\frac{1}{m} D_q H(\xi_i^m(s,z), m\eta_i^m(s,z)) + D_{q_i} F^{(m)}(\xi^m(s,z)), & s \in (0,t), \\ \xi_i^m(0,z) = z_i,\ \eta_i^m(0,z) = D_{q_i} U_0^{(m)}(z), \end{cases} \tag{2.7}$$

for $i \in \{1,\dots,m\}$, where $z = (z_1,\dots,z_m) \in \mathbb{M}^m$.

We next introduce functions on $\mathbb{M}^m$ and list some of their special properties which are useful for our study.

Property 2.2. For a permutation-invariant function $G^{(m)} : \mathbb{M}^m \to \mathbb{R}$ we define the following properties by assuming for each $r > 0$ that there is a $C \equiv C(r)$ increasing in r such that the following hold:

(1) (a) $G^{(m)} \in C^{0,1}_{\rm loc}(\mathbb{M}^m) \cap C^1(\mathbb{M}^m)$, and for every $m \in \mathbb{N}$ and $q \in \mathbb{B}_r^m(0)$ we have

$$|D_{q_i} G^{(m)}(q)| \le Cm^{-1},\ \forall i \in \{1,\dots,m\}. \tag{2.8}$$

(b) $G^{(m)} \in C^{0,1}_{\rm loc}(\mathbb{M}^m) \cap C^1(\mathbb{M}^m)$, and for every $m \in \mathbb{N}$ and $q \in \mathbb{B}_r^m(0)$ we have

$$\sum_{i=1}^m m|D_{q_i} G^{(m)}(q)|^2 \le C. \tag{2.9}$$

(2) $G^{(m)} \in C^{1,1}_{\rm loc}(\mathbb{M}^m) \cap C^2(\mathbb{M}^m)$, and for every $m \in \mathbb{N}$ and $q \in \mathbb{B}_r^m(0)$ we have

$$|D^2_{q_i q_j} G^{(m)}(q)|_\infty \le \begin{cases} Cm^{-1}, & i = j;\ i \in \{1,\dots,m\}, \\ Cm^{-2}, & i \ne j;\ i,j \in \{1,\dots,m\}. \end{cases} \tag{2.10}$$

Here for $A = (A_{ij})_{i,j=1}^m$, we use the notation $|A|_\infty := \max_{(i,j)} |A_{ij}|$.

(3) $G^{(m)} \in C^{2,1}_{\rm loc}(\mathbb{M}^m) \cap C^3(\mathbb{M}^m)$, and for every $m \in \mathbb{N}$ and $q \in \mathbb{B}_r^m(0)$ we have

$$\begin{aligned} &\left|D^3_{q_i q_j q_k} G^{(m)}(q)\right|_\infty \\ &\quad \le \begin{cases} Cm^{-1}, i = j = k;\ i \in \{1,\dots,m\}, \\ Cm^{-2}, (i = j \ne k) \text{ or } (i \ne j = k) \text{ or } (i = k \ne j), \\ Cm^{-3}, i \ne j \ne k,\ i,j,k \in \{1,\dots,m\}, \end{cases} \end{aligned} \tag{2.11}$$

for $i,j,k \in \{1,\dots,m\}$. Here for $A = (A_{ijk})_{i,j,k=1}^m$, we use the notation $|A|_\infty := \max_{(i,j,k)} |A_{ijk}|$.

We present now the main theorem of this section.

THEOREM 2.3. *Let* $U^{(m)} : (0,T) \times \mathbb{M}^m \to \mathbb{R}$ *be the unique viscosity solution of* (2.5)*, which is constructed by the discretization approach described in Remark* 1.4. *Let* $r > 0$*. Then for all* $t \in (0,T)$ *there exists* $C(t,r) > 0$ *such that the following hold for all* $m \in \mathbb{N}$.

(1) $U^{(m)}(t,\cdot)$ *satisfies the estimates in Property* 2.2(2) *in* $\mathbb{B}^m_r(0)$ *with constant* $C(t,r)$.

(2) *Further assume that* $U_0^{(m)}$ *and* $F^{(m)}$ *satisfy Property* 2.2(3) *and* (H13) *takes place. Then* $U^{(m)}(t,\cdot)$ *satisfies the estimates in Property* 2.2(3) *in* $\mathbb{B}^m_r(0)$ *with constant* $C(t,r)$.

(3) *We assume that the assumptions from* (1) *and* (H15) *take place. Then* $\partial_t U^{(m)}(t,\cdot)$ *satisfies the estimates in Property* 2.2(1)(b) *in* $\mathbb{B}^m_r(0)$ *with constant* $C(t,r)$.

Remark 2.4. Since the proof of the previous theorem is quite technical, we summarize its main ideas. First, as a consequence of the results in Section 1 (in particular in Proposition 1.5), $U^{(m)}$ is actually a classical solution to (2.5), which is of class $C^{1,1}_{\text{loc}}$. Then classical results from the literature will imply that it is as smooth as the data $H, F^{(m)}$ and $U_0^{(m)}$ (cf. [12]). Therefore, it remains to obtain the precise uniform derivative estimates as claimed in the statement of the theorem.

A key observation is the well-known representation formula for $D_q U^{(m)}$, i.e.,

$$D_q U^{(m)}(t,q) = \eta^m(t,\cdot) \circ (\xi^m)^{-1}(t,q),$$

where (ξ^m, η^m) is the Hamiltonian flow, the solution to (2.7). Therefore, the precise derivative estimates on $U^{(m)}$ can be obtained by differentiating the previous formula and relying on careful derivative estimates of the flow (ξ^m, η^m) and of its inverse. We obtain these necessary estimates by studying the linearized system (and its derivative) associated to (2.7). Since these computations will be quite delicate, we identify two simplified systems in Lemma 2.5 and Lemma 2.6, which carry the main structure of the original linearized systems. Estimates on these simpler systems will essentially be enough to deduce the estimates on the linearized systems we are aiming for. Finally, the derivative estimates on $\partial_t U^{(m)}$ are obtained by directly differentiating the Hamilton–Jacobi equation and using the previously established estimates on spatial derivatives of $U^{(m)}$.

PROOF OF THEOREM 2.3. We aim to obtain precise upper bounds on expressions depending on m (with respect to m when m is large). For this, we use the standard big-O notation. For instance, if α is an integer and $A(m)$ is a real number depending on m, by

$$A(m) = O(m^\alpha)$$

we mean that there exists $C > 0$ independent of m such that $|A(m)| \le Cm^\alpha$ for all m large. If $A(m) = (a_{ij}(m))_{ij}$ is a matrix whose elements are real numbers depending on m, by abuse of the notation, by $A(m) = O(m^\alpha)$ we mean that there exists a constant $C > 0$ independent of m such that $|a_{ij}(m)| \le Cm^\alpha$ for all i, j.

When $A(m) = (a_{ij}(m))_{ij}$ and $B(m) = (b_{ij}(m))_{ij}$ are matrices, by $A(m) = O(B(m))$ we mean that $a_{ij}(m) = O(b_{ij}(m))$ for all i, j. To simplify the notation, we sometimes write $A(m) \sim B(m)$ for $A(m) = O(B(m))$ and $B(m) = O(A(m))$.

First, let us notice that by Proposition 1.5, $U^{(m)}$ is a $C^{1,1}_{\mathrm{loc}}((0,T)\times\mathbb{M}^m)$ classical solution of (2.5); therefore in particular any point $(t,q) \in (0,T)\times\mathbb{M}^m$ is regular and not conjugate (by the proof of Lemma 2.1) in the sense of definition 6.3.4 of [12].

Furthermore, we notice that Lemma 2.1 asserts that $\xi^m(s,\cdot)$ is a C^N diffeomorphism, and theorem 6.4.11 from [12] yields that $U^{(m)} \in C^3((0,T)\times\mathbb{M}^m)$. In what follows we aim to obtain quantitative derivative estimates on $U^{(m)}$ with respect to the discretization parameter m.

Step 0. *Basic bounds on $\xi^m(t,z)$ when $q := \xi^m_t(z) \in \mathbb{B}^m_r(0)$.*

By Proposition C.2, $\xi^m(s,z) = S^{t,m}_s[q]$ since $q = \xi^m(t,z)$. By the same proposition, for $i \in \{1,\dots,m\}$ and $z \in \mathbb{M}^m$, we have

$$\begin{cases} \dot\xi^m_i(t,z) = D_pH(\xi^m_i(t,z), mD_{q_i}U^{(m)}(t,\xi^m(t,z))), & t\in(0,T),\\ \xi^m(0,z) = z, \end{cases} \tag{2.12}$$

and

$$\begin{aligned} \eta^m_i(t,z) &= D_{q_i}U^{(m)}(t,\xi^m(t,z)) = D_{q_i}U^{(m)}(t,x)\\ \eta^m_i(0,z) &= D_{q_i}U^{(m)}_0(z). \end{aligned} \tag{2.13}$$

By Proposition 1.5 there exists $\beta(t,r) > 0$ (independent of m); for any $q \in \mathbb{B}^m_r(0)$ we have

$$S^{t,m}_s[q] \equiv \xi^m(s,z) \in \mathbb{B}^m_{\beta(t,r)} \quad \text{for all } s\in[0,t]. \tag{2.14}$$

Proposition 1.2 ensures $\widetilde{\mathscr{U}}$ is locally Lipschitz on $[0,\infty)\times\mathbb{H}$, and so there exists $C_1(t,r) > 0$ (depending on $\beta(t,r)$) such that $\|\nabla\widetilde{\mathscr{U}}(t,\xi(t,M^z)\| \le C_1(t,r)$. Using the relation between $\nabla\widetilde{\mathscr{U}}$ and η provided by Proposition C.2(iv) we conclude

$$\sum_{i=1}^m m|\eta^m_i(t,z)|^2 \le C_1(t,r). \tag{2.15}$$

We are now well equipped to start the proof of the assertion (1) of the theorem.

Step 1. *Estimates on $(D_{z_j}\xi_i(t,\cdot), D_{z_j}\eta_i(t,\cdot))^m_{i,j=1}$.*

CLAIM 1. There exists a constant $C_2(t,r) > 0$ (independent of m) such that if $\xi(t,z) = q \in \mathbb{B}^m_r(0)$, then for all $i, j \in \{1,\dots,m\}$ we have

$$\left|D_{z_j}\xi^m_i(t,\cdot)\right|_\infty \le \begin{cases} C_2(t,r), & i=j,\\ \frac{C_2(t,r)}{m}, & i\neq j, \end{cases}$$

and

$$\big|D_{z_j}\eta_i^m(t,\cdot)\big|_\infty \le \begin{cases} \frac{C_2(t,r)}{m}, & i=j,\\ \frac{C_2(t,r)}{m^2}, & i\neq j.\end{cases} \tag{2.16}$$

PROOF OF CLAIM 1. By differentiating the Hamiltonian system (2.7) with respect to the z_j, we get

$$\begin{cases} \partial_t D_{z_j}\xi_i^m = D^2_{qp}H(\xi_i^m, m\eta_i)D_{z_j}\xi_i^m + mD^2_{pp}H(\xi_i^m, m\eta_i^m)D_{z_j}\eta_i^m,\\ \partial_t D_{z_j}\eta_i^m = -\frac{1}{m}\left(D^2_{qq}H(\xi_i^m, m\eta_i)D_{z_j}\xi_i^m + mD^2_{pq}H(\xi_i^m, m\eta_i)D_{z_j}\eta_i^m\right)\\ \qquad\qquad + \sum_{l=1}^m D^2_{q_l q_i}F^{(m)}(\xi^m)D_{z_j}\xi_l^m,\\ D_{z_j}\xi_i^m(0,\cdot) = \begin{cases} I_{d\times d}, & i=j,\\ 0_{d\times d}, & i\neq j,\end{cases} \quad D_{z_j}\eta_i^m(0,z) = D^2_{q_j q_i}U_0^{(m)}(z).\end{cases} \tag{2.17}$$

$$\begin{cases} \partial_t D_{z_j}\xi_i^m = D^2_{qp}H(\xi_i^m, m\eta_i)D_{z_j}\xi_i^m + mD^2_{pp}H(\xi_i^m, m\eta_i^m)D_{z_j}\eta_i^m,\\ \partial_t D_{z_j}\eta_i^m = -\frac{1}{m}\left(D^2_{qq}H(\xi_i^m, m\eta_i)D_{z_j}\xi_i^m + mD^2_{pq}H(\xi_i^m, m\eta_i)D_{z_j}\eta_i^m\right)\\ \qquad\qquad + \sum_{l=1}^m D^2_{q_l q_i}F^{(m)}(\xi^m)D_{0z_j}\xi_l^m,\\ D_{z_j}\xi_i^m(0,\cdot) = \begin{cases} I_{d\times d}, & i=j,\\ 0_{d\times d}, & i\neq j,\end{cases} \quad D_{z_j}\eta_i^m(0,z) = D^2_{q_j q_i}U_0^{(m)}(z).\end{cases} \tag{2.18}$$

Let us set

$$\bar{C}_2 := \max\{|\partial_q^a\partial_p^b H(q,p)| : (q,p)\in\mathbb{R}^d\times\mathbb{R}^d,\ |a|+|b|=2\}.$$

If $\xi^m(t,z) = q \in \mathbb{B}_r^m(0)$, then in the same way, there exists $\widetilde{C}_2(t,r) > 0$ (depending on $\beta(t,r)$) such that $D^2_{q_l q_i}F^{(m)}(\xi_1,\dots,\xi_m)$ and $D^2_{q_j q_i}U_0^{(m)}(z)$ satisfy the estimate (2.10) with $\widetilde{C}_2(t,r)$. Set

$$\widehat{C}_2 = \widehat{C}_2(t,r) := \max\{\bar{C}_2, \widetilde{C}_2(t,r)\}.$$

We plan to use the bounds

$$\big|D^2_{qp}H(\xi_i^m, m\eta_i)\big|_\infty,\quad \big|D^2_{pq}H(\xi_i^m, m\eta_i)\big|_\infty \le \bar{C}_2,$$
$$\big|(1/m)D^2_{qq}H(\xi_i^m, m\eta_i)\big|_\infty \le \bar{C}_2/m,\quad \big|mD^2_{pp}H(\xi_i^m, m\eta_i^m)\big|_\infty \le \bar{C}_2 m,$$

and

$$\big|D^2_{q_l q_i}F^{(m)}(\xi^m)\big|_\infty \le \begin{cases} \widetilde{C}_2(t,r)m^{-1}, & i=l,\\ \widetilde{C}_2(t,r)m^{-2}, & i\neq l,\end{cases}$$

$$\big|D^2_{q_j q_i}U_0^{(m)}(z)\big|_\infty \le \begin{cases} \widetilde{C}_2(t,r)m^{-1}, & i=j,\\ \widetilde{C}_2(t,r)m^{-2}, & i\neq j.\end{cases}$$

Thus, to obtain the precise bounds (in terms of m) on the solution to the system (2.18), it is enough to obtain bounds on the solution

$$(\widehat{X}(s), \widehat{Y}(s)) = \big((\widehat{X}_{ij}(s))_{i,j=1}^m, (\widehat{Y}_{ij}(s))_{i,j=1}^m\big)$$

to

$$\begin{cases} \partial_t \widehat{X}_{ij} = \widehat{C}_2 \widehat{X}_{ij} + m\widehat{C}_2 \widehat{Y}_{ij}, \\ \partial_t \widehat{Y}_{ij} = (\widehat{C}_2/m)\widehat{X}_{ij} + \widehat{C}_2 \widehat{Y}_{ij} + \sum_{l=1, l\neq i}^m (\widehat{C}_2/m^2)\widehat{X}_{lj}, \\ \widehat{X}_{ij}(0) = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}, \quad \widehat{Y}_{ij}(0) = \begin{cases} \widehat{C}_2 m^{-1}, & i = j, \\ \widehat{C}_2 m^{-2}, & i \neq j, \end{cases} \end{cases}$$

The constant $\widehat{C}_2 > 0$ can be simply factorized out from the previous system, and since this is independent of m, when studying the solution, without loss of generality it is enough to study the modified system with coefficients 1, instead of $\widehat{C}_2$. Thus, when writing the system in a closed form, one can clearly identify the blocks $B_1, \ldots, B_4$ defined in (2.25) and the system appearing in Lemma 2.6. Therefore, by the precise estimates on $(X_{ij}, Y_{ij})_{i,j=1}^m$ in Lemma 2.6, we conclude that there exists $C > 0$ (independent of m) such that Claim 1 follows by setting

$$C_2(t,r) := e^{tC\widehat{C}(t,r)}.$$

Now, let us denote by $\zeta^m = (\zeta_1^m(t,\cdot), \ldots, \zeta_m^m(t,\cdot)) := S_0^{t,m}[q]$ the inverse of $\xi^m(t,\cdot)$; in particular, we have that if $\xi_i^m(t,z) = q_i$, then $\zeta_i^m(t,q) = z_i$. Next, we derive estimates for $D_{q_j}\zeta_i^m(t,\cdot)$.

Step 2. *Estimates on* $(D_{q_j}\zeta_i^m)_{i,j=1}^m$.

CLAIM 2. There exists $C_3(t,r) > 0$ (independent of m) such that for all $i, j \in \{1, \ldots, m\}$ we have

$$\big|D_{q_j}\zeta_i^m(t,\cdot)\big|_\infty \leq \begin{cases} C_3(t,r), & i = j, \\ \frac{C_3(t,r)}{m}, & i \neq j, \end{cases} \quad \text{in } \mathbb{B}_r^m(0).$$

Since $\xi^m(t,\cdot) : \mathbb{M} \to \mathbb{M}$ is a diffeomorphism, we have

$$D_q\zeta^m(t,q) = \big(D_z\xi^m(t,\cdot)\big)^{-1} \circ \zeta^m(t,q). \tag{2.19}$$

Since we have a uniform lower bound on $\det(D_z\xi(t,\cdot))$ in $\mathbb{M}^m$, we can simply study the asymptotic behavior of $D_q\zeta^m(t,q)$ with respect to m via the asymptotic behavior of $(D_z\xi^m(t,\cdot))^{-1}$. By the previous uniform local estimates on

$D_z\xi^m(t,\cdot)$ (from Claim 1), we have that there exists a constant $C(t,r) > 0$ depending on $C_2(t,r)$ such that

$$D_z\xi^m(t,\cdot) \sim C(t,r) \begin{bmatrix} A_d & \frac{1}{m}A_d & \frac{1}{m}A_d & \cdots & \frac{1}{m}A_d \\ \frac{1}{m}A_d & A_d & \frac{1}{m}A_d & \cdots & \frac{1}{m}A_d \\ \cdots & \cdots & \cdots & \ddots & \cdots \\ \frac{1}{m}A_d & \frac{1}{m}A_d & \frac{1}{m}A_d & \cdots & A_d \end{bmatrix} \tag{2.20}$$

for some invertible $(d \times d)$-blocks A_d. Therefore,

$$(D_z\xi(t,\cdot))^{-1} \sim \frac{1}{C(t,r)} \begin{bmatrix} \frac{m}{m-\frac{1}{2}}A_d^{-1} & \frac{-m}{(2m-1)(m-1)}A_d^{-1} & \frac{-m}{(2m-1)(m-1)}A_d^{-1} & \cdots & \frac{-m}{(2m-1)(m-1)}A_d^{-1} \\ \frac{-m}{(2m-1)(m-1)}A_d^{-1} & \frac{m}{m-\frac{1}{2}}A_d^{-1} & \frac{-m}{(2m-1)(m-1)}A_d^{-1} & \cdots & \frac{-m}{(2m-1)(m-1)}A_d^{-1} \\ \cdots & \cdots & \cdots & \ddots & \cdots \\ \frac{-m}{(2m-1)(m-1)}A_d^{-1} & \frac{-m}{(2m-1)(m-1)}A_d^{-1} & \frac{-m}{(2m-1)(m-1)}A_d^{-1} & \cdots & \frac{m}{m-\frac{1}{2}}A_d^{-1} \end{bmatrix}$$

and so Claim 2 follows by setting $C_3(t,r) := C(t,r)^{-1}$.

Going forward to conclude the proof of assertion (1) of the theorem, we recall that by (2.13),

$$\eta_i^m(t,\zeta^m(t,q)) = D_{q_i}U^{(m)}(t,q).$$

Differentiating this expression with respect to q_j yields

$$\begin{aligned} D_{q_jq_i}U^{(m)}(t,q) &= \sum_{l=1}^{m} D_{q_l}\big(\eta_i(t,\zeta^m(t,q))\big)D_{q_j}\zeta_l^m(t,q) \\ &= D_{q_j}\eta_i^m(t,\zeta^m(t,q))D_{q_j}\zeta_j^m(t,q) \\ &\quad + D_{q_i}\eta_i^m(t,\zeta^m(t,q))D_{q_j}\zeta_i^m(t,q) \\ &\quad + \sum_{l\neq i, l\neq j} D_{q_l}\eta_i^m(t,\zeta(t,q))D_{q_j}\zeta_l^m(t,q). \end{aligned}$$

The previous estimates established in Claim 1 and Claim 2 yield assertion (1).

Step 3. *Estimates on* $(D^2_{z_kz_j}\xi_i^m(t,\cdot), D^2_{z_kz_j}\eta_i^m(t,\cdot))_{i,j,k=1}^m$.

CLAIM 3. There exists a constant $C_4(t,r) > 0$ depending on all the previous ones, but independent of m such that if $\xi(t,z) = q \in \mathbb{B}_r^m(0)$; then for all $i,j,k \in \{1,\dots,m\}$ we have

$$\left|D^2_{z_kz_j}\xi_i^m(t,\cdot)\right|_\infty \le \begin{cases} C_4(t,r), & i = j = k, \\ \dfrac{C_4(t,r)}{m}, & i = j \neq k,\ i \neq j = k,\ i = k \neq j, \\ \dfrac{C_4(t,r)}{m^2}, & i \neq j \neq k, \end{cases}$$

and

$$\left|D^2_{z_k z_j}\eta_i^m(t,\cdot)\right|_\infty \le \begin{cases} \dfrac{C_4(t,r)}{m}, & i=j=k,\\ \dfrac{C_4(t,r)}{m^2}, & i=j\neq k,\ i\neq j=k,\ i=k\neq j,\\ \dfrac{C_4(t,r)}{m^3}, & i\neq j\neq k. \end{cases}$$

PROOF OF CLAIM 3. Differentiating the system (2.18) with respect to z_k, we obtain for the first equation

$$\begin{aligned}\partial_t D^2_{z_k z_j}\xi_i^m &\\ &= D_{z_k}\xi_i^m D^3_{qqp}H(\xi_i^m, m\eta_i)D_{z_j}\xi_i^m\\ &\quad + mD_{z_k}\eta_i^m D^3_{pqp}H(\xi_i^m, m\eta_i)D_{z_j}\xi_i^m\\ &\quad + D^2_{qp}H(\xi_i^m, m\eta_i^m)D^2_{z_k z_j}\xi_i^m + mD_{z_k}\xi_i D^3_{qpp}H(\xi_i^m, m\eta_i)D_{z_j}\eta_i^m\\ &\quad + m^2 D_{z_k}\eta_i^m D^3_{ppp}H(\xi_i^m, m\eta_i)D_{z_j}\eta_i^m + mD^2_{pp}H(\xi_i^m, m\eta_i)D^2_{z_k z_j}\eta_i^m\end{aligned}\tag{2.21}$$

together with the initial condition $D^2_{z_k z_j}\xi_i^m(0,\cdot) = 0_{d\times d\times d}$. From the differentiation of the second equation with respect to z_k, we obtain

$$\begin{aligned}\partial_t D^2_{z_k z_j}\eta_i^m &\\ &= -\frac{1}{m}\big(D_{z_k}\xi_i D^3_{qqq}H(\xi_i^m, m\eta_i^m)D_{z_j}\xi_i^m + mD_{z_k}\eta_i D^3_{pqq}H(\xi_i^m, m\eta_i^m)D_{z_j}\xi_i^m\big)\\ &\quad - \frac{1}{m}\big(D^2_{qq}H(\xi_i^m, m\eta_i^m)D^2_{z_k z_j}\xi_i^m D_{z_k}\xi_i^m + D^3_{qpq}H(\xi_i^m, m\eta_i^m)D_{z_j}\eta_i^m\big)\\ &\quad - \frac{1}{m}\big(m^2 D_{z_k}\eta_i D^3_{ppx}H(\xi_i^m, m\eta_i^m)D_{z_j}\eta_i^m + mD^2_{pq}H(\xi_i, m\eta_i^m)D^2_{z_k z_j}\eta_i^m\big)\\ &\quad + \sum_{l_1,l_2=1}^m D_{z_k}\xi^m_{l_1}D^3_{q_{l_1}q_{l_2}q_i}F^{(m)}(\xi^m)D_{z_j}\xi^m_{l_2} + \sum_{l=1}^m D^2_{q_l q_i}F^{(m)}(\xi^m)D^2_{z_k z_j}\xi^m_l\end{aligned}\tag{2.22}$$

with the initial condition

$$D^2_{z_k z_j}\eta_i^m(0,z) = D^3_{q_k q_j q_i}U_0^{(m)}(z).\tag{2.23}$$

Let us fix k, j. The asymptotic behavior of $(D_{z_k z_j}\xi_i^m(t,\cdot), D_{z_k z_j}\eta_i^m(t,\cdot))$, as the solution to the system (2.21)–(2.22), can be studied in the same way as for (2.18) in Step 1. For this, one needs to identify the precise bounds on the coefficient matrices in (2.21)–(2.22). Let us set

$$\bar C_4 := \max\{|\partial_q^\alpha\partial_p^\beta H(q,p)| : (q,p)\in\mathbb{R}^d\times\mathbb{R}^d,\ 2\le|\alpha|+|\beta|\le 3\};$$

then we notice that by the assumptions on H, we have that if $\xi^m(t,z) = q\in\mathbb{B}^m_r(0)$, then

$$\left|\partial_q^\alpha\partial_p^\beta H(\xi_i^m(t,z), m\eta_i^m(t,z))\right| \le \bar C_4.$$

In the same way, there exists $\widetilde C_4(t,r) > 0$ (depending on $\beta(t,r)$) such that

$$D^3_{q_k q_j q_i}F^{(m)}(\xi^m) \quad\text{and}\quad D^3_{q_k q_j q_i}U_0^{(m)}(q)$$

satisfy the estimate (2.11) with $\widetilde{C}_4(t,r)$. Set

$$\widehat{C}_4(t,r) := \max\{\overline{C}_4, \widetilde{C}_4(t,t)\} \max\{C_2(t,r), 1\}^2.$$

Now, system (2.21)–(2.22) has the same structure as (2.24), where the quantities $(D^2_{q_k q_j}\xi_i^m, D^2_{q_k q_j}\eta_i^m)$ play the role of (X_i, Y_i). The blocks $B_1, \dots, B_4$, the coefficient blocks appearing in (2.24), can be identified in the same way as in Step 1. It remains to study the bounds on the corresponding A_1, A_2, and Y_0 appearing in this system, where

$$\begin{aligned}(A_1)_i := {} & D_{z_k}\xi_i D^3_{qqp} H(\xi_i^m, m\eta_i) D_{z_j}\xi_i^m + m D_{z_k}\eta_i^m D^3_{pqp} H(\xi_i^m, m\eta_i^m) D_{z_j}\xi_i^m \\ & + m D_{z_k}\xi_i^m D^3_{qpp} H(\xi_i^m, m\eta_i^m) D_{z_j}\eta_i^m + m^2 D_{z_k}\eta_i^m D^3_{ppp} H(\xi_i^m, m\eta_i^m) D_{z_j}\eta_i^m, \\ (A_2)_i := {} & -\frac{1}{m}\big(D_{z_k}\xi_i D^3_{qqq} H(\xi_i^m, m\eta_i^m) D_{z_j}\xi_i^m + m D_{z_k}\eta_i^m D^3_{pqq} H(\xi_i^m, m\eta_i) D_{z_j}\xi_i^m\big) \\ & - \frac{1}{m}\big(D_{z_k}\xi_i^m m D^3_{qpq} H(\xi_i^m, m\eta_i^m) D_{z_j}\eta_i^m + m^2 D_{z_k}\eta_i^m D^3_{ppq} H(\xi_i^m, m\eta_i^m) D_{z_j}\eta_i^m\big) \\ & + \sum_{l_1,l_2=1}^m D_{z_k}\xi_{l_1}^m D^3_{q_{l_1} q_{l_2} q_i} F^{(m)}(\xi^m) D_{z_j}\xi_{l_2}^m,\end{aligned}$$

and we set

$$(Y_0)_i := D^3_{q_k q_j q_i} U_0^{(m)}.$$

Using the obtained bounds on $(D_{z_j}\xi_i, D_{z_j}\eta_i)$ in Step 1 and the assumptions on $U_0^{(m)}$ in (2.11), one checks the following asymptotic properties with respect to m.

Subclaim 3.

(1) If $k = j = i$, then $(A_1)_i = O(\widehat{C}_4(t,r))$, $(A_2)_i = O(\frac{\widehat{C}_4(t,r)}{m})$, and $(Y_0)_i = O(\frac{\widehat{C}_4(t,r)}{m})$.

(2) If $k = j \neq i$, then $(A_1)_i = O(\frac{\widehat{C}_4(t,r)}{m^2})$, $(A_2)_i = O(\frac{\widehat{C}_4(t,r)}{m^2})$, and $(Y_0)_i = O(\frac{\widehat{C}_4(t,r)}{m^2})$.

(3) If $k = i \neq j$ or $i = j \neq k$, $(A_1)_i = O(\frac{\widehat{C}_4(t,r)}{m})$, $(A_2)_i = O(\frac{\widehat{C}_4(t,r)}{m^2})$, and $(Y_0)_i = O(\frac{\widehat{C}_4(t,r)}{m^2})$.

(4) If $k \neq j \neq i$, then $(A_1)_i = O(\frac{\widehat{C}_4(t,r)}{m^2})$, $(A_2)_i = O(\frac{\widehat{C}_4(t,r)}{m^3})$, and $(Y_0)_i = O(\frac{\widehat{C}_4(t,r)}{m^3})$,

Now, one considers two cases when studying the desired properties. Let us recall that k, j are fixed.

Case 1. If $k = j$, (1) and (2) of Subclaim 3 can be combined with Lemma 2.5(1) to conclude the proof of the claim.

Case 2. If $k \neq j$, (3) and (4) of Subclaim 3 can be combined with Lemma 2.5(2) to conclude the proof of the claim.

Therefore there exists a constant $C > 0$ such that Claim 3 holds for $C_4(t,r) := e^{tC\widehat{C}_4(t,r)}$.

Step 4. *Estimates on* $(D^2_{q_k q_j}\zeta_i(t,\cdot))^m_{i,j,k=1}$.

CLAIM 4. There exists a constant $C_5(t,r) > 0$ depending on all the previous ones but independent of m such that for all $i,j,k \in \{1,\dots,m\}$, we have

$$\left|D^2_{x_k x_j}\zeta_i(t,\cdot)\right|_\infty \le \begin{cases} C_5(t,r), & i=j=k, \\ \dfrac{C_5(t,r)}{m}, & i=j\neq k,\ i\neq j=k,\ i=k\neq j, \\ \dfrac{C_5(t,r)}{m^2}, & i\neq j\neq k, \end{cases} \quad \text{in } \mathbb{B}^m_r.$$

PROOF OF CLAIM 4. It is enough to differentiate the expression (2.19) and use all the previous estimates on $(D^2_{z_k z_j}\xi_i)^m_{i,j,k=1}$ and on $(D_{q_j}\zeta_i)^m_{i,j=1}$ from Step 3 and Step 2, respectively.

We have

$$D^2_{qq}\zeta(t,q) = -\left\{\left[\left(D_z\xi(t,\cdot)\right)^{-1} D^2_{zz}\xi(t,\cdot)D_q\zeta(t,q)\left(D_z\xi(t,\cdot)\right)^{-1}\right]\circ\zeta(t,q)\right\}.$$

The previous writing is used for the following shorthand notation: we have

$$D_{q_k}D_q\zeta(t,q) = -\left\{\left[\left(D_z\xi(t,\cdot)\right)^{-1}\left(\sum_{l=1}^m D_{z_l}D_z\xi(t,\cdot)D_{q_k}\zeta_l(t,q)\right)\left(D_z\xi(t,\cdot)\right)^{-1}\right]\circ\zeta(t,q)\right\},$$

for $k\in\{1,\dots,m\}$, and in particular for $i,j\in\{1,\dots,m\}$, we have

$$\left(\sum_{l=1}^m D_{z_l}D_z\xi(t,\cdot)D_{q_k}\zeta_l(t,q)\right)_{ij} = \sum_{l=1}^m D^2_{z_l z_j}\xi_i(t,\cdot)D_{q_k}\zeta_l(t,q) =: A_{ij}.$$

For $k\in\{1,\dots,m\}$ fixed, by the definition of A_{ij} and by Steps 2 and 3, this last matrix can be bounded as follows: by setting $\widetilde{C}_5(t,r) := C_4(t,r)C_3(t,r)$, we have

$$|A_{ij}|_\infty \le \begin{cases} \widetilde{C}_5(t,r), & i=j=k, \\ \dfrac{\widetilde{C}_5(t,r)}{m}, & i=j\neq k,\ i\neq j=k,\ i=k\neq j, \\ \dfrac{\widetilde{C}_5(t,r)}{m^2}, & i\neq j\neq k. \end{cases}$$

Now, using the bounds on $(D_{\mathbf{z}}\xi(t,\cdot))^{-1}$ from (2.20), by setting

$$C_5(t,r) := \widetilde{C}_5(t,r)C(t,r)^2,$$

we conclude the statement of Claim 4.

Final Step. Let us recall that from (2.13) we have

$$\eta_i(t,\zeta(t,q)) = D_{q_i}U^{(m)}(t,q).$$

Differentiating this expression with respect to q_j and q_k, we obtain

$$D^3_{q_k q_j q_i} U^{(m)}(t,\cdot) = \sum_{l_1,l_2=1}^{m} D_{q_k}\zeta_{l_2}(t,\cdot) D^2_{z_{l_2} z_{l_1}} \eta_i(t,\zeta(t,\cdot)) D_{q_j}\zeta_{l_1}(t,\cdot) + \sum_{l=1}^{m} D_{z_l}\eta_i(t,\zeta(t,\cdot)) D^2_{q_k q_j}\zeta_l(t,\cdot)$$

from where by using the estimates from Steps 1–4, we obtain

$$\begin{aligned}
&\big|D^3_{q_k q_j q_i} U^{(m)}(t,\cdot)\big|_\infty \\
&\quad \le \frac{1}{m}\big(|D_{q_k}\zeta_i|_\infty |D_{q_j}\zeta_i|_\infty + |D^2_{q_k q_j}\zeta_i|_\infty\big) \\
&\qquad + \frac{1}{m^2}\Big(\sum_{l=1,l\neq i}^{m} |D_{q_k}\zeta_l|_\infty |D_{q_j}\zeta_i|_\infty + \sum_{l=1,l\neq i}^{m} |D_{q_k}\zeta_i|_\infty |D_{q_j}\zeta_i|_\infty + \sum_{l=1,l\neq i}^{m} |D^2_{q_k q_j}\zeta_l|_\infty\Big) \\
&\qquad + \frac{1}{m^3}\sum_{\substack{l_1,l_2=1\\ l_1\neq l_2\neq i}}^{m} |D_{q_k}\zeta_{l_1}|_\infty |D_{q_j}\zeta_{l_2}|_\infty
\end{aligned}$$

Using again the estimates from the previous steps, we obtain (1) and (2) of the theorem.

The statement in (3) can be easily shown by differentiating the Hamilton–Jacobi equation satisfied by $U^{(m)}$ with respect to the variable q_j and by using the estimates on $U^{(m)}$ provided in (1) and (2). Indeed, we have

$$\begin{aligned}
\big|D_{q_j}\partial_t U^{(m)}\big| &\le \frac{1}{m}\big|D_q H(q_j, m D_{q_j} U^{(m)})\big| \\
&\quad + \frac{1}{m}\big|D_p H(q_j, m D_{q_j} U^{(m)})|m|D^2_{q_j q_j} U^{(m)}\big| \\
&\quad + \sum_{i\neq j}\frac{1}{m}\big|D_p H(q_i, m D_{q_i} U^{(m)})\big|m\big|D^2_{q_j q_i} U^{(m)}\big| + \big|D_{q_j} F^{(m)}\big| \\
&\le \frac{1}{m}\big|D_q H\big(q_j, m D_{q_j} U^{(m)}\big)\big| + \frac{1}{m}\big|D_p H\big(q_j, m D_{q_j} U^{(m)}\big)\big| \\
&\quad + \frac{C}{m} + |D_{q_j} F^{(m)}|.
\end{aligned}$$

Thus

$$\begin{aligned}\sum_{j=1}^{m} m|D_{q_j}\partial_t U^{(m)}|^2 &\le \sum_{j=1}^{m}\frac{1}{m}|D_q H(q_j, mD_{q_j}U^{(m)})|^2 + \sum_{j=1}^{m}\frac{1}{m}|D_p H(q_j, mD_{q_j}U^{(m)})|^2\\ &\quad + C + \sum_{j=1}^{m} m|D_{q_j}F^{(m)}|^2 \le C,\end{aligned}$$

where we used the assumption on $F^{(m)}$, (H15), and the fact that since

$$\mathscr{U} \in C^{1,1}_{\mathrm{loc}}([0,T]\times\mathscr{P}_2(\mathbb{M}))$$

and D_pH is Lipschitz, we have $\sum_{j=1}^{m}\frac{1}{m}|D_pH(q_j, mD_{q_j}U^{(m)})|^2 \le C$. The claim follows, which concludes the proof of the theorem. □

LEMMA 2.5. *Let* $[X\ Y]^\top = [X_1\ \dots\ X_m\ Y_1\ \dots\ Y_m]^\top \in \mathbb{R}^{2m}$ *be the solution of the ODE system*

$$\partial_t\begin{bmatrix}X\\Y\end{bmatrix} = \begin{bmatrix}A_1\\A_2\end{bmatrix} + \begin{bmatrix}B_1 & B_2\\B_3 & B_4\end{bmatrix}\begin{bmatrix}X\\Y\end{bmatrix},\quad \begin{bmatrix}X(0)\\Y(0)\end{bmatrix} = \begin{bmatrix}0_m\\Y_0\end{bmatrix}, \tag{2.24}$$

where $A_1, A_2, Y_0 \in \mathbb{R}^m$, $0_m \in \mathbb{R}^m$ *is the zero vector, and the* $(m\times m)$*-dimensional blocks* B_i *are such that*

$$B_1 = B_4 = I_m,\ B_2 = mI_m,\ \text{and} B_3 = \begin{bmatrix}\frac{1}{m} & \frac{1}{m^2} & \cdots & \frac{1}{m^2}\\ \frac{1}{m^2} & \frac{1}{m} & \cdots & \frac{1}{m^2}\\ \cdots & \cdots & \ddots & \cdots\\ \frac{1}{m^2} & \cdots & \frac{1}{m^2} & \frac{1}{m}\end{bmatrix}. \tag{2.25}$$

Then there exists a constant $C > 0$ (*independent of* m) *such that*

(1) *If for* $i_0 \in \{1,\dots,m\}$ *fixed*

$$(A_1)_{i_0} = 1,\ (A_1)_i = \frac{1}{m}\ \ \forall i \ne i_0$$

and

$$(A_2)_{i_0} = (Y_0)_{i_0} = \frac{1}{m},\ (A_2)_i = (Y_0)_i = \frac{1}{m^2}\ \ \forall i \ne i_0,$$

then

$$|X_i(t)| \le \begin{cases} e^{tC}, & i = i_0,\\ \frac{e^{tC}}{m}, & i \in \{1,\dots,m\},\ i \ne i_0,\end{cases}$$

and

$$|Y_i(t)| \le \begin{cases} \frac{e^{tC}}{m}, & i = i_0, \\ \frac{e^{tC}}{m^2}, & i \in \{1,\dots,m\},\ i \neq i_0. \end{cases}$$

(2) *If for some* $k, j \in \{1,\dots,m\}$ *fixed,* $k \neq j$*, we have*

$$(A_1)_j = (A_1)_k = \frac{1}{m},\ \ (A_1)_i = \frac{1}{m^2}\ \ \forall i \neq j,\ i \neq k,$$

and

$$(A_2)_j = (A_2)_k = (Y_0)_j = (Y_0)_k = \frac{1}{m^2},$$
$$(A_2)_i = (Y_0)_i = \frac{1}{m^3}\ \ \forall i \neq j,\ i \neq k,$$

then

$$|X_i(t)| \le \begin{cases} \frac{e^{tC}}{m}, & i = j,\ i = k, \\ \frac{e^{tC}}{m^2}, & i \in \{1,\dots,m\},\ i \neq j,\ i \neq k, \end{cases}$$

and

$$|Y_i(t)| \le \begin{cases} \frac{e^{tC}}{m^2}, & i = j,\ i = k, \\ \frac{e^{tC}}{m^3}, & i \in \{1,\dots,m\},\ i \neq j,\ i \neq k. \end{cases}$$

PROOF. We analyze the representation formula for (2.24) in the different cases. Since we are only interested in the asymptotic properties of the solution with respect to m, first let us study the asymptotic behavior of the exponential and the inverse of the coefficient matrix.

Let $B := \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}$ and for $n \in \mathbb{N}$, let us denote the powers of B as $B^n := \begin{bmatrix} B_{1,n} & B_{2,n} \\ B_{3,n} & B_{4,n} \end{bmatrix}$.

CLAIM. We have the following properties for the blocks $B_{i,n}$ for all $n \in \mathbb{N}$ and for $i, j \in \{1,\dots,m\}$:

(1) $(B_{1,n})_{ii} = O(1)$, $(B_{1,n})_{ij} = O(\frac{1}{m})$, if $i \neq j$.
(2) $(B_{2,n})_{ii} = O(m)$, $(B_{2,n})_{ij} = O(1)$, if $i \neq j$.
(3) $(B_{3,n})_{ii} = O(\frac{1}{m})$, $(B_{3,n})_{ij} = O(\frac{1}{m^2})$, if $i \neq j$.
(4) $(B_{4,n})_{ii} = O(1)$, $(B_{4,n})_{ij} = O(\frac{1}{m})$, if $i \neq j$.

PROOF OF THE CLAIM. This follows from a mathematical induction argument in n.

Since we have a characterization of the asymptotic properties in terms of m of the elements of the powers $n \in \mathbb{N}$ of the block matrix (which are uniform in n), the property from the Claim will also hold true for the blocks of the matrix exponential of B. Setting $A := [A_1^\top \ A_2^\top]^\top$, the representation formula for the solutions of (2.24) reads as

$$\begin{bmatrix} X(t) \\ Y(t) \end{bmatrix} = \exp(tB)\big(\big[0_m^\top \ Y_0^\top\big]^\top + B^{-1}A\big) - B^{-1}A.$$

It remains to compute B^{-1} (which exists, since B is nonsingular), for which we have the formula (using the blocks from (2.25))

$$\begin{aligned} B^{-1} &= \begin{bmatrix} (I_m - mB_3)^{-1} & -m(I_m - mB_3)^{-1} \\ -B_3(I_m - mB_3)^{-1} & I_m + mB_3(I_m - mB_3)^{-1} \end{bmatrix} \\ &= \begin{bmatrix} M & -mM \\ -B_3M & I_m + mB_3M \end{bmatrix}, \end{aligned}$$

where we have used the notation

$$\begin{aligned} M := (I_m - mB_3)^{-1} &= m \begin{bmatrix} 0 & -1 & \cdots & -1 \\ -1 & 0 & \cdots & -1 \\ \cdots & \cdots & \ddots & \cdots \\ -1 & \cdots & -1 & 0 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} m\frac{m-2}{m-1} & \frac{-m}{m-1} & \cdots & \frac{-m}{m-1} \\ \frac{-m}{m-1} & m\frac{m-2}{m-1} & \cdots & \frac{-m}{m-1} \\ \cdots & \cdots & \ddots & \cdots \\ \frac{-m}{m-1} & \cdots & \frac{-m}{m-1} & m\frac{m-2}{m-1} \end{bmatrix}. \end{aligned}$$

Now, in the case of (1), we have that $(B^{-1}A)_i = 0$ if $i \in \{1, \dots, m\}$, and $(B^{-1}A)_{m+i_0} = \frac{1}{m}$ and $(B^{-1}A)_i = \frac{1}{m^2}$ if $i \in \{m+1, \dots, 2m\}$, $i \neq m + i_0$. Furthermore, there exists a constant $C > 0$ (independent of m) such that

$$\big(\exp(tB)[0_m^\top \ Y_0^\top]^\top\big)_i \sim \begin{cases} e^{tC}, & i = i_0, \\ \frac{e^{tC}}{m}, & i \in \{1, \dots, m\},\ i \neq i_0, \\ \frac{e^{tC}}{m}, & i = m + i_0, \\ \frac{e^{tC}}{m^2}, & i \in \{m+1, \dots, 2m\},\ i \neq m + i_0. \end{cases}$$

(1) from the thesis of the lemma follows.

In the case of (2), we compute similarly $(B^{-1}A)_i = 0$ if $i \in \{1, \dots, m\}$, $(B^{-1}A)_i = \frac{1}{m^2}$ if $i = m + j$ or $j = m + k$, and $(B^{-1}A)_i = \frac{1}{m^3}$ otherwise.

Furthermore, there exists a constant $C > 0$ (independent of m) such that

$$\left(\exp(tB)[0_m^\top\ Y_0^\top]^\top\right)_i \sim \begin{cases} \frac{e^{tC}}{m}, & i = j,\ i = k, \\ \frac{e^{tC}}{m^2}, & i \in \{1,\dots,m\},\ i \neq j,\ i \neq k, \\ \frac{e^{tC}}{m^2}, & i = m+j,\ i = m+k, \\ \frac{e^{tC}}{m^3}, & i \in \{m+1,\dots,2m\},\ i \neq m+j,\ i \neq m+k. \end{cases}$$

And finally, (2) from the thesis of the lemma follows. □

LEMMA 2.6. *Let $X = (X_{ij})_{i,j=1}^m$ and $Y = (X_{ij})_{i,j=1}^m$ be such that $[X\ Y]^\top \in \mathbb{R}^{2m\times m}$ is the solution of the ODE system*

$$\partial_t \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix}, \quad \begin{bmatrix} X(0) \\ Y(0) \end{bmatrix} = \begin{bmatrix} I_m \\ Y_0 \end{bmatrix}, \tag{2.26}$$

where $Y_0 \in \mathbb{R}^{m\times m}$ is set to $Y_0 := B_3$ and the $(m\times m)$-dimensional blocks B_i are defined in (2.25). *Then there exists $C > 0$ (independent of m) such that*

$$|X_{ij}(t)| \le \begin{cases} e^{tC}, & i = j, \\ \frac{e^{tC}}{m}, & i \neq j, \end{cases} \quad \text{and} \quad |Y_{ij}(t)| \le \begin{cases} \frac{e^{tC}}{m}, & i = j, \\ \frac{e^{tC}}{m^2}, & i \neq j. \end{cases}$$

PROOF. This result is a consequence of the asymptotic behavior of the matrix exponential $\exp(tB)$, where $B := \begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}$. Using the asymptotic result from the Claim in Lemma 2.5 and from the representation formula

$$\begin{bmatrix} X(t) \\ Y(t) \end{bmatrix} = \exp(tB)[I_m\ Y_0]^\top, \tag{2.27}$$

the result follows. □

3 Comparing Regularity Properties of Functions Defined on $\mathscr{P}_2(\mathbb{M})$, $\mathbb{H}$, and $\mathbb{M}^m$

Throughout this section, we lift any given function $\mathscr{U} : \mathscr{P}_2(\mathbb{M}) \to \mathbb{R}$ to $\mathbb{H}$ to obtain the function $\widetilde{\mathscr{U}} : \mathbb{H} \to \mathbb{R}$ defined as $\widetilde{\mathscr{U}}(x) := \mathscr{U}(\sharp(x))$. Recall $(\Omega_j)_{j=1}^m$ is the Borel partition in Section 1. We set

$$U^{(m)}(q) := \mathscr{U}(\mu_q^{(m)}) = \widetilde{\mathscr{U}}(M^q).$$

3.1 Semiconvex and semiconcave functions on Hilbert spaces

DEFINITION 3.1 (Semiconvexity and semiconcavity on $\mathbb{H}$). Let $\mathbb{B} \subseteq \mathbb{H}$ be a convex open set. We say that $\widetilde{\mathscr{U}} : \mathbb{B} \to \mathbb{R}$ is semiconvex (or λ-convex) on $\mathbb{B}$, if there exists $\lambda \in \mathbb{R}$ and for all $x \in \mathbb{B}$ there exists a continuous linear form θ_x on $\mathbb{H}$ such that

$$\widetilde{\mathscr{U}}(y) \ge \widetilde{\mathscr{U}}(x) + \theta_x(y-x) + \frac{\lambda}{2}\|x-y\|^2 \quad \forall y \in \mathbb{B}.$$

We say that a function $\widetilde{\mathscr{U}} : \mathbb{B} \to \mathbb{R}$ is λ-concave, if $-\widetilde{\mathscr{U}}$ is $(-\lambda)$-semiconvex.

Remark 3.2. The previous definition has an equivalent reformulation. Let $\mathbb{B} \subseteq \mathbb{H}$ be a convex open set. Then $\widetilde{\mathscr{U}} : \mathbb{B} \to \mathbb{R}$ is λ-convex if and only if

$$\widetilde{\mathscr{U}}((1-t)x + ty) \le (1-t)\widetilde{\mathscr{U}}(x) + t\widetilde{\mathscr{U}}(y) - \frac{\lambda}{2}t(1-t)\|x-y\|^2 \quad \forall t \in [0,1],\ \forall x, y \in \mathbb{B}.$$

DEFINITION 3.3 ($C^{1,1}$ functions). We say that $\widetilde{\mathscr{U}} : \mathbb{B} \to \mathbb{R}$ is $C^{1,1}$ on an open set $\mathbb{B} \subseteq \mathbb{H}$ if it is Fréchet differentiable on $\mathbb{B}$ and its Fréchet differential is Lipschitz-continuous; i.e., there exists $C > 0$ such that

$$\|\nabla\widetilde{\mathscr{U}}(x) - \nabla\widetilde{\mathscr{U}}(y)\| \le C\|x-y\| \forall x, y \in \mathbb{B}.$$

Inspired by similar results on finite-dimensional smooth manifolds (see, for instance, in [27]), we can state the following characterization of $C^{1,1}$ functions defined on subsets of $\mathbb{H}$.

Remark 3.4. In fact $\widetilde{\mathscr{U}} : \mathbb{B} \to \mathbb{R}$ is $C^{1,1}$ on a convex set $\mathbb{B} \subseteq \mathbb{H}$ if and only if it is Fréchet differentiable on $\mathbb{B}$ and there exists $K \ge 0$ such that

$$|\widetilde{\mathscr{U}}(y) - \widetilde{\mathscr{U}}(x) - \nabla\widetilde{\mathscr{U}}(x)(y-x)| \le K\|x-y\|^2,\ \forall\ x, y \in \mathbb{B}.$$

3.2 Notions of convexity on $(\mathscr{P}_2(\mathbb{M}), W_2)$

There are various notions of convexity for functionals defined on the Wasserstein space. The concept of so-called *displacement convexity* [6, 40] is expressed in terms of W_2-geodesics. Recall that given $\mu_0, \mu_1 \in \mathscr{P}_2(\mathbb{M})$, for any geodesics $[0,1] \ni t \mapsto \mu_t \in \mathscr{P}_2(\mathbb{M})$, of constant speed connecting μ_0 to μ_1 in $\mathscr{P}_2(\mathbb{M})$ is of the form $\mu_t = \mu_t := ((1-t)\pi^1 + t\pi^1)_\sharp\gamma$ for some $\gamma \in \Gamma_o(\mu_0, \mu_1)$, then:

DEFINITION 3.5 (Semiconvexity and semiconcavity on $(\mathscr{P}_2(\mathbb{M}), W_2)$). Let $\mathscr{U} : \mathscr{P}_2(\mathbb{M}) \to \mathbb{R}$.

(1-i) We say that $\mathscr{U}$ is *semiconvex* (or λ-convex) in the classical sense if there is $\lambda \in \mathbb{R}$ such that

$$\mathscr{U}((1-t)\mu_0 + t\mu_1) \le (1-t)\mathscr{U}(\mu_0) + t\mathscr{U}(\mu_1) - \frac{\lambda}{2}t(1-t)W_2^2(\mu_0, \mu_1),$$

$\forall \mu_0, \mu_1 \in \mathscr{P}_2(\mathbb{M}),\ \forall t \in [0,1].$

(1-ii) We say that $\mathscr{U} : \mathscr{P}_2(\mathbb{M}) \to \mathbb{R}$ is semiconcave (or λ-concave) in the classical sense if $-\mathscr{U}$ is $(-\lambda)$-convex. We refer to 0-convex and 0-concave functions simply as convex and concave functions, respectively.

(2-i) We say $\mathscr{U} : \mathscr{P}_2(\mathbb{M}) \to \mathbb{R}$ is *displacement semiconvex* (or displacement λ-convex) if there exists $\lambda \in \mathbb{R}$ such that for any $[0,1] \ni t \mapsto \mu_t \in \mathscr{P}_2(\mathbb{M})$ geodesic of constant speed connecting μ_0 to μ_1 we have

$$\mathscr{U}(\mu_t) \le (1-t)\mathscr{U}(\mu_0) + t\mathscr{U}(\mu_1) - \frac{\lambda}{2}t(1-t)W_2^2(\mu_0, \mu_1)$$

$\forall \mu_0, \mu_1 \in \mathscr{P}_2(\mathbb{M}),\ \forall t \in [0,1].$

(2-ii) We say that $\mathscr{U} : \mathscr{P}_2(\mathbb{M}) \to \mathbb{R}$ is *displacement semiconcave* (or displacement λ-concave) if $-\mathscr{U}$ is displacement $(-\lambda)$-convex. We refer to displacement 0-convex and displacement 0-concave as simply *displacement convex* and *displacement concave*, respectively.

The following results link λ-convexity on the Wasserstein, the Hilbert, and the finite-dimensional space $\mathbb{M}^m$. This is a generalization of proposition 5.79 in [16].

LEMMA 3.6. *Let $\mathscr{U} : \mathscr{P}_2(\mathbb{M}) \to \mathbb{R}$ be a continuous function, and let $\widetilde{\mathscr{U}} : \mathbb{H} \to \mathbb{R}$ be defined as $\widetilde{\mathscr{U}} := \mathscr{U} \circ \sharp$ so that $\widetilde{\mathscr{U}}$ is continuous. As above, for a natural number m consider $U^{(m)} : \mathbb{M}^m \to \mathbb{R}$. Finally, fix $\lambda \in \mathbb{R}$. Then the following are equivalent.*

(1) *$\widetilde{\mathscr{U}}$ is λ-convex on $\mathbb{H}$.*
(2) *$\mathscr{U}$ is displacement λ-convex on $(\mathscr{P}_2(\mathbb{M}), W_2)$.*
(3) *For any natural number m, we have that $U^{(m)}$ is $\frac{\lambda}{m}$-convex on $\mathbb{M}^m$.*

PROOF. (1)$\Rightarrow$(2). Let us suppose $\widetilde{\mathscr{U}}$ is λ-convex, let $\mu, \nu \in \mathscr{P}(\mathbb{M})$, and let $\gamma \in \Gamma_o(\mu, \nu)$. Then, there exist $x, y \in \mathbb{H}$ such that $(x, y)_\sharp \mathscr{L}^d_\Omega = \gamma$. In particular, we have $\sharp(x) = \mu$, $\sharp(y) = \nu$ and $W_2(\mu, \nu) = \|x - y\|$. For $[0, 1] \ni t \mapsto \mu_t := \left[(1-t)\pi^1 + t\pi^2\right]_\sharp \gamma$ is a geodesic of constant speed connecting μ to ν. Actually, any geodesic between μ and ν has this representation. By the λ-convexity of $\widetilde{\mathscr{U}}$ we have

$$
\begin{aligned}
\mathscr{U}(\mu_t) &= \mathscr{U}\left(\sharp\left[(1-t)x + ty\right]\right) = \widetilde{\mathscr{U}}((1-t)x + ty) \\
&\leq (1-t)\widetilde{\mathscr{U}}(x) + t\widetilde{\mathscr{U}}(y) - \frac{\lambda}{2}t(1-t)\|x-y\|^2 \\
&= (1-t)\mathscr{U}(\mu) + tU(\nu) - \frac{\lambda}{2}t(1-t)W_2^2(\mu, \nu).
\end{aligned}
$$

Thus, $\mathscr{U}$ is displacement λ-convex.

(2)$\Rightarrow$(3). Let us suppose that $\mathscr{U}$ is displacement λ-convex and we show that $U^{(m)}$ is $\frac{\lambda}{m}$-convex on $\mathbb{M}^m$. Let us fix $(q_1, \dots, q_m) \in \mathbb{M}^m$. It is enough to show the $\frac{\lambda}{m}$-convexity of $U^{(m)}$ in a small neighborhood of this fixed point. Therefore, let $(q'_1, \dots, q'_m) \in \mathbb{M}^m$ be such that $\max\{|q_i - q'_i| : i \in \{1, \dots, m\}\}$ is small so that $W_2^2(\mu_q^{(m)}, \mu_{q'}^{(m)}) = \frac{1}{m}\sum_{i=1}^m |q_i - q'_i|^2$. By this assumption, we also have that the constant speed geodesic connecting $\mu_q^{(m)}$ to $\mu_{q'}^{(m)}$ in a unit time is given by $[0, 1] \ni t \mapsto \mu_t^{(m)} = \frac{1}{m}\sum_{i=1}^m \delta_{(1-t)q_i + tq'_i}$.

By this construction, for $t \in [0,1]$ we have

$$\begin{aligned} &U^{(m)}((1-t)q + tq') \\ &= \mathscr{U}(\mu_t^{(m)}) \\ &\leq (1-t)\mathscr{U}(\mu_q^{(m)}) + t\mathscr{U}(\mu_{q'}^{(m)}) - \frac{\lambda}{2}t(1-t)W_2\big(\mu_q^{(m)}, \mu_{q'}^{(m)}\big) \\ &= (1-t)U^{(m)}(q) + tU^{(m)}(q') - \frac{\lambda}{2m}t(1-t)\sum_{i=1}^{m}|q_i - q_i'|^2. \end{aligned}$$

Therefore, the $\frac{\lambda}{m}$-convexity of $U^{(m)}$ in a small neighborhood of q follows.

(3)⇒(1) We suppose $U^{(m)}$ is $\frac{\lambda}{m}$-convex for all natural numbers m. We plan to show the λ-convexity of $\widetilde{\mathscr{U}}$ on $\mathbb{H}$. Note the $\frac{\lambda}{m}$-convexity of $U^{(m)}$ is equivalent to the λ-convexity of the restriction of $\widetilde{\mathscr{U}}$ to $\{M^q : q \in \mathbb{R}^{md}\} \subset \mathbb{H}$. In particular, the local Lipschitz constants of these restrictions are bounded from above by a number that is independent of m. These finite-dimensional functions then have a unique extension $\widetilde{\mathscr{V}}$ on $\mathbb{H}$, which is λ-convex and coincides with $\widetilde{\mathscr{U}}$ on a dense subset of $\mathbb{H}$. It suffices to know that $\widetilde{\mathscr{U}}$ is continuous to conclude that it is nothing but $\widetilde{\mathscr{V}}$. □

3.3 $C^{1,1}$ functions on $(\mathscr{P}_2(\mathbb{M}), W_2)$ versus $C^{1,1}$ functions on $\mathbb{H}$

Given a differentiable function $\mathscr{U} : \mathscr{P}_2(\mathbb{M}) \to \mathbb{R}$ (cf. [6]), we denote as $\nabla_w \mathscr{U}$ the Wasserstein gradient field of $\mathscr{U}$. This subsection exploits the connection between the differential of $\mathscr{U} : \mathscr{P}_2(\mathbb{M}) \to \mathbb{R}$ and the differential of its lift $\widetilde{\mathscr{U}} : \mathbb{H} \to \mathbb{R}$ [32]. More precisely, we have the following result.

Remark 3.7. Let $x \in \mathbb{H}$ and set $\mu := \sharp(x)$. Then $\mathscr{U}$ is differentiable at μ if and only if $\widetilde{\mathscr{U}}$ is differentiable at x and in this case, we have the factorization $\nabla\widetilde{\mathscr{U}}(x) = \nabla_w \mathscr{U}(\mu) \circ x$.

DEFINITION 3.8. Let $\mathscr{B} \subseteq \mathscr{P}_2(\mathbb{M})$ be open and geodesically convex. Let $\alpha \in (0,1]$. We say that $\mathscr{U} \in C^{1,\alpha}(\mathscr{B})$ if it is continuously differentiable on $\mathscr{B}$ and there exists a constant $C \geq 0$ such that

(1) $\mathrm{spt}(\mu) \ni q_1 \mapsto \nabla_w \mathscr{U}(\mu)(q_1)$ is α-Hölder-continuous (or simply Lipschitz-continuous if $\alpha = 1$) with constant C for any $\mu \in \mathscr{B}$.

(2)
$$\begin{aligned} &\left|\mathscr{U}(\nu) - \mathscr{U}(\mu) - \int_{\mathbb{M}^2} \nabla_w \mathscr{U}(\mu)(q_1)\cdot(q_2 - q_1)\,d\gamma(q_1,q_2)\right| \\ &\qquad \leq C W_2^{1+\alpha}(\mu,\nu), \quad \forall \mu,\nu \in \mathscr{B},\ \forall \gamma \in \Gamma_o(\mu,\nu). \end{aligned}$$

DEFINITION 3.9. Similarly to the previous definition, let $\mathscr{B} \subseteq \mathscr{P}_2(\mathbb{M})$ be open and geodesically convex and let $K \subseteq \mathbb{M}$ be a convex open set. Let $\alpha \in (0,1]$. We say that $u \in C^{1,\alpha}(K \times \mathscr{B})$, if it is continuously differentiable on $K \times \mathscr{B}$ and there exists a constant $C \geq 0$ such that

(1) $\mathrm{spt}(\mu) \ni q_1 \mapsto \nabla_w u(q,\mu)(q_1)$ is α-Hölder continuous (or simply Lipschitz-continuous if $\alpha = 1$) with constant C for any $(q,\mu) \in K \times \mathscr{B}$.

(2)
$$\Big| u(\overline{q},\nu) - u(q,\mu) - D_q u(q,\mu)\cdot(\overline{q}-q) - \int_{\mathbb{M}^2} \nabla_w u(q,\mu)(q_1)\cdot(q_2-q_1)\,d\gamma(q_1,q_2)\Big| \le C\left(|\overline{q}-q|^{1+\alpha} + W_2^{1+\alpha}(\mu,\nu)\right),$$
$$\forall \overline{q}, q \in K, \mu,\nu \in \mathscr{B},\ \forall \gamma \in \Gamma_o(\mu,\nu).$$

Remark 3.10. (i) Let us notice that Definition 3.8(2) implies that $\nabla_w \mathscr{U}$ is 'α-Hölder- continuous' in the following sense. We have
$$\Big| \int_{\mathbb{M}^2} \nabla_w \mathscr{U}(\mu)(q_1)\cdot(q_1-q_2)\,d\gamma(q_1,q_2) - \int_{\mathbb{M}^2} \nabla_w \mathscr{U}(\nu)(q_2)\cdot(q_1-q_2)\,d\widetilde{\gamma}(q_2,q_1)\Big| \le 2C W_2^{1+\alpha}(\mu,\nu),$$
for any $\mu,\nu \in \mathscr{B}$ and $\gamma \in \Gamma_o(\mu,\nu)$, $\widetilde{\gamma} \in \Gamma_o(\nu,\mu)$.

(ii) Let us underline that the inequality in Definition 3.8(2) naturally encodes also the fact that $\mathscr{U}$ is locally Lipschitz-continuous. Indeed, that inequality implies that
$$\begin{aligned}|\mathscr{U}(\nu) - \mathscr{U}(\mu)| &\le C W_2^{1+\alpha}(\mu,\nu) + \int_{\mathbb{M}^2} |\nabla_w \mathscr{U}(\mu)(q_1)|\cdot|q_2-q_1|\,d\gamma(q_1,q_2)\\ &\le C W_2^{1+\alpha}(\mu,\nu) + \|\nabla_w \mathscr{U}(\mu)\|_{L^2(\mu)} W_2(\mu,\nu)\\ &= \left(C W_2^{\alpha}(\mu,\nu) + \|\nabla_w \mathscr{U}(\mu)\|_{L^2(\mu)}\right) W_2(\mu,\nu),\end{aligned}$$
so the local Lipschitz property follows.

(iii) Definition 3.9(2) naturally encodes that $K \ni q \mapsto u(q,\mu)$ is of class $C^{1,\alpha}$, uniformly with respect to μ.

LEMMA 3.11. *$\mathscr{U} \in C^{1,1}\big(\mathscr{P}_2(\mathbb{M})\big)$ if and only if $\widetilde{\mathscr{U}} \in C^{1,1}(\mathbb{H})$.*

PROOF.

Part 1. Suppose first that $\widetilde{\mathscr{U}} \in C^{1,1}(\mathbb{H})$ so that by Remark 3.4 there exists a constant $C \ge 0$ such that
$$|\widetilde{\mathscr{U}}(y) - \widetilde{\mathscr{U}}(x) - \nabla\widetilde{\mathscr{U}}(x)(y-x)| \le \frac{C}{2}\|x-y\|^2 \quad \forall x,y \in \mathbb{H}. \tag{3.1}$$
This implies in particular that $\mathscr{U} \in C^1(\mathscr{P}_2(\mathbb{M}))$, and for any $x \in \mathbb{H}$ such that $\sharp(x) = \mu \in \mathscr{P}_2(\mathbb{M})$, we have $\nabla\widetilde{\mathscr{U}}(x) = \nabla_w \mathscr{U}(\mu)\circ x$.

CLAIM. For any $\mu \in \mathscr{P}_2(\mathbb{M})$, $q \mapsto \nabla_w \mathscr{U}(\mu)(q)$ is Lipschitz-continuous on $\mathrm{spt}(\mu)$ uniformly in μ, with Lipschitz constant at most C.

PROOF OF THE CLAIM. Let $\mu \in \mathscr{P}_2(\mathbb{M})$ and consider $x, y \in \mathbb{H}$ such that $\sharp(x) = \sharp(y) = \mu$ and $\|x - y\| > 0$. Since $\nabla\widetilde{\mathscr{U}}$ is Lipschitz-continuous, one has that

$$\|\nabla\widetilde{\mathscr{U}}(x) - \nabla\widetilde{\mathscr{U}}(y)\| \le C\|x - y\|.$$

This is equivalent to

$$\|\nabla_w \mathscr{U}(\mu)(x) - \nabla_w \mathscr{U}(\mu)(y)\| \le C\|x - y\|. \tag{3.2}$$

Suppose that $\mathrm{spt}(\mu)$ contains more than one element; otherwise the statement is trivial. Although x is defined up to a set of measure zero, we are going to choose a representative which is Borel. Set

$$\Omega_0 := \{\omega \in \Omega \mid \omega \ \text{ is a Lebesgue point for } \ x, \nabla\widetilde{\mathscr{U}}(x)\} \cap x^{-1}(\mathrm{spt}\,(\mu)).$$

Note that Ω_0 is a set of full measure in Ω, and so $x(\Omega_0)$ is a set of full μ-measure. In fact, we do not know that $x(\Omega_0)$ is Borel, but we can find a Borel set $A \subset x(\Omega_0)$ of full μ-measure.

We suppose that A has more than one element; otherwise the statement is trivial. Let $q_1, q_2 \in A$ with $q_1 \neq q_2$ and let $q_1^0, q_2^0 \in \Omega_0$ such that $x(q_1^0) = q_1$ and $x(q_2^0) = q_2$. Let $r > 0$ small such that $B_r(q_1^0) \cap B_r(q_2^0) = \varnothing$. Set

$$S_r(\omega) := \begin{cases} \omega & \text{if } \omega \in \Omega \setminus \big(B_r(q_1^0) \cup B_r(q_2^0)\big), \\ \omega - q_1^0 + q_2^0 & \text{if } \omega \in B_r(q_1^0), \\ \omega - q_2^0 + q_1^0 & \text{if } \omega \in B_r(q_2^0). \end{cases} \tag{3.3}$$

Since S_r preserves $\mathscr{L}^d \llcorner \Omega$, x and $y := x \circ S_r$ have the same law μ. We notice that in particular

$$y = x\chi_{\mathbb{M}\setminus(B_r(q_1^0)\cup B_r(q_2^0))} + x\big(\cdot + q_2^0 - q_1^0\big)\chi_{B_r(q_1^0)} + x\big(\cdot + q_1^0 - q_2^0\big)\chi_{B_r(q_2^0)}.$$

Since q_1 and q_2 are distinct image points of x for $r > 0$ sufficiently small,

$$\begin{aligned}\|x - y\|^2 = &\int_{B_r(q_1^0)} \big|x(z) - x\big(z + q_2^0 - q_1^0\big)\big|^2\, dz \\ &+ \int_{B_r(q_2^0)} |x(z) - x(z + q_1^0 - q_2^0)|^2\, dz > 0.\end{aligned}$$

Similarly, (3.2) yields

$$
\begin{aligned}
&\|\nabla_w \mathscr{U}(\mu)(x) - \nabla_w \mathscr{U}(\mu)(y)\|^2 \\
&\quad = \int_{B_r(q_1^0)} \left|\nabla_w \mathscr{U}(\mu)(x(z)) - \nabla_w \mathscr{U}(\mu)\big(x\big(z + q_2^0 - q_1^0\big)\big)\right|^2 dz \\
&\qquad + \int_{B_r(q_2^0)} \left|\nabla_w \mathscr{U}(\mu)(x(z)) - \nabla_w \mathscr{U}(\mu)\big(x\big(z + q_1^0 - q_2^0\big)\big)\right|^2 dz \\
&\quad \le C^2 \bigg(\int_{B_r(q_1^0)} |x(z) - x(z + q_2^0 - q_1^0)|^2 dz \\
&\qquad\qquad + \int_{B_r(q_2^0)} |x(z) - x(z + q_1^0 - q_2^0)|^2 \, dz \bigg)
\end{aligned}
$$

Now, dividing the inequality by $\mathscr{L}^d(B_r(q_1^0))$ and sending $r \downarrow 0$, since q_1^0 and q_2^0 are Lebesgue points of x with $x(q_1^0) = q_1$ and $x(q_2^0) = q_2$, one obtains that

$$
|\nabla_w \mathscr{U}(\mu)(q_1) - \nabla_w \mathscr{U}(\mu)(q_2)| \le C|q_1 - q_2|,
$$

as desired. The claim follows.

Now, let $\mu, \nu \in \mathscr{P}(\mathbb{M})$ and $x, y \in \mathbb{H}$ such that $\sharp(x) = \mu$, $\sharp(y) = \nu$, and $W_2(\mu, \nu) = \|x - y\|$. Let us note that $\gamma := \sharp(x, y) \in \Gamma_o(\mu, \nu)$. We have

$$
\begin{aligned}
\nabla \widetilde{\mathscr{U}}(x)(y - x) &= \int_\Omega \nabla_w \mathscr{U}(\mu)(x(\omega)) \cdot (y(\omega) - x(\omega)) d\omega \\
&= \int_{\mathbb{M}^2} \nabla_w \mathscr{U}(\mu)(q_1) \cdot (q_2 - q_1) d\gamma(q_1, q_2).
\end{aligned}
$$

Thus, by (3.1)

$$
\left| \mathscr{U}(\nu) - \mathscr{U}(\mu) - \int_{\mathbb{M}^2} \nabla_w \mathscr{U}(\mu)(q_1) \cdot (q_2 - q_1) d\gamma(q_1, q_2) \right| \le \frac{C}{2} W_2^2(\mu, \nu),
$$

which by the arbitrariness of μ, ν implies the statement.

Part 2. We now need to prove the reversed implication and start by assuming that $\mathscr{U}$ is $C^{1,1}(\mathscr{P}_2(\mathbb{M}))$. In particular, $\nabla_w \mathscr{U}(\mu)(\cdot)$ is C-Lipschitz-continuous on $\mathrm{spt}(\mu)$ (uniformly in μ) and increasing the value of C if necessary, we assume the inequality in Definition 3.8(2) to hold with the same constant C. Take $x, y \in \mathbb{H}$ and set $\mu := \sharp(x)$ and $\nu := \sharp(y)$. Recall

$$
\widetilde{\mathscr{U}} \in C^1(\mathbb{H}) \quad \text{and} \quad \nabla \widetilde{\mathscr{U}}(x) = \nabla_w \mathscr{U}(\mu) \circ x.
$$

Let $\gamma := \sharp(x, y)$ and let $\gamma_0 \in \Gamma_o(\mu, \nu)$. We have

$$
\begin{aligned}
&\left|\widetilde{\mathscr{U}}(y) - \widetilde{\mathscr{U}}(x) - \nabla\widetilde{\mathscr{U}}(x)(y-x)\right| \\
&\quad = \left|\mathscr{U}(\nu) - \mathscr{U}(\mu) - \int_{\mathbb{M}^2} \nabla_w \mathscr{U}(\mu)(q_1)\cdot(q_2-q_1)d\gamma(q_1,q_2)\right| \\
&\quad \le \left|\mathscr{U}(\nu) - \mathscr{U}(\mu) - \int_{\mathbb{M}^2} \nabla_w \mathscr{U}(\mu)(q_1)\cdot(q_2-q_1)d\gamma_0(q_1,q_2)\right| \\
&\qquad + \left|\int_{\mathbb{M}^2} \nabla_w \mathscr{U}(\mu)(q_1)\cdot(q_2-q_1)d(\gamma_0-\gamma)(q_1,q_2)\right| \\
&\quad \le C W_2^2(\mu,\nu) \\
&\qquad + \frac{1}{2}\left\|D_q\nabla_w\mathscr{U}(\mu)\right\|_{L^\infty} \\
&\qquad \cdot \left(\int_{\mathbb{M}^2} |q_1-q_2|^2 d\gamma(q_1,q_2) + \int_{\mathbb{M}^2} |q_1-q_2|^2 d\gamma_0(q_1,q_2)\right) \\
&\quad \le C W_2^2(\mu,\nu) + \frac{1}{2}C\left(\|x-y\|^2 + W_2^2(\mu,\nu)\right) \le 2C\|x-y\|^2,
\end{aligned}
$$

where in the penultimate line we used an inequality from lemma 3.3 in [32]. Indeed, according to this lemma if $\gamma_1, \gamma_2 \in \Gamma(\mu,\nu)$ and $\xi \in C_c^2(\mathbb{M})$, then

$$
\begin{aligned}
&\left|\int_{\mathbb{M}^2} D\xi(q_1)\cdot(q_2-q_1)d(\gamma_1-\gamma_2)(q_1,q_2)\right| \\
&\quad \le \frac{1}{2}\|D^2\xi\|_{L^\infty}\left(\int_{\mathbb{M}^2}|q_1-q_2|^2 d(\gamma_1+\gamma_2)(q_1,q_2)\right).
\end{aligned}
$$

Since $\nabla_w\mathscr{U}(\mu)$ is the limit of $(D\xi_n)_{n\in\mathbb{N}}$ (where $(\xi_n)_{n\in\mathbb{N}} \in C_c^\infty(\mathbb{M})$) in $L^2_\mu(\mathbb{M};\mathbb{R}^d)$ and $\nabla_w\mathscr{U}(\mu)$ has a global Lipschitz-continuous extension to $\mathbb{M}$, it is easy to see that the previous inequality is still valid for $D\xi = \nabla_w\mathscr{U}(\mu)$ (for which we use its Lipschitz-continuous extension to $\mathbb{M}$).

This completes the verification of the proof of the lemma. □

Remark 3.12.

(i) It seems an interesting open problem whether the equivalence in Lemma 3.11 holds for $C^{1,\alpha}$ functions for $\alpha \in (0,1)$.

(ii) The uniform Lipschitz continuity property of $q \mapsto \nabla_w\mathscr{U}(\mu)(q)$, from the proof of Lemma 3.11, appeared already in [15, lemma 3.3] and in [16, prop. 5.36]. However, not only is our proof based on a different approach, it is considerably shorter and will be useful in the proof of Lemma A.1.

DEFINITION 3.13. Let $\mathscr{B} \subseteq \mathscr{P}_2(\mathbb{M})$ be open and geodesically convex and let $\alpha \in (0,1]$. We say that $\mathscr{U} \in C^{2,\alpha,w}(\mathscr{B})$ if $\mathscr{U} \in C^{1,\alpha}(\mathscr{B})$, and if there exist a constant $C > 0$ and functions

$$
\Lambda_0 : \mathbb{R}^d \times \mathscr{B} \to \mathbb{R}^{d\times d}, \quad \Lambda_1 : \mathbb{M}^2 \times \mathscr{B} \to \mathbb{R}^{d\times d},
$$

such that

(1) $$\Lambda_0 \in L^\infty(\mathbb{M};\mu), \quad \Lambda_1 \in L^\infty(\mathbb{M}^2;\mu\otimes\mu),$$

$$\Big|\nabla_w \mathscr{U}(\nu)(\bar{q}_1) - \nabla_w \mathscr{U}(\mu)(q_1) - \Lambda_0(q_1,\mu)(\bar{q}_1 - q_1) - \int_{\mathbb{M}^2} \Lambda_1(q_1,a,\mu)(b-a)\,d\gamma(a,b)\Big| \le C\left(|q_1-\bar{q}_1|^{1+\alpha} + W_2(\mu,\nu)^{1+\alpha}\right).$$

(2) Λ_0 and Λ_1 are α-Hölder continuous, i.e.,

$$|\Lambda_0(q_1,\mu) - \Lambda_0(\bar{q}_1,\nu)|_\infty \le C\left(|q_1-\bar{q}_1|^\alpha + W_2^\alpha(\mu,\nu)\right)$$

and

$$|\Lambda_1(q_1,q_2,\mu) - \Lambda_1(\bar{q}_1,\bar{q}_2,\nu)|_\infty \le C(|q_1-\bar{q}_1|^\alpha + |q_2-\bar{q}_2|^\alpha + W_2^\alpha(\mu,\nu)),$$

for any $\mu,\nu \in \mathscr{B}$, $(q_1,\bar{q}_1),(q_2,\bar{q}_2) \in \mathrm{spt}(\mu)\times\mathrm{spt}(\nu)$, and $\gamma \in \Gamma_o(\mu,\nu)$.

We say that $\mathscr{U} \in C^{2,\alpha,w}_{\mathrm{loc}}(\mathscr{P}_2(\mathbb{M}))$ if $\mathscr{U} \in C^{2,\alpha,w}(\mathscr{B}_r)$ for all $r>0$.

Remark 3.14. Let Λ_0 and Λ_1 be as above.

(1) By abuse of notation we write

$$D_{q_1}\big(\nabla_w \mathscr{U}(\mu)(q_1)\big) := \Lambda_0(q_1,\mu) \text{ and } \bar{\nabla}^2_{ww}\mathscr{U}(\mu)(q_1,q_2) := \Lambda_1(q_1,q_2,\mu),$$

for all $\mu \in \mathscr{P}_2(\mathbb{M})$ and $x,y \in \mathrm{spt}(\mu)$. The bar is to recall that Λ_1 is not exactly the second Wasserstein gradient as introduced in [21].

(2) Note that if we choose any matrix $\Lambda(a,\mu)$ such that any of its rows w is such that $\nabla\cdot(w\mu) = 0$ and $w \in L^2(\mu)$, then the matrix defined as $\bar{\Lambda}_1(q,a,\mu) := \Lambda_1(q,a,\mu) + \Lambda(a,\mu)$ also satisfies Definition 3.13(1). We could determine $\Lambda_1(q,\cdot,\mu)$ uniquely by imposing that the i^{th} row of $(\Lambda_0(q,\mu),\Lambda_1(q,\cdot,\mu))$ is the unique element of minimal norm of the subdifferential of $(q,\mu)\mapsto \nabla_w\mathscr{U}(\mu)(q)$. The i^{th} row of the element of minimal norm belongs to $\mathbb{M}\times T_\mu\mathscr{P}_2(\mathbb{M})$, and the new matrix will be denoted as $\nabla^2_{ww}\mathscr{U}(\mu)$. This new matrix is selected at the expense of giving up the property that Λ_1 is uniformly bounded. Increasing C if necessary, we can instead ensure

$$\left\|\nabla^2_{ww}\mathscr{U}(\mu)(q_1,\cdot)\right\|_{L^2_\mu} \le C(r) \quad \forall \mu\in\mathscr{B},\ \forall q_1 \in \mathrm{spt}(\mu).$$

(3) In the spirit of the terminology used in [21], we refer to $\bar{\nabla}^2_{ww}\mathscr{U}$ as an "extended Wasserstein Hessian" of $\mathscr{U}$. In contrast with the assumptions in [21], in Definition 3.13(1), we assume slightly different conditions: the expansion here is required only on $\mathrm{spt}(\mu)\times\mathrm{spt}(\nu)$, Λ_0 and Λ_1 are supposed to be essentially bounded only on $\mathrm{spt}(\mu)$, and in addition we require the Hölder/Lipschitz property in Definition 3.13(2) to be fulfilled.

(4) We shall now compare our definition of $C^{2,\alpha,w}_{\rm loc}(\mathscr{P}_2(\mathbb{M}))$ regularity of $\mathscr{U}$ to $C^{2,\alpha}_{\rm loc}(\mathbb{H})$ regularity of $\widetilde{\mathscr{U}}$ (where $\widetilde{\mathscr{U}}(x) = \mathscr{U}(\sharp(x))$). If $\widetilde{\mathscr{U}} \in C^{2,\alpha}_{\rm loc}(\mathbb{H})$, then $\widetilde{\mathscr{U}}$ is twice continuously differentiable in the Fréchet sense and for each $r > 0$ there exists $C = C(r)$ such that

$$\|\nabla\widetilde{\mathscr{U}}(y) - \nabla\widetilde{\mathscr{U}}(x) - \nabla^2\widetilde{\mathscr{U}}(x)(y-x,\cdot)\| \le C\|x-y\|^{1+\alpha} \quad \forall x, y \in \mathbb{B}_r. \tag{3.4}$$

To heuristically compare this inequality to the setting of $\mathscr{P}_2(\mathbb{M})$ we proceed as follows. Let $\sharp(x) = \mu$ and $\sharp(y) = \nu$ with $\|x-y\| = W_2(\mu,\nu)$. Then we know (see [32]) that

$$\nabla\widetilde{\mathscr{U}}(x) = \nabla_w\mathscr{U}(\mu)\circ x, \quad \nabla\widetilde{\mathscr{U}}(y) = \nabla_w\mathscr{U}(\nu)\circ y,$$

and

$$\nabla^2\widetilde{U}(x)(h,h_*) = \int_\Omega D_q\big(\nabla_w\mathscr{U}(\mu)\big)\circ x\ h\cdot h_* d\omega + \int_{\Omega^2} \nabla^2_{ww}\mathscr{U}(\mu)\big(x(\omega), x(\omega_*)\big)h(\omega)\cdot h_*(\omega_*)d\omega\, d\omega_*,$$

if $\xi, \xi_* \in T_\mu\mathscr{P}_2(\mathbb{M})$ and $h = \xi\circ x$ and $h_* = \xi_*\circ x$. Thus, (3.4) would read as

$$\begin{aligned}\sup_{\|h_*\|\le 1}\Bigg|&\int_\Omega [\nabla_w\mathscr{U}(\nu)(y(\omega))\cdot h_*(\omega) - \nabla_w\mathscr{U}(\mu)(x(\omega))\cdot h_*(\omega)]\,d\omega\\ &-\int_\Omega D_q\big(\nabla_w\mathscr{U}(\mu)\big)\circ x\ (y-x)\cdot h_*\,d\omega\\ &-\int_{\Omega^2}\nabla^2_{ww}\mathscr{U}(\mu)\big(x(\omega),x(\omega_*)\big)(y-x)(\omega)\cdot h_*(\omega_*)d\omega\,d\omega_*\Bigg|\\ \le C\,W_2(\mu,\nu)^{1+\alpha}.&\end{aligned} \tag{3.5}$$

From here we see that a necessary condition to obtain inequality (1) in Definition 3.13 is to have (3.5) hold when we maximize over the set of h such that $\|h_*\|_{L^1} \le 1$ rather than maximizing over the set of h such that $\|h_*\| \le 1$. In other words, we have not been able to show that if $\widetilde{\mathscr{U}} \in C^{2,\alpha}_{\rm loc}(\mathbb{H})$ then $\mathscr{U} \in C^{2,\alpha,w}_{\rm loc}(\mathscr{P}_2(\mathbb{M}))$. Moreover, in Appendix A we show that imposing $\mathscr{U} \in C^{2,\alpha,w}_{\rm loc}(\mathscr{P}_2(\mathbb{M}))$ in general does not imply that $\widetilde{\mathscr{U}} \in C^{2,\alpha}_{\rm loc}(\mathbb{H})$.

(5) Let us point out that using an extrinsic approach, [11] introduced spaces of the type $C^{2,1}(\mathscr{P}_2(\mathbb{M}))$ via the differentials of their lifts on a Hilbert space. In this work, we define $C^{2,1,w}(\mathscr{P}_2(\mathbb{M}))$ in an intrinsic way, i.e., directly via the differential calculus on the Wasserstein space. As a result, our derivatives are always defined on the supports of the corresponding

measures, while in [11] the authors work with global extensions. Similarly, we require essential boundedness of the Wasserstein Hessian only on the support of the corresponding measures, while [11] requires boundedness of the global extensions. The work [32] allows us to assert that both the intrinsic and extrinsic approaches are essentially the same. However, $C^{2,1,w}(\mathscr{P}_2(\mathbb{M}))$ has the advantage that it can be seen as an increasing 'limit' of the spaces $C^{2,1}(\mathbb{M}^m)$ when $m \to +\infty$, as we show in Section 3.4 below.

(6) [11, sec. 2] constructs an example of $\mathscr{U} \in C^{2,1}(\mathscr{P}_2(\mathbb{M}))$ for which its lifted version $\widetilde{\mathscr{U}}$ fails to be twice Fréchet differentiable at any point. More discussions can be found in [11, 14, 16, 17, 19].

3.4 Regularity of $\mathscr{U}$ as a by-product of regularity estimates on $U^{(m)}$

This subsection implies regularity properties on functions $\mathscr{U}$ defined on $\mathscr{P}_2(\mathbb{M})$ from estimates on their restrictions $U^{(m)}$. Recall that for $r > 0$ $\mathbb{B}_r^m$ is a ball in $\mathbb{M}^m$ while $\mathscr{B}_r$ is a ball in $\mathscr{P}_2(\mathbb{M})$. We assume that we have at hand a constant $C = C(r) > 0$.

LEMMA 3.15. *Suppose for each $m \in \mathbb{N}$ fixed, $U^{(m)} : \mathbb{M}^m \to \mathbb{R}$ is permutation invariant with respect to its m-variables and $|U^{(m)}|$ is bounded on $\mathbb{B}_r^m$ by a constant that depends on $r > 0$ but is independent of m. Then there exists $C = C(r) > 0$ such that the followings hold true*:

(i) *If $U^{(m)}$ satisfies Property* 2.2 *(1)-(b), then for any $q, b \in \mathbb{B}_r^m$, we have*

$$|U^{(m)}(q) - U^{(m)}(b)| \le C W_2\big(\mu_q^{(m)}, \mu_b^{(m)}\big).$$

(ii) *If $U^{(m)}$ satisfies Property* 2.2 *(2), then for any $q, b \in \mathbb{B}_r^m$, we have*

$$\left| U^{(m)}(b) - U^{(m)}(q) - \sum_{i=1}^{m} D_{q_i} U^{(m)}(q) \cdot (b_i - q_i) \right| \le C W_2^2(\mu_q^{(m)}, \mu_b^{(m)}).$$

(iii) *The assumption in* (ii) *implies for any $q, b \in \mathbb{B}_r^m$*

(a)

$$m|D_{q_i} U^{(m)}(q) - D_{q_i} U^{(m)}(b)| \le C\big(|q_i - b_i| + W_2\big(\mu_q^{(m)}, \mu_b^{(m)}\big)\big).$$

(b) *We have*

$$m|D_{q_i} U^{(m)}(q) - D_{q_j} U^{(m)}(b)| \le C\left(|q_i - b_j| + W_2(\mu_q^{(m)}, \mu_b^{(m)}) + \frac{1}{\sqrt{m}}\right),\ i \ne j.$$

(iv) *Suppose that $U^{(m)}$ satisfies Property* 2.2(3). *If $i \in \{1,\dots,m\}$ and $q,b \in \mathbb{B}_r^m$, then*

$$m\left|D_{q_i}U^{(m)}(b) - D_{q_i}U^{(m)}(q) - \sum_{j=1}^m D^2_{q_iq_j}U^{(m)}(q)(b_j-q_j)\right| \le C\big(|q_i-b_i|^2 + W_2^2(\mu_q^{(m)},\mu_b^{(m)})\big).$$

(v) *The assumption in* (iv) *implies $q,b \in \mathbb{B}_r^m$,*

(a) *If $i \neq j$ then*

$$m^2\big|D^2_{q_iq_j}U^{(m)}(q) - D^2_{q_iq_j}U^{(m)}(b)\big| \le C\big(|q_i-b_i| + |q_j-b_j| + W_2\big(\mu_q^{(m)},\mu_b^{(m)}\big)\big).$$

(b) *If $(i,j)\neq(k,l), i\neq j, k\neq l$, then*

$$m^2\big|D^2_{q_iq_j}U^{(m)}(q) - D^2_{q_kq_l}U^{(m)}(b)\big| \le C\left(|q_i-b_k| + |q_j-b_l| + W_2\big(\mu_q^{(m)},\mu_b^{(m)}\big) + \frac{1}{\sqrt{m}}\right).$$

(c) *We have*

$$m\big|D^2_{q_iq_i}U^{(m)}(q) - D^2_{q_iq_i}U^{(m)}(b)\big| \le C\big(|q_i-b_i| + W_2\big(\mu_q^{(m)},\mu_b^{(m)}\big)\big).$$

(d) *We have*

$$m\big|D^2_{q_iq_i}U^{(m)}(q) - D^2_{q_jq_j}U^{(m)}(b)\big| \le C\left(|q_i-b_j| + W_2(\mu_q^{(m)},\mu_b^{(m)}) + \frac{1}{\sqrt{m}}\right).$$

PROOF. Since $U^{(m)}$ is permutation invariant, reordering q and b if necessary, we may assume

$$\gamma^{(m)} := \frac{1}{m}\sum_{i=1}^m \delta_{(q_i,b_i)} \in \Gamma_o(\mu_q^{(m)},\mu_b^{(m)}).$$

Below, using Taylor's expansion, we may find $\xi \in \mathbb{B}_r^m$ on the line segment connecting q to b such that (using the shorthand notation $\|\cdot\|_\infty$ to denote $\|\cdot\|_{L^\infty(\mathbb{B}_r^m)}$)

(i) we have

$$|U^{(m)}(b) - U^{(m)}(q)| \le \left|\sum_{i=1}^m D_{q_i}U^{(m)}(\xi)\cdot(b_i-q_i)\right| \le \left(\sum_{i=1}^m m|D_{q_i}U^{(m)}|^2\right)^{\frac12}\left(\sum_{i=1}^m \frac{1}{m}|q_i-b_i|^2\right)^{\frac12}.$$

Using the fact that

$$\sum_{i=1}^{m} m\left|D_{q_i}U^{(m)}(q)\right|^2 \le C^2 \quad \text{and} \quad \sum_{i=1}^{m}\frac{1}{m}|q_i-b_i|^2 = W_2^2\big(\mu_q^{(m)},\mu_b^{(m)}\big),$$

we verify the statement in (i).

(ii) A second-order Taylor expansion yields

$$\begin{aligned}
&U^{(m)}(b)-U^{(m)}(q)-\sum_{i=1}^{m}D_{q_i}U^{(m)}(q)\cdot(b_i-q_i)\\
&\quad=\frac{1}{2}\sum_{i,j=1}^{m}\big\langle(b_i-q_i),D^2_{q_iq_j}U^{(m)}(\xi)(b_j-q_j)\big\rangle\\
&\quad=\frac{1}{2}\sum_{i=1}^{m}\big\langle(b_i-q_i),D^2_{q_iq_i}U^{(m)}(\xi)(b_i-q_i)\big\rangle\\
&\qquad+\frac{1}{2}\sum_{i\neq j}\big\langle(b_i-q_i),D^2_{q_iq_j}U^{(m)}(\xi)(b_j-q_j)\big\rangle.
\end{aligned}$$

Thus, under the assumption in (ii), we have

$$\begin{aligned}
&\left|U^{(m)}(b)-U^{(m)}(q)-\sum_{i=1}^{m}D_{q_i}U^{(m)}(q)\cdot(b_i-q_i)\right|\\
&\quad\le\frac{C}{2m}\sum_{i=1}^{m}|q_i-b_i|^2+\frac{1}{4}\sum_{i\neq j}\big\|D^2_{q_iq_j}U^{(m)}\big\|_\infty|q_i-b_i|^2\\
&\qquad+\frac{1}{4}\sum_{i\neq j}\|D^2_{q_iq_j}U^{(m)}\|_\infty|q_j-b_j|^2\\
&\quad\le\left(\frac{C}{2}+\frac{C}{4}+\frac{C}{4}\right)\int_{\mathbb{M}^2}|z-w|^2d\gamma^{(m)}(z,w)=C\,W_2^2\big(\mu_q^{(m)},\mu_b^{(m)}\big).
\end{aligned}$$

(iii)-(a) Performing again a first-order Taylor expansion, we find

$$\begin{aligned}
&D_{q_i}U^{(m)}(q)-D_{q_i}U^{(m)}(b)\\
&\quad=\sum_{k=1}^{m}D^2_{q_kq_i}U^{(m)}(q)(q_k-b_k)\\
&\quad=D^2_{q_iq_i}U^{(m)}(\xi)(q_i-b_i)+\sum_{k\neq i}D^2_{q_kq_i}U^{(m)}(\xi)(q_k-b_k).
\end{aligned}$$

Thus using the assumptions, we find

$$\begin{aligned}
&\left|D_{q_i}U^{(m)}(q) - D_{q_i}U^{(m)}(b)\right| \\
&\le \frac{C}{m}|q_i - b_i| + \Big(\sum_{k\neq i} m^3 \|D^2_{q_k q_i}U^{(m)}\|^2_\infty\Big)^{\frac12}\Big(\sum_{k\neq i}\frac{1}{m^3}|q_k - b_k|^2\Big)^{\frac12} \\
&\le \frac{C}{m}\big(|q_i - b_i| + W_2\big(\mu_q^{(m)}, \mu_b^{(m)}\big)\big).
\end{aligned}$$

(iii)-(b) Without loss of generality, let us suppose that $i < j$. By the permutation invariance of $U^{(m)}$, we observe that $D_{q_i}U^{(m)}(q) = D_{q_1}U^{(m)}(q^{ij})$ and a similar identity holds for $D_{q_j}U^{(m)}(b)$ if we set

$$q^{ij} := (q_i, q_j, q_1, \dots, q_{i-1}, q_{i+1}, \dots, q_{j-1}, q_{j+1}, \dots, q_m). \tag{3.6}$$

Using a similar identity for $D_{q_j}U^{(m)}(b)$ we obtain

$$\begin{aligned}
\left|D_{q_i}U^{(m)}(q) - D_{q_j}U^{(m)}(b)\right| &= |D_{q_1}U^{(m)}(q^{ij}) - D_{q_1}U^{(m)}(b^{ij})| \\
&\le \|D^2_{q_1 q_1}U^{(m)}\|_\infty |q_i - b_j| \\
&\quad + \sum_{k=1}^{i-1} \|D^2_{q_{k+2}\, q_1}U^{(m)}\|_\infty |q_k - b_k| \\
&\quad + \sum_{k=i+1}^{j-1} \|D^2_{q_{k+1}\, q_1}U^{(m)}\|_\infty |q_k - b_k| \\
&\quad + \sum_{k=j+1}^{m} \|D^2_{q_k q_1}U^{(m)}\|_\infty |q_k - b_k|.
\end{aligned}$$

Thus,

$$\begin{aligned}
&\left|D_{q_i}U^{(m)}(q) - D_{q_j}U^{(m)}(b)\right| \\
&\le \frac{C}{m}|q_i - b_j| + \frac{C}{m^2}(|q_j| + |b_i|) + \frac{C}{m^2}\sum_{k=1}^{m}|q_k - b_k| \\
&\le \frac{C}{m}\left(|q_i - b_j| + W_2(\mu_q^{(m)}, \mu_b^{(m)}) + \frac{2r\sqrt{m}}{m}\right) \\
&\le \frac{C}{m}\left(|q_i - b_j| + W_2\big(\mu_q^{(m)}, \mu_b^{(m)}\big) + \frac{1}{\sqrt{m}}\right),
\end{aligned}$$

where we have used the assumptions on $D^2_{q_i q_j}U^{(m)}$, and in the last two rows we used the facts that since $q, b \in \mathbb{B}_r^m$, we have that $|q_i|, |b_j| \le r\sqrt{m}$ for all $i, j \in \{1, \dots, m\}$.

(iv) Similarly to the previous points, we perform a Taylor expansion (of order 2) to obtain

$$D_{q_i}U^{(m)}(b) - D_{q_i}U^{(m)}(x) - \sum_{j=1}^{m} D^2_{q_iq_j}U^{(m)}(q)(b_j - q_j)$$
$$= \frac{1}{2}\sum_{j,k=1}^{m} \langle (b_k - q_k), D^3_{q_iq_jq_k}U^{(m)}(q)(b_j - q_j)\rangle,$$

and thus

$$\left| D_{q_i}U^{(m)}(b) - D_{q_i}U^{(m)}(q) - \sum_{j=1}^{m} D^2_{q_iq_j}U^{(m)}(q)(b_j - q_j)\right|$$
$$\le \frac{1}{2}\| D^3_{q_iq_iq_i}U^{(m)}\|_\infty |q_i - b_i|^2 + \frac{1}{2}\sum_{j\neq i}\| D^3_{q_iq_jq_j}U^{(m)}\|_\infty |q_j - b_j|^2$$
$$+ \frac{1}{2}\sum_{j\neq k\neq i}\| D^3_{q_iq_jq_j}U^{(m)}\|_\infty |q_j - b_j|\cdot|q_k - b_k|.$$

We conclude

$$\left| D_{q_i}U^{(m)}(b) - D_{q_i}U^{(m)}(q) - \sum_{j=1}^{m} D^2_{q_iq_j}U^{(m)}(q)(b_j - q_j)\right|$$
$$\le \frac{C}{2m}|q_i - b_i|^2 + \frac{C}{2m}\sum_{j=1}^{m}\frac{1}{m}|q_j - b_j|^2$$
$$+ \frac{C}{2m}\left(\sum_{j=1}^{m}\frac{1}{m}|q_j - b_j|\right)\left(\sum_{k=1}^{m}\frac{1}{m}|q_k - b_k|\right)$$
$$\le \frac{C}{2m}\big(|q_i - b_i|^2 + W_2^2\big(\mu_q^{(m)}, \mu_q^{(m)}\big)\big),$$

(v) We write again

$$D^2_{q_iq_j}U^{(m)}(q) - D^2_{q_iq_j}U^{(m)}(b)$$
$$= \sum_{k=1}^{m} D^3_{q_iq_jq_k}U^{(m)}(q)(q_k - b_k)$$
$$= D^3_{q_iq_jq_i}U^{(m)}(q)(q_i - q_i) + D^3_{q_iq_jq_j}U^{(m)}(q)(q_j - b_j)$$
$$+ \sum_{k=1,k\neq i,k\neq j}^{m} D^3_{q_iq_jq_k}U^{(m)}(q)(q_k - q_k).$$

Thus in the case of (a) using the assumptions, we find

$$\begin{aligned}
&\left|D^2_{q_iq_j}U^{(m)}(q) - D^2_{q_iq_j}U^{(m)}(b)\right| \\
&\quad \le \frac{C}{m^2}(|q_i - b_i| + |q_j - b_j|) + C\sum_{k=1}^{m}\frac{1}{m^3}|q_k - b_k| \\
&\quad \le \frac{C}{m^2}\big(|q_i - b_i| + |q_j - b_j| + W_2\big(\mu_q^{(m)}, \mu_b^{(m)}\big)\big).
\end{aligned}$$

In the case of (c), since $i = j$ in the above expansion, we find

$$\begin{aligned}
\left|D^2_{q_iq_i}U^{(m)}(q) - D^2_{q_iq_i}U^{(m)}(b)\right| &\le \left\|D^3_{q_iq_iq_i}U^{(m)}\right\|_\infty |q_i - b_i| \\
&\quad + \sum_{k\neq i}\|D^3_{q_iq_iq_k}U^{(m)}\|_\infty |q_k - b_k| \\
&\le \frac{C}{m}\big(|q_i - b_i| + W_2\big(\mu_q^{(m)}, \mu_b^{(m)}\big)\big).
\end{aligned}$$

To show (b), let us suppose without loss of generality that $i < j < k < l$. By the permutation invariance of $U^{(m)}$ we have the identities

$$D^2_{q_iq_j}U^{(m)}(q) = D^2_{q_1q_2}U^{(m)}(q_i, q_j, q_k, q_l, \overline{q})$$

and

$$D^2_{q_kq_l}U^{(m)}(b) = D^2_{q_1q_2}U^{(m)}(b_k, b_l, b_i, b_j, \overline{b}),$$

where $\overline{q}, \overline{b} \in \mathbb{R}^{d\times(m-4)}$ obtained from q and b, respectively, by deleting the vectors indexed by i, j, k, l. Therefore, using the local bounds on the third-order derivatives of $U^{(m)}$, we have

$$\begin{aligned}
&\left|D^2_{q_iq_j}U^{(m)}(q) - D^2_{q_kq_l}U^{(m)}(b)\right| \\
&\quad = \left|D^2_{q_1q_2}U^{(m)}(q_i, q_j, q_k, q_l, \overline{q}) - D^2_{q_1q_2}U^{(m)}(q_k, q_l, q_i, q_j, \overline{b})\right|,
\end{aligned}$$

and so

$$
\begin{aligned}
&\big|D^2_{q_i q_j} U^{(m)}(q) - D^2_{q_k q_l} U^{(m)}(b)\big| \\
&\quad \leq \big\|D^3_{q_1 q_2 q_1} U^{(m)}\big\|_\infty |q_i - b_k| \\
&\qquad + \big\|D^3_{q_1 q_2 q_2} U^{(m)}\big\|_\infty |q_j - b_l| + \big\|D^3_{q_1 q_2 q_3} U^{(m)}\big\|_\infty |q_k - b_i| \\
&\qquad + \big\|D^3_{q_1 q_2 q_4} U^{(m)}\big\|_\infty |q_l - b_j| + \sum_{\alpha=1}^{i-1} \big\|D^3_{q_1 q_2 q_{\alpha+4}} U^{(m)}\big\|_\infty |q_\alpha - b_\alpha| \\
&\qquad + \sum_{\alpha=i+1}^{j-1} \big\|D^3_{q_1 q_2 q_{\alpha+3}} U^{(m)}\big\|_\infty |q_\alpha - b_\alpha| \\
&\qquad + \sum_{\alpha=j+1}^{k-1} \big\|D^3_{q_1 q_2 q_{\alpha+2}} U^{(m)}\big\|_\infty |q_\alpha - b_\alpha| \\
&\qquad + \sum_{\alpha=k+1}^{l-1} \big\|D^3_{q_1 q_2 q_{\alpha+1}} U^{(m)}\big\|_\infty |q_\alpha - b_\alpha| \\
&\qquad + \sum_{\alpha=l+1}^{m} \big\|D^3_{q_1 q_2 q_\alpha} U^{(m)}\big\|_\infty |q_\alpha - b_\alpha|.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
&|D^2_{q_i q_j} U^{(m)}(q) - D^2_{q_k q_l} U^{(m)}(b)| \\
&\quad \leq \frac{C}{m^2}\big(|q_i - b_k| + |q_j - b_l|\big) + \frac{C}{m^3}(|q_k| + |b_i| + |q_l| + |b_j|) \\
&\qquad + \frac{C}{m^3} \sum_{\alpha=1}^{m} |q_\alpha - b_\alpha| \\
&\quad \leq \frac{C}{m^2}\left(|q_i - b_k| + |q_j - b_l| + W_2(\mu_q^{(m)}, \mu_b^{(m)}) + \frac{1}{\sqrt{m}}\right),
\end{aligned}
$$

where we have used again that since $q, b \in \mathbb{B}_r^m$, we have $|q_\alpha|, |b_\alpha| \leq C\sqrt{m}$ for all $\alpha \in \{1, \dots, m\}$.

In the case of (d), we proceed similarly as for (b). Let us suppose without loss of generality that $i < j$. Then, by the permutation invariance of $U^{(m)}$, we use the expression in (3.6) to obtain

$$
D^2_{q_i q_i} U^{(m)}(q) = D^2_{q_1 q_1} U^{(m)}(q^{ij}),
$$

Using the analogous identity with $D^2_{q_jq_j}U^{(m)}(b)$ we conclude

$$\begin{aligned}
&\big|D^2_{q_iq_i}U^{(m)}(q) - D^2_{q_jq_j}U^{(m)}(b)\big| \\
&\quad = \big|D^2_{q_1q_1}U^{(m)}(q^{ij}) - D^2_{q_1q_1}U^{(m)}(b^{ij})\big| \\
&\quad \le \big\|D^3_{q_1q_1q_1}U^{(m)}\big\|_\infty |q_i - b_j| + \big\|D^3_{q_1q_1q_2}U^{(m)}\big\|_\infty |q_j - b_i| \\
&\qquad + \sum_{k=1}^{i-1}\big\|D^3_{q_1q_1q_{k+2}}\big\|_\infty |q_k - b_k| + \sum_{k=i+1}^{j-1}\big\|D^3_{q_1q_1q_{k+1}}\big\|_\infty |q_k - b_k| \\
&\qquad + \sum_{k=j+1}^{m}\big\|D^3_{q_1q_1q_k}\big\|_\infty |q_k - b_k|.
\end{aligned}$$

Thus,

$$\begin{aligned}
&\big|D^2_{q_iq_i}U^{(m)}(q) - D^2_{q_jq_j}U^{(m)}(b)\big| \\
&\quad \le \frac{C}{m}|q_i - b_j| + \frac{C}{m^2}(|q_j| + |b_i|) + \frac{C}{m^2}\sum_{k=1}^{m}|q_k - b_k| \\
&\quad \le \frac{C}{m}\left(|q_i - b_j| + W_2(\mu_q^{(m)}, \mu_b^{(m)}) + \frac{1}{\sqrt{m}}\right),
\end{aligned}$$

where we have used again that since $q, b \in \mathbb{B}_r^m$, we have $|q_\alpha|, |b_\alpha| \le C\sqrt{m}$ for all $\alpha \in \{1, \dots, m\}$. □

The following two theorems show how the quantified regularity estimates on the restrictions of functions $u : \mathbb{M} \times \mathscr{P}_2(\mathbb{M}) \to \mathbb{R}$ and $\mathscr{U} : \mathscr{P}_2(\mathbb{M}) \to \mathbb{R}$ to $\mathbb{M} \times \mathbb{M}^m$ and $\mathbb{M}^m$, respectively, will imply the corresponding regularity of the original functions.

THEOREM 3.16. *Let $u : \mathbb{M} \times \mathscr{P}_2(\mathbb{M}) \to \mathbb{R}$ be a continuous function. For $m \in \mathbb{N}$, we define $u^{(m)} : \mathbb{M} \times (\mathbb{M})^m \to \mathbb{R}$ as*

$$u^{(m)}(q_0, q) := u\big(q_0, \mu_q^{(m+1)}\big),$$

where $(q_0, q) = (q_0, q_1, \dots, q_m) \in (\mathbb{M})^{m+1}$ and $\mu_q^{(m+1)} = \frac{1}{m+1}\sum_{i=0}^m \delta_{q_i}$. Suppose that $u^{(m)} \in C^{1,1}_{\rm loc}(\mathbb{M} \times (\mathbb{M})^m)$ and that for $K \subset \mathbb{M}$ compact and $r > 0$, $u^{(m)}(q_0, \cdot)$ satisfies the estimates of Property 2.2(1)-(a) and (2) for all $q_0 \in K$, with a constant $C = C(K, r) > 0$. Let us moreover assume that for any $K \subset \mathbb{M}$ compact and $r > 0$, there exists $C = C(K, r) > 0$ such that

$$\begin{aligned}
&|D_{q_0}u^{(m)}(q_0, q)| \le C, \quad |D^2_{q_0q_0}u^{(m)}(q_0, q)|_\infty \le C, \\
&\sum_{i=1}^m m|D^2_{q_iq_0}u^{(m)}(q_0, q)|^2_\infty \le C,
\end{aligned} \tag{3.7}$$

and

$$\left|D^2_{q_i q_j} u^{(m)}(q_0, q)\right|_\infty \le \begin{cases} \dfrac{C}{m}, & i = j \text{ and } i > 0, \\ \dfrac{C}{m^2}, & i \neq j,\ i, j > 0, \end{cases}$$

for any $q_0 \in K$ *and* $q = (q_1, \dots, q_m) \in \mathbb{B}^m_r$.

Then, there exists $\Phi_1 : \mathbb{M} \times \mathscr{P}_2(\mathbb{M}) \times \mathbb{M} \to \mathbb{R}^d$*, a locally Lipschitz-continuous function such that for any* $r > 0$ *and* $K \subset \mathbb{M}$ *compact, there exists*

$$C = C(K, r) > 0$$

such that for any $q_0, y_0 \in K$*, any* $\mu, \nu \in \mathscr{P}_2(\mathbb{M})$*, and* $\gamma \in \Gamma_o(\mu, \nu)$*,* u *satisfies*

$$\begin{aligned} &\bigg| u(y_0, \nu) - u(q_0, \mu) - D_{q_0} u(q_0, \mu) \cdot (y_0 - q_0) \\ &- \int_{\mathbb{M}^2} \Phi_1(q_0, \mu, q) \cdot (y - q) d\gamma(q, y) \bigg| \\ &\le C \left(|q_0 - y_0|^2 + W_2^2(\mu, \nu) \right). \end{aligned}$$

This implies in particular that $u \in C^{1,1}_{\text{loc}}(\mathbb{M} \times \mathscr{P}_2(\mathbb{M}))$*,* $\nabla_w u(q_0, \mu)(\cdot)$ *can be obtained as the projection of* $\Phi_1(q_0, \mu, \cdot)$ *onto* $T_\mu \mathscr{P}_2(\mathbb{M})$ *and*

$$\begin{aligned} &\bigg| u(y_0, \nu) - u(q_0, \mu) - D_{q_0} u(q_0, \mu) \cdot (y_0 - q_0) \\ &- \int_{\mathbb{M}^2} \nabla_w u(q_0, \mu)(q) \cdot (y - q) d\gamma(q, y) \bigg| \\ &\le C \left(|q_0 - y_0|^2 + W_2^2(\mu, \nu) \right). \end{aligned}$$

PROOF. Our construction is inspired by [31, lemma 8.10].

For $m \in \mathbb{N}$ we define $\Phi_0^{(m)} : \mathbb{M} \times \mathscr{P}_2^{(m)}(\mathbb{M}) \to \mathbb{R}^d$ and $\Phi_1^{(m)} : \mathbb{M} \times \bigcup_{\mu \in \mathscr{P}_2^{(m)}(\mathbb{M})} \operatorname{spt}(\mu) \times \{\mu\} \to \mathbb{R}^d$ as

$$\Phi_0^{(m)}(q_0, \mu_q^{(m)}) := D_{q_0} u^{(m)}(q_0, q)$$

and

$$\Phi_1^{(m)}(q_0, q_i, \mu_q^{(m)}) := m D_{q_i} u^{(m)}(q_0, q) \quad \forall i \in \{1, \dots, m\}.$$

Here

$$q = (q_1, \dots, q_m) \quad \text{and} \quad \mu_q^{(m)} := \frac{1}{m} \sum_{i=1}^m \delta_{q_i} \in \mathscr{P}_2^{(m)}(\mathbb{M}).$$

From the assumptions of this theorem, as a consequence of Lemma 3.15(i), when restricted to $K \times \mathscr{P}_2^{(m)}(\mathbb{M}) \cap \mathscr{B}_r$ where $K \subseteq \mathbb{M}$ is compact and $r > 0$, $\Phi_0^{(m)}$ is uniformly bounded and uniformly Lipschitz-continuous, with respect to m (and the Lipschitz constant depends solely on K and r).

Let $\mathscr{K}$ be the collection of compact sets in $\mathbb{M}$. We assume there exists a positive function C defined on $\mathscr{K} \times (0,\infty)$ such that $C(K,r) \le C(K',r')$ $K \subset K'$ and $r \le r'$.

We assume to be given a family of functions

$$f^{(m)} : \mathbb{M} \times \mathscr{P}_2^{(m)}(\mathbb{M}) \to \mathbb{R}$$

such that for each $r > 0$ and each $K \in \mathscr{K}$, the restriction of $f^{(m)}$ to $K \times (\mathscr{P}_2^{(m)}(\mathbb{M}) \cap \mathscr{B}_r)$ is $C(K,r)$-Lipschitz. We assume there exists a compact subset in the real line which contains all the $f^{(m)}(0,\delta_0)$.

In what follows, we will perform Lipschitz extensions of various functions using the Kirszbraun extension formula. For $r > 0$, $q_0 \in \mathbb{M}$, and $K \in \mathscr{K}$, we define the Kirszbraun-Valentine extension $f_{K,r}^{(m)}(q_0,\cdot) : \mathscr{P}_2(\mathbb{M}) \to \mathbb{R}$ as

$$\begin{aligned} &f_{K,r}^{(m)}(q_0,\mu) \\ &\quad = \inf_{\nu}\bigl\{f^{(m)}(q_0,\nu) + C(K,r)W_2(\mu,\nu) : \nu \in \mathscr{P}_2^{(m)}(\mathbb{M}) \cap \mathscr{B}_r\bigr\}. \end{aligned} \tag{3.8}$$

We have that $f_{K,r}^{(m)}(q_0,\cdot)$ is $C(K,r)$-Lipschitz for all $q_0 \in \mathbb{M}$ and $f_{K,r}^{(m)}$ coincides with $f^{(m)}$ on $K \times (\mathscr{P}_2^{(m)}(\mathbb{M}) \cap \mathscr{B}_r)$. Furthermore, for any $K' \in \mathscr{K}$, $f_{K,r}^{(m)}(\cdot,\mu)$ is $C(K',r)$-Lipschitz on $K' \times \mathscr{P}_2(\mathbb{M})$.

Let $\bar{B}_R(0)$ denote the closed ball of radius $R > 0$, centered at the origin in $\mathbb{M}$, and let $\mathscr{P}_{\mathrm{c}}(\mathbb{M})$ be the union of all the $\mathscr{P}_2(\bar{B}_R(0))$. Since $\mathscr{P}_2(\bar{B}_R(0))$ is a compact subset of $\mathscr{P}_2(\mathbb{M})$, we apply the Ascoli-Arzelà theorem and use a diagonalization argument to obtain a function

$$f_{K,r}^{\infty} : \mathbb{M} \times \mathscr{P}_{\mathrm{c}}(\mathbb{M}) \to \mathbb{R}$$

such that a subsequence of $(f_{K,r}^{(m)})_m$ converges locally uniformly to $f_{K,r}^{\infty}$ on compact sets. We have that $f_{K,r}^{\infty}(q_0,\cdot)$ is $C(K,r)$-Lipschitz on $\mathscr{P}_{\mathrm{c}}(\mathbb{M})$ for all $q_0 \in \mathbb{M}$, and $f_{K,r}^{\infty}(\cdot,\mu)$ is $C(K',r)$-Lipschitz on K' for $\mu \in \mathscr{P}_{\mathrm{c}}(\mathbb{M})$. In fact,

$$\bigl|f_{K,r}^{\infty}(q_0,\mu) - f_{K,r}^{\infty}(a_0,\nu)\bigr| \le C(K',r)\bigl(|q_0 - a_0| + W_2(\mu,\nu)\bigr) \tag{3.9}$$

for all $q_0, a_0 \in K'$ and $\mu,\nu \in \mathscr{B}_r$.

The function $f_{K,r}^{\infty}$ admits a unique $C(K,r)$-Lipschitz extension to $K \times \mathscr{B}_r$, which we continue to denote as $f_{K,r}^{\infty}$. Using the construction (3.8) for each coordinate function of $\Phi_0^{(m)}$, we construct

$$\Phi_{0,K,r}^{\infty} : \mathbb{M} \times \mathscr{P}_2(\mathbb{M}) \to \mathbb{R}^d.$$

Similarly, assume we are given a family of functions $\Phi_1^{(m)}$ defined on

$$\mathbb{M} \times \left\{\left(q_i, \frac{1}{m}\sum_{j=1}^{m}\delta_{q_j}\right) : q \in (\mathbb{M})^m\right\}.$$

As a consequence of the assumptions and Lemma 3.15(iii)-(b), we assume for each $r > 0$ and $K \in \mathscr{K}$,

$$\begin{aligned}&\left|\Phi_1^{(m)}\big(q_0, q_1, \mu_q^{(m)}\big) - \Phi_1^{(m)}\big(\overline{q}_0, \overline{q}_1, \mu_{\overline{q}}^{(m)}\big)\right| \\ &\qquad \le C(K,r)\left(|q_0 - \overline{q}_0| + |q_1 - \overline{q}_1| + W_2(\mu_q^{(m)}, \mu_{\overline{q}}^{(m)}) + \frac{1}{\sqrt{m}}\right)\end{aligned}$$

for all $q_0, \overline{q}_0 \in K$, and all $q, \overline{q} \in \mathbb{B}_r^m$.

For each $k \in \{1, \dots, d\}$, $\Phi_1^{(m),k}$ and $q_0, q_* \in \mathbb{M}$, define

$$\begin{aligned}&\Phi_{1,K,r}^{(m),k}(q_0, q_*, \mu) \\ &\quad := \inf_{\overline{q}} \left\{\Phi_1^{(m),k}(q_0, \overline{q}_i, \mu_{\overline{q}}^{(m)})) + C(K,r)\Big(|q_* - \overline{q}_i| + W_2(\mu, \mu_{\overline{q}}^{(m)})\Big) : \overline{q} \in \mathbb{B}_r^m\right\}.\end{aligned}$$

Note

$$\left|\Phi_{1,K,r}^{(m),k}\big(q_0, q_i, \mu_q^{(m)}\big) - \Phi_1^{(m),k}\big(q_0, q_i, \mu_q^{(m)}\big)\right| \le \frac{C}{\sqrt{m}} \quad \forall (q_0, q) \in K \times \mathbb{B}_r^m. \tag{3.10}$$

As done earlier, there is a function

$$\Phi_{1,K,r}^{\infty,k} : \mathbb{M} \times \mathbb{M} \times \mathscr{P}_2(\mathbb{M}) \to \mathbb{R}$$

and a subsequence (which we may assume to be the same as the ones above) such that $(\Phi_{1,K,r}^{(m),k})_m$ converges locally uniformly to $\Phi_{1,K,r}^{\infty,k}$ on compact sets. Increasing the value of $C(K', r)$ if necessary, we have

$$\left|\Phi_{1,K,r}^{\infty}(q_0, q_1, \mu) - \Phi_{1,K,r}^{\infty}(\overline{q}_0, \overline{q}_1, \nu)\right| \le C(K', r)\big(|q_0 - \overline{q}_0| + |q_1 - \overline{q}_1| + W_2(\mu, \nu)\big) \tag{3.11}$$

if $q_0, q_1, \overline{q}_0, \overline{q}_1 \in K'$ and $\mu, \nu \in \mathscr{B}_r$.

Let $q_0, \overline{q}_0 \in \mathbb{M}$ and let $K \subset \mathbb{M}$ be the closure of a bounded open set containing the line segment $[q_0, \overline{q}_0]$. Let furthermore $q, \overline{q} \in \mathbb{B}_r^m$. By the regularity assumptions on $u^{(m)}$, one can write the following Taylor expansion:

$$\begin{aligned}&u^{(m)}(\overline{q}_0, \overline{q}) - u^{(m)}(q_0, q) - D_{q_0} u^{(m)}(q_0, q) \cdot (\overline{q}_0 - q_0) \\ &\quad - \sum_{i=1}^m D_{q_i} u^{(m)}(q_0, q) \cdot (\overline{q}_i - q_i) = \\ &\quad = \frac{1}{2}(\overline{q}_0 - q_0) \cdot D^2_{q_0 q_0} u^{(m)}(z_0, z)(\overline{q}_0 - q_0) \\ &\qquad + \sum_{i=1}^m (\overline{q}_i - q_i) \cdot D^2_{q_i q_0} u^{(m)}(z_0, z)(\overline{q}_0 - q_0) \\ &\qquad + \frac{1}{2}\sum_{i=1}^m (\overline{q}_i - q_i) D^2_{q_i q_i} u^{(m)}(z_0, z)(\overline{q}_i - q_i) \\ &\qquad + \frac{1}{2}\sum_{i \ne j = 1}^m (\overline{q}_j - q_j) D^2_{q_i q_j} u^{(m)}(z_0, z)(\overline{q}_i - q_i),\end{aligned}$$

where $(z_0, z) \in \mathbb{M} \times (\mathbb{M})^m$ is a point on the line segment connecting (q_0, q) to $(\overline{q}_0, \overline{q})$. If $q, \overline{q} \in \mathbb{B}_r^m$, by convexity, we also have that $z \in \mathbb{B}_r^m$. Now, using the uniform bounds on $D^2_{q_i q_j} u^{(m)}$ from the assumptions of this theorem, increasing the value of $C = C(K, r) > 0$ if necessary, we have

$$
\begin{aligned}
&\Big| u^{(m)}(\overline{q}_0, \overline{y}) - u^{(m)}(q_0, q) - D_{q_0} u^{(m)}(q_0, q) \cdot (\overline{q}_0 - q_0) \\
&\quad - \sum_{i=1}^m D_{q_i} u^{(m)}(q_0, q) \cdot (\overline{q}_i - q_i) \Big| \\
&\quad \le C|\overline{q}_0 - q_0|^2 + C|\overline{q}_0 - q_0| \sum_{i=1}^m \frac{1}{\sqrt{m}} |\overline{q}_i - q_i| \sqrt{m} |D^2_{q_i q_0} u^{(m)}| \\
&\qquad + \frac{C}{2m} \sum_{i=1}^m |\overline{q}_i - q_i|^2 + \frac{C}{2} \left(\sum_{j=1}^m \frac{1}{m} |\overline{q}_j - q_j|^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^m \frac{1}{m} |\overline{q}_i - q_i|^2 \right)^{\frac{1}{2}} \\
&\quad \le C\Big(|q_0 - \overline{q}_0|^2 + W_2^2\big(\mu_q^{(m)}, \mu_{\overline{q}}^{(m)}\big) \Big),
\end{aligned} \tag{3.12}
$$

where in the last inequality we have used a Cauchy-Schwarz and a Young inequality, i.e.,

$$
\begin{aligned}
&|\overline{q}_0 - q_0| \left| \sum_{i=1}^m \frac{1}{\sqrt{m}} |\overline{q}_i - q_i| \sqrt{m} |D^2_{q_i q_0} u^{(m)} \right| \\
&\quad \le |\overline{q}_0 - q_0| \left(\sum_{i=1}^m \frac{1}{m} |\overline{q}_i - q_i|^2 \right)^{\frac{1}{2}} \big(m |D^2_{q_i q_0} u^{(m)}|^2 \big)^{\frac{1}{2}} \\
&\quad \le \frac{1}{2} |\overline{q}_0 - q_0|^2 + \frac{C}{2} \sum_{i=1}^m \frac{1}{m} |\overline{q}_i - q_i|^2.
\end{aligned}
$$

Now, using the previous constructions, the first line in the chain of inequalities (3.12) can be rewritten as

$$
\begin{aligned}
&u^{(m)}(\overline{q}_0, \overline{q}) - u^{(m)}(q_0, q) - D_{q_0} u^{(m)}(q_0, q) \cdot (\overline{q}_0 - q_0) \\
&\quad - \sum_{i=1}^m D_{q_i} u^{(m)}(q_0, q) \cdot (\overline{q}_i - q_i) \\
&\quad = u\big(\overline{q}_0, \mu_{\overline{q}}^{(m+1)}\big) - u\big(q_0, \mu_q^{(m+1)}\big) - \Phi_0^{(m)}\big(q_0, \mu_q^{(m)}\big) \cdot (\overline{q}_0 - q_0) \\
&\qquad - \int_{\mathbb{M}^2} \Phi_1^{(m)}\big(q_0, q, \mu_q^{(m)}\big) \cdot (\overline{q} - q) \gamma^{(m)}(dq, d\overline{q}),
\end{aligned} \tag{3.13}
$$

where $(q_i)_{i=1}^m$ and $(\overline{q}_i)_{i=1}^m$ are ordered in such a way that

$$W_2^2\big(\mu_q^{(m)}, \mu_{\overline{q}}^{(m)}\big) = \frac{1}{m}\sum_{i=1}^m |q_i - \overline{q}_i|^2,$$

$$\gamma^{(m)} := \frac{1}{m}\sum_{i=1}^m \delta_{(q_i,\overline{q}_i)} \in \Gamma_o\big(\mu_q^{(m)}, \mu_{\overline{q}}^{(m)}\big).$$

In what follows, we pass to the limit all the terms in the previous line, keeping in mind that only the integral term needs some additional effort. We have

$$\begin{aligned}
&\int_{\mathbb{M}^2} \Phi_1^{(m)}(q_0, e, \mu_q^{(m)}) \cdot (\overline{e} - e)\gamma^{(m)}(de, d\overline{e}) \\
&\quad = \int_{\mathbb{M}^2} \Phi_{1,K,r}^{(m)}(q_0, e, \mu_q^{(m)}) \cdot (\overline{e} - e)\gamma^{(m)}(de, d\overline{e}) \\
&\qquad + \int_{\mathbb{M}^2} \big(\Phi_1^{(m)}(q_0, e, \mu_q^{(m)}) - \Phi_{1,K,r}^{(m)}(q_0, e, \mu_q^{(m)})\big) \cdot (\overline{e} - e)\gamma^{(m)}(de, d\overline{e}).
\end{aligned} \tag{3.14}$$

Let us observe that

$$\begin{aligned}
&\left|\int_{\mathbb{M}^2} \big(\Phi_1^{(m)}(q_0, e, \mu_q^{(m)}) - \Phi_{1,K,r}^{(m)}(q_0, e, \mu_q^{(m)})\big) \cdot (\overline{e} - e)\gamma^{(m)}(de, d\overline{e})\right| \\
&\quad \leq \frac{C}{\sqrt{m}}\int_{\mathbb{M}^2} |e - \overline{e}|\gamma^{(m)}(de, d\overline{e}) \leq \frac{2rC}{\sqrt{m}}.
\end{aligned} \tag{3.15}$$

The next step in our argument to pass to the limit in the remaining integral in the first line of (3.15) works as follows. Fix a compact set $K \subset \mathbb{M}$, $R > 0$, $q_0 \in K$, and let $\mu, \nu \in \mathscr{P}(\overline{B}_R(0))$ and $\gamma \in \Gamma_o(\mu, \nu)$. Moreover, let $x, y \in \mathbb{H}$ be such that $\sharp(x, y) = \gamma$, which implies $\sharp(x) = \mu$, $\sharp(y) = \nu$. For $m \in \mathbb{N}$, recall $(\Omega_j^m)_{j=1}^m$ is the partition introduced in Section 1. Let us notice that for a.e. $\omega \in \Omega$, $(x(\omega), y(\omega)) \in \operatorname{spt}(\gamma)$. Let $(\omega_i)_{i=1}^m$ be Lebesgue points of (x, y) such that $\omega_i \in \Omega_i$ for all $i \in \{1, \dots, m\}$. Let us define

$$q_i := x(\omega_i),\ \overline{q}_i := y(\omega_i),\ q := (q_1, \dots, q_m), \overline{q} := (\overline{q}_1, \dots, \overline{q}_m) \in \mathbb{B}_r^m,$$

for all $i \in \{1, \dots, m\}$. We will assume that we have chosen the Lebesgue points such that $M_m^q \to x$, $M_m^{\overline{q}} \to y$ as $m \to +\infty$, strongly in $\mathbb{H}$. We have that $\{(q_i, \overline{q}_i)\}_{i=1}^m$ is contained in $\operatorname{spt}(\gamma)$ and so, it is cyclical monotone. This implies that if we define $\gamma^{(m)} := 1/m \sum_{i=1}^m \delta_{(q_i,\overline{q}_i)}$, then by the monotonicity of the set of these points, one has that

$$\gamma^{(m)} \in \Gamma_o(\mu_q^{(m)}, \mu_{\overline{q}}^{(m)}).$$

Let us underline that in our construction it is very important that $\gamma^{(m)}$ be an optimal plan and a necessary and sufficient condition, for this is the cyclical monotonicity of its support (cf. [43, 44]).

Furthermore, as the supports of the measure involved are contained in the compact set $\overline{B}_R(0)$, we have the following narrow convergence

$$\gamma^{(m)} \rightharpoonup \gamma,\ m \to +\infty, \qquad \lim_{m\to\infty} W_2\big(\mu_q^{(m)}, \mu\big) = \lim_{m\to\infty} W_2\big(\mu_{\overline{q}}^{(m)}, \nu\big) = 0.$$

As

$$\sharp(M_m^q) = \mu_q^{(m)},\ \sharp(M_m^{\overline{q}}) = \mu_{\overline{q}}^{(m)},\ \text{and}\ \sharp(M_m^q, M_m^{\overline{q}}) = \gamma^{(m)},$$

we have in particular

$$W_2^2\big(\mu_q^{(m)}, \mu_{\overline{q}}^{(m)}\big) = \sum_{i=1}^m \frac{1}{m}|q_i - \overline{q}_i|^2 = \|M_m^q - M_m^{\overline{q}}\|^2.$$

By the uniform Lipschitz property of $\Phi_{1,K,r}^{(m)}$, we have

$$\lim_{m\to\infty} \Phi_{1,K,r}^{(m)}\big(q_0, M_m^q(\omega), \mu_q^{(m)}\big) = \Phi_{1,K,r}^{\infty}(q_0, x(\omega), \mu)$$

and

$$\lim_{m\to\infty} \Phi_{1,K,r}^{(m)}\big(q_0, M_m^{\overline{q}}(\omega), \mu_{\overline{q}}^{(m)}\big) = \Phi_{1,K,r}^{\infty}(q_0, y(\omega), \nu),$$

for a.e. ω in Ω. Also, since for a.e. $\omega \in \Omega$, (3.10) implies

$$\Phi_{1,K,r}^{(m)}\big(q_0, M_m^q(\omega), \mu_q^{(m)}\big) = m D_{q_i} u^{(m)}(q_0, q) + O(1/\sqrt{m}),$$

for some $i \in \{1, \dots, m\}$, by the assumption Property 2.2(1)(a), we have that $(\Phi_{1,K,r}^m(q_0, M_m^q(\cdot), \mu_q^{(m)}))_m$ is a uniformly bounded sequence. Therefore, using all these facts, Lebesgue's dominated convergence theorem yields that up to passing to a suitable subsequence, that we do not relabel, we obtain

$$\begin{aligned}\lim_{m\to\infty} &\big\|\Phi_{1,K,r}^{(m)}\big(q_0, M_m^q, \mu_q^{(m)}\big) - \Phi_{1,K,r}^{\infty}(q_0, x, \mu)\big\| \\ &= \lim_{m\to\infty} \big\|\Phi_{1,K,r}^{(m)}\big(q_0, M_m^{\overline{q}}, \mu_{\overline{q}}^{(m)}\big) - \Phi_{1,K,r}^{\infty}(q_0, y, \nu)\big\| = 0.\end{aligned}$$

Now, using a suitable subsequence that we do not relabel, we conclude

$$\begin{aligned}\lim_{m\to\infty} &\int_{\mathbb{M}^2} \Phi_{1,K,r}^{(m)}\big(q_0, q, \mu_q^{(m)}\big) \cdot (\overline{e} - e)\gamma^{(m)}(de, d\overline{e}) \\ &= \lim_{m\to\infty} \int_\Omega \Phi_{1,K,r}^{(m)}\big(q_0, M_m^q(\omega), \mu_q^{(m)}\big) \cdot \big(M_m^{\overline{q}}(\omega) - M_m^q(\omega)\big) d\omega \\ &= \int_\Omega \Phi_{1,K,r}^{\infty}(q_0, x(\omega), \mu) \cdot (y(\omega) - x(\omega)) d\omega \\ &= \int_{\mathbb{M}^2} \Phi_{1,K,r}^{\infty}(q_0, e, \mu) \cdot (\overline{e} - e)\gamma(de, d\overline{e}).\end{aligned}$$

We combine (3.12) and (3.13) to obtain

$$\begin{aligned}&\Big|u(\bar{q}_0,\nu)-u(q_0,\mu)-\Phi^{\infty}_{0,K,r}(q_0,\mu)\cdot(\bar{q}_0-q_0)\\&\quad-\int_{\mathbb{M}^2}\Phi^{\infty}_{1,K,r}(q_0,e,\mu)\cdot(\bar{e}-e)\gamma(de,d\bar{e})\Big|\\&\quad\le C(K,r)\left(|q_0-\bar{q}_0|^2+W_2^2(\mu,\nu)\right).\end{aligned}$$

We underline that the previous inequality has only been established under the condition that $\mu,\nu\in\mathscr{B}_r$ have compact support. Since u is continuous, we combine (3.9) and (3.11) to conclude

$$\begin{aligned}&\Big|u(\bar{q}_0,\nu)-u(q_0,\mu)-\Phi^{\infty}_{0,K,r}(q_0,\mu)\cdot(\bar{q}_0-q_0)\\&\quad-\int_{\mathbb{M}^2}\Phi^{\infty}_{1,K,r}(q_0,e,\mu)\cdot(\bar{e}-e)\gamma(de,d\bar{e})\Big|\\&\quad\le C(K,r)\left(|q_0-\bar{q}_0|^2+W_2^2(\mu,\nu)\right)\end{aligned}\tag{3.16}$$

for any $q_0,\bar{q}_0\in K$ and $\mu,\nu\in\mathscr{B}_r$.

Note that in (3.16), $\Phi^{\infty}_{0,K,r}$ and $\Phi^{\infty}_{1,K,r}$ depend a priori on K and r. However since K and r are arbitrary, u is differentiable at every $(q_0,\mu)\in\mathbb{M}\times\mathscr{P}_2(\mathbb{M})$. We have that $\Phi^{\infty}_{0,K,r}(q_0,\mu)$ must coincide with $D_{q_0}u(q_0,\mu)$ which is uniquely determined and so, it is independent of K and r. Furthermore, the Wasserstein sub- and super-differentials of $u(q_0,\cdot)$ at μ coincide and contain a unique element of minimal norm $\nabla_w u(q_0,\mu)$. We do not know that $\Phi^{\infty}_{1,K,r}(q_0,\cdot,\mu)$ equals to $\nabla_w u(q_0,\mu)(\cdot)$, however, for $\gamma\in\Gamma_o(\mu,\nu)$, (3.16) implies

$$\begin{aligned}&\Big|u(\bar{q}_0,\nu)-u(q_0,\mu)-D_{q_0}u(q_0,\mu)\cdot(\bar{q}_0-q_0)\\&\quad-\int_{\mathbb{M}^2}\nabla_w u(q_0,\mu)(e)\cdot(\bar{e}-e)\gamma(de,d\bar{e})\Big|\\&\quad\le C(K,r)\left(|q_0-\bar{q}_0|^2+W_2^2(\mu,\nu)\right)\end{aligned}\tag{3.17}$$

for any $q_0,\bar{q}_0\in K$ and $\mu,\nu\in\mathscr{B}_r$. In fact, we notice that $\nabla_w u(q_0,\mu)$ is the projection of $\Phi^{\infty}_{1,K,r}(q_0,\cdot,\mu)$ onto $T_\mu\mathscr{P}_2(\mathbb{R}^d)$. □

Using the exact same steps as in the proof of Theorem 3.16, we can show an analogous result for functions depending on time as well. We formulate this in the following:

COROLLARY 3.17. *Let $u:(0,+\infty)\times\mathbb{M}\times\mathscr{P}_2(\mathbb{M})\to\mathbb{R}$ be a continuous function. For $m\in\mathbb{N}$, we define $u^{(m)}:(0,+\infty)\times\mathbb{M}\times(\mathbb{M})^m\to\mathbb{R}$ as*

$$u^{(m)}(t_0,q_0,q):=u(t_0,q_0,\mu_q^{(m+1)}),$$

where $(q_0, q) = (q_0, q_1, \dots, q_m) \in (\mathbb{M})^{m+1}$ *and* $\mu_q^{(m+1)} = \frac{1}{m+1}\sum_{i=0}^m \delta_{q_i}$. *Suppose that* $u^{(m)} \in C^{1,1}_{\mathrm{loc}}((0,+\infty)\times\mathbb{M}\times(\mathbb{M})^m)$ *and that for* $I \subset (0,+\infty)$ *and* $K \subset \mathbb{M}$ *compacts and* $r > 0$, $u^{(m)}(t_0, q_0, \cdot)$ *satisfies the estimates of Property* 2.2*(1)-(a) and (2) for all* $(t_0, q_0) \in I \times K$, *with a constant* $C = C(I, K, r) > 0$. *We assume moreover that for any* $I \subset (0,+\infty)$ *and* $K \subset \mathbb{M}$ *compacts and* $r > 0$, *there exists* $C = C(I, K, r) > 0$ *such that*

$$
\begin{gathered}
\left|D_{q_0} u^{(m)}(t_0, q_0, q)\right| \le C, \quad \left|D^2_{q_0 q_0} u^{(m)}(t_0, q_0, q)\right|_\infty \le C,\\
\sum_{i=1}^m m|D^2_{q_i q_0} u^{(m)}(t_0, q_0, q)|^2_\infty \le C,\\
\left|D^2_{q_i q_j} u^{(m)}(t_0, q_0, q)\right|_\infty \le \begin{cases} \dfrac{C}{m}, & i = j,\ i > 0,\\ \dfrac{C}{m^2}, & i \ne j,\ i, j > 0, \end{cases}
\end{gathered}
\tag{3.18}
$$

and

$$
\begin{gathered}
\left|\partial_{t_0} u^{(m)}(t_0, q_0, q)\right| \le C, \quad \left|\partial^2_{t_0 t_0} u^{(m)}(t_0, q_0, q)\right| \le C,\\
|\partial_{t_0} D_{q_0} u^{(m)}(t_0, q_0, q)| \le C, \quad \sum_{i=1}^m m|D_{q_i}\partial_{t_0} u^{(m)}(t_0, q_0, q)|^2 \le C,
\end{gathered}
\tag{3.19}
$$

for any $(t_0, q_0) \in I \times K$ *and* $q = (q_1, \dots, q_m) \in \mathbb{B}_r^m$.

Then, there exists $\Phi_1 : (0,+\infty)\times\mathbb{M}\times\mathscr{P}_2(\mathbb{M})\times\mathbb{M} \to \mathbb{R}^d$ *locally Lipschitz-continuous function such that for any* $r > 0$ *and* $I \subset (0,+\infty)$ *and* $K \subset \mathbb{M}$ *compacts, there exists* $C = C(I, K, r) > 0$ *such that for any* $s_0, t_0 \in I$, $q_0, y_0 \in K$, *any* $\mu, \nu \in \mathscr{P}_2(\mathbb{M})$, *and* $\gamma \in \Gamma_o(\mu, \nu)$, u *satisfies*

$$
\begin{aligned}
&\Big| u(y_0, \nu) - u(q_0, \mu) - D_{q_0} u(q_0, \mu)\cdot(y_0 - q_0)\\
&\quad - \int_{\mathbb{M}^2} \Phi_1(q_0, \mu, q)\cdot(y - q)\, d\gamma(q, y)\Big|\\
&\quad \le C\left(|q_0 - y_0|^2 + W_2^2(\mu, \nu)\right).
\end{aligned}
$$

This implies in particular that

$$
u \in C^{1,1}_{\mathrm{loc}}((0,+\infty)\times\mathbb{M}\times\mathscr{P}_2(\mathbb{M})), \nabla_w u(t_0, q_0, \mu)(\cdot)
$$

is the projection of $\Phi_1(t_0, q_0, \mu, \cdot)$ *onto* $T_\mu\mathscr{P}_2(\mathbb{M})$ *and*

$$
\begin{aligned}
&\Big| u(s_0, y_0, \nu) - u(t_0, q_0, \mu) - D_{q_0} u(t_0, q_0, \mu)\cdot(y_0 - q_0)\\
&\quad \partial_{t_0} u(t_0, q_0, \mu)(s_0 - t_0) - \int_{\mathbb{M}^2} \Phi_1(t_0, q_0, \mu, q)\cdot(y - q)\, d\gamma(q, y)\\
&\quad \le C\left(|s_0 - t_0|^2 + |q_0 - y_0|^2 + W_2^2(\mu, \nu)\right).
\end{aligned}
$$

THEOREM 3.18. *Let* $\mathscr{U} \in C^{1,1}_{\rm loc}(\mathscr{P}_2(\mathbb{M}))$. *Let* $U^{(m)} : (\mathbb{M})^m \to \mathbb{R}$ *be defined as* $U^{(m)}(q) := \mathscr{U}(\mu^{(m)}_q)$ *for* $q \in \mathbb{M}^m$ *such that Property* 2.2(2–3) *are satisfied. Then* $\mathscr{U} \in C^{2,1,w}_{\rm loc}(\mathscr{P}_2(\mathbb{M}))$ *in the sense of Definition* 3.13 *such that the following hold. There exist* $C : (0,\infty) \to (0,\infty)$ *monotone nondecreasing and*

(i) *there are continuous maps*

$$\Lambda_0 : \mathbb{M} \times \mathscr{P}_2(\mathbb{M}) \to \mathbb{R}^{d\times d} \quad \text{and} \quad \Lambda_1 : \mathbb{M} \times \mathbb{M} \times \mathscr{P}_2(\mathbb{M}) \to \mathbb{R}^{d\times d}$$

such that for $\mu \in \mathscr{P}_2(\mathbb{M})$ *we have*

$$\sup_{\mu\in\mathscr{B}_r} \|\Lambda_0(\cdot,\mu)\|_{L^\infty(\mu)}, \quad \sup_{\mu\in\mathscr{B}_r} \|\Lambda_1(\cdot,\cdot,\mu)\|_{L^\infty(\mu\otimes\mu)} \le C(r).$$

(ii) *Let* $\mu,\nu \in \mathscr{B}_r$ *and* $\gamma \in \Gamma_o(\mu,\nu)$. *We have*

$$\begin{aligned} &\Big|\nabla_w \mathscr{U}(\nu)(\bar q) - \nabla_w \mathscr{U}(\mu)(q) - \Lambda_0(q,\mu)(\bar q - q) \\ &\quad - \int_{\mathbb{M}^2} \Lambda_1(q,a,\mu)(b-a)\,d\gamma(a,b)\Big| \\ &\quad \le C\left(|q-\bar q|^2 + W_2^2(\mu,\nu)\right) \end{aligned} \tag{3.20}$$

and

$$\left|\nabla_w \mathscr{U}(\mu)(q) - \nabla_w \mathscr{U}(\nu)(\bar q)\right| \le C\left(|q-\bar q| + W_2(\mu,\nu)\right) \quad \forall \mu,\nu \in \mathscr{B}_r, \tag{3.21}$$

for all $(q,\bar q) \in \mathrm{spt}(\mu) \times \mathrm{spt}(\nu)$.

PROOF. We follow ideas similar to those presented in the proof of Theorem 3.16. Recall that for $q \in \mathbb{B}^m_r$, we use the notation $\mu^{(m)}_q := 1/m \sum_{i=1}^m \delta_{q_i}$ and use a similar notation for $\bar q \in \mathbb{B}^m_r$. Let us define the matrix-valued functions

$$\Lambda^{(m)}_0 : \bigcup_{q\in\mathbb{B}^m_r} \mathrm{spt}\big(\mu^{(m)}_q\big) \times \big\{\mu^{(m)}_q\big\} \to \mathbb{R}^{d\times d}$$

and

$$\Lambda^{(m)}_1 : \bigcup_{q\in\mathbb{B}^m_r} \Big(\big(\mathrm{spt}\big(\mu^{(m)}_q\big) \times \mathrm{spt}\big(\mu^{(m)}_q\big)\big) \setminus \{(q_i,q_i) : i = 1,\dots,m\}\Big) \times \big\{\mu^{(m)}_q\big\} \to \mathbb{R}^{d\times d}$$

as

$$\begin{aligned} \Lambda^{(m)}_0\big(q_i,\mu^{(m)}_q\big) &:= m D^2_{q_iq_i} U^{(m)}(q), \\ \Lambda^{(m)}_1(q_i,q_j,\mu^{(m)}_q) &:= m^2 D^2_{q_iq_j} U^{(m)}(q), \quad i \ne j. \end{aligned}$$

Let us underline that we have not defined $\Lambda^{(m)}_1(q_i,q_i,\mu^{(m)}_q)$ for $i = j$. Because of this, later we will need special care when one passes to the limit the corresponding objects as $m \to +\infty$.

We observe that as a consequence of the assumptions and Lemma 3.15(v)-(b,d), we have that for any $r > 0$, there exists a constant $C = C(r) > 0$ such that

$$\left|\Lambda_0^{(m)}(q_i, \mu_q^{(m)}) - \Lambda_0^{(m)}(\overline{q}_j, \mu_{\overline{q}}^{(m)})\right| \le C\left(|q_i - \overline{q}_j| + W_2(\mu_q^{(m)}, \mu_{\overline{q}}^{(m)}) + \frac{1}{\sqrt{m}}\right)$$

and

$$\begin{aligned}\left|\Lambda_1^{(m)}\big(q_i, q_k, \mu_q^{(m)}\big) - \Lambda_1^{(m)}\big(\overline{q}_j, \overline{q}_l, \mu_{\overline{q}}^{(m)}\big)\right| \\ \le C\left(|q_i - \overline{q}_j| + |q_k - \overline{q}_l| + W_2\big(\mu_q^{(m)}, \mu_{\overline{q}}^{(m)}\big) + \frac{1}{\sqrt{m}}\right)\end{aligned}$$

for any $q, \overline{q} \in \mathbb{B}_r^m$, and for any $i, j, k, l \in \{1, \dots, m\}$, $i \ne k$, $j \ne l$. For every coordinate function $(\Lambda_0^{(m)})_{\alpha\beta}, (\Lambda_1^{(m)})_{\alpha\beta}$ $(\alpha, \beta \in \{1, \dots, d\})$, we define the extensions

$$\big(\Lambda_{0,r}^{(m)}\big)_{\alpha\beta} : \mathbb{M} \times \mathscr{P}_2(\mathbb{M}) \to \mathbb{R} \quad \text{and} \quad \big(\Lambda_{1,r}^{(m)}\big)_{\alpha\beta} : \mathbb{M} \times \mathbb{M} \times \mathscr{P}_2(\mathbb{M}) \to \mathbb{R}$$

as follows. For $z, z_1, z_2 \in \mathbb{M}$, $\mu \in \mathscr{P}_2(\mathbb{M})$ we set

$$\big(\Lambda_{0,r}^{(m)}\big)_{\alpha\beta}(z, \mu) := \inf\big\{(\Lambda_0^{(m)})_{\alpha\beta}(q_i, \mu_q^{(m)}) + C\big(|q_i - z| + W_2(\mu_q^{(m)}, \mu)\big\}$$

and

$$\begin{aligned}&\big(\Lambda_{1,r}^{(m)}\big)_{\alpha\beta}(z_1, z_2, \mu) := \\ &\quad \inf\big\{\big(\Lambda_1^{(m)}\big)_{\alpha\beta}\big(q_i, q_k, \mu_q^{(m)}\big) + C(|q_i - z_1| + |q_k - z_2|) + W_2\big(\mu_q^{(m)}, \mu\big)\big\},\end{aligned}$$

where both infima are taken over $q \in \mathbb{B}_r^m$, $i, k \in \{1, \dots, m\}, i \ne k$.

Recall that $\Lambda_{0,r}^{(m)}$ and $\Lambda_{1,r}^{(m)}$ are $C(r)$-Lipschitz, and we have

$$\left|\Lambda_{0,r}^{(m)}\big(q_i, \mu_q^{(m)}\big) - \Lambda_0^{(m)}\big(q_i, \mu_q^{(m)}\big)\right|_\infty \le \frac{C}{\sqrt{m}} \quad \forall q \in \mathbb{B}_r^m,\ i \in \{1, \dots, m\} \tag{3.22}$$

and

$$\begin{aligned}\left|\Lambda_{1,r}^{(m)}\big(q_i, q_k, \mu_q^{(m)}\big) - \Lambda_1^{(m)}\big(q_i, q_k, \mu_q^{(m)}\big)\right|_\infty \le \frac{C}{\sqrt{m}} \\ \forall q \in \mathbb{B}_r^m,\ i, k \in \{1, \dots, m\},\ i \ne k.\end{aligned} \tag{3.23}$$

If $R > 0$, $z_1, z_2 \in B_R(0)$, and μ is supported by $B_R(0)$, then for all $\alpha, \beta \in \{1, \dots, d\}$

$$-C \le \big(\Lambda_{1,r}^{(m)}\big)_{\alpha\beta}(z_1, z_2, \mu) \le C + C\big(|z_1| + |z_2| + W_2(0, \mu)\big) \le C(3R).$$

We obtain a similar uniform bound on $(\Lambda_{0,r}^{(m)})_m$. As in the proof of Theorem 3.16, there are C-Lipschitz functions

$$\Lambda_{0,r} : \mathbb{M} \times \mathscr{P}_2(\mathbb{M}) \to \mathbb{R}^{d\times d}, \quad \Lambda_{1,r} : \mathbb{M} \times \mathbb{M} \times \mathscr{P}_2(\mathbb{M}) \to \mathbb{R}^{d\times d},$$

locally bounded, respectively, on $\mathbb{M} \times \mathscr{P}_2(\mathbb{M})$ and $\mathbb{M}^2 \times \mathscr{P}_2(\mathbb{M})$ by a constant depending only on r and R. Up to a subsequence, as $m \to +\infty$, $(\Lambda_{0,r}^{(m)})_m$

and $(\Lambda_{1,r}^{(m)})_m$ converge to $\Lambda_{0,r}$ and $\Lambda_{1,r}$, uniformly on $\bar{B}_R(0) \times \mathscr{P}(\bar{B}_R(0))$ and $\bar{B}_R(0) \times \bar{B}_R(0) \times \mathscr{P}(\bar{B}_R(0))$, respectively.

Our next task is to show that

$$\Lambda_{0,r}(\cdot,\mu) \in L^\infty(\mathbb{M};\mu), \quad \Lambda_{1,r}(\cdot,\cdot,\mu) \in L^\infty(\mathbb{M}\times\mathbb{M};\mu\otimes\mu), \quad \forall \mu \in \mathscr{B}_r \cap \mathscr{P}(\bar{B}_R(0)). \tag{3.24}$$

CLAIM 1. $\Lambda_{1,r}(\cdot,\cdot,\mu) \in L^\infty(\mathbb{M}^2;\mu\otimes\mu)$.

PROOF OF CLAIM 1. Let $r > 0$, $R > 0$, and first let $\mu \in \mathscr{B}_R \cap \mathscr{P}(\bar{B}_R(0))$. Let $z_1, z_2 \in B_R(0)$. As we plan to let m tend to ∞, there is no loss of generality to assume $R \le r\sqrt{m}$. Since $q = (z_1, z_2, 0, \dots, 0) \in \mathbb{B}_r^m$, we have

$$\begin{aligned}
-C \le \big(\Lambda_{1,r}^{(m)}\big)_{\alpha\beta}(z_1,z_2,\mu) &\le \big(\Lambda_1^{(m)}\big)_{\alpha\beta}(z_1,z_2,\mu_q^{(m)}) \\
&\quad + C(r)\big(|z_1 - z_1| + |z_2 - z_2| + W_2\big(\mu_q^{(m)},\mu\big)\big) \\
&\le C(r) + 2rC(r).
\end{aligned}$$

Letting m tend to ∞ we conclude that $\big|\big(\Lambda_{1,r}\big)_{\alpha\beta}(z_1,z_2,\mu)\big| \le C(r) + 2rC(r)$ first on $\mathbb{M}^2 \times \mathscr{P}_c(\mathbb{M})$ and by continuity, this holds on $\mathbb{M}^2 \times \mathscr{P}_2(\mathbb{M})$.

CLAIM 2. $\Lambda_{0,r}(\cdot,\mu) \in L^\infty(\mathbb{M};\mu)$.

PROOF OF CLAIM 2. The proof is similar to but simpler than that of Claim 1.

For $q, \bar{q} \in \mathbb{B}_r^m$ we have the expansion

$$\begin{aligned}
&mD_{q_1}U^{(m)}(\bar{q}) - mD_{q_1}U^{(m)}(q) - mD^2_{q_1q_1}U^{(m)}q)(\bar{q}_1 - q_1) \\
&\quad - m\sum_{k=2}^m D^2_{q_1q_k}U^{(m)}(q)(\bar{q}_k - q_k) \\
&\quad = \frac{m}{2}\sum_{k,l=1}^m (\bar{q}_l - q_l)D^3_{q_1q_kq_l}U^{(m)}(z)(\bar{q}_k - q_k)
\end{aligned} \tag{3.25}$$

where z is a point on the line segment connecting q to $\bar{q}$.

Let $\mu, \nu \in \mathscr{B}_r$, $\gamma \in \Gamma_o(\mu,\nu)$, and let $(q_1, \bar{q}_1) \in \operatorname{spt}(\mu) \times \operatorname{spt}(\nu)$ (which is not necessarily in $\operatorname{spt}(\gamma)$). Suppose that both $\operatorname{spt}(\mu)$ and $\operatorname{spt}(\nu)$ contain more than one element. We choose $x, y \in \mathbb{H}$ such that $\sharp(x,y) = \gamma$ and so, $\sharp(x) = \mu$, $\sharp(y) = \nu$. Let $(\Omega_i^{m-1})_{i=1}^{m-1}$ be the partition of Ω introduced in Section 1. We are going to choose special values of $m := 2^l + 1$ and choose Lebesgue points $\omega_{i+1} \in \Omega_i^{2^l}$ such that all the points in $\Omega_i^{2^l}$ are kept in $\Omega_i^{2^{l+1}}$. We set $q_i := x(\omega_i)$, $\bar{q}_i := y(\omega_i)$

for $i = 2, \cdots, m$ Set

$$\gamma^{(m-1)} := \frac{1}{m-1}\sum_{i=2}^{m} \delta_{(q_i,\overline{q}_i)}, \quad \mu_q^{(m-1)} := \frac{1}{m-1}\sum_{i=2}^{m} \delta_{q_i},$$
$$\mu_{\overline{q}}^{(m-1)} := \frac{1}{m-1}\sum_{i=2}^{m} \delta_{\overline{q}_i}.$$

Since, $(q_i, \overline{q}_i)_{i=2}^{\infty}$ is cyclically monotone,

$$\gamma^{(m-1)} \in \Gamma_o\big(\mu_q^{(m-1)}, \mu_{\overline{q}}^{(m-1)}\big).$$

By construction $(\gamma^{(m-1)})_m$ converges narrowly to γ. Let $M^q_{(m-1)}, M^{\overline{q}}_{(m-1)} \in \mathbb{H}$, the random variables corresponding to the previously chosen points $(q_2, \ldots, q_m)$ and $(\overline{q}_2, \ldots, \overline{q}_m)$, respectively. We have

$$\begin{aligned} &\lim_{m\to+\infty} W_2(\mu_q^{(m)}, \mu) \\ &\quad = \lim_{m\to+\infty} W_2(\mu_q^{(m-1)}, \mu) = \lim_{m\to+\infty} W_2(\mu_{\overline{q}}^{(m)}, \nu) \\ &\quad = \lim_{m\to+\infty} W_2(\mu_{\overline{q}}^{(m-1)}, \nu) = 0. \end{aligned} \tag{3.26}$$

Furthermore,

$$\sharp\big(M^q_{(m-1)}, M^{\overline{q}}_{(m-1)}\big) = \gamma^{(m-1)},$$

and

$$\lim_{m\to+\infty} \big\| M^q_{(m-1)} - x \big\| = \lim_{m\to+\infty} \big\| M^{\overline{q}}_{(m-1)} - y \big\| = 0.$$

Using the assumptions on $D^3_{q_j q_k q_l} U^{(m)}$, since $z \in \mathbb{B}^m_r$, increasing the value of C if necessary, we have

$$
\begin{aligned}
&\left| m \sum_{k,l=1}^{m} (y_l - x_l) D^3_{q_1 q_k q_l} U^{(m)}(z)(\overline{q}_k - q_k) \right| \\
&\le m |D^3_{q_1 q_1 q_1} U^{(m)}(z)|_\infty |\overline{q}_1 - q_1|^2 \\
&\quad + m \sum_{k=2}^{m} |D^3_{q_1 q_k q_1} U^{(m)}(z)|_\infty |\overline{q}_k - q_k||\overline{q}_1 - q_1| \\
&\quad + m \sum_{l=2}^{m} |D^3_{q_1 q_1 q_l} U^{(m)}(z)|_\infty |\overline{q}_1 - q_1||\overline{q}_l - q_l| \\
&\quad + m \sum_{k=2}^{m} |D^3_{q_1 q_k q_k} U^{(m)}(z)|_\infty |\overline{q}_k - q_k|^2 + \\
&\quad + m \sum_{k \neq l = 2}^{m} |\overline{q}_l - q_l||D^3_{q_1 q_k q_l} U^{(m)}(z)|_\infty |\overline{q}_k - q_k| \\
&\le C \left(|\overline{q}_1 - q_1|^2 + |\overline{q}_1 - q_1| \sum_{k=2}^{m} \frac{1}{m} |\overline{q}_k - q_k| + \sum_{k=2}^{m} \frac{1}{m} |\overline{q}_k - q_k|^2 \right) \\
&\quad + \frac{C}{m^2} \sum_{k \neq l = 2}^{m} |\overline{q}_l - q_l||\overline{q}_k - q_k| \\
&\le C \big(|\overline{q}_1 - q_1|^2 + W_2^2 \big(\mu_q^{(m-1)}, \mu_{\overline{q}}^{(m-1)} \big) \big).
\end{aligned}
$$

Thus, this together with (3.25) implies

$$
\begin{aligned}
m &\Bigg| D_{q_1} U^{(m)}(\overline{q}) - D_{q_1} U^{(m)}(q) - D^2_{q_1 q_1} U^{(m)}(q)(\overline{q}_1 - q_1) \\
&- \sum_{k=2}^{m} D^2_{q_1 q_k} U^{(m)}(q)(\overline{q}_k - q_k) \Bigg| \le C \big(|\overline{q}_1 - q_1|^2 + W_2^2 \big(\mu_q^{(m-1)}, \mu_{\overline{q}}^{(m-1)} \big) \big).
\end{aligned}
$$

Using the definition of $\Lambda_0^{(m)}$ and $\Lambda_1^{(m)}$ we read off

$$
\begin{aligned}
&\Big| \nabla_w \mathscr{U} \big(\mu_{\overline{q}}^{(m)} \big)(\overline{q}_1) - \nabla_w \mathscr{U} \big(\mu_q^{(m)} \big)(q_1) - \Lambda_0^{(m)} \big(q_1, \mu_q^{(m)} \big)(\overline{q}_1 - q_1) \\
&- \frac{m-1}{m} \int_{\mathbb{M}^2} \Lambda_1^{(m)}(q_1, a, \mu_q^{(m)})(b - a) \gamma^{(m-1)}(da, db) \Big| \\
&\quad \le C \big(|\overline{q}_j - q_i|^2 + W_2^2 \big(\mu_q^{(m-1)}, \mu_{\overline{q}}^{(m-1)} \big) \big),
\end{aligned} \tag{3.27}
$$

Now, first by the continuity of $\nabla_w \mathscr{U}$, (3.26) implies

$$\lim_{m\to\infty} \nabla_w \mathscr{U}(\mu_q^{(m)})(q_1) = \nabla_w \mathscr{U}(\mu)(q_1),$$
$$\lim_{m\to\infty} \nabla_w \mathscr{U}(\mu_q^{(m)})(\overline{q}_1) = \nabla_w \mathscr{U}(\nu)(\overline{q}_1).$$

Before passing to the limit in the other terms, let us *further* suppose that $\mu, \nu \in \mathscr{P}(\overline{B}_R(0))$ for some $R > 0$. In light of (3.22), $\Lambda_0^{(m)}(q_1, \mu_q^{(m)})$ and $\Lambda_{0,r}^{(m)}(q_1, \mu_q^{(m)})$ have the same limit. By the local uniform convergence property of $\Lambda_{0,r}^{(m)}$, we have that $\lim_{m\to\infty} \Lambda_0^{(m)}(q_1, \mu_q^{(m)}) = \Lambda_{0,r}(q_1, \mu)$.

To handle the limit in the last term on the left-hand side of the inequality (3.27), we observe that

$$\begin{aligned}
&\int_{\mathbb{M}^2} \Lambda_1^{(m)}\big(q_1, a, \mu_q^{(m)}\big)(b-a)\gamma^{(m-1)}(da, db) \\
&\quad = \int_{\mathbb{M}^2} \Lambda_{1,r}^{(m)}\big(q_1, a, \mu_q^{(m)}\big)(b-a)\gamma^{(m-1)}(da, db) \\
&\qquad + \int_{\mathbb{M}^2} \big(\Lambda_1^{(m)}(q_1, a, \mu_q^{(m)}) - \Lambda_{1,r}^{(m)}(q_1, a, \mu_q^{(m)})\big)(b-a)\gamma^{(m-1)}(da, db)
\end{aligned}$$

and by (3.23), increasing C if necessary, we have that

$$\begin{aligned}
&\left| \int_{\mathbb{M}^2} \Big(\Lambda_1^{(m)}\big(q_1, a, \mu_q^{(m)}\big) - \Lambda_{1,r}^{(m)}\big(q_1, a, \mu_q^{(m)}\big)\Big)(b-a)\gamma^{(m-1)}(da, db)\right| \\
&\leq \frac{C}{\sqrt{m}} \iint_{\mathbb{M}^2} |b-a|\gamma^{(m-1)}(da, db) \\
&\leq \frac{Cr}{\sqrt{m}}.
\end{aligned}$$

Therefore, it is enough to study the limit of

$$\int_{\mathbb{M}^2} \Lambda_{1,r}^{(m)}\big(q_1, a, \mu_q^{(m)}\big)(b-a)\gamma^{(m-1)}(da, db).$$

Since

$$\left|\Lambda_{1,r}^{(m)}\big(q_1, M_{(m-1)}^q(\omega), \mu_q^{(m)}\big) - \Lambda_1^{(m)}\big(q_1, M_{(m-1)}^q(\omega), \mu_q^{(m)}\big)\right| \leq \frac{C}{\sqrt{m}}$$

and since

$$\Lambda_1^{(m)}\big(q_1, M_{(m-1)}^q(\omega), \mu_q^{(m)}\big) = \Lambda_1^{(m)}\big(q_1, q_i, \mu_q^{(m)}\big)$$

for some $i \in \{2, \dots, m\}$ for a.e. $\omega \in \Omega$, we have that

$$\omega \mapsto \Lambda_{1,r}^{(m)}\big(q_1, M_{(m-1)}^q(\omega), \mu_q^{(m)}\big)$$

is uniformly bounded with respect to $m \in \{2, 3, \dots\}$. Thus by the previous convergences and by Lebesgue's dominated convergence theorem, up to passing to a subsequence that we do not relabel, we have that

$$\lim_{m\to\infty} \left\| \Lambda_{1,r}^{(m)}(q_1, M^q_{(m-1)}, \mu_q^{(m)}) - \Lambda_1(q_1, x, \mu) \right\| = 0.$$

Thus, up to a subsequence,

$$\begin{aligned}
&\lim_{m\to\infty} \int_{\mathbb{M}^2} \Lambda_{1,r}^{(m)}\big(q_1, a, \mu_q^{(m)}\big)(b-a)\gamma^{(m-1)}(da, bb) \\
&= \lim_{m\to\infty} \int_{\Omega} \Lambda_{1,r}^{(m)}\big(q_1, M^q_{(m-1)}(\omega), \mu_q^{(m)}\big)\big(M^{\bar q}_{(m-1)}(\omega) - M^q_{(m-1)}(\omega)\big)d\omega \\
&= \int_{\Omega} \Lambda_{1,r}(q_1, x(\omega), \mu)(y(\omega) - x(\omega))d\omega \\
&= \int_{\mathbb{M}^2} \Lambda_{1,r}(q_1, a, \mu)(b-a)\gamma(da, db).
\end{aligned}$$

We have all the ingredients to conclude that up to subsequence (3.27) implies

$$\begin{aligned}
&\bigg| \nabla_w \mathscr{U}(\nu)(\bar q_1) - \nabla_w \mathscr{U}(\mu)(q_1) - \Lambda_{0,r}(q_1, \mu)(\bar q_1 - q_1) \\
&\quad - \int_{\mathbb{M}^2} \Lambda_{1,r}(q_1, a, \mu)(b-a)\gamma(da, db) \bigg| \\
&\quad \le C\big(|q_1 - \bar q_1|^2 + W_2^2(\mu, \nu)\big).
\end{aligned}$$

As C is independent of R, we extend the previous inequality to all $\mu, \nu \in \mathscr{B}_r$ without imposing that they lie in $\mathscr{P}(B_R(0))$. We also notice that by the assumptions, i.e., Property 2.2(3), the map $q \mapsto \nabla_w \mathscr{U}(\mu)(q)$ is Lipschitz-continuous uniformly with respect to $\mu \in \mathscr{B}_r$. More precisely, Lemma 3.15 (iii)-(b) yields that there exists $C = C(r) > 0$ such that for all $\mu, \nu \in \mathscr{B}_r$ and $(q_1, \bar q_1) \in \mathrm{spt}(\mu) \times \mathrm{spt}(\nu)$, we have

$$|\nabla_w \mathscr{U}(t, \mu)(q_1) - \nabla_w \mathscr{U}(t, \nu)(\bar q_1)| \le C(|q_1 - \bar q_1| + W_2(\mu, \nu)),$$

so (3.21) follows. □

Remark 3.19. Note that Λ_0 is a symmetric matrix, as a limit of symmetric matrices.

4 Global Well-Posedness of Master Equations

Throughout this section, we fix $T > 0$ and impose (H1)–(H7). We further assume

(H8) $\mathscr{U}_0, \mathscr{F} \in C^{2,1,w}_{\mathrm{loc}}(\mathscr{P}_2(\mathbb{M}))$ and $U_0^{(m)}$, $F^{(m)}$ satisfy Property 2.2(3).

Let $\widetilde{\mathscr{U}}$ be the solution obtained in Proposition 1.5 and define $\mathscr{U} : [0, T] \times \mathscr{P}_2(\mathbb{M}) \to \mathbb{R}$ as $\mathscr{U}(t, \mu) := \widetilde{\mathscr{U}}(t, x)$ where $\mu = \sharp(x)$. By Lemma 3.11, the regularity property obtained on $\widetilde{\mathscr{U}}$ in Proposition 1.5 ensures that

$$\mathscr{U}(t, \cdot) \text{ is } C^{1,1}_{\mathrm{loc}}(\mathscr{P}_2(\mathbb{M})).$$

We use Remark 3.7 to obtain that $\mathscr{U} \in C^{1,1}_{\mathrm{loc}}([0,T] \times \mathscr{P}_2(\mathbb{M}))$ (in the sense of Definition 3.8), and it is a classical solution to the Hamilton–Jacobi equation

$$\begin{cases} \partial_t \mathscr{U} + \mathscr{H}(\mu, \nabla_w \mathscr{U}) = \mathscr{F}(\mu) & \text{in } (0,T) \times \mathscr{P}_2(\mathbb{M}), \\ \mathscr{U}(0,\mu) = \mathscr{U}_0(\mu), & \text{in } \mathscr{P}_2(\mathbb{M}). \end{cases} \tag{4.1}$$

4.1 The vectorial master equation

Let $\mathscr{V} : \mathscr{P}_2(\mathbb{M}) \times \mathbb{M} \to \mathbb{R}^d$ and define

$$\bar{\mathscr{N}_\mu}\big[\mathscr{V}, \nabla_w^\top \mathscr{V}\big](t,\mu,q) := \int_{\mathbb{M}} \nabla_w^\top \mathscr{V}(t,\mu,q)(b) D_p H\big(b, \mathscr{V}(t,\mu,b)\big)\mu(db).$$

We plan to obtain existence of $\mathscr{V} : [0,T] \times \mathscr{P}_2(\mathbb{M}) \times \mathbb{M} \to \mathbb{R}^d$, a solution to the so-called *vectorial master equation*

$$\begin{cases} \partial_t \mathscr{V} + D_q H(q, \mathscr{V}(t,\mu,q)) + D_q \mathscr{V}(t,\mu,q) \nabla_p H(q, \mathscr{V}(t,\mu,q)) \\ \quad + \bar{\mathscr{N}_\mu}\big[\mathscr{V}, \nabla_w^\top \mathscr{V}\big](t,\mu,q) = \nabla_w \mathscr{F}(\mu)(q) \\ \mathscr{V}(0,\mu,\cdot) = \mathscr{V}_0(\mu), \end{cases} \tag{4.2}$$

as a by-product of the regularity properties of the solution to (4.1). The lower-order regularity results in the Hilbert setting are starting points to improve to higher-order regularity results in the Wasserstein space. First, let us discuss the existence and regularity of solutions of (4.1).

THEOREM 4.1. *The equation* (4.1) *has a unique classical solution*

$$\mathscr{U} \in C^{1,1}_{\mathrm{loc}}([0,T] \times \mathscr{P}_2(\mathbb{M}))$$

such that $\mathscr{U}(t,\cdot) \in C^{2,1,w}_{\mathrm{loc}}(\mathscr{P}_2(\mathbb{M}))$*, which has to be understood in the sense of Definition* 3.13.

PROOF. First, we notice that Proposition 1.5 asserts existence and uniqueness of a solution $\mathscr{U} \in C^{1,1}_{\mathrm{loc}}([0,T] \times \mathscr{P}_2(\mathbb{M}))$. Then, Theorem 2.3 will imply that

$$U^{(m)}(t,q) := \mathscr{U}(t, \mu_q^{(m)}) \ \text{ for } t \in (0,T),\ m \in \mathbb{N},\ q \in (\mathbb{M})^m,$$

satisfies the regularity estimates from Property 2.2 in $\mathbb{B}_r^m(0)$ with constant $C(t,r)$. We apply Theorem 3.18 to infer $\mathscr{U}(t,\cdot)$ is of class $C^{2,1,w}_{\mathrm{loc}}(\mathscr{P}_2(\mathbb{M}))$. □

Remark 4.2. In this subsection we discuss existence of *weak solutions* to (4.2). The regularity of solutions $\mathscr{U}$ to the Hamilton–Jacobi equation (4.1) established in Theorem 4.1 are enough to differentiate this equation with respect to the measure variable. This procedure gives us a notion of weak solution to the vectorial master equation. Better regularity properties of this solution are subtle, and we need additional effort to obtain these. We postpone this analysis to Section 5.1, where we point out a deep connection between the vectorial and the scalar master equations as well.

DEFINITION 4.3. We say that $\mathscr{V} : [0,T]\times\bigcup_{\mu\in\mathscr{P}_2(\mathbb{M})}\{\mu\}\times\mathrm{spt}(\mu) \to \mathbb{R}^d$ is a *weak solution* to (4.2) if it is locally Lipschitz on its domain of definition, $\mathscr{V}(\cdot,\mu,q)$ is differentiable on $(0,T)$ for all $\mu \in \mathscr{P}_2(\mathbb{M})$, and $q \in \mathrm{spt}(\mu)$,

$$\mathscr{V}(t,\cdot,\cdot) \in C^{1,1}_{\mathrm{loc}}\Big(\bigcup_{\mu\in\mathscr{P}_2(\mathbb{M})}\{\mu\}\times\mathrm{spt}(\mu)\Big),$$

$\mathscr{V}(t,\mu,\cdot)$ is differentiable on $\mathrm{spt}(\mu)$ for all $t \in [0,T]$, and $\mu \in \mathscr{P}_2(\mathbb{M})$ and the equation (4.2) is satisfied pointwise on $[0,T]\times\bigcup_{\mu\in\mathscr{P}_2(\mathbb{M})}\{\mu\}\times\mathrm{spt}(\mu)$.

THEOREM 4.4. *Suppose* $\mathscr{U}(t,\cdot) \in C^{2,1,w}_{\mathrm{loc}}(\mathscr{P}_2(\mathbb{M}))$ (*in the sense of Definition* 3.13). *Using the notation in Remark* 3.14*, we have assumed*

$$D_q\big(\nabla_w\mathscr{U}(t,\mu)(\cdot)\big) \in L^\infty(\mathbb{M};\mu),\quad \bar\nabla^2_{ww}\mathscr{U}(t,\mu)(\cdot,\cdot) \in L^\infty(\mathbb{M}\times\mathbb{M};\mu\otimes\mu),$$

$\forall\mu \in \mathscr{P}_2(\mathbb{M})$, *and a.e.* $t \in (0,T)$. *Then the vector field*

$$\mathscr{V}(t,\mu,q) := \nabla_w\mathscr{U}(t,\mu)(q)$$

defined on $[0,T]\times\bigcup_{\mu\in\mathscr{P}_2(\mathbb{M})}\{\mu\}\times\mathrm{spt}(\mu)$ *solves the vectorial master equation* (4.2) *with initial data* $\mathscr{V}_0 = \nabla_w\mathscr{U}_0$ *in the sense of Definition* 4.3.

PROOF OF THEOREM 4.4. Let $\mu \in \mathscr{P}_2(\mathbb{M})$, let $\varphi \in C^\infty_c(\mathbb{M})$ be arbitrary, and set $\xi := D\varphi$. Choose $\varepsilon > 0$ small enough such that for all $s \in [0,\varepsilon]$, $X_s := \mathrm{id} + s\xi$ is a diffeomorphism of $\mathbb{M}$ into $\mathbb{M}$ and $|\mathrm{id}|^2/2 + s\varphi$ is convex. For any $q \in \mathrm{spt}(\mu)$ we have

$$\begin{aligned}\nabla_w\mathscr{U}(t,\sigma_s)(X_s(q)) = {}& \nabla_w\mathscr{U}(t,\mu)(q) + sD_q\nabla_w\mathscr{U}(t,\mu)(q)\xi(q)\\ &+ s\int_{\mathbb{M}}\nabla^2_{ww}\mathscr{U}(t,\mu)(q,a)\xi(a)\mu(da)\\ &+ o(s).\end{aligned}\tag{4.3}$$

Since

$$\int_{\mathbb{M}} H\big(z,\nabla_w\mathscr{U}(t,\sigma_s)(z)\big)\sigma_s(dz) = \int_{\mathbb{M}} H\big(X_s(q),\nabla_w\mathscr{U}(t,\sigma_s)\big(X_s(q)\big)\big)\mu(dq),$$

(4.3) implies

$$\begin{aligned}&\mathscr{H}\big(\sigma_s,\nabla_w\mathscr{U}(t,\sigma_s)\big) = \mathscr{H}\big(\mu,\nabla_w\mathscr{U}(t,\mu)\big)\\ &\quad + s\int_{\mathbb{M}} D_qH\big(q,\nabla_w\mathscr{U}(t,\mu(q))\big)\cdot\xi(q)\mu(dq)\\ &\quad + s\int_{\mathbb{M}} D_pH\big(q,\nabla_w\mathscr{U}(t,\mu)(q)\big)\cdot\Big(D_q\nabla_w\mathscr{U}(t,\mu)(q)\xi(q)\Big)\mu(dq)\\ &\quad + s\int_{\mathbb{M}^2} D_pH\big(q,\nabla_w\mathscr{U}(t,\mu)(q)\big)\cdot\Big(\nabla^2_{ww}\mathscr{U}(t,\mu)(q,a)\xi(a)\mu(da)\Big)\mu(dq)\\ &\quad - \mathscr{F}(\mu) - s\int_{\mathbb{M}}\nabla_w\mathscr{F}(\mu)(q)\cdot\xi(q)\mu(dq) + o(s).\end{aligned}\tag{4.4}$$

Similarly,

$$\partial_t \mathscr{U}(t,\sigma_s) = \partial_t \mathscr{U}(t,\mu) + s\int_{\mathbb{M}} \partial_t \nabla_w \mathscr{U}(t,\mu)(q)\cdot\xi(q)\mu(dq) + o(s). \tag{4.5}$$

Note that since $\mathscr{U}$ is a $C^{1,1}_{\mathrm{loc}}([0,T]\times\mathscr{P}_2(\mathbb{M}))$ solution to (4.1), $\nabla_w\mathscr{U}(\,\cdot\,,\mu)(q)$ is Lipschitz-continuous on $[0,T]$. Moreover, from equation (4.1) and since $\mathscr{U}(t,\cdot)\in C^{2,1,w}_{\mathrm{loc}}(\mathscr{P}_2(\mathbb{M}))$, we get that $\partial_t\mathscr{U}(t,\cdot)$ is differentiable for all $t\in(0,T)$. Therefore, $\partial_t\nabla_w\mathscr{U}(t,\mu)(q)=\nabla_w\partial_t\mathscr{U}(t,\mu)(q)$ for all $(t,\mu)\in(0,T)\times\mathscr{P}_2(\mathbb{M})$ and $q\in\mathrm{spt}(\mu)$.

Since

$$\partial_t\mathscr{U}(t,\sigma_s) + \mathscr{H}\big(\sigma_s, \nabla_w\mathscr{U}(t,\sigma_s)\big) = 0,$$

(4.4) and (4.5) imply

$$\begin{aligned}
&\int_{\mathbb{M}}\Big(\partial_t\nabla_w\mathscr{U}(t,\mu)(q) + D_qH\big(q,\nabla_w\mathscr{U}(t,\mu)(q)\big) - \nabla_w\mathscr{F}(\mu)(q)\Big)\cdot\xi(q)\mu(dq)\\
&\quad+\int_{\mathbb{M}} D_pH\big(q,\nabla_w\mathscr{U}(t,\mu)(q)\big)\cdot\Big(D_q\nabla_w\mathscr{U}(t,\mu)(q)\xi(q)\Big)\mu(dq)\\
&\quad+\int_{\mathbb{M}^2} D_pH\big(q,\nabla_w\mathscr{U}(t,\mu)(q)\big)\cdot\Big(\nabla^2_{ww}\mathscr{U}(t,\mu)(q,a)\xi(a)\mu(da)\Big)\mu(dq) = 0.
\end{aligned}\tag{4.6}$$

Since we asserted in Remark 3.19 that $D_q\nabla_w\mathscr{U}(t,\mu)(\cdot)$ is symmetric, (4.6) can be rewritten as

$$\begin{aligned}
&\int_{\mathbb{M}}\big[\partial_t\nabla_w\mathscr{U}(t,\mu)(q) + D_qH\big(q,\nabla_w\mathscr{U}(t,\mu)(q)\big) - \nabla_w\mathscr{F}(\mu)(q)\big]\cdot\xi(q)\mu(dq)\\
&\quad+\int_{\mathbb{M}} D_q\nabla_w\mathscr{U}(t,\mu)(q)D_pH\big(q,\nabla_w\mathscr{U}(t,\mu)(q)\big)\cdot\xi(q)\mu(dq)\\
&\quad+\int_{\mathbb{M}^2}\big(\nabla^2_{ww}\mathscr{U}(t,\mu)(q,a)^{\top}D_pH\big(q,\nabla_w\mathscr{U}(t,\mu)(q)\big)\big)\mu(dq)\cdot\xi(a)\mu(da) = 0.
\end{aligned}$$

Note that

$$\begin{aligned}
D_qH\big(\,\cdot\,,\nabla_w\mathscr{U}(t,\mu)\big) &+ D_q\nabla_w\mathscr{U}(t,\mu)D_pH\big(\,\cdot\,,\nabla_w\mathscr{U}(t,\mu)\big)\\
&= D_q\big(H\big(\,\cdot\,,\nabla_w\mathscr{U}(t,\mu)\big)\big)\in T_\mu\mathscr{P}_2(\mathbb{M}).
\end{aligned}$$

Since the rows of $\nabla^2_{ww}\mathscr{U}(t,\mu)(q,a)$ belong to $T_\mu\mathscr{P}_2(\mathbb{M})$, so does

$$\nabla^2_{ww}\mathscr{U}(t,\mu)(q,a)^{\top}D_pH\big(q,\nabla_w\mathscr{U}(t,\mu)(q)\big)$$

(as linear combinations of these rows). By the arbitrariness of ξ and the previous claims, we conclude

$$\begin{aligned}
\partial_t\nabla_w\mathscr{U}(t,\mu) + D_qH\big(\,\cdot\,,\nabla_w\mathscr{U}(t,\mu)\big) &+ D_q\nabla_w\mathscr{U}(t,\mu)D_pH\big(\,\cdot\,,\nabla_w\mathscr{U}(t,\mu)\big)\\
&+\bar{\mathscr{N}_\mu}\big[\mathscr{V},\nabla_w^T\mathscr{V}\big](t,\mu,\cdot) = \nabla_w\mathscr{F}(\mu),
\end{aligned}$$

μ-almost everywhere on $q\in\mathbb{M}$. □

Remark 4.5. At this point we do not know whether all the terms appearing in (4.2) could be extended to (at least $\mathscr{L}^d$-a.e.) $q \in \mathbb{M}$. We have good pointwise continuity properties of $\bar{\nabla}_{ww}^{\top}\mathscr{U}(t,\cdot)(\cdot,\cdot)$, but we do not know much about the continuity properties of $\nabla_{ww}^{\top}\mathscr{U}(t,\cdot)(\cdot,\cdot)$. If we knew

$$\bar{\mathscr{N}_\mu}\big[\mathscr{V}, \nabla_{ww}^{\top}\mathscr{U}\big](t,\mu,q) = \bar{\mathscr{N}_\mu}\big[\mathscr{V}, \bar{\nabla}_{ww}^{\top}\mathscr{U}\big](t,\mu,q)$$

we could deduce that $q \mapsto \bar{\mathscr{N}_\mu}\big[\mathscr{V}, \nabla_{ww}^{\top}\mathscr{U}\big](t,q,\mu)$ is continuous. In the same time, we do not know whether $\partial_t\mathscr{V}$ admits a continuous extension.

As a remark, despite the fact that $\mathscr{V}(t,\mu,\cdot)$ itself is defined only on $\mathrm{spt}(\mu)$, we know that it is Lipschitz-continuous there, uniformly with respect to t and μ. But it is not clear at all whether any Lipschitz-continuous extension of this at the same time would produce a valid extension for $\partial_t\mathscr{V}$ and $\nabla_w^{\top}\mathscr{V}$. As highlighted before, we revisit this question in Section 5.1, and in particular there we produce a solution to the vectorial master equation that is defined for (Lebesgue) a.e. $q \in \mathbb{M}$.

4.2 The scalar master equation

In this subsection we assume there exists a function C which assigns to each compact set $K \subset \mathbb{M}$ and each real number $r > 0$ a positive value $C(K,r)$. We assume to be given

(H9) $$u_0, f \in C^{1,1}_{\mathrm{loc}}(\mathbb{M} \times \mathscr{P}_2(\mathbb{M}))$$

such that

(H10) $$\begin{aligned} \nabla_w \mathscr{U}_0(\mu)(q) &= D_q u_0(q,\mu), \\ \nabla_w \mathscr{F}(\mu)(q) &= D_q f(q,\mu), \end{aligned} \quad \forall (q,\mu) \in \mathbb{M} \times \mathscr{P}_2(\mathbb{M}).$$

Since we can modify L or $\tilde{\tilde{\mathscr{F}}}$ as follows,

$$\tilde{\mathscr{L}}(x,a) = \int_\Omega \big(L(x(\omega),a(\omega)) - r|x(\omega)|^2\big)d\omega + \tilde{\tilde{\mathscr{F}}}(x) + r\|x\|^2,$$

we learn from Proposition B.6 that (H2) and (H7) imply that

(4.7) $$\begin{aligned} &\mathbb{M} \ni q \mapsto u_0(q,\mu) \text{ is convex and } \mathbb{M}\times\mathbb{R}^d \ni (q,v) \mapsto \\ &L(q,v) + f(q,\mu) \text{ is strictly convex } \forall \mu \in \mathscr{P}_2(\mathbb{M}). \end{aligned}$$

Let us remark that by the fact that $u_0, f \in C^{1,1}_{\mathrm{loc}}(\mathbb{M} \times \mathscr{P}_2(\mathbb{M}))$, we have that u_0 and f are locally bounded, i.e., $\forall K \subset \mathbb{M}$ compact and $r > 0$, $\exists C = C(K,r)$: $|u_0(q_0,\mu)|, |f(q_0,\mu)| \le C \ \forall (q_0,\mu) \in K \times \mathscr{B}_r$.

We are to find a function $u : [0,T] \times \mathbb{M} \times \mathscr{P}_2(\mathbb{M}) \to \mathbb{R}$ that satisfies the *scalar master equation*

(4.8) $$\begin{cases} \partial_t u(t,q,\mu) + H(q, D_q u(t,q,\mu)) \\ \quad + \mathscr{N}_\mu\big[D_q u(t,\cdot,\mu), \nabla_w u(t,q,\mu)(\cdot)\big] = f(q,\mu), \\ \quad \text{in } (0,T) \times \mathbb{M} \times \mathscr{P}_2(\mathbb{M}), \\ u(0,\cdot,\cdot) = u_0 \text{ in } \mathbb{M} \times \mathscr{P}_2(\mathbb{M}), \end{cases}$$

where the nonlocal operator $\mathscr{N}_\mu$ is defined as in (0.1). We define the notion of classical solution to (4.8) as follows.

DEFINITION 4.6. We say that u is a classical solution to (4.8) if the following holds. It is continuously differentiable on $(0,T)\times\mathbb{M}\times\mathscr{P}_2(\mathbb{M})$, continuous up to the initial time 0, and the PDE is satisfied pointwise. The vector field $\mathbb{M}\ni q\mapsto D_q u(t,q,\nu)$ is Lipschitz, uniformly with respect to $(t,\nu)\in[0,T]\times\mathscr{B}_r$ $(r>0)$.

Furthermore, for all $\nu\in\mathscr{P}_2(\mathbb{M})$ and for $\mathscr{L}^1\otimes\mathscr{L}^d$-a.e.

$$(s,q)\in(0,T)\times\mathbb{M},\quad D_q\nabla_w u(s,q,\nu)(\cdot),\quad \nabla_w D_q u(s,q,\nu)(\cdot)$$

exist, belong to $L^2(\nu)$, and satisfy additionally

$$\int_{\mathbb{M}}\Big(\big(D_q\nabla_w-\nabla_w D_q\big)u(s,q,\nu)(y)\Big)D_pH(y,D_qu(s,y,\nu))\nu(dy)=0. \tag{4.9}$$

Remark 4.7. The condition (4.9) in the previous definitions needs some comments. In Theorem 4.19 we will actually show existence of the $C^{1,1}_{\text{loc}}([0,T]\times\mathbb{M}\times\mathscr{P}_2(\mathbb{M}))$ solution to (4.8). Let us notice that for functions $w\in C^{1,1}_{\text{loc}}(\mathbb{M}\times\mathscr{P}_2(\mathbb{M}))$, $D_q\nabla_w w(q,\nu)(\cdot)$ is meaningful for all $\nu\in\mathscr{P}_2(\mathbb{M})$ and for a.e. $q\in\mathbb{M}$ (see Section 5.1). But since $D_q w$ is only Lipschitz-continuous with respect to the measure variable, $\nabla_w D_q w(q,\nu)(\cdot)$ might not be meaningful in general (since Rademacher-type theorems in $(\mathscr{P}_2(\mathbb{M}),W_2)$ are more subtle; cf. [26]). So the $C^{1,1}$ regularity in general is not enough to ensure (4.9).

Nevertheless, as the discussion in Section 5.1 shows, the solution that we construct for the master equation (4.8) naturally satisfies (4.9). This condition in particular will imply uniqueness of the solution as well.

For $m\in\mathbb{N}$, we define

$$u_0^{(m)},f^{(m)}:\mathbb{M}\times(\mathbb{M})^m\to\mathbb{R},\quad U_0^{(m)},F^{(m)}:(\mathbb{M})^m\to\mathbb{R}$$

as

$$\begin{aligned}u_0^{(m)}(y,q)&:=u_0\big(y,\mu_q^{(m)}\big),\quad & f^{(m)}(y,q)&:=f\big(y,\mu_q^{(m)}\big),\\ U_0^{(m)}(q)&:=\mathscr{U}_0\big(\mu_q^{(m)}\big),\quad & F^{(m)}(q)&:=\mathscr{F}\big(\mu_q^{(m)}\big),\end{aligned}$$

where for $q=(q_1,\dots,q_m)\in(\mathbb{M})^m$, $\mu_q^{(m)}$ is defined as in (1.1).

We impose the following hypotheses on $u_0^{(m)}$ and $f^{(m)}$.

(H11) $u_0^{(m)}(y,\cdot)$, $f^{(m)}(y,\cdot)$ satisfy Properties 2.2(1)(a) and 2.2(2), locally uniformly with respect to $y\in\mathbb{M}$.

(H12) $D_yu_0^{(m)}(y,\cdot)$, $D_yf^{(m)}(y,\cdot)$ satisfy Property 2.2(1)(a), locally uniformly with respect to $y\in\mathbb{M}$.

Notice that based on the previous assumptions, we have that $D_y u_0^{(m)}$ and $D_y f^{(m)}$ are locally uniformly bounded, i.e., $\forall r > 0$, $K \subset \mathbb{M}$ compact,

$$\exists C = C(K,r) : \left|D_y u_0^{(m)}(y,q)\right|,\ \left|D_y f^{(m)}(y,q)\right| \le C \text{ if } (y,q) \in K \times \mathbb{B}_r^m.$$

At the same time, by the assumption (H5), $D_q L$ and $\partial_y^a \partial_v^b L$ (for all a, b multi-indices with $|a| + |b| = 2$) are locally uniformly bounded.

We assume that there exists a constant $C > 0$ such that

(H13) $\quad \|\partial_q^a \partial_p^b H\|_{L^\infty(\mathbb{M}\times\mathbb{R}^d)} \le C \quad$ for a, b multi-indices with $|a| + |b| = 3$.

We also assume there exists a locally bounded continuous function $\theta : \mathscr{P}_2(\mathbb{M}) \to [0,\infty)$ such that

(H14)

$L(q,v) + f(q,\mu) \ge \lambda_1 |v|^2 - \theta(\mu)(|q|+1) \quad \forall (q,v) \in \mathbb{M}\times\mathbb{R}^d,\ \forall \mu \in \mathscr{P}_2(\mathbb{M})$.

Note that it suffices to impose that $f(\cdot,\mu)$ is convex to have that (H6) implies (H14).

Recall that Remark 1.1 (iii) ensures there exists a constant C such that

We assume that there exists $C > 0$ such that

(H15) $\quad \begin{aligned} &|D_q H(q,p)| \le C(1 + |q| + |p|) \text{ and} \\ &|D_q L(q,v)| \le C(1+|q|+|v|)\ \forall (q,p,v) \in \mathbb{M}\times\mathbb{R}^{2d}. \end{aligned}$

4.3 Examples of data functions

We pause for a moment to give examples of initial data $\mathscr{U}_0$ and u_0, which satisfy the standing assumptions of this manuscript. Similar examples can be constructed for $\mathscr{F}$ and f as well.

Let $\phi_0, \phi_1 : \mathbb{M} \to \mathbb{R}$ be smooth bounded functions with uniformly bounded derivatives up to order 3. For simplicity, we assume also that they are positive and ϕ_1 is even. Fix $\lambda > 0$ and let $\phi : \mathbb{M} \to \mathbb{R}$ be defined as $\phi(q) := \frac{\lambda}{2}|q|^2 + \phi_0(q)$ and assume λ is large enough such that $D^2\phi + D^2\phi_1 \ge 0$ on $\mathbb{M}$. Then, let us define $\mathscr{U}_0 : \mathscr{P}_2(\mathbb{M}) \to \mathbb{R}$ as

$$\mathscr{U}_0(\mu) := \int_{\mathbb{M}} \phi(q)\mu(dq) + \frac{1}{2}\int_{\mathbb{M}} \phi_1 * \mu(q)\mu(dq), \quad \widetilde{\mathscr{U}_0}(x) = \mathscr{U}_0\big(x_\sharp \mathscr{L}_\Omega^d\big),$$

$\forall \mu \in \mathscr{P}_2(\mathbb{M}), x \in \mathbb{H}$. Then $\widetilde{\mathscr{U}_0}$ fulfills the assumptions (H1) and (H2).

Set

$$u_0(q_0,\mu) = \phi(q_0) + (\phi_1 * \mu)(q_0).$$

For $q := (q_1,\ldots,q_m) \in \mathbb{M}^m$ and $q_0 \in \mathbb{M}$, we have

$$u_0^{(m)}(q_0,q) = \phi(q_0) + \sum_{i=1}^m \frac{1}{m}\phi_1(q_0 - q_i)$$

and

$$U_0^{(m)}(q) = \frac{1}{m}\sum_{i=1}^{m}\phi(q_i) + \frac{1}{2m^2}\sum_{i,j=1}^{m}\phi_1(q_i - q_j),$$

and so for $1 \le i \le m$,

$$D_{q_i} u_0^{(m)}(q_0, q) = \frac{1}{m} D\phi_1(q_0 - q_i) \ \text{ and } \ D^2_{q_0 q_i} u_0^{(m)}(q_0, q)) = \frac{1}{m} D^2\phi_1(y - x_i).$$

We have

$$D_{q_0} u_0^{(m)}(q_0, q) = D\phi(y) + \sum_{i=1}^{m}\frac{1}{m} D\phi_1(q_0 - q_i).$$

>From these computations, one can easily verify that (H9) through (H12) are satisfied.

Under appropriate conditions on functions L_0, l, and g, Lagrangians of the form

$$L(q, v) := L_0(v) + l(q, v) + g(q)$$

and Hamiltonians defined as $H(q, \cdot) := L^*(q, \cdot)$ satisfy (H3) through (H7) and (H13) through (H15).

We are ready now to define the candidate for the solution to the scalar master equation. Given $t \in [0, T]$, $q \in \mathbb{M}$, and $\mu \in \mathscr{P}_2(\mathbb{M})$ we define

$$u(t, q, \mu) := \inf_{\gamma}\Big\{ u_0(\gamma_0, \sigma_0^t[\mu]) + \int_0^t \big(L(\gamma_s, \dot{\gamma}_s) + f(\gamma_s, \sigma_s^t[\mu])\big) ds : \gamma \in W^{1,2}([0, t], \mathbb{M}), \gamma_t = q \Big\}. \tag{4.10}$$

Here the curve $(\sigma_s^t[\mu])_{s\in[0,t]}$ is defined in (C.3). Define

$$M_*(r) := \sup_{B_r(0)\times\mathscr{B}_{e_T(r)}} |\theta| + |u_0| + T(|f| + |L(0, \cdot)|), \quad c_*(r) := \sup_{\bar{B}_1(0)\times\bar{\mathscr{B}}_r} |u_0|.$$

Remark 4.8. Let $r > 0$.

(i) As $u_0(\cdot, \nu)$ is convex, if $D_q u(0, \nu) \neq 0$, then

$$u_0\left(\frac{D_q u(0, \nu)}{|D_q u(0, \nu)|}, \nu\right) \ge u_0(0, \nu) + \frac{D_q u(0, \nu)}{|D_q u(0, \nu)|} \cdot D_q u(0, \nu) = u_0(0, \nu) + \frac{|D_q u(0, \nu)|^2}{|D_q u(0, \nu)|}.$$

Thus, if $\nu \in \mathscr{B}_r$, we conclude that

$$|D_q u(0, \nu)| \le 2c_*(r).$$

Clearly, the previous inequality still holds when $D_q u(0, \nu) = 0$. Consequently,

$$u_0(q, \nu) \ge u_0(0, \nu) + D_q u(0, \nu) \cdot q \ge -c_*(r)(1 + |q|).$$

(ii) Suppose $(t, q, \mu) \in [0, T] \times B_r(0) \times \mathscr{B}_r$. Then

$$u(t, q, \mu) \leq M_*(r),$$

and so, if γ is the unique minimizer in (4.10), we use (H14) and Remark C.6 (ii) to obtain

$$\begin{aligned} M_*(r) &\geq u(t, q, \mu) \\ &\geq -c_*\big(e_T(r)\big)(1 + |\gamma(0)|) - M_*(r)T - M_*(r)\int_0^t |\gamma| ds \\ &\quad + \lambda_1 \int_0^t |\dot{\gamma}|^2 \, ds. \end{aligned}$$

We conclude there exists a constant $\bar{M}(r)$ independent of t such that

$$\int_0^t |\dot{\gamma}|^2 ds \leq \bar{M}(r).$$

Hence,

$$|\gamma_{\tau_1} - \gamma_{\tau_2}|^2 \leq \bar{M}(r)|\tau_2 - \tau_1| \quad \text{if } 0 \leq \tau_1 \leq \tau_2 \leq t. \tag{4.11}$$

(iii) By (ii), there is a constant $M^*(r)$ such that

$$|u(t, q, \mu)| \leq M^*(r) \quad (t, q, \mu) \in [0, T] \times B_r(0) \times \mathscr{B}_r$$

Since

$$(q, v) \mapsto L_{s,t}(q, v) := L(q, v) + f(q, \sigma_s^t[\mu]), \quad q \mapsto u_0(q, \sigma_0^t[\mu]),$$

are convex, we obtain that $u(t, \cdot, \mu)$ is a convex function and so as argued above,

$$|D_q u(t, q, \mu)| \leq u\left(t, q + \frac{D_q u(t, q, \mu)}{|D_q u(t, q, \mu)|}\right) - u(t, q, \mu) \leq M^*(r) + M^*(r + 1).$$

LEMMA 4.9. *Let* $(t, q, \mu) \in [0, T] \times B_r(0) \times \mathscr{B}_r$ *and let* $\gamma : [0, t] \to \mathbb{M}$ *be the unique optimizer in* (4.10). *Suppose that the assumptions* (H4), (H5), (H6), (H10), *and* (H15) *take place. Then* $\gamma \in C^{1,1}([0, t])$.

PROOF. The proof follows the same lines as the one of [12, theorem 6.2.5]. □

PROPOSITION 4.10. *Let* $\mu \in \mathscr{P}_2(\mathbb{M})$ *and* $t \in [0, T]$. *Recall* $[0, t] \ni s \mapsto \sigma_s^t[\mu]$ *is defined in* (C.3) *in Lemma* C.5.

(i) *We have* $u(t, \cdot, \mu) \in C^{1,1}_{\text{loc}}(\mathbb{M})$. *Furthermore, there exists a unique* γ *minimizer in* (4.10) *which we denote as* $s \mapsto S_s^t[\mu](q)$.
(ii) *If* $\omega \in \Omega$, $x \in \mathbb{H}$, $\mu = \sharp(x)$, *and* $q = x(\omega)$ *(meaning in particular that* $q \in \text{spt}(\mu)$*), then* $\widetilde{S}_s^t[x](\omega) = S_s^t[\mu](q)$.
(iii) *Under the assumptions in* (ii) *we have* $D_q u(t, q, \mu) = \nabla_w \mathscr{U}(t, \mu)(q)$.
(iv) $[0, t] \ni s \mapsto D_q u(s, S_s^t[\mu](q), \sigma_s^t[\mu])$ *is Lipschitz-continuous, for all* $(q, \mu) \in \mathbb{M} \times \mathscr{P}_2(\mathbb{M})$.

(v) *We have that* $u(\cdot,\cdot,\mu) \in C^{0,1}_{\mathrm{loc}}([0,T]\times\mathbb{M})$*, with Lipschitz constants depending on* $r>0$*, where* $\mu\in\mathscr{B}_r$.

PROOF. By Remark 4.8(iii), $u(t,\cdot,\mu)$ is a convex function. The fact that $u(t,\cdot,\mu)$ is locally semiconcave is a standard property. Thus, $u(t,\cdot,\mu)$ is $C^{1,1}_{\mathrm{loc}}(\mathbb{M})$. Since the action

$$\gamma \mapsto A_t[\gamma] := u_0(\gamma_0, \sigma_0^t[\mu]) + \int_0^t L_{s,t}(\gamma_s,\dot\gamma_s)ds$$

is strictly convex, $S_s^t[\mu](q)$ is uniquely defined.

(ii) By the convexity of A_t, any critical point of A_t on the set $\{\gamma\in C^1([0,t],\mathbb{M}) : \gamma_t = q\}$ is a minimizer. Set

$$p_s := P_s^t[\mu](q).$$

The Hamiltonian associated to $L_{s,t}$ is $H_{s,t}(q,p) := H(q,p) - f(q,\sigma_s^t[\mu])$. Since

$$D_pH_{s,t}(q,p) \equiv D_pH(q,p),$$

in light of Proposition C.2(iv) we have

$$D_pH_{s,t}(\gamma_s,p_s) = D_pH\big(\widetilde S_s^t[x](\omega), \widetilde P_s^t[x](\omega)\big) = \partial_s\widetilde S_s^t[x](\omega) = \dot\gamma_s. \tag{4.12}$$

By (H10)

$$D_qH_{s,t}(q,p) = D_qH(q,p) - D_qf(q,\sigma_s^t[\mu]) = D_qH(q,p) - \nabla_w\mathscr{F}(\sigma_s^t[\mu])(q).$$

Thus, by Remark 3.7

$$\begin{aligned} D_qH_{s,t}(\gamma_s,p_s) &= D_qH\big(\widetilde S_s^t[x](\omega), \widetilde P_s^t[x](\omega)\big) - \nabla\widetilde{\mathscr{F}}(\widetilde S_s^t[\mu])(\omega) \\ &= -\partial_s\widetilde P_s^t[x](\omega) = -\dot p_s. \end{aligned} \tag{4.13}$$

We use first (H10), second Remark 3.7, and third the last identity in (1.26) to obtain

$$D_qu_0(\gamma_0,\sigma_0^t[\mu]) = \nabla_w\mathscr{U}_0\big(\sigma_0^t[\mu]\big)(\gamma_0) = \nabla\widetilde{\mathscr{U}_0}(\widetilde S_0^t[\mu])(\omega) = \widetilde P_0^t[x](\omega)) = p_0.$$

This, together with (4.12) and (4.13), implies γ is a critical point of A_t on the set

$$\{\gamma\in C^1([0,t],\mathbb{M}) : \gamma_t = q\}.$$

Hence, γ is the unique minimizer, which verifies (ii).

(iii) By the optimality property of γ, the standard Hamilton–Jacobi theory ensures that

$$\dot\gamma_s = D_pH(\gamma_s, D_qu(s,\gamma_s,\sigma_s^t[\mu])) \quad \forall s\in(0,t). \tag{4.14}$$

First, by the strict convexity of H in the second variable, we have that

$$D_qu(s,\gamma_s,\sigma_s^t[\mu]) = D_vL(\gamma_s,\dot\gamma_s) \quad \forall s\in(0,t),$$

from where, by Lemma 4.9 and by the regularity of D_vL, one obtains that $[0,t]\ni s\mapsto D_qu(s,\gamma_s,\sigma_s^t[\mu])$ is Lipschitz-continuous. This shows (iv).

Then, by Proposition C.2 (iv),

$$\dot\gamma_s = D_pH\big(\gamma_s, \nabla_w\mathscr{U}(s,\sigma_s^t[\mu])(\gamma_s)\big),$$

which, together with (4.14), implies

$$D_p H\big(\gamma_s, \nabla_w \mathscr{U}(s, \sigma_s^t[\mu])(\gamma_s)\big) = D_p H\big(\gamma_s, D_q u(s, \gamma_s, \sigma_s^t[\mu])\big) \quad \forall s \in (0,t).$$

Thus, by (H4), one has

$$\nabla_w \mathscr{U}(s, \sigma_s^t[\mu])(\gamma_s) = D_q u(s, \gamma_s, \sigma_s^t[\mu]) \quad \forall s \in (0,t).$$

Letting s increase to t we verify (iii).

(v) What remains to be shown is the Lipschitz regularity of u with respect to the variable t. But this follows from the dynamic programming principle and from the time Lipschitz continuity of $(\gamma_s)_{s\in[0,t]}$ and $(\sigma_s^t[\mu])_{s\in[0,t]}$ (see Lemma C.7(ii) and Lemma 4.9). □

Remark 4.11. (i) Let $\mu \in \mathscr{P}_2(\mathbb{M})$, $t \in [0,T]$. Note that in Proposition 4.10 $S_s^t[\mu]$ is defined on the whole set $\mathbb{M}$ and not just on the support of μ. When $x \in \mathbb{H}$ is such that $\mu = \sharp(x)$, Proposition 4.10 (ii) reads as

$$\widetilde{S}_s^t[x] = S_s^t[\mu] \circ x.$$

Also,

$$\begin{cases} \partial_s S_s^t[\mu] = D_p H(S_s^t[\mu], \nabla_w \mathscr{U}(s, \sigma_s^t[\mu])(S_s^t[\mu])), & s \in (0,t), \\ S_t^t[\mu] = \mathrm{id}. \end{cases} \tag{4.15}$$

(ii) It is very important to underline the fact that by Proposition 4.10(iii) we have that for all $(t,\mu) \in (0,T) \times \mathscr{P}_2(\mathbb{M})$, $D_q u(t,\cdot,\mu) = \nabla_w \mathscr{U}(t,\mu)(\cdot)$ on $\mathrm{spt}(\mu)$. Since $D_q u(t,\cdot,\mu)$ is defined on the whole $\mathbb{M}$ (and we will see below that it is locally Lipschitz-continuous), this produces a very natural extension for $\nabla_w \mathscr{U}(t,\mu)(\cdot)$ to the whole $\mathbb{M}$. This observation will also help us to improve the previous notion of weak solution to the vectorial master equation, as we will see in Section 5.1.

(iii) Since $\mathscr{U}$ is of class $C^{1,1}_{\mathrm{loc}}$ (cf. Definition 3.8) [16, cor. 3.38] yields the existence of a Lipschitz-continuous extension of $\nabla_w \mathscr{U}(t,\mu)(\cdot)$ to the whole $\mathbb{M}$, with a Lipschitz constant independent of μ. This extension has the property that it is continuous at (μ, q) for $q \in \mathrm{spt}(\mu)$. Our result, as described above, because of the local Lipschitz continuity of $D_q u$ (cf. Lemma 4.13) provides a slightly better extension.

PROPOSITION 4.12. *For all $t \in [0,T]$ and $q \in \mathbb{M}$, the function $u(t,q,\cdot)$ is continuous on $\mathscr{P}_2(\mathbb{M})$.*

We skip the proof of this proposition since it is obtained by standard arguments, similar to those appearing in the proof of Proposition C.1.

LEMMA 4.13. *When* (H1)–(H15) *hold, then u defined in* (4.10) *is of class*

$$C^{1,1}_{\mathrm{loc}}([0,T] \times \mathbb{M} \times \mathscr{P}_2(\mathbb{M})).$$

PROOF. We proceed by a discretization approach. Let $\mu \in \mathscr{P}_2(\mathbb{M})$, $t > 0$, $m \in \mathbb{N}$, and $q_0 \in \operatorname{spt}(\mu)$ be fixed. Moreover, given $\{q_1, \dots, q_m\} \subset \operatorname{spt}(\mu)$ we shall use the notation of $q = (q_1, \dots, q_m) \in (\mathbb{M})^m$. We define

$$\mu_q^{(m+1)} = \frac{1}{m+1} \sum_{i=0}^{m} \delta_{q_i}, \quad \sigma_s^{(m+1)} := \sigma_s^t\big[\mu_q^{(m+1)}\big],$$

so that $\sigma^{(m+1)}$ is the solution to the continuity equation (C.4) with $\mu_q^{(m+1)}$ as terminal condition. Note

$$\sigma_s^{(m+1)} = \frac{1}{m+1} \sum_{i=0}^{m} \delta_{S_s^t[\mu_q^{(m+1)}](q_i)} \quad \forall s \in (0,t).$$

We define

$$u_0^{(m+1)}, f^{(m+1)} : \mathbb{M} \times (\mathbb{M})^{(m+1)} \to \mathbb{R}, \;\; U^{(m+1)}, u^{(m)} : (0,T) \times (\mathbb{M})^{(m+1)} \to \mathbb{R}$$

as

$$u_0^{(m+1)}(y_0, q_0, q) := u_0\big(y_0, \mu_q^{(m+1)}\big), \quad f^{(m+1)}(y_0, q_0, q) := f\big(y_0, \mu_q^{(m+1)}\big),$$

and

$$U^{(m+1)}(s, q_0, q) := \mathscr{U}(s, \mu_q^{(m+1)}), \quad u^{(m)}(t, q_0, q) := u(t, q_0, \mu_q^{(m+1)}). \tag{4.16}$$

Observe

$$\begin{aligned} u^{(m)}(t, q_0, q) &= u_0\Big(Q_0(0, q_0, q), \sigma_0^{(m+1)}\Big) \\ &+ \int_0^t L\Big(Q_0(s, q_0, q), D_p H\Big(Q_0(s, q_0, q), \nabla_w \mathscr{U}\big(s, \sigma_s^{(m+1)}\big)\big(Q_0(s, q_0, q)\big)\Big)\Big) ds \\ &+ \int_0^t f\Big(Q_0(s, q_0, q), \sigma_s^{(m+1)}\Big) ds \\ &+ u_0^{(m+1)}\Big(Q_0(0, q_0, q), Q_0(0, q_0, q), Q(0, q_0, q)\Big)\Big) \\ &+ \int_0^t L\Big(Q_0(s, q_0, q), D_p H(Q_0(s, q_0, q), (m+1) D_{q_0} U^{(m+1)}(s, Q_0(s, q_0, q), Q(s, q_0, q)\Big) ds \\ &+ \int_0^t f^{(m+1)}\Big(Q_0(s, q_0, q), Q_0(s, q_0, q), Q(s, q_0, q)\Big) ds \end{aligned} \tag{4.17}$$

where we have set

$$\begin{aligned} Q_i(s, q_0, q) &:= S_s^t\big[\mu_q^{(m+1)}\big](q_i), \\ Q(s, q_0, q) &:= (Q_1(s, q_0, q), \dots, Q_m(s, q_0, q)). \end{aligned} \tag{4.18}$$

Now our first goal is to obtain derivative estimates on $u^{(m)}$ with respect to the 'distinguished' variable q_0 and second, with respect to all the other variables q. Finally, we also derive the necessary estimates involving the time variable t as well. It is convenient to introduce the notation

$$\tilde{u}_0^{(m+1)}, \tilde{f}^{(m+1)}, V^{(m+1)} : \mathbb{M} \times (\mathbb{M})^m \to \mathbb{R}$$

defined as

(4.19)
$$\begin{aligned}\widetilde{u}_0^{(m+1)}(q_0,q) &:= u_0^{(m+1)}(Q_0(0,q_0,q), Q_0(0,q_0,q), Q(0,q_0,q)),\\ \widetilde{f}^{(m+1)}(q_0,q) &:= \int_0^t f(Q_0(s,q_0,q), Q_0(s,q_0,q), Q(s,q_0,q))ds\\ V^{(m+1)}(q_0,q) &:= \int_0^t L\big(Q_0(s,q_0,q), D_pH(Q_0(s,q_0,q),\\ &\qquad (m+1)\nabla_{q_0}U^{(m+1)}(s, Q_0(s,q_0,q), Q(s,q_0,q)))\big)ds.\end{aligned}$$

In Lemma 4.15 and Lemma 4.18 below we establish the necessary derivative estimates on these new quantities. These imply in particular that there exists a constant $C = C(T,r,K) > 0$ such that for any $(q_0,q) \in \mathbb{B}_r^{(m+1)}$, $q_0 \in K$ (where $K \subset \mathbb{M}$ is compact), and for all $t \in [0,T]$ and $i, j \in \{0,\dots,m\}$, we have

$$\left|D_{q_i}u^{(m)}(t,q_0,q)\right| \leq \begin{cases} C, & i = 0,\\ \dfrac{C}{m+1}, & i > 0,\end{cases} \tag{4.20}$$

(4.21)
$$|D^2_{q_iq_j}u^{(m)}(t,q_0,q)|_\infty \leq \begin{cases} C, & i = j = 0,\\ \dfrac{C}{m+1}, & (i = j \text{ and } i > 0) \text{ or } (i\cdot j = 0,\ \max\{i,j\} > 0),\\ \dfrac{C}{(m+1)^2}, & i \neq j,\ i,j > 0.\end{cases}$$

and

$$|D_{q_0}\partial_t u^{(m)}(t,q_0,q)| \leq C, \quad \sum_{k=1}^m (m+1)|D_{q_k}\partial_t u^{(m)}|^2 \leq C, \tag{4.22}$$

and

$$\left|\partial_t u^{(m)}(t,q_0,q)\right| \leq C, \ \ \left|\partial^2_{tt}u^{(m)}(t,q_0,q)\right| \leq C. \tag{4.23}$$

Let us notice that by definition and the assumption (H10), u is bounded on $[0,T]\times K\times\mathscr{B}_r$ for any $K \subseteq \mathbb{M}$ compact and $r > 0$. Therefore, $u^{(m)}$ is uniformly bounded (with respect to m) on $[0,T]\times K\times\mathbb{B}_r^m$.

Now, all these properties allow us to verify the assumptions of Corollary 3.17 and conclude by this that there exists $\widetilde{u} : [0,T]\times\mathbb{M}\times\mathscr{P}_2(\mathbb{M}) \to \mathbb{R}$ such that after passing to a suitable subsequence $(u^{(m)})_{m\in\mathbb{N}}$ converges to $\widetilde{u}$ in the sense as described in Corollary 3.17. Let us notice furthermore that $\widetilde{u}(t,q_0,\mu)$ has to be the limit of $u(t,q_0,\mu_q^{(m+1)})$ (since by Proposition 4.12 $u(t,q_0,\cdot)$ is continuous) and therefore $\widetilde{u}$ and u must coincide. Thus, as a consequence of Corollary 3.17, $u \in C^{1,1}_{\text{loc}}([0,T]\times\mathbb{M}\times\mathscr{P}_2(\mathbb{M}))$. □

COROLLARY 4.14. *Under the assumptions of Lemma 4.13, we have that the vector field $\mathbb{M} \ni q \mapsto D_q u(t,q,\mu)$ is globally Lipschitz, uniformly with respect to $(t,\mu) \in [0,T] \times \mathscr{B}_r$ for any $r > 0$.*

PROOF. Let $r > 0$, $t \in [0,T]$, and $\mu \in \mathscr{B}_r$. Let $q_1, q_2 \in \mathbb{M}$. Let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence in $\mathscr{B}_r$ such that $W_2(\mu_n,\mu) \to 0$ as $n \to +\infty$ and $\operatorname{spt}(\mu_n) = \mathbb{M}$ for all $n \in \mathbb{N}$. By Proposition 4.10(iii) we have $D_q u(t,q_i,\mu_n) = \nabla_w \mathscr{U}(t,\mu_n)(q_i)$, $i = 1,2$. In light of Proposition 1.3 and Lemma 3.11 there exists $C = C(r,T) > 0$ independent of n such that

$$\begin{aligned}|D_q u(t,q_1,\mu_n) - D_q u(t,q_2,\mu_n)| &= |\nabla_w \mathscr{U}(t,\mu_n)(q_1) - \nabla_w \mathscr{U}(t,\mu_n)(q_2)| \\ &\le C|q_1 - q_2|.\end{aligned}$$

By the continuity of $D_q u(t,q_i,\cdot)$ provided in Lemma 4.13, one can pass to the limit with $n \to +\infty$ to obtain

$$|D_q u(t,q_1,\mu_n) - D_q u(t,q_2,\mu_n)| \le C|q_1 - q_2|.$$

The result follows. □

LEMMA 4.15. *Let $\widetilde{u}_0^{(m+1)}$, $\widetilde{f}^{(m+1)}$, and $V^{(m+1)}$ be defined as in (4.19) and suppose the assumptions of Lemma 4.13 are fulfilled. Then, for $T, r > 0$ and $K \subset \mathbb{M}$ compact, there exists a constant $C = C(T,r,K) > 0$ such that for any $(q_0,q) \in \mathbb{B}_r^{(m+1)}$ with $q_0 \in K$ and $i,j \in \{0,\dots,m\}$, we have*

(1)

$$\left|D_{q_i}\widetilde{u}_0^{(m+1)}(q_0,q)\right| \le \begin{cases} C, & i = 0,\\ \dfrac{C}{m+1}, & i > 0,\end{cases}$$

and

$$\left|D_{q_i}\widetilde{f}^{(m+1)}(q_0,q)\right| \le \begin{cases} C, & i = 0,\\ \dfrac{C}{m+1}, & i > 0.\end{cases}$$

(2)

$$\begin{aligned}&\left|D^2_{q_iq_j}\widetilde{u}_0^{(m+1)}(q_0,q)\right|_\infty \\ &\quad\le \begin{cases} C, & i = j = 0,\\ \dfrac{C}{m+1}, & (i = j \text{ and } i > 0) \text{ or } (i\cdot j = 0 \text{ and } \max\{i,j\} > 0),\\ \dfrac{C}{(m+1)^2}, & i \ne j,\ i,j > 0,\end{cases}\end{aligned}$$

and

$$
|D^2_{q_iq_j}\widetilde{f}^{(m+1)}(q_0,q)|_\infty \le \begin{cases} C, & i=j=0,\\ \dfrac{C}{m+1}, & (i=j \text{ and } i>0) \text{ or } (i\cdot j=0 \text{ and } \max\{i,j\}>0),\\ \dfrac{C}{(m+1)^2}, & i\neq j,\ i,j>0. \end{cases}
$$

(3)

$$
\left|D_{q_i}V^{(m+1)}(q_0,q)\right| \le \begin{cases} C, & \text{if } i=0,\\ \dfrac{C}{m+1}, & \text{if } i>0. \end{cases}
$$

(4)

$$
\left|D^2_{q_iq_j}V^{(m+1)}(q_0,q)\right|_\infty \le \begin{cases} C, & i=j=0,\\ \dfrac{C}{m+1}, & (i=j \text{ and } i>0) \text{ or } (i\cdot j=0 \text{ and } \max\{i,j\}>0),\\ \dfrac{C}{(m+1)^2}, & i\neq j. \end{cases}
$$

As a consequence, $u^{(m)}$ *defined in* (4.17) *satisfied the estimates* (4.20) *and* (4.21) *from Lemma* 4.13.

PROOF. As the computations to obtain the corresponding estimates in the case of $\widetilde{u}_0^{(m+1)}$ and $\widetilde{f}^{(m+1)}$ are completely parallel, we perform these only in the case of $\widetilde{u}_0^{(m+1)}$.

(1) In the computations below, to facilitate the reading, we will display neither the time nor the space variables in Q_i. For $i \ge 0$, we have

$$
\begin{aligned}
&D_{q_i}\widetilde{u}_0^{(m+1)}(q_0,q)\\
&= D_y u_0^{(m+1)}(Q_0,Q_0,Q)D_{q_i}Q_0 + D_{q_i}u_0^{(m+1)}(Q_0,Q_0,Q)D_{q_i}Q_i\\
&\quad + \sum_{k=0,k\neq i}^{m} D_{q_k}u_0^{(m+1)}(Q_0,Q_0,Q)D_{q_i}Q_k.
\end{aligned} \tag{4.24}
$$

Now, let us recall that by assumption (H10) we have

$$
D_y u_0(y,\mu) = \nabla_w \mathscr{U}_0(\mu)(y), \quad u_0^{(m+1)}(y,q_0,q_1,\dots,q_m) = u_0(y,\mu_q^{(m+1)}),
$$

for all $\mu \in \mathscr{P}_2(\mathbb{M})$, all $y \in \operatorname{spt}(\mu)$, and all $q_0,q_1,\dots,q_m \in \mathbb{M}$. This implies

$$
D_y u_0^{(m+1)}(y,q_0,q) = D_y u_0(y,\mu_q^{(m+1)}) = \nabla_w \mathscr{U}_0(\mu_q^{(m+1)})(y),
$$

and so

$$D_y u_0^{(m+1)}(q_i, q_0, q) = D_y u_0\big(q_i, \mu_q^{(m+1)}\big) = \nabla_w \mathscr{U}_0(\mu_q^{(m+1)})(q_i) = (m+1) D_{q_i} U_0^{(m+1)}(q_0, q) \tag{4.25}$$

for all $i \in \{0, \dots, m\}$.

Let us notice that by (H11)–(H12), Lemma 4.16, and Lemma 4.17 provide precise regularity estimates on the discrete flow $(Q_i)_{i=0}^m$), with a positive constant $C = C(T, r, K)$ such that

$$(m+1)|D_{q_k} u_0^{(m+1)}(Q_0, Q_0, Q_1, \dots, Q_m)| \le C,$$

and

$$|D_y u_0^{(m+1)}(Q_0, Q_0, Q_1, \dots, Q_m)| \le C.$$

so (1) follows by combining the previous arguments with Lemma 4.16.

(2) Differentiating (4.24) with respect to q_j one obtains

$$\begin{aligned} D^2_{q_i q_j} \widetilde{u}_0^{(m+1)}(q_0, q) = {} & D_{q_j} Q_0 D^2_{yy} u_0^{(m+1)}(Q_0, Q_0, Q) D_{q_i} Q_0 \\ & + \sum_{k=0}^m D_{q_j} Q_k D^2_{y q_k} u_0^{(m+1)}(Q_0, Q_0, Q) D_{q_i} Q_0 \\ & + D_y u_0^{(m+1)}(Q_0, Q_0, Q) D^2_{q_i q_j} Q_0 \\ & + \sum_{k,l=0}^m D_{q_j} Q_l D^2_{q_k q_l} u_0^{(m+1)}(Q_0, Q_0, Q) D_{q_i} Q_k \\ & + \sum_{k=0}^m D_{q_k} u_0^{(m+1)}(Q_0, Q_0, Q) D^2_{q_i q_j} Q_k. \end{aligned}$$

>From (4.25) we observe again for any $i \in \{0, \dots, m\}$,

$$\begin{aligned} D^2_{yy} u_0^{(m+1)}(q_i, q_0, q) &= D^2_{yy} u_0\big(q_i, \mu_q^{(m+1)}\big) = D_y \nabla_w \mathscr{U}_0\big(\mu_q^{(m+1)}\big)(q_i) \\ &= (m+1) D^2_{q_i q_i} U_0^{(m+1)}(q_0, q). \end{aligned}$$

Thus, if $i, j > 0$ and $i \neq j$,

$$\begin{aligned}
&\left|D^2_{q_i q_j}\widetilde{u}_0^{(m+1)}(q_0,q)\right|_\infty \\
&\quad \leq \frac{C}{m+1}(m+1)\left|D^2_{q_0q_0}U_0^{(m+1)}(Q_0,Q)\right|_\infty \frac{C}{m+1} \\
&\qquad + \sum_{k=0}^{m} |D_{q_j}Q_k|_\infty |D^2_{yq_k}u_0^{(m+1)}(Q_0,Q_0,Q)|_\infty |D_{q_i}Q_0|_\infty \\
&\qquad + (m+1)|D_{q_0}U_0^{(m+1)}(Q_0,Q)|\frac{C}{(m+1)^2} \\
&\qquad + \sum_{k=0}^{m} |D_{q_j}Q_k|_\infty |D^2_{q_kq_k}u_0^{(m+1)}(Q_0,Q_0,Q)|_\infty |D_{q_i}Q_k|_\infty \\
&\qquad + \sum_{k\neq l}^{m} |D_{q_j}Q_l|_\infty \left|D^2_{q_kq_l}u_0^{(m+1)}(Q_0,Q_0,Q)\right|_\infty |D_{q_i}Q_k|_\infty \\
&\qquad + \sum_{k=0}^{m} \left|D_{q_k}u_0^{(m+1)}(Q_0,Q_0,Q)\right|\left|D^2_{q_iq_j}Q_k\right|_\infty.
\end{aligned}$$

Let us recall that by our assumptions, there exists $C = C(T,r,K)$ such that

$$\left|D^2_{q_0q_0}U_0^{(m+1)}(Q_0,Q)\right|_\infty \leq \frac{C}{m+1}, \quad \left|D^2_{yq_k}u_0^{(m+1)}(Q_0,Q_0,Q)\right|_\infty \leq \frac{C}{m+1},$$

$$\left|D^2_{q_kq_l}u_0^{(m+1)}(Q_0,Q_0,Q)\right|_\infty \leq \begin{cases} \frac{C}{m+1}, & k=l, \\ \frac{C}{(m+1)^2}, & k \neq l, \end{cases}$$

$$\left|D_{q_k}u_0^{(m+1)}(Q_0,Q_0,Q)\right| \leq \frac{C}{m+1}$$

and by Lemma 4.17 and by the assumptions on $U_0^{(m+1)}$,

$$\left|D_{q_0}U_0^{(m+1)}(Q_0,Q)\right| \leq \frac{C}{m+1}.$$

Therefore, combining the previous arguments and computations, we conclude that

$$\left|D^2_{q_iq_j}\widetilde{u}_0^{(m+1)}(q_0,q)\right|_\infty \leq \frac{C}{(m+1)^2}.$$

Similar arguments yield that if $i = j$, we have

$$\left|D^2_{q_iq_i}\widetilde{u}_0^{(m+1)}(q_0,q)\right|_\infty \leq \frac{C}{m+1}.$$

The computations and arguments given above yield that

$$\left|D^2_{q_0q_0}\widetilde{u}_0^{(m+1)}(q_0,q)\right|_\infty \leq C \ \text{ and } \ \left|D^2_{q_0q_k}\widetilde{u}_0^{(m+1)}(q_0,q)\right|_\infty \leq \frac{C}{m+1} \quad \text{if } k > 0,$$

and so the thesis of the claim follows.

(3) Let us set $v_0 := D_p H(Q_0, (m+1)\nabla_{x_0} U^{(m+1)}(s, Q_0, Q))$. First, we have

$$
\begin{aligned}
D_{q_i} v_0 = {} & D^2_{pq} H(Q_0, (m+1)\nabla_{x_0} U^{(m+1)}(s, Q_0, Q)) D_{q_i} Q_0 \\
& + D^2_{pp} H(Q_0, (m+1)\nabla_{q_0} U^{(m+1)}(s, Q_0, Q))(m+1) \\
& \cdot \sum_{k=0}^{m} D^2_{q_0 q_k} U^{(m+1)}(s, Q_0, Q)) D_{q_i} Q_k,
\end{aligned}
\tag{4.26}
$$

by using the assumptions (H3) and (H5) on H, Lemma 4.16, and the properties of $D^2_{x_0 x_k} U^{(m+1)}$, we obtain

$$
\begin{aligned}
|D_{q_i} v_0|_\infty \le {} & \frac{C}{m+1} + \frac{C}{m+1} \\
& + (m+1) \sum_{k=1}^{m} \big| D^2_{q_0 q_k} U^{(m+1)}(s, Q_0, \dots, Q_m)) \big|_\infty \big| D_{q_i} Q_k \big|_\infty \\
\le {} & \frac{C}{m+1} \quad \text{if } i > 0.
\end{aligned}
$$

The same computation and arguments yield that $|D_{q_0} v_0|_\infty \le C$.

Now, we compute

$$
\begin{aligned}
& D_{q_i} V^{(m+1)}(q_0, q) \\
& \qquad = \int_0^t \big(D_y L(Q_0, v_0) D_{q_i} Q_0 + D_v L(Q_0, v_0) D_{q_i} v_0 \big) ds.
\end{aligned}
\tag{4.27}
$$

Using the smoothness property and the assumptions (H3) and (H5) on L, together with Lemma 4.17, we have that there exists a positive constant $C = C(T, r, K)$ such that $|Q_0(s, \cdot)| \le C$ and $|\dot{Q}_0(s, \cdot)| \le C$ for all $s \in (0, t)$, and so

$$
|D_y L(Q_0, v_0)| \le C \quad \text{and} \quad |D_v L(Q_0, v_0)| \le C.
$$

Therefore, by combining all the previous arguments, the thesis of the claim follows.

(4) From (4.27) one obtains

$$
\begin{aligned}
& D^2_{q_i q_j} V^{(m+1)}(q_0, q) \\
& \quad = \int_0^t \big(D_{q_j} Q_0 D^2_{yy} L(Q_0, v_0) D_{q_i} Q_0 + D_{q_j} v_0 D^2_{yv} L(Q_0, v_0) D_{q_i} Q_0 + D_y L(Q_0, v_0) D^2_{q_i q_j} Q_0 \big) ds \\
& \qquad + \int_0^t \big(D_{q_j} Q_0 D^2_{vy} L(Q_0, v_0) D_{x_i} v_0 + D_{q_j} v_0 D^2_{vv} L(Q_0, v_0) D_{q_i} v_0 + D_v L(Q_0, v_0) D^2_{q_i q_j} v_0 \big) ds.
\end{aligned}
\tag{4.28}
$$

We first notice that by the arguments from (3), we have that there exists a constant $C = C(T, r, K)$ such that $|Q_0(s, \cdot)| \le C$ and $|v_0(s, \cdot)| \le C$ for all $s \in (0, t)$, and so $|D^2_{yy} L(Q_0, v_0)| \le C$, $|D^2_{yv} L(Q_0, v_0)| \le C$, and $|D^2_{vv} L(Q_0, v_0)| \le C$.

To conclude, from (4.26) we compute

$$
\begin{aligned}
D^2_{q_i q_j} v_0 &= D_{q_j} Q_0 D^3_{pqq} H(Q_0, (m+1) D_{q_0} U^{(m+1)}(s, Q_0, Q)) D_{q_i} Q_0 \\
&+ (m+1) \sum_{k=0}^{m} D^2_{q_0 q_k} U^{(m+1)}(s, Q_0, Q)) D_{q_j} Q_k D^3_{pqp} H(Q_0, (m+1) D_{q_0} U^{(m+1)}(s, Q_0, Q)) D_{q x_i} Q_0 \\
&+ D^2_{pq} H(Q_0, (m+1) D_{q_0} U^{(m+1)}(s, Q_0, Q)) D^2_{q_i q_j} Q_0 \\
&+ D_{q_j} Q_0 D^3_{ppq} H(Q_0, (m+1) D_{q_0} U^{(m+1)}(s, Q_0, Q))(m+1) \sum_{k=0}^{m} D^2_{q_0 q_k} U^{(m+1)}(s, Q_0, Q)) D_{q_i} Q_k \\
&+ D^2_{pp} H(Q_0, (m+1) D_{q_0} U^{(m+1)}(s, Q_0, Q))(m+1) \sum_{k,l=0}^{m} D_{q_j} Q_l D^3_{q_0 q_k q_l} U^{(m+1)}(s, Q_0, Q)) D_{q_i} Q_k \\
&+ D^2_{pp} H(Q_0, (m+1) D_{q_0} U^{(m+1)}(s, Q_0, Q))(m+1) \sum_{k=0}^{m} D^2_{q_0 q_k} U^{(m+1)}(s, Q_0, \dots, Q_m)) D^2_{q_i q_j} Q_k .
\end{aligned}
$$

>From here, using the assumptions (H5) and (H13) on H, the estimates on the quantities $D^2_{q_0 q_k} U^{(m+1)}$ and $D^3_{q_0 q_k q_l} U^{(m+1)}$ and Lemma 4.16, we obtain that there exists $C = C(T, r, K) > 0$ such that

$$
\begin{aligned}
&\left| D^2_{q_i q_j} v_0(q_0, q) \right|_\infty \\
&\quad \leq \begin{cases} C, & i = j = 0, \\ \dfrac{C}{m+1}, & (i = j \text{ and } i > 0) \text{ or } (i \cdot j = 0 \text{ and } \max\{i, j\} > 0), \\ \dfrac{C}{(m+1)^2}, & i \neq j. \end{cases}
\end{aligned}
$$

Combining this with the previous arguments and with (4.28) the thesis of the claim follows. □

LEMMA 4.16. *For $m \in \mathbb{N}$ and $q = (q_0, \dots, q_m) \in (\mathbb{M})^{m+1}$, let*

$$
\mu_q^{(m+1)} := \frac{1}{(m+1)} \sum_{i=0}^{m} \delta_{q_i}, \ \ Q_i(s, q) := S_s^t \big[\mu_q^{(m+1)}\big](q_i),
$$

and

$$
P_i(s, q) := \frac{1}{(m+1)} P_s^t \big[\mu_q^{(m+1)}\big](q_i) \quad 0 \leq i \leq m.
$$

We set $U_0^{(m+1)}(q) := \mathscr{U}_0(\mu_q^{(m+1)})$ and $F^{(m+1)}(q) := \mathscr{F}(\mu_q^{(m+1)})$. Further assume $U_0^{(m+1)}$ and $F^{(m+1)}$ satisfy Property 2.2*(3). Then* (*as in Theorem* 2.3) *for $r > 0$ and $t > 0$, there exists $C = C(t, r)$ such that for all $q \in \mathbb{B}_r^{(m+1)}$, $s \in (0, t)$, and $i, j \in \{0, \dots, m\}$ we have*

$$
|D_{q_j} Q_i(s, q)|_\infty \leq \begin{cases} C, & i = j, \\ \frac{C}{(m+1)}, & i \neq j. \end{cases} \tag{4.29}
$$

and

$$(4.30)\quad \left|D^2_{q_k q_j} Q_i(s,q)\right|_\infty \le \begin{cases} C, & i=j=k,\\ \frac{C}{(m+1)}, & i=j\neq k,\ i\neq j=k,\ i=k\neq j,\\ \frac{C}{(m+1)^2}, & i\neq j\neq k. \end{cases}$$

PROOF. Let $\xi(\cdot,z)=(\xi_0(\cdot,z),\dots,\xi_m(\cdot,z))$ be defined as in (1.16) (see also the systems in (1.17) and (2.7)). By Proposition C.2 we first observe that

$$\xi(t,\cdot)^{-1}=S_0^{t,m}.$$

To facilitate the writing, as it is done in Appendix 2, we denote $\zeta(t,\cdot):=\xi^{-1}(t,\cdot)$, and so we have

$$Q_i(s,q)=\xi_i(s,\zeta(t,q)).$$

Thus, by differentiating and using the estimates on $(\xi_0,\dots,\xi_m)$ and $(\zeta_0,\dots,\zeta_m)$ from Theorem 2.3, by denoting $|\cdot|_\infty:=\|\cdot\|_{L^\infty(\mathbb{B}_r^{(m+1)})}$, we have that there exists $C=C(t,r)$ such that

$$\begin{aligned} |D_{q_j}Q_i(s,\cdot)|_\infty &\le \sum_{k=0}^{m}|D_{z_k}\xi_i(s,\zeta_0(t,\cdot),\dots,\zeta_m(t,\cdot))|_\infty|D_{q_j}\zeta_k(t,\cdot)|_\infty\\ &= |D_{z_i}\xi_i(s,\zeta_0(t,\cdot),\dots,\zeta_m(t,\cdot))|_\infty|D_{q_j}\zeta_i(t,\cdot)|_\infty\\ &\quad+\sum_{k\neq i}|D_{z_k}\xi_i(s,\zeta_0(t,\cdot),\dots,\zeta_m(t,\cdot))|_\infty|D_{q_j}\zeta_k(t,\cdot)|_\infty\\ &\le \begin{cases} C, & i=j,\\ \frac{C}{m+1}, & i\neq j.\end{cases} \end{aligned}$$

Therefore, (4.29) follows. Furthermore, since

$$\begin{aligned} D^2_{q_kq_j}Q_i(s,\cdot) &= \sum_{l_1,l_2=0}^{m} D^2_{q_{l_2}q_{l_1}}\xi_i(s,\zeta_0(t,\cdot),\dots,\zeta_m(t,\cdot)D_{q_k}\zeta_{l_2}(t,\cdot)D_{q_j}\zeta_{l_1}(t,\cdot)\\ &\quad+\sum_{l_1=0}^{m} D_{z_{l_1}}\xi_i(s,\zeta_0(t,\cdot),\dots,\zeta_m(t,\cdot))D^2_{q_kq_j}\zeta_{l_1}(t,\cdot)=\\ &= \sum_{l_1\neq l_2}^{m} D^2_{q_{l_2}q_{l_1}}\xi_i(s,\zeta_0(t,\cdot),\dots,\zeta_m(t,\cdot)D_{q_k}\zeta_{l_2}(t,\cdot)D_{q_j}\zeta_{l_1}(t,\cdot)\\ &\quad+\sum_{l=0}^{m} D^2_{q_lq_l}\xi_i(s,\zeta_0(t,\cdot),\dots,\zeta_m(t,\cdot)D_{q_k}\zeta_l(t,\cdot)D_{q_j}\zeta_l(t,\cdot)\\ &\quad+\sum_{l_1=0}^{m} D_{z_{l_1}}\xi_i(s,\zeta_0(t,\cdot),\dots,\zeta_m(t,\cdot))D^2_{q_kq_j}\zeta_{l_1}(t,\cdot), \end{aligned}$$

we have that (4.30) follows. □

LEMMA 4.17. *Let us suppose that we are in the setting of Lemma* 4.13 *and in particular all of its assumptions are in place. Let* $(Q_i)_{i=0}^m$ *be defined in* (4.18). *Let* $(q_0, q) \in \mathbb{M}^{(m+1)}$. *Then* $(0,t) \ni s \mapsto Q_0(s, q_0, q)$ *is Lipschitz-continuous with a Lipschitz constant independent of m and for all* $r > 0$ *and* $K \subset \mathbb{M}$ *compact. Let us notice that* $(Q_0(s, q_0, q))_{s\in(0,t)}$ *solves* (4.15), *with data* $\sigma_s^t[\mu_q^{(m+1)}]$ *and final condition* q_0. *Furthermore, since* $(\sigma_s^t[\mu_q^{(m+1)}])_{s\in(0,t)}$ *belongs to* $\mathscr{B}_{\beta(t,r)}$, *for some* $\beta(t,r) > 0$, *the velocity field*

$$(0,t) \times \mathbb{M} \ni (s, y) \mapsto D_p H\big(y, \nabla_w \mathscr{U}(s, \sigma_s^t[\mu_q^{(m+1)}](y))\big)$$

is globally Lipschitz-continuous after a suitable extension of the velocity field

$$\nabla_w \mathscr{U}(s, \sigma_s^t[\mu_q^{(m+1)}](\cdot)).$$

Therefore, classical results in the theory of ODEs imply the thesis of the lemma, and the bound on $Q_0(s,\cdot,\cdot)$ *depends only on* t, K, *and the Lipschitz constant of the previously mentioned velocity field* (*hence on* r).

LEMMA 4.18. *Under the assumptions of Theorem* 4.19, $u^{(m)}$ *defined in* (4.10) *satisfies the estimates* (4.22) *and* (4.23) *from Lemma* 4.13.

PROOF. In Lemma 4.15 we showed that $u^{(m)}(t,\cdot,\cdot) \in C^{1,1}_{\mathrm{loc}}(\mathbb{M}^{m+1})$ with the corresponding derivative estimates (4.20) and (4.21), uniformly with respect to $t \in [0,T]$. Furthermore, since by Proposition 4.10(v), $u(\cdot, q, \mu)$ is Lipschitz-continuous for all $q, \mu \in \mathbb{M} \times \mathscr{P}_2(\mathbb{M})$, this property is inherited by $u^{(m)}$, and therefore $u^{(m)}(\cdot, q_0, q)$ is Lipschitz-continuous on $[0,T]$ for all $(q_0, q) \in \mathbb{M}^{m+1}$.

Let us recall now the representation formula (4.17) of $u^{(m)}(t, q_0, q)$. We fix K to be the closure of a bounded open set in $\mathbb{M}$ and $r > 0$ such that $\mu_q^{(m+1)} \in \mathbb{B}_r^{m+1}$. The regularity properties of $u^{(m)}$ and (4.17) for almost every $t \in (0,T)$ and all $(q_0, q) \in \mathbb{M}^{m+1}$ yield

$$\begin{aligned}
&\partial_t u^{(m)}(t, q_0, q) + D_{q_0} u^{(m)}(t, q_0, q) \cdot D_p H\big(q_0, (m+1) D_{q_0} U^{(m+1)}(t, q_0, q)\big) \\
&\quad + \sum_{j=1}^m D_{q_j} u^{(m)}(t, q_0, q) \cdot D_p H(q_j, (m+1) D_{q_j} U^{(m+1)}(t, q_0, q)) \\
&\quad = L(q_0, D_p H(q_0, (m+1) D_{q_0} U^{(m+1)}(t, q_0, q))) + f^{(m+1)}(q_0, q_0, q).
\end{aligned} \tag{4.31}$$

Proposition 4.10(iii) and (4.16) yield

$$(m+1) D_{q_0} U^{(m+1)}(t, q_0, q) = \nabla_w \mathscr{U}\big(t, \mu_q^{(m+1)}\big)(q_0) = D_{q_0} u(t, q_0, \mu_q^{(m+1)}).$$

Now, let us notice that by the definition of $u^{(m)}$, one has the identity

$$D_{q_0} u^{(m)}(t, q_0, q) = D_{q_0} u\big(t, q_0, \mu_q^{(m+1)}\big) + \frac{1}{m+1} \nabla_w u\big(t, q_0, \mu_q^{(m+1)}\big)(q_0).$$

For an arbitrary $a \in \mathbb{M}$, if we set in $\widehat{u}^{(m+1)}(t, a, q_0, q) := u(t, a, \mu_q^{(m+1)})$, we have that

$$\frac{1}{m+1} \nabla_w u(t, q_0, \mu_q^{(m+1)})(q_0) = D_{q_0} \widehat{u}^{(m+1)}(t, a, q_0, q)\Big|_{a=q_0}$$

and so

(4.32)

$$\begin{aligned}(m+1) D_{q_0} U^{(m+1)}(t, q_0, q) &= D_{q_0} u(t, q_0, \mu_q^{(m+1)}) \\ &= D_{q_0} u^{(m)}(t, q_0, q) - D_{q_0} \widehat{u}^{(m+1)}(t, q_0, q_0, q).\end{aligned}$$

We notice furthermore that $\widehat{u}^{(m+1)}$ (with respect to the regularity and derivative estimates) essentially behaves as $u^{(m+1)}(t, q_0, q_0, q)$, and in particular by (4.20) and (4.21) there exists a constant $C = C(K, r) > 0$ such that

$$\left| D_{q_0} \widehat{u}^{(m+1)}(t, q_0, q_0, q) \right| \le \frac{C}{m+2}.$$

All these arguments allow us to conclude that

$$\left| (m+1) D_{q_0} U^{(m+1)}(t, q_0, q) \right| \le C.$$

Now, we differentiate (4.31) with respect to the spatial variables.

Differentiating with respect to q_0, denoting the variables of $f^{(m+1)}$ as (y_0, q_0, q), we find that there exists $C = C(T, K, r)$ such that if $(t, q_0, q) \in [0, T] \times \mathbb{B}_r^{(m+1)}$ with $q_0 \in K$, then

$$\begin{aligned}&\left| D_{q_0} \partial_t u^{(m)} \right| \\ &\quad \le \left| D^2_{q_0 q_0} u^{(m)} \right| \left| D_p H(q_0, (m+1) D_{q_0} U^{(m+1)}) \right| \\ &\qquad + \left| D_{q_0} u^{(m)} \right| \left| D^2_{qp} H\big(q_0, (m+1) D_{q_0} U^{(m+1)}\big) \right| \\ &\qquad + (m+1) \left| D_{q_0} u^{(m)} \right| \left| D^2_{pp} H(q_0, (m+1) D_{q_0} U^{(m+1)}) \right| \left| D^2_{q_0 q_0} U^{(m+1)} \right| \\ &\qquad + \sum_{j=1}^{m} \left| D^2_{q_0 q_j} u^{(m)} \right| \left| D_p H(q_j, (m+1) D_{q_j} U^{(m+1)}) \right| + \mathrm{I} + \mathrm{II}\end{aligned}$$

where

$$\begin{aligned}\mathrm{I} := &\sum_{j=1}^{m} \left| D_{q_j} u^{(m)} \right| \left| D^2_{pp} H(q_j, (m+1) D_{q_j} U^{(m+1)}) \right| (m+1) \left| D^2_{q_0 q_j} U^{(m+1)} \right| \\ &+ \left| D_{q_0} L\big(q_0, D_p H(q_0, (m+1) D_{q_0} U^{(m+1)})\big) \right| \\ &+ \left| D_v L\big(q_0, D_p H\big(q_0, (m+1) D_{q_0} U^{(m+1)}\big)\big) \right| \left| D^2_{qp} H \right|\end{aligned}$$

and

$$\begin{aligned}\mathrm{II} :=& \left|D_v L\big(q_0, D_p H\big(q_0,(m+1)D_{q_0}U^{(m+1)}\big)\big)\right|\left|D^2_{pp}H\right|(m+1)\left|D^2_{q_0q_0}U^{(m+1)}\right|\\ &+\left|D_{y_0}f^{(m+1)}(q_0,q_0,q)\right|+\left|D_{q_0}f^{(m+1)}(q_0,q_0,q)\right|.\end{aligned}$$

Thus, using (4.20), (4.21), and the estimates on $U^{(m+1)}$ from Theorem 2.3, as well as the hypotheses on the data H and $f^{(m+1)}$, we have

$$\begin{aligned}&\left|D_{q_0}\partial_t u^{(m)}\right|\\ &\quad\le C + C\left(\sum_{j=1}^m m|D^2_{q_0q_j}u^{(m)}|^2\right)^{\frac12}\left(\sum_{j=1}^m \frac{1}{m}\left|D_pH(q_j,(m+1)D_{q_j}U^{(m+1)})\right|^2\right)^{\frac12}\\ &\qquad + C + \sum_{i=0}^m \frac{1}{\sqrt{m+1}}\sqrt{m+1}|D_{q_i}f^{(m+1)}|\\ &\quad\le C + \left(\sum_{i=0}^m \frac{1}{m+1}\right)^{\frac12}\left(\sum_{i=1}^m (m+1)|D_{q_i}f^{(m+1)}|^2\right)^{\frac12}\le C,\end{aligned}$$

This yields the first part of (4.22), since

$$D_pH\Big(\cdot,\nabla\mathscr{U}\big(t,\mu_q^{(m+1)}\big)(\cdot)\Big)\in L^2\big(\mu_q^{(m+1)}\big),$$

with an $L^2(\mu_q^{(m+1)})$-norm uniformly bounded with respect to m.

If $k\in\{1,\dots,m\}$, completely parallel computation gives

$$\begin{aligned}&\left|D_{q_k}\partial_t u^{(m)}\right|\\ &\quad\le \left|D^2_{q_kq_0}u^{(m)}\right|\left|D_pH(q_0,(m+1)D_{q_0}U^{(m+1)})\right|\\ &\qquad+(m+1)\left|D_{q_0}u^{(m)}\right|\left|D^2_{pp}H\big(q_0,(m+1)D_{q_0}U^{(m+1)}\big)\right|\left|D^2_{q_kq_0}U^{(m+1)}\right|\\ &\qquad+\sum_{j=1}^m\left|D^2_{q_kq_j}u^{(m)}\right|\left|D_pH\big(q_j,(m+1)D_{q_j}U^{(m+1)}\big)\right|\\ &\qquad+\left|D_{q_k}u^{(m)}\right|w\left|D^2_{qp}H\big(q_k,(m+1)D_{q_k}U^{(m+1)}\big)\right|\\ &\qquad+\sum_{j=1}^m\left|D_{q_j}u^{(m)}\right|(m+1)\left|D^2_{pp}H\big(q_j,(m+1)D_{q_j}U^{(m+1)}\big)\right|\left|D^2_{q_kq_j}U^{(m+1)}\right|\\ &\qquad+\left|D_vL\big(q_0,D_pH\big(q_0,(m+1)D_{q_0}U^{(m+1)}\big)\big)\right|\left|D^2_{pp}H\right|(m+1)\left|D^2_{q_kq_0}U^{(m+1)}\right|\\ &\qquad+\left|D_{q_k}f^{(m+1)}\right|\\ &\quad\le C\left|D^2_{q_kq_0}u^{(m)}\right|+\frac{C}{(m+1)}\left|D_pH\big(q_k,(m+1)D_{q_k}U^{(m+1)}\big)\right|+\frac{C}{(m+1)}\\ &\qquad+\left|D_{q_k}f^{(m+1)}\right|,\end{aligned}$$

from where, using the same arguments as for the conclusion of the first part of (4.22), we find $\sum_{k=1}^m(m+1)|D_{q_k}\partial_t u^{(m)}|^2\le C$, as desired.

To show (4.23), we argue similarly. First, from (4.31) we simply have

$$
\begin{aligned}
&\left|\partial_t u^{(m)}\right| \\
&\quad \le \left|D_{q_0} u^{(m)}\right|\left|D_p H(q_0, (m+1) D_{q_0} U^{(m+1)})\right| \\
&\qquad + \sum_{j=1}^{m} |D_{q_j} u^{(m)}||D_p H(q_j, (m+1) D_{q_j} U^{(m+1)})| \\
&\qquad + |L(q_0, D_p H(q_0, (m+1) D_{q_0} U^{(m+1)}))| + |f^{(m+1)}| \\
&\quad \le C + \left(\sum_{j=1}^{m} m |D_{q_j} u^{(m)}|^2\right)^{\frac12} \left(\sum_{j=1}^{m} \frac{1}{m} |D_p H(q_j, (m+1) D_{q_j} U^{(m+1)})|^2\right)^{\frac12} \\
&\quad \le C,
\end{aligned}
$$

where we used the previous estimates and the fact that $H(q_0, D_{q_0} u^{(m)})$ and $f^{(m+1)}$ are locally bounded.

Second, differentiating (4.31) with respect to t, we find

$$
\begin{aligned}
&\left|\partial^2_{tt} u^{(m)}\right| \\
&\quad \le \left|\partial_t D_{q_0} u^{(m)}\right|\left|D_p H\left(q_0, (m+1) D_{q_0} U^{(m+1)}\right)\right| \\
&\qquad + \left|D_{q_0} u^{(m)}\right|\left|D^2_{pp} H\left|(m+1)\right|\partial_t D_{q_0} U^{(m+1)})\right| + \\
&\qquad + \sum_{j=1}^{m} \left|\partial_t D_{q_j} u^{(m)}\right|\left|D_p H\left(q_j, (m+1) D_{q_j} U^{(m+1)}\right)\right| \\
&\qquad + \sum_{j=1}^{m} \left|D_{q_j} u^{(m)}\right|\left|D^2_{pp} H\left(q_j, (m+1) D_{q_j} U^{(m+1)}\right)\right|(m+1)\left|\partial_t D_{q_j} U^{(m+1)}\right| \\
&\qquad + |(m+1) D_{q_0} U^{(m+1)}||D^2_{pp} H|(m+1)|\partial_t D_{q_0} U^{(m+1)})| \\
&\quad \le C + \left(\sum_{j=1}^{m} (m+1) |\partial_t D_{q_j} u^{(m)}|^2\right)^{\frac12} \left(\sum_{j=1}^{m} \frac{1}{(m+1)} |D_p H(q_j, (m+1) D_{q_j} U^{(m+1)})|^2\right)^{\frac12} \\
&\qquad + C(m+1)|\partial_t D_{q_0} U^{(m+1)})| \\
&\qquad + C \left(\sum_{j=1}^{m} (m+1) |\partial_t D_{q_j} U^{(m+1)}|^2\right)^{\frac12}.
\end{aligned}
$$

Let us notice that by (4.32) we have that

$$
(m+1)|\partial_t D_{q_0} U^{(m+1)}| \le |\partial_t D_{q_0} u^{(m)}| + |\partial_t D_{q_0} \hat{u}^{(m+1)}| \le C + \frac{C}{\sqrt{m+2}},
$$

where we have used that $\sum_{j=0}^{m} (m+2) |\partial_t D_{q_0} \hat{u}^{(m+1)}|^2 \le C$. Relying on the previously obtained estimates and on the fact that by Theorem 2.3(3),

$$
\sum_{j=1}^{m} (m+1) |\partial_t D_{q_j} U^{(m+1)}|^2 \le C,
$$

the claim in (4.23) follows. □

Recall that throughout this section, we have imposed that (H1)–(H7) and (H8) hold. We are ready to state and prove the main theorem of this section.

THEOREM 4.19. *Suppose the assumptions* (H1) *through* (H15) *are satisfied. Then, the scalar master equation* (4.8) *has a unique global-in-time classical solution of class*

$$C^{1,1}_{\rm loc}([0,+\infty)\times\mathbb{M}\times\mathscr{P}_2(\mathbb{M}))$$

in the sense of Definition 4.6.

PROOF. Let $T>0$ be a fixed time horizon. Notice that Theorem 4.1 yields that u defined in (4.10) is of class $C^{1,1}_{\rm loc}([0,T]\times\mathbb{M}\times\mathscr{P}_2(\mathbb{M}))$.

Let $\mu\in\mathscr{P}_2(\mathbb{M})$, $q\in\mathbb{M}$, and $t\in(0,T)$. Using the representation formula (4.10), by the dynamic programming principle, we have that for $s\in(0,t)$

$$\begin{aligned}
&u(t,q,\mu)\\
&\quad= u\big(s,S^t_s[\mu](q),\sigma^t_s[\mu]\big)\\
&\qquad+\int_s^t L\big(S^t_\tau[\mu](q),D_pH\big(S^t_\tau[\mu](q),D_qu(\tau,S^t_\tau[\mu](q),\sigma^t_\tau[\mu])\big)\big)d\tau\\
&\qquad+\int_s^t f\big(S^t_\tau[\mu](q),\sigma^t_s[\mu]\big)d\tau.
\end{aligned}$$

Hence,

$$\begin{aligned}
&\lim_{s\to t}\frac{u(t,q,\mu)-u\big(s,S^t_s[\mu](q),\sigma^t_s[\mu]\big)}{t-s}\\
&\quad=\lim_{s\to t}\bigg\{\int_s^t L\big(S^t_\tau[\mu](q),D_pH\big(S^t_\tau[\mu](q),D_qu(\tau,S^t_\tau[\mu](q),\sigma^t_\tau[\mu])\big)\big)d\tau\\
&\qquad\qquad+\int_s^t f\big(S^t_\tau[\mu](q),\sigma^t_s[\mu]\big)d\tau\bigg\},
\end{aligned}$$

where both limits exist and are finite, due to the continuity of the integrand on the right-hand side. Using the chain rule with respect to the measure variable (provided in Lemma 4.20), this is equivalent to

$$\begin{aligned}
&\partial_t u(t,q,\mu)+D_qu(t,q,\mu)\cdot D_pH(q,D_qu(t,q,\mu))\\
&\quad+\int_{\mathbb{M}}\nabla_w u(t,q,\mu)(y)\cdot D_pH\big(y,\nabla_w\mathscr{U}(s,\mu)(y)\big)\mu(dy)\\
&\qquad= L\big(q,D_pH\big(q,D_qu(t,q,\mu)\big)\big)+f(q,\mu)
\end{aligned}$$

Here we used that the optimal curve $\tau\mapsto S^t_\tau[\mu](q)$ satisfies (4.14), while the curve $\tau\mapsto\sigma^t_\tau[\mu]$ solves the continuity equation (C.4).

Using that by Proposition 4.10(ii)

$$D_qu(t,\cdot,\mu)=\nabla_w\mathscr{U}(t,\mu)(\cdot)\ \ \mu-\text{a.e.},$$

one obtains

$$\begin{aligned} f(q,\mu) &= \partial_t u(t,q,\mu) + D_q u(t,q,\mu)\cdot D_p H(q, D_q u(t,q,\mu)) \\ &\quad + \int_{\mathbb{M}} \nabla_w u(t,q,\mu)(y)\cdot D_p H(y, D_q u(t,y,\mu))d\mu(y) \\ &\quad - L(q, D_p H(q, D_q u(t,q,\mu))) \\ &= \partial_t u(t,q,\mu) + H(q, D_q u(t,q,\mu)) \\ &\quad + \int_{\mathbb{M}} \nabla_w u(t,q,\mu)(y)\cdot D_p H(y, D_q u(t,y,\mu))\mu(dy), \end{aligned}$$

where we have used the Legendre duality in the last equation. The arguments in Section 5.1 imply in particular that u also satisfies the condition (4.9). This completes the existence part of the theorem.

Uniqueness. Let $u \in C^{1,1}_{\rm loc}([0,T]\times\mathbb{M}\times\mathscr{P}_2(\mathbb{M}))$ be a solution to (4.8). Let $t\in(0,T)$, $\mu\in\mathscr{P}_2(\mathbb{M})$, and $z\in\mathbb{H}$ be fixed such that $\sharp(z)=\mu$. Using the vector field $D_pH(\cdot, D_qu(\cdot,\cdot,\cdot))$, let $(\sigma_s)_{s\in(0,t)}$ be the unique solution to the continuity equation

$$\begin{cases} \partial_s\sigma_s + \nabla\cdot(\sigma_s D_pH(\cdot, D_qu(s,\cdot,\sigma_s))) = 0 & \text{in } \mathscr{D}'((0,t)\times\mathbb{M}), \\ \sigma_t = \mu. \end{cases} \tag{4.33}$$

Since D_qu is locally Lipschitz on $[0,T]\times\mathbb{M}\times\mathscr{P}_2(\mathbb{M})$ and the vector field $\mathbb{M}\ni q\mapsto D_qu(t,q,\nu)$ is Lipschitz, uniformly with respect to $(t,\nu)\in[0,T]\times\mathscr{B}_r$, the existence and uniqueness of σ above follows from standard arguments and from the adaptation of theorem 3.3 from [30].

Then, in $\mathbb{H}$ we consider the ODE

$$\begin{cases} x_s' = D_pH(x_s, D_qu(s,x_s,\sigma_s)), & s\in(0,t), \\ x_t = z. \end{cases} \tag{4.34}$$

This has a unique continuously differentiable solution $x:(0,t)\to\mathbb{H}$.

CLAIM 1. We have that $\sharp(x_s)=\sigma_s$.

PROOF OF CLAIM 1. Indeed, let us denote $\overline{\sigma}_s := \sharp(x_s)$; we have

$$\partial_s\overline{\sigma}_s + \nabla\cdot(\overline{\sigma}_s D_pH(\cdot, D_qu(s,\cdot,\sigma_s))) = 0,$$

in the sense of distributions. But the vector field $(s,q)\mapsto D_pH(q, D_qu(s,q,\sigma_s))$ induces a unique solution to the continuity equation; therefore σ and $\overline{\sigma}$ must coincide and the claim follows.

CLAIM 2. The unique solution x to (4.34) satisfies the Euler–Lagrange equations

$$D_qL(x_s,x_s') + \nabla\widetilde{\mathscr{F}}(x_s) = \frac{d}{ds}D_vL(x_s,x_s') \text{ and } D_vL(x(0),x'(0)) = \nabla\widetilde{\mathscr{U}_0}(x(0)),$$

a.e. in Ω.

PROOF OF CLAIM 2. Let us notice first that by our assumptions $D_v L(q,\cdot)$ and $D_p H(q,\cdot)$ are inverses of each other for all $q \in \mathbb{M}$. Furthermore, we have

$$D_q L(q, D_p H(q,p)) = -D_q H(q,p) \ \ \forall (q,p) \in \mathbb{M} \times \mathbb{R}^d.$$

Indeed, this last equation is a consequence of the Legendre-Fenchel identity

$$H(q,p) = p \cdot D_p H(q,p) - L(q, D_p H(q,p)).$$

Now, from (4.34) by continuity, by (H10), and by the fact

$$\nabla_w \mathscr{U}_0(\sigma_s)(x_s) = \nabla \widetilde{\mathscr{U}_0}(x_s),$$

one can deduce that

$$\begin{aligned} x'(0) = D_p H(x(0), D_q u_0(x(0),\sigma_0)) &= D_p H(x(0), \nabla_w \mathscr{U}_0(\sigma_0)(x(0))) \\ &= D_p H(x(0), \nabla \widetilde{\mathscr{U}_0}(x(0))), \end{aligned}$$

which by inversion of $D_p H(x(0),\cdot)$ is equivalent to

$$D_v L(x(0), x'(0)) = \nabla \widetilde{\mathscr{U}_0}(x(0)).$$

0

Then, from (4.34), again by inversion of $D_p H(x_s,\cdot)$ we have

$$D_v L(x_s, x_s') = D_q u(s, x_s, \sigma_s).$$

Since $u \in C^{1,1}_{\mathrm{loc}}([0,T] \times \mathbb{M} \times \mathscr{P}_2(\mathbb{M}))$, for a.e. $s \in (0,t)$ we have

(4.35)
$$\begin{aligned} \frac{d}{ds} D_v L(x_s, x_s') &= \partial_s D_q u(s, x_s, \sigma_s) + D^2_{qq} u(s, x_s, \sigma_s) D_p H(x_s, D_q u(s, x_s, \sigma_s)) \\ &\quad + \int_{\mathbb{M}} \nabla_w D_q u(s, x_s, \sigma_s)(a) \cdot D_p H(a, D_q u(s, a, \sigma_s)) \sigma_s(da) \\ &= \partial_s D_q u(s, x_s, \sigma_s) + D^2_{qq} u(s, x_s, \sigma_s) D_p H(x_s, D_q u(s, x_s, \sigma_s)) \\ &\quad + \int_{\mathbb{M}} D_q \nabla_w u(s, x_s, \sigma_s)(a) \cdot D_p H(a, D_q u(s, a, \sigma_s)) \sigma_s(da), \end{aligned}$$

a.e. in Ω, where we have used (4.9) in the last equation. Let us note that the previous computation is meaningful. Indeed, by the regularity on u (see also the arguments in Section 5.1), we can differentiate the master equation (4.8) with respect to q, and so for $\mathscr{L}^1 \otimes \mathscr{L}^d$-a.e. $(s,q) \in (0,t) \times \mathbb{M}$ and for all $\nu \in \mathscr{P}_2(\mathbb{M})$ we have

$$\begin{aligned} &\partial_s D_q u(s,q,\nu) + D^2_{qq} u(s,q,\nu) D_p H(q, D_q u(s,q,\nu)) \\ &\quad + \int_{\mathbb{M}} D_q \nabla_w u(s,q,\nu)(a) D_p H(a, D_q u(s,a,\nu)) \nu(da) \\ &\quad = D_q f(q,\nu) - D_q H(q, D_q u(s,q,\nu)). \end{aligned} \tag{4.36}$$

We notice that (H10) implies that $D_q f(q,\nu) = \nabla_w \mathscr{F}(\nu)(q)$ and so, by combining (4.35) and (4.36) one deduces

$$\begin{aligned}\frac{d}{ds} D_v L(x_s, x_s') &= D_q f(x_s, \sigma_s) - D_q H(x_s, D_q u(s, x_s, \sigma_s)) \\ &= \nabla_w \mathscr{F}(\sigma_s)(x_s) + D_q L(x_s, D_q u(s, x_s, \sigma_s)) \\ &= \nabla \widetilde{\mathscr{F}}(x_s) + D_q L(x_s, D_q u(s, x_s, \sigma_s)),\end{aligned}$$

and so the claim follows.

CLAIM 3. For each $t \in [0,T]$ and $\mu \in \mathscr{P}_2(\mathbb{M})$, $u(t,\cdot,\mu)$ is uniquely determined on $\mathrm{spt}(\mu)$.

PROOF OF CLAIM 3. By the strict convexity of the action, the previous claims show that $(x_s)_{s\in(0,t)}$ is the unique solution in the action minimization problem (1.5) for $\widetilde{\mathscr{U}}(t,z)$ Butsince $\widetilde{\mathscr{U}} \in C^{1,1}_{\mathrm{loc}}([0,T]\times\mathbb{H})$ (as we showed in Proposition 1.5(ii)), we have at the same time that the optimal velocity for this curve is $D_p H(x_s, \nabla\widetilde{\mathscr{U}}(s,x_s))$, and so, by the convexity of H in the second variable, one deduces that

$$D_q u(s, x_s(\omega), \sigma_s) = \nabla \widetilde{\mathscr{U}}(s, x_s)(\omega),$$

for a.e. $\omega \in \Omega$. This further yields that the vector field $q \mapsto D_q u(s,q,\sigma_s)$ is unique (i.e., does not depend on the solution u) on $\mathrm{spt}(\sigma_s)$ for all $s \in [0,t]$. From here we also deduce that for each $\mu \in \mathscr{P}_2(\mathbb{M})$, the solution to the continuity equation (4.33) is unique (independent of the solution u) and this corresponds to the unique minimizer in the action minimization problem, i.e., to the solution to (C.4).

Now let $q_1 \in \mathrm{spt}(\mu)$ and let $(q_s)_{s\in(0,t)}$ be the unique solution to

$$\begin{cases} q_s' = D_p H(q_s, D_q u(s, q_s, \sigma_s)), & s \in (0,t), \\ q_t = q_1. \end{cases} \tag{4.37}$$

It is clear that $q_s \in \mathrm{spt}(\sigma_s)$ for all $s \in [0,t]$. Moreover, for each fixed q_1, the curve solving (4.37) is unique (independent of the solution u).

Using the Legendre duality, the master equation for u can be rewritten as

$$\begin{aligned}&\partial_s u(s,q,\nu) + D_q u(s,q,\nu)\cdot D_p H(q, D_q u(s,q,\nu)) \\ &\quad + \int_{\mathbb{M}} \nabla_w u(s,q,\nu)(a)\cdot D_p H(a, D_q u(s,a,\nu))\nu(da) \\ &\quad = f(q,\nu) + L\big(q, D_p H(q, D_q u(s,q,\nu))\big),\end{aligned}$$

and replacing in $(q,\nu) = (q_s,\sigma_s)$ the chain rule gives us

$$\frac{d}{ds}\left(u(s,q_s,\sigma_s)\right) = f(q_s,\sigma_s) + L\big(q_s, D_p H(q_s, D_q u(s,q_s,\sigma_s))\big). \tag{4.38}$$

Now, let $\bar{u} \in C^{1,1}_{\mathrm{loc}}([0,T]\times\mathbb{M}\times\mathscr{P}_2(\mathbb{M}))$ be another solution to (4.8) in the sense of Definition 4.6. By the previous arguments one has $D_q\bar{u}(s,q,\sigma_s) =$

$D_q u(s, q, \sigma_s)$ for all $s \in [0, t]$ and $q \in \operatorname{spt}(\sigma_s)$. Then, similarly to (4.38), one has that

$$\frac{d}{ds}\left(\overline{u}(s, q_s, \sigma_s)\right) = f(q_s, \sigma_s) + L\big(q_s, D_p H(q_s, D_q u(s, q_s, \sigma_s))\big). \tag{4.39}$$

By defining now $w : [0, t] \to \mathbb{R}$ as $w(s) := u(s, q_s, \sigma_s) - \overline{u}(s, q_s, \sigma_s)$, we have that $w'(s) = 0$ (by subtracting (4.39) from (4.38)) and $w(0) = 0$. Therefore one must have $w \equiv 0$ and so $u(s, q_s, \sigma_s) = \overline{u}(s, q_s, \sigma_s)$. By continuity one has also that

$$u(t, q_1, \mu) = \overline{u}(t, q_1, \mu) \ \ \forall q_1 \in \operatorname{spt}(\mu).$$

CLAIM 4. u is a unique solution to (4.8).

PROOF OF CLAIM 4. It remains to show that if u and $\overline{u}$ are two solutions to (4.8), one has $u(t, q, \mu) = \overline{u}(t, q, \mu)$ for all $q \in \mathbb{M} \setminus \operatorname{spt}(\mu)$. Suppose that μ does not have full support; otherwise there is nothing to prove. Let $q_0 \in \mathbb{M} \setminus \operatorname{spt}(\mu)$. For $\varepsilon > 0$ let ρ_ε stand for the heat kernel centered at 0 with variance $\varepsilon > 0$, and define $\mu_\varepsilon := \mu * \rho_\varepsilon$. Then one obtained a fully supported smooth probability measure μ_ε such that $W_2(\mu, \mu_\varepsilon) \to 0$ as $\varepsilon \downarrow 0$. Therefore, we have

$$u(t, q_0, \mu_\varepsilon) = \overline{u}(t, q_0, \mu_\varepsilon).$$

By the continuity of both u and μ_ε with respect to the measure variable, one can pass to the limit as $\varepsilon \downarrow 0$ to obtain that

$$u(t, q_0, \mu) = \overline{u}(t, q_0, \mu),$$

as desired. □

Despite the fact that the velocity field $v(t, \cdot) := D_p H\big(\cdot, \nabla_w \mathscr{U}\big(t, \mu\big)\big)$ appearing in the continuity equation (C.4) typically does not belong to $T_\mu \mathscr{P}_2(\mathbb{M})$, we have the following chain rule (cf. [39] in the compact setting).

LEMMA 4.20. *We assume that the hypotheses of Theorem* 4.19 *take place. Let* $T > 0$, $t_0, t \in (0, T)$, $s \in (0, t)$, $q \in \mathbb{M}$, *and* $\mu \in \mathscr{P}_2(\mathbb{M})$ *and let* $(0, t) \ni s \mapsto \sigma_s^t[\mu]$ *be the solution to the continuity equation* (C.4). *Then*

$$\lim_{s \to t} \frac{u(t_0, q, \mu) - u\big(t_0, q, \sigma_s^t[\mu]\big)}{t - s} = \int_{\mathbb{M}} \nabla_w u(t_0, q, \mu)(y) \cdot D_p H\big(y, \nabla_w \mathscr{U}\big(t, \mu\big)(y)\big)\mu(dy).$$

5 Further Implications of the Scalar Master Equation

5.1 Improvements on the notion of weak solution to the vectorial master equation

Let us recall that the first part of Theorem 4.19 asserts the existence of $u \in C^{1,1}_{\rm loc}([0,T]\times\mathbb{M}\times\mathscr{P}_2(\mathbb{M}))$, which satisfies the scalar master equation

$$\begin{aligned}&\partial_t u(t,q,\mu)+H(q,D_q u(t,q,\mu))\\&\quad+\int_{\mathbb{M}}\nabla_w u(t,q,\mu)(y)\cdot D_pH(y,D_qu(t,y,\mu))\mu(dy)=f(q,\mu).\end{aligned}\tag{5.1}$$

Let us observe that all the terms in the previous equation are locally Lipschitz-continuous with respect to the q-variable. Indeed, except the nonlocal term, the Lipschitz continuity of the others is a consequence of the regularity of u and the data. Setting $v(t,y):=D_pH\big(y,\nabla_w\mathscr{U}\big(t,\mu\big)(y)\big)$ and denoting by $\overline{v}(t,\cdot)$ the projection of $v(t,\cdot)$ onto $T_\mu\mathscr{P}_2(\mathbb{M})$, we have that

$$\int_{\mathbb{M}}\nabla_w u(t,q,\mu)(y)\cdot v(t,y)\mu(dy)=\int_{\mathbb{M}}\Phi_1(t,q,\mu,y)\cdot\overline{v}(t,y)\mu(dy),$$

where Φ_1 is defined in Corollary 3.17. This relationship holds because we have that $\nabla_w u(t,q,\mu)(\cdot)$ is the projection of $\Phi_1(t,q,\mu,\cdot)$ onto $T_\mu\mathscr{P}_2(\mathbb{M})$. Since

$$\Phi_1\in C^{1,1}_{\rm loc}([0,T]\times\mathbb{M}\times\mathscr{P}_2(\mathbb{M})\times\mathbb{M}),$$

the function $q\mapsto\int_{\mathbb{M}}\Phi_1(t,q,\mu,y)\cdot\overline{v}(t,y)\mu(dy)$ is locally Lipschitz-continuous and for (Lebesgue) a.e. $q\in\mathbb{M}$, we have

$$\int_{\mathbb{M}}D_q\nabla_w u(t,q,\mu)(y)\cdot v(t,y)\mu(dy)=\int_{\mathbb{M}}D_q\Phi_1(t,q,\mu,y)\cdot\overline{v}(t,y)\mu(dy).$$

Therefore, we are allowed to differentiate (5.1) for (Lebesgue) a.e. $q\in\mathbb{M}$ to obtain

$$\begin{aligned}&\partial_t D_qu(t,q,\mu)+D_qH\big(q,D_qu(t,q,\mu)\big)+D^2_{qq}u(t,q,\mu)D_pH\big(q,D_qu(t,q,\mu)\big)\\&\quad+\int_{\mathbb{M}}D_q\nabla_w u(t,q,\mu)(y)\cdot D_pH\big(y,D_qu(t,y,\mu)\big)\mu(dy)=D_qf(q,\mu).\end{aligned}$$

By Proposition 4.10(iii) we know that for all $(t,\mu)\in(0,T)\times\mathscr{P}_2(\mathbb{M})$,

$$D_qu(t,\cdot,\mu)=\nabla_w\mathscr{U}(t,\mu)(\cdot)\ \text{ on spt}(\mu),$$

where $\mathscr{U}$ is the unique solution to (4.1). Since D_qu is locally Lipschitz-continuous with respect to all of its variables, it serves a natural extension for $\nabla_w\mathscr{U}(t,\mu)(\cdot)$ to the whole space, and so we have

$$\begin{aligned}&\partial_t\nabla_w\mathscr{U}(t,\mu)(q)+D_qH(q,\nabla_w\mathscr{U}(t,\mu)(q))\\&\quad+D_q\nabla_w\mathscr{U}(t,\mu)(q)D_pH(q,\nabla_w\mathscr{U}(t,\mu)(q))\\&\quad+\int_{\mathbb{M}}D_q\nabla_w u(t,q,\mu)(y)\cdot D_pH(y,\nabla_w\mathscr{U}(t,\mu)(y))\mu(dy)\\&\quad=D_qf(q,\mu)=\nabla_w\mathscr{F}(\mu)(q),\end{aligned}\tag{5.2}$$

for all $(t,\mu) \in (0,T) \times \mathscr{P}_2(\mathbb{M})$ and for (Lebesgue) a.e. $q \in \mathbb{M}$.

In Theorem 4.4 we have seen that $\mathscr{V} := \nabla_w \mathscr{U}$ solves the vectorial master equation (4.2) when the variable q needs to be taken in $\mathrm{spt}(\mu)$. Since we have a correspondence between all terms in (4.2) and (5.2), except the nonlocal ones, we can deduce that we must have

$$\begin{aligned}
&\bar{\mathscr{N}_\mu}\big[\mathscr{V}, \nabla_w^\top \mathscr{V}\big](t,\mu,q) \\
&= \bar{\mathscr{N}_\mu}\big[\nabla_w \mathscr{U}, \nabla_{ww}^2 \mathscr{U}^\top\big](t,\mu,q) \\
&= \int_{\mathbb{M}} D_q \nabla_w u(t,q,\mu)(y) \cdot D_p H(y, \nabla_w \mathscr{U}(t,\mu)(y))\mu(dy)
\end{aligned}$$

for $\mathscr{L}^d$-a.e. $q \in \mathbb{M}$.

This fact implies furthermore that

$$\begin{aligned}
&\int_{\mathbb{M}} D_q \nabla_w u(t,q,\mu)(y) D_p H(y, D_q u(t,y,\mu))\mu(dy) \\
&\quad = \int_{\mathbb{M}} \nabla_w D_q u(t,q,\mu)(y) D_p H(y, D_q u(t,y,\mu))\mu(dy) \\
&\quad = \int_{\mathbb{M}} \nabla_{ww}^2 \mathscr{U}(t,\mu)(q,y) D_p H(y, D_q u(t,y,\mu))\mu(dy)
\end{aligned} \tag{5.3}$$

for all $\mu \in \mathscr{P}_2(\mathbb{M})$ and for $\mathscr{L}^1 \otimes \mathscr{L}^d$-a.e. $(t,q) \in (0,T) \times \mathbb{M}$, which shows in particular that the function u constructed in the first part of the proof of Theorem 4.19 satisfies also (4.9).

All the previous arguments allow to formulate the following:

PROPOSITION 5.1. *The weak solution $\mathscr{V}$ to the vectorial master equation* (4.2) *provided in Theorem* 4.4 *can be extended in a Lipschitz-continuous way to* $[0,T] \times \mathscr{P}_2(\mathbb{M}) \times \mathbb{M}$ *such that this extension still solves* (4.2) *at every* $(t,\mu) \in (0,T) \times \mathscr{P}_2(\mathbb{M})$ *and at* $\mathscr{L}^d$*-a.e.* $q \in \mathbb{M}$.

Remark 5.2. Relying on the very same procedure as in Theorems 2.3 and 3.16, if we assume higher regularity properties on the data (as $H, L \in C^4$ with uniformly bounded fourth-order derivatives, $\mathscr{F}, \mathscr{U} \in C^{3,1,w}_{\mathrm{loc}}$ and $f, u_0 \in C^{2,1}_{\mathrm{loc}}$), one can improve further the regularity of both u and $\mathscr{U}$ (as $u \in C^{2,1}_{\mathrm{loc}}$ and $\mathscr{U} \in C^{3,1,w}_{\mathrm{loc}}$). Such improvements would imply furthermore that one could have the vectorial master equation satisfied for all $q \in \mathbb{M}$ (rather than $\mathscr{L}^d$-a.e.). We do not pursue the realistic goal of improving the regularity of u only to avoid writing a longer paper.

Appendix A Hilbert Regularity Is Too Stringent for Rearrangement Invariant Functions

Let $\Phi \in C^2(\mathscr{P}_2(\mathbb{M}))$ and let $\widetilde{\Phi} \in C^2(\mathbb{H})$ be such that $\Phi(\mu) = \widetilde{\Phi}(x)$ if μ is the law of x. Recall that

$$\begin{aligned}\nabla^2\widetilde{\Phi}(x)(h,h_*) &= \int_\Omega D_q\big(\nabla_w\Phi(\mu)\big)\circ x\; h\cdot h_*\,d\omega \\ &\quad + \int_{\Omega^2}\nabla^2_{ww}\Phi(\mu)\big(x(\omega),x(\omega_*)\big)h(\omega)\cdot h_*(\omega_*)d\omega\,d\omega_*\end{aligned} \tag{A.1}$$

if $\xi,\xi_* \in T_\mu\mathscr{P}_2(\mathbb{M})$ and $h = \xi\circ x$ and $h_* = \xi_*\circ x$.

For $k \in \mathbb{N}$ and $g \in C^2(\mathbb{M}^k)$, we define

$$\widetilde{\Phi}^{(k)}_g(x) := \int_{\Omega^k} g\big(x(\omega_1),\dots,x(\omega_k)\big)d\omega_1\cdots d\omega_k \quad \forall x \in \mathbb{H}$$

and

$$\Phi^{(k)}_g(\mu) := \int_{\mathbb{M}^k} g(q_1,\dots,q_k)\mu(dq_1)\cdots\mu(dq_k) \quad \forall \mu \in \mathscr{P}_2(\mathbb{M}).$$

Let P_k be the set of permutations of k letters. Replacing g by its symmetrization

$$\widetilde{g}(x_1,\dots,x_k) = \frac{1}{k!}\sum_{\tau\in P_k} g(x_{\tau(1)},\dots,x_{\tau(k)}),$$

we have $\widetilde{\Phi}^{(k)}_g = \widetilde{\Phi}^{(k)}_{\widetilde{g}}$. Therefore, there no loss of generality to assume g is symmetric.

We do not know how to write (A.1) for general $h,h_* \in \mathbb{H}\setminus\{\xi\circ x : \xi \in T_\mu\mathscr{P}_2(\mathbb{M})\}$. In some cases such as when $\widetilde{\Phi} = \widetilde{\Phi}^{(k)}_g$ for some smooth g, then (A.1) extends to $h,h_* \in \mathbb{H}\setminus\{\xi\circ x : \xi \in T_\mu\mathscr{P}_2(\mathbb{M})\}$. This can be checked by hand by writing the Taylor expansion of second order of

$$g\big(x(\omega_1)+h(\omega_1),\dots,x(\omega_k)+h(\omega_k)\big).$$

Another example is when

$$\Phi(\mu) = \theta\left(\frac{1}{2}\int_{\mathbb{M}}|q|^2\mu(dq)\right) \;\; \forall \mu \in \mathscr{P}_2(\mathbb{M}), \tag{A.2}$$

and so $\widetilde{\Phi}(x) = \theta\left(\frac{1}{2}\|x\|^2\right) \;\forall x \in \mathbb{H}$. Writing the second-order Taylor expansion, we have

$$\nabla\widetilde{\Phi}(x)(h) = \theta'\left(\frac{1}{2}\|x\|^2\right)(x,h)$$

and

$$\nabla^2\widetilde{\Phi}(x)(h,h) = \theta'\left(\frac{1}{2}\|x\|^2\right)\|h\|^2 + \theta''\left(\frac{1}{2}\|x\|^2\right)(x,h)^2 \quad \forall x,h \in \mathbb{H}. \tag{A.3}$$

We conclude

$$D_q\big(\nabla_w \Phi(\mu)\big) = \theta'\bigg(\frac{1}{2}\int_{\mathbb{M}} |q|^2 \mu(dq)\bigg) I_d \quad \forall \mu \in \mathscr{P}_2(\mathbb{M}) \tag{A.4}$$

and

$$\nabla^2_{ww}\Phi(\mu)(q,b) = \theta''\bigg(\frac{1}{2}\int_{\mathbb{M}} |q|^2 \mu(dq)\bigg) q\otimes b \quad \forall \mu \in \mathscr{P}_2(\mathbb{M})\ \ \forall q,b \in \operatorname{spt}(\mu).$$

Thus, when Φ is of the form (A.2), (A.1) continues to hold for all $h, h_* \in \mathbb{H}$. Note that the expression in (A.4) is constant on $\mathbb{M}$. In fact, we shall see this is not a coincidence, which is the aim of these notes.

Our goal is to show that if $\widetilde{\Phi} \in C^{2,\alpha}_{\text{loc}}\big(\mathbb{H}\big)$, then $D_q\big(\nabla_w \Phi(\mu)\big)$ must be a constant function on $\operatorname{spt}(\mu)$. This will allow us to make inference about the dimension of $C^{2,\alpha}_{\text{loc}}\big(\mathbb{H}\big) \cap \{\widetilde{\Phi}^{(k)}_g\}$ for any natural number k. In conclusion, the set of $\widetilde{\Phi} \in C^{2,\alpha}_{\text{loc}}\big(\mathbb{H}\big)$ may be too small in some sense and a theory of mean field games for functions $\widetilde{\Phi} \in C^{2,\alpha}_{\text{loc}}\big(\mathbb{H}\big)$ may be too restrictive. Hence, $C^{2,\alpha,w}_{\text{loc}}\big(\mathscr{P}_2(\mathbb{M})\big)$ (cf. def. 3.13) is a better space for a general theory.

LEMMA A.1. *Let* $\alpha \in (0,1]$ *and assume* $\widetilde{\Phi} \in C^{2,\alpha}_{\text{loc}}(\mathbb{H})$ *is rearrangement invariant so that it is the lift of a function* Φ*. If* (A.1) *holds for all* $h, h_* \in \mathbb{H}$*, then* $D_q\big(\nabla_w \Phi(\mu)\big)$ *is constant function on* $\operatorname{spt}(\mu)$.

PROOF. Let $x \in \mathbb{H}$ and let μ be the law of x. Fix an open ball $\mathbb{B} \subset \mathbb{H}$ that contains x and choose $\kappa_{\mathbb{B}} > 0$ such that

$$\Big(\nabla^2\widetilde{\Phi}(x) - \nabla^2\widetilde{\Phi}(y)\Big)(h,h_*) \le \kappa_{\mathbb{B}}\|x-y\|^{\alpha} \tag{A.5}$$

for all $y \in \mathbb{B}$ and all $h, h_* \in \mathbb{H}$ such that $\|h\|, \|h_*\| \le 1$.

Let $\varrho \in C^\infty_c(\mathbb{M})$ be a probability density function whose support is the unit ball in $\mathbb{R}^d$. For $z, z_* \in \mathbb{R}^d$ unit vectors and for $\omega, o \in \Omega$, we set

$$h^\epsilon = z\sqrt{\varrho^o_\epsilon}, \quad h^\epsilon_* = z_*\sqrt{\varrho^o_\epsilon}, \quad \varrho^o_\epsilon(\omega) := \epsilon^{-d}\varrho\Big(\frac{\omega - o}{\epsilon}\Big).$$

Let $y \in \mathbb{H}$ have the same law with x. We have

$$\begin{aligned}
&\big(\nabla^2\widetilde{\Phi}(y) - \nabla^2\widetilde{\Phi}(x)\big)(h^\epsilon, h^\epsilon_*)\\
&\quad= \int_\Omega \Big(D_q\big(\nabla_w\Phi(\mu)\big)\big(y(\omega)\big) - D_q\big(\nabla_w\Phi(\mu)\big)\big(x(\omega)\big)\Big)h(\omega)\cdot h_*(\omega)\\
&\qquad+ \int_{\Omega^2}\Big(\nabla^2_{ww}\Phi(\mu)\big(y(\omega), y(\omega_*)\big) - \nabla^2_{ww}\Phi(\mu)\big(x(\omega), x(\omega_*)\big)\Big)h(\omega)\cdot h_*(\omega_*)d\omega\, d\omega_*\\
&\quad= \int_\Omega \Big(D_q\big(\nabla_w\Phi(\mu)\big)\big(y(o+\epsilon a)\big) - D_q\big(\nabla_w\Phi(\mu)\big)\big(x(o+\epsilon a)\big)\Big)z\cdot z_*\varrho(a)da\\
&\quad+ \epsilon^d\int_{\Omega^2}\nabla^2_{ww}\Phi(\mu)\big(y(o+\epsilon a), y(o+\epsilon b)\big)z\cdot z_*\sqrt{\varrho(a)\varrho(b)}da\, db\\
&\quad- \epsilon^d\int_{\Omega^2}\nabla^2_{ww}\Phi(\mu)\big(x(o+\epsilon a), x(o+\epsilon b)\big)z\cdot z_*.\sqrt{\varrho(a)\varrho(b)}da\, db.
\end{aligned} \tag{A.6}$$

Since $\widetilde{\Phi} \in C^{1,1}(\mathbb{B})$, $\nabla^2_{ww}\Phi(\mu)$ is bounded, we use (A.6) to obtain that if o is a Lebesgue point for $\big(D_q\nabla_w\Phi(\mu)\big)\circ y$ and $\big(D_q\nabla_w\Phi(\mu)\big)\circ x$, then

$$\lim_{\epsilon\to 0}\Big(\nabla^2\Phi(y)-\nabla^2\Phi(x)\Big)(h^\epsilon, h^\epsilon_*) = \Big(D_q\big(\nabla_w\Phi(\mu)\big)\big(y(o)\big) - D_q\big(\nabla_w\Phi(\mu)\big)\big(x(o)\big)\Big)z\cdot z_*.$$

This, together with (A.5) implies that if $y\in\mathbb{B}$ then

$$|D_q\big(\nabla_w\Phi(\mu)\big)\circ y(o) - D_q\big(\nabla_w\Phi(\mu)\big)\circ x(o)| \le \kappa_{\mathbb{B}}\|x-y\|^\alpha. \tag{A.7}$$

In the spirit of the proof of Lemma 3.11, set

$$\Omega_0 := \big\{\omega\in\Omega \mid \omega \text{ is a Lebesgue point for } x, D_q\nabla_w\Phi(\mu)\circ x\big\}\cap x^{-1}(\operatorname{spt}(\mu)).$$

Note that Ω_0 is a set of full measure in Ω and so, $x(\Omega_0)$ is a set of full μ-measure. In fact, we do not know that $x(\Omega_0)$ is Borel, but we can find a Borel set $A\subset x(\Omega_0)$ of full μ-measure.

Assume in the sequel that $o\in A$ and set $q_1 := x(o)$. Assume we can find $\bar{o}\in A$ such that $q_2 = x(\bar{o}) \neq q_1$. Let $r>0$ small such that $B_r(o)\cap B_r(\bar{o}) = \varnothing$. Set

$$S_r(\omega) := \begin{cases}\omega, & \text{if } \omega\in\Omega\setminus\big(B_r(o)\cup B_r(\bar{o})\big),\\ \omega - o + \bar{o}, & \text{if } \omega\in B_r(o),\\ \omega - \bar{o} + o, & \text{if } \omega\in B_r(\bar{o}).\end{cases}$$

Since S_r preserves Lebesgue measure, x and $y := x\circ S_r$ have the same law μ. We notice

$$\|x-y\|^2 = 2\int_{B_r(o)} |x(\omega) - x(\omega+\bar{o}-o)|^2\,dz$$

and so, for r small enough, $y\in\mathbb{B}$. By (A.7) implies

$$\begin{aligned}\Big|D_q\big(\nabla_w\Phi(\mu)\big)(q_2) - D_q\big(\nabla_w\Phi(\mu)\big)(q_1)\Big| &= \Big|D_q\big(\nabla_w\Phi(\mu)\big)\circ y(o) - D_q\big(\nabla_w\Phi(\mu)\big)\circ x(o)\Big|\\ &\le \kappa_{\mathbb{B}}\bigg(2\int_{B_r(o)}|x(z)-x(z+\bar{o}-o)|^2dz\bigg)^{\frac{\alpha}{2}}.\end{aligned}$$

We let r tend to 0 to conclude the proof. □

PROPOSITION A.2. *For any $\alpha\in(0,1]$ and $k\in\mathbb{N}$, we have*

$$\dim\Big(C^{2,\alpha}_{\mathrm{loc}}(\mathbb{H})\cap\big\{\widetilde{\Phi}_g : g\in C^{2,\alpha}_{\mathrm{loc}}(\mathbb{M}^k),\ \ \|D^2g\|_{L^\infty}<\infty\big\}\Big)<\infty.$$

PROOF. We aim to use Lemma A.1, since this asserts that $D_q\nabla_w\Phi_g(\mu)(q)$ is a constant matrix $C(\mu)$ which depends only on μ.

In particular, in the case of $k=1$, we have $D_q\nabla_w\Phi_g(\mu)(q) = D^2g(q)$, and this being constant implies that g is a polynomial of degree 2, so the claim follows.

For $k \in \mathbb{N}$ general we have

$$D_q \nabla_w \Phi_g(\mu)(q) = \int_{\mathbb{M}^{k-1}} D^2_{q_1 q_1} g(q, q_2, \dots, q_k)\mu(dq_2)\dots\mu(dq_k) + \cdots + \int_{\mathbb{M}^{k-1}} D^2_{q_k q_k} g(q_1, q_2, \dots, q_{k-1}, q)\mu(dq_1)\dots\mu(dq_{k-1}).$$

In fact, by [21]

$$\begin{aligned} C(\mu) &= D_q \nabla_w \Phi_g(\mu)(q) \\ &= k\int_{\mathbb{M}^{k-1}} D^2_{q_1 q_1} g(q, q_2, \dots, q_k)\mu(dq_2)\dots\mu(dq_k) \qquad \text{(A.8)} \\ &= \cdots \\ &= k\int_{\mathbb{M}^{k-1}} D^2_{q_k q_k} g(q_1, q_2, \dots, q_{k-1}, q)\mu(dq_1)\dots\mu(dq_{k-1}), \qquad \text{(A.9)} \end{aligned}$$

μ-a.e. For simplicity, let us set $k = 2$ (the proof of the result for general $k \in \mathbb{N}$ follows along the same lines). Let $a \in \mathbb{M}$ and $\varrho \in C_b(\mathbb{M})$ be a fully supported probability measure and let ϱ_ϵ be its standard rescaled function. The measures $\varrho_\epsilon(q-a)$ have the whole $\mathbb{M}$ as their support, and so

$$\int_{\mathbb{M}} D^2_{q_1 q_1} g(q, q_2)\varrho_\epsilon(q_2 - a)dq_2 = \int_{\mathbb{M}} D^2_{q_1 q_1} g(\overline{q}, q_2)\varrho_\epsilon(q_2 - a)dq_2 \quad \forall q, \overline{q} \in \mathbb{M}.$$

Letting ϵ tend to 0 we conclude

$$D^2_{q_1 q_1} g(q, a) = D^2_{q_1 q_1} g(\overline{q}, a).$$

In fact,

$$D^2_{q_1 q_1} g(q, a) = D^2_{q_1 q_1} g(\overline{q}, a) = D^2_{q_2 q_2} g(a, q) = D^2_{q_2 q_2} g(a, \overline{q}) = C(a).$$

>From these arguments, one can conclude that both $q_1 \mapsto D^2_{q_1 q_1} g(q_1, a)$ and $q_2 \mapsto D^2_{q_2 q_2} g(a, q_2)$ are constants for all $a \in \mathbb{M}$; therefore the $q_1 \mapsto g(q_1, a)$ and $q_2 \mapsto g(a, q_2)$ are polynomials of degree at most 2 for all $a \in \mathbb{M}$. By an adaptation of the result of [18] we conclude that g needs to be a polynomial of degree at most 2. The result follows. □

COROLLARY A.3. *Similarly, for the example in* (A.2)*, if* $\widetilde{\Phi} \in C^{2,\alpha}_{\text{loc}}(\mathbb{H})$*, then by Lemma* A.1 *and* (A.4) *we have that* $\theta(t) = c_0 t$ *for some* $c_0 \in \mathbb{R}$.

The result from Proposition A.2 in case of $k = 1$ is the consequence of the proposition below, where we show that assuming even only C^2 regularity (instead of $C^{2,\alpha}$) for functionals on $\mathbb{H}$ having local representations might result in trivialities.

PROPOSITION A.4.

$$C^2(\mathbb{H}) \cap \left\{\widetilde{\Phi}_g : g \in C^3(\mathbb{M}),\ \|D^2 g\|_{L^\infty} < \infty,\ \|D^3 g\|_{L^\infty} < \infty,\ D^3 g \not\equiv 0\right\} = \varnothing,$$

and so

$$C^2(\mathbb{H}) \cap \left\{\widetilde{\Phi}_g : g \in C^3(\mathbb{M}),\ \|D^2 g\|_{L^\infty} < \infty,\ \|D^3 g\|_{L^\infty} < \infty\right\}$$

is a finite-dimensional space.

PROOF. For simplicity, let us suppose that $d = 1$ and so $\Omega = [0,1]$. The result in higher dimensions follows from similar arguments.

For $x, y \in \mathbb{H}$ we can write the following expansion for $\widetilde{\Phi}_g$:

(A.10)
$$\begin{aligned}
&\int_\Omega g(y(\omega))d\omega - \int_\Omega g(x(\omega))d\omega - \int_\Omega g'(x(\omega))(y(\omega)-x(\omega))d\omega\\
&\quad - \frac{1}{2}\int_\Omega g''(x(\omega))(y(\omega)-x(\omega))^2 d\omega\\
&\qquad = \int_\Omega \int_0^1\int_0^1\int_0^1 t^2 s g'''(x(\omega)+ts\tau(y(\omega)-x(\omega)))(y(\omega)-x(\omega))^3\, d\tau\, ds\, dt\, d\omega.
\end{aligned}$$

By the assumptions on g''', there exist constants c_0, c_1, having the same sign, such that on a bounded open interval $c_0 \le g''' \le c_1$. Without loss of generality, let us suppose that this open interval is $(-1,1)$ and $0 < c_0 < c_1$.

CLAIM. The right-hand side of (A.10) is not of order $o(\|x-y\|^2)$ when $x \equiv 0$.

PROOF OF THE CLAIM. Let $x(\omega) = 0$ and $y_n(\omega) = \omega^n$ for $\omega \in \Omega$ and $n \in \mathbb{N}$. Then clearly $\|y_n\|^2 = \frac{1}{2n+1} \to 0$ as $n \to +\infty$. We write the previous expansion for y_n and x. In particular, the remainder satisfies

(A.11)
$$\begin{aligned}
\frac{c_0}{6}\int_\Omega y_n^3(\omega)d\omega &\le \int_\Omega\int_0^1\int_0^1\int_0^1 t^2 s g'''(ts\tau y_n(\omega))y_n^3(\omega)d\tau\, ds\, dt\, d\omega\\
&\le \frac{c_1}{6}\int_\Omega y_n^3(\omega)d\omega.
\end{aligned}$$

We easily find $\int_0^1 y_n^3(\omega)d\omega = \frac{1}{3n+1}$. Therefore dividing (A.11) by $\|y_n\|^2$ and taking $n \to +\infty$, we find

$$\frac{2c_0}{18} \le \lim_{n\to+\infty}\frac{1}{\|y_n\|^2}\int_\Omega\int_0^1\int_0^1\int_0^1 t^2 s g'''(ts\tau y_n(\omega))y_n^3(\omega)d\tau\, ds\, dt\, d\omega \le \frac{2c_1}{18}.$$

The claim follows and so does the thesis of the proposition. □

Appendix B Convexity Versus Displacement Convexity

B.1 Displacement convexity versus classical convexity

Using the terminology of [14], in this section will consider weakly Fréchet continuously differentiable functions $\mathscr{V} : \mathscr{P}_2(\mathbb{M}) \to \mathbb{R}$ and denote their weak Fréchet differentials as $\frac{\delta\mathscr{V}}{\delta\mu} : \mathbb{R}^d \times \mathscr{P}_2(\mathbb{M}) \to \mathbb{R}$. Let $\phi_1, \phi \in C^2(\mathbb{M})$ be functions of bounded second derivatives such that ϕ_1 is even. Set

$$\mathscr{V}_1(\mu) := \frac{1}{2}\int_{\mathbb{R}^d}\phi_1 * \mu(q)\mu(dq), \quad \mu \in \mathscr{P}_2(\mathbb{M})$$

and

$$\mathscr{V}(\mu) := \mathscr{V}_1(\mu) + \int_{\mathbb{R}^d} \phi(q)\mu(dq), \quad \mu \in \mathscr{P}_2(\mathbb{M}).$$

Remark B.1. Recall from [14] that $\frac{\delta \mathscr{V}}{\delta \mu}$ is monotone if and only if $\mathscr{V}$ is convex in the classical sense. Furthermore, the function $\mathscr{V}_1$ is a twice weakly Fréchet continuously differentiable function, and

$$\frac{\delta \mathscr{V}_1}{\delta \mu}(q,\mu) = (\phi_1 * \mu)(q), \quad \frac{\delta \mathscr{V}}{\delta \mu}(q,\mu) = (\phi_1 * \mu)(q) + \phi(q),$$

and

$$\frac{\delta^2 \mathscr{V}_1}{\delta \mu^2}(q,y,\mu) = \frac{\delta^2 \mathscr{V}}{\delta \mu^2}(q,y,\mu) = \phi_1(q-y).$$

LEMMA B.2. *If we further assume $\phi_1 \in L^1(\mathbb{M})$, then $\frac{\delta \mathscr{V}}{\delta \mu}$ is monotone if and only if the Fourier transform ϕ_1 is nonnegative.*

PROOF. Denote the Fourier transform of ϕ_1 by $\widehat{\phi}_1$. Note that for any $f \in L^2(\mathbb{M})$, by Young's inequality we have $\phi_1 * f \in L^2(\mathbb{M})$, and so $f(\phi_1 * f) \in L^1(\mathbb{M})$. By the Riemann-Lebesgue lemma, $\widehat{\phi}_1 \in C_0(\mathbb{M})$. Furthermore, $\widehat{\phi}_1$ is even and has its range contained in the set of real numbers. By Remark B.1 $\delta\mathscr{V}/\delta\mu$ is monotone if and only if $\mathscr{V}_1$ is convex. Thus, using the expression of $\delta^2\mathscr{V}_1/\delta\mu^2$ in Remark B.1, we conclude that $\delta\mathscr{V}/\delta\mu$ is monotone if and only if for any $f \in C(\mathbb{M}) \cap L^2(\mathbb{M})$ such that $\int_{\mathbb{M}} f(q)dq = 0$ we have $0 \le \int_{\mathbb{R}^d} (\phi_1 * f)(q) f(q) dq$. Thanks to Plancherel theorem, $\delta\mathscr{V}/\delta\mu$ is monotone if and only if

$$0 \le \int_{\mathbb{R}^d} \widehat{\phi_1 * f}(\xi) \hat{f}^*(\xi) d\xi = \int_{\mathbb{R}^d} \widehat{\phi}_1(\xi) \hat{f}(\xi) \hat{f}^*(\xi) d\xi = \int_{\mathbb{R}^d} \widehat{\phi}_1(\xi) |\hat{f}(\xi)|^2 \, d\xi.$$

This concludes the proof of the lemma. □

LEMMA B.3. *Assume $\lambda > 0$, $\lambda_1 \in (-\lambda/2, \lambda/2)$, ϕ is λ-convex, and ϕ_1 is λ_1-convex. Then*

(i) *$\mathscr{V}$ is κ-displacement convex, hence displacement convex, where $\kappa := \lambda - 2|\lambda_1| > 0$.*
(ii) *If we further assume ϕ_1 is nonnegative, $\phi_1 \equiv 1$ on the unit ball, and $\phi_1 \equiv 0$ outside the ball of radius 2 centered at the origin, then $\mathscr{V}$ fails to be convex in the classical sense.*

PROOF. (i) As above, denote the Fourier transform of ϕ_1 as $\widehat{\phi}_1$. Let us consider $\sigma \in AC_2(0,1;\mathscr{P}_2(\mathbb{M}))$ to be a geodesic such that its velocity v is not identically null. Since $\|v_t\|_{\sigma_t}$ is independent of t, it is then positive. We have

$$\begin{aligned}\frac{d^2}{dt^2}\mathscr{V}(\sigma_t) &= \int_{\mathbb{M}} D^2\phi(q)v_t(q)\cdot v_t(q)\sigma_t(dq)\\ &\quad+\int_{\mathbb{M}^2} D^2\phi_1(q-w)v_t(q)\cdot v_t(q)\sigma_t(dq)\sigma_t(dw)\\ &\quad+\int_{\mathbb{M}^2} D^2\phi_1(q-w)v_t(q)\cdot v_t(w)\sigma_t(dq)\sigma_t(dw)\\ &\geq \lambda\|v_t\|^2_{\sigma_t}+\lambda_1\|v_t\|^2_{\sigma_t}-|\lambda_1|\|v_t\|^2_{\sigma_t}\geq \kappa\|v_t\|^2_{\sigma_t}.\end{aligned}$$

This completes the verification of (i).

(ii) Since ϕ_1 is even, the range of its Fourier transform is contained in the set of real numbers (including negative ones). Assume on the contrary that the range of $\widehat{\phi}_1$ is contained in $[0,\infty)$. By the Fourier inversion theorem we have for $x \in \mathbb{M}$,

$$|\phi_1(x)| = \left|\int_{\mathbb{M}} \widehat{\phi}_1(\xi)e^{2\pi i x\cdot\xi}d\xi\right| \leq \int_{\mathbb{M}} |\widehat{\phi}_1(\xi)|d\xi = \int_{\mathbb{M}} \widehat{\phi}_1(\xi)d\xi = \phi_1(0).$$

Since $\phi_1(x) \equiv 1 = \phi_1(0)$ on $B_1(0)$, the ball of center 0 and radius 1, we must have

$$\widehat{\phi}_1(\xi)\cos(2\pi x\cdot\xi) \equiv |\widehat{\phi}_1(\xi)| \equiv \widehat{\phi}_1(\xi) \quad \forall (x,\xi)\in B_1(0)\times\mathbb{M}. \tag{B.1}$$

Since ϕ_1 is not the null function, $\widehat{\phi}_1$ cannot be the null function. Choose ξ_0 such that $\widehat{\phi}_1(\xi_0) > 0$, and since $\widehat{\phi}_1$ is continuous, we can assume without loss of generality that $\xi_0 \neq 0$. By (B.1), $\cos(2\pi x\cdot\xi_0) = 1$ for all $x \in B_1(0)$, which yields a contradiction. One concludes the proof of (ii) by Lemma B.2. □

B.2 Convexity versus displacement convexity of the action

Here we would like to emphasize the fact that imposing the joint convexity assumption on the Lagrangian action, as in (H7), comes as a natural assumption for displacement convex potential mean field games, which are considered in this manuscript. We compare this to the more standard monotonicity assumption in potential MFG.

Assume $L, H \in C^1(\mathbb{M}\times\mathbb{R}^d)$ are such that $H(q,\cdot)$ and $L(q,\cdot)$ are Legendre transforms of each other. We consider the actions

$$\mathscr{A}_0^T(\sigma,v) := \int_0^T \left(\int_{\mathbb{M}} L(q,v_t(q))\sigma_t(dq) + \mathscr{F}(\sigma_t)\right)dt$$

over the set of pairs (σ, v) such that

$$\partial_t\sigma + \nabla\cdot(\sigma v) = 0 \quad \mathscr{D}'\big((0,T)\times\mathbb{M}\big). \tag{B.2}$$

Recall that if we set $\nabla_q f(q,\mu) := \nabla_w F(\mu)(q)$ then f monotone means $\mathscr{F}$ is convex.

We can rewrite $\mathscr{A}_0^T(\sigma, v)$ in terms of the momentum by setting

$$\bar{\mathscr{A}}_0^T(\sigma,\eta) := \int_0^T \left(\int_{\mathbb{M}} L\Big(q, \frac{d\eta_t}{d\sigma_t}(q)\Big)\sigma_t(dq) + \mathscr{F}(\sigma_t) \right) dt$$

over the set of pairs (σ, η) such that $|\eta_t| \ll \sigma_t$ and

$$\partial_t \sigma + \nabla \cdot \eta = 0 \quad \mathscr{D}'\big((0,T)\times\mathbb{M}\big). \tag{B.3}$$

In fact, for each $q \in \mathbb{M}$ we introduce the function $\bar{L}_q : \mathbb{R}\times\mathbb{R}^d \to \mathbb{R}\cup\{\infty\}$, defined as

$$\bar{L}_q(\rho, e) := \begin{cases} \rho L(q, \frac{e}{\rho}) & \text{if } \rho > 0, \\ 0 & \text{if } \rho = 0, e = \vec{0}, \\ +\infty & \text{otherwise.} \end{cases} \tag{B.4}$$

Here $\vec{0} := (0, \dots, 0)$. Since L_q is homogeneous of degree 1, whenever μ is a probability measure and $\xi_1, \dots, \xi_d$ are signed Borel measures, the following function is well-defined:

$$(\mu,\xi) \mapsto A(\mu,\xi) := \begin{cases} \int_{\mathbb{M}} \bar{L}_q(\mu(dq), d\xi) & \text{if } |\xi| \ll \mu, \\ +\infty & \text{if } |\xi| \not\ll \mu. \end{cases}$$

Let $\mathscr{C}$ be the set of (σ,η) such that $\sigma \in AC_2(0,T;\mathscr{P}_2(\mathbb{M}))$ and $t \mapsto \eta_t \in \mathscr{M}(\mathbb{M})\times\cdots\times\mathscr{M}(\mathbb{M})$ is a Borel path of vector fields such that each one of its d components is a signed Borel measure on $\mathbb{M}$ and

$$\partial_t \sigma + \nabla \cdot \eta = 0, \quad \mathscr{D}'\big((0,T)\times\mathbb{M}\big). \tag{B.5}$$

We can now extend the definition of $\bar{\mathscr{A}}_0^T$ over $\mathscr{C}$ to obtain

$$\bar{\mathscr{A}}_0^T(\sigma,\eta) := \int_0^T \big(A(\sigma_t,\eta_t) + \mathscr{F}(\sigma_t)\big)dt.$$

LEMMA B.4. *If $\mathscr{F}$ is convex on $\mathscr{P}_2(\mathbb{M})$, then $(\mu,\xi)\mapsto A(\mu,\xi)+\mathscr{F}(\mu)$ is convex* (*we do not assume L is jointly convex*).

PROOF. It suffices to show that $(\mu,\xi)\mapsto A(\mu,\xi)$ is convex. The proof of this well-known fact can be found in [43, prop. 5.18]. □

Remark B.5. (i) Note that the classical theory of potential mean field games in which it is assumed that f is monotone and $L, H \in C^1(\mathbb{M}\times\mathbb{R}^d)$ are such that $H(q,\cdot)$ and $L(q,\cdot)$ are Legendre transforms of each other, This ensures that $(\mu,\xi)\mapsto A(\mu,\xi)+\mathscr{F}(\mu)$ is a convex function. Therefore, if we extend the definition of $\bar{\mathscr{A}}_0^T$ to obtain

$$\bar{\mathscr{A}}_0^T(\sigma,\eta) := \int_0^T \big(A(\sigma_t,\eta_t) + \mathscr{F}(\sigma_t)\big)dt$$

over $\mathscr{C}$, the action $\bar{\mathscr{A}}_0^T$ is a convex function in the variables (σ,η).

(ii) When replacing the assumption of convexity on the action by an assumption of displacement convexity, as it is done in this manuscipt, it seems natural to impose that $\mathscr{A}_0^T(\sigma, v)$ is displacement convex on the set of pairs (σ, v) satisfying (B.2). This means that

$$\mathbb{H}\times\mathbb{H} \ni (X,V) \mapsto \int_\Omega L(X,V)d\omega + \widetilde{\mathscr{F}}(X) \quad \text{is convex},$$

and thus the Lagrangian L is assumed to be jointly convex on $\mathbb{M}\times\mathbb{R}^d$.

B.3 Convexity of $f(\cdot,\mu)$ is a consequence of the displacement convexity of $\mathscr{F}$

To study the scalar master equation, among others we have imposed the assumptions (4.7) and (H10) on the functions f and $\mathscr{F}$. As we have detailed in the previous couple of lines, in our setting it is natural for the Lagrangian L to impose joint λ-convexity, and we impose that $\mathscr{F}$ is displacement λ-convex. We show below that in this sense, imposing (4.7), i.e. that $f(\cdot,\mu)$ is λ-convex, is also natural, and it is a consequence of the displacement λ-convexity of $\mathscr{F}$.

PROPOSITION B.6. *Let $\mathscr{F}: \mathscr{P}_2(\mathbb{M}) \to \mathbb{R}$ and $f: \mathbb{M}\times\mathscr{P}_2(\mathbb{M}) \to \mathbb{R}$ be of class C^2 such that they are related via* (H10). *We assume that $\mathscr{F}$ is displacement λ-convex*; $\mathbb{M}\times\mathscr{P}_2(\mathbb{M}) \ni (q,\mu) \mapsto D_q\nabla_w\mathscr{F}(\mu)(q) = D^2_{qq}f(q,\mu)$ *is continuous and that for any $\mathscr{K}\subset\mathscr{P}_2(\mathbb{M})$ compact, there exists $C = C(\mathscr{K}) > 0$ such that $|D^2_{ww}\mathscr{F}(\mu)(q_1,q_2)| \le C$ for any $\mu\in\mathscr{K}$ and for any $q_1,q_2\in$* spt(μ).

Then, for any $\mu\in\mathscr{P}_2(\mathbb{M})$, the function $\mathrm{spt}(\mu)\ni q\mapsto f(q,\mu)$ *is λ-convex, i.e.,*

$$D^2_{qq}f(x,\mu) \ge \lambda I_d \quad \forall\, q\in \mathrm{spt}(\mu).$$

PROOF. Let $m\in\mathbb{N}$ and define $F^{(m)}: (\mathbb{M})^m \to \mathbb{R}$ as

$$F^{(m)}(q_1,\ldots,q_m) := \mathscr{F}(\mu_q^{(m)}).$$

By the assumptions on $\mathscr{F}$, we have that $F^{(m)}$ is twice differentiable on $(\mathbb{M})^m$, and by Lemma 3.6, it is $\frac{\lambda}{m}$-convex on $(\mathbb{M})^m$. This means in particular that

$$D^2F^{(m)}(q_1,\ldots,q_m) \ge \frac{\lambda}{m}I_{md} \quad \forall (q_1,\ldots,q_m)\in(\mathbb{M})^m$$

or equivalently

$$a^\top D^2F^{(m)}(q_1,\ldots,q_m)a \ge \frac{\lambda}{m}|a|^2_{md} \quad \forall a\in\mathbb{M}^m, (q_1,\ldots,q_m)\in(\mathbb{M})^m,$$

where $|\cdot|_{md}$ stands for the standard Euclidean norm on $\mathbb{M}^m$. For $i\in\{1,\ldots,m\}$, let us choose the vector $a\in\mathbb{M}^m$ such that its coordinates between the indices $d(i-1)+1$ and di are not all zero, while all the others are zero. Then, the previous inequality implies that

$$D^2_{q_iq_i}F^{(m)}(q_1,\ldots,q_m) \ge \frac{\lambda}{m}I_d \quad \forall (q_1,\ldots,q_m)\in(\mathbb{M})^m. \tag{B.6}$$

We also have (see, for instance, in [21, remark 3.5(iv)]) that

$$mD^2_{q_iq_i}F^{(m)}(q_1,\dots,q_m) = D_q\nabla_w\mathscr{F}\big(\mu_q^{(m)}\big)(q_i) + \frac{1}{m}\nabla^2_{ww}\mathscr{F}\big(\mu_q^{(m)}\big)(q_i,q_i),$$

$\forall m\in\mathbb{N},\ \{q_1,\dots,q_m\}\subseteq \mathrm{spt}(\mu_q^m)$.

Let $b\in\mathbb{M}$. By (B.6), one has that

$$b^\top D_q\nabla_w\mathscr{F}(\mu_q^{(m)})(q_i)b + \frac{1}{m}b^\top\nabla^2_{ww}\mathscr{F}(\mu_q^{(m)})(q_i,q_i)b \ge \lambda|b|^2_d,$$

$\forall m\in\mathbb{N},\ \{q_1,\dots,q_m\}\subseteq \mathrm{spt}(\mu_q^m)$. Now let us fix $\mu\in\mathscr{P}_2(\mathbb{M})$ and $q_1\in\mathrm{spt}(\mu)$. For $m\ge 2$ a natural number, let $q_i\in\mathrm{spt}(\mu)$, $i\in\{2,\dots,m\}$, and let us build $\mu_q^{(m)} := \sum_{i=1}^m \delta_{q_i}$, as an approximation of μ.

We have that

$$b^\top D_q\nabla_w\mathscr{F}\big(\mu_q^{(m)}\big)(q_1)b + \frac{1}{m}b^\top\nabla^2_{ww}\mathscr{F}(\mu_1^{(m)})(q_1,q_1)b \ge \lambda|b|^2_d.$$

Since $\mathscr{K} := \{\mu_q^{(m)} : m\in\mathbb{N}\}\cup\{\mu\}$ is a compact set, by the assumptions we have that $\nabla^2_{ww}\mathscr{F}(\mu_q^{(m)})(q_1,q_1)$ is uniformly bounded by a constant $C = C(\mathscr{K}) > 0$ independently of m. By the continuity of $D_q\nabla_w\mathscr{F}$, one can pass to the limit in the previous inequality to obtain

$$b^\top D_q\nabla_w\mathscr{F}(\mu)(q_1)b \ge \lambda|b|^2_d,$$

and equivalently

$$b^\top D^2_{qq}f(q_1,\mu)b \ge \lambda|b|^2_d.$$

By the arbitrariness of $b\in\mathbb{R}^d$ and $q_1\in\mathrm{spt}(\mu)$, the thesis of the proposition follows. □

B.4 Failure of smoothness of solutions to the Hamilton–Jacobi equation for monotone initial data

It is well-known in the theory of Hamilton–Jacobi equations on finite-dimensional spaces that typically one cannot expect global existence of smooth solutions. This led to the development of the notion of viscosity solution by Crandall–Lions and Evans. We emphasize below that this phenomenon of existence of nonsmooth solutions to Hamilton–Jacobi equations is also present on $\mathscr{P}_2(\mathbb{M})$.

Let us consider $d=1$. Let $L:\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ and $\phi:\mathbb{R}\to\mathbb{R}$ be defined as

$$L(q,v) := \frac{|v|^2}{2},\quad \phi(q) := -\sqrt{1+q^2}.$$

Set

$$\mathscr{U}_*(\mu) := \int_{\mathbb{R}}\phi(q)\mu(dq),\ \ u_*(q,\mu)=\phi(q),\quad \mathscr{L}(\mu,\xi) := \int_{\mathbb{R}} L(q,\xi(q))\mu(dq).$$

Note that $\mathscr{U}_*$ is convex, and so u_* is monotone.

Let $\mathscr{U} : [0,\infty)\times\mathscr{P}_2(\mathbb{R})$ be the unique viscosity solution to the Hamilton–Jacobi equation

$$\partial_t \mathscr{U} + \frac{1}{2}\int_{\mathbb{R}} |\nabla_w \mathscr{U}|^2 \mu(dq) = 0, \quad \mathscr{U}(0,\cdot) = \mathscr{U}_*. \tag{B.7}$$

Assume on the contrary that $\mathscr{U}$ is of class C^1. Then $\mathscr{U}$ must satisfy (B.7) pointwise, and so its restriction defined as

$$u(t,q) = \mathscr{U}(t,\delta_q)$$

must be a C^1 function satisfying

$$\partial_t u + \frac{1}{2}|\partial_q u|^2 = 0, \quad u(0,\cdot) = \phi. \tag{B.8}$$

Thus,

$$u(t,q) = \min_y \left\{ \frac{|y-q|^2}{2t} + \phi(y) : y \in \mathbb{R} \right\}. \tag{B.9}$$

Given q, the minimum in (B.9) is attained by y such that

$$\frac{y-q}{t} - \frac{y}{\sqrt{1+y^2}} = 0. \tag{B.10}$$

When $q = 0$, (B.10) has three solutions that are

$$y_0 = 0, \quad y_1 = \sqrt{t^2-1}, \quad y_2 = -\sqrt{t^2-1}.$$

They produce in (B.9) the values

$$-1 \quad \text{and} \quad -\frac{t}{2} - \frac{1}{2t}.$$

Therefore for $t > 1$, we have

$$u(t,0) = -\frac{t}{2} - \frac{1}{2t}.$$

Since for $i \in \{1,2\}$ we have

$$u(t,q) - u(t,0) \le \frac{|y_i - q|^2}{2t} + \phi(y_i) - \left(\frac{|y_i|^2}{2t} + \phi(y_i) \right) = \frac{-y_i \cdot q}{t} + \frac{|q|^2}{2t},$$

$\pm y_i/t = \pm\frac{\sqrt{t^2-1}}{t}$ belong to the superdifferential of $u(t,\cdot)$ at $q = 0$. Thus, $u(t,\cdot)$ is not differentiable at 0.

Appendix C Hamiltonian Flows and Minimizers of the Lagrangian Action

Most of the results of this section are expected to be known in some communities. We include them here for the sake of completeness and because of a lack of a precise reference.

C.1 Hamiltonian flows on the Hilbert space

Throughout this subsection, we impose (H1)–(H6). Showing that the value function of our Hamilton–Jacobi equation is of class $C^{1,1}$ on the Hilbert space is the starting point before improving the regularity property via a discretization method. We underline that in Section 1.3, using 'direct techniques' relying on the convexity of the Lagrangian action, we have shown already that the value function $\widetilde{\mathscr{U}}$ is of class $C^{1,1}_{\text{loc}}$. In this section, we discuss the regularity properties of the infinite-dimensional Hamiltonian flow (0.5), which could also be transfered to the value function.

Let $\widetilde{\xi}, \widetilde{\eta} : [0,\infty) \times \mathbb{H} \to \mathbb{H}$ be given by (0.6). Using (1.6) and the last inequality in Remark 1.1 (iii), we have

$$\left\| \left(\widetilde{\xi}(t,x), \widetilde{\eta}(t,x) \right) \right\| + 1 \leq \left(\sqrt{\|x\|^2 + \bar{\kappa}^2 (\|x\|^2 + 1)} + 1 \right) e^{\widetilde{\kappa} t} \tag{C.1}$$

for any $t > 0$ and $x \in \mathbb{H}$. We can formulate the following result.

PROPOSITION C.1. *Let $t \in (0,T)$, $\mu \in \mathscr{P}_2(\mathbb{M})$, and $q \in \mathbb{M}$. Suppose $(t_n)_n \subset [0,T]$ converges to t, $(\mu_n)_n \subset \mathscr{P}_2(\mathbb{M})$ converges to μ, and $(q_n)_n \subset \mathbb{M}$ converges to q. Then for every compact set $K \subset [0,t)$, we have*

$$\lim_{n\to\infty} \left\| S_s^{t_n}[\mu_n](q_n) - S_s^t[\mu](q) \right\|_{C(K)} = 0.$$

PROOF. To alleviate the notation, we set $\gamma^n(s) := S_s^{t_n}[\mu_n](q_n)$. It is characterized by the property that

$$u(t_n, q_n, \mu_n) = u_0\left(\gamma_0^n, \sigma_0^{t_n}[\mu_n]\right) + \int_0^{t_n} \left(L\left(\gamma_\tau^n, \dot{\gamma}_\tau^n\right) + f\left(\gamma_\tau^n, \sigma_\tau^{t_n}[\mu_n]\right) \right) d\tau, \tag{C.2}$$

with $\gamma_{t_n}^n = q_n$.

We assume without loss of generality that there exists $r > 0$ such that $(\mu_n)_n \subset \mathscr{B}_r$ and $(q_n) \subset B_r(0)$. By Remark C.6 (ii)

$$\left\{ \sigma_s^{t_n}[\mu_n] : n \in \mathbb{N}, s \in [0, t_n] \right\} \subset \mathscr{B}_{e_T(r)}.$$

In light of Remark 4.8 (ii), we may apply the Ascoli–Arzelà lemma to obtain a subsequence that we continue to denote as $(\gamma^n)_n$ which converges uniformly in $C([0, t-\delta]; \mathbb{M})$ for every $\delta \in (0,t)$. We have $\gamma \in W^{1,2}(0,t;\mathbb{M})$ and may also assume $(\gamma^n)_n$ converges weakly to γ in $W^{1,2}(0,t;\mathbb{M})$. We use (4.11) to obtain that $\gamma_t = q$. We would like to replace t_n by $t-\delta$. Since the integrand there is not known to be nonnegative, we use (H14) to write

$$\begin{aligned} &u(t_n, q_n, \mu_n) \\ &\quad = u_0\left(\gamma_0^n, \sigma_0^{t_n}[\mu_n]\right) + \int_0^{t_n} \theta\left(\sigma_\tau^{t_n}[\mu_n]\right)\left(|\gamma_\tau^n| + 1\right) d\tau \\ &\quad\quad \leq + \int_0^{t_n} \left(L\left(\gamma_\tau^n, \dot{\gamma}_\tau^n\right) + f\left(\gamma_\tau^n, \sigma_\tau^{t_n}[\mu_n]\right) - \theta(\sigma_\tau^{t_n}[\mu_n])(|\gamma_\tau^n| + 1) \right) d\tau. \end{aligned}$$

Thus, since all the integrands are nonnegative, we have

$$\begin{aligned}\liminf_{n\to\infty} u(t_n, q_n, \mu_n) &\\ \geq \liminf_{n\to\infty} u_0\big(\gamma_0^n, \sigma_0^{t_n}[\mu_n]\big) &+ \liminf_{n\to\infty} \int_0^{t-\delta} \theta(\sigma_\tau^{t_n}[\mu_n])(|\gamma_\tau^n| + 1)d\tau \\ + \liminf_{n\to\infty} \int_0^{t-\delta} \Big(L\big(\gamma_\tau^n, \dot\gamma_\tau^n\big) &+ f\big(\gamma_\tau^n, \sigma_\tau^{t_n}[\mu_n]\big) - \theta\big(\sigma_\tau^{t_n}[\mu_n]\big)\big(|\gamma_\tau^n| + 1\big)\Big)d\tau.\end{aligned}$$

We invoke the uniform convergence of $(\gamma^n)_n$, the pointwise convergence of the curves $(\sigma_\tau^{t_n}[\mu_n])_n$ provided in (C.7), and the convexity of the functions in (4.7) to conclude that

$$\begin{aligned}\liminf_{n\to\infty} u(t_n, q_n, \mu_n) \geq u_0\big(\gamma_0, \sigma_0^t[\mu]\big) &\\ + \int_0^{t-\delta} \Big(L\big(\gamma_\tau, \dot\gamma_\tau\big) &+ f\big(\gamma_\tau, \sigma_\tau^t[\mu]\big) - \theta(\sigma_\tau^t[\mu])(|\gamma_\tau| + 1)\Big)d\tau \\ &+ \int_0^{t-\delta} \theta\big(\sigma_\tau^t[\mu]\big)(|\gamma_\tau| + 1)d\tau.\end{aligned}$$

We let δ tend to 0 to conclude that

$$\begin{aligned}\liminf_{n\to\infty} u(t_n, q_n, \mu_n) &\geq u_0\big(\gamma_0, \sigma_0^t[\mu]\big) + \int_0^t \Big(L\big(\gamma_\tau, \dot\gamma_\tau\big) + f\big(\gamma_\tau, \sigma_\tau^t[\mu]\big)\Big) \\ &\geq u(t, q, \mu).\end{aligned}$$

Since Proposition 4.12 asserts that u is continuous, we infer

$$u(t, q, \mu) = u_0\big(\gamma_0, \sigma_0^t[\mu]\big) + \int_0^t \Big(L\big(\gamma_\tau, \dot\gamma_\tau\big) + f\big(\gamma_\tau, \sigma_\tau^t[\mu]\big)\Big)d\tau,$$

and so $\gamma_s \equiv S_s^t[\mu](q)$.

In conclusion, we have proven that every subsequence of $\big(S_s^t[\mu_n](q_n)\big)_n$ admits itself a subsequence which converges uniformly on every compact subset of $[0, t)$. This is enough to conclude the proof. □

PROPOSITION C.2. *Let $t > 0$. Then the following hold*:

(i) $\Sigma(t, \cdot)$ *given in* (0.5) *is of class* $C^{0,1}_{\mathrm{loc}}$.

(ii) $\tilde\xi_t : \mathbb{H} \to \mathbb{H}$ *is a bijection and its inverse is* $\tilde S_0^t$. *For each natural number* m, $\tilde\xi_t$ *is a homeomorphism* $\{M^q : q \in \mathbb{M}^m\}$ *onto* $\{M^q : q \in \mathbb{M}^m\}$. *This means* $S_s^{t,m} : \mathbb{M}^m \to \mathbb{M}^m$ *is a homeomorphism.*

(iii) $\tilde S_s^t \circ \tilde\xi_t = \tilde\xi_s$ *and* $\tilde P_s^t \circ \tilde\xi_t = \tilde\eta_s$ *for* $s \in [0, t]$.

(iv) *We have* $\nabla\tilde{\mathscr{U}}(t, \tilde\xi(t, \cdot)) = \tilde\eta(t, \cdot)$. *Furthermore, the vector field* B *in* (1.28) *is a velocity for the flow* $\tilde\xi$ *in the sense that* $\dot{\tilde\xi} = \nabla_b \tilde{\mathscr{H}}(\tilde\xi, \tilde\nabla\mathscr{U}(\cdot, \tilde\xi))$

Remark C.3. Although $\tilde\xi_t$ is a homeomorphism, let us underline that in Proposition C.2(ii) we state that the image of $\{M^q : q \in \mathbb{M}^m\}$ through $\tilde\xi_t$ is not an arbitrarily closed space but is exactly $\{M^q : q \in \mathbb{M}^m\}$. Such special vector spaces are mapped onto themselves. Otherwise, we would not be able to conclude that the finite-dimensional ODEs are restrictions of the infinite-dimensional ones.

PROOF OF PROPOSITION C.2. (i) Since $\widetilde{\mathscr{H}}$ is of class $C^{1,1}$, Σ is Lipschitz-continuous. Let κ^* be the Lipschitz constant of $\nabla\widetilde{\mathscr{H}}$. We have

$$\mathrm{Lip}(\Sigma(t,\cdot)) \leq \mathrm{Lip}(\Sigma(0,\cdot))e^{t\kappa^*}$$

for all $t > 0$. Here, $\mathrm{Lip}(\Sigma(t,\cdot))$ stands for the Lipschitz constant of $\Sigma(t,\cdot)$.

Since Σ satisfies (0.5), we conclude that Σ is of class $C^{0,1}_{\mathrm{loc}}$.

(ii) *Surjectivity.* Given any $x \in \mathbb{H}$. Set $z := \widetilde{S}^t_0[x]$ and define

$$\gamma(s) = \widetilde{S}^t_s[x], \quad b(s) = \nabla_a\widetilde{\mathscr{L}}(\gamma(s),\dot{\gamma}(s)).$$

We have that (γ, b) satisfies the same system of differential equations as $(\widetilde{\xi},\widetilde{\eta})$ on $(0,t)$. Furthermore, $\gamma(0) = z$ and

$$b(0) = \nabla_b\mathscr{L}(\widetilde{S}^t_s[x], \partial_s\widetilde{S}^t_s[x]|_{s=0}) = \nabla\widetilde{\mathscr{U}}_0(z).$$

Thus, (γ, b) have the same initial conditions as $(\widetilde{\xi},\widetilde{\eta})$. Hence, conclude that $\gamma \equiv \widetilde{\xi}(\cdot, z)$ on $[0,t]$. In particular, $x = \widetilde{S}^t_t[x] = \widetilde{\xi}(t,z) = \widetilde{\xi}(t,\widetilde{S}^t_0[x])$. This shows the surjectivity property.

Injectivity. The above shows that $\widetilde{S}^t_0$ is injective and $\widetilde{\xi}(t,\cdot)$ is its inverse. To show that $\widetilde{\xi}(t,\cdot)$ is injective, it suffices to show that $\mathbb{H}$ is the range of $\widetilde{S}^t_0$. Let $z_0 \in \mathbb{H}$. Set $x_0 := \widetilde{\xi}(t,z_0)$ set

$$\gamma(s) = \widetilde{\xi}(s,z_0), \quad g(s) = \widetilde{\eta}(s,z_0).$$

Then (γ, g) satisfies the same system of differential equations as

$$[0,t] \ni s \mapsto (\widetilde{S}^t_s[x_0], \widetilde{P}^t_s[x_0]) \quad \text{on } (0,t).$$

We have $\gamma(t) = x_0$ and

$$g(0) = \widetilde{\eta}(0,z_0) = \nabla\widetilde{\mathscr{U}}_0(z_0) = \nabla\widetilde{\mathscr{U}}_0(\gamma(0)).$$

Thus, $(\gamma, g)(s) \equiv (\widetilde{S}^t_s[x_0], \widetilde{P}^t_s[x_0])$ on $[0,t]$. In particular, $z_0 = \gamma(0) = \widetilde{S}^t_0[x_0]$. Thus, $\widetilde{S}^t_0$ is surjective.

Continuity. Since $\widetilde{\xi}_t$ is a bijection of $\mathbb{H}$ onto $\mathbb{H}$, (1.27) and the invariance of domain theorem imply that $\widetilde{\xi}_t$ is a homeomorphism of $\{M^q : q \in \mathbb{M}^m\}$ onto $\{M^q : q \in \mathbb{M}^m\}$.

(iii) By (ii)

$$\widetilde{S}^t_0 \circ \widetilde{\xi}_t = \mathrm{id}_{\mathbb{H}} = \widetilde{\xi}_0 \quad \text{and} \quad \widetilde{P}^t_0 \circ \widetilde{\xi}_t = \nabla\widetilde{\mathscr{U}}_0\big(\widetilde{S}^t_0 \circ \xi_t\big) = \nabla\widetilde{\mathscr{U}}_0 = \widetilde{\eta}_0.$$

Since $s \mapsto (\widetilde{S}^t_s \circ \widetilde{\xi}_t, \widetilde{P}^t_s \circ \widetilde{\xi}_t)$ and $s \mapsto (\widetilde{\xi}_s, \widetilde{\eta}_s)$ satisfy the same system of differential equations on $(0,t)$, we obtain the assertions in (iii).

(iv) We use first Proposition 1.5 (iv) and then (i) of the current proposition to obtain that $\nabla\widetilde{\mathscr{U}}(t,\widetilde{\xi}(t,\cdot)) = \widetilde{\eta}(t,\cdot)$. We use the identity $\dot{\widetilde{\xi}} = \nabla_b\mathscr{H}(\widetilde{\xi},\widetilde{\eta})$ to conclude the proof. $\square$

Remark C.4. (i) We notice that Proposition C.1, which imposes (4.7), allows us to improve the continuity property of $\widetilde{\xi}_t$ and its inverse to the infinite-dimensional space; i.e., this implies that $\widetilde{\xi}_t$ is a homeomorphism of $\mathbb{H}$ onto itself.

(ii) We observe that by Proposition C.2(iv) we have that $\nabla\widetilde{\mathscr{U}}(t,\cdot) = \widetilde{\eta}(t, \widetilde{S}_0^t[\cdot])$, and since both $\widetilde{\eta}$ and $\widetilde{S}_0^t$ are locally Lipschitz-continuous (by (i) of the previous proposition and Lemma C.7, respectively) we have that $\nabla\widetilde{\mathscr{U}}(t,\cdot)$ is locally Lipschitz-continuous, just as in Section 1.3; by a different perspective one obtains that $\widetilde{\mathscr{U}}(t,\cdot) \in C^{1,1}_{\mathrm{loc}}(\mathbb{H})$.

C.2 Flows on $\mathbb{H}$ and on $\mathscr{P}_2(\mathbb{M})$ and their properties

LEMMA C.5. *Let $x, y \in \mathbb{H}$ be such that $\sharp(x) = \sharp(y)$. Then for $0 \le s \le t$, we have $\sharp(\widetilde{S}_s^t[x]) = \sharp(\widetilde{S}_s^t[y])$. As a consequence, given $\mu \in \mathscr{P}_2(\mathbb{M})$, the following measures are well-defined*

$$\sigma_s^t[\mu] := \sharp\big(\widetilde{S}_s^t[x]\big) \tag{C.3}$$

where $\sharp(x) = \mu$ depends only on μ and is independent of the choice of x.

PROOF. Since $\sharp(x) = \sharp(y)$, there exist Borel bijective maps $S_n : \Omega \to \Omega$ such that (cf. [13, 32])

$$\sharp(S_n) = \sharp(S_n^{-1}) = \mathscr{L}^d_\Omega, \quad \lim_{n\to\infty} \|y - x\circ S_n\| = 0.$$

Thus,

$$\lim_{n\to\infty} \big\|\widetilde{S}_s^t[y] - \widetilde{S}_s^t[x]\circ S_n\big\| = \lim_{n\to\infty} \big\|\widetilde{S}_s^t[y] - \widetilde{S}_s^t[x\circ S_n]\big\| = 0.$$

This proves

$$W_2\Big(\sharp\big(\widetilde{S}_s^t[y]\big), \sharp\big(\widetilde{S}_s^t[x]\big)\Big) = \lim_{n\to\infty} W_2\Big(\sharp\big(\widetilde{S}_s^t[x]\circ S_n\big), \sharp\big(\widetilde{S}_s^t[x]\big)\Big) = 0. \qquad \square$$

Remark C.6. The following hold.

(i) By Proposition 1.5, there exists $e_T : [0,\infty) \to [0,\infty)$, monotone nondecreasing such that

$$\big\|\widetilde{S}_s^t[x]\big\|,\ \big\|\partial_s \widetilde{S}_s^t[x]\big\| \le e_T(\|x\|) \quad \forall s \in [0,t], \forall t \in [0,T].$$

(ii) By (i)

$$\big\{\sigma_s^t[\mu] : \mu \in \mathscr{B}_r, 0 \le s \le t \le T\big\} \subset \mathscr{B}_{e_T(r)}.$$

(iii) By Proposition 1.5 again, there exists $C_T : (0,\infty) \to (0,\infty)$ monotone nondecreasing such that

$$\big\|\nabla\widetilde{\mathscr{U}}(t,x)\big\| \le C_T(r)(1 + \|x\|), \quad \forall x \in \mathbb{B}_r(0), \forall t \in [0,T].$$

(iv) By Lemma 3.11, the regularity property obtained on $\widetilde{\mathscr{U}}$ in Proposition 1.5, we have that $\mathscr{U}$ is differentiable. We use Proposition C.2 (iv) to conclude that $(s,q)\mapsto D_pH(q,\nabla_w\mathscr{U}(s,\sigma_s^t[\mu])(q))$ is a velocity for $s\mapsto\sigma_s^t[\mu]$. In other words

(C.4)
$$\partial_s\sigma_s^t[\mu]+\nabla\cdot\Big(D_pH\big(\,\cdot\,,\nabla_w\mathscr{U}\big(s,\sigma_s^t[\mu]\big)\big)\sigma_s^t[\mu]\Big)=0 \text{ in } \mathscr{D}'((0,t)\times\mathbb{M})\ \ \sigma_t^t[\mu]=\mu.$$

LEMMA C.7. *Suppose* $0<t\le\bar{t}\le T$ *and* $r>0$. *Then there exists a constant* $C(r,T)$ *monotone increasing in* r *such that the following hold*:

(i) *If* $x,y\in\mathbb{B}_r(0)$ *then*
$$\big\|\widetilde{S}_s^{\bar{t}}[x]-\widetilde{S}_s^t[y]\big\|\le e^{C(r,T)(t-s)}\big(|\bar{t}-t|e_T(\|x\|)+\|x-y\|\big)\quad\forall s\in[0,t].$$
and
$$\big\|\widetilde{S}_s^{\bar{t}}[x]-\widetilde{S}_t^{\bar{t}}[x]\big\|\le(s-t)e_T(r)\quad\forall s\in[t,\bar{t}].$$

(ii) *If* $\mu,\nu\in\mathscr{B}_r$ *then*
$$W_2\big(\sigma_s^{\bar{t}}[\mu],\sigma_s^t[\nu]\big)\le e^{C(r,T)(t-s)}\big(|\bar{t}-t|e_T(r)+W_2(\mu,\nu)\big)\quad\forall s\in[0,t]. \tag{C.5}$$
and
$$W_2\Big(\sigma_s^{\bar{t}}[\mu],\sigma_t^{\bar{t}}[\mu]\Big)\le(s-t)e_T(r)\qquad\forall s\in[t,\bar{t}].$$

PROOF. (i) Let $x,y\in\mathbb{B}_r(0)$.
We have
$$\big\|x-\widetilde{S}_t^{\bar{t}}[x]\big\|=\left\|\int_t^{\bar{t}}\partial_s\widetilde{S}_s^{\bar{t}}[x]ds\right\|\le\int_t^{\bar{t}}\big\|\partial_s\widetilde{S}_s^{\bar{t}}[x]\big\|ds.$$
We use Remark C.6 (i) to infer
$$\big\|x-\widetilde{S}_t^{\bar{t}}[x]\big\|\le|\bar{t}-t|e_T(\|x\|). \tag{C.6}$$
Set
$$h(s):=\frac{1}{2}\big\|\widetilde{S}_s^{\bar{t}}[x]-\widetilde{S}_s^t[x]\big\|\quad\forall s\in[0,t].$$
We have
$$h'(s)=\int_\Omega\big(\widetilde{S}_s^{\bar{t}}[x]-\widetilde{S}_s^t[x]\big)\cdot\Big(D_pH\big(\widetilde{S}_s^{\bar{t}}[x],\nabla\widetilde{U}(s,\widetilde{S}_s^{\bar{t}}[x])\big)-D_pH\big(\widetilde{S}_s^t[x],\nabla\widetilde{U}(s,\widetilde{S}_s^t[x])\big)\Big)d\omega.$$
By the fact that DH is Lipschitz, we have
$$\Big|D_pH\big(\widetilde{S}_s^{\bar{t}}[x],\nabla\widetilde{U}(s,\widetilde{S}_s^{\bar{t}}[x])\big)-D_pH\big(\widetilde{S}_s^t[x],\nabla\widetilde{U}(s,\widetilde{S}_s^t[x])\big)\Big|^2\le\kappa_0^2\Big(\big|\widetilde{S}_s^{\bar{t}}[x]-\widetilde{S}_s^t[x]\big|^2+\big|\nabla\widetilde{U}(s,\widetilde{S}_s^{\bar{t}}[x])-\nabla\widetilde{U}(s,\widetilde{S}_s^t[x])\big|^2\Big).$$
We use Proposition 1.5 to obtain a constant $C(r,T)$ which increases in r and such that

$$\left\| D_p H\big(\widetilde{S}_s^{\bar t}[x], \nabla \widetilde{U}(s, \widetilde{S}_s^{\bar t}[x])\big) - D_p H\big(\widetilde{S}_s^{t}[x], \nabla \widetilde{U}(s, \widetilde{S}_s^{t}[x])\big) \right\| \le C(r,T) \| \widetilde{S}_s^{\bar t}[x] - \widetilde{S}_s^{t}[x] \|.$$

This implies $h' \ge -2C(r,T)h$, and so Grönwall's inequality yields

$$h(s) \le e^{2C(r,T)(t-s)} h(t) \quad \forall s \in [0,t].$$

Thus,

$$\left\| \widetilde{S}_s^{\bar t}[x] - \widetilde{S}_s^{t}[x] \right\| \le e^{C(r,T)(t-s)} \left\| \widetilde{S}_t^{\bar t}[x] - \widetilde{S}_t^{t}[x] \right\| = e^{C(r,T)(t-s)} \left\| \widetilde{S}_t^{\bar t}[x] - x \right\|.$$

This, together with (C.6), implies

$$\left\| \widetilde{S}_s^{\bar t}[x] - \widetilde{S}_s^{t}[x] \right\| \le e^{C(r,T)(t-s)} |\bar t - t| e_T(\|x\|). \tag{C.7}$$

We use arguments similar to the ones above to obtain

$$\left\| \widetilde{S}_s^{t}[x] - \widetilde{S}_s^{t}[y] \right\| \le e^{C(r,T)(t-s)} \|x - y\| \quad \forall s \in [0,t]. \tag{C.8}$$

We combine (C.7) and (C.8) to verify the first identity in (i). The second identity follows from direct integration.

(ii) Let $\mu, \nu \in \mathscr{B}_r$ and choose $x, y \in \mathbb{H}$ such that $\sharp(x) = \mu$ and $\sharp(y) = \nu$ and $W_2(\mu,\nu) = \|x - y\|$. Since $\sharp(\widetilde{S}_s^{\bar t}[x]) = \sigma_s^{\bar t}[\mu]$ and $\sharp(\widetilde{S}_s^{t}[y]) = \sigma_s^{t}[\nu]$, (i) implies (ii). □

C.3 Proof of Proposition 1.5

Let $y \in \mathbb{B}_r(0)$.

(i) By Remark 1.4, $U^{(m)}$ is a viscosity solution to (1.20), and so the standard theory of Hamilton–Jacobi equations in finite-dimensional spaces yields the pointwise identity

$$U^{(m)}(t_2, q) - U^{(m)}(t_1, q) = -\int_{t_1}^{t_2} \mathscr{H}^m\big(q, D_q U^{(m)}(\tau, q)\big) d\tau$$

for $q \in \mathbb{M}^m$. We use (1.10) to infer

$$\widetilde{\mathscr{U}}(t_2, M^q) - \widetilde{\mathscr{U}}(t_1, M^q) = -\int_{t_1}^{t_2} \widetilde{\mathscr{H}}\big(M^q, \nabla \widetilde{\mathscr{U}}(\tau, M^q)\big) d\tau.$$

By Proposition 1.3(ii), when $r > 1$, $\nabla \widetilde{\mathscr{U}}$ is bounded on $[t_1, t_2] \times \mathbb{B}_r(y)$. Observe that $\nabla \widetilde{\mathscr{U}}(\tau, \cdot)$ is continuous when $\tau \in [t_1, t_2]$ and $\widetilde{\mathscr{H}}$ is continuous. Since

$$\{M^q : q \in \mathbb{M}^m,\ m \in \mathbb{N}\}$$

is dense in $\mathbb{H}$, (i) holds.

(ii) First, one obtains a finite number $c(r,T)$ increasing in the variables r and T such that

$$\left| \widetilde{\nabla} \mathscr{U}(t_2, y) - \widetilde{\nabla} \mathscr{U}(t_1, y) \right| \le 2c(r,T)|t_2 - t_1|. \tag{C.9}$$

This together with the space Lipschitz property of $\nabla\widetilde{\mathscr{U}}$ implies $\nabla\widetilde{\mathscr{U}}$ is Lipschitz on $[0,T]\times\mathbb{B}_r(0)$. As a composition of locally Lipschitz functions, $(\tau,x)\mapsto\widetilde{\mathscr{H}}(x,\nabla\widetilde{\mathscr{U}}(\tau,x))$ is Lipschitz on $[0,T]\times\mathbb{B}_r(0)$. Hence since by (i) we have that $\partial_t\widetilde{\mathscr{U}}=-\widetilde{\mathscr{H}}\big(\cdot,\nabla\widetilde{\mathscr{U}}\big)$, we conclude $\partial_t\widetilde{\mathscr{U}}$ is Lipschitz on $[0,T]\times\mathbb{B}_r(0)$.

(iii)–(v) We refer the reader to [33]. □

Acknowledgment. The research of WG was supported by National Science Foundation Grant DMS–1700202. Both authors acknowledge the support of U.S. Air Force Grant FA9550-18-1-0502. The authors would like to express their gratitude to P. Cannarsa for the discussions and for pointing out important references on regularity properties of solutions to Hamilton–Jacobi equations on $\mathbb{R}^d$. They wish to thank A. Swiech for the discussions on regularity of solutions to Hamilton–Jacobi equations on Hilbert spaces. The feedback of P. Cardaliaguet on the manuscript is also greatly appreciated. The authors wish to thank the anonymous referee for making pertinent suggestions which improved the manuscript.

Bibliography

[1] Ahuja, S. Wellposedness of mean field games with common noise under a weak monotonicity condition. *SIAM J. Control Optim.* **54** (2016), no. 1, 30–48. doi:10.1137/140974730

[2] Ahuja, S.; Ren, W.; Yang, T.-W. Forward-backward stochastic differential equations with monotone functionals and mean field games with common noise. *Stochastic Process. Appl.* **129** (2019), no. 10, 3859–3892. doi:10.1016/j.spa.2018.11.005

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

[4] Ambrose, D. Existence theory for a time-dependent mean field games model of household wealth. *Appl. Math. Optim.* **83** (2021), no. 3, 2051–2081. doi:10.1007/s00245-019-09619-5

[5] Ambrosio, L.; Gangbo, W. Hamiltonian ODEs in the Wasserstein space of probability measures. *Comm. Pure Appl. Math.* **61** (2008), no. 1, 18–53. doi:10.1002/cpa.20188

[6] Ambrosio, L.; Gigli, N.; Savaré, G. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser, Basel, 2008.

[7] Bensoussan, A.; Graber, P.J.; Yam, S.C.P. Stochastic control on space of random variables. Preprint, 2019. arXiv:1903.12602 [math.OC]

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

[9] Bensoussan, A.; Yam, S.C.P. Control problem on space of random variables and master equation. *ESAIM Control Optim. Calc. Var.* **25** (2019), Paper No. 10, 36 pp. doi:10.1051/cocv/2018034

[10] Bessi, U. Existence of solutions of the master equation in the smooth case. *SIAM J. Math. Anal.* **48** (2016), no. 1, 204–228. doi:10.1137/15M1018782

[11] Buckdahn, R.; Li, J.; Peng, S.; Rainer, C. Mean-field stochastic differential equations and associated PDEs. *Ann. Probab.* **45** (2017), no. 2, 824–878. doi:10.1214/15-AOP1076

[12] Cannarsa, P.; Sinestrari, C. *Semiconcave functions, Hamilton–Jacobi equations and optimal control*. Progress in Nonlinear Differential Equations and their Applications, 58. Birkhäuser Boston, Boston, 2004.

[13] Cardaliaguet, P. *Notes on Mean-Field Games*. Lectures by P. L. Lions, Collège de France, 2010.

[14] Cardaliaguet, P.; Delarue, F.; Lasry, J-M.; Lions, P-L. *The master equation and the convergence problem in mean field games*. Annals of Mathematics Studies, 201. Princeton University Press, Princeton, NJ, 2019. doi:10.2307/j.ctvckq7qf

[15] Carmona, R.; Delarue, F. Forward-backward stochastic differential equations and controlled McKean-Vlasov dynamics. *Ann. Probab.* **43** (2015), no. 5, 2647–2700. doi:10.1214/14-AOP946

[16] Carmona, R.; Delarue, F. *Probabilistic theory of mean field games with applications. I. Mean field FBSDEs, control, and games*. Probability Theory and Stochastic Modelling, 83. Springer, Cham, 2018.

[17] Carmona, R.; Delarue, F. *Probabilistic theory of mean field games with applications. II. Mean field games with common noise and master equations*. Probability Theory and Stochastic Modelling, 84. Springer, Cham, 2018.

[18] Carroll, F. W. A polynomial in each variable separately is a polynomial. *Amer. Math. Monthly* **68** (1961), 42. doi:10.2307/2311361

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

[20] Chassagneux, J.-F.; Szpruch, L.; Tse, A. Weak quantitative propagation of chaos via differential calculus on the space of measures. *Ann. Appl. Probab.*, to appear. Preprint, 2019. arXiv:1901.02556 [math.PR]

[21] Chow, Y. T.; Gangbo, W. A partial Laplacian as an infinitesimal generator on the Wasserstein space. *J. Differential Equations* **267** (2019), no. 10, 6065–6117. doi:10.1016/j.jde.2019.06.012

[22] Crandall, M. G.; Lions, P.-L. Hamilton–Jacobi equations in infinite dimensions I. Uniqueness of viscosity solutions. *J. Funct. Anal.* **62** (1985), no. 3, 379–396. doi:10.1016/0022-1236(85)90011-4

[23] Crandall, M. G.; Lions, P.-L. Hamilton–Jacobi equations in infinite dimensions II. Existence of viscosity solutions. *J. Funct. Anal.* **65** (1986), no. 3, 368–405. doi:10.1016/0022-1236(86)90026-1

[24] Delarue, F.; Lacker, D.; Ramanan, K. From the master equation to mean field game limit theory: a central limit theorem. *Electron. J. Probab.* **24** (2019), Paper No. 51, 54 pp. doi:10.1214/19-EJP298

[25] Delarue, F.; Lacker, D.; Ramanan, K. 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

[26] Dello Schiavo, L. A Rademacher-type theorem on L^2-Wasserstein spaces over closed Riemannian manifolds. *J. Funct. Anal.* **278** (2020), no. 6, 108397, 57 pp. doi:10.1016/j.jfa.2019.108397

[27] Fathi, A. *Weak KAM theorem in Lagrangian dynamics*.Cambridge Studies in Advanced Mathematics, to appear.

[28] Fischer, M.; Silva, F. J. On the asymptotic nature of first order mean field games. *Appl. Math. Optim.* **84** (2021), no. 2, 2327–2357. doi:10.1007/s00245-020-09711-1

[29] Gangbo, W.; Nguyen, T.; Tudorascu, A. Hamilton–Jacobi equations in the Wasserstein space. *Meth. Appl. Anal.* **15** (2008), no. 2, 155–184. doi:10.4310/MAA.2008.v15.n2.a4

[30] Gangbo, W.; Swiech, A. Optimal transport and large number of particles. *Discrete Contin. Dyn. Syst.* (2014) **34** (2014), no. 4, 1397–1441. doi:10.3934/dcds.2014.34.1397

[31] Gangbo, W.; Swiech, A. 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

[32] Gangbo, W.; Tudorascu, A. On differentiability in the Wasserstein space and well-posedness for Hamilton–Jacobi equations. *J. Math. Pures Appl. (9)* **125** (2019), 119–174. doi:10.1016/j.matpur.2018.09.003

[33] Gomes, D.; Nurbekyan, L. On the minimizers of calculus of variations problems in Hilbert spaces. *Calc. Var. Partial Differential Equations* **52** (2015), no. 1-2, 65–93. doi:10.1007/s00526-013-0705-6

[34] Lasry, J.-M.; Lions, P.-L. Jeux à champ moyen. I. Le cas stationnaire. *C. R. Math. Acad. Sci. Paris* **343** (2006), no. 9, 619–625. doi:10.1016/j.crma.2006.09.019

[35] Lasry, J.-M.; Lions, P.-L. Jeux à champ moyen. II. Horizon fini et contrôle optimal. *C. R. Math. Acad. Sci. Paris* **343** (2006), no. 10, 679–684. doi:10.1016/j.crma.2006.09.018
[36] Lasry, J.-M.; Lions, P.-L. Mean field games. *Jpn. J. Math.* **2** (2007), no. 1, 229–260. doi:10.1007/s11537-007-0657-8
[37] Lions, P.-L. Lectures at Collège de France. 2007–2013.
[38] Lions, P.-L.; Lasry, J.-M. Large investor trading impacts on volatility. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **24** (2007), no. 2, 311–323. doi:10.1016/j.anihpc.2005.12.006
[39] Mayorga, S. 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
[40] McCann, R. J. A convexity principle for interacting gases. *Adv. Math.* **128** (1997), no. 1, 153–179. doi:10.1006/aima.1997.1634
[41] Mischler, S.; Mouhot, C. Kac's program in kinetic theory. *Invent. Math.* **193** (2013), no. 1, 1–147. doi:10.1007/s00222-012-0422-3
[42] Mou, C.; Zhang, J. Wellposedness of second order master equations for mean field games with nonsmooth data. Preprint, 2019. arXiv:1903.09907 [math.AP]
[43] Santambrogio, F. *Optimal transport for applied mathematicians. Calculus of variations, PDEs, and modeling.* Progress in Nonlinear Differential Equations and Their Applications, 87. Birkhäuser/Springer, Cham, 2015. doi:10.1007/978-3-319-20828-2
[44] Villani, C. *Topics in optimal transportation.* Graduate Studies in Mathematics, 58. American Mathematical Society, Providence, RI, 2003. doi:10.1090/gsm/058

WILFRID GANGBO
Department of Mathematics
UCLA
520 Portola Plaza
Los Angeles, CA 90095
USA
E-mail: wgangbo@math.ucla.edu

ALPÁR R. MÉSZÁROS
Department of Mathematical Sciences
Durham University
Upper Mountjoy
Durham DH1 3LE
ENGLAND
E-mail: alpar.r.meszaros@durham.ac.uk

Received April 2020.