

# OPTIMAL CONTROL PROBLEMS DRIVEN BY NONLINEAR DEGENERATE FOKKER-PLANCK EQUATIONS

FRANCESCA ANCESCHI, GIACOMO ASCIONE, DANIELE CASTORINA,  
AND FRANCESCO SOLOMBRINO

ABSTRACT. The well-posedness of a class of optimal control problems is analysed, where the state equation couples a nonlinear degenerate Fokker-Planck equation with a system of Ordinary Differential Equations (ODEs). Such problems naturally arise as mean-field limits of Stochastic Differential models for multipopulation dynamics, where a large number of agents (followers) is steered through parsimonious intervention on a selected class of leaders. The proposed approach combines stability estimates for measure solutions of nonlinear degenerate Fokker-Planck equations with a general framework of assumptions on the cost functional, ensuring compactness and lower semicontinuity properties. The Lie structure of the state equations allows one for considering non-Lipschitz nonlinearities, provided some suitable dissipativity assumptions are considered in addition to non-Euclidean Hölder and sublinearity conditions.

## CONTENTS

1. Introduction	2
2. Useful notation and preliminary results	7
3. Proof of Theorem 1.3 and Corollary 1.4	13
4. The PDE-ODE system and a related optimal control problem	27
Appendix A. Some sufficient conditions to the validity of assumption $(v_3)$	39
Appendix B. Preliminaries for the well-posedness of the PDE-ODE system	41
Appendix C. Other Possible Control Classes	46
Appendix D. An example of application of Theorem 1.5	52
References	61

---

2020 *Mathematics Subject Classification.* 49J20, 49J55, 60H10.

*Key words and phrases.* Mean-field limit,  $\Gamma$ -limit, optimal control with ODE-SDE constraints.

FA is partially supported by INdAM-GNAMPA Project "Problemi ellittici e sub-ellittici: non linearità, singolarità e crescita critica" CUP E53C23001670001. GA is supported by the GNAMPA-INdAM Project "Deterministic Control of Stochastic Dynamics", CUP E53C23001670001 and the PRIN 2022XZSAFN Project "Anomalous Phenomena on Regular and Irregular Domains: Approximating Complexity for the Applied Sciences" CUP E53D23005970006. DC and FS are supported by the PRIN 2022HKBF5C Project "Variational Analysis of Complex Systems in Material Science, Physics and Biology" CUP E53D23005720006. PRIN projects are part of PNRR Italia Domani, financed by European Union through NextGenerationEU. FS additionally acknowledges support from the GNAMPA-INdAM Project "Problemi di controllo ottimo nello spazio di Wasserstein delle misure definite su spazi di Banach" CUP E53C23001670001.

## 1. INTRODUCTION

**1.1. Presentation of the problem.** The aim of this work is to study a class of optimal control problems driven by a dynamic of the type

$$\begin{cases} \partial_t \mu_t = -v \cdot \nabla_x \mu_t + \sigma \Delta_v \mu_t - \operatorname{div}_v(\mathbf{v}[t, \boldsymbol{\mu}](z) \mu_t) & (t, x, v) \in (0, T] \times \mathbb{R}^{2d}, \\ \mu_0 = \bar{\mu} & (x, v) \in \mathbb{R}^{2d}, \end{cases} \quad (1.1)$$

where  $\sigma \in \mathbb{R}$  and the drift field  $\mathbf{v}$  is allowed to depend explicitly on a solution  $\boldsymbol{\mu}$  of the problem. Hence, the above is a nonlinear Fokker-Planck equation driven by a nonlocal drift field  $\mathbf{v}[t, \boldsymbol{\mu}]$  depending on the state of the system, a typical example being the choice

$$\mathbf{v}[t, \boldsymbol{\mu}](z) = H(t, z) * \mu_t.$$

Furthermore, (1.1) is the natural mean field counterpart of second order multi-agents systems with additive noise, see (1.3) below. In fact, in this case a drift field of the form  $\mathbf{v}[t, \boldsymbol{\mu}]$  naturally arises if one assumes that the natural dynamic is driven by mutual interactions (see Appendix D). These type of mean field systems are in connection with a great number of applications. Indeed, multiagent systems have been used in several contexts, such as in biology to model cell aggregation [15, 19, 23, 26], in chemistry to model chemical networks [30, 37, 38], or even to describe human interaction in social sciences [14, 20, 31, 45]. Furthermore, controlled leader-follower multiagent systems have been used, for instance, in the context of gene regulation in microbial consortia [34, 44], traffic control [41], swarms control [7], or even oil cleaning through controlled robots [49]. The controllability of this kind of leader-follower systems, with a fixed finite (sufficiently small) number of followers, has been widely studied, see, for instance, [9, 28, 29]. On the other hand, the mean-field approach for a large number of followers has been considered for instance in [3] for crowd control and in [40] for prevention of maritime crime. A general study of these mean-field control problems arising from first-order leader-follower systems with additive noise has been presented in [8], while the notion of *continuification of controlled systems* has been further explored for instance in [32, 33]. It is clear that this list of applications and references is far from being exhaustive. Furthermore, let us stress that one could be interested in the control of a second-order system, as in [1, 2], in which the followers are subject to an additive noise only in the velocity component. Hence, we consider (1.1) as a second-order generalization of the mean-field system obtained in [8].

Before describing our point of view in detail, it is useful to give a closer look to the notion of solution we need to consider when dealing with (1.1). Indeed, the well-posedness of the class of optimal control problems we consider heavily relies on stability estimates for solutions in suitable spaces. Thus, this structure immediately suggests that it is natural to look for solutions to (1.1) in Wasserstein spaces (see Section 2), and hence the definition of solution to (1.1) we are interested into is the following.

**Definition 1.1.** Let  $p \geq 1, T > 0$  (possibly  $T = \infty$ ) and  $\bar{\mu} \in \mathcal{W}_p(\mathbb{R}^{2d})$ . A continuous curve of probability measures  $\boldsymbol{\mu} \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$  is a solution of (1.1) if and only if for all

$\psi \in C_c^\infty(\mathbb{R}^{2d})$  and for all  $t \in [0, T]$  it holds

$$\int_{\mathbb{R}^{2d}} \psi d\mu_t - \int_{\mathbb{R}^{2d}} \psi d\bar{\mu} = \int_0^t \int_{\mathbb{R}^{2d}} (v \cdot \nabla_x \psi + \mathbf{v}[s, \boldsymbol{\mu}](z) \cdot \nabla_v \psi + \sigma \Delta_v \psi) d\mu_s ds. \quad (1.2)$$

If  $T = \infty$ , we say that  $\boldsymbol{\mu}$  is a global solution of (1.1), otherwise we call it a local solution.

As a first goal of our work, we will focus on uniqueness and stability issues for (1.1). In doing so, we consider a set of assumptions on the drift term  $\mathbf{v}[t, \boldsymbol{\mu}]$  which is as general as possible and allows for possibly unbounded non-globally Lipschitz drift fields provided some Hölder continuity combined with a dissipativity condition are satisfied (see Assumptions (v) below). Indeed, we refer to  $\mathbf{v}_3$  as a dissipativity condition since for  $p = 2$  it is implied by the metrical dissipativity condition considered in [17], see also [16], as shown in Appendix A.

Despite the presence of a nonlinearity in the structure of (1.1), the results we are going to show strongly rely on well-posedness results for a linear counterpart, that has been analyzed in [5]. Furthermore a key point for uniqueness and stability issues consists in showing that (1.1) is indeed equivalent to its McKean-Vlasov counterpart, in the sense that a solution  $\mu$  is always the law of a stochastic process of the type

$$\begin{cases} dX(t) = V(t)dt \\ dV(t) = \mathbf{v}[t, \boldsymbol{\mu}](X(t), V(t))dt + \sqrt{2\sigma}dB(t) \\ X(0) = X_0 \quad V(0) = V_0 \quad \boldsymbol{\mu} = \text{Law}(X, V), \end{cases} \quad (1.3)$$

where  $B$  is a  $d$ -dimensional Brownian motion on  $(\Omega, \Sigma, \mathbb{P})$ ,  $\mathbf{v} : [0, T] \times C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d})) \times \mathbb{R}^{2d} \rightarrow \mathbb{R}^d$  is a measurable function,  $\sigma > 0$  and  $Z_0 = (X_0, V_0) \in L^p(\Omega; \mathbb{R}^{2d})$ . For this reason, throughout this work we fix a probability space  $(\Omega, \Sigma, \mathbb{P})$  and we consider the following adapted formalism for McKean-Vlasov SDEs to the case of our interest. For simplicity, by  $Z \in L^p(\Omega; C(\mathbb{R}_0^+; \mathbb{R}^{2d}))$  we actually mean that  $\{Z(t)\}_{t \in [0, T]} \in L^p(\Omega; C([0, T]; \mathbb{R}^{2d}))$  for all  $T > 0$ .

**Definition 1.2.** Let  $p \geq 1$ ,  $T > 0$  (possibly  $T = \infty$ ) and consider the McKean-Vlasov SDE (1.3), for  $t \in [0, T]$ . We say that a stochastic process  $Z = (X, V)$  is a strong solution of (1.3) if  $Z \in L^p(\Omega; C([0, T]; \mathbb{R}^{2d}))$ ,  $\int_0^t |\mathbf{v}[s, \boldsymbol{\mu}](Z(s))| ds < \infty$  for all  $t \in [0, T]$  and

$$X(t) = X_0 + \int_0^t V(s)ds \quad \text{and} \quad V(t) = V_0 + \int_0^t \mathbf{v}[s, \boldsymbol{\mu}](Z(s))ds + \sqrt{2\sigma}B(t), \quad (1.4)$$

for a.a.  $t \in [0, T]$  a.s., where  $\boldsymbol{\mu} = \text{Law}(Z)$ . We say that uniqueness in law holds for (1.3) if for any two strong solutions  $Z_1, Z_2$  we have  $\text{Law}(Z_1(t)) = \text{Law}(Z_2(t))$  for all  $t \in [0, T]$ . We say that pathwise uniqueness holds for (1.3) if for any two strong solutions  $Z_1, Z_2$  we have  $\mathbb{P}(Z_1 = Z_2) = 1$ . Clearly pathwise uniqueness implies uniqueness in law, while the converse is also true by [18, Theorem 3.2]. For such a reason, we just state that the strong solution is unique.

**1.2. Our assumptions.** Specifically, in order to state our main results we need to introduce a suitable set of assumptions to consider. For any  $p \geq 1$ ,  $t \geq 0$  and  $\boldsymbol{\mu} \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$  we set

$$\overline{M}_p(t, \boldsymbol{\mu}) := \sup_{0 \leq s \leq t} M_p(\mu_s) := \sup_{0 \leq s \leq t} \int_{\mathbb{R}^{2d}} |z|^p d\mu_s,$$

then the assumptions we will consider throughout this work are the following ones.

**Assumptions on  $\mathbf{v}$ : (v)**

(v<sub>0</sub>)  $\mathbf{v} : [0, +\infty) \times C(\mathbb{R}_0^+; \mathcal{W}_p(\mathbb{R}^{2d})) \times \mathbb{R}^{2d} \rightarrow \overline{\mathbb{R}}^d$  is a Carathéodory map, i.e. it is continuous in the variable  $(\boldsymbol{\mu}, z) \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d})) \times \mathbb{R}^{2d}$  and measurable in the variable  $t \in [0, T]$ .

(v<sub>1</sub>) For all  $T > 0$  there exist two constants  $K > 0$  and  $\beta \in [0, 1)$  such that

$$|\mathbf{v}[t, \boldsymbol{\mu}](z)| \leq K \left( 1 + |x|^{\frac{\beta}{3}} + |v|^\beta + (\overline{M}_p(T, \boldsymbol{\mu}))^{\frac{1}{p}} \right)$$

for every  $z \in \mathbb{R}^{2d}$ ,  $t \in [0, T]$  and for every  $\boldsymbol{\mu} \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$ .

(v<sub>2</sub>) For all  $T > 0$  there exists an exponent  $\alpha \in (\beta, 1]$  with the property that for all  $R > 0$  there exists a constant  $L > 0$  and such that for all  $t \in [0, T]$ , for all  $\boldsymbol{\mu} \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$  with  $\overline{M}_p(T, \boldsymbol{\mu}) \leq R$  and for all  $z_1, z_2 \in \mathbb{R}^{2d}$  with  $|z_1|, |z_2| \leq R$  it holds

$$|\mathbf{v}[t, \boldsymbol{\mu}](z_1) - \mathbf{v}[t, \boldsymbol{\mu}](z_2)| \leq L (|x_1 - x_2|^{\frac{\alpha}{3}} + |v_1 - v_2|^\alpha).$$

(v<sub>3</sub>) For all  $T > 0$  there exists a constant  $D \geq 0$  such that for any couple of stochastic processes  $Z_j = (X_j, V_j) \in L^p(\Omega; C([0, T]; \mathbb{R}^{2d}))$  with  $X'_j = V_j$  in  $[0, T]$ ,  $Z_1(0) = Z_2(0)$  and  $\boldsymbol{\mu}^j = \text{Law}(X_j, V_j)$ ,  $j = 1, 2$ , it holds

$$\begin{aligned} & \int_0^t \mathbb{E} [ |V_1(s) - V_2(s)|^{p-2} (\mathbf{v}[s, \boldsymbol{\mu}^1](Z_1(s)) - \mathbf{v}[s, \boldsymbol{\mu}^2](Z_2(s))) \cdot (V_1(s) - V_2(s)) ] ds \\ & \leq D \int_0^t \sup_{0 \leq z \leq s} \mathbb{E} [ |Z_1(z) - Z_2(z)|^p ] ds \quad \forall t \in [0, T]. \end{aligned}$$

(v<sub>4</sub>) For all  $T > 0$ ,  $t \in [0, T]$  and  $\boldsymbol{\mu}^1, \boldsymbol{\mu}^2 \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$  with  $\mu_s^1 = \mu_s^2$  for all  $s \in [0, t]$  it holds  $\mathbf{v}[t, \boldsymbol{\mu}^1] = \mathbf{v}[t, \boldsymbol{\mu}^2]$ .

Despite its generality, it is clear that the dissipativity condition (v<sub>3</sub>) is not easily verified. However, we can still provide some sufficient conditions for its validity, see Appendix A. In the notation for  $\mathbf{v}$ , we separate the variables  $t$  and  $\boldsymbol{\mu}$  and  $(X(t), V(t))$  for the following reason: throughout the proof of the main theorem, we will freeze the variable  $\boldsymbol{\mu}$  and study (1.3) as a simple SDE, and we only need measurability in  $t$ .

**1.3. Main results.** Our first main result concerns well-posedness of (1.1) in the sense discussed above of which we give a precise statement here, while for the proof we refer to Section 3.

**Theorem 1.3.** *Let  $p \geq 1$ ,  $\bar{\mu} \in \mathcal{W}_p(\mathbb{R}^{2d})$  and suppose Assumptions (v) hold. Then the non linear Cauchy problem (1.1) admits a unique global solution  $\boldsymbol{\mu} \in C_{\text{loc}}^{\gamma_p}(\mathbb{R}_0^+; \mathcal{W}_p(\mathbb{R}^{2d}))$ , where  $\gamma_p = \frac{1}{\max\{2,p\}}$ , that can be characterized as follows. Let  $B$  be a  $d$ -dimensional Brownian motion and  $(X_0, V_0) \in L^p(\Omega; \mathbb{R}^{2d})$  be independent of  $B$  with  $\bar{\mu} = \text{Law}(X_0, V_0)$ . Then there exists a unique global strong solution  $(X, V) \in L^p(\Omega; C(\mathbb{R}_0^+; \mathbb{R}^{2d}))$  of (1.3) and  $\boldsymbol{\mu}$  is the unique solution of (1.1). Furthermore, for all  $T > 0$  there exist two functions  $\mathcal{C} : \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+$  and  $\tilde{\Phi} : \mathbb{R}_0^+ \times \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+$  such that  $\mathcal{C}$  is non-decreasing,  $\tilde{\Phi}(t; \cdot)$  is non-increasing for all  $t \geq 0$ ,  $\tilde{\Phi}(\cdot; h)$  is a Young function for all  $h \geq 0$  and it holds*

$$\overline{M}_p(T, \boldsymbol{\mu}) + \overline{M}_{\tilde{\Phi}_p(\cdot; K)}(T, \boldsymbol{\mu}) + \sup_{\substack{0 \leq t, s \leq T \\ t \neq s}} \frac{\mathcal{W}_p(\mu_t, \mu_s)}{|t - s|^{\gamma_p}} \leq \mathcal{C}(K), \quad (1.5)$$

where  $K$  is the constant defined in (v<sub>1</sub>).

Furthermore, we derive the following stability result for solutions to the nonlinear system.

**Corollary 1.4.** *Fix  $\bar{\mu} \in \mathcal{W}_p(\mathbb{R}^{2d})$  and consider a function  $\mathbf{v} : [0, T] \times C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d})) \times \mathbb{R}^{2d} \rightarrow \mathbb{R}^d$  sequence  $\mathbf{v}_j : [0, T] \times C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d})) \times \mathbb{R}^{2d} \rightarrow \mathbb{R}^d$  satisfying Assumptions (v), where the constants  $K$  and  $D$  in (v<sub>1</sub>) and (v<sub>3</sub>) are independent of  $j$ . Denote by  $\mathcal{K}$  the set of curves of probability measures  $\boldsymbol{\mu} \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$  satisfying (1.5). Assume that*

$$\lim_{j \rightarrow \infty} \mathbf{v}_j[t, \boldsymbol{\mu}](z) = \mathbf{v}[t, \boldsymbol{\mu}](z), \quad \text{for any } t \in [0, T], \boldsymbol{\mu} \in \mathcal{K}, z \in \mathbb{R}^{2d}.$$

Let  $\boldsymbol{\mu}^j, \boldsymbol{\mu}$  be the unique solutions of (1.1) with  $\mathbf{v}_j$  and  $\mathbf{v}$ . Then

$$\lim_{j \rightarrow \infty} \sup_{0 \leq t \leq T} \mathcal{W}_p(\mu_t^j, \mu_t) = 0.$$

Eventually, we will turn our attention to optimal control problems connected with nonlinear Fokker-Planck equations. In the situation we consider, a two population system is given. A discrete leader population  $Y_i(t)$ , with  $i = 1, \dots, m$ , is controlled by a policy maker whose aim is steering a population of followers towards a decided goal through mutual interactions. As naturally done in the kinetic approximation of multi-agent systems, the followers population is described by a distribution  $\mu(t)$ , whereas the dynamic is a second order one. Indeed, control fields act on the leaders  $Y_i$ , with  $i = 1, \dots, m$ . Furthermore, the diffusion term  $\Delta_v \mu$  reflects the presence of additive noise on the velocity of the generic follower as in (1.3). Specifically, the control dynamic we

consider is of the form

$$\begin{cases} \partial_t \mu_t = -v \cdot \nabla_x \mu_t + \sigma \Delta_v \mu_t - \operatorname{div}_v((\mathbf{v}[t, \boldsymbol{\mu}](z) \\ \quad + \mathfrak{w}[t, \mathbf{Y}(t), \mathbf{W}(t)](z)) \mu_t)) & (t, x, v) \in (0, T] \times \mathbb{R}^{2d} \\ \dot{\mathbf{Y}}(t) = \mathbf{W}(t) = F[t, \boldsymbol{\mu}](\mathbf{Y}(t)) + \mathbf{u}(t, \boldsymbol{\mu}) & t \in (0, T] \\ \mu_0 = \bar{\mu}, \quad \mathbf{Y}(0) = \bar{\mathbf{Y}}, \quad \mathbf{W}(0) = \bar{\mathbf{W}}, \end{cases} \quad (1.6)$$

under some specific assumptions thoroughly described in Section 4 and that allow for admissible controls  $\mathbf{u}(t, \boldsymbol{\mu})$  to allow the policy-maker to tune its action on the state of the system. In this sense, our results can be seen as a generalization to second order systems of control problems considered in [5]. However, a relevant difference is that we allow for a way more general structure of the drift  $\mathbf{v}[t, \boldsymbol{\mu}]$  considering unbounded Lipschitz frameworks. Then equation (4.1) is coupled with the cost functional

$$\mathcal{F}[\mathbf{u}] = \int_0^T \mathcal{L}(t, \mathbf{Y}, \boldsymbol{\mu}) dt + \int_0^T \Psi(\mathbf{u}(t, \boldsymbol{\mu})) dt,$$

where  $(\mathbf{Y}, \boldsymbol{\mu})$  is the solution of (4.1) with control  $\mathbf{u}$ ,  $\mathcal{L}$  is a Lagrangian functional accounting for closedness to the decided target and  $\Psi$  is a convex control cost. The specific assumptions on admissible controls and on the functional  $\mathcal{F}$  are discussed in Section 4. Thus, our second main result consists in showing the existence of a minimizer as it is stated below.

**Theorem 1.5.** *Suppose that Assumptions (v), (w), (F) and (F) hold and let  $\mathcal{A}$  be the class of admissible controls described via Assumptions (A). Then the optimal control problem*

<b>Problem 1</b>	
Find $\mathbf{u}^* \in \mathcal{A}$ such that	$\mathcal{F}[\mathbf{u}^*] = \min_{\mathbf{u} \in \mathcal{A}} \mathcal{F}[\mathbf{u}].$

*admits at least a solution.*

These control type results can be applied to various problems arising from real life phenomena. A model application is presented in Appendix D. It can be considered as a second order generalization of multiagent control problem analysed in [8, 9], as proposed, for instance, in [1, 2].

**1.4. Structure of our work.** The paper is structured as follows. In Section 2 we introduce the notation and we give some preliminaries on the involved spaces of functions and measures. In particular, in Section 2.5 we recall the statement of the well-posedness result in the linear case, as shown in [5], and, at the same time, we prove, in Proposition 2.3, a moment estimate that will be crucial in the remainder of the work.

Section 3 is entirely devoted to the proofs of Theorem 1.3 and Corollary 1.4. The proof of Theorem 1.3 is divided in 6 steps: first we prove the existence of the solution in case  $\mathbf{v}$  is *uniformly* sublinear (i.e., when  $\mathbf{v}_1$  holds without the additional term involving the

measure) for a finite time horizon; this is right after extended to the general case, still for a finite time horizon; next, uniqueness of the local solution is shown and then this is used to construct the unique global solution; once existence and uniqueness has been shown, the Hölder continuity (in the Wasserstein distance) and finally the higher moment estimate is shown. Corollary 1.4 represent a stability result in terms of suitable perturbations of the velocity field  $\mathbf{v}$ . Notice that, since  $\mathbf{v}_3$  could be not trivial to verify, sufficient conditions, including the metric dissipativity condition, are given in Appendix A.

In Section 4 we discuss the optimal control Problem 1 on the PDE-ODE system 1.6. Precisely, we first prove the existence and uniqueness of the solution of (1.6) for any given control  $\mathbf{u}$ , together with a stability result with respect to suitable perturbations of the control function. Technical results related to the PDE-ODE systems are left in Appendix B. Right after, we set the control problem and then we prove Theorem 1.5. Possible different classes of admissible controls are described in Appendix C while a specific example is given in Appendix D.

## 2. USEFUL NOTATION AND PRELIMINARY RESULTS

Throughout the paper,  $C$  will denote any constant depending at most on  $d, \alpha, \beta, \sigma, p, T$  and  $\bar{\mu}$  whose value is not relevant.

**2.1. Wasserstein spaces.** Let  $(\mathfrak{B}, d_{\mathfrak{B}})$  be a complete separable metric space and  $\mathcal{B}(\mathfrak{B})$  be the Borel  $\sigma$ -algebra on  $\mathfrak{B}$ . We denote by  $\mathcal{P}(\mathfrak{B})$  the space of (Borel) probability measures on  $\mathfrak{B}$ . For any  $\mu, \nu \in \mathcal{P}(\mathfrak{B})$ , let  $\Pi(\mu, \nu)$  be the set of (measure) couplings of  $\mu$  and  $\nu$  (see [46, Definition 1.1]), i.e. Borel probability measures  $\gamma \in \mathcal{P}(\mathfrak{B} \times \mathfrak{B})$  such that for any  $A, B \in \mathcal{B}(\mathfrak{B})$  we have

$$\gamma(A \times \mathfrak{B}) = \mu(A), \quad \gamma(\mathfrak{B} \times B) = \nu(B).$$

For any  $p \geq 1$  and  $x_0 \in \mathfrak{B}$  we denote by

$$M_p(\mu; x_0) = \int_{\mathfrak{B}} (d_{\mathfrak{B}}(x, x_0))^p d\mu(x)$$

the  $p$ -th moment (with respect to  $x_0 \in \mathfrak{B}$ ) of  $\mu \in \mathcal{P}(\mathfrak{B})$ . It is clear that if  $M_p(\mu; x_0) < \infty$  for some  $x_0 \in \mathfrak{B}$ , then it is finite for all  $x_0 \in \mathfrak{B}$ . The  $p$ -th Wasserstein space over  $\mathfrak{B}$  is defined as (see [46, Definition 6.4])

$$\mathcal{W}_p(\mathfrak{B}) := \{\mu \in \mathcal{P}(\mathfrak{B}) : M_p(\mu; x_0) < \infty, \forall x_0 \in \mathfrak{B}\},$$

equipped with the distance (see [46, Definition 6.1])

$$\mathcal{W}_p(\mu, \nu) = \inf_{\gamma \in \Pi(\mu, \nu)} \left( \int_{\mathfrak{B} \times \mathfrak{B}} (d_{\mathfrak{B}}(x, y))^p d\gamma(x, y) \right)^{\frac{1}{p}}.$$

It is well-known (see [46, Theorem 6.18]) that  $\mathcal{W}_p(\mathfrak{B})$  is a complete separable metric space. If  $\mathfrak{B}$  is a Banach space, then we will denote  $M_p(\mu) := M_p(\mu; 0)$ .

**2.2. Spaces of continuous functions and related Wasserstein spaces.** For any  $T > 0$ , we denote by  $C([0, T]; \mathfrak{B})$  the space of continuous functions  $f : [0, T] \rightarrow \mathfrak{B}$  equipped with the uniform distance, i.e.

$$d_\infty(f, g) = \sup_{t \in [0, T]} d_{\mathfrak{B}}(f(t), g(t)).$$

Clearly,  $C([0, T]; \mathfrak{B})$  is a complete separable metric space. Furthermore, we denote by  $C(\mathbb{R}_0^+; \mathfrak{B})$ , where  $\mathbb{R}_0^+ := [0, +\infty)$ , the space of continuous functions  $f : \mathbb{R}_0^+ \rightarrow \mathfrak{B}$ . In general, for reader's convenience, we will directly refer to  $C(\mathbb{R}_0^+; \mathfrak{B})$  as  $C([0, T]; \mathfrak{B})$  with  $T = \infty$ .

Furthermore, for any  $t \in [0, T]$ , we denote by  $\text{ev}_t$  the *evaluation map* defined on  $f \in C([0, T]; \mathfrak{B})$  as  $\text{ev}_t(f) = f(t)$ . For any given measure  $\mu \in \mathcal{P}(C([0, T]; \mathfrak{B}))$  and any  $t \in [0, T]$ , we denote

$$\mu_t = \text{ev}_t \# \mu,$$

where  $\#$  denotes the pushforward.

In what follows, we will use the notation  $\mu \in C([0, T]; \mathcal{W}_p(\mathfrak{B}))$  to denote continuous curves of probability measures, with  $\mu = \{\mu_t\}_{t \in [0, T]}$ . In particular, if  $\mathfrak{B}$  is a Banach space and  $T < \infty$ , also  $C([0, T]; \mathcal{W}_p(\mathfrak{B}))$  is a Banach space under the norm

$$(\overline{M}_p(T, \mu))^{\frac{1}{p}} := \sup_{0 \leq s \leq T} (M_p(\mu_s))^{\frac{1}{p}}.$$

We write  $\mu \in \mathcal{W}_p(C([0, T]; \mathfrak{B}))$  to state that there exists a Borel probability measure (that we still denote  $\mu$ ) on  $C([0, T]; \mathfrak{B})$  such that  $\mu_t = \text{ev}_t \# \mu$ . With this identification, we can say that  $\mathcal{W}_p(C([0, T]; \mathfrak{B})) \subset C([0, T]; \mathcal{W}_p(\mathfrak{B}))$ , since, clearly,

$$\overline{M}_p(T, \mu) \leq M_p(\mu),$$

where the second  $p$ -th order moment is taken in  $\mathcal{W}_p(C([0, T]; \mathfrak{B}))$ .

Notice that for any  $T_1 > T_2$  and any  $\mu = \{\mu_t\}_{t \in [0, T_1]} \in C([0, T_1]; \mathcal{W}_p(\mathfrak{B}))$  we still denote  $\mu = \{\mu_t\}_{t \in [0, T_2]}$  (i.e. truncating the flow at  $T_2$ ), so that with such an identification we have  $C([0, T_1]; \mathcal{W}_p(\mathfrak{B})) \subset C([0, T_2]; \mathcal{W}_p(\mathfrak{B}))$ .

For any  $\gamma \in (0, 1]$  and  $T > 0$  (with  $T \neq \infty$ ), we can consider the set  $C^\gamma([0, T]; \mathcal{W}_p(\mathfrak{B}))$  of  $\gamma$ -Hölder continuous curves, i.e.  $\mu \in C^\gamma([0, T]; \mathcal{W}_p(\mathfrak{B}))$  if and only if

$$\sup_{\substack{0 \leq t, s \leq T \\ t \neq s}} \frac{\mathcal{W}_p(\mu_t, \mu_s)}{|t - s|^\gamma} < \infty.$$

Thanks to the previous identification, we can also define the set  $C_{\text{loc}}^\gamma(\mathbb{R}_0^+; \mathcal{W}_p(\mathfrak{B}))$  of locally  $\gamma$ -Hölder continuous curves, that is to say that  $\mu \in C_{\text{loc}}^\gamma(\mathbb{R}_0^+; \mathcal{W}_p(\mathfrak{B}))$  if and only if  $\mu \in C^\gamma([0, T]; \mathcal{W}_p(\mathfrak{B}))$  for all  $T > 0$ .

**2.3. Random variables and couplings.** Throughout the paper, we will fix a probability space  $(\Omega, \Sigma, \mathbb{P})$  on which all the random quantities will be defined. In general, we call any measurable function  $X : \Omega \rightarrow \mathfrak{B}$  a  $\mathfrak{B}$ -valued random variable and we denote the set of all  $\mathfrak{B}$ -valued random variables as  $\mathcal{M}(\Omega; \mathfrak{B})$ .



The law of a random variable  $X \in \mathcal{M}(\Omega; \mathfrak{B})$  is defined as  $\text{Law}(X) = X\#P$ , i.e. for any  $A \in \mathfrak{B}(\mathfrak{B})$

$$\text{Law}(X)(A) = P(X^{-1}(A)).$$

We denote by  $L^p(\Omega; \mathfrak{B})$  the subspace of  $\mathcal{M}(\Omega; \mathfrak{B})$  of random variables  $X$  with the property that there exists  $x_0 \in \mathfrak{B}$  such that

$$E[(d_{\mathfrak{B}}(X, x_0))^p] < \infty, \quad (2.1)$$

where  $E$  is the expected value. It is clear that if (2.1) holds for a specific choice of  $x_0 \in \mathfrak{B}$ , then it holds for any  $x_0 \in \mathfrak{B}$ . We define a coupling of two probability measures  $\mu, \nu \in \mathcal{P}(\mathfrak{B})$  as any random variable  $(X, Y) \in \mathcal{M}(\Omega; \mathfrak{B} \times \mathfrak{B})$  such that  $\text{Law}(X) = \mu$  and  $\text{Law}(Y) = \nu$  (see [46, Definition 1.1]). It is clear that for any  $\mu, \nu \in \mathcal{P}(\mathfrak{B})$  and any measure coupling  $\gamma \in \Pi(\mu, \nu)$ , there exists a coupling  $(X, Y)$  such that  $\text{Law}(X, Y) = \gamma$ . Also note that  $\mu \in \mathcal{W}_p(\mathfrak{B})$  if and only if for any  $X \in \mathcal{M}(\Omega; \mathfrak{B})$  with  $\text{Law}(X) = \mu$  it holds  $X \in L^p(\Omega; \mathfrak{B})$ . In particular, this means that we can rewrite the Wasserstein distance between  $\mu$  and  $\nu$  as

$$\mathcal{W}_p^p(\mu, \nu) = \inf_{\substack{(X, Y) \in L^p(\Omega; \mathfrak{B} \times \mathfrak{B}) \\ \text{Law}(X) = \mu, \text{Law}(Y) = \nu}} E[(d_{\mathfrak{B}}(X, Y))^p].$$

For  $T > 0$  (possibly  $T = \infty$ ) and  $X \in \mathcal{M}(\Omega; C([0, T]; \mathfrak{B}))$ , we can define for any  $t \in [0, T]$

$$\begin{aligned} X(t) : \Omega &\mapsto \mathfrak{B} \\ \omega &\mapsto X(t)(\omega) = \text{ev}_t(X(\omega)). \end{aligned}$$

Hence,  $X(t) \in \mathcal{M}(\Omega; \mathfrak{B})$  for any  $t \in [0, T]$ , and if  $\text{Law}(X) = \mu$ , then  $\text{Law}(X(t)) = \mu_t$ . In particular, we can identify  $X$  as a  $\mathfrak{B}$ -valued stochastic process with continuous trajectories.

**2.4. Young functions.** Let us recall the definition of Young function: we say that  $\Phi$  is a Young function if for any  $x \geq 0$

$$\Phi(x) = \int_0^x \varphi(y) dy,$$

where

- (i)  $\varphi(0) = 0$  and  $\varphi(y) > 0$  for any  $y > 0$ ;
- (ii)  $\varphi$  is non-decreasing;
- (iii)  $\varphi$  is right-continuous;
- (iv)  $\lim_{y \rightarrow +\infty} \varphi(y) = +\infty$ .

We say that a Young function  $\Phi$  satisfies the  $\Delta_2$  condition if there exists a constant  $C_{\Phi} > 1$  such that for any  $x \geq 0$

$$\Phi(2x) \leq C_{\Phi} \Phi(x).$$

It is known (see [42, Remark 4.4.6]) that any Young function  $\Phi$  satisfying the  $\Delta_2$  condition is such that there exist two constants  $C_0 > 0$  and  $p > 1$  such that

$$\Phi(x) \leq C_0 x^p, \quad \forall x \geq 1.$$

The converse is not always true. By definition, Young functions are non-decreasing, convex, and satisfy

$$\lim_{x \rightarrow +\infty} \frac{\Phi(x)}{x} = +\infty$$

For further details on Young functions, we refer to [42, Chapter 4].

For our purposes, it will be useful to consider moments with respect to Young functions. Indeed, let  $\Phi : [0, +\infty) \rightarrow [0, +\infty)$  be a Young function and  $(\mathfrak{B}, d_{\mathfrak{B}})$  be a complete separable metric space. Then for any  $\mu \in \mathcal{P}(\mathfrak{B})$ , we define

$$M_{\Phi}(\mu; x_0) = \int_{\mathfrak{B}} \Phi(d_{\mathfrak{B}}(x, x_0)) d\mu(x)$$

and we say that  $\mu \in \mathcal{W}_{\Phi}(\mathfrak{B})$  if and only if there exists  $x_0 \in \mathfrak{B}$  such that  $M_{\Phi}(\mu; x_0) < \infty$  (or, equivalently,  $M_{\Phi}(\mu; x_0) < \infty$  for any  $x_0 \in \mathfrak{B}$ ). If  $\mathfrak{B}$  is a Banach space, then we set  $M_{\Phi}(\mu) := M_{\Phi}(\mu; 0)$ . Observe that, for instance, if we set  $\Phi_p(x) = x^p$  for  $p > 1$ , then  $\mathcal{W}_{\Phi}(\mathfrak{B}) = \mathcal{W}_p(\mathfrak{B})$ . Similarly to what we did for the moments of order  $p$ , we set

$$\overline{M}_{\Phi}(t, \boldsymbol{\mu}) = \sup_{0 \leq s \leq t} M_{\Phi}(\mu_s)$$

for any  $\boldsymbol{\mu} \in C([0, T]; \mathcal{W}_{\Phi}(\mathfrak{B}))$ , where  $\mathfrak{B}$  is a Banach space.

An important result that will be frequently employed is the de la Vallée-Poussin Theorem (see [36, Theorem T22]). More precisely, it states that a sequence  $X_n$  of random variables is uniformly integrable if and only if there exists a Young function  $\Phi : [0, +\infty) \rightarrow [0, +\infty)$  such that  $\Phi(x) \leq x^2$  and

$$\sup_n \mathbb{E} [\Phi(|X_n|)] < \infty.$$

We use this result in combination with the Dunford-Pettis Theorem (see [13, Theorem 4.30]). Indeed, if for some  $p \geq 1$  we have  $X \in L^p(\Omega)$ , then the set  $\{|X|^p\}$  is clearly a weakly compact set in  $L^1$ . Hence, it must be uniformly integrable and thus by the de la Vallée-Poussin Theorem we know that there exists a Young function such that  $\mathbb{E}[\Phi(|X|^p)] < \infty$ . This proves the following result.

**Proposition 2.1.** *Let  $\mu \in \mathcal{W}_p(\mathfrak{B})$ . Then there exists a Young function  $\Phi$  with the property that  $\Phi(x) \leq x^2$  and defining  $\Phi_p(x) = \Phi(x^p)$  we have  $\mu \in \mathcal{W}_{\Phi_p}(\mathfrak{B})$ .*

**2.5. The linear equation.** Given a function  $F : [0, +\infty) \times \mathbb{R}^{2d} \rightarrow \mathbb{R}^d$  we will consider the following set of assumptions:

- (A<sub>0</sub>)  $F : [0, +\infty) \times \mathbb{R}^{2d} \rightarrow \mathbb{R}^d$  is a Carathéodory map, i.e. it is measurable in the first variable and continuous in the second one;
- (A<sub>1</sub>) For any  $T > 0$ , there exist two constants  $C > 0$  and  $\beta \in (0, 1)$  such that

$$F(t, x, v) \leq C(1 + |x|^{\frac{\beta}{3}} + |v|^{\beta}),$$

for all  $t \in [0, T]$  and  $(x, v) \in \mathbb{R}^{2d}$ ;

(A<sub>2</sub>) For any  $T > 0$  and any compact set  $K \subset \mathbb{R}^{2d}$  there exist two constants  $L > 0$  and  $\alpha \in (\beta, 1]$  such that

$$|F(t, x_1, v_1) - F(t, x_2, v_2)| \leq L(|x_1 - x_2|^{\frac{\alpha}{3}} + |v_1 - v_2|^\alpha)$$

for all  $t \in [0, T]$  and  $(x_1, v_1), (x_2, v_2) \in K$ .

We also define

$$[F]_{\beta, T} := \sup_{\substack{t \in [0, T] \\ (x, v) \in \mathbb{R}^{2d}}} \frac{F(t, x, v)}{1 + |x|^{\frac{\beta}{3}} + |v|^\beta}.$$

We recall the following existence and uniqueness result for an associated linear system proved in [5].

**Theorem 2.2.** *Under Assumptions (A<sub>0</sub>), (A<sub>1</sub>) and (A<sub>2</sub>), for any  $\mu_0 \in \mathcal{W}_1(\mathbb{R}^{2d})$  and  $T > 0$ , the equation*

$$\begin{cases} \partial_t \mu_t = -v \cdot \nabla_x \mu_t + \sigma \Delta_v \mu_t - \operatorname{div}_v (F(t, z) \mu_t) & (t, x, v) \in \mathbb{R}_0^+ \times \mathbb{R}^{2d}, \\ \mu_0 = \bar{\mu} & (x, v) \in \mathbb{R}^{2d}, \end{cases} \quad (2.2)$$

admits exactly one global solution  $\boldsymbol{\mu} := \{\mu_t\}_{t \in \mathbb{R}_0^+} \in C(\mathbb{R}_0^+; \mathcal{W}_1(\mathbb{R}^{2d}))$ , which can be characterized as follows. Let  $B$  be a  $d$ -dimensional Brownian motion and  $(X_0, V_0) \in \mathcal{M}(\Omega; \mathbb{R}^{2d})$  be independent of  $B$  and such that  $\operatorname{Law}(X_0, V_0) = \bar{\mu}$ , then there exists a stochastic process  $(X, V) \in L^1(\Omega; C([0, T]; \mathbb{R}^{2d}))$  for all  $T > 0$  such that  $\boldsymbol{\mu} = \operatorname{Law}(X, V)$  and  $(X, V)$  is the pathwise unique strong solution to

$$\begin{cases} dX(t) = V(t) dt & t \in \mathbb{R}_0^+ \\ dV(t) = F(t, X(t), V(t)) dt + \sqrt{2\sigma} dB(t) & t \in \mathbb{R}_0^+ \\ X(0) = X_0, \quad V(0) = V_0. \end{cases} \quad (2.3)$$

Actually, we can show that the solution  $\boldsymbol{\mu}$  is slightly more regular, depending on the regularity of the initial data  $\bar{\mu}$ . Before doing this, we state an inequality that we will use often throughout the paper, whose interest is not only related to the proof of the next proposition. Precisely, for any  $T > 0$  and  $p > 1$  it holds

$$\mathbb{E} \left[ \sup_{t \in [0, T]} |B(t)|^p \right] \leq \frac{1}{\sqrt{\pi}} \left( \frac{2p\sqrt{T}}{p-1} \right)^p \Gamma \left( \frac{p+1}{2} \right), \quad (2.4)$$

that follows by Doob's maximal inequality (see [43, Theorem II.1.7]) and the formula for the absolute  $p$ -moments of a Gaussian random variable [.....].

We have the following result.

**Proposition 2.3.** *Let Assumptions (A<sub>0</sub>), (A<sub>1</sub>) and (A<sub>2</sub>) hold true. Let  $\Phi : [0, +\infty) \rightarrow [0, +\infty)$  be either the identity or a Young function satisfying  $\Phi(x) \leq x^2$  and define  $\Phi_p(x) = \Phi(x^p)$  for any  $x \geq 0$  and  $p \geq 1$ . Let  $\boldsymbol{\mu} \in C([0, T]; \mathcal{W}_1(\mathbb{R}^d))$  be the unique solution of (2.2). If*

$\bar{\mu} \in \mathcal{W}_{\Phi_p}(\mathbb{R}^{2d})$  for some  $p \geq 1$ , then  $\boldsymbol{\mu} \in \mathcal{W}_{\Phi_p}(C([0, T]; \mathbb{R}^{2d}))$  and there exist two constants  $\bar{C}_1, \bar{C}_2$  depending only on  $p, \beta, T$  and  $\sigma$  such that

$$\bar{M}_{\tilde{\Phi}_p}(T, \boldsymbol{\mu}) \leq M_{\tilde{\Phi}_p}(\boldsymbol{\mu}) \leq \bar{C}_2 (1 + M_{\Phi_p}(\bar{\mu}) + [F]_{\beta, T}^{2p}), \quad (2.5)$$

where

$$\tilde{\Phi}(r) = \Phi \left( \frac{e^{-\bar{C}_1 [F]_{\beta, T}^p r}}{3\bar{C}_1} r \right).$$

If, furthermore,  $\Phi$  satisfies the  $\Delta_2$  condition, then there exists a constant  $\bar{C}_3$  depending only on  $\Phi, p, \beta, T$  and  $\sigma$  such that

$$\bar{M}_{\Phi_p}(T, \boldsymbol{\mu}) \leq \bar{C}_3 (1 + M_{\Phi_p}(\bar{\mu}) + [F]_{\beta, T}^{2p}) e^{\bar{C}_3 [F]_{\beta, T}^p}. \quad (2.6)$$

Finally, if we consider the solution  $Z = (X, V)$  of (2.3), it also holds, almost surely,

$$\sup_{0 \leq s \leq T} |Z(s)| \leq C_4 \left( |Z_0| + [F]_{\beta, T} + \sigma \sup_{0 \leq s \leq T} |B(s)| \right) e^{C_4 ([F]_{\beta, T} + 1)T}. \quad (2.7)$$

*Proof.* Let us first show that if  $\Phi$  satisfies the  $\Delta_2$  condition, then (2.5) implies (2.6). Indeed, in such a case, since  $\Phi(2x) \leq C_\Phi \Phi(x)$  for a constant  $C_\Phi > 1$ , we have

$$\Phi(x^p) \leq C_\Phi^{|\log_2(2\bar{C}_1)|} e^{\bar{C}_1 \log(C_\Phi) [F]_{\beta, T}^p} \tilde{\Phi}(x^p).$$

This is enough to prove the desired implication.

Now we prove (2.5). Let  $Z = (X, V)$  be any strong solution of (2.3) and recall that  $\boldsymbol{\mu} = \text{Law}(Z)$ . By Assumption (A<sub>1</sub>) and the second equation in (2.3) we know that there exists a constant  $C_1$  depending on  $p$  and  $T$  such that

$$|V(s)|^p \leq C_1 \left( |V_0|^p + [F]_{\beta, T}^p + [F]_{\beta, T}^p \int_0^t \left( |X(\tau)|^{\frac{p\beta}{3}} + |V(\tau)|^{p\beta} \right) d\tau + \sup_{0 \leq s \leq t} \sigma^p |B(s)|^p \right)$$

By using Young's inequality on  $|X(\tau)|^{\frac{p\beta}{3}}$  and  $|V(\tau)|^{p\beta}$  with exponents respectively  $\frac{3}{\beta}$  and  $\frac{1}{\beta}$  and taking the supremum inside the integral, we get

$$|V(s)|^p \leq C_2 \left( |V_0|^p + [F]_{\beta, T}^p + [F]_{\beta, T}^p \int_0^t \sup_{0 \leq z \leq \tau} (|X(z)|^p + |V(z)|^p) d\tau + \sigma^p \sup_{0 \leq s \leq t} |B(s)|^p \right) \quad (2.8)$$

for some constant  $C_2$  depending only on  $p, T$  and  $\beta$ . Moreover, we also have, by the first equation in (2.3),

$$|X(s)|^p \leq C_3 \left( |X_0|^p + \int_0^t \sup_{0 \leq z \leq \tau} |V(z)|^p d\tau \right), \quad (2.9)$$

where  $C_3$  is a constant depending only on  $p$  and  $T$ . Summing (2.9) and (2.8) and taking the supremum, it is clear that

$$\sup_{0 \leq s \leq t} |Z(s)|^p \leq C_4 \left( |Z_0|^p + [F]_{\beta, T}^p + ([F]_{\beta, T}^p + 1) \int_0^t \sup_{0 \leq z \leq \tau} |Z(z)|^p d\tau + \sigma^p \sup_{0 \leq s \leq t} |B(s)|^p \right),$$

where  $C_4 = \max\{C_2, C_3\}$  and we can use Grönwall's inequality to get

$$\sup_{0 \leq s \leq t} |Z(s)|^p \leq C_4 \left( |Z_0|^p + [F]_{\beta, T}^p + \sigma^p \sup_{0 \leq s \leq t} |B(s)|^p \right) e^{C_4([F]_{\beta, T}^p + 1)T}. \quad (2.10)$$

For  $p = 1$ , (2.10) implies (2.7).

Now we apply the function  $\Phi$  on both sides of (2.10), obtaining, by convexity and the bound  $\Phi(x) \leq x^2$ , setting  $\bar{C}_1 = \frac{1}{3C_4}$ ,

$$\tilde{\Phi} \left( \sup_{0 \leq s \leq t} |Z(s)|^p \right) \leq \frac{1}{3} \Phi(|Z_0|^p) + [F]_{\beta, T}^{2p} + \sigma^{2p} \sup_{0 \leq s \leq t} |B(s)|^{2p}. \quad (2.11)$$

Finally, we take the expectation on both sides of (2.11) and we use (2.4) to achieve (2.5).  $\square$

### 3. PROOF OF THEOREM 1.3 AND COROLLARY 1.4

**3.1. Proof of Theorem 1.3.** We proceed in four steps: first we prove the existence of a solution under a stronger assumption; then we weaken such assumption to get the actual existence result; next we prove uniqueness of the local solution; finally, we prove existence and uniqueness of the global solution.

**3.1.1. Existence with  $\mathfrak{v}$  uniformly sublinear.** We first assume that  $\mathfrak{v}$  satisfies  $(\mathfrak{v}_0)$ ,  $(\mathfrak{v}_2)$  and  $(\mathfrak{v}_3)$ , while we substitute  $(\mathfrak{v}_1)$  with the following further assumption:

$(\bar{\mathfrak{v}}_1)$  For all  $T > 0$  there exist two constants  $K > 0$  and  $\beta \in [0, 1)$  such that

$$|\mathfrak{v}[t, \mu](z)| \leq K \left( 1 + |x|^{\frac{\beta}{3}} + |v|^\beta \right)$$

for every  $z \in \mathbb{R}^{2d}$ ,  $t \in [0, T]$  and for every  $\mu \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$ .

Let  $T > 0$  (with  $T \neq +\infty$ ) and  $(X_0, V_0)$  be any random variable independent of  $B$  with  $\text{Law}(X_0, V_0) = \bar{\mu}$ . Fix any curve of probability measures  $\nu \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$  and consider the following SDE.

$$\begin{cases} dX(t) = V(t)dt & t \in (0, T] \\ dV(t) = F_\nu(t, X(t), V(t))dt + \sqrt{2\sigma}dB(t) & t \in (0, T] \\ X(0) = X_0, \quad V(0) = V_0, \end{cases} \quad (3.1)$$

where  $F_\nu(t, x, v) = \mathfrak{v}[t, \nu](x, v)$ . Notice that, by Assumptions  $(\mathfrak{v}_0)$ ,  $(\bar{\mathfrak{v}}_1)$  and  $(\mathfrak{v}_2)$ , we know that the function  $F_\nu$  satisfies Assumptions  $(A_0)$ ,  $(A_1)$  and  $(A_2)$ , hence, by Theorem 2.2, there exists a unique strong solution  $(X_\nu, V_\nu) \in L^1(\Omega; C([0, T]; \mathbb{R}^{2d}))$  of (3.1). Let  $\tilde{\nu} = \text{Law}(X_\nu, V_\nu)$ . By Proposition 2.1, we know that there exists a Young functions  $\Phi$  such that  $\Phi(x) \leq x^2$  and  $\bar{\mu} \in \mathcal{W}_{\Phi_p}(\mathbb{R}^{2d})$ , where  $\Phi_p(r) = \Phi(r^p)$ . Thus, by Proposition 2.3, in particular (2.6), we know that

$$M_p(\tilde{\nu}) \leq C(1 + [F_\nu]_{\beta, T}^{2p})e^{C[F_\nu]_{\beta, T}^p} \leq C(1 + K^{2p})e^{CK^p}, \quad (3.2)$$

where we also used Assumption  $(\bar{\mathbf{v}}_1)$ , and

$$M_{\tilde{\Phi}_p}(\tilde{\nu}) \leq C(1 + [F_\nu]_{\beta,T}^{2p}) \leq C(1 + K^{2p}), \quad (3.3)$$

where

$$\tilde{\Phi}^\nu(r) = \Phi \left( C e^{-C[F_\nu]_{\beta,T}^p r} \right).$$

It is clear, since  $\Phi$  is increasing, that for all  $r \geq 0$

$$\bar{\Phi}(r) := \Phi \left( C e^{-CK^p r} \right) \leq \Phi \left( C e^{-C[F_\nu]_{\beta,T}^p r} \right) = \tilde{\Phi}^\nu(r),$$

hence, by (3.3) we have

$$M_{\bar{\Phi}_p}(\tilde{\nu}) \leq C(1 + K^{2p}). \quad (3.4)$$

Notice that both bounds (3.2) and (3.3) are uniform with respect to  $\nu$ . It is also clear, by (3.2), that  $(X_\nu, Y_\nu) \in L^p(\Omega; C([0, T]; \mathbb{R}^{2d}))$ . Then we can consider the map

$$\mathcal{T} : \nu \in \mathcal{W}_p(C([0, T]; \mathbb{R}^{2d})) \mapsto \mathcal{T}\nu = \text{Law}(X_\nu, Y_\nu) \in \mathcal{W}_p(C([0, T]; \mathbb{R}^{2d})).$$

We want to show that  $\mathcal{T}$  admits a fixed point. To do this, we first prove that

$$\text{Im}(\mathcal{T}) = \{ \mathcal{T}\nu : \nu \in \mathcal{W}_p(C([0, T]; \mathbb{R}^{2d})) \}$$

is relatively (sequentially) compact in  $\mathcal{W}_p(C([0, T]; \mathbb{R}^{2d}))$ . First, we claim that the family  $\{(X_\nu, V_\nu)\}_{\nu \in \mathcal{W}_p(C([0, T]; \mathbb{R}^{2d}))}$  is uniformly tight, i.e. for any  $\varepsilon > 0$  there exists a compact set  $K_\varepsilon \subset C([0, T]; \mathbb{R}^{2d})$  such that

$$\sup_{\nu \in \mathcal{W}_p(C([0, T]; \mathbb{R}^{2d}))} \mathbb{P}((X_\nu, V_\nu) \notin K_\varepsilon) < \varepsilon.$$

To prove this, we first fix  $\gamma < 1/2$  and consider the random variable

$$L_{B,\gamma} = \sup_{\substack{t,s \in [0,T] \\ t \neq s}} \frac{|B(t) - B(s)|}{|t - s|^\gamma},$$

that is a.s. finite by the  $\gamma$ -Hölder continuity of the Brownian motion. By Fernique's theorem (see [21, Theorem 1.3.2]), we know that there exists  $\alpha > 0$  such that  $\mathbb{E} [e^{\alpha L_{B,\gamma}}] < \infty$ , which in turn implies that  $\mathbb{E} [|L_{B,\gamma}|^n] < \infty$  for all  $n \in \mathbb{N}$ . Furthermore, we observe that

$$\begin{aligned} |V_\nu(t) - V_\nu(s)| &\leq \int_s^t |\mathbf{v}[\tau, \nu](X_\nu(\tau), V_\nu(\tau))| d\tau + \sqrt{2\sigma} L_{B,\gamma} |t - s|^\gamma \\ &\leq K|t - s| + K \int_s^t (|X_\nu(\tau)|^{\frac{\beta}{3}} + |V_\nu(\tau)|^\beta) d\tau + \sqrt{2\sigma} L_{B,\gamma} |t - s|^\gamma \\ &\leq CK(1 + \sup_{\tau \in [0,T]} |Z_\nu(\tau)|) |t - s| + \sqrt{2\sigma} L_{B,\gamma} |t - s|^\gamma \\ &\leq CK(1 + |Z_0| + K + \sup_{\tau \in [0,T]} |B(\tau)| + L_{B,\gamma}) e^{CK} |t - s|^\gamma, \end{aligned} \quad (3.5)$$

where we used Assumption  $(\bar{v}_1)$  in the second inequality, Young's inequality with exponents  $\frac{3}{\beta}$  and  $\frac{1}{\beta}$  in the third inequality and (2.7) in the fourth inequality. Furthermore, we easily have

$$|X_\nu(t) - X_\nu(s)| \leq \left( \sup_{\tau \in [0, T]} |V_\nu(\tau)| \right) |t - s| \leq C \left( 1 + |Z_0| + K + \sup_{0 \leq \tau \leq T} |B(\tau)| \right) |t - s|, \quad (3.6)$$

where we used again (2.7). Consider, for any  $\delta \in (0, T]$  and  $n \in \mathbb{N}$ , the (random) modulus of continuity

$$m_\nu(\delta) = \sup_{\substack{t, s \in (0, T] \\ |t - s| < \delta}} |(X_\nu(t), V_\nu(t)) - (X_\nu(s), V_\nu(s))|.$$

Then we have, by (3.5) and (3.6),

$$m_\nu(\delta) \leq C \left( 1 + |Z_0| + K + \sup_{\tau \in [0, T]} |B(\tau)| + L_{B, \gamma} \right) \delta^\gamma$$

and taking the expectation

$$\mathbf{E} [m_\nu(\delta)] \leq C \left( 1 + M_1(\bar{\mu}) + K + \mathbf{E} \left[ \sup_{\tau \in [0, T]} |B(\tau)| \right] + \mathbf{E} [L_{B, \gamma}] \right) \delta^\gamma \leq C(1 + K) \delta^\gamma$$

where we used the inequality  $\mathbf{E}[\sup_{\tau \in [0, T]} |B(\tau)|] \leq \sqrt{\mathbf{E}[\sup_{\tau \in [0, T]} |B(\tau)|^2]}$  and (2.4). Now fix any  $\varepsilon > 0$  and any  $\eta \in (0, 1]$ . By Markov's inequality we have

$$\mathbf{P}(m_\nu(\delta) \geq \varepsilon) \leq \frac{\mathbf{E}[m_\nu(\delta)]}{\varepsilon} \leq C(1 + K) \frac{\delta^\gamma}{\varepsilon}.$$

Hence, if we set  $\delta_0 = \frac{(\varepsilon\eta)^{\frac{1}{\gamma}}}{C(1+K)}$ , we get, for any  $\delta < \delta_0$  and any  $n \in \mathbb{N}$ ,

$$\mathbf{P}(m_n(\delta) \geq \varepsilon) \leq \eta.$$

Combining this with the fact that  $\text{Law}(X_\nu(0), V_\nu(0)) = \bar{\mu}$  for any  $n \in \mathbb{N}$ , we conclude that the sequence  $\{(X_\nu, V_\nu)\}_{n \in \mathbb{N}}$  is uniformly tight by [11, Theorem 7.3].

Next, we claim that  $\{(X_\nu, V_\nu)\}_{\nu \in \mathcal{W}_p(C([0, T]; \mathbb{R}^{2d}))}$  are uniformly  $L^p$ -integrable, i.e., setting  $\tilde{\nu} = \text{Law}(X_\nu, Y_\nu)$  and  $B_R = \{Z \in C([0, T]; \mathbb{R}^{2d}), \sup_{t \in [0, T]} |Z(t)| < R\}$ , we have

$$\sup_{\nu \in \mathcal{W}_p(C([0, T]; \mathbb{R}^{2d}))} \lim_{R \rightarrow +\infty} \int_{C([0, T]; \mathbb{R}^{2d}) \setminus B_R} \sup_{t \in [0, T]} |Z(t)|^p d\tilde{\nu}(Z) = 0.$$

Indeed, (3.4) guarantees that there exists a Young function  $\bar{\Phi}$  such that

$$\sup_{\nu \in \mathcal{W}_p(C([0, T]; \mathbb{R}^{2d}))} \mathbf{E} \left[ \bar{\Phi} \left( \sup_{\tau \in [0, T]} |(X_\nu(\tau), V_\nu(\tau))|^p \right) \right] \leq C(1 + K^{2p}),$$

which in turn implies the uniform  $L^p$ -integrability by the de la Vallée-Poussin Theorem. Thus, by the characterization of relatively compact sets in Lebesgue-Bochner spaces, shown in [47, Theorem 2.1], we know that

$$\text{Im}^*(\mathcal{T}) := \{(X_\nu, Y_\nu) \in L^p(\Omega; C([0, T]; \mathbb{R}^{2d})), \nu \in \mathcal{W}_p(C([0, T]; \mathbb{R}^{2d}))\}$$

is relatively compact in  $L^p(\Omega; C([0, T]; \mathbb{R}^{2d}))$ , while by [4, Proposition 7.1.5] we get that  $\text{Im}(\mathcal{T})$  is relatively compact.

Now we prove that  $\mathcal{T}$  is continuous. Indeed, consider any sequence  $\{\nu^n\}_{n \in \mathbb{N}} \subset \mathcal{W}_p(C([0, T]; \mathbb{R}^{2d}))$  converging towards  $\nu \in \mathcal{W}_p(C([0, T]; \mathbb{R}^{2d}))$  and denote by  $(X_n, V_n)$  the solutions of (3.1) with respect to  $\nu^n$  for each  $n \in \mathbb{N}$ . Assume, without loss of generality, that

$$\limsup_{n \rightarrow +\infty} \mathcal{W}_p(\mathcal{T}\nu^n, \mathcal{T}\nu) = \lim_{n \rightarrow +\infty} \mathcal{W}_p(\mathcal{T}\nu^n, \mathcal{T}\nu)$$

and recall that there exists a subsequence  $\{(X_{n_k}, V_{n_k})\}_{k \in \mathbb{N}}$  converging towards  $(X_\infty, V_\infty) \in L^p(\Omega; C([0, T]; \mathbb{R}^{2d}))$  both in the  $L^p$ -norm and almost surely. Now we show that the limit  $(X_\infty, V_\infty)$  is the solution of (3.1) associated with  $\nu$ . Indeed, by Assumption (v<sub>0</sub>), we know that for any  $t \in [0, T]$  and almost surely

$$\lim_{k \rightarrow \infty} \mathfrak{v}[t, \nu^{n_k}](X_{n_k}(t), V_{n_k}(t)) = \mathfrak{v}[t, \nu](X_\infty(t), V_\infty(t)). \quad (3.7)$$

Now fix any  $\omega \in \Omega$  such that (3.7) holds. In the following we will omit the dependence on  $\omega$  for the ease of the reader. By Assumption (v<sub>1</sub>), it holds

$$\begin{aligned} \mathfrak{v}[t, \nu^{n_k}](X_{n_k}(t), V_{n_k}(t)) &\leq K(1 + |X_{n_k}(t)|^{\frac{\beta}{3}} + |V_{n_k}(t)|^\beta) \leq CK(1 + \sup_{0 \leq s \leq T} |Z_{n_k}(t)|) \\ &\leq CK(1 + |Z_0| + K + \sup_{0 \leq s \leq T} |B(s)|)e^{CK}, \end{aligned}$$

where the upper bound is independent of  $t \in [0, T]$  and  $n \in \mathbb{N}$ . Hence, by the dominated convergence theorem, we get

$$\begin{aligned} V_\infty(t) &= \lim_{k \rightarrow \infty} V_{n_k}(t) = V_0 + \lim_{k \rightarrow \infty} \int_0^t \mathfrak{v}[s, \nu^{n_k}](X_{n_k}(s), V_{n_k}(s)) ds + \sqrt{2\sigma}B(s) \\ &= V_0 + \int_0^t \mathfrak{v}[s, \nu](X_\infty(s), V_\infty(s)) ds + \sqrt{2\sigma}B(s), \end{aligned}$$

while, since  $V_{n_k} \rightarrow V_\infty$  uniformly almost surely,

$$X_\infty(t) = \lim_{k \rightarrow \infty} X_{n_k}(t) = X_0 + \lim_{k \rightarrow \infty} \int_0^t V_{n_k}(s) ds = X_0 + \int_0^t V_\infty(s) ds,$$

i.e.  $(X_\infty, V_\infty)$  solves (3.1) with respect to  $\nu$ . This shows that for all  $k \in \mathbb{N}$   $(X_{n_k}, V_{n_k})$  and  $(X_\infty, V_\infty)$  constitute a coupling of  $\mathcal{T}\nu^{n_k}$  and  $\mathcal{T}\nu$  and then

$$\begin{aligned} \lim_{n \rightarrow +\infty} \mathcal{W}_p(\mathcal{T}\nu^n, \mathcal{T}\nu) &= \lim_{k \rightarrow +\infty} \mathcal{W}_p(\mathcal{T}\nu^{n_k}, \mathcal{T}\nu) \\ &\leq \lim_{k \rightarrow +\infty} \mathbb{E} \left[ \sup_{t \in [0, T]} |(X_{n_k}(t), V_{n_k}(t)) - (X_\infty(t), V_\infty(t))|^p \right] = 0. \end{aligned}$$



Since  $\nu \in W_p(C([0, T]; \mathbb{R}^{2d}))$  is arbitrary, this implies that  $\mathcal{T}$  is continuous.

We can then use Schauder's fixed point theorem to ensure that there exists  $\mu \in \mathcal{W}_p(C([0, T]; \mathbb{R}^{2d}))$  such that  $\mu = \mathcal{T}\mu$ . We only need to show that  $\mu$  is a solution of (1.1). To do this, let  $(X, V)$  be the solution of (3.1) associated with  $\mu$ . Then, since  $\mu = \mathcal{T}\mu = \text{Law}(X, V)$ , the process  $(X, V)$  solves the McKean-Vlasov SDE (1.3). To obtain (1.2), let  $\psi \in C_c^\infty(\mathbb{R}^{2d})$ . By Itô's formula [39, Theorem 4.1.2] we have

$$\begin{aligned} \psi(X(t), V(t)) &= \psi(X_0, V_0) + \int_0^t (V(s) \cdot \nabla_x \psi(X(s), V(s)) \\ &\quad + \mathbf{v}[s, \mu](X(s), V(s)) \cdot \nabla_v \psi(X(s), V(s)) + \sigma \Delta_v \psi(X(s), V(s))) ds \\ &\quad + \sqrt{2\sigma} \sum_{j=1}^d \int_0^s \frac{\partial \psi}{\partial v_j}(X(s), V(s)) dB_j(s), \end{aligned}$$

where  $B_j$  is the  $j$ -th coordinate of the  $d$ -dimensional Brownian motion  $B = (B_1, \dots, B_d)$ . Since  $\psi \in C_c^\infty(\mathbb{R}^{2d})$ , it is clear that for all  $j = 1, \dots, d$  it holds

$$\left| \frac{\partial \psi}{\partial v_j}(X(s), V(s)) \right| \leq \|\nabla \psi\|_{L^\infty(\mathbb{R}^{2d})}, \quad \forall s \geq 0,$$

hence  $\int_0^s \frac{\partial \psi}{\partial v_j}(X(s), V(s)) dB_j(s)$  is a martingale [39, Corollary 3.2.6] and then

$$\mathbb{E} \left[ \int_0^s \frac{\partial \psi}{\partial v_j}(X(s), V(s)) dB_j(s) \right] = 0, \quad j = 1, \dots, d.$$

As a consequence we get

$$\begin{aligned} \mathbb{E}[\psi(X(t), V(t))] &= \mathbb{E}[\psi(X_0, V_0)] + \mathbb{E} \left[ \int_0^t (V(s) \cdot \nabla_x \psi(X(s), V(s)) \right. \\ &\quad \left. + \mathbf{v}[s, \mu](X(s), V(s)) \cdot \nabla_v \psi(X(s), V(s)) + \sigma \Delta_v \psi(X(s), V(s))) ds \right] \end{aligned} \quad (3.8)$$

Next, notice that

$$\begin{aligned} |V(s) \cdot \nabla_x \psi(X(s), V(s))| &\leq |Z(s)| \|\nabla \psi\|_{L^\infty(\mathbb{R}^{2d})} \\ |\mathbf{v}[s, \mu](X(s), V(s)) \cdot \nabla_v \psi(X(s), V(s))| &\leq CK(1 + |Z(s)|) \|\nabla \psi\|_{L^\infty(\mathbb{R}^{2d})} \\ \sigma \Delta_v \psi(X(s), V(s)) &\leq \sigma \|\Delta \psi\|_{L^\infty(\mathbb{R}^{2d})}, \end{aligned}$$

hence

$$\begin{aligned}
& \mathbb{E} \left[ \int_0^t |V(s) \cdot \nabla_x \psi(X(s), V(s)) \right. \\
& \quad \left. + \mathbf{v}[s, \boldsymbol{\mu}](X(s), V(s)) \cdot \nabla_v \psi(X(s), V(s)) + \sigma \Delta_v \psi(X(s), V(s))| ds \right] \\
& \leq C(1+K)(\|\nabla \psi\|_{L^\infty(\mathbb{R}^{2d})} + \|\Delta \psi\|_{L^\infty(\mathbb{R}^{2d})}) \int_0^t \mathbb{E}[1 + |Z(s)|] ds \\
& \leq C(1+K)(\|\nabla \psi\|_{L^\infty(\mathbb{R}^{2d})} + \|\Delta \psi\|_{L^\infty(\mathbb{R}^{2d})})(1 + \overline{M}_1(T; \boldsymbol{\mu}))T,
\end{aligned}$$

where we recall, by Hölder's inequality, that  $\overline{M}_1(T; \boldsymbol{\mu}) \leq (\overline{M}_p(T; \boldsymbol{\mu}))^{\frac{1}{p}} < \infty$ . Thus we can use Fubini's theorem in (3.8) to achieve

$$\begin{aligned}
\mathbb{E}[\psi(X(t), V(t))] &= \mathbb{E}[\psi(X_0, V_0)] + \int_0^t \mathbb{E}[(V(s) \cdot \nabla_x \psi(X(s), V(s)) \\
& \quad + \mathbf{v}[s, \boldsymbol{\mu}](X(s), V(s)) \cdot \nabla_v \psi(X(s), V(s)) + \sigma \Delta_v \psi(X(s), V(s)))] ds.
\end{aligned}$$

Now we use the fact that  $\mu_t = \text{Law}(X(t), V(t))$  and  $\bar{\mu} = \text{Law}(X_0, V_0)$  to finally get

$$\int_{\mathbb{R}^{2d}} \psi d\mu_t = \int_{\mathbb{R}^{2d}} \psi d\bar{\mu} + \int_0^t \int_{\mathbb{R}^{2d}} (v \cdot \nabla_x \psi + \mathbf{v}[s, \boldsymbol{\mu}](z) \cdot \nabla_v \psi + \sigma \Delta_v \psi) d\mu_s ds,$$

that is (1.2).

**3.1.2. Existence for a general  $\mathbf{v}$ .** Now, let us consider  $\mathbf{v}$  as in the statement of Theorem 1.3. For any  $N > 0$  let  $\eta_N \in C_c^\infty(\mathbb{R})$  such that  $\text{supp}(\eta_N) \subset [-1, N+1]$ ,  $\eta_N(r) = 1$  for any  $r \in [0, N]$  and  $\eta_N(x) \in [0, 1]$  for any  $r \in \mathbb{R}$ . Define, for any  $t \in [0, T]$ ,  $\boldsymbol{\mu} \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$  and  $z \in \mathbb{R}^{2d}$ ,

$$\mathbf{v}_N[t, \boldsymbol{\mu}](z) = \mathbf{v}[t, \boldsymbol{\mu}](z) \eta_N(\overline{M}_p(t, \boldsymbol{\mu}))$$

By definition,  $\mathbf{v}_N$  satisfies  $(\mathbf{v}_0)$ ,  $(\mathbf{v}_2)$  and  $(\mathbf{v}_4)$ . Furthermore, notice that

$$|\mathbf{v}_N[t, \boldsymbol{\mu}](z)| \leq K \left( 1 + |x|^{\frac{\beta}{3}} + |y|^\beta + (\overline{M}_p(T, \boldsymbol{\mu}))^{\frac{1}{p}} \right) \eta_N(\overline{M}_p(t, \boldsymbol{\mu})).$$

If  $\overline{M}_p(t, \boldsymbol{\mu}) \leq N+1$ , then

$$|\mathbf{v}_N[t, \boldsymbol{\mu}](z)| \leq K(N+1)^{\frac{1}{p}} \left( 1 + |x|^{\frac{\beta}{3}} + |y|^\beta \right),$$

otherwise

$$|\mathbf{v}_N[t, \boldsymbol{\mu}](z)| = 0 \leq K(N+1)^{\frac{1}{p}} \left( 1 + |x|^{\frac{\beta}{3}} + |y|^\beta \right),$$

hence  $\mathbf{v}_N$  satisfies  $(\bar{\mathbf{v}}_1)$ . To use the first claim of this proof, it is not necessary to check whether  $(\mathbf{v}_3)$  still holds. For any  $N \in \mathbb{N}$ , we can find a solution  $\boldsymbol{\mu}^N \in \mathcal{W}_p(C([0, T]; \mathbb{R}^{2d}))$  of (1.1) with drift  $\mathbf{v}_N$  and a stochastic process  $(X_N, Y_N) \in L^p(\Omega; C([0, T]; \mathbb{R}^{2d}))$  solving (1.3) such that  $\boldsymbol{\mu}^N = \text{Law}(X_N, V_N)$ . Now we need to prove a further uniform bound

on the moments  $\overline{M}_p(t, \boldsymbol{\mu}^N)$ . To do this we argue as in the proof of Proposition 2.3. Fix  $t \in [0, T]$  and observe that for any  $s \leq t$

$$\begin{aligned} |V_N(s)|^p &\leq C \left( |V_0|^p + \int_0^s |\mathbf{v}_N[\tau, \boldsymbol{\mu}^N](X_N(\tau), V_N(\tau))|^p d\tau + \sup_{0 \leq \tau \leq T} |B(\tau)|^p \right) \\ &\leq C \left( |V_0|^p + K + K \int_0^s \left( |X_N(\tau)|^{\frac{p\beta}{3}} + |V_N(\tau)|^{p\beta} + \overline{M}_p(\tau; \boldsymbol{\mu}^N) \right) d\tau + \sup_{0 \leq \tau \leq T} |B(\tau)|^p \right) \\ &\leq C \left( |V_0|^p + K + K \int_0^s (|X_N(\tau)|^p + |V_N(\tau)|^p) d\tau + K \int_0^t \overline{M}_p(\tau; \boldsymbol{\mu}^N) d\tau + \sup_{0 \leq \tau \leq T} |B(\tau)|^p \right), \end{aligned}$$

where we used the fact that  $|\mathbf{v}_N[\tau, \boldsymbol{\mu}^N](z)| \leq |\mathbf{v}[\tau, \boldsymbol{\mu}^N](z)|$ , Assumption  $(\mathbf{v}_1)$  and Young's inequality with exponents  $\frac{3}{\beta}$  and  $\frac{1}{\beta}$ . On the other hand, for any  $s \leq t$

$$|X_N(t)|^p \leq C \left( |X_0|^p + \int_0^s |V_N(\tau)|^p d\tau \right)$$

and then, setting  $Z_N = (X_N, V_N)$ ,

$$|Z_N(s)|^p \leq C \left( |Z_0|^p + K + K \int_0^s |Z_N(\tau)|^p d\tau + K \int_0^t \overline{M}_p(\tau; \boldsymbol{\mu}^N) d\tau + \sup_{0 \leq \tau \leq T} |B(\tau)|^p \right).$$

By Grönwall's inequality, this implies

$$|Z_N(s)|^p \leq C \left( |Z_0|^p + K + K \int_0^t \overline{M}_p(\tau; \boldsymbol{\mu}^N) d\tau + \sup_{0 \leq \tau \leq T} |B(\tau)|^p \right) e^{KT}.$$

Now take the expectation to achieve

$$M_p(\mu_s^N) \leq C \left( 1 + M_p(\bar{\mu}) + K + K \int_0^t \overline{M}_p(\tau; \boldsymbol{\mu}^N) d\tau \right) e^{KT}, \quad (3.9)$$

where we also used (2.4) (and Hölder's inequality if  $p = 1$ ). Since (3.9) holds for all  $s \in [0, t]$ , we can take the supremum and get

$$\overline{M}_p(t, \boldsymbol{\mu}^N) \leq C \left( 1 + M_p(\bar{\mu}) + K + K \int_0^t \overline{M}_p(\tau; \boldsymbol{\mu}^N) d\tau \right) e^{KT},$$

which in turn implies, by Grönwall's inequality,

$$\overline{M}_p(t, \boldsymbol{\mu}^N) \leq C(1 + K) e^{KT(1+e^{KT})}. \quad (3.10)$$

Now consider any  $N > C(1 + K) e^{KT(1+e^{KT})}$ . Then

$$\mathbf{v}_N[t, \boldsymbol{\mu}^N](X_N(t), V_N(t)) = \mathbf{v}[t, \boldsymbol{\mu}^N](X_N(t), V_N(t)),$$

thus  $(X_N, V_N)$  solves (1.3), while  $\boldsymbol{\mu}^N = \text{Law}(X_N, V_N)$  solves (1.1). This also proves the first bound in (1.5).

**3.1.3. Uniqueness of the local solution.** Let us first show that if  $\mu$  is a solution of (1.1), then there exists a strong solution  $(X, V)$  of (1.3) such that  $\mu = \text{Law}(X, V)$ . To do this, fix a solution  $\mu$  of (1.3) and define  $F_\mu(t, x, v) = \mathbf{v}[t, \mu](x, v)$ . Then  $\mu$  solves the linear PDE

$$\begin{cases} \partial_t \mu_t = -v \cdot \nabla_x \mu_t + \sigma \Delta_v \mu_t - \text{div}_v(F_\mu(t, z)\mu_t) & (t, x, v) \in \mathbb{R}_0^+ \times \mathbb{R}^{2d}, \\ \mu_0 = \bar{\mu} & (x, v) \in \mathbb{R}^{2d}, \end{cases}$$

hence, by Theorem 2.2 we know that  $\mu = \text{Law}(X, V)$  where  $(X, V)$  is the strong solution of (3.1). The fact that  $(X, V) \in L^p(\Omega; C([0, T]; \mathbb{R}^{2d}))$  follows by (2.6) where  $\Phi$  is the identity. Hence, in particular, by definition of  $F_\mu$ ,  $(X, V)$  solves (1.3) as required.

Now assume that  $\mu^1, \mu^2 \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$  are both solutions of 1.1 and let  $Z_j = (X_j, V_j) \in L^p(\Omega; C([0, T]; \mathbb{R}^{2d}))$  be the respective solutions of (1.3) with  $\text{Law}(X_j, V_j) = \mu^j$ ,  $j = 1, 2$ . Define the auxiliary processes  $\delta X(t) = X_1(t) - X_2(t)$ ,  $\delta V(t) = V_1(t) - V_2(t)$  and  $\delta Z(t) = Z_1(t) - Z_2(t)$ . Consider any  $t \in [0, T]$  and  $s \in [0, t]$ . For a fixed  $\omega \in \Omega$  for which (1.4) holds, we have

$$\delta V(s, \omega) = \int_0^s (\mathbf{v}[\tau, \mu^1](Z_1(\tau, \omega)) - \mathbf{v}[\tau, \mu^2](Z_2(\tau, \omega))) d\tau,$$

i.e.  $\delta V(\cdot, \omega)$  is absolutely continuous. As a consequence also  $|\delta V(\cdot, \omega)|^p$  is absolutely continuous and by the chain rule and taking the expectation it holds

$$\mathbb{E}[|\delta V(s)|^p] = p\mathbb{E}\left[\int_0^s |\delta V(\tau)|^{p-2} (\mathbf{v}[\tau, \mu^1](Z_1(\tau)) - \mathbf{v}[\tau, \mu^2](Z_2(\tau))) \cdot \delta V(\tau) d\tau\right].$$

Now we want to use Fubini's theorem to exchange the expectation with the integral. To do this, we observe that

$$\begin{aligned} & |\delta V(\tau)|^{p-1} |\mathbf{v}[\tau, \mu^1](Z_1(\tau)) - \mathbf{v}[\tau, \mu^2](Z_2(\tau))| \\ & \leq CK(|V_1(\tau)|^{p-1} + |V_2(\tau)|^{p-1}) \left(1 + |X_1(\tau)|^{\frac{\beta}{3}} + |V_1(\tau)|^\beta + |X_2(\tau)|^{\frac{\beta}{3}} + |V_2(\tau)|^\beta\right) \\ & \quad + (\overline{M}_p(\tau, \mu^1))^{\frac{1}{p}} + (\overline{M}_p(\tau, \mu^2))^{\frac{1}{p}} \\ & \leq CK(|V_1(\tau)|^{p-1} + |V_2(\tau)|^{p-1})(|X_1(\tau)|^{\frac{\beta}{3}} + |X_2(\tau)|^{\frac{\beta}{3}}) \\ & \quad + CK(|V_1(\tau)|^{p-1} + |V_2(\tau)|^{p-1} + |V_1(\tau)|^{p-1+\beta} + |V_2(\tau)|^{p-1+\beta}) \\ & \quad + CK(|V_1(\tau)|^{p-1} + |V_2(\tau)|^{p-1}) \left((\overline{M}_p(\tau, \mu^1))^{\frac{1}{p}} + (\overline{M}_p(\tau, \mu^2))^{\frac{1}{p}}\right) \\ & \leq CK \left[1 + (|V_1(\tau)|^p + |V_2(\tau)|^p) + (|X_1(\tau)|^p + |X_2(\tau)|^p) + ((\overline{M}_p(\tau, \mu^1)) + (\overline{M}_p(\tau, \mu^2)))\right], \end{aligned}$$

where in the last inequality we used Young's inequality. Taking the expectation it is not difficult to check that

$$\mathbb{E}[|\delta V(\tau)|^{p-1} |\mathbf{v}[\tau, \mu^1](Z_1(\tau)) - \mathbf{v}[\tau, \mu^2](Z_2(\tau))|] \leq CK(1 + \overline{M}_p(T, \mu^1) + \overline{M}_p(T, \mu^2))$$

where the right-hand side is independent of  $\tau$ . Hence, we can use Fubini's theorem to get

$$\begin{aligned} \mathbb{E} [|\delta V(s)|^p] &= p \int_0^s \mathbb{E} [|\delta V(\tau)|^{p-2} (\mathbf{v}[\tau, \boldsymbol{\mu}^1](Z_1(\tau)) - \mathbf{v}[\tau, \boldsymbol{\mu}^2](Z_2(\tau))) \cdot \delta V(\tau)] d\tau \\ &\leq CD \int_0^s \sup_{0 \leq r \leq \tau} \mathbb{E} [|\delta Z(r)|^p] d\tau, \end{aligned} \quad (3.11)$$

where we used the dissipativity assumption  $(\mathbf{v}_3)$ . We also have, clearly,

$$\mathbb{E} [|\delta X(s)|^p] \leq C \int_0^s \mathbb{E} [|\delta V(\tau)|^p] d\tau \leq C \int_0^s \sup_{0 \leq r \leq \tau} \mathbb{E} [|\delta Z(r)|^p] d\tau. \quad (3.12)$$

Combining (3.11) and (3.12) and taking the supremum over  $[0, t]$  we get

$$\sup_{0 \leq s \leq t} \mathbb{E} [|\delta Z(s)|^p] \leq C(D+1) \int_0^t \sup_{0 \leq r \leq \tau} \mathbb{E} [|\delta Z(r)|^p] d\tau.$$

Since this holds for all  $t \in [0, T]$ , we can use Grönwall's inequality to achieve  $\sup_{0 \leq s \leq t} \mathbb{E} [|\delta Z(s)|^p] = 0$  for all  $t \in [0, T]$ , which in turn implies  $Z_1 = Z_2$  a.s. Hence  $\boldsymbol{\mu}^1 = \boldsymbol{\mu}^2$ . Observe that this also shows that (1.3) admits a unique strong solution.

**3.1.4. Constructon of the global solution.** Now consider any increasing sequence  $T_n \uparrow +\infty$  and denote by  $\boldsymbol{\mu}^n$  the local solution of (1.1) on  $[0, T_n]$ . Let  $n_1 < n_2$  and observe that  $\boldsymbol{\mu}^{n_2}$  still solves (1.1) on  $[0, T_1]$ . Hence, by uniqueness of the local solution,  $\boldsymbol{\mu}^{n_1} = \boldsymbol{\mu}^{n_2}$  on  $[0, T_1]$ . For such a reason, if we denote  $n(t) = \min\{n \geq 0 : t \geq T_n\}$ , we can define the continuous curve of probability measures  $\boldsymbol{\mu} \in C(\mathbb{R}_0^+; \mathcal{W}_p(\mathbb{R}^{2d}))$  by setting

$$\mu_t = \mu_t^{n(t)}, \quad \forall t \in \mathbb{R}_0^+.$$

It is clear that  $\boldsymbol{\mu}$  is the unique global solution we are searching for. Analogously, if we set  $(X_n, V_n)$  to be the respective solutions of (1.3), a similar argument shows that the stochastic process  $(X, V) \in L^p(\Omega; C(\mathbb{R}_0^+; \mathbb{R}^{2d}))$  defined as

$$(X(t), V(t)) = (X_{n(t)}(t), V_{n(t)}(t))$$

still solves (1.3) and  $\boldsymbol{\mu} = \text{Law}(X, V)$ .

**3.1.5. Hölder continuity of the solution.** For any  $t \geq 0$ , let  $\Sigma_t$  be the  $\sigma$ -algebra generated by  $(X_0, V_0)$  and  $\{B(s); s \leq t\}$ . Then it is not difficult to check that  $(X, V)$  is a stochastic process adapted to the filtration  $\{\Sigma_t\}_{t \geq 0}$ . Now let  $\boldsymbol{\mu}$  be the solution fo (1.1) and  $Z = (X, V)$  the solution of (1.3) with  $\boldsymbol{\mu} = \text{Law}(X, V) = \text{Law}(Z)$ . Let  $T > 0, t \geq 0$  and consider  $h > 0$  such that  $0 \leq t < t+h \leq T$ , without loss of generality. Clearly,  $(Z(t+h), Z(t))$  constitute a coupling of  $\mu_{t+h}$  and  $\mu_t$ , hence

$$\mathcal{W}_p(\mu_{t+h}, \mu_t) \leq (\mathbb{E} [ |Z(t+h) - Z(t)|^p ])^{\frac{1}{p}}$$

Set  $\tilde{V}(h) = V(t+h) - V(t)$ ,  $\tilde{X}(h) = X(t+h) - X(t)$  and  $\tilde{Z}(h) = Z(t+h) - Z(t)$ . Observe in particular that

$$\tilde{X}(h) = \int_t^{t+h} V(s) ds$$

and then

$$\mathbb{E} \left[ |\tilde{X}(h)|^p \right] \leq h^{p-1} \mathbb{E} \left[ \int_t^{t+h} |V(s)|^p ds \right] \leq h^p \overline{M}_p(T; \boldsymbol{\mu}) \leq C(1+K)e^{KT(1+e^{KT})} h^p, \quad (3.13)$$

where we used (3.10). As a consequence, we get

$$\begin{aligned} \mathcal{W}_p(\mu_{t+h}, \mu_t) &\leq \left( \mathbb{E} \left[ |\tilde{Z}(h)|^p \right] \right)^{\frac{1}{p}} \leq C \left( \left( \mathbb{E} \left[ |\tilde{X}(h)|^p \right] \right)^{\frac{1}{p}} + \left( \mathbb{E} \left[ |\tilde{V}(h)|^p \right] \right)^{\frac{1}{p}} \right) \\ &\leq C \left( (1+K)^{\frac{1}{p}} e^{\frac{KT(1+e^{KT})}{p}} h + \left( \mathbb{E} \left[ |\tilde{V}(h)|^p \right] \right)^{\frac{1}{p}} \right), \end{aligned} \quad (3.14)$$

hence we only need to estimate  $\mathbb{E} \left[ |\tilde{V}(h)|^p \right]$ . We have

$$\tilde{V}(h) = \int_0^h \tilde{\mathbf{v}}[s, \boldsymbol{\mu}](\tilde{X}(s) + X(t), \tilde{V}(s) + V(t)) ds + \sqrt{2\sigma} \tilde{B}(h).$$

Where  $\tilde{\mathbf{v}}[h, \boldsymbol{\mu}](z) = \mathbf{v}[t+h, \boldsymbol{\mu}](z)$  and  $\tilde{B}(h) = B(t+h) - B(t)$ . Fix  $\varepsilon > 0$  and set  $H_{p,\varepsilon}(x) = (|x|^2 + \varepsilon)^{\frac{p}{2}}$ . Once we recall that  $\tilde{B}$  is still a Brownian motion and  $H_{p,\varepsilon} \in C^2(\mathbb{R}^d)$ , we can use Itô's formula to write

$$\begin{aligned} H_{p,\varepsilon}(\tilde{V}(h)) &= \varepsilon^{\frac{p}{2}} \\ &+ \int_0^h \left( \nabla H_{p,\varepsilon}(\tilde{V}(\tau)) \cdot \tilde{\mathbf{v}}[\tau, \boldsymbol{\mu}](\tilde{X}(\tau) + X(t), \tilde{V}(\tau) + V(t)) + \sigma \Delta H_{p,\varepsilon}(\tilde{V}(\tau)) \right) d\tau \\ &+ \sqrt{2\sigma} \int_0^h \nabla H_{p,\varepsilon}(\tilde{V}(\tau)) d\tilde{B}(\tau). \end{aligned}$$

Next, for any  $h \geq 0$ , let  $\tilde{\Sigma}_h^t$  be the  $\sigma$ -algebra generated by  $\{(X(s), V(s)), s \in [0, t]\}$  and  $\{\tilde{B}(\tau); \tau \in [0, h]\}$ . Then  $(\tilde{X}, \tilde{V})$  is adapted to the filtration  $\{\tilde{\Sigma}_h^t\}_{h \geq 0}$  and then, for any  $N \in \mathbb{N}$ , the random variable

$$\mathcal{T}_N(h) := \min\{\inf\{\tau > 0 : |\tilde{V}(\tau)| \geq N\}, h\}$$

is a  $\{\tilde{\Sigma}_h^t\}_{h \geq 0}$ -stopping time. Observe further that

$$\begin{aligned} H_{p,\varepsilon}(\tilde{V}(\mathcal{T}_N(h))) &= \varepsilon^{\frac{p}{2}} \\ &+ \int_0^{\mathcal{T}_N(h)} \left( \nabla H_{p,\varepsilon}(\tilde{V}(\tau)) \cdot \tilde{\mathbf{v}}[\tau, \boldsymbol{\mu}](\tilde{X}(\tau) + X(t), \tilde{V}(\tau) + V(t)) + \sigma \Delta H_{p,\varepsilon}(\tilde{V}(\tau)) \right) d\tau \\ &+ \sqrt{2\sigma} \int_0^{\mathcal{T}_N(h)} \nabla H_{p,\varepsilon}(\tilde{V}(\min\{\tau, \mathcal{T}_N(h)\})) d\tilde{B}(\tau). \end{aligned} \quad (3.15)$$

Before taking the expectation on both sides, observe that

$$\nabla H_{p,\varepsilon}(x) = p(|x|^2 + \varepsilon)^{\frac{p-2}{2}} x \quad (3.16)$$

and then

$$|\nabla H_{p,\varepsilon}(\tilde{V}(\min\{\tau, \mathcal{T}_N(h)\}))| \leq pN(N^2 + \varepsilon)^{\frac{p-2}{2}}.$$

Thus the process  $\int_0^h \nabla H_{p,\varepsilon}(\tilde{V}(\min\{\tau, \mathcal{T}_N(h)\})) d\tilde{B}(\tau)$  is a martingale [39, Corollary 3.2.6]. Hence, since  $\mathcal{T}_N(h)$  is a bounded stopping time, we can use the optional stopping theorem [43, Theorem II.3.2] to get

$$\mathbb{E} \left[ \int_0^{\mathcal{T}_N(h)} \nabla H_{p,\varepsilon}(\tilde{V}(\min\{\tau, \mathcal{T}_N(h)\})) d\tilde{B}(\tau) \right] = 0.$$

Taking the expectation in (3.15) and then the absolute value we achieve

$$\begin{aligned} & \mathbb{E} \left[ H_{p,\varepsilon}(\tilde{V}(\mathcal{T}_N(h))) \right] \leq \varepsilon^{\frac{p}{2}} \\ & + \mathbb{E} \left[ \int_0^h \left( \left| \nabla H_{p,\varepsilon}(\tilde{V}(\tau)) \cdot \tilde{\mathbf{v}}[\tau, \boldsymbol{\mu}](\tilde{X}(\tau) + X(t), \tilde{V}(\tau) + V(t)) \right| + \sigma \left| \Delta H_{p,\varepsilon}(\tilde{V}(\tau)) \right| \right) dt \right] \end{aligned} \quad (3.17)$$

Now we use  $(\mathbf{v}_1)$  and (3.16) to get

$$\begin{aligned} & \left| \nabla H_{p,\varepsilon}(\tilde{V}(\tau)) \cdot \tilde{\mathbf{v}}[\tau, \boldsymbol{\mu}](\tilde{X}(\tau) + X(t), \tilde{V}(\tau) + V(t)) \right| \\ & \leq pK(|\tilde{V}(\tau)|^2 + \varepsilon)^{\frac{p-2}{2}} |\tilde{V}(\tau)| (1 + |\tilde{X}(\tau) + X(t)|^{\frac{\beta}{3}} + |\tilde{V}(\tau) + V(t)|^\beta + (\overline{M}_p(t + \tau; \boldsymbol{\mu}))^{\frac{1}{p}}) \\ & \leq CK(|\tilde{V}(\tau)|^2 + \varepsilon)^{\frac{p-1}{2}} (1 + |\tilde{X}(\tau)|^{\frac{\beta}{3}} + |X(t)|^{\frac{\beta}{3}} + |\tilde{V}(\tau)|^\beta + |V(t)|^\beta + (\overline{M}_p(t + \tau; \boldsymbol{\mu}))^{\frac{1}{p}}) \\ & \leq CKH_{p,\varepsilon}(\tilde{V}(\tau)) + CK(|\tilde{V}(\tau)|^2 + \varepsilon)^{\frac{p-1}{2}} (1 + |\tilde{X}(\tau)|^{\frac{\beta}{3}} + |X(t)|^{\frac{\beta}{3}} + |V(t)|^\beta + (\overline{M}_p(t + \tau; \boldsymbol{\mu}))^{\frac{1}{p}}). \end{aligned}$$

To handle the second summand, we use Young's inequality with exponent  $p$ , and then we achieve

$$\begin{aligned} & \left| \nabla H_{p,\varepsilon}(\tilde{V}(\tau)) \cdot \tilde{\mathbf{v}}[\tau, \boldsymbol{\mu}](\tilde{X}(\tau) + X(t), \tilde{V}(\tau) + V(t)) \right| \\ & \leq CKH_{p,\varepsilon}(\tilde{V}(\tau)) + CK^p (1 + |\tilde{X}(\tau)|^{\frac{\beta p}{3}} + |X(t)|^{\frac{\beta p}{3}} + |V(t)|^{\beta p} + \overline{M}_p(t + \tau; \boldsymbol{\mu})) \\ & \leq CKH_{p,\varepsilon}(\tilde{V}(\tau)) + CK(1 + |\tilde{X}(\tau)|^p + |Z(t)|^p + \overline{M}_p(T; \boldsymbol{\mu})). \end{aligned}$$

Taking the expectation this leads to

$$\begin{aligned} & \mathbb{E} \left[ \left| \nabla H_{p,\varepsilon}(\tilde{V}(\tau)) \cdot \tilde{\mathbf{v}}[\tau, \boldsymbol{\mu}](\tilde{X}(\tau) + X(t), \tilde{V}(\tau) + V(t)) \right| \right] \\ & \leq CK \mathbb{E} \left[ H_{p,\varepsilon}(\tilde{V}(\tau)) \right] + CK(1 + (1 + K)e^{KT(1+e^{KT})}), \end{aligned} \quad (3.18)$$

where we also used (3.10) and (3.13). Plugging (3.18) into (3.17) we have

$$\begin{aligned} \mathbb{E} \left[ H_{p,\varepsilon}(\tilde{V}(\mathcal{T}_N(h))) \right] &\leq \varepsilon^{\frac{p}{2}} + CK \int_0^h \mathbb{E} \left[ H_{p,\varepsilon}(\tilde{V}(\tau)) \right] d\tau \\ &\quad + CK(1 + (1 + K)e^{KT(1+e^{KT})})h + \sigma \int_0^h \mathbb{E} \left[ \left| \Delta H_{p,\varepsilon}(\tilde{V}(\tau)) \right| d\tau \right] \end{aligned} \quad (3.19)$$

Next, observe that since  $V$  is a global solution of (1.3), it cannot blow up in finite time, hence it must hold  $\lim_{N \rightarrow \infty} \mathcal{T}_N(h) = h$ . Furthermore, for almost all  $\omega \in \Omega$ ,  $V$  (and then  $\tilde{V}$ ) is continuous. Hence, by Fatou's Lemma, taking the limit inferior in (3.19) we get

$$\begin{aligned} \mathbb{E} \left[ H_{p,\varepsilon}(\tilde{V}(h)) \right] &\leq \varepsilon^{\frac{p}{2}} + CK \int_0^h \mathbb{E} \left[ H_{p,\varepsilon}(\tilde{V}(\tau)) \right] d\tau \\ &\quad + CK(1 + (1 + K)e^{KT(1+e^{KT})})h + \sigma \int_0^h \mathbb{E} \left[ \left| \Delta H_{p,\varepsilon}(\tilde{V}(\tau)) \right| d\tau \right]. \end{aligned} \quad (3.20)$$

Furthermore, notice that

$$\Delta H_{p,\varepsilon}(x) = p(|x|^2 + \varepsilon)^{\frac{p-4}{2}} [(p+d-2)|x|^2 + d\varepsilon],$$

hence

$$\left| \Delta H_{p,\varepsilon}(\tilde{V}(h)) \right| \leq C(|\tilde{V}(h)|^2 + \varepsilon)^{\frac{p-2}{2}}$$

Now we have to distinguish among two cases. If  $p > 2$  we get, by Young's inequality with exponent  $\frac{p}{p-2}$ ,

$$\left| \Delta H_{p,\varepsilon}(\tilde{V}(h)) \right| \leq C(1 + H_{p,\varepsilon}(|\tilde{V}(h)|)). \quad (3.21)$$

Hence, using (3.21) into (3.20), we achieve

$$\begin{aligned} \mathbb{E} \left[ H_{p,\varepsilon}(\tilde{V}(h)) \right] &\leq \varepsilon^{\frac{p}{2}} + C(1 + K) \int_0^h \mathbb{E} \left[ H_{p,\varepsilon}(\tilde{V}(\tau)) \right] d\tau \\ &\quad + CK(1 + (1 + K)e^{KT(1+e^{KT})})h. \end{aligned}$$

Since  $h \in [0, T - t]$  is arbitrary, we can use Grönwall's inequality to get

$$\mathbb{E} \left[ H_{p,\varepsilon}(\tilde{V}(h)) \right] \leq \left( \varepsilon^{\frac{p}{2}} + CK(1 + (1 + K)e^{KT(1+e^{KT})})h \right) e^{C(1+K)h}.$$

It remains to take the limit as  $\varepsilon \rightarrow 0$  to finally achieve, by the dominated convergence theorem,

$$\mathbb{E} \left[ |\tilde{V}(h)|^p \right] \leq CK(1 + (1 + K)e^{KT(1+e^{KT})})e^{C(1+K)h}.$$

Using the latter inequality into (3.14), we have the local  $1/p$ -Hölder continuity of  $\mu$  and the third bound in (1.5).

If  $p \leq 2$  instead we use

$$\left| \Delta H_{p,\varepsilon}(\tilde{V}(h)) \right| \leq C\varepsilon^{\frac{p-2}{2}}$$



and then

$$\begin{aligned} \mathbb{E} \left[ H_{p,\varepsilon}(\tilde{V}(h)) \right] &\leq \varepsilon^{\frac{p}{2}} + CK \int_0^h \mathbb{E} \left[ H_{p,\varepsilon}(\tilde{V}(\tau)) \right] d\tau \\ &\quad + CK(1 + (1 + K)e^{KT(1+e^{KT})})h + C\varepsilon^{\frac{p-2}{2}}h. \end{aligned}$$

Again, by Grönwall's inequality we have

$$\mathbb{E}[|\tilde{V}(h)|^p] \leq \mathbb{E} \left[ H_{p,\varepsilon}(\tilde{V}(h)) \right] \leq \left( \varepsilon^{\frac{p}{2}} + CK(1 + (1 + K)e^{KT(1+e^{KT})})h + C\varepsilon^{\frac{p-2}{2}}h \right) e^{CK}.$$

The latter inequality holds for any  $\varepsilon > 0$  and any  $h \in [0, T - t]$ , hence, we can take  $\varepsilon = h$  so that

$$\mathbb{E}[|\tilde{V}(h)|^p] \leq \left( 1 + CK(1 + (1 + K)e^{KT(1+e^{KT})}) \right) h^{\frac{p}{2}} e^{CK}.$$

Again, we get the local 1/2-Hölder continuity of  $\mu$  and the third bound in (1.5) by (3.14).

**3.1.6. Higher moment estimate.** It remains to prove the second bound in (1.5). Again, consider the Young function  $\Phi$  such that  $\bar{\mu} \in \mathcal{W}_{\Phi_p}(\mathbb{R}^{2d})$ . Observe that, by (v<sub>1</sub>) and (3.10), we have, for all  $t \in [0, T]$  and  $z \in \mathbb{R}^{2d}$

$$\mathbf{v}[t, \mu](z) \leq C(1 + K)e^{KT(1+e^{KT})}(1 + |x|^{\frac{p}{3}} + |v|^\beta).$$

Hence we can use (2.5) to get

$$\overline{M}_{\tilde{\Phi}_p(\cdot; K)}(T, \mu) \leq C(1 + (1 + K)e^{KT(1+e^{KT})}),$$

where

$$\tilde{\Phi}(r; K) = \Phi \left( Ce^{-C(1+K)e^{KT(1+e^{KT})}} r \right).$$

This ends the proof. □

*Remark 3.1.* Notice that, to prove the existence of local solutions, Assumption (v<sub>3</sub>) is not needed. Furthermore, for uniqueness, one only needs that Assumption (v<sub>3</sub>) holds on solutions of (1.3).

**3.2. Proof of Corollary 1.4.** From now on, we fix the time horizon  $T > 0$ . Let  $Z_j = (X_j, V_j)$  and  $Z = (X, V)$  be the solutions of (1.3) associated with  $\mathbf{v}_j$  and  $\mathbf{v}$  respectively. Then, clearly,  $\mu_t^j = \text{Law}(X_j(t), V_j(t))$  and  $\mu_t = \text{Law}(X(t), V(t))$ , hence

$$\sup_{0 \leq t \leq T} (\mathcal{W}_p(\mu_t^j, \mu_t))^p \leq C \sup_{0 \leq t \leq T} (\mathbb{E}[|Z_j(t) - Z(t)|^p]).$$

Let  $0 \leq t \leq T$ . Once we notice that  $\delta V_j(t) := V_j(t) - V(t)$  is absolutely continuous, we can use the chain rule to get for any  $s \in [0, t]$

$$\mathbb{E}[|\delta V_j(s)|^p] = p \mathbb{E} \left[ \int_0^s |\delta V_j(\tau)|^{p-2} (\mathbf{v}_j[\tau, \mu^j](Z_j(\tau)) - \mathbf{v}[\tau, \mu](Z(\tau))) \cdot \delta V_j(\tau) d\tau \right].$$

Now notice that, by Young's inequality and (v<sub>1</sub>) we get

$$\begin{aligned} |\delta V_j(\tau)|^{p-1} |\mathbf{v}_j[\tau, \boldsymbol{\mu}^j](Z_j(\tau)) - \mathbf{v}[\tau, \boldsymbol{\mu}](Z(\tau))| \\ \leq C(K+1) (1 + |Z_j(\tau)|^p + |Z(\tau)|^p + \overline{M}_p(T; \boldsymbol{\mu}_j) + \overline{M}_p(T; \boldsymbol{\mu})), \end{aligned}$$

and then

$$\begin{aligned} \int_0^s \mathbb{E} [|\delta V_j(\tau)|^{p-1} |\mathbf{v}_j[\tau, \boldsymbol{\mu}^j](Z_j(\tau)) - \mathbf{v}[\tau, \boldsymbol{\mu}](Z(\tau))|] d\tau \\ \leq C(K+1)T (1 + \overline{M}_p(T; \boldsymbol{\mu}_j) + \overline{M}_p(T, \boldsymbol{\mu})) \\ \leq C(K+1)T (1 + H_T(K)), \end{aligned}$$

where we also used (1.5). Hence we can employ Fubini's theorem to get

$$\begin{aligned} \mathbb{E} [|\delta V_j(s)|^p] &= p \int_0^s \mathbb{E} [|\delta V_j(\tau)|^{p-2} (\mathbf{v}_j[\tau, \boldsymbol{\mu}^j](Z_j(\tau)) - \mathbf{v}[\tau, \boldsymbol{\mu}](Z(\tau))) \cdot \delta V_j(\tau)] d\tau \\ &= p \int_0^s \mathbb{E} [|\delta V_j(\tau)|^{p-2} (\mathbf{v}_j[\tau, \boldsymbol{\mu}^j](Z_j(\tau)) - \mathbf{v}_j[\tau, \boldsymbol{\mu}](Z(\tau))) \cdot \delta V_j(\tau)] d\tau \\ &\quad + p \int_0^s \mathbb{E} [|\delta V_j(\tau)|^{p-2} (\mathbf{v}_j[\tau, \boldsymbol{\mu}](Z(\tau)) - \mathbf{v}[\tau, \boldsymbol{\mu}](Z(\tau))) \cdot \delta V_j(\tau)] d\tau \\ &= I_1 + I_2. \end{aligned} \tag{3.22}$$

For the first term, we use (v<sub>3</sub>) to get

$$I_1 \leq D \int_0^s \sup_{0 \leq r \leq \tau} \mathbb{E} [|\delta Z_j(r)|^p] d\tau, \tag{3.23}$$

where  $\delta Z_j(r) := Z_j(r) - Z(r)$ . Next, we move to  $I_2$ . To handle it we use Young's inequality as follows

$$\begin{aligned} I_2 &\leq p \int_0^s \mathbb{E} [|\delta V_j(\tau)|^{p-1} |\mathbf{v}_j[\tau, \boldsymbol{\mu}](Z(\tau)) - \mathbf{v}[\tau, \boldsymbol{\mu}](Z(\tau))|] d\tau \\ &\leq C \left( \int_0^s \sup_{0 \leq r \leq \tau} \mathbb{E} [|\delta Z_j(r)|^p] d\tau + \int_0^s \mathbb{E} [|\mathbf{v}_j[\tau, \boldsymbol{\mu}](Z(\tau)) - \mathbf{v}[\tau, \boldsymbol{\mu}](Z(\tau))|^p] d\tau \right) \end{aligned} \tag{3.24}$$

We denote

$$\mathcal{R}_j(s) = \int_0^s \mathbb{E} [|\mathbf{v}_j[\tau, \boldsymbol{\mu}](Z(\tau)) - \mathbf{v}[\tau, \boldsymbol{\mu}](Z(\tau))|^p] d\tau$$

and we notice that it is a nondecreasing function. Hence we get, combining (3.22), (3.23) and (3.24)

$$\mathbb{E} [|\delta V_j(s)|^p] \leq C(1+D) \int_0^s \sup_{0 \leq r \leq \tau} \mathbb{E} [|\delta Z_j(r)|^p] d\tau + C\mathcal{R}_j(s). \tag{3.25}$$

Next, we observe that

$$\mathbb{E} [|\delta X_j(s)|^p] \leq T^{p-1} \int_0^s \mathbb{E} [|\delta V_j(\tau)|^p] d\tau \leq T^{p-1} \int_0^s \sup_{0 \leq r \leq \tau} \mathbb{E} [|\delta Z_j(r)|^p] d\tau, \tag{3.26}$$

where  $\delta X_j(t) := X_j(t) - X(t)$ . Combining (3.25) and (3.26) and taking the supremum on  $[0, t]$  we have

$$\sup_{0 \leq s \leq t} \mathbb{E} [|\delta Z_j(s)|^p] \leq C(1+D) \int_0^t \sup_{0 \leq r \leq \tau} \mathbb{E} [|\delta Z_j(r)|^p] d\tau + C\mathcal{R}_j(t),$$

that, by Grönwall's inequality, implies

$$\sup_{0 \leq s \leq T} \mathbb{E} [|\delta Z_j(s)|^p] \leq e^{C(1+D)} \mathcal{R}_j(T)$$

and then

$$\sup_{0 \leq t \leq T} (\mathcal{W}_p(\mu_t^j, \mu_t))^p \leq e^{C(1+D)} \mathcal{R}_j(T).$$

It remains to show that the right-hand side converges to 0. To do this, it is sufficient to observe that, by (v<sub>1</sub>) and Young's inequality

$$|\mathbf{v}_j[\tau, \boldsymbol{\mu}](Z(\tau)) - \mathbf{v}_j[\tau, \boldsymbol{\mu}](Z(\tau))| \leq CK (1 + |Z(\tau)|^p + \overline{M}_p(T, \boldsymbol{\mu})),$$

where

$$\int_0^T \mathbb{E} [CK (1 + |Z(\tau)|^p + \overline{M}_p(T, \boldsymbol{\mu}))] d\tau \leq CK(1 + \overline{M}_p(T, \boldsymbol{\mu})) < \infty.$$

Hence, by the dominated convergence theorem and the fact that  $\boldsymbol{\mu}$  satisfies (1.5),  $\lim_{j \rightarrow \infty} \mathcal{R}_j(T) = 0$ . This concludes the proof.  $\square$

#### 4. THE PDE-ODE SYSTEM AND A RELATED OPTIMAL CONTROL PROBLEM

In this section we apply the previous results to a mean field sparse optimal control problem. Precisely, we consider the following PDE-ODE system [1, 2]

$$\begin{cases} \partial_t \mu_t = -v \cdot \nabla_x \mu_t + \sigma \Delta_v \mu_t - \operatorname{div}_v((\mathbf{v}[t, \boldsymbol{\mu}](z) \\ \quad + \mathbf{w}[t, \mathbf{H}](z))\mu_t) & (t, x, v) \in (0, T] \times \mathbb{R}^{2d} \\ \dot{\mathbf{Y}}(t) = \mathbf{W}(t) = F[t, \boldsymbol{\mu}](\mathbf{Y}) + \mathbf{u}(t, \boldsymbol{\mu}) & t \in (0, T] \\ \mu_0 = \bar{\mu}, \quad \mathbf{Y}(0) = \bar{\mathbf{Y}}, \end{cases} \quad (4.1)$$

where  $\boldsymbol{\mu} \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$ ,  $\mathbf{H} = (\mathbf{Y}, \mathbf{W}) : [0, T] \rightarrow \mathbb{R}^{2m}$ ,  $\mathbf{v} : [0, T] \times C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d})) \times \mathbb{R}^{2d} \rightarrow \mathbb{R}^d$ ,  $\mathbf{w} : [0, T] \times C([0, T]; \mathbb{R}^{2m}) \times \mathbb{R}^{2d} \rightarrow \mathbb{R}^d$ ,  $F : [0, T] \times C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d})) \times C([0, T]; \mathbb{R}^m) \rightarrow \mathbb{R}^m$ ,  $\mathbf{u} : [0, T] \times C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d})) \rightarrow \mathbb{R}^m$ ,  $\bar{\mu} \in \mathcal{W}_p(\mathbb{R}^{2d})$ , and  $\bar{\mathbf{H}} = (\bar{\mathbf{Y}}, \bar{\mathbf{W}}) \in \mathbb{R}^{2m}$ .

**Definition 4.1.** We say that  $(\boldsymbol{\mu}, \mathbf{Y}) \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}) \times \mathbb{R}^{2m})$  is a solution of (4.1) if and only if  $\boldsymbol{\mu}$  is solution of

$$\begin{cases} \partial_t \mu_t = -v \cdot \nabla_x \mu_t + \sigma \Delta_v \mu_t - \operatorname{div}_v((\mathbf{v}[t, \boldsymbol{\mu}](z) + \mathbf{w}[t, \mathbf{H}](z))\mu_t) & (t, x, v) \in (0, T] \times \mathbb{R}^{2d} \\ \mu_0 = \bar{\mu}, \end{cases}$$

in the sense of Definition 1.1 and  $\mathbf{Y}$  satisfies

$$\mathbf{Y}(t) = \bar{\mathbf{Y}} + \int_0^t (F[s, \boldsymbol{\mu}](\mathbf{Y}) + \mathbf{u}(s, \boldsymbol{\mu})) ds \quad t \in (0, T]$$

From now on, we assume that  $\mathbf{v}$  satisfies Assumptions (v). Furthermore we consider the following assumptions of  $\mathfrak{w}$  and  $F$ :

### Assumptions on $\mathfrak{w}$ : (w)

- (w<sub>0</sub>)  $\mathfrak{w} : [0, T] \times C([0, T]; \mathbb{R}^{2m}) \times \mathbb{R}^{2d} \rightarrow \mathbb{R}^d$  is a Carathéodory map, i.e., it is measurable in the variable  $t \in [0, T]$  and continuous in  $(\mathbf{H}, z) \in C([0, T]; \mathbb{R}^{2m}) \times \mathbb{R}^{2d}$ .
- (w<sub>1</sub>) There exists a constant  $K_{\mathfrak{w}} > 0$  such that for any  $t \in [0, T]$ ,  $z \in \mathbb{R}^{2d}$  and  $\mathbf{H} \in C([0, T]; \mathbb{R}^{2m})$  it holds

$$|\mathfrak{w}[t, \mathbf{H}](z)| \leq K_{\mathfrak{w}}(1 + \sup_{0 \leq s \leq t} |\mathbf{H}(s)| + |z|^{\frac{\beta}{3}} + |v|^{\beta}),$$

where  $\beta$  is the exponent in (v<sub>1</sub>).

- (w<sub>2</sub>) There exists a constant  $L_{\mathfrak{w}}$  such that for any  $t \in [0, T]$ ,  $\mathbf{H}^j \in C([0, T]; \mathbb{R}^{2m})$  and  $z_j \in \mathbb{R}^{2d}$ ,  $j = 1, 2$ , it holds

$$|\mathfrak{w}[t, \mathbf{H}^1](z_1) - \mathfrak{w}[t, \mathbf{H}^2](z_2)| \leq L_{\mathfrak{w}} \left( \sup_{0 \leq s \leq t} |\mathbf{H}^1 - \mathbf{H}^2| + |z_1 - z_2| \right).$$

### Assumptions on $F$ : (F)

- (F<sub>0</sub>)  $F : [0, T] \times C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d})) \times C([0, T]; \mathbb{R}^m) \rightarrow \mathbb{R}^m$  is a Carathéodory map, i.e. is measurable in  $t \in [0, T]$  and continuous in  $(\boldsymbol{\mu}, \mathbf{Y}) \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d})) \times C([0, T]; \mathbb{R}^m)$ .

- (F<sub>1</sub>) There exists a constant  $K_F > 0$  such that

$$|F(t, \boldsymbol{\mu})(\mathbf{Y})| \leq K_F \left( 1 + \sup_{0 \leq s \leq t} |\mathbf{Y}(s)| + (\bar{M}_p(T; \boldsymbol{\mu}))^{\frac{1}{p}} \right),$$

for all  $\boldsymbol{\mu} \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$ ,  $\mathbf{Y} \in C([0, T]; \mathbb{R}^m)$  and  $t \in [0, T]$ .

- (F<sub>2</sub>) There exists a constant  $L_F > 0$  such that

$$|F[t, \boldsymbol{\mu}](\mathbf{Y}^1) - F[t, \boldsymbol{\nu}](\mathbf{Y}^2)| \leq L_F \left( \sup_{0 \leq s \leq t} |\mathbf{Y}^1(s) - \mathbf{Y}^2(s)| + \sup_{0 \leq s \leq t} \mathcal{W}_p(\mu_s, \nu_s) \right),$$

for all  $\boldsymbol{\mu}, \boldsymbol{\nu} \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$ ,  $\mathbf{Y}^j \in C([0, T]; \mathbb{R}^m)$ ,  $j = 1, 2$  and  $t \in [0, T]$ .

Lastly, concerning the controls  $\mathbf{u} = (u_j)_{j=0, \dots, m}$ , we fix two constants  $M_{\mathbf{u}}$  and  $L_{\mathbf{u}}$  and we denote by  $\mathcal{A}$  the set of admissible controls, characterized as follows:

**Assumptions on  $u \in \mathcal{A}$ : ( $\mathcal{A}$ )**

( $\mathcal{A}_0$ )  $\mathbf{u} : [0, T] \times C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d})) \rightarrow \mathbb{R}^m$  is a Carathéodory map, i.e., it is measurable in  $t \in [0, T]$  and continuous in  $\boldsymbol{\mu} \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$ .

( $\mathcal{A}_1$ ) For all  $t \in [0, T]$  it holds

$$|\mathbf{u}(t, \boldsymbol{\delta}_0)| \leq M_u,$$

where  $\boldsymbol{\delta}_0 \in \mathcal{W}_p(C([0, T]; \mathbb{R}^{2d}))$  is the Dirac delta measure concentrated in the constant function  $0 \in C([0, T]; \mathbb{R}^{2d})$ .

( $\mathcal{A}_2$ ) For all  $\boldsymbol{\mu}, \boldsymbol{\nu} \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$ ,  $j = 1, \dots, m$  and  $t \in [0, T]$ , it holds

$$|u_j(t, \boldsymbol{\mu}) - u_j(t, \boldsymbol{\nu})| \leq \frac{L_u}{m} \sup_{0 \leq s \leq t} \mathcal{W}_p(\mu_s, \nu_s).$$

By Assumptions ( $\mathcal{A}_1$ ) and ( $\mathcal{A}_2$ ), for all  $u \in \mathcal{A}$  it holds

$$|\mathbf{u}(t, \boldsymbol{\mu})| \leq |\mathbf{u}(t, \boldsymbol{\mu}) - \mathbf{u}(t, \boldsymbol{\delta}_0)| + |\mathbf{u}(t, \boldsymbol{\delta}_0)| \leq L_u (\overline{M}_p(t; \boldsymbol{\mu}))^{\frac{1}{p}} + M_0,$$

where we used the fact that  $(\overline{M}_p(t; \boldsymbol{\mu}))^{\frac{1}{p}} = \sup_{0 \leq s \leq t} \mathcal{W}_p(\mu_s, \delta_0)$ . Hence, if we set

$$K_u = \max\{L_u, M_u\},$$

then we have that for all  $\boldsymbol{\mu} \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$  and all  $t \in [0, T]$  it holds

$$|\mathbf{u}(t, \boldsymbol{\mu})| \leq K_u (1 + (\overline{M}_p(t; \boldsymbol{\mu}))^{\frac{1}{p}}). \quad (4.2)$$

**4.1. Well-posedness of the PDE-ODE system (4.1).** Before setting the control problem, let us show that the system (4.1) is well-posed. Our proof relies on some preliminary results that are proved in Appendix B. First of all, for any fixed  $T > 0$ ,  $u \in \mathcal{A}$  and  $\boldsymbol{\mu} \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$  we introduce the system

$$\begin{cases} \dot{\mathbf{Y}}_\mu(t) = \mathbf{W}_\mu(t) = F[t, \boldsymbol{\mu}](\mathbf{Y}_\mu) + \mathbf{u}(t, \boldsymbol{\mu}) & t \in (0, T) \\ \mu_0 = \bar{\mu}, \quad \mathbf{Y}_\mu(0) = \bar{\mathbf{Y}}, \end{cases} \quad (4.3)$$

which admits a unique solution  $\mathbf{Y}_\mu$  thanks to Lemma B.1. Then we define the map  $\mathcal{S} : C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d})) \times \mathcal{A} \times [0, T] \rightarrow \mathbb{R}^{2m}$  as follows:

$$\mathcal{S}[\boldsymbol{\mu}, \mathbf{u}](t) = \mathbf{H}_\mu(t) = (\mathbf{Y}_\mu(t), \mathbf{W}_\mu(t)),$$

where  $\mathbf{Y}_\mu$  is the unique solution of (4.3). For any fixed  $u \in \mathcal{A}$ , we also define the map

$$G_u : [0, T] \times C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d})) \times \mathbb{R}^{2d} \rightarrow \mathbb{R}^d$$

as follows:

$$G_u[t, \boldsymbol{\mu}](z) = \mathfrak{v}[t, \boldsymbol{\mu}](z) + \mathfrak{w}[t, \mathcal{S}[\boldsymbol{\mu}, \mathbf{u}]](z),$$

which satisfies assumptions ( $\mathfrak{v}_0$ ), ( $\mathfrak{v}_1$ ), ( $\mathfrak{v}_2$ ) and ( $\mathfrak{v}_3$ ) thanks to Lemma B.2.

Once this is established, we set

$$K_G := K_v + K_w(1 + K_F)(1 + C_1 + K_F + K_u),$$

where  $C_1$  is defined in Lemma B.1, and we denote by

$$\mathcal{K} := \left\{ \boldsymbol{\mu} \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d})) : \overline{M}_p(T, \boldsymbol{\mu}) + \overline{M}_{\tilde{\Phi}_p(\cdot; K_G)}(T; \boldsymbol{\mu}) + \sup_{\substack{0 \leq s, t \leq T \\ t \neq s}} \frac{\mathcal{W}_p(\mu_t, \mu_s)}{|t - s|^{\gamma_p}} \leq \mathcal{C}(K_G) \right\},$$

where  $\tilde{\Phi}$  and  $\mathcal{C}$  are the functions defined in Theorem 1.3 and  $\gamma_p = \frac{1}{\max\{2, p\}}$ . We also set

$$C_1(K) := (C_1 + C_3) \left( 1 + (\mathcal{C}(K_G))^{\frac{1}{p}} \right),$$

where  $C_1$  and  $C_3$  are defined in Lemma B.1, and

$$\mathcal{K}_1 := \left\{ \mathbf{Y} \in C([0, T]; \mathbb{R}^m) : \sup_{0 \leq s \leq T} |\mathbf{Y}(s)| + \sup_{\substack{0 \leq t, s \leq T \\ t \neq s}} \frac{|\mathbf{Y}(t) - \mathbf{Y}(s)|}{|t - s|} \leq C_1(K_G) \right\}.$$

By Lemma B.3 we know both  $\mathcal{K}$  and  $\mathcal{K}_1$  are compact subsets and  $\mathcal{K}' = \mathcal{K} \times \mathcal{K}_1$ . Then we are in a position to prove (4.1) is well-posed.

**Theorem 4.2.** *For any fixed  $\mathbf{u} \in \mathcal{A}$  there exists a unique solution  $(\boldsymbol{\mu}, \mathbf{Y}) \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d})) \times C([0, T]; \mathbb{R}^m)$  of (4.1). In particular, if  $B$  is a  $d$ -dimensional Brownian motion and  $(X_0, V_0) \in L^p(\Omega; \mathbb{R}^{2d})$  is independent of  $B$  and  $\bar{\boldsymbol{\mu}} = \text{Law}(X_0, V_0)$ , then  $\boldsymbol{\mu} = \text{Law}(X, V)$ , where  $(X, V)$  is the unique global strong solution in  $L^p(\Omega; C(\mathbb{R}_0^+; \mathbb{R}^{2d}))$  of*

$$\begin{cases} dX(t) = V(t)dt & t \in [0, T] \\ dV(t) = (\mathfrak{v}[t, \boldsymbol{\mu}](X(t), V(t)) + \mathfrak{w}[t, \mathbf{H}](X(t), V(t))) dt + \sqrt{2\sigma} dB(t) & t \in [0, T] \\ \dot{\mathbf{Y}}(t) = \mathbf{W}(t) = F[t, \boldsymbol{\mu}](\mathbf{Y}) + \mathbf{u}(t, \boldsymbol{\mu}) & t \in (0, T] \\ X(0) = X_0, \quad V(0) = V_0, \quad \boldsymbol{\mu} = \text{Law}(X, V) \\ \mathbf{Y}(0) = \bar{\mathbf{Y}} \end{cases} \quad (4.4)$$

Moreover,  $(\boldsymbol{\mu}, \mathbf{Y}) \in \mathcal{K}'$ . Finally, if  $\{\mathbf{u}_j\}_{j \in \mathbb{N}} \subset \mathcal{A}$  is a sequence of admissible controls such that for all  $t \in [0, T]$  and  $\boldsymbol{\nu} \in \mathcal{K}$

$$\lim_{j \rightarrow \infty} \int_0^t \mathbf{u}_j(s, \boldsymbol{\nu}) ds = \int_0^t \mathbf{u}(s, \boldsymbol{\nu}) ds$$

and  $(\boldsymbol{\mu}^j, \mathbf{Y}^j)$  are the respective solutions of (4.1), then,

$$\lim_{j \rightarrow \infty} \sup_{0 \leq t \leq T} (\mathcal{W}_p(\mu_t^j, \mu_t) + |\mathbf{H}^j(t) - \mathbf{H}(t)|) = 0.$$

*Proof.* Fix  $\mathbf{u} \in \mathcal{A}$  and consider the equation

$$\begin{cases} \partial_t \mu_t = -v \cdot \nabla_x \mu_t + \sigma \Delta_v \mu_t - \text{div}_v (G_{\mathbf{u}}[t, \boldsymbol{\mu}](z) \mu_t) & (t, z) \in (0, T] \times \mathbb{R}^{2d} \\ \mu_0 = \bar{\boldsymbol{\mu}} & z \in \mathbb{R}^{2d}. \end{cases} \quad (4.5)$$

Since  $G_{\mathbf{u}}$  satisfies Assumptions (v), then (4.5) admits a unique solution  $\boldsymbol{\mu} \in C^{\gamma_p}([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$  that can be expressed as  $\boldsymbol{\mu} = \text{Law}(X, V)$  as in the statement, by Theorem 1.3. In particular,  $\boldsymbol{\mu}$  belongs to  $\mathcal{K}$  by (1.5). Furthermore, setting  $\mathbf{H} = \mathcal{S}[\boldsymbol{\mu}, \mathbf{u}]$ , we get that  $(\boldsymbol{\mu}, \mathbf{Y})$

clearly satisfies (4.1) and  $\mathbf{Y} \in \mathcal{K}_1$  by Lemma B.1. This shows the existence of a solution for (4.1). To prove uniqueness, let  $(\boldsymbol{\mu}', \mathbf{Y}')$  be another solution. Then, by Definition 4.1 and uniqueness of the solution of (4.3), we know that  $\mathbf{H}' = \mathcal{S}[\boldsymbol{\mu}', \mathbf{u}]$ . However, this means that  $\boldsymbol{\mu}'$  is solution of (4.5), hence by Lemma B.1, then  $\boldsymbol{\mu}' = \boldsymbol{\mu}$  and  $\mathbf{H}' = \mathcal{S}[\boldsymbol{\mu}', \mathbf{u}] = \mathcal{S}[\boldsymbol{\mu}, \mathbf{u}] = \mathbf{H}$ .

Now, let  $\{\mathbf{u}_j\}_{j \in \mathbb{N}} \subset \mathcal{A}$  be a sequence of controls assumed as in the assumption. Then, by Lemma B.4 and Corollary 1.4, we get that, for all  $t \in [0, T]$ ,

$$\lim_{j \rightarrow \infty} \sup_{0 \leq s \leq t} \mathcal{W}_p(\boldsymbol{\mu}_s^j, \boldsymbol{\mu}_s) = 0. \quad (4.6)$$

Furthermore, we observe that by definition

$$\begin{aligned} |\mathbf{H}^j(t) - \mathbf{H}(t)| &\leq |\mathcal{S}[\boldsymbol{\mu}^j, \mathbf{u}^j](t) - \mathcal{S}[\boldsymbol{\mu}, \mathbf{u}^j](t)| + |\mathcal{S}[\boldsymbol{\mu}, \mathbf{u}^j](t) - \mathcal{S}[\boldsymbol{\mu}, \mathbf{u}](t)| \\ &\leq C \sup_{0 \leq s \leq T} \mathcal{W}_p(\boldsymbol{\mu}_s^j, \boldsymbol{\mu}_s) + |\mathcal{S}[\boldsymbol{\mu}, \mathbf{u}^j](t) - \mathcal{S}[\boldsymbol{\mu}, \mathbf{u}](t)|, \end{aligned}$$

where we applied (B.3) and the assumptions on  $F$  and  $\mathcal{A}$ . We get the desired statement by taking the supremum, then the limit on both sides of the previous inequality and using (B.4) and (4.6).  $\square$

**4.2. The Control Problem.** Now, we want to set a control problem on (4.1). Precisely, we consider the cost functional

$$\mathcal{F}[\mathbf{u}] = \int_0^T \mathcal{L}(t, \boldsymbol{\mu}, \mathbf{Y}) dt + \int_0^T \Psi(\mathbf{u}(t, \boldsymbol{\mu})) dt$$

where  $(\boldsymbol{\mu}, \mathbf{Y})$  is the solution of (4.1) with given control  $\mathbf{u} \in \mathcal{A}$ ,  $\mathcal{L}$  is a lagrangian functional accounting for closedness to the decided target and  $\Psi$  is a convex control cost. We consider the following assumptions on the cost functional  $\mathcal{F}$ .

#### Assumptions on $\mathcal{F}$ : ( $\mathcal{F}$ )

( $\mathcal{F}_0$ )  $\mathcal{L} : [0, T] \times C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d})) \times C([0, T]; \mathbb{R}^m)$  and  $\Psi : \mathbb{R}^m \rightarrow \mathbb{R}$  are measurable and bounded from below.

( $\mathcal{F}_1$ )  $\mathcal{L}$  is continuous in the variables  $(\boldsymbol{\mu}, \mathbf{Y})$ .

( $\mathcal{F}_2$ ) There exists a function  $M_{\mathcal{K}'} \in L^1(0, T)$  such that for a.a.  $t \in [0, T]$  and all  $(\boldsymbol{\mu}, \mathbf{Y}) \in \mathcal{K}'$

$$|\mathcal{L}(t, \boldsymbol{\mu}, \mathbf{Y})| \leq M_{\mathcal{K}'}(t)$$

( $\mathcal{F}_3$ )  $\Psi : \mathbb{R}^m \rightarrow \mathbb{R}$  is convex.

We are now ready to prove Theorem 1.5.

**4.3. Proof of Theorem 1.5.** By Theorem 4.2, for any fixed  $\mathbf{u} \in \mathcal{A}$  the solution  $(\boldsymbol{\mu}, \mathbf{Y})$  of (4.1) belongs to  $\mathcal{K}'$ . First, we show that  $\mathcal{F}[\mathbf{u}] < \infty$  for all  $\mathbf{u} \in \mathcal{A}$ . Indeed, by (4.2) and letting  $(\boldsymbol{\mu}, \mathbf{Y})$  be the solution of (4.1) with control  $\mathbf{u}$ , we have

$$|\mathbf{u}(t, \boldsymbol{\mu})| \leq K_u(1 + \mathcal{C}(K_G)^{\frac{1}{p}}).$$

Now, let  $M_\Psi = \sup_{|x| \leq K_u(1 + \mathcal{C}(K_G)^{\frac{1}{p}})} \Psi(x)$ , which exists since  $\Psi$  is convex and thus continuous. Hence

$$\mathcal{F}[\mathbf{u}] \leq \int_0^T M_{\mathcal{L}}(t) dt + M_\Psi T < \infty.$$

Now, we show that for any two controls  $\mathbf{u}_1, \mathbf{u}_2 \in \mathcal{A}$  such that for all  $\boldsymbol{\nu} \in \mathcal{K}$  and  $t \in [0, T]$  it holds  $\mathbf{u}_1(t, \boldsymbol{\nu}) = \mathbf{u}_2(t, \boldsymbol{\nu})$ , we have  $\mathcal{F}[\mathbf{u}_1] = \mathcal{F}[\mathbf{u}_2]$ . To do this, let  $(\boldsymbol{\mu}^1, \mathbf{Y}^1)$  and  $(\boldsymbol{\mu}^2, \mathbf{Y}^2)$  be the solutions of (4.1) with controls  $\mathbf{u}_1$  and  $\mathbf{u}_2$ . Since  $\boldsymbol{\mu}^1 \in \mathcal{K}$ , then  $\mathbf{u}_1(t, \boldsymbol{\mu}^1) = \mathbf{u}_2(t, \boldsymbol{\mu}^1)$ , and hence we have

$$\mathbf{Y}^1(t) = \bar{\mathbf{Y}} + \int_0^t (F[s, \boldsymbol{\mu}^1](\mathbf{Y}^1) + \mathbf{u}_1(s, \boldsymbol{\mu}^1)) ds = \bar{\mathbf{Y}} + \int_0^t (F[s, \boldsymbol{\mu}^1](\mathbf{Y}^1) + \mathbf{u}_2(s, \boldsymbol{\mu}^1)) ds.$$

Thus,  $\mathbf{Y}^1$  also solves (4.3) with control  $\mathbf{u}_2$  and measure  $\boldsymbol{\mu}^1$ , i.e.

$$\mathcal{S}[\boldsymbol{\mu}^1, \mathbf{u}_2] = \mathbf{H}^1 = \mathcal{S}[\boldsymbol{\mu}^1, \mathbf{u}_1].$$

However, this means that

$$\begin{aligned} G_{\mathbf{u}_1}[t, \boldsymbol{\mu}^1](z) &= \mathbf{v}[t, \boldsymbol{\mu}^1](z) + \mathbf{w}[t, \mathcal{S}[\boldsymbol{\mu}^1, \mathbf{u}_1]](z) \\ &= \mathbf{v}[t, \boldsymbol{\mu}^1](z) + \mathbf{w}[t, \mathcal{S}[\boldsymbol{\mu}^1, \mathbf{u}_2]](z) = G_{\mathbf{u}_2}[t, \boldsymbol{\mu}^1](z) \end{aligned}$$

and then  $\boldsymbol{\mu}^1$  solves (4.5) with field  $G_{\mathbf{u}_2}$ . However, also  $\boldsymbol{\mu}^2$  solves (4.5) with the field  $G_{\mathbf{u}_2}$ . Hence, since for a given field  $G_{\mathbf{u}_2}$  (4.5) admits a unique solution,  $\boldsymbol{\mu}^1 = \boldsymbol{\mu}^2$  and also  $\mathbf{H}^1 = \mathbf{H}^2$ . Then, it is clear, by definition of  $\mathcal{F}$ , that  $\mathcal{F}[\mathbf{u}_1] = \mathcal{F}[\mathbf{u}_2]$ . Thus we are in a position to consider the equivalence relation  $\sim$  over  $\mathcal{A}$  defined as

$$\mathbf{u}_1 \sim \mathbf{u}_2 \quad \Leftrightarrow \quad \mathbf{u}_1(t, \boldsymbol{\mu}) = \mathbf{u}_2(t, \boldsymbol{\mu}) \quad \forall t \in [0, T], \quad \forall \boldsymbol{\mu} \in \mathcal{K},$$

the quotient set  $\tilde{\mathcal{A}} = \mathcal{A} / \sim$  and the functional  $\tilde{\mathcal{F}} : \tilde{\mathcal{A}} \rightarrow \mathbb{R}$  defined as

$$\tilde{\mathcal{F}}[[\mathbf{u}]_\sim] = \mathcal{F}[\mathbf{u}], \quad \forall [\mathbf{u}]_\sim \in \tilde{\mathcal{A}}.$$

To actually work with  $\tilde{\mathcal{F}}$ , we need to provide a suitable representation of the quotient set  $\tilde{\mathcal{A}}$ . To do this, first notice that for any  $\mathbf{u} \in \mathcal{A}$  we consider a function  $\mathbf{U} : [0, T] \rightarrow C(\mathcal{K}; \mathbb{R}^m)$  defined as  $\mathbf{U}(t) = \mathbf{u}(t, \cdot)$  for  $t \in [0, T]$ , so that  $\mathbf{U}(t)(\boldsymbol{\mu}) = \mathbf{u}(t, \boldsymbol{\mu})$  for any  $\boldsymbol{\mu} \in \mathcal{K}$ . Since

$$|\mathbf{u}(t, \boldsymbol{\mu})| \leq K_u(1 + (\mathcal{C}(K_G))^{\frac{1}{p}}),$$

and then

$$\int_0^T \sup_{\boldsymbol{\mu} \in \mathcal{K}} |\mathbf{u}(t, \boldsymbol{\mu})| dt \leq K_u(1 + (\mathcal{C}(K_G))^{\frac{1}{p}})T,$$

it is clear that  $\mathbf{U} \in L^1([0, T]; C(\mathcal{K}; \mathbb{R}^m))$ . Let  $\Pi : \mathcal{A} \rightarrow L^1([0, T]; C(\mathcal{K}; \mathbb{R}^m))$  be the function that maps  $\mathbf{u}$  into  $\mathbf{U}$  as before. If  $\mathbf{u}_1 \sim \mathbf{u}_2$ , then  $\Pi \mathbf{u}_1 = \Pi \mathbf{u}_2$ , i.e.,  $\Pi$  is compatible with the equivalence relation  $\sim$ . In particular, this means that we can define the map  $\tilde{\Pi} : \tilde{\mathcal{A}} \rightarrow \Pi \mathcal{A}$  such that  $\tilde{\Pi}[[\mathbf{u}]_\sim] = \Pi \mathbf{u}$ . By definition, this map is surjective. Furthermore, assume that  $\tilde{\Pi}[[\mathbf{u}_1]_\sim] = \tilde{\Pi}[[\mathbf{u}_2]_\sim]$ . Then  $\Pi \mathbf{u}_1 = \Pi \mathbf{u}_2$  hence  $\mathbf{u}_1(t, \boldsymbol{\mu}) = \mathbf{u}_2(t, \boldsymbol{\mu})$  for all  $t \in [0, T]$  and  $\boldsymbol{\mu} \in \mathcal{K}$ ,



which in turn implies  $\mathbf{u}_1 \sim \mathbf{u}_2$  and  $[\mathbf{u}_1]_{\sim} = [\mathbf{u}_2]_{\sim}$ . In particular, we have shown that  $\tilde{\Pi}$  is a bijection. From now on, without loss of generality, we identify  $\tilde{\mathcal{A}}$  with  $\Pi\mathcal{A}$  and

$$\mathcal{F}[\mathbf{u}] = \tilde{\mathcal{F}}[\Pi\mathbf{u}].$$

We can then rewrite Problem 1 as follows:

<b>Problem 1*</b>	
Find $\mathbf{U}^* \in \tilde{\mathcal{A}}$ such that	$\tilde{\mathcal{F}}[\mathbf{U}^*] = \min_{\mathbf{U} \in \tilde{\mathcal{A}}} \tilde{\mathcal{F}}[\mathbf{U}].$

Now let us consider a minimizing sequence  $\{\mathbf{U}^n\}_{n \in \mathbb{N}} \subset \tilde{\mathcal{A}}$  for  $\tilde{\mathcal{F}}$ . We will prove that this minimizing sequence admits (up to a subsequence) a limit  $\mathbf{U}^*$  in some sense that will be specified later. To do this, however, we preliminarily need to prove that  $\{\mathbf{U}^n\}_{n \in \mathbb{N}}$  is uniformly integrable and uniformly tight, i.e.

$$\lim_{R \rightarrow +\infty} \sup_{n \in \mathbb{N}} \int_{\{t \in [0, T] : \sup_{\mu \in \mathcal{K}} |\mathbf{U}^n(t)(\mu)| > R\}} \sup_{\mu \in \mathcal{K}} |\mathbf{U}^n(t)(\mu)| dt = 0, \quad (4.7)$$

and for all  $\varepsilon > 0$  there exists a compact-valued multifunction  $\Gamma_\varepsilon : [0, T] \rightarrow 2^{C(\mathcal{K}; \mathbb{R}^m)}$  with measurable graph such that

$$\sup_{n \in \mathbb{N}} |\{t \in [0, T] : \mathbf{U}^n(t) \notin \Gamma_\varepsilon(t)\}| < \varepsilon,$$

respectively. Let us first show the uniform integrability. Indeed, for all  $n \in \mathbb{N}$ , there exists  $\mathbf{u}^n \in \mathcal{A}$  such that  $\mathbf{U}^n = \Pi\mathbf{u}^n$  and then, for all  $t \in [0, T]$  and  $\mu \in \mathcal{K}$ ,

$$|\mathbf{U}^n(t)(\mu)| = |\mathbf{u}^n(t, \mu)| \leq K_u(1 + (\mathcal{C}(K_G))^{\frac{1}{p}}).$$

Hence, for  $R > K_u(1 + (\mathcal{C}(K_G))^{\frac{1}{p}})$ , we clearly have

$$\int_{\{t \in [0, T] : \sup_{\mu \in \mathcal{K}} |\mathbf{U}^n(t)(\mu)| > R\}} \sup_{\mu \in \mathcal{K}} |\mathbf{U}^n(t)(\mu)| dt = 0$$

and then (4.7) holds.

Now we show the uniform tightness. To do this, we fix  $t \in [0, T]$  and notice that, by Assumption  $(\mathcal{A}_2)$  it holds

$$|\mathbf{U}^n(t)(\boldsymbol{\nu}^1) - \mathbf{U}^n(t)(\boldsymbol{\nu}^2)| \leq L_u \sup_{0 \leq t \leq T} \mathcal{W}_p(\mu_t^1, \mu_t^2), \quad \forall \boldsymbol{\mu}^1, \boldsymbol{\mu}^2 \in \mathcal{K}.$$

Notice that the uniform topology of  $C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$  is generated by the metric

$$\mathfrak{d}(\boldsymbol{\mu}^1, \boldsymbol{\mu}^2) = \sup_{t \in [0, T]} \mathcal{W}_p(\mu_t^1, \mu_t^2), \quad \forall \boldsymbol{\mu}^1, \boldsymbol{\mu}^2 \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d})).$$

Hence by Lemma B.3,  $\mathcal{K}$  equipped with  $\mathfrak{d}$  is a compact Hausdorff space. Defining

$$\tilde{\mathcal{K}} = \left\{ f \in C(\mathcal{K}; \mathbb{R}^m) : \sup_{\mu \in \mathcal{K}} |f(\mu)| + \sup_{\substack{\mu^1, \mu^2 \\ \mu^1 \neq \mu^2}} \frac{|f(\mu^1) - f(\mu^2)|}{\mathfrak{d}(\mu^1, \mu^2)} \leq L_u + K_u(1 + (\mathcal{C}(K_G))^{\frac{1}{p}}) \right\},$$

which is a compact subset of  $C(\mathcal{K}; \mathbb{R}^m)$  by the Arzelá-Ascoli Theorem, we have  $\mathbf{U}^n(t) \in \tilde{\mathcal{K}}$  for all  $t \in [0, T]$  and  $n \in \mathbb{N}$ . Hence, to get the uniform tightness of  $\{\mathbf{U}^n\}_{n \in \mathbb{N}}$  it is sufficient to set  $\Gamma_\varepsilon(t) = \tilde{\mathcal{K}}$  for all  $\varepsilon > 0$  and  $t \in [0, T]$ .

Now, we can use a suitable generalization of the Dunford-Pettis theorem, as in [10, Theorem 1.3], which tells us that there exists a function  $\mathbf{U}^* \in L^1([0, T]; C(\mathcal{K}, \mathbb{R}^m))$  such that, up to a non-relabelled subsequence, for all measurable subsets  $E \subset [0, T]$

$$\lim_{n \rightarrow \infty} \sup_{\mu \in \mathcal{K}} \frac{1}{|E|} \left| \int_E (\mathbf{U}^n(t)(\mu) - \mathbf{U}^*(t)(\mu)) dt \right| = 0. \quad (4.8)$$

Then we need to show that  $\mathbf{U}^* \in \tilde{\mathcal{A}}$ . To do this, first observe that for any  $\mu^1, \mu^2 \in \mathcal{K}$  it holds, by Assumption  $(\mathcal{A}_2)$ ,

$$\frac{1}{|E|} \left| \int_E (U^n(t)_j(\mu^1) - U_j^n(t)(\mu^2)) dt \right| \leq \frac{L_u}{m} \sup_{0 \leq s \leq \sup E} \mathcal{W}_p(\mu_s^1, \mu_s^2),$$

for any  $j = 1, \dots, m$ , where  $\mathbf{U}^n = (U_1^n, \dots, U_m^n)$ . Taking the limit as  $n \rightarrow \infty$  and using (4.8) we get

$$\frac{1}{|E|} \left| \int_E (U_j^*(t)(\mu^1) - U_j^*(t)(\mu^2)) dt \right| \leq \frac{L_u}{m} \sup_{0 \leq s \leq \sup E} \mathcal{W}_p(\mu_s^1, \mu_s^2), \quad (4.9)$$

where  $\mathbf{U}^* = (U_1^*, \dots, U_m^*)$ . Since  $\mathbf{U}^* \in L^1([0, T]; C(\mathcal{K}; \mathbb{R}^m))$ , by [6, Proposition 1.2.2], we know that almost any  $t \in [0, T]$  is a Lebesgue point for  $\mathbf{U}^*$ , i.e. there exists a set  $E_{\text{Leb}} \subset [0, T]$  such that  $|[0, T] \setminus E_{\text{Leb}}| = 0$  and for all  $t \in E_{\text{Leb}}$  it holds

$$\lim_{h \rightarrow 0} \frac{1}{h} \int_t^{t+h} \sup_{\mu \in \mathcal{K}} |\mathbf{U}^*(s)(\mu) - \mathbf{U}^*(t)(\mu)| ds = 0.$$

In particular, if we fix  $\mu \in \mathcal{K}$  for all  $t \in E_{\text{Leb}}$  we have

$$\lim_{h \rightarrow 0} \left| \frac{1}{h} \int_t^{t+h} \mathbf{U}^*(s)(\mu) ds - \mathbf{U}^*(t)(\mu) \right| \leq \lim_{h \rightarrow 0} \frac{1}{h} \int_t^{t+h} \sup_{\mu \in \mathcal{K}} |\mathbf{U}^*(s)(\mu) - \mathbf{U}^*(t)(\mu)| ds = 0.$$

Fix  $t \in E_{\text{Leb}}$  and observe that by (4.9) for any  $h > 0$  it holds

$$\left| \frac{1}{h} \int_t^{t+h} (U_j^*(s)(\mu^1) - U_j^*(s)(\mu^2)) ds \right| \leq \frac{L_u}{m} \sup_{0 \leq s \leq t+h} \mathcal{W}_p(\mu_s^1, \mu_s^2).$$

Now we take the limit as  $h \rightarrow 0$  on both sides: on the left-hand side, we use the fact that  $t \in E_{\text{Leb}}$ , while on the right-hand side we use the continuity of  $\mu^1, \mu^2 \in \mathcal{K}$ . Finally, for

any  $j = 1, \dots, m$  we achieve

$$|U_j^*(t)(\boldsymbol{\mu}^1) - U_j^*(t)(\boldsymbol{\mu}^2)| \leq \frac{L_u}{m} \sup_{0 \leq s \leq t} \mathcal{W}_p(\mu_s^1, \mu_s^2). \quad (4.10)$$

For  $t \notin E_{\text{Leb}}$ , without loss of generality, we set  $\mathbf{U}^*(t) \equiv 0$ , since  $|[0, T] \setminus E_{\text{Leb}}| = 0$  and (4.10) still holds. Hence, we have  $\mathbf{U}^*$  satisfies (4.10) for all  $t \in [0, T]$ . Next, notice that, since  $\delta_0 \in \mathcal{K}$  and by Assumption  $(\mathcal{A}_1)$ , for any measurable subset  $E \subset [0, T]$  we have

$$\frac{1}{|E|} \left| \int_E \mathbf{U}^n(t)(\delta_0) dt \right| \leq M_u.$$

With the same argument as before, we get, for all  $t \in E_{\text{Leb}}$ ,

$$|\mathbf{U}^*(t)(\delta_0)| \leq M_u. \quad (4.11)$$

Combining this with the fact that we set  $\mathbf{U}^*(t) \equiv 0$  whenever  $t \notin E_{\text{Leb}}$ , we infer (4.11) holds for all  $t \in [0, T]$ .

Now, we fix  $t \in [0, T]$  and introduce another equivalence relation  $\sim_t$  on  $\mathcal{K}$ , as follows:

$$\boldsymbol{\mu}^1 \sim_t \boldsymbol{\mu}^2 \Leftrightarrow \mu_s^1 = \mu_s^2, \forall s \in [0, t].$$

Notice that, by (4.10), if  $\boldsymbol{\mu}^1 \sim_t \boldsymbol{\mu}^2$  then  $U_j^*(t)(\boldsymbol{\mu}^1) = U_j^*(t)(\boldsymbol{\mu}^2)$  for all  $j = 1, \dots, m$ . This means that  $U_j^*(t)$  is compatible with the equivalence relation  $\sim_t$  and then we can define the function  $\tilde{U}_j^*(t) : \tilde{\mathcal{K}}_t \rightarrow \mathbb{R}$ , where  $\tilde{\mathcal{K}}_t = \mathcal{K} / \sim_t$  is the quotient of  $\mathcal{K}$  with respect to the equivalence relation  $\sim_t$ , as

$$\tilde{U}_j^*(t)([\boldsymbol{\mu}]_{\sim_t}) = U_j^*(t)(\boldsymbol{\mu}).$$

Let us also denote by  $\pi_t : \mathcal{K} \mapsto \tilde{\mathcal{K}}_t$  the projection, i.e.  $\pi_t \boldsymbol{\mu} = [\boldsymbol{\mu}]_{\sim_t}$ , and endow  $\tilde{\mathcal{K}}_t$  with the quotient topology. Notice also that the function  $\Pi_t : \mathcal{K} \rightarrow C([0, t]; \mathcal{W}_p(\mathbb{R}^{2d}))$  defined as  $(\Pi_t \boldsymbol{\mu})_s = \mu_s$  for all  $s \in [0, t]$  (i.e.,  $\Pi_t$  is the function that maps  $\boldsymbol{\mu}$  to its restriction on the interval  $[0, t]$ ) is compatible with  $\sim_t$ , i.e., if  $\boldsymbol{\mu}^1 \sim_t \boldsymbol{\mu}^2$  then  $\Pi_t \boldsymbol{\mu}^1 = \Pi_t \boldsymbol{\mu}^2$ . We endow  $\Pi_t \mathcal{K}$  with the uniform metric  $\mathfrak{d}_t$  on  $C([0, t]; \mathcal{W}_p(\mathbb{R}^{2d}))$ , i.e.,

$$\mathfrak{d}_t(\boldsymbol{\mu}^1, \boldsymbol{\mu}^2) = \sup_{0 \leq s \leq t} \mathcal{W}_p(\mu_s^1, \mu_s^2), \boldsymbol{\mu}^1, \boldsymbol{\mu}^2 \in \Pi_t \mathcal{K}$$

It is clear that  $\Pi_t$  is non-expansive, i.e., for all  $\boldsymbol{\mu}^1, \boldsymbol{\mu}^2 \in \mathcal{K}$  it holds

$$\mathfrak{d}_t(\Pi_t \boldsymbol{\mu}^1, \Pi_t \boldsymbol{\mu}^2) \leq \mathfrak{d}(\boldsymbol{\mu}^1, \boldsymbol{\mu}^2).$$

Since  $\Pi_t$  is compatible with  $\sim_t$ , we define the map  $\tilde{\Pi}_t : \tilde{\mathcal{K}}_t \rightarrow \Pi_t \mathcal{K}$  as  $\tilde{\Pi}_t([\boldsymbol{\mu}]_{\sim_t}) = \Pi_t \boldsymbol{\mu}$ , which is clearly continuous with respect of the quotient topology on  $\tilde{\mathcal{K}}_t$ . Furthermore, notice that if  $\Pi_t \boldsymbol{\mu}^1 = \Pi_t \boldsymbol{\mu}^2$ , then, by definition,  $\boldsymbol{\mu}^1 \sim_t \boldsymbol{\mu}^2$ . Hence,  $\tilde{\Pi}_t$  is bijective. Finally, consider an open set  $E \subset \tilde{\mathcal{K}}_t$ . We want to prove that  $\tilde{\Pi}_t E$  is an open set. To do this, fix  $[\boldsymbol{\mu}]_{\sim_t} \in E$ . We want to fix a special representative  $\boldsymbol{\mu} \in [\boldsymbol{\mu}]_{\sim_t}$ . Precisely, let  $\tilde{\boldsymbol{\mu}} = \tilde{\Pi}_t([\boldsymbol{\mu}]_{\sim_t})$  and define  $\boldsymbol{\mu} \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$  by setting  $\mu_s = \tilde{\mu}_{\min\{s, t\}}$  for  $s \in [0, T]$ . It is not difficult to check that  $\boldsymbol{\mu} \in \mathcal{K}$ . In particular  $\boldsymbol{\mu} \in \pi_t^{-1}(U)$ , and notice that since  $\pi_t^{-1}(E)$  is an open set in  $\mathcal{K}$ , there exists  $\varepsilon > 0$  such that  $B_\varepsilon(\boldsymbol{\mu}) := \{\boldsymbol{\nu} \in \mathcal{K} : \mathfrak{d}(\boldsymbol{\mu}, \boldsymbol{\nu}) < \varepsilon\} \subset \pi_t^{-1}(E)$ .

Now, we define  $B_\varepsilon(\Pi_t \boldsymbol{\mu}; t) := \{\boldsymbol{\nu} \in \Pi_t \mathcal{K} : \mathfrak{d}_t(\Pi_t \boldsymbol{\mu}, \boldsymbol{\nu}) < \varepsilon\}$ . Since  $\Pi_t$  is non-expansive, we have  $\Pi_t B_\varepsilon(\boldsymbol{\mu}) \subseteq B_\varepsilon(\Pi_t \boldsymbol{\mu}; t)$ . On the other hand, consider any  $\tilde{\boldsymbol{\nu}} \in B_\varepsilon(\Pi_t \boldsymbol{\mu}; t)$  and define  $\boldsymbol{\nu} \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$  by setting  $\nu_s = \tilde{\nu}_{\min\{s, t\}}$  for  $s \in [0, T]$ . Again, it is not difficult to check that  $\boldsymbol{\nu} \in \mathcal{K}$  and  $\mathfrak{d}(\boldsymbol{\nu}, \boldsymbol{\mu}) < \varepsilon$ . Hence,  $\tilde{\boldsymbol{\nu}} = \Pi_t \boldsymbol{\nu}$  with  $\boldsymbol{\nu} \in B_\varepsilon(\boldsymbol{\mu})$  and then  $\tilde{\boldsymbol{\nu}} \in \Pi_t B_\varepsilon(\boldsymbol{\mu})$ . This shows that  $\Pi_t B_\varepsilon(\boldsymbol{\mu}) = B_\varepsilon(\Pi_t \boldsymbol{\mu}; t)$ . Recalling that  $\Pi_t \boldsymbol{\mu} = \tilde{\Pi}_t[\boldsymbol{\mu}]_{\sim_t}$ , we get

$$B_\varepsilon(\Pi_t \boldsymbol{\mu}; t) = \Pi_t B_\varepsilon(\boldsymbol{\mu}) = \tilde{\Pi}_t \pi_t B_\varepsilon(\boldsymbol{\mu}) \subseteq \tilde{\Pi}_t E,$$

where we used the fact that since  $B_\varepsilon(\boldsymbol{\mu}) \subseteq \pi_t^{-1}(E)$ . Since  $[\boldsymbol{\mu}]_{\sim_t} \in E$  is arbitrary,  $\tilde{\Pi}_t E$  is an open set. In conclusion,  $\tilde{\Pi}_t : \tilde{\mathcal{K}}_t \rightarrow \Pi_t \mathcal{K}$  is a homeomorphism.

Now we notice that  $\tilde{U}_j^*(t) \circ \tilde{\Pi}_t^{-1}$  is  $\frac{L_u}{m}$ -Lipschitz. Indeed, for any  $\tilde{\boldsymbol{\mu}}^1, \tilde{\boldsymbol{\mu}}^2 \in \Pi_t \mathcal{K}$ , letting  $\boldsymbol{\mu}^j \in \Pi_t^{-1} \tilde{\boldsymbol{\mu}}^j$  for  $j = 1, 2$ , it holds

$$\begin{aligned} \left| \tilde{U}_j^*(t)(\tilde{\Pi}_t^{-1}(\tilde{\boldsymbol{\mu}}^1)) - \tilde{U}_j^*(t)(\tilde{\Pi}_t^{-1}(\tilde{\boldsymbol{\mu}}^2)) \right| &= \left| \tilde{U}_j^*(t)([\boldsymbol{\mu}^1]_{\sim_t}) - \tilde{U}_j^*(t)([\boldsymbol{\mu}^2]_{\sim_t}) \right| \\ &= \left| U_j^*(t)(\boldsymbol{\mu}^1) - U_j^*(t)(\boldsymbol{\mu}^2) \right| \\ &\leq \frac{L_u}{m} \sup_{0 \leq s \leq t} \mathcal{W}_p(\mu_s^1, \mu_s^2) = \frac{L_u}{m} \mathfrak{d}(\tilde{\boldsymbol{\mu}}^1, \tilde{\boldsymbol{\mu}}^2). \end{aligned}$$

Let us recall that also  $\Pi_t \mathcal{K}$  is a compact subset of  $C([0, t]; \mathcal{W}_p(\mathbb{R}^{2d}))$ . Now we can use McShane's Extension Theorem (see [35]) to state that the function  $\tilde{u}_j^*(t, \cdot)$  defined as

$$\tilde{u}_j^*(t, \boldsymbol{\nu}) = \min_{\boldsymbol{\mu} \in \Pi_t \mathcal{K}} \left\{ \tilde{U}_j^*(t)(\tilde{\Pi}_t^{-1} \boldsymbol{\mu}) + \frac{L_u}{m} \mathfrak{d}(\boldsymbol{\mu}, \boldsymbol{\nu}) \right\}, \quad \boldsymbol{\nu} \in C([0, t]; \mathcal{W}_p(\mathbb{R}^{2d})),$$

is the maximal  $\frac{L_u}{m}$ -Lipschitz extension of  $\tilde{U}_j^*(t) \circ \tilde{\Pi}_t^{-1}$  on  $C([0, t]; \mathcal{W}_p(\mathbb{R}^{2d}))$ . Finally, we set, for all  $t \in [0, T]$  and  $\boldsymbol{\mu} \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$ ,

$$u_j^*(t, \boldsymbol{\mu}) = \tilde{u}_j^*(t, \Pi_t \boldsymbol{\mu})$$

and  $\boldsymbol{u}^* = (u_j^*)_{j=1, \dots, m}$ . We can clearly rewrite  $u_j^*$  as follows:

$$u_j^*(t, \boldsymbol{\nu}) = \min_{\boldsymbol{\mu} \in \mathcal{K}} \left\{ U_j^*(t)(\boldsymbol{\mu}) + \frac{L_u}{m} \sup_{0 \leq s \leq t} \mathcal{W}_p(\boldsymbol{\mu}, \boldsymbol{\nu}) \right\}, \quad \boldsymbol{\nu} \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d})),$$

so that it is clear that  $u_j^*$  is measurable in the variable  $t$ , as it is the minimum of a family of measurable functions. Concerning the continuity on  $\boldsymbol{\mu}$ , notice that for all  $\boldsymbol{\mu}^1, \boldsymbol{\mu}^2 \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$ ,

$$\left| u_j^*(t, \boldsymbol{\mu}^1) - u_j^*(t, \boldsymbol{\mu}^2) \right| = \left| \tilde{u}_j^*(t, \Pi_t \boldsymbol{\mu}^1) - \tilde{u}_j^*(t, \Pi_t \boldsymbol{\mu}^2) \right| \leq \frac{L_u}{m} \sup_{0 \leq s \leq t} \mathcal{W}_p(\mu_s^1, \mu_s^2).$$

This shows that  $\mathbf{u}^*$  is continuous in the  $\boldsymbol{\mu}$  variable, so that  $(\mathcal{A}_0)$  holds, and that  $(\mathcal{A}_2)$  also holds. Furthermore, notice that

$$\begin{aligned} |\mathbf{u}^*(t, \boldsymbol{\delta}_0)|^2 &= \sum_{j=1}^m |\tilde{u}_j^*(t, \Pi_t \boldsymbol{\delta}_0)|^2 = \sum_{j=1}^m \left| \tilde{U}_j^*(t)(\tilde{\Pi}_t^{-1} \Pi_t \boldsymbol{\delta}_0) \right|^2 \\ &= \sum_{j=1}^m \left| \tilde{U}_j^*(t)([\boldsymbol{\delta}_0]_{\sim_t}) \right|^2 = \sum_{j=1}^m |U_j^*(t)(\boldsymbol{\delta}_0)|^2 \leq M_0^2, \end{aligned}$$

by (4.11), thus  $(\mathcal{A}_1)$  also holds. As a consequence,  $\mathbf{u}^* \in \mathcal{A}$ . However, if  $\boldsymbol{\mu} \in \mathcal{K}$ , we have

$$u_j^*(t, \boldsymbol{\mu}) = \tilde{u}_j^*(t, \Pi_t \boldsymbol{\mu}) = \tilde{U}_j^*(t)(\tilde{\Pi}_t^{-1} \Pi_t \boldsymbol{\mu}) = \tilde{U}_j^*(t)([\boldsymbol{\mu}]_{\sim_t}) = U_j^*(t)(\boldsymbol{\mu}),$$

i.e.  $\mathbf{U}^* = \Pi \mathbf{u}^* \in \tilde{\mathcal{A}}$ .

It remains to show that  $\mathbf{U}^*$  is solution of Problem 1\*. Recall that  $\{\mathbf{U}^n\}_{n \in \mathbb{N}} \subset \tilde{\mathcal{A}}$ , hence there exists a sequence of controls  $\{\mathbf{u}^n\}_{n \in \mathbb{N}} \subset \mathcal{A}$  such that  $\mathbf{U}^n = \Pi \mathbf{u}^n$ . Let  $(\boldsymbol{\mu}^n, \mathbf{Y}^n)$  be the respective solutions of (4.1). Furthermore, let  $(\boldsymbol{\mu}^*, \mathbf{Y}^*)$  be the solution of (4.1) with control  $\mathbf{u}^*$ . Observe that  $\Psi : \mathbb{R}^m \rightarrow \mathbb{R}$  is a convex function and let  $\Psi^*$  be its Legendre transform, i.e.

$$\Psi^*(x) = \sup_{y \in \mathbb{R}^m} (y \cdot x - \Psi(y)), \quad x \in \mathbb{R}^m.$$

Recall that  $\Psi^* : \mathbb{R}^m \rightarrow \mathbb{R}$  is still a convex function and that the Legendre transform is an involution, i.e.,

$$\Psi(x) = \Psi^{**}(x) = \sup_{y \in \mathbb{R}^m} (y \cdot x - \Psi^*(y)), \quad x \in \mathbb{R}^m.$$

Fix  $y \in \mathbb{R}^m$  and let  $\Psi(x; y) = y \cdot x - \Psi^*(y)$ , which is an affine function. Then we have, for any measurable subset  $E \subset [0, T]$

$$\begin{aligned} \int_E \Psi(\mathbf{U}^n(t)(\boldsymbol{\mu}^n); y) dt &= y \cdot \int_E \mathbf{U}^n(t)(\boldsymbol{\mu}^n) dt - \Psi^*(y)|E| \\ &= y \cdot \int_E \mathbf{U}^n(t)(\boldsymbol{\mu}^*) dt - \Psi^*(y)|E| + y \cdot \mathcal{R}_n(E), \end{aligned} \quad (4.12)$$

where

$$\mathcal{R}_n(E) = \int_E (\mathbf{U}^n(t)(\boldsymbol{\mu}^n) - \mathbf{U}^n(t)(\boldsymbol{\mu}^*)) dt.$$

Let us show that  $\mathcal{R}_n \rightarrow 0$ . To do this, first notice that

$$|\mathcal{R}_n(E)| \leq \int_E |\mathbf{u}^n(t, \boldsymbol{\mu}^n) - \mathbf{u}^n(t, \boldsymbol{\mu}^*)| dt \leq L_u |E| \sup_{0 \leq s \leq T} \mathcal{W}_p(\boldsymbol{\mu}, \boldsymbol{\mu}^*).$$

By (4.8), we know that

$$\lim_{n \rightarrow \infty} \int_E \mathbf{u}^n(t, \boldsymbol{\mu}) dt = \int_E \mathbf{u}^*(t, \boldsymbol{\mu}) dt, \quad \forall \boldsymbol{\mu} \in \mathcal{K},$$

hence we can use Theorem 4.2 to conclude that

$$\lim_{n \rightarrow +\infty} |\mathcal{R}_n(E)| = 0.$$

Taking the limit as  $n \rightarrow \infty$  in (4.12) we get

$$\lim_{n \rightarrow +\infty} \int_E \Psi(\mathbf{U}^n(t)(\boldsymbol{\mu}^n); y) dt = y \cdot \int_E \mathbf{U}^*(t)(\boldsymbol{\mu}^*) dt - \Psi^*(y)|E| = \int_E \Psi(\mathbf{U}^*(t)(\boldsymbol{\mu}^*); y) dt.$$

In particular, taking the supremum only on the left-hand side, we have

$$\begin{aligned} \int_E \Psi(\mathbf{U}^*(t)(\boldsymbol{\mu}^*); y) dt &\leq \sup_{y \in \mathbb{R}^m} \lim_{n \rightarrow +\infty} \int_E \Psi(\mathbf{U}^n(t)(\boldsymbol{\mu}^n); y) dt \\ &\leq \liminf_{n \rightarrow +\infty} \sup_{y \in \mathbb{R}^m} \int_E \Psi(\mathbf{U}^n(t)(\boldsymbol{\mu}^n); y) dt \\ &\leq \liminf_{n \rightarrow +\infty} \int_E \Psi(\mathbf{U}^n(t)(\boldsymbol{\mu}^n)) dt =: \mathcal{G}(E). \end{aligned}$$

Notice that  $\mathcal{G}$  is superadditive. Indeed, if we consider two measurable sets  $E_1, E_2 \subset [0, T]$  with  $E_1 \cap E_2 = \emptyset$ , then

$$\int_{E_1 \cup E_2} \Psi(\mathbf{U}^n(t)(\boldsymbol{\mu}^n)) dt = \int_{E_1} \Psi(\mathbf{U}^n(t)(\boldsymbol{\mu}^n)) dt + \int_{E_2} \Psi(\mathbf{U}^n(t)(\boldsymbol{\mu}^n)) dt$$

and taking the limit inferior we achieve

$$\mathcal{G}(E_1 \cup E_2) \geq \mathcal{G}(E_1) + \mathcal{G}(E_2).$$

Hence, by the Localization Lemma (see [12, Proposition 1.16]), we have

$$\begin{aligned} \int_E \Psi(\mathbf{U}^*(t)(\boldsymbol{\mu}^*)) dt &= \int_E \sup_{y \in \mathbb{R}^m} \Psi(\mathbf{U}^*(t)(\boldsymbol{\mu}^*); y) dt \\ &\leq \mathcal{G}(E) = \liminf_{n \rightarrow +\infty} \int_E \Psi(\mathbf{U}^n(t)(\boldsymbol{\mu}^n)). \end{aligned} \tag{4.13}$$

Next, recall that by Theorem 4.2 we have

$$\lim_{j \rightarrow \infty} \sup_{0 \leq t \leq T} (\mathcal{W}_p(\mu_t^j, \mu_y^*) + |\mathbf{H}^j(t) - \mathbf{H}^*(t)|) = 0$$

hence, by  $(\mathcal{F}_1)$ ,

$$\lim_{j \rightarrow \infty} \mathcal{L}(t, \boldsymbol{\mu}^j, \mathbf{Y}^j) = \mathcal{L}(t, \boldsymbol{\mu}^*, \mathbf{Y}^*).$$

We also recall that, still by Theorem 4.2,  $(\boldsymbol{\mu}^j, \mathbf{H}^j) \in \mathcal{K}'$  for all  $j \in \mathbb{N}$ . Thus, by  $(\mathcal{F}_2)$  and dominated convergence,

$$\lim_{j \rightarrow \infty} \int_0^T \mathcal{L}(t, \boldsymbol{\mu}^j, \mathbf{Y}^j) dt = \int_0^T \mathcal{L}(t, \boldsymbol{\mu}^*, \mathbf{Y}^*) dt. \tag{4.14}$$

Finally, combining (4.13) and (4.14) we achieve

$$\tilde{\mathcal{F}}[\mathbf{U}^*] \leq \lim_{n \rightarrow \infty} \tilde{\mathcal{F}}[\mathbf{U}^n] = \inf_{\mathbf{U} \in \tilde{\mathcal{A}}} \mathcal{F}[\mathbf{U}],$$

i.e.,  $\mathbf{U}^*$  is solution of Problem 1\*. □

APPENDIX A. SOME SUFFICIENT CONDITIONS TO THE VALIDITY OF ASSUMPTION  $(\mathbf{v}_3)$ 

As already underlined in the introduction, it could be difficult to verify that  $(\mathbf{v}_3)$  actually holds. For such reason, let us give some sufficient conditions for  $(\mathbf{v}_3)$  to be true.

**Proposition A.1.** *Let Assumptions  $(\mathbf{v})$  (except for  $(\mathbf{v}_3)$ ) hold true and denote*

$$D([0, T]; \mathbb{R}^{2d}) = \{(X, V) \in C([0, T]; \mathbb{R}^{2d}) : X' = V\}.$$

Consider the following properties:

$(\mathbf{v}'_3)$  For any  $T > 0$  there exists a constant  $D \geq 0$  such that for any  $t \in [0, T]$ , any couple of measures  $\mu^1, \mu^2 \in \mathcal{W}_p(C([0, T]; \mathbb{R}^{2d}))$  supported on  $D([0, T]; \mathbb{R}^{2d})$  with  $\mu_0^1 = \mu_0^2$  and any  $\gamma \in \Pi(\mu^1, \mu^2)$  it holds

$$\begin{aligned} \int_{\mathbb{R}^{2d} \times \mathbb{R}^{2d}} |v_1 - v_2|^{p-2} (\mathbf{v}[t, \mu^1](z_1) - \mathbf{v}[t, \mu^2](z_2)) \cdot (v_1 - v_2) d\gamma_t(z_1, z_2) \\ \leq D \sup_{0 \leq s \leq t} \left( \int_{\mathbb{R}^{2d} \times \mathbb{R}^{2d}} |z_1 - z_2|^p d\gamma_s(z_1, z_2) \right) \end{aligned}$$

$(\mathbf{v}''_3)$  For any  $T > 0$  there exists a constant  $D \geq 0$  such that for any  $t \in [0, T]$ , any  $\mu^1, \mu^2 \in \mathcal{W}_p(C([0, T]; \mathbb{R}^{2d}))$  supported on  $D([0, T]; \mathbb{R}^{2d})$  with  $\mu_0^1 = \mu_0^2$  and  $z_1, z_2 \in \mathbb{R}^{2d}$

$$|\mathbf{v}[t, \mu^1](z_1) - \mathbf{v}[t, \mu^2](z_2)| \leq D \left( \sup_{0 \leq s \leq t} \mathcal{W}_p(\mu_s^1, \mu_s^2) + |z_1 - z_2| \right).$$

Then  $(\mathbf{v}''_3) \Rightarrow (\mathbf{v}'_3) \Rightarrow (\mathbf{v}_3)$ .

*Proof.* To show that  $(\mathbf{v}'_3) \Rightarrow (\mathbf{v}_3)$ , consider any two processes  $Z_j \in L^p(\Omega; C([0, T]; \mathbb{R}^{2d}))$  as in Assumption  $(\mathbf{v}_3)$  and let  $\mu^j = \text{Law}(Z_j)$  for  $j = 1, 2$ . Then  $\mu^j$  is supported on  $D([0, T]; \mathbb{R}^{2d})$  and  $\mu_0^1 = \mu_0^2$ . Set also  $\gamma = \text{Law}(Z_1, Z_2) \in \Pi(\mu^1, \mu^2)$ . Then, by  $(\mathbf{v}'_3)$ , we have

$$\begin{aligned} \mathbb{E} [ |V_1(t) - V_2(t)|^{p-2} (\mathbf{v}[t, \mu^1](Z_1(t)) - \mathbf{v}[t, \mu^1](Z_2(t))) \cdot (V_1(t) - V_2(t)) ] \\ \leq D \sup_{0 \leq s \leq t} \mathbb{E} [|Z_1(s) - Z_2(s)|^p]. \end{aligned}$$

Integrating on  $[0, T]$  we get  $(\mathbf{v}_3)$ .

Now let us show that  $(\mathbf{v}_3'') \Rightarrow (\mathbf{v}_3')$ . Fix  $\boldsymbol{\mu}^1, \boldsymbol{\mu}^2 \in \mathcal{W}_p(C([0, T]; \mathbb{R}^{2d}))$  and  $\gamma \in \Pi(\boldsymbol{\mu}^1, \boldsymbol{\mu}^2)$ . First consider the case  $p = 1$ . Observe that

$$\begin{aligned}
& \int_{\mathbb{R}^{2d} \times \mathbb{R}^{2d}} |v_1 - v_2|^{-1} (\mathbf{v}[t, \boldsymbol{\mu}^1](z_1) - \mathbf{v}[t, \boldsymbol{\mu}^2](z_2)) \cdot (v_1 - v_2) d\gamma_t(z_1, z_2) \\
& \leq \int_{\mathbb{R}^{2d} \times \mathbb{R}^{2d}} |\mathbf{v}[t, \boldsymbol{\mu}^1](z_1) - \mathbf{v}[t, \boldsymbol{\mu}^2](z_2)| d\gamma_t(z_1, z_2) \\
& \leq D \int_{\mathbb{R}^{2d} \times \mathbb{R}^{2d}} \sup_{0 \leq s \leq t} \mathcal{W}_1(\mu_s^1, \mu_s^2) d\gamma_t(z_1, z_2) + D \int_{\mathbb{R}^{2d} \times \mathbb{R}^{2d}} |z_1 - z_2| d\gamma_t(z_1, z_2) \\
& \leq D \sup_{0 \leq s \leq t} \mathcal{W}_1(\mu_s^1, \mu_s^2) + D \sup_{0 \leq s \leq t} \int_{\mathbb{R}^{2d} \times \mathbb{R}^{2d}} |z_1 - z_2| d\gamma_s(z_1, z_2) \\
& \leq 2D \sup_{0 \leq s \leq t} \int_{\mathbb{R}^{2d} \times \mathbb{R}^{2d}} |z_1 - z_2| d\gamma_s(z_1, z_2)
\end{aligned}$$

On the other hand, if  $p > 1$  we get

$$\begin{aligned}
& \int_{\mathbb{R}^{2d} \times \mathbb{R}^{2d}} |v_1 - v_2|^{p-2} (\mathbf{v}[t, \boldsymbol{\mu}^1](z_1) - \mathbf{v}[t, \boldsymbol{\mu}^2](z_2)) \cdot (v_1 - v_2) d\gamma_t(z_1, z_2) \\
& \leq \int_{\mathbb{R}^{2d} \times \mathbb{R}^{2d}} |v_1 - v_2|^{p-1} |\mathbf{v}[t, \boldsymbol{\mu}^1](z_1) - \mathbf{v}[t, \boldsymbol{\mu}^2](z_2)| d\gamma_t(z_1, z_2) \\
& \leq D \int_{\mathbb{R}^{2d} \times \mathbb{R}^{2d}} |v_1 - v_2|^{p-1} \sup_{0 \leq s \leq t} \mathcal{W}_p(\mu_s^1, \mu_s^2) d\gamma_t(z_1, z_2) \\
& \quad + D \int_{\mathbb{R}^{2d} \times \mathbb{R}^{2d}} |v_1 - v_2|^{p-1} |z_1 - z_2| d\gamma_t(z_1, z_2) \\
& \leq D \frac{p-1}{p} \int_{\mathbb{R}^{2d} \times \mathbb{R}^{2d}} |v_1 - v_2|^p d\gamma_t(z_1, z_2) + \frac{D}{p} \int_{\mathbb{R}^{2d} \times \mathbb{R}^{2d}} \sup_{0 \leq s \leq t} \mathcal{W}_p^p(\mu_s^1, \mu_s^2) d\gamma_t(z_1, z_2) \\
& \quad + D \int_{\mathbb{R}^{2d} \times \mathbb{R}^{2d}} |z_1 - z_2|^p d\gamma_t(z_1, z_2) \leq 2D \sup_{0 \leq s \leq t} \int_{\mathbb{R}^{2d} \times \mathbb{R}^{2d}} |z_1 - z_2|^p d\gamma_s(z_1, z_2).
\end{aligned}$$

This ends the proof.  $\square$



## APPENDIX B. PRELIMINARIES FOR THE WELL-POSEDNESS OF THE PDE-ODE SYSTEM

In this subsection, we collect some useful preliminary results to prove the well-posedness of the PDE-ODE system in (4.1). First of all, we consider an auxiliary system and we show the existence and uniqueness of its solution, alongside with some additional qualitative properties.

**Lemma B.1.** *Let Assumptions (v), (w), (F) and (A) hold. Let  $T > 0$ ,  $\mathbf{u} \in \mathcal{A}$  and  $\boldsymbol{\mu} \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$  be fixed, then the system*

$$\begin{cases} \dot{\mathbf{Y}}_{\boldsymbol{\mu}}(t) = \mathbf{W}_{\boldsymbol{\mu}}(t) = F[t, \boldsymbol{\mu}](\mathbf{Y}_{\boldsymbol{\mu}}) + \mathbf{u}(t, \boldsymbol{\mu}) & t \in (0, T) \\ \mu_0 = \bar{\mu}, \quad \mathbf{Y}_{\boldsymbol{\mu}}(0) = \bar{\mathbf{Y}}, \end{cases} \quad (\text{B.1})$$

admits a unique solution  $\mathbf{Y}_{\boldsymbol{\mu}}$ . Furthermore, setting  $\mathbf{H}_{\boldsymbol{\mu}} = (\mathbf{Y}_{\boldsymbol{\mu}}, \mathbf{W}_{\boldsymbol{\mu}})$ :

(i) *there exist three constants  $C_j$ ,  $j = 1, 2, 3$ , depending on  $p, K_F, K_{\mathbf{u}}, L_F, L_{\mathbf{u}}, T$  and  $\bar{\mathbf{H}}$  such that*

$$\sup_{0 \leq s \leq T} |\mathbf{Y}_{\boldsymbol{\mu}}(s)| \leq C_1(1 + \bar{M}_p(T; \boldsymbol{\mu})^{\frac{1}{p}}), \quad (\text{B.2})$$

$$\sup_{0 \leq s \leq T} |\mathbf{Y}_{\boldsymbol{\mu}}(s) - \mathbf{Y}_{\boldsymbol{\nu}}(s)| \leq C_2 \sup_{0 \leq s \leq T} \mathcal{W}_p(\mu_s, \nu_s), \quad (\text{B.3})$$

$$|\mathbf{Y}_{\boldsymbol{\mu}}(s) - \mathbf{Y}_{\boldsymbol{\mu}}(t)| \leq C_3(1 + \bar{M}_p(T; \boldsymbol{\mu})^p)|t - s|, \quad \forall t \in [0, T];$$

(ii) *if there exists a sequence  $\{\mathbf{u}_j\}_{j \in \mathbb{N}} \subset \mathcal{A}$  such that*

$$\lim_{j \rightarrow \infty} \int_0^t \mathbf{u}_j(s, \boldsymbol{\mu}) dt = \int_0^t \mathbf{u}(s, \boldsymbol{\mu}) dt, \quad \forall t \in [0, T],$$

then

$$\lim_{j \rightarrow \infty} \sup_{t \in [0, T]} |\mathbf{H}_{\boldsymbol{\mu}}^j(t) - \mathbf{H}_{\boldsymbol{\mu}}(t)| = 0, \quad (\text{B.4})$$

where  $\mathbf{H}_{\boldsymbol{\mu}}^j$  is the solution of (B.1) with  $\mathbf{u}_j$  in place of  $\mathbf{u}$ ;

(iii) *for any  $t \in [0, T]$ , if  $\boldsymbol{\mu}^1, \boldsymbol{\mu}^2 \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$  are such that  $\mu_s^1 = \mu_s^2$  for all  $s \in [0, t]$ , then  $\mathbf{H}_{\boldsymbol{\mu}^1}(t) = \mathbf{H}_{\boldsymbol{\mu}^2}(t)$ .*

*Proof.* First of all, existence and uniqueness of the solution of (B.1) is proved by means of the Picard iteration method, and for this reason we do not recall the proof here.

(i) Now, by (F<sub>1</sub>) and (4.2), it follows

$$\sup_{0 \leq s \leq t} |\mathbf{Y}_{\boldsymbol{\mu}}(s)| \leq C \left( 1 + (K_F + K_{\mathbf{u}})(1 + \bar{M}_p(T; \boldsymbol{\mu})^{\frac{1}{p}}) + K_F \int_0^t \sup_{0 \leq z \leq \tau} |\mathbf{Y}_{\boldsymbol{\mu}}(z)| d\tau \right)$$

and then, by Grönwall inequality,

$$\sup_{0 \leq s \leq t} |\mathbf{Y}_{\boldsymbol{\mu}}(s)| \leq C \left( 1 + (K_F + K_{\mathbf{u}})(1 + \bar{M}_p(T; \boldsymbol{\mu})^{\frac{1}{p}}) \right) e^{(K_F+1)T}.$$

leading to (B.2).

Next, let us consider any  $\boldsymbol{\mu}, \boldsymbol{\nu} \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$ . Then, by  $(F_2)$  and  $(A_2)$  we have

$$\sup_{0 \leq s \leq t} |\mathbf{Y}_{\boldsymbol{\mu}}(s) - \mathbf{Y}_{\boldsymbol{\nu}}(s)| \leq L_F \int_0^t \sup_{0 \leq r \leq \tau} |\mathbf{Y}_{\boldsymbol{\mu}}(r) - \mathbf{Y}_{\boldsymbol{\nu}}(r)| d\tau + (L_F + L_u) \int_0^t \sup_{0 \leq r \leq \tau} \mathcal{W}_p(\mu_r, \nu_r) d\tau,$$

Hence, once again, by Grönwall's inequality, we infer

$$\sup_{0 \leq s \leq t} |\mathbf{Y}_{\boldsymbol{\mu}}(s) - \mathbf{Y}_{\boldsymbol{\nu}}(s)| \leq T(L_F + L_u)e^{L_F T} \sup_{0 \leq s \leq t} \mathcal{W}_p(\mu_s, \nu_s).$$

that is (B.3).

Finally, we fix  $0 \leq t \leq s \leq T$  and observe that

$$\begin{aligned} |\mathbf{Y}_{\boldsymbol{\mu}}(t) - \mathbf{Y}_{\boldsymbol{\mu}}(s)| &\leq \int_s^t |F[\tau, \boldsymbol{\mu}](\mathbf{Y}_{\boldsymbol{\mu}})| d\tau + \int_s^t |\mathbf{u}(\tau, \boldsymbol{\mu})| d\tau \\ &\leq (K_F + K_u + C_1(K_F + 1))(1 + (\overline{M}_p(T, \boldsymbol{\mu}))^{\frac{1}{p}})|t - s|. \end{aligned}$$

(ii) By the assumptions of  $F$  and  $\mathbf{u}$ , it is sufficient to verify the uniform convergence on  $\mathbf{Y}$ . We consider a fixed  $\boldsymbol{\mu}$  and  $\mathbf{u}$ , and we assume there exists a sequence  $\{\mathbf{u}^j\}_{j \in \mathbb{N}}$  satisfying the assumptions. Furthermore, we denote by  $(\mathbf{Y}_{\boldsymbol{\mu}}^j, \mathbf{W}_{\boldsymbol{\mu}}^j)$  and  $(\mathbf{Y}_{\boldsymbol{\mu}}, \mathbf{W}_{\boldsymbol{\mu}})$  the solutions of the ODE (B.1) respectively with  $\mathbf{u}_j$  and  $\mathbf{u}$ . Then, for all  $0 \leq s \leq t \leq T$

$$|\mathbf{Y}_{\boldsymbol{\mu}}^j(s) - \mathbf{Y}_{\boldsymbol{\mu}}(s)| \leq L_F \int_0^s \sup_{0 \leq r \leq \tau} |\mathbf{Y}_{\boldsymbol{\mu}}^j(r) - \mathbf{Y}_{\boldsymbol{\mu}}(r)| d\tau + \left| \int_0^s (\mathbf{u}^j(\tau, \boldsymbol{\mu}) - \mathbf{u}(\tau, \boldsymbol{\mu})) ds \right|,$$

Now we take the supremum over  $[0, t]$  and we achieve

$$\sup_{0 \leq s \leq t} |\mathbf{Y}_{\boldsymbol{\mu}}^j(s) - \mathbf{Y}_{\boldsymbol{\mu}}(s)| \leq (L_F + 1) \int_0^t \sup_{0 \leq r \leq \tau} |\mathbf{Y}_{\boldsymbol{\mu}}^j(r) - \mathbf{Y}_{\boldsymbol{\mu}}(r)| d\tau + \mathcal{R}_j(t),$$

where

$$\mathcal{R}_j(t) := \sup_{0 \leq s \leq t} \left| \int_0^s (\mathbf{u}^j(s, \boldsymbol{\mu}) - \mathbf{u}(s, \boldsymbol{\mu})) ds \right|,$$

which, by Grönwall's inequality, implies

$$\sup_{0 \leq s \leq t} |\mathbf{Y}_{\boldsymbol{\mu}}^j(s) - \mathbf{Y}_{\boldsymbol{\mu}}(s)| \leq \mathcal{R}_j(t)e^{(L_F+1)T}.$$

We only need to prove that  $\lim_{j \rightarrow \infty} \mathcal{R}_j(T) = 0$ . To do this, consider the sequence of functions

$$\mathbf{J}_j(t) = \int_0^t (\mathbf{u}^j(s, \boldsymbol{\mu}) - \mathbf{u}(s, \boldsymbol{\mu})) ds.$$

By assumption, we have that for all  $t \in [0, T]$  it holds  $\lim_{j \rightarrow \infty} \mathbf{J}_j(t) = 0$ . Furthermore, it is clear that  $\mathcal{R}_j(T) = \sup_{0 \leq t \leq T} |\mathbf{J}_j(t)|$ , hence we want to prove that  $\mathbf{J}_j \rightarrow 0$  uniformly

in  $[0, T]$ . However, this is clear since the sequence  $\mathbf{J}_j$  is equi-Lipschitz. Indeed, for all  $0 \leq s \leq t \leq T$ ,

$$\begin{aligned} |\mathbf{J}_j(t) - \mathbf{J}_j(s)| &= \left| \int_s^t (\mathbf{u}^j(\tau, \boldsymbol{\mu}) - \mathbf{u}(\tau, \boldsymbol{\mu})) d\tau \right| \leq \int_s^t |\mathbf{u}^j(\tau, \boldsymbol{\mu}) - \mathbf{u}(\tau, \boldsymbol{\mu})| d\tau \\ &\leq 2K_u \left( 1 + (\overline{M}_p(t; \boldsymbol{\mu}))^{\frac{1}{p}} \right) |t - s|. \end{aligned}$$

This shows (B.4).

(iii) The last statement follows by uniqueness of the solution of (B.1), together with assumptions  $(F_2)$  and  $(A_2)$ , which in turn imply that if  $\mu_s^1 = \mu_s^2$  for all  $s \in [0, t]$ , then  $F[s, \boldsymbol{\mu}^1](y) = F[s, \boldsymbol{\mu}^2](y)$  and  $\mathbf{u}(s, \boldsymbol{\mu}^1) = \mathbf{u}(s, \boldsymbol{\mu}^2)$  for all  $s \in [0, t]$  and  $h \in \mathbb{R}^{2m}$ .  $\square$

Then, for reader's convenience, we recall the definition of the map  $\mathcal{S} : C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d})) \times \mathcal{A} \times [0, T] \rightarrow \mathbb{R}^{2m}$ :

$$\mathcal{S}[\boldsymbol{\mu}, \mathbf{u}](t) = \mathbf{H}_\mu(t),$$

where  $\mathbf{H}_\mu$  is the unique solution of (B.1). Furthermore, as already done in Section 4, for any fixed  $\mathbf{u} \in \mathcal{A}$ , we also define the map

$$G_{\mathbf{u}} : [0, T] \times C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d})) \times \mathbb{R}^{2d} \rightarrow \mathbb{R}^d$$

as follows:

$$G_{\mathbf{u}}[t, \boldsymbol{\mu}](z) = \mathbf{v}[t, \boldsymbol{\mu}](z) + \mathbf{w}[t, \mathcal{S}[\boldsymbol{\mu}, \mathbf{u}]](z).$$

Concerning  $G_{\mathbf{u}}$ , we can prove the following result.

**Lemma B.2.** Fix  $\mathbf{u} \in \mathcal{A}$ . Then  $G_{\mathbf{u}}$  satisfies Assumptions (v), where the constants are independent of the choice of  $\mathbf{u} \in \mathcal{A}$ .

*Proof.* Since  $\mathbf{v}$  already satisfies Assumptions (v), it is sufficient to verify that  $\mathbf{w}[\cdot, \mathcal{S}[\cdot, \mathbf{u}]](\cdot)$  satisfies the same assumptions. First, we notice that  $(\mathbf{v}_0)$  is satisfied. Indeed, by (B.3) and the assumptions on  $F$ , we know that  $\mathcal{S}[\cdot, \mathbf{u}]$  is continuous. Since by  $(\mathbf{w}_0)$  is a Carathéodory map, the composition  $\mathbf{w}[\cdot, \mathcal{S}[\cdot, \mathbf{u}]](\cdot)$  is also a Carathéodory map. Next, we prove  $(\mathbf{v}_1)$ . To do this, observe that by  $(\mathbf{w}_1)$

$$|\mathbf{w}[t, \mathcal{S}[\boldsymbol{\mu}, \mathbf{u}]](z)| \leq K_{\mathbf{w}}(1 + \sup_{0 \leq s \leq t} |\mathcal{S}[\boldsymbol{\mu}, \mathbf{u}](s)| + |x|^{\frac{\beta}{3}} + |v|^\beta),$$

and then, by (B.2) and the assumptions on  $F$  and  $\mathcal{A}$  we have

$$|\mathbf{w}[t, \mathcal{S}[\boldsymbol{\mu}, \mathbf{u}]](z)| \leq K_{\mathbf{w}}(1 + K_F)(1 + C_1 + K_F + K_u)(1 + \overline{M}_p(T; \boldsymbol{\mu})^{\frac{1}{p}} + |x|^{\frac{\beta}{3}} + |v|^\beta),$$

Assumption  $(\mathbf{v}_2)$  clearly follows by  $(\mathbf{w}_2)$ . Now we prove  $(\mathbf{v}_3)$ . To do this, let  $\boldsymbol{\mu}^1, \boldsymbol{\mu}^2 \in \mathcal{W}_p(C([0, T]; \mathbb{R}^{2d}))$ ,  $z_1, z_2 \in \mathbb{R}^{2d}$  and observe that, by  $(\mathbf{w}_2)$

$$\begin{aligned} |\mathbf{w}[t, \mathcal{S}[\boldsymbol{\mu}^1, \mathbf{u}]](z_1) - \mathbf{w}[t, \mathcal{S}[\boldsymbol{\mu}^2, \mathbf{u}]](z_2)| &\leq L_{\mathbf{w}} \left( \sup_{0 \leq s \leq t} |\mathcal{S}[\boldsymbol{\mu}^1, \mathbf{u}](s) - \mathcal{S}[\boldsymbol{\mu}^2, \mathbf{u}](s)| + |z_1 - z_2| \right) \\ &\leq L_{\mathbf{w}}(1 + L_F + L_u(1 + C_2)) \left( \sup_{0 \leq s \leq T} \mathcal{W}_p(\mu_s^1, \mu_s^2) + |z_1 - z_2| \right), \end{aligned}$$

where we also used (B.3) and the assumptions on  $F$  and  $\mathcal{A}$ . The latter inequality implies (v<sub>3</sub>) (see Appendix A for more details). Finally, (v<sub>4</sub>) follows once we observe that if  $\mu_s^1 = \mu_s^2$  for all  $s \in [0, t]$ , then  $\mathcal{S}[\boldsymbol{\mu}^1, \mathbf{u}](s) = \mathcal{S}[\boldsymbol{\mu}^2, \mathbf{u}](s)$  for all  $s \in [0, t]$  by Lemma B.1.  $\square$

The following lemma is a consequence of the Arzelá-Ascoli Theorem (see [27, Page 81]) and the characterization of compact subsets of  $\mathcal{W}_p(\mathbb{R}^{2d})$  (see [4, Proposition 7.1.5]). It will be useful in the setting of the control problem.

**Lemma B.3.**  $\mathcal{K}$  and  $\mathcal{K}_1$  are compact subsets respectively of  $C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$  and  $C([0, T]; \mathbb{R}^m)$ .

*Proof.* Concerning  $\mathcal{K}_1$ , we notice that the functions belonging to it are equibounded and equi-Lipschitz, hence we can apply the Arzelá-Ascoli Theorem. Concerning  $\mathcal{K}$ , let us first observe that the curves of probability measures belonging to it are equi-Hölder with respect to the Wasserstein distance, hence equi-continuous. Furthermore, it is clear that  $\mathcal{K}$  is closed. Hence, it is only necessary to show that for all  $t \in [0, T]$  the set

$$\mathcal{K}_t := \{\mu_t \in \mathcal{W}_p(\mathbb{R}^{2d}) : \boldsymbol{\mu} \in \mathcal{K}\}$$

is relatively compact. Notice that  $\mathcal{K}_t \subset \mathcal{W}_p(\mathbb{R}^{2d})$ . We now prove that  $\mathcal{K}_t$  is uniformly tight. Fix  $\varepsilon > 0$  and let  $\delta = \frac{2(\mathcal{C}(K_G))^{1/p}}{\varepsilon}$ . Then, by Markov's inequality, setting  $B_\delta := \{z \in \mathbb{R}^{2d} : |z| < \delta\}$ ,

$$\mu_t(\mathbb{R}^{2d} \setminus B_\delta) \leq \frac{M_1(\mu_t)}{\delta} \leq \frac{M_p(\mu_t)^{1/p}}{\delta} \leq \frac{(\mathcal{C}(K_G))^{1/p}}{\delta} < \varepsilon.$$

Since  $\delta$  is independent of the choice of  $\mu_t \in \mathcal{K}_t$ , this shows that  $\mathcal{K}_t$  is uniformly tight. Now we show that  $\mathcal{K}_t$  has uniformly integrable  $p$ -moments. To do this, we claim that there exists a probability measure  $\tilde{\mathbb{P}}$  on  $(\mathbb{R}^{2d}, \mathcal{B}(\mathbb{R}^{2d}))$  such that for all  $\mu_t \in \mathcal{K}_t$  there exists a random variable  $Z$  on the probability space  $(\mathbb{R}^{2d}, \mathcal{B}(\mathbb{R}^{2d}), \tilde{\mathbb{P}})$  with  $\text{Law}(Z) = \mu_t$ . Despite this property is well-known in measure theory, we provide here a proof for completeness. Recall that  $(\mathbb{R}^{2d}, \mathcal{B}(\mathbb{R}^{2d}))$  and  $([0, 1], \mathcal{B}([0, 1]))$  are uncountable Polish spaces, hence by Kuratowski's Isomorphism Theorem (see [25, Theorem 15.6]) we know that there exists a Borel isomorphism  $f : [0, 1] \rightarrow \mathbb{R}^{2d}$ . We equip  $([0, 1], \mathcal{B}([0, 1]))$  with the Lebesgue measures  $\text{Leb}$  and we define  $\tilde{\mathbb{P}} = f\#\text{Leb}$ , i.e. the pushforward of  $\text{Leb}$  on  $(\mathbb{R}^{2d}, \mathcal{B}(\mathbb{R}^{2d}))$  through the Borel isomorphism  $f$ . Furthermore, for any  $\mu_t \in \mathcal{K}_t$ , denote  $\tilde{\mu}_t = f^{-1}\#\mu_t$  the pushforward of  $\mu_t$  on  $([0, 1], \mathcal{B}([0, 1]))$  through the Borel isomorphism  $f^{-1}$ . Notice that both  $\tilde{\mathbb{P}}$  and  $\tilde{\mu}_t$  are probability measures. In particular, by Skorokhod's representation theorem (see [48, Section 3.12]) we know that there exists a random variable  $\tilde{Z}$  on  $([0, 1], \mathcal{B}([0, 1]), \text{Leb})$  such that  $\tilde{\mu}_t = \text{Law}(\tilde{Z})$ . We set  $Z = f \circ \tilde{Z}$ , which is measurable since it is composition of measurable functions. Furthermore, for any Borel set  $B \in \mathcal{B}(\mathbb{R}^{2d})$ , we have

$$\tilde{\mathbb{P}}(Z \in B) = \text{Leb}(\tilde{Z} \in f^{-1}(B)) = \tilde{\mu}_t(f^{-1}(B)) = \mu_t(B),$$

i.e.  $\text{Law}(Z) = \mu_t$ . Let

$$\tilde{\mathcal{K}}_t := \{Z \in \mathcal{M}(\mathbb{R}^{2d}, \mathcal{B}(\mathbb{R}^{2d}), \tilde{\mathbb{P}}) : \text{Law}(Z) \in \mathcal{K}_t\}.$$

Then, by the previous argument,  $\tilde{\mathcal{K}}_t \neq \emptyset$  and for all  $\mu_t \in \mathcal{K}_t$  there exists  $Z \in \tilde{\mathcal{K}}_t$  such that  $\text{Law}(Z) = \mu_t$ . Hence, to prove that  $\mathcal{K}_t$  has uniformly integrable  $p$ -moments, it is sufficient to show that  $\tilde{\mathcal{K}}_t$  is uniformly  $L^p$ -integrable, or, equivalently, that  $\{|Z|^p, Z \in \tilde{\mathcal{K}}_t\}$  is uniformly integrable. However, it is clear that, for any  $Z \in \tilde{\mathcal{K}}_t$  with  $\text{Law}(Z) = \mu_t$  for some  $\mu \in \mathcal{K}$ ,

$$\mathbb{E} \left[ \tilde{\Phi}_p(|Z|; K_G) \right] \leq \overline{M}_{\tilde{\Phi}_p(\cdot; K_G)}(T, \mu) \leq \mathcal{C}(K_G),$$

where the right-hand side of the previous inequality is independent of  $Z$ . Hence, by the de la Vallée-Poussin Theorem,  $\{|Z|^p, Z \in \tilde{\mathcal{K}}_t\}$  is uniformly integrable and then  $\mathcal{K}_t$  has uniformly integrable  $p$ -moments. By [4, Proposition 7.1.5], we know that  $\mathcal{K}_t$  is relatively compact, hence  $\mathcal{K}$  is equicontinuous, closed and pointwise relatively compact, which in turn implies, by the Arzelá-Ascoli theorem, that  $\mathcal{K}$  is compact.  $\square$

Finally, we observe that  $\mathcal{K}' := \mathcal{K} \times \mathcal{K}_1$  is a compact subset of  $C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d})) \times C([0, T]; \mathbb{R}^m)$ . Then we are able to describe the behaviour of the field  $G_u$  with respect to suitable variations of the control  $u \in \mathcal{A}$ .

**Lemma B.4.** *Let  $\{\mathbf{u}_j\}_{j \in \mathbb{N}} \subset \mathcal{A}$  and  $\mathbf{u} \in \mathcal{A}$  be such that*

$$\lim_{j \rightarrow \infty} \int_0^t \mathbf{u}_j(s, \mu) ds = \int_0^t \mathbf{u}(s, \mu) ds, \quad \forall t \in [0, T], \forall \mu \in \mathcal{K}.$$

*Then, for any  $t \in [0, T]$ ,  $\mu \in \mathcal{K}$  and  $z \in \mathbb{R}^{2d}$*

$$\lim_{j \rightarrow \infty} G_{\mathbf{u}_j}[t, \mu](z) = G_{\mathbf{u}}[t, \mu](z).$$

*Proof.* Fix  $\mu \in \mathcal{K}$ ,  $t \in [0, T]$  and  $z \in \mathbb{R}^{2d}$ . Then we have

$$\begin{aligned} |G_{\mathbf{u}_j}[t, \mu](z) - G_{\mathbf{u}}[t, \mu](z)| &= |\mathfrak{w}[t, \mathcal{S}[\mu, \mathbf{u}_j]](z) - \mathfrak{w}[t, \mathcal{S}[\mu, \mathbf{u}]](z)| \\ &\leq L_{\mathfrak{w}} \sup_{0 \leq s \leq t} |\mathcal{S}[\mu, \mathbf{u}_j](s) - \mathcal{S}[\mu, \mathbf{u}](s)|. \end{aligned}$$

Taking the limit as  $j \rightarrow \infty$ , we get the statement by Lemma B.1.  $\square$

## APPENDIX C. OTHER POSSIBLE CONTROL CLASSES

It could be possible, for modelling restrictions, that we can apply controls with some more restrictive assumptions than the ones considered for  $\mathcal{A}$ . Hence, in this section, we consider three possible subclasses of  $\mathcal{A}$ . Since the proofs of the existence of the solutions to the related optimal control problems differ from Theorem 1.5 only of some technical details, we will only underline the parts in which the proof are actually different.

**C.1. Controls satisfying the Oscillation Restriction Criterion.** First, let us assume that the controls cannot *oscillate too much*. Precisely, we consider the class of admissible controls  $\mathcal{A}_{\text{ORC}} \subset \mathcal{A}$  satisfying the additional condition

( $\mathcal{A}_{\text{ORC}}$ ) For all compact sets  $\mathcal{K} \subset C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$ , all measurable sets  $A \subseteq [0, T]$  and all  $\varepsilon > 0$ , there exists a measurable set  $B \subset A$  such that

$$\frac{1}{|B|} \int_B \sup_{\mu \in \mathcal{K}} \left| \mathbf{u}(t, \mu) - \frac{1}{|B|} \int_B \mathbf{u}(s, \mu) ds \right| dt < \varepsilon.$$

Assumption ( $\mathcal{A}_{\text{ORC}}$ ) can be recognized as a special form of the *Oscillation Reduction Criterion*, introduced in [22] for real-valued functions and then discussed in [24] for Bochner integrals. With this assumption, we can actually relax the requirements on the cost functional  $\mathcal{F}$ . Precisely, we assume that  $\mathcal{F}$  satisfies ( $\mathcal{F}_0$ ), ( $\mathcal{F}_1$ ), ( $\mathcal{F}_2$ ) and

( $\mathcal{F}_{\text{ORC}}$ )  $\Psi : \mathbb{R}^m \rightarrow \mathbb{R}$  is continuous.

We are interested in the following optimal control problem:

<b>Problem 2</b>
<p>Find <math>\mathbf{u}^* \in \mathcal{A}_{\text{ORC}}</math> such that</p> $\mathcal{F}[\mathbf{u}^*] = \min_{\mathbf{u} \in \mathcal{A}_{\text{ORC}}} \mathcal{F}[\mathbf{u}].$

**Theorem C.1.** Under ( $\mathcal{F}_0$ ), ( $\mathcal{F}_1$ ), ( $\mathcal{F}_2$ ) and ( $\mathcal{F}_{\text{ORC}}$ ), Problem 2 admits at least a solution.

*Proof.* We proceed as in the proof of Theorem 1.5 except that this time  $\mathbf{U}^n \rightarrow \mathbf{U}^*$  in measure by [10, Theorem 1.4]. Up to a non-relabelled subsequence, we can assume  $\mathbf{U}^n(t) \rightarrow \mathbf{U}^*(t)$  in  $C(\mathcal{K}; \mathbb{R}^m)$  for a.a.  $t \in [0, T]$ . Moreover, if we denote by  $(\mu^n, \mathbf{Y}^n)$  the solutions of (4.1) with controls  $\mathbf{u}^n$  and  $(\mu^*, \mathbf{Y}^*)$  the solution of (4.1) with control  $\mathbf{u}^*$ , we also know that, by Theorem 4.2,  $\mu^n \rightarrow \mu^*$  in  $C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$ . Hence  $\lim_{n \rightarrow \infty} \mathbf{U}^n(t)(\mu^n) = \mathbf{U}^*(t)(\mu^*)$  for a.a.  $t \in [0, T]$ . Furthermore, recall that  $|\Psi(\mathbf{U}^n(t)(\mu^n))| \leq M_\Psi$ , where  $M_\Psi$  is defined in the proof of Theorem 1.5. Hence, by dominated convergence, we get

$$\lim_{n \rightarrow \infty} \int_0^T \Psi(\mathbf{U}^n(t)(\mu^n)) dt = \int_0^T \Psi(\mathbf{U}^*(t)(\mu^*)) dt.$$

The remainder of the proof is exactly the same as the one of Theorem 1.5. □

**C.2. Controls with current-state dependence on  $\mu$ .** It could be the case that the control we want to apply cannot depend on the whole trajectory of  $\mu$ , but only on the current state. To model this case, we consider the class of admissible controls  $\mathcal{A}_{\text{EV}} \subset \mathcal{A}$  satisfying the additional condition:

( $\mathcal{A}_{\text{EV}}$ ) There exists a function  $\mathbf{u}^\sharp : [0, T] \times \mathcal{W}_p(\mathbb{R}^{2d}) \rightarrow \mathbb{R}^m$  such that  $\mathbf{u}(t, \mu) = \mathbf{u}^\sharp(t, \mu_t)$  for all  $\mu \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$  and

$$|\mathbf{u}^\sharp(t, \mu) - \mathbf{u}^\sharp(t, \nu)| \leq \frac{L_{\mathbf{u}}}{m} \mathcal{W}_p(\mu, \nu)$$

for all  $t \in [0, T]$  and  $\mu, \nu \in \mathcal{W}_p(\mathbb{R}^{2d})$ .

We consider then the optimal control problem:

<b>Problem 3</b>	
Find $\mathbf{u}^* \in \mathcal{A}_{\text{EV}}$ such that	$\mathcal{F}[\mathbf{u}^*] = \min_{\mathbf{u} \in \mathcal{A}_{\text{EV}}} \mathcal{F}[\mathbf{u}].$

**Theorem C.2.** Under  $(\mathcal{F}_0)$ ,  $(\mathcal{F}_1)$ ,  $(\mathcal{F}_2)$  and  $(\mathcal{F}_3)$ , Problem 3 admits at least a solution.

*Proof.* Recall that, arguing as in Theorem 1.5, if we consider a minimizing sequence  $\{\mathbf{u}^n\}_{n \in \mathbb{N}} \subset \mathcal{A}_{\text{EV}}$ , then there exists a function  $\mathbf{u}^* \in \mathcal{A}$  such that

$$\limsup_{n \rightarrow \infty} \frac{1}{|E|} \left| \int_E (\mathbf{u}^n(t, \mu) - \mathbf{u}^*(t, \mu)) \right| dt = 0.$$

Hence, to prove the statement, it is sufficient to show that  $\mathbf{u}^* \in \mathcal{A}_{\text{EV}}$ . To do this, consider  $\mu, \nu \in \mathcal{K}$  and observe that, by  $(\mathcal{A}_{\text{EV}})$ , for any  $j = 1, \dots, m$ ,  $t \in E_{\text{Leb}}$  and  $h > 0$  small enough,

$$\frac{1}{h} \left| \int_t^{t+h} (u_j^n(s, \mu) - u_j^*(s, \nu)) ds \right| \leq \frac{L_{\mathbf{u}}}{m} \frac{1}{h} \int_t^{t+h} \mathcal{W}_p(\mu_s, \nu_s) ds.$$

Taking the limit as  $n \rightarrow \infty$  and then as  $h \downarrow 0$ , we get

$$|u_j^*(t, \mu) - u_j^*(t, \nu)| \leq \frac{L_{\mathbf{u}}}{m} \mathcal{W}_p(\mu_t, \nu_t). \quad (\text{C.1})$$

This holds for all  $t \in [0, t]$  once we assume that  $u_j^*(t, \cdot) \equiv 0$  for  $t \notin E_{\text{Leb}}$ . Now observe that if  $\mu \in \mathcal{K}$ , then  $\mu_t \in \mathcal{K}^\sharp$ , where

$$\mathcal{K}^\sharp = \left\{ \mu \in \mathcal{W}_p(\mathbb{R}^{2d}) : M_p(\mu) + M_{\tilde{\Phi}(\cdot; K_G)}(\mu) \leq \mathcal{C}(K_G) \right\}.$$

On the other hand, if  $\mu \in \mathcal{K}^\sharp$ , then  $\mu^\sharp \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$  defined as  $\mu_t^\sharp \equiv \mu$  for all  $t \in [0, t]$  belongs to  $\mathcal{K}$ . Hence, we define  $\mathbf{u}^{\sharp, *} : \mathcal{K}^\sharp \rightarrow \mathbb{R}^m$  as follows

$$\mathbf{u}^{\sharp, *}(t, \mu) = \mathbf{u}^*(t, \mu^\sharp).$$

Then, for all  $\boldsymbol{\mu} \in \mathcal{K}$  and  $t \in [0, T]$ , by a simple application of (C.1), it must hold

$$\mathbf{u}^*(t, \boldsymbol{\mu}) = \mathbf{u}^{\sharp,*}(t, \mu_t). \quad (\text{C.2})$$

Using McShane's Extension Theorem as in the proof of Theorem 1.5, we extend  $\mathbf{u}^{\sharp,*}(t, \cdot)$  to the whole space  $\mathcal{W}_p(\mathbb{R}^{2d})$  and then we define  $\mathbf{u}^*(t, \cdot)$  on  $C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$  by means of (C.2). As a consequence,  $\mathbf{u}^* \in \mathcal{A}_{\text{EV}}$ , concluding the proof.  $\square$

**C.3. Controls with separated variables.** Among the controls in  $\mathcal{A}_{\text{EV}}$ , one could consider the ones whose dependence on  $t$  and  $\mu_t$  is *separated*. To do this, let  $\ell \in \mathbb{N}$  and fix three positive constants  $M_h$ ,  $M_g$  and  $L_g$ . We define the set  $\mathcal{A}'$  as the class of couples  $(\mathbf{h}, \mathbf{g})$  satisfying the following assumptions:

<b>Assumptions on <math>(\mathbf{h}, \mathbf{g}) \in \mathcal{A}'</math>: <math>(\mathcal{A}')</math></b>	
( $\mathcal{A}'_0$ )	$\mathbf{h} \in L^1([0, T]; \mathbb{R}^{m \times \ell})$ and $\mathbf{g} \in C(\mathcal{W}_p(\mathbb{R}^{2d}); \mathbb{R}^\ell)$ .
( $\mathcal{A}'_1$ )	For all $t \in [0, T]$ it holds $ \mathbf{g}(\delta_0)  \leq M_g,$ where $\delta_0 \in \mathcal{W}_p(\mathbb{R}^{2d})$ is the Dirac delta measure concentrated on $0 \in \mathbb{R}^{2d}$ .
( $\mathcal{A}'_2$ )	It holds $ g_j(\mu) - g_j(\nu)  \leq \frac{L_g}{m} \mathcal{W}_p(\mu, \nu),$ for all $\mu, \nu \in \mathcal{W}_p(\mathbb{R}^{2d})$ , $j = 1, \dots, m$ and $t \in [0, T]$ .
( $\mathcal{A}'_3$ )	For almost all $t \in [0, T]$ it holds $ \mathbf{h}(t)  \leq M_h,$ where $ \mathbf{h}(t) $ is the Frobenius norm of $\mathbf{h}(t)$ .

Now fix  $(\mathbf{h}, \mathbf{g}) \in \mathcal{A}'$  and consider the function  $\mathbf{u} : (t, \boldsymbol{\mu}) \mapsto \mathbf{h}(t)\mathbf{g}(\mu_t)$ . Clearly,  $\mathbf{u}$  is a Carathéodory map. Furthermore, since the Frobenius norm is sub-multiplicative, for any  $t \in [0, T]$

$$|\mathbf{u}(t, \delta_0)| \leq |\mathbf{h}(t)| |\mathbf{g}(\delta_0)| \leq M_h M_g,$$

while, denoting by  $\mathbf{h}_j$  the  $j$ -th row of  $\mathbf{h}$ , for all  $t \in [0, T]$  and  $\boldsymbol{\mu}, \boldsymbol{\nu} \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$ ,

$$|u_j(t, \boldsymbol{\mu}) - u_j(t, \boldsymbol{\nu})| \leq |\mathbf{h}_j(t)| |\mathbf{g}(\mu_t) - \mathbf{g}(\nu_t)| \leq M_h L_g \sup_{0 \leq t \leq T} \mathcal{W}_p(\mu_t, \nu_t),$$

hence for  $M_u = M_h M_g$  and  $L_u = m M_h L_g$ , we have that  $\mathbf{u} \in \mathcal{A}_{\text{EV}}$ . Setting  $M_u$  and  $L_u$  as declared, we can define the call of admissible controls  $\mathcal{A}_{\text{SV}} \subset \mathcal{A}_{\text{EV}}$  satisfying the additional condition

( $\mathcal{A}_{\text{SV}}$ ) There exists  $(\mathbf{h}, \mathbf{g}) \in \mathcal{A}'$  such that  $\mathbf{u}(t, \boldsymbol{\mu}) = \mathbf{h}(t)\mathbf{g}(\mu_t)$ .

We are interested in the following optimal control problem:



**Problem 4**Find  $\mathbf{u}^* \in \mathcal{A}_{\text{SV}}$  such that

$$\mathcal{F}[\mathbf{u}^*] = \min_{\mathbf{u} \in \mathcal{A}_{\text{SV}}} \mathcal{F}[\mathbf{u}].$$

Before proceeding with the proof of the existence of a solution for Problem 4, let us notice that, with a similar argument as in the proof of Theorem C.2, it is possible to reduce the problem on  $\mathcal{A}_{\text{SV}}$  to an analogous one on  $\mathcal{A}'$ . Indeed, consider the function  $\Psi' : \mathbb{R}^{m \times \ell} \times \mathbb{R}^\ell \rightarrow \mathbb{R}$  and the functional  $\mathcal{F}' : \mathcal{A}' \rightarrow \mathbb{R}$  defined as follows:

$$\Psi'(\mathbf{h}, \mathbf{g}) = \Psi(\mathbf{h}\mathbf{g}), \quad \forall (\mathbf{h}, \mathbf{g}) \in \mathbb{R}^{m \times \ell} \times \mathbb{R}^\ell \quad (\text{C.3})$$

and

$$\mathcal{F}'[\mathbf{h}, \mathbf{g}] = \int_0^T \mathcal{L}(t, \boldsymbol{\mu}, \mathbf{Y}) dt + \int_0^T \Psi'(\mathbf{h}(t), \mathbf{g}(\mu_t)) dt, \quad \forall (\mathbf{h}, \mathbf{g}) \in \mathcal{A}',$$

where  $(\boldsymbol{\mu}, \mathbf{Y})$  is the solution of (4.1) with control  $\mathbf{u}(t, \boldsymbol{\mu}) = \mathbf{h}(t)\mathbf{g}(\mu_t)$ . Notice that if  $\mathbf{u} \in \mathcal{A}_{\text{SV}}$ , then  $\mathbf{u}(t, \boldsymbol{\mu}) = \mathbf{h}(t)\mathbf{g}(\mu_t)$  for some  $(\mathbf{h}, \mathbf{g}) \in \mathcal{A}'$  and  $\mathcal{F}[\mathbf{u}] = \mathcal{F}'[\mathbf{h}, \mathbf{g}]$ . On the other hand, if  $(\mathbf{h}, \mathbf{g}) \in \mathcal{A}'$ , then  $\mathbf{u}(t, \boldsymbol{\mu}) = \mathbf{h}(t)\mathbf{g}(\mu_t)$  belongs to  $\mathcal{A}_{\text{SV}}$  and  $\mathcal{F}[\mathbf{u}] = \mathcal{F}'[\mathbf{h}, \mathbf{g}]$ . Hence, solving Problem 4 is equivalent to solving

**Problem 5**Find  $(\mathbf{h}^*, \mathbf{g}^*) \in \mathcal{A}'$  such that

$$\mathcal{F}'[\mathbf{h}^*, \mathbf{g}^*] = \min_{(\mathbf{h}, \mathbf{g}) \in \mathcal{A}'} \mathcal{F}'[\mathbf{u}].$$

Notice that if we start from Problem 4, then  $\Psi'$  must satisfy (C.3) in Problem 5. However, we could directly consider Problem 5 and be less restrictive on the dependence of  $\Psi'$  on the separated variables  $\mathbf{h}$  and  $\mathbf{g}$ . Indeed, we can assume that  $\mathcal{F}'$  satisfies the following assumptions:

**Assumptions on  $\mathcal{F}'$ : ( $\mathcal{F}'$ )**

( $\mathcal{F}'_0$ )  $\mathcal{L} : [0, T] \times C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d})) \times C([0, T]; \mathbb{R}^{2m})$  and  $\Psi' : \mathbb{R}^{m \times \ell} \times \mathbb{R}^\ell \rightarrow \mathbb{R}$  are measurable and bounded from below;

( $\mathcal{F}'_1$ )  $\mathcal{L}$  is continuous in the variables  $(\boldsymbol{\mu}, \mathbf{Y})$ .

( $\mathcal{F}'_2$ ) There exists a function  $M_{\mathcal{L}} \in L^1(0, T)$  such that for a.a.  $t \in [0, T]$  and all  $(\boldsymbol{\mu}, \mathbf{Y}) \in \mathcal{K}'$

$$|\mathcal{L}(t, \boldsymbol{\mu}, \mathbf{Y})| \leq M_{\mathcal{K}'}(t)$$

( $\mathcal{F}'_3$ )  $\Psi : \mathbb{R}^{m \times \ell} \times \mathbb{R}^\ell \rightarrow \mathbb{R}$  is convex in the first variable and continuous in the second one.

**Theorem C.3.** *Problem 5 admits at least a solution.*

*Proof.* Arguing as in Theorem 1.5, since for all  $\mu \in \mathcal{K}$  it holds  $\mu_t \in \mathcal{K}'$  for any  $t \in [0, T]$ , solving Problem 5 is equivalent to solving

**Problem 5\***

Find  $(\mathbf{h}^*, \mathbf{g}^*) \in \tilde{\mathcal{A}}'$  such that

$$\tilde{\mathcal{F}}'[\mathbf{h}^*, \mathbf{g}^*] = \min_{(\mathbf{h}, \mathbf{g}) \in \tilde{\mathcal{A}}'} \tilde{\mathcal{F}}'[\mathbf{h}, \mathbf{g}].$$

where  $\tilde{\mathcal{A}}' = \Pi\mathcal{A}'$ ,  $\Pi : C([0, T]; \mathbb{R}^{m \times \ell}) \times C(\mathcal{W}_p(\mathbb{R}^{2d}); \mathbb{R}^\ell) \rightarrow C([0, T]; \mathbb{R}^{m \times \ell}) \times C(\mathcal{K}'; \mathbb{R}^\ell)$  is the domain restriction map, i.e. for all  $\mu \in \mathcal{K}'$  and  $t \in [0, T]$  it holds  $\Pi(\mathbf{h}, \mathbf{g})(t, \mu) = (\mathbf{h}(t), \mathbf{g}(\mu))$ , and for all  $\Pi(\mathbf{h}, \mathbf{g}) \in \tilde{\mathcal{A}}'$  it holds

$$\tilde{\mathcal{F}}'[\Pi(\mathbf{h}, \mathbf{g})] = \mathcal{F}'[\mathbf{h}, \mathbf{g}].$$

Now consider a minimizing sequence  $\{(\mathbf{h}^n, \mathbf{g}^n)\}_{n \in \mathbb{N}} \subseteq \tilde{\mathcal{A}}'$ . Since  $|\mathbf{h}^n(t)| \leq M_h$  for all  $n \in \mathbb{N}$  and almost all  $t \in [0, T]$ , then by the Dunford-Pettis theorem (see [13, Theorem 4.30]) we know that there exists a function  $\mathbf{h}^* \in L^1([0, T]; \mathbb{R}^{m \times \ell})$  such that, up to a non-relabelled subsequence,  $\mathbf{h}^n \rightharpoonup \mathbf{h}^*$  in  $L^1([0, T]; \mathbb{R}^{m \times \ell})$  and it is clear, arguing again as in Theorem 1.5, that  $|\mathbf{h}^*(t)| \leq M_h$  for almost all  $t \in [0, T]$ . Next, notice that for all  $\mu \in \mathcal{K}'$

$$|\mathbf{g}^n(\mu)| \leq L_g(M_p(\mu))^{\frac{1}{p}} + M_g \leq L_g(\mathcal{C}(K_G))^{\frac{1}{p}} + M_g =: \tilde{M}_g.$$

Hence  $\{\mathbf{g}^n\}_{n \in \mathbb{N}} \subset C(\mathcal{K}'; \mathbb{R}^\ell)$  are equibounded and equiLipschitz (by  $(\mathcal{A}'_2)$ ). Hence, up to a non-relabelled subsequence, by the Arzelá-Ascoli theorem, there exists  $\mathbf{g}^* \in C(\mathcal{K}'; \mathbb{R}^\ell)$  such that  $\mathbf{g}^n \rightarrow \mathbf{g}^*$ . With the same arguments as in Theorem 1.5, one can use the McShane extension theorem to extend  $\mathbf{g}^*$  to a function in  $C(\mathcal{W}_p(\mathbb{R}^{2d}); \mathbb{R}^\ell)$  whose components are still  $\frac{L_g}{m}$ -Lipschitz, hence proving that  $(\mathbf{h}^*, \mathbf{g}^*) \in \tilde{\mathcal{A}}'$ . Furthermore, for all  $\mu \in \mathcal{K}$ ,  $t \in [0, T]$  and  $n \in \mathbb{N}$ , set  $\mathbf{u}^n(t, \mu) = \mathbf{h}(t)\mathbf{g}^n(\mu_t)$  and  $\mathbf{u}^*(t, \mu) = \mathbf{h}(t)\mathbf{g}^*(\mu_t)$ . Then, for all  $t \in [0, T]$  and  $\mu \in \mathcal{K}$  we have

$$\begin{aligned} & \left| \int_0^t \mathbf{u}^n(s, \mu) ds - \int_0^t \mathbf{u}^*(s, \mu) ds \right| \\ & \leq \int_0^t |\mathbf{h}^n(s)| |\mathbf{g}^n(\mu_s) - \mathbf{g}^*(\mu_s)| ds + \left| \int_0^t \mathbf{g}^*(\mu_s) (\mathbf{h}^n(s) - \mathbf{h}^*(s)) ds \right| \\ & \leq M_h T \sup_{\mu \in \mathcal{K}} |\mathbf{g}^n(\mu) - \mathbf{g}^*(\mu)| + \left| \int_0^t \mathbf{g}^*(\mu_s) (\mathbf{h}^n(s) - \mathbf{h}^*(s)) ds \right|. \end{aligned}$$

Hence, recalling that the function  $s \in [0, T] \mapsto \mathbf{g}^*(\mu_s)$  belongs to  $L^\infty([0, T]; \mathbb{R}^\ell)$ , we can take the limit to get

$$\lim_{n \rightarrow \infty} \left| \int_0^t \mathbf{u}^n(s, \mu) ds - \int_0^t \mathbf{u}^*(s, \mu) ds \right| = 0.$$

As a consequence, if we denote by  $(\boldsymbol{\mu}^n, \mathbf{Y}^n)$  the solutions of (4.1) with controls  $\mathbf{u}^n$  and  $(\boldsymbol{\mu}^*, \mathbf{Y}^*)$  the solution of (4.1) with control  $\mathbf{u}^*$ , it holds, by Theorem 4.2,  $(\boldsymbol{\mu}^n, \mathbf{Y}^n) \rightarrow (\boldsymbol{\mu}^*, \mathbf{Y}^*)$  in  $C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d})) \times C([0, T]; \mathbb{R}^m)$ . Furthermore, notice that

$$\begin{aligned} \int_0^T \Psi'(\mathbf{h}^n(t), \mathbf{g}^n(\mu_t^n)) dt &= \int_0^T (\Psi'(\mathbf{h}^n(t), \mathbf{g}^n(\mu_t^n)) - \Psi'(\mathbf{h}^n(t), \mathbf{g}^*(\mu_t^*))) dt \\ &\quad + \int_0^T \Psi'(\mathbf{h}^n(t), \mathbf{g}^*(\mu_t^*)) dt. \end{aligned}$$

To handle the first integral, we observe that

$$\int_0^T |\Psi'(\mathbf{h}^n(t), \mathbf{g}^n(\mu_t^n)) - \Psi'(\mathbf{h}^n(t), \mathbf{g}^*(\mu_t^*))| dt \leq \int_0^T \sup_{|\mathbf{h}| \leq M_h} |\Psi'(\mathbf{h}, \mathbf{g}^n(\mu_t^n)) - \Psi'(\mathbf{h}, \mathbf{g}^*(\mu_t^*))| dt$$

and recall that  $|\mathbf{g}^n(\mu_t^n)|, |\mathbf{g}^*(\mu_t^*)| \leq \widetilde{M}_g$ , so that, by dominated convergence,

$$\begin{aligned} &\lim_{n \rightarrow \infty} \int_0^T |\Psi'(\mathbf{h}^n(t), \mathbf{g}^n(\mu_t^n)) - \Psi'(\mathbf{h}^n(t), \mathbf{g}^*(\mu_t^*))| dt \\ &= \lim_{n \rightarrow \infty} \int_0^T \sup_{|\mathbf{h}| \leq M_h} |\Psi'(\mathbf{h}, \mathbf{g}^n(\mu_t^n)) - \Psi'(\mathbf{h}, \mathbf{g}^*(\mu_t^*))| dt = 0. \end{aligned} \quad (\text{C.4})$$

On the other hand, arguing exactly as in Theorem 4.2, we have

$$\int_0^T \Psi'(\mathbf{h}^*(t), \mathbf{g}^*(\mu_t^*)) dt \leq \liminf_{n \rightarrow \infty} \int_0^T \Psi'(\mathbf{h}^n(t), \mathbf{g}^*(\mu_t^*)) dt. \quad (\text{C.5})$$

Combining (C.4) and (C.5) we get

$$\int_0^T \Psi'(\mathbf{h}^*(t), \mathbf{g}^*(\mu_t^*)) dt \leq \liminf_{n \rightarrow \infty} \int_0^T \Psi'(\mathbf{h}^n(t), \mathbf{g}^*(\mu_t^*)) dt$$

and then we proceed as in the proof of Theorem 4.2.  $\square$

*Remark C.4.* Notice that these are the controls considered in [8] for the first order case, under more restrictive assumptions on the system (4.1).

## APPENDIX D. AN EXAMPLE OF APPLICATION OF THEOREM 1.5

In this appendix, we provide an example of second order mean field control problem of the form of Problem 1 arising from a second order multi-agent system. Precisely, let  $N, m \in \mathbb{N}$  be respectively the number of followers and leaders. Denote by  $(X_N^j, V_N^j)$  the position and velocity of the generic follower, for  $j = 1, \dots, N$ , and by  $(Y_N^i, W_N^i)$  the position and velocity of the generic leader, for  $i = 1, \dots, m$ . We assume the dynamics of the leader-follower system is described by the following set of *controlled* SDEs:

$$\left\{ \begin{array}{l} dX_N^j(t) = V_N^j(t)dt \\ dV_N^j(t) = \left( \frac{1}{N} \sum_{i=1}^N K_{1,1}(X_N^i(t) - X_N^j(t), V_N^i(t) - V_N^j(t)) \right. \\ \quad \left. + \frac{1}{m} \sum_{i=1}^m K_{1,2}(Y_N^i(t) - X_N^j(t), W_N^i(t) - V_N^j(t)) \right) dt \\ \quad + \sqrt{2\sigma} dB^j(t) \\ \dot{Y}_N^i(t) = W_N^i(t) = \frac{1}{N} \sum_{k=1}^N K_{2,1}(X_N^k(t) - Y_N^i(t)) \\ \quad + \frac{1}{m} \sum_{k=1}^m K_{2,2}(Y_N^k(t) - Y_N^i(t)) + u_N^i(t, \mathbf{X}_N(t), \mathbf{V}_N(t)) \\ X_N^j(0) = X_0^j, V_N^j(0) = V_0^j \\ Y_N^i(0) = Y_0^i \end{array} \right. \begin{array}{l} j = 1, \dots, N, \\ t \in [0, T] \\ j = 1, \dots, N, \\ t \in [0, T] \\ i = 1, \dots, m, \\ t \in [0, T] \\ j = 1, \dots, N \\ i = 1, \dots, m, \end{array} \quad (\text{D.1})$$

where

- $\mathbf{X}_N := (X_N^1, \dots, X_N^N)$ ,  $\mathbf{V}_N := (V_N^1, \dots, V_N^N)$  and  $\mathbf{Z}_N := (\mathbf{X}_N, \mathbf{V}_N)$ ;
- $K_{1,j} : \mathbb{R}^{2d} \rightarrow \mathbb{R}^d$  and  $K_{2,j} : \mathbb{R}^d \rightarrow \mathbb{R}^d$  are  $L_{\text{Ker}}$ -Lipschitz and  $M_{\text{Ker}}$ -bounded interaction kernels such that  $K_{i,j}(-x) = K_{i,j}(x)$ ;
- the initial data  $(X_0^j, V_0^j)_{j \in \mathbb{N}}$  are independent and identically distributed (i.i.d.) random variables belonging to  $\mathcal{W}_p(\mathbb{R}^{2d})$  for some  $p > 1$ ;
- for all  $i = 1, \dots, m$  there exists a function  $u^i : [0, T] \times \mathcal{W}_1(\mathbb{R}^{2d}) \rightarrow \mathbb{R}^d$  such that

$$u_N^i(t, \mathbf{x}, \mathbf{v}) = u^i \left( \frac{1}{N} \sum_{j=1}^N \delta_{x_j, v_j} \right),$$

where  $\delta_{x_j, v_j}$  are Dirac delta measures concentrated in  $(x_j, v_j) \in \mathbb{R}^{2d}$  for all  $j = 1, \dots, N$ . In particular,  $u^i$  is such that  $|u^i(t, \delta_0)| \leq M_u$  for every  $i = 1, \dots, m$  and

$$|u_k^i(t, \mu) - u_k^i(t, \nu)| \leq \frac{L_u}{md} \mathcal{W}_1(\mu, \nu),$$

for all  $k = 1, \dots, N, t \in [0, t]$  and  $\mu, \nu \in \mathcal{W}_1(\mathbb{R}^{2d})$ ;

- the initial data  $\{Y_0^i\}_{i=1, \dots, m} \subset \mathbb{R}^d$  are deterministic;

- $\mathbf{B} = (B^1(t), \dots, B^N(t))$  is a  $dN$ -dimensional standard Brownian motion.

Notice that by standard theory of SDEs (see [39, Theorem 5.2.1]) for all  $N \in \mathbb{N}$  there exists a unique strong solution  $(\mathbf{X}_N, \mathbf{V}_N, \mathbf{Y}_N, \mathbf{W}_N)$ , where  $\mathbf{X}_N = (X_N^j)_{j=1, \dots, N}$ ,  $\mathbf{V}_N = (V_N^j)_{j=1, \dots, N}$ ,  $\mathbf{Y}_N = (Y_N^j)_{j=1, \dots, m}$  and  $\mathbf{W}_N = (W_N^j)_{j=1, \dots, m}$ . We would like to send, in some sense,  $N \rightarrow \infty$ . Heuristically, we expect the dynamic of a single follower to converge towards the solution of the following McKean-Vlasov system of SDEs/ODEs:

$$\left\{ \begin{array}{l} d\bar{X}(t) = \bar{V}(t)dt \quad t \in [0, T] \\ d\bar{V}(t) = \left( (K_{1,1} * \mu_t)(\bar{X}(t), \bar{V}(t)) \right. \\ \quad \left. + \frac{1}{m} \sum_{i=1}^m K_{1,2}(Y_N^i(t) - \bar{X}(t), W_N^i(t) - \bar{V}(t)) \right) dt \\ \quad + \sqrt{2\sigma} dB(t) \quad t \in [0, T] \\ \dot{\bar{Y}}^j(t) = W_N^j(t) = (K_{2,1} * \mu_t)(\bar{Y}^j(t)) \\ \quad + \frac{1}{m} \sum_{i=1}^m K_{2,2}(\bar{Y}^i(t) - \bar{Y}^j(t)) + u^j(t, \mu_t) \quad \begin{array}{l} j = 1, \dots, m, \\ t \in [0, T] \end{array} \\ \bar{X}(0) = \bar{X}_0, \bar{V}(0) = \bar{V}_0 \\ \bar{Y}^j(0) = Y_0^j \quad j = 1, \dots, m, \\ \boldsymbol{\mu} = \text{Law}(\bar{X}, \bar{V}), \end{array} \right. \quad (\text{D.2})$$

where

$$\begin{aligned} (K_{1,1} * \mu_t)(x, v) &= \int_{\mathbb{R}^{2d}} K_{1,j}(\xi - x, \nu - v) d\mu_t(\xi, \nu), \\ (K_{2,1} * \mu_t)(y) &= \int_{\mathbb{R}^{2d}} K_{2,j}(\xi - y) d\mu_t(\xi, \nu). \end{aligned}$$

We set now  $\bar{Z} = (\bar{X}, \bar{V})$ ,  $\bar{\mathbf{W}}(t) = (\bar{W}^j(t))_{j=1, \dots, m}$ ,  $\bar{\mathbf{Y}}(t) = (\bar{Y}^j(t))_{j=1, \dots, m}$ ,  $\bar{\mathbf{H}} = (\bar{\mathbf{Y}}, \bar{\mathbf{W}})$ ,  $\mathbf{u}(t, \boldsymbol{\mu}) = (u^j(t, \mu_t))_{j=1, \dots, m}$ ,

$$\begin{aligned} F[t, \boldsymbol{\mu}](\bar{\mathbf{Y}}) &:= \left( (K_{2,1} * \mu_t)(\bar{Y}^j(t)) + \frac{1}{m} \sum_{i=1}^m K_{2,2}(\bar{Y}^i(t) - \bar{Y}^j(t)) \right)_{j=1, \dots, m}, \\ \mathbf{v}[t, \boldsymbol{\mu}](\bar{Z}) &:= (K_{1,1} * \mu_t)(\bar{X}(t), \bar{V}(t)), \\ \mathbf{w}[t, \bar{\mathbf{H}}](\bar{Z}) &:= \frac{1}{m} \sum_{i=1}^m K_{1,2}(Y_N^i(t) - \bar{X}(t), W_N^i(t) - \bar{V}(t)), \end{aligned}$$

and  $\mathbf{Y}_0 = (Y_0^j)_{j=1,\dots,m}$ . With this notation, we get exactly equation (4.4). Notice that, clearly,  $\mathbf{v}$  satisfies  $(\mathcal{B}_0)$  and  $(\mathcal{B}_4)$  and in particular

$$|\mathbf{v}[t, \boldsymbol{\mu}](z)| \leq \int_{\mathbb{R}^{2d}} |K_{1,1}(\xi - x, \nu - v)| d\mu_t(\xi, \nu) \leq M_{\text{Ker}}, \quad (\text{D.3})$$

that implies  $(\mathcal{B}_1)$ , while

$$\begin{aligned} |\mathbf{v}[t, \boldsymbol{\mu}](z_1) - \mathbf{v}[t, \boldsymbol{\mu}](z_2)| &\leq \int_{\mathbb{R}^{2d}} |K_{1,1}(\xi - x_1, \nu - v_1) - K_{1,1}(\xi - x_2, \nu - v_2)| d\mu_t(\xi, \nu) \\ &\leq L_{\text{Ker}}(|x_1 - x_2| + |v_1 - v_2|), \end{aligned} \quad (\text{D.4})$$

which, combined with (D.3), leads to  $(\mathcal{B}_2)$ . Furthermore, for any  $\boldsymbol{\mu}^1, \boldsymbol{\mu}^2 \in \mathcal{W}_1(C([0, T]; \mathbb{R}^{2d}))$  let  $\gamma \in \Pi(\boldsymbol{\mu}^1, \boldsymbol{\mu}^2)$  and denote by  $(Z_1, Z_2) \in L^1(\Omega; C([0, T]; \mathbb{R}^{4d}))$  a random variable such that  $\gamma = \text{Law}(Z_1, Z_2)$ . It holds

$$\begin{aligned} |\mathbf{v}[t, \boldsymbol{\mu}^1](z) - \mathbf{v}[t, \boldsymbol{\mu}^2](z)| &\leq \mathbb{E} [|K_{1,1}(X_1(t) - x, V_1(t) - v) - K_{1,1}(X_2(t) - x, V_2(t) - v)|] \\ &\leq L_{\text{Ker}} \mathbb{E} [|Z_1(t) - Z_2(t)|]. \end{aligned}$$

Since  $\gamma \in \Pi(\boldsymbol{\mu}^1, \boldsymbol{\mu}^2)$  is arbitrary, taking the infimum we get

$$|\mathbf{v}[t, \boldsymbol{\mu}^1](z) - \mathbf{v}[t, \boldsymbol{\mu}^2](z)| \leq L_{\text{Ker}} \mathcal{W}_1(\boldsymbol{\mu}^1, \boldsymbol{\mu}^2). \quad (\text{D.5})$$

Hence, (D.5) combined with (D.4) and Proposition A.1 shows that  $\mathbf{v}$  satisfies  $(\mathbf{v}_3)$ . It is also clear that  $\mathbf{v}$  satisfies Assumptions  $(\mathbf{w})$ . Concerning  $F$ , the fact that it satisfies Assumptions  $(F)$  follows in the same way as we did for  $\mathbf{v}$ . Finally, notice that the control  $\mathbf{u} \in \mathcal{A}_{\text{EV}}$  for some fixed constants  $M_{\mathbf{u}}$  and  $L_{\mathbf{u}}$ . Hence, by Theorem 4.2, there exist a unique solution  $(\bar{\boldsymbol{\mu}}, \bar{\mathbf{Y}})$  of (4.1) and  $\bar{\boldsymbol{\mu}} = \text{Law}(\bar{X}, \bar{V})$  where  $\bar{Z} = (\bar{X}, \bar{V})$  solves (D.2) together with  $\bar{\mathbf{Y}}$ . The link between the solutions of (D.1) and (D.2) is underlined by the following *propagation of chaos* result, whose proof is omitted since it is exactly the same as the one of [8, Theorem 3.9].

**Theorem D.1.** *Assume that  $\text{Law}(X_0^j, V_0^j) = \text{Law}(\bar{X}, \bar{V})$  for all  $j \in \mathbb{N}$  and  $(\bar{Z}^j)_{j \in \mathbb{N}} = (\bar{X}^j, \bar{Y}^j)_{j \in \mathbb{N}}$  be a sequence of i.i.d. copies of the solution  $\bar{Z} = (\bar{X}, \bar{Y})$  of (D.2). Then it holds*

$$\mathbb{E} \left[ \max_{1 \leq n \leq N} \max_{0 \leq t \leq T} |Z_N^n(t) - \bar{Z}^n(t)| + \max_{0 \leq t \leq T} |\mathbf{H}_N(t) - \bar{\mathbf{H}}(t)| \right] \leq P_{d,p}(N),$$

where  $Z_N^n = (X_N^n, V_N^n)$  for all  $n = 1, \dots, N$ ,  $H_N^n = (Y_N^n, W_N^n)$  for all  $n = 1, \dots, m$ ,  $\mathbf{H}_N = (H_N^n)_{n=1,\dots,m}$  and  $P_{d,p}$  is a function depending only on  $d$  and  $p$  such that  $\lim_{N \rightarrow \infty} P_{d,p}(N) = 0$ . Furthermore, it holds

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[ \sup_{0 \leq t \leq T} \mathcal{W}_1(\mu_t^N, \bar{\mu}_t) \right] = 0, \quad \text{where} \quad \mu_t^N := \frac{1}{N} \sum_{j=1}^N \delta_{(X_N^j(t), V_N^j(t))}. \quad (\text{D.6})$$

Hence, we also have, in (D.6), a *Law of Large Numbers* which guarantees that the empirical measure  $\mu^N$  of the followers in multiagent system (4.12) converges towards the

solution of mean field limit PDE-ODE system (4.1). Furthermore, with the notation introduced in Theorem D.1, we can rewrite (D.1) as follows:

$$\begin{cases} d\mathbf{X}_N(t) = \mathbf{V}_N(t)dt & t \in [0, T], \\ d\mathbf{V}_N(t) = (\mathbf{v}[t, \boldsymbol{\mu}^N](Z_N^j(t)) + \mathbf{w}[t, \mathbf{H}_N](Z_N^j(t)))_{j=1, \dots, N} dt \\ \quad + \sqrt{2\sigma}d\mathbf{B}(t) & t \in [0, T], \\ \dot{\mathbf{Y}}_N(t) = \mathbf{W}_N(t) = F[t, \boldsymbol{\mu}^N](\mathbf{H}_N) + \mathbf{u}(t, \boldsymbol{\mu}^N) & t \in [0, T], \\ \mathbf{X}_N(0) = (X_0^j)_{j=1, \dots, N}, \mathbf{V}_N(0) = (V_0^j)_{j=1, \dots, N}, \\ \mathbf{Y}_N(0) = \mathbf{Y}_0. \end{cases} \quad (\text{D.7})$$

Now assume that we wanted to control (D.7). Precisely, for any  $N \in \mathbb{N}$  we consider the functional  $\mathcal{F}^N : \mathcal{A}_{\text{EV}} \rightarrow \mathbb{R}$  defined as

$$\mathcal{F}^N[\mathbf{u}] = \mathbb{E} \left[ \int_0^T \mathcal{L}(t, \boldsymbol{\mu}^N, \mathbf{Y}^N) dt + \int_0^T \Psi(\mathbf{u}(t, \boldsymbol{\mu}^N)) \right]$$

under assumptions  $(\mathcal{F}_0)$ ,  $(\mathcal{F}_1)$ ,  $(\mathcal{F}_2)$  and  $(\mathcal{F}_3)$ . We are interested in the following optimal control problem:

<b>Problem MA<sub>N</sub></b>	
Find $\mathbf{u}^N \in \mathcal{A}_{\text{EV}}$ such that	$\mathcal{F}^N[\mathbf{u}^N] = \min_{\mathbf{u} \in \mathcal{A}_{\text{EV}}} \mathcal{F}^N[\mathbf{u}].$

To show that Problem MA<sub>N</sub> is well-posed, one needs some preliminaries. First of all, let us define

$$\mathcal{K}_n := \left\{ \boldsymbol{\mu} \in C([0, T]; \mathcal{W}_1(\mathbb{R}^{2d})) : \overline{M}_p(T; \boldsymbol{\mu}) + \sup_{\substack{0 \leq s, t \leq T \\ s \neq t}} \frac{\mathcal{W}_1(\mu_s, \mu_t)}{|t - s|^{\frac{1}{4}}} \leq n \right\}.$$

The fact that  $\mathcal{K}_n \subset C([0, T]; \mathcal{W}_1(\mathbb{R}^{2d}))$  is compact follows as in Lemma B.3. Then, we need the following stability result.

**Lemma D.2.** *Let  $\{\mathbf{u}^j\}_{j \in \mathbb{N}} \subset \mathcal{A}_{\text{EV}}$  and  $\mathbf{u}^* \in \mathcal{A}_{\text{EV}}$  such that*

$$\lim_{j \rightarrow \infty} \int_0^t \mathbf{u}^j(s, \boldsymbol{\mu}) ds = \int_0^t \mathbf{u}(s, \boldsymbol{\mu}) ds \quad \forall t \in [0, T], \quad \forall \boldsymbol{\mu} \in \bigcup_{n \in \mathbb{N}} \mathcal{K}_n.$$

Denote by  $\boldsymbol{\mu}^{N,j}$  and  $\boldsymbol{\mu}^{N,*}$  the empirical measures of the followers in (D.7) with controls respectively  $\mathbf{u}^j$  and  $\mathbf{u}^*$  and by  $\mathbf{H}_N^j$  and  $\mathbf{H}^*$  the dynamics of the respective leaders. Then

$$\lim_{j \rightarrow \infty} \mathbb{E} \left[ \sup_{0 \leq t \leq T} \mathcal{W}_1(\boldsymbol{\mu}_t^{N,j}, \boldsymbol{\mu}_t^{N,*}) + \sup_{0 \leq t \leq T} |\mathbf{H}_N^j(t) - \mathbf{H}^*(t)| \right] = 0. \quad (\text{D.8})$$

*Proof.* Arguing as in Lemma B.1, we have

$$\sup_{0 \leq s \leq t} |\mathbf{Y}_N^j(s) - \mathbf{Y}_N^*(s)| \leq \mathcal{R}^j(t) e^{(L_F+1)T}, \quad (\text{D.9})$$

where this time

$$\begin{aligned} \mathcal{R}^j(t) &:= \sup_{0 \leq s \leq t} \left| \int_0^s (\mathbf{u}^j(\tau, \boldsymbol{\mu}^{N,j}) - \mathbf{u}^*(\tau, \boldsymbol{\mu}^{N,*})) ds \right| \\ &\leq \int_0^t |\mathbf{u}^j(s, \boldsymbol{\mu}^{N,j}) - \mathbf{u}^j(s, \boldsymbol{\mu}^{N,*})| ds + \sup_{0 \leq s \leq t} \left| \int_0^s (\mathbf{u}^j(\tau, \boldsymbol{\mu}^{N,*}) - \mathbf{u}^*(\tau, \boldsymbol{\mu}^{N,*})) ds \right| \\ &\leq L_\mu \int_0^t \mathcal{W}_1(\boldsymbol{\mu}_s^{N,j}, \boldsymbol{\mu}_s^{N,*}) ds + \sup_{0 \leq s \leq t} \left| \int_0^s (\mathbf{u}^j(\tau, \boldsymbol{\mu}^{N,*}) - \mathbf{u}^*(\tau, \boldsymbol{\mu}^{N,*})) ds \right| \\ &=: L_\mu \int_0^t \mathcal{W}_1(\boldsymbol{\mu}_s^{N,j}, \boldsymbol{\mu}_s^{N,*}) ds + \mathcal{R}_1^j(t). \end{aligned} \quad (\text{D.10})$$

Furthermore, notice that, by using (w<sub>2</sub>), (D.4) and (D.5),

$$\begin{aligned} |\mathbf{Z}_N^j(t) - \mathbf{Z}_N^*(t)| &\leq C \left( \int_0^t \mathcal{W}_1(\boldsymbol{\mu}_s^{N,j}, \boldsymbol{\mu}_s^{N,*}) ds \right. \\ &\quad \left. + \int_0^t |\mathbf{Z}_N^j(s) - \mathbf{Z}_N^*(s)| ds + \sup_{0 \leq s \leq t} |\mathbf{H}_N^j(s) - \mathbf{H}_N^*(s)| \right) \end{aligned}$$

that, by Grönwall's inequality, (D.9) and (D.10), leads to

$$\mathcal{W}_1(\boldsymbol{\mu}_t^{N,j}, \boldsymbol{\mu}_t^{N,*}) \leq |\mathbf{Z}_N^j(t) - \mathbf{Z}_N^*(t)| \leq C \left( \int_0^t \mathcal{W}_1(\boldsymbol{\mu}_s^{N,j}, \boldsymbol{\mu}_s^{N,*}) ds + \mathcal{R}_1^j(t) \right).$$

Using again Grönwall's inequality, combining it with (D.9) and taking the supremum and the expectation, we finally achieve

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} \mathcal{W}_1(\boldsymbol{\mu}_t^{N,j}, \boldsymbol{\mu}_t^{N,*}) + \sup_{0 \leq t \leq T} |\mathbf{Y}_N^j(t) - \mathbf{Y}_N^*(t)| \right] \leq C \mathbb{E} [\mathcal{R}_1^j(T)].$$

Once we show that  $\boldsymbol{\mu}^{N,*}$  is Hölder-continuous, the fact that  $\lim_{j \rightarrow \infty} \mathcal{R}_1^j(T) = 0$  follows by the same arguments as in Theorem 4.2. On the other hand, by (4.2) and Proposition 2.3, we know that we can use the dominated convergence theorem to deduce

$$\lim_{j \rightarrow \infty} \mathbb{E} \left[ \sup_{0 \leq t \leq T} \mathcal{W}_1(\boldsymbol{\mu}_t^{N,j}, \boldsymbol{\mu}_t^{N,*}) + \sup_{0 \leq t \leq T} |\mathbf{Y}_N^j(t) - \mathbf{Y}_N^*(t)| \right] = 0.$$

Finally, (D.8) follows by observing that

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} |\mathbf{W}_N^j(t) - \mathbf{W}_N^*(t)| \right] \leq C \mathbb{E} \left[ \sup_{0 \leq t \leq T} \mathcal{W}_1(\boldsymbol{\mu}_t^{N,j}, \boldsymbol{\mu}_t^{N,*}) + \sup_{0 \leq t \leq T} |\mathbf{Y}_N^j(t) - \mathbf{Y}_N^*(t)| \right]$$

for some constant  $C > 0$ .



Hence, it only remains to prove the Hölder continuity of  $\mu^{N,*}$ . To do this, let  $0 \leq s \leq t \leq T$  and recall that

$$|\mathbf{V}_N^*(t)| \leq |\mathbf{V}_N(0)| + 2M_{\text{Ker}}T + \sqrt{2\sigma}M_B(T), \quad \text{where} \quad M_B(t) = \sup_{0 \leq s \leq t} |B(s)|.$$

and

$$|\mathbf{V}_N^*(t) - \mathbf{V}_N^*(s)| \leq 2M_{\text{Ker}}|t-s| + \sqrt{2\sigma}L_B|t-s|^{\frac{1}{4}}, \quad \text{where} \quad L_B := \sup_{0 \leq s < t \leq T} \frac{|B(t) - B(s)|}{|t-s|^\gamma},$$

so that

$$|\mathbf{Z}_N^*(t) - \mathbf{Z}_N^*(s)| \leq (4M_{\text{Ker}} + \sqrt{2\sigma}M_B(T))|t-s| + \sqrt{2\sigma}L_B|t-s|^\gamma.$$

Taking the expectation and recalling that  $E[M_B(T)], E[L_B] < \infty$  we get that  $\mu^{N,*} \in \bigcup_{n \in \mathbb{N}} \mathcal{K}_n$  and we conclude the proof.  $\square$

Now we are ready to prove the existence of a solution for Problem  $\text{MA}_N$ . Since the proof is really similar to Theorem C.2, we only underline the main differences.

**Theorem D.3.** *Under  $(\mathcal{F}_0), (\mathcal{F}_1), (\mathcal{F}_2)$  and  $(\mathcal{F}_3)$ , Problem  $\text{MA}_N$  admits at least a solution.*

*Proof.* First, we recall that  $\mathcal{W}_p(\mathbb{R}^{2d})$  for  $p > 1$  is a hemicompact dense subset of  $\mathcal{W}_1(\mathbb{R}^{2d})$ . Precisely, for all  $n \in \mathbb{N}$  define

$$\mathcal{K}_n^\# := \{\mu \in \mathcal{W}_1(\mathbb{R}^{2d}) : M_p(\mu) \leq n\}$$

and observe that  $\bigcup_{n \in \mathbb{N}} \mathcal{K}_n^\# = \mathcal{W}_p(\mathbb{R}^{2d})$ . The fact that  $\mathcal{K}_n^\#$  are compact subsets of  $\mathcal{W}_1(\mathbb{R}^{2d})$  is shown as in Lemma B.3. Consider a minimizing sequence  $\{\mathbf{u}^j\}_{j \in \mathbb{N}} \subset \mathcal{A}_{\text{EV}}$  for  $\mathcal{F}^N$ . Arguing as in Theorem C.2 and using a diagonal argument we know that there exists a function  $\mathbf{u}^* : [0, T] \times \bigcup_{n \in \mathbb{N}} \mathcal{K}_n \rightarrow \mathbb{R}^m$  such that, up to a non-relabelled subsequence, for all  $\mu \in \bigcup_{n \in \mathbb{N}} \mathcal{K}_n$

$$\lim_{n \rightarrow \infty} \frac{1}{|E|} \left| \int_E (\mathbf{u}^n(t, \mu) - \mathbf{u}^*(t, \mu)) dt \right| = 0 \quad (\text{D.11})$$

and, for all  $j = 1, \dots, md, t \in [0, T]$  and  $\mu, \nu \in \bigcup_{n \in \mathbb{N}} \mathcal{K}_n$ ,

$$|u_j^*(t, \mu) - u_j^*(t, \nu)| \leq \frac{L_u}{md} \mathcal{W}_1(\mu_t, \nu_t).$$

The fact that  $\mathbf{u}^* \in (\mathcal{A}_{\text{EV}})$  follows as in Theorem C.2. Furthermore, by (D.11), denoting by  $(\mathbf{Z}^{N,j}, \mathbf{Y}_N^j)$  and  $(\mathbf{Z}^{N,*}, \mathbf{Y}_N^*)$  the solutions of (D.7) with controls  $\mathbf{u}^j$  and  $\mathbf{u}^*$  respectively and by  $\mu^{N,j}$  and  $\mu^{N,*}$  the corresponding empirical measures of the followers, we have (D.8) and then, up to a subsequence, almost surely

$$\lim_{j \rightarrow \infty} \sup_{0 \leq t \leq T} \mathcal{W}_1(\mu_t^{N,j}, \mu_t^{N,*}) + \sup_{0 \leq t \leq T} |\mathbf{H}_N^j(t) - \mathbf{H}_N^*(t)| = 0.$$

Hence, by  $(\mathcal{F}_1), (\mathcal{F}_2)$  and the dominated convergence theorem, we have

$$\lim_{j \rightarrow +\infty} \mathbb{E} \left[ \int_0^T \mathcal{L}(t, \mu^{N,j}, \mathbf{Y}_N^j) dt \right] = \mathbb{E} \left[ \int_0^T \mathcal{L}(t, \mu^{N,*}, \mathbf{Y}_N^*) dt \right].$$

On the other hand, the same arguments as in Theorem 1.5 and a simple application of Fatou's Lemma lead to

$$\mathbb{E} \left[ \int_0^T \Psi(\mathbf{u}^{N,*}(t), \boldsymbol{\mu}^{N,*}) dt \right] \leq \liminf_{j \rightarrow \infty} \mathbb{E} \left[ \int_0^T \Psi(\mathbf{u}^{N,j}(t), \boldsymbol{\mu}^{N,j}) dt \right].$$

The remainder of the proof follows as in Theorem 1.5.  $\square$

However, it is really difficult to provide (even numerically) a solution of Problem  $\text{MA}_N$  for a big value of  $N$ . For such a reason, one tries to use a mean field limit approach, justified by Theorem D.1, to provide a suitable approximation of the optimal control. Indeed, on (4.4), one can consider the optimal control Problem 3, whose solution exists by Theorem C.2 and is much more affordable, at least from the numerical point of view, since it involves a system of a PDE and  $2md$  ODEs, as in (4.1), in place of a system of  $2(N+m)d$  SDEs as in (D.7). However, it remains to show that a solution of Problem 3 actually approximates, in some sense, a solution of Problem  $\text{MA}_N$ . We will do this for a slightly different problem. First we consider the following assumptions, that are more restrictive than the ones we considered in Section 4.2.

**Assumptions on  $\mathcal{F}^N$ : ( $\mathcal{F}^N$ )**

( $\mathcal{F}_0$ )  $\mathcal{L} : [0, T] \times C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d})) \times C([0, T]; \mathbb{R}^m)$  and  $\Psi : \mathbb{R}^m \rightarrow \mathbb{R}$  are measurable and bounded from below;

( $\mathcal{F}_1^N$ ) There exists a non-negative increasing concave function  $\varpi_{\mathcal{L}} : \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+$  such that

$$|\mathcal{L}(t, \boldsymbol{\mu}^1, \mathbf{Y}^1) - \mathcal{L}(t, \boldsymbol{\mu}^2, \mathbf{Y}^2)| \leq \varpi_{\mathcal{L}} \left( \sup_{0 \leq s \leq t} \mathcal{W}_1(\boldsymbol{\mu}_s^1, \boldsymbol{\mu}_s^2) + \sup_{0 \leq s \leq t} |\mathbf{Y}^1(s) - \mathbf{Y}^2(s)| \right)$$

( $\mathcal{F}_2$ ) There exists a function  $M'_{\mathcal{L}} \in L^1(0, T)$  such that for a.a.  $t \in [0, T]$ , for all  $\boldsymbol{\mu} \in C([0, T]; \mathcal{W}_p(\mathbb{R}^{2d}))$  and for all  $\mathbf{Y} \in C([0, T]; \mathbb{R}^m)$ ,

$$|\mathcal{L}(t, \boldsymbol{\mu}, \mathbf{Y})| \leq M'_{\mathcal{L}}(t).$$

( $\mathcal{F}_3$ )  $\Psi : \mathbb{R}^m \rightarrow \mathbb{R}$  is convex.

Furthermore, we have to further reduce the set of controls we are working on. Precisely, we define the class of controls  $\mathcal{A}_{\text{EVB}} \subset \mathcal{A}_{\text{EV}}$  satisfying ( $\mathcal{A}_0$ ), ( $\mathcal{A}_2$ ), ( $\mathcal{A}_{\text{EV}}$ ) and

( $\mathcal{A}_B$ ) For all  $t \in [0, T]$  and  $\boldsymbol{\mu} \in C([0, T]; \mathcal{W}_1(\mathbb{R}^{2d}))$  it holds

$$|\mathbf{u}(t, \boldsymbol{\mu})| \leq M_u.$$

Assumption ( $\mathcal{A}_B$ ) clearly implies ( $\mathcal{A}_1$ ), but it is much more restrictive. We consider then the following optimal control problems:

**Problem 3<sub>B</sub>**Find  $\mathbf{u}^* \in \mathcal{A}_{\text{EVB}}$  such that

$$\mathcal{F}[\mathbf{u}^*] = \min_{\mathbf{u} \in \mathcal{A}_{\text{EVB}}} \mathcal{F}[\mathbf{u}].$$

**Problem MAB<sub>N</sub>**Find  $\mathbf{u}^* \in \mathcal{A}_{\text{EVB}}$  such that

$$\mathcal{F}^N[\mathbf{u}^*] = \min_{\mathbf{u} \in \mathcal{A}_{\text{EVB}}} \mathcal{F}^N[\mathbf{u}].$$

Notice that in both cases the optimal solution  $\mathbf{u}^*$  is obtained by considering a minimizing sequence  $\{\mathbf{u}^j\}_{j \in \mathbb{N}} \subset \mathcal{A}_{\text{EVB}}$  such that for any measurable  $E \subset [0, T]$  and  $\boldsymbol{\mu} \in \mathcal{K}$  in case of Problem 3<sub>B</sub> or  $\boldsymbol{\mu} \in \bigcup_{n \in \mathbb{N}} \mathcal{K}_n$  in case of Problem MAB<sub>N</sub>

$$\lim_{j \rightarrow \infty} \frac{1}{|E|} \int_E \mathbf{u}^j(t, \boldsymbol{\mu}) dt = \frac{1}{|E|} \int_E \mathbf{u}^*(t, \boldsymbol{\mu}) dt.$$

Theorem C.2 already guarantees that  $\mathbf{u}^* \in \mathcal{A}_{\text{EV}}$ . Furthermore, for all  $t \in E_{\text{Leb}}$  and  $h > 0$  small enough

$$\frac{1}{h} \left| \int_t^{t+h} \mathbf{u}^j(t, \boldsymbol{\mu}) dt \right| \leq M_{\mathbf{u}}.$$

Hence, taking the limit as  $j \rightarrow \infty$  and then  $h \downarrow 0$ , one show that  $\mathbf{u}^* \in \mathcal{A}_{\text{EVB}}$ . As a consequence, we have the following result:

**Theorem D.4.** *There exists at least a solution for both Problems 3<sub>B</sub> and MAB<sub>N</sub>.*

Now, we can prove the following result:

**Theorem D.5.** *Let  $\{\mathbf{u}^N\}_{N \in \mathbb{N}} \subset \mathcal{A}_{\text{EVB}}$  and  $\mathbf{u}^* \in \mathcal{A}_{\text{EVB}}$  be such that for all  $t \in [0, T]$  and  $\boldsymbol{\mu} \in \bigcup_{n \in \mathbb{N}} \mathcal{K}_n$*

$$\lim_{N \rightarrow \infty} \int_0^t \mathbf{u}^N(s, \boldsymbol{\mu}) ds = \int_0^t \mathbf{u}^*(s, \boldsymbol{\mu}) ds.$$

Then

$$\mathcal{F}[\mathbf{u}^*] \leq \liminf_{N \rightarrow \infty} \mathcal{F}^N[\mathbf{u}^N].$$

Furthermore for all  $\mathbf{u} \in \mathcal{A}_{\text{EVB}}$  it holds

$$\lim_{N \rightarrow \infty} \mathcal{F}^N[\mathbf{u}] = \mathcal{F}[\mathbf{u}].$$

Finally, it holds

$$\lim_{N \rightarrow \infty} \min_{\mathbf{u} \in \mathcal{A}_{\text{EVB}}} \mathcal{F}^N[\mathbf{u}] = \min_{\mathbf{u} \in \mathcal{A}_{\text{EVB}}} \mathcal{F}[\mathbf{u}].$$

*Proof.* Let  $\{\mathbf{u}^N\}_{N \in \mathbb{N}} \subset \mathcal{A}_{\text{EVB}}$  and  $\mathbf{u}^* \in \mathcal{A}_{\text{EVB}}$  as in the statement. Denote by  $(\mathbf{Z}^N, \mathbf{H}_N)$  the solutions of (D.7) with controls  $\mathbf{u}^N$ , by  $\boldsymbol{\mu}^N$  the empirical measures of the followers  $\mathbf{Z}^N$  and by  $(\bar{\boldsymbol{\mu}}^N, \bar{\mathbf{Y}}^N)$  the solutions of (4.1) with controls  $\mathbf{u}^N$ . Arguing as in the proof of [8, Theorem 4.6] and using Theorem D.1, we have

$$\lim_{N \rightarrow \infty} \left| \mathbb{E} \left[ \int_0^T \mathcal{L}(t, \boldsymbol{\mu}^N, \mathbf{Y}_N) dt \right] - \int_0^T \mathcal{L}(t, \bar{\boldsymbol{\mu}}^N, \bar{\mathbf{Y}}_N) dt \right| = 0. \quad (\text{D.12})$$

On the other hand, by Theorem 4.2 and Assumptions  $(\mathcal{F}_1^N)$  (that implies  $(\mathcal{F}_1)$ ) and  $(\mathcal{F}_2)$ , we have

$$\lim_{N \rightarrow \infty} \left| \int_0^T \mathcal{L}(t, \bar{\boldsymbol{\mu}}^N, \bar{\mathbf{Y}}_N) dt - \int_0^T \mathcal{L}(t, \bar{\boldsymbol{\mu}}^*, \bar{\mathbf{Y}}_*) dt \right| = 0, \quad (\text{D.13})$$

where  $(\bar{\boldsymbol{\mu}}^*, \bar{\mathbf{Y}}_*)$  is the solution of (4.1) with control  $\mathbf{u}^*$ . Concerning  $\Psi$ , the same arguments as in the proof of Theorem 1.5, together with Theorem D.1, and Fatou's lemma lead to

$$\int_0^T \Psi(\mathbf{u}^*(t, \boldsymbol{\mu}^*)) dt \leq \liminf_{N \rightarrow \infty} \mathbb{E} \left[ \int_0^T \Psi(\mathbf{u}^N(t, \boldsymbol{\mu}^N)) dt \right].$$

This, together with (D.12) and (D.13) prove that

$$\mathcal{F}[\mathbf{u}^*] \leq \liminf_{N \rightarrow \infty} \mathcal{F}^N[\mathbf{u}^N].$$

Now consider a control  $\mathbf{u} \in \mathcal{A}_{\text{EVB}}$ , denote by  $(\mathbf{Z}^N, \mathbf{Y}_N)$  the solutions of (D.7) with controls  $\mathbf{u}$ , by  $\boldsymbol{\mu}^N$  the empirical measures of the followers  $\mathbf{Z}^N$  and by  $(\bar{\boldsymbol{\mu}}, \bar{\mathbf{Y}})$  the solutions of (4.1) with controls  $\mathbf{u}$ . Since  $\Psi$  is uniformly continuous on  $\{z \in \mathbb{R}^m : |z| \leq M_{\mathbf{u}}\}$ , we can consider a concave modulus of continuity  $\varpi_2$  and argue as in [8, Theorem 4.6], by using also  $(\mathcal{A}_2)$  and Theorem D.1,

$$\lim_{N \rightarrow \infty} \left| \mathbb{E} \left[ \int_0^T \Psi(\mathbf{u}(t, \boldsymbol{\mu}^N)) dt \right] - \int_0^T \Psi(\mathbf{u}(t, \bar{\boldsymbol{\mu}})) dt \right| = 0. \quad (\text{D.14})$$

Combining (D.12) and (D.14) we achieve

$$\lim_{N \rightarrow \infty} \mathcal{F}^N[\mathbf{u}] = \mathcal{F}[\mathbf{u}].$$

Finally, the last statement follows as in [8, Proposition 4.7].  $\square$

*Remark D.6.* It is worth noticing that the solution  $\mathbf{u}^*$  of Problem  $3_B$  is an *approximate solution* of Problem  $\text{MAB}_N$  in the sense that it can be used to provide a control whose cost is *nearly minimal*. However, if  $\mathbf{u}^N$  is a solution of Problem  $\text{MAB}_N$ , one cannot guarantee that  $\mathbf{u}^*$  approximate  $\mathbf{u}^N$ , but only that  $\mathcal{F}^N[\mathbf{u}^*]$  is near  $\mathcal{F}^N[\mathbf{u}^N]$ .

## REFERENCES

- [1] G. Albi, M. Bongini, E. Cristiani, and D. Kalise. Invisible control of self-organizing agents leaving unknown environments. *SIAM Journal on Applied Mathematics*, 76(4):1683–1710, 2016.
- [2] G. Albi, Y.-P. Choi, M. Fornasier, and D. Kalise. Mean field control hierarchy. *Applied Mathematics & Optimization*, 76(1):93–135, 2017.
- [3] G. Albi, E. Cristiani, L. Pareschi, and D. Peri. Mathematical models and methods for crowd dynamics control. *Crowd Dynamics, Volume 2: Theory, Models, and Applications*, pages 159–197, 2020.
- [4] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2005.
- [5] F. Anceschi, G. Ascione, D. Castorina, and F. Solombrino. Well-posedness of Kolmogorov-Fokker-Planck equations with unbounded drift. *Journal of Mathematical Analysis and Applications*, 543(2):Article number 128909, 2025.
- [6] W. Arendt, C. J. K. Batty, M. Hieber, and F. Neubrander. *Vector-valued Laplace Transforms and Cauchy Problems*, volume 96. Springer Science & Business Media, 2011.
- [7] D. Armbruster, S. Martin, and A. Thatcher. Elastic and inelastic collisions of swarms. *Physica D: Nonlinear Phenomena*, 344:45–57, 2017.
- [8] G. Ascione, D. Castorina, and F. Solombrino. Mean-field sparse optimal control of systems with additive white noise. *SIAM Journal on Mathematical Analysis*, 55(6):6965 – 6990, 2023.
- [9] F. Auletta, D. Fiore, M. J. Richardson, and M. di Bernardo. Herding stochastic autonomous agents via local control rules and online target selection strategies. *Autonomous Robots*, 46(3):469–481, 2022.
- [10] E. Balder. An extension of the Dunford-Pettis theorem with applications to a Bocce-type oscillation restriction criterion. *Rijksuniversiteit Utrecht. Mathematisch Instituut : preprint*, 1021, 1997.
- [11] P. Billingsley. *Convergence of Probability Measures*. John Wiley & Sons, 2013.
- [12] A. Braides. *Approximation of free-discontinuity problems*. Number 1694. Springer Science & Business Media, 1998.
- [13] H. Brézis. *Functional analysis, Sobolev spaces and partial differential equations*, volume 2. Springer, 2011.
- [14] C. Calabrese, M. Lombardi, E. Bollt, P. De Lellis, B. G. Bardy, and M. Di Bernardo. Spontaneous emergence of leadership patterns drives synchronization in complex human networks. *Scientific Reports*, 11(1):18379, 2021.
- [15] S. Camazine, J.-L. Deneubourg, N. R. Franks, J. Sneyd, G. Theraula, and E. Bonabeau. *Self-organization in biological systems*. Princeton University Press, 2020.
- [16] G. Cavagnari, G. Savaré, and G. E. Sodini. A Lagrangian approach to totally dissipative evolutions in Wasserstein spaces. *arXiv preprint arXiv:2305.05211*, 2023.
- [17] G. Cavagnari, G. Savaré, and G. E. Sodini. Dissipative probability vector fields and generation of evolution semigroups in Wasserstein spaces. *Probability Theory and Related Fields*, 185(3):1087–1182, 2023.
- [18] A. S. Cherny. On the uniqueness in law and the pathwise uniqueness for stochastic differential equations. *Theory of Probability & Its Applications*, 46(3):406–419, 2002.
- [19] F. Cucker and S. Smale. Emergent behavior in flocks. *IEEE Transactions on automatic control*, 52(5):852–862, 2007.
- [20] B. Düring, P. Markowich, J.-F. Pietschmann, and M.-T. Wolfram. Boltzmann and Fokker–Planck equations modelling opinion formation in the presence of strong leaders. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 465(2112):3687–3708, 2009.
- [21] X. Fernique. Regularité des trajectoires des fonctions aleatoires gaussiennes. In *Ecole d’Eté de Probabilités de Saint-Flour IV—1974*.
- [22] M. Girardi. Weak vs. norm compactness in  $L_1$ : the Bocce criterion. *Studia Mathematica*, 98(1):95–97, 1991.
- [23] J. Hofbauer, K. Sigmund, et al. *Evolutionary games and population dynamics*. Cambridge University Press, 1998.

- [24] E. J. Balder, M. Girardi, and V. Jalby. From weak to strong types of  $\mathcal{L}_E^1$ -convergence by the Bocce criterion. *Studia Mathematica*, 111(3):241–262, 1994.
- [25] A. Kechris. *Classical Descriptive Set Theory*, volume 156. Springer Science & Business Media, 2012.
- [26] E. F. Keller and L. A. Segel. Initiation of slime mold aggregation viewed as an instability. *Journal of Theoretical Biology*, 26(3):399–415, 1970.
- [27] J. L. Kelley and I. Namioka. *Linear Topological Spaces*. Springer, 1963.
- [28] D. Ko and E. Zuazua. Asymptotic behavior and control of a “guidance by repulsion” model. *Mathematical Models and Methods in Applied Sciences*, 30(04):765–804, 2020.
- [29] A. Lama and M. di Bernardo. Shepherding and herdability in complex multiagent systems. *Physical Review Research*, 6(3):L032012, 2024.
- [30] T. S. Lim, Y. Lu, and J. H. Nolen. Quantitative propagation of chaos in a bimolecular chemical reaction-diffusion model. *SIAM Journal on Mathematical Analysis*, 52(2):2098–2133, 2020.
- [31] M. Lombardi, W. H. Warren, and M. di Bernardo. Nonverbal leadership emergence in walking groups. *Scientific reports*, 10(1):18948, 2020.
- [32] G. C. Maffettone, A. Boldini, M. Di Bernardo, and M. Porfiri. Continuification control of large-scale multiagent systems in a ring. *IEEE Control Systems Letters*, 7:841–846, 2022.
- [33] G. C. Maffettone, M. Porfiri, and M. Di Bernardo. Continuification control of large-scale multiagent systems under limited sensing and structural perturbations. *IEEE Control Systems Letters*, 7:2425–2430, 2023.
- [34] V. Martinelli, D. Salzano, D. Fiore, and M. di Bernardo. Multicellular PI control for gene regulation in microbial consortia. *IEEE Control Systems Letters*, 6:3373–3378, 2022.
- [35] E. McShane. Extension of range of functions. *Bull. Amer. Math. Soc.*, 40(12):837–842, 1934.
- [36] P. A. Meyer. *Probability and Potentials*. Blaisdell Publishing Company, 1966.
- [37] P. Mozgunov, M. Beccuti, A. Horvath, T. Jaki, R. Sirovich, and E. Bibbona. A review of the deterministic and diffusion approximations for stochastic chemical reaction networks. *Reaction Kinetics, Mechanisms and Catalysis*, 123(2):289–312, 2018.
- [38] K. Oelschläger. On the derivation of reaction-diffusion equations as limit dynamics of systems of moderately interacting stochastic processes. *Probability Theory and Related Fields*, 82(4):565–586, 1989.
- [39] B. Oksendal. *Stochastic differential equations: an introduction with applications*. Springer Science & Business Media, 2013.
- [40] G. Orlando. Mean-field optimal control in a multi-agent interaction model for prevention of maritime crime. *Advances in Continuous and Discrete Models*, 2023(1):24, 2023.
- [41] G. Piacentini, P. Goatin, and A. Ferrara. Traffic control via moving bottleneck of coordinated vehicles. *IFAC-PapersOnLine*, 51(9):13–18, 2018.
- [42] L. Pick, A. Kufner, O. John, and S. Fucík. *Function spaces*, 1. Walter de Gruyter, 2012.
- [43] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.
- [44] D. Salzano, B. Shannon, C. Grierson, L. Marucci, N. J. Savery, and M. di Bernardo. In-vivo distributed multicellular control of gene expression in microbial consortia. *bioRxiv*, pages 2023–11, 2023.
- [45] G. Toscani. Kinetic models of opinion formation. *Communications in mathematical sciences*, 4(3):481–496, 2006.
- [46] C. Villani. *Optimal transport*, volume 338 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2009. Old and new.
- [47] Y. Wang, X. Zhu, and P. Kloeden. Compactness in Lebesgue–Bochner spaces of random variables and the existence of mean-square random attractors. *Stochastics and Dynamics*, 19(04):1950032, 2019.
- [48] D. Williams. *Probability with Martingales*. Cambridge University Press, 1991.
- [49] E. M. H. Zahugi, M. M. Shanta, and T. Prasad. Oil spill cleaning up using swarm of robots. In *Advances in Computing and Information Technology: Proceedings of the Second International Conference on*

*Advances in Computing and Information Technology (ACITY) July 13-15, 2012, Chennai, India-Volume 3, pages 215–224. Springer, 2013.*

FRANCESCA ANCESCHI, DIPARTIMENTO DI INGEGNERIA INDUSTRIALE E SCIENZE MATEMATICHE,  
UNIVERSITÀ POLITECNICA DELLE MARCHE, VIA BRECCIE BIANCHE 12, 60131 ANCONA, ITALY  
*Email address:* f.anceschi@staff.unipvm.it

GIACOMO ASCIONE, SCUOLA SUPERIORE MERIDIONALE, UNIVERSITÀ DI NAPOLI, LARGO SAN MAR-  
CELLINO 10, 80138 NAPOLI, ITALY  
*Email address:* giacomo.ascione@unina.it

DANIELE CASTORINA, DIPARTIMENTO DI MATEMATICA E APPLICAZIONI, UNIVERSITÀ DI NAPOLI,  
VIA CINTIA, MONTE S. ANGELO 80126 NAPOLI, ITALY  
*Email address:* daniele.castorina@unina.it

FRANCESCO SOLOMBRINO, DIPARTIMENTO DI MATEMATICA E APPLICAZIONI, UNIVERSITÀ DI NAPOLI,  
VIA CINTIA, MONTE S. ANGELO 80126 NAPOLI, ITALY  
*Email address:* francesco.solombrino@unina.it