



**HAL**  
open science

# Carathéodory Theory and A Priori Estimates for Continuity Inclusions in the Space of Probability Measures

Benoît Bonnet-Weill, Hélène Frankowska

► **To cite this version:**

Benoît Bonnet-Weill, Hélène Frankowska. Carathéodory Theory and A Priori Estimates for Continuity Inclusions in the Space of Probability Measures. *Nonlinear Analysis: Theory, Methods and Applications*, 2024, 247, pp.113595. 10.1016/j.na.2024.113595 . hal-03969118v2

**HAL Id: hal-03969118**

**<https://hal.science/hal-03969118v2>**

Submitted on 4 Nov 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Carathéodory Theory and A Priori Estimates for Continuity Inclusions in the Space of Probability Measures\*

Benoît Bonnet-Weill<sup>†</sup> and Héléne Frankowska<sup>‡</sup>

November 4, 2024

## Abstract

In this article, we extend the foundations of the theory of differential inclusions in the space of compactly supported probability measures, introduced recently by the authors, to the setting of general Wasserstein spaces. In this context, we prove a novel existence result à la Peano for this class of dynamics under mere Carathéodory regularity assumptions. The latter is based on a natural, yet previously unexplored set-valued adaptation of the semi-discrete Euler scheme proposed by Filippov to study ordinary differential equations whose right-hand sides are measurable in the time variable. By leveraging some of the underlying methods along with new estimates for solutions of continuity equations, we also bring substantial improvements to the earlier versions of the Filippov estimates, compactness and relaxation properties of the solution sets of continuity inclusions, which are derived in the Cauchy-Lipschitz framework.

**Keywords :** Continuity Inclusions, Optimal Transport, Set-Valued Analysis, Peano Existence Theorem, Filippov Estimates, Compactness and Relaxation.

**MSC2020 Subject Classification :** 28B20, 34A60, 34G20, 46N20, 49Q22

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Preliminaries</b>	<b>4</b>
2.1	Measure theory and optimal transport in Banach spaces . . . . .	5
2.2	Set-valued analysis and topological properties of continuous functions . . . . .	7
2.3	Continuity equations and inclusions in the space of probability measures . . . . .	10
<b>3</b>	<b>Existence à la Peano in the Carathéodory framework</b>	<b>14</b>
<b>4</b>	<b>A priori estimates, compactness and relaxation in the Cauchy-Lipschitz framework</b>	<b>22</b>
4.1	Existence of solutions and Filippov estimates . . . . .	22
4.2	Compactness of the solution set and relaxation property . . . . .	28
	<b>Appendices</b>	<b>34</b>
A	Proof of Lemma 2.10 . . . . .	34
B	Proof of Lemma 2.15 . . . . .	35
C	Proof of Proposition 2.21 . . . . .	36
D	Proof of Proposition 2.23 . . . . .	37

---

\*This material is based upon work supported by the Air Force Office of Scientific Research under award number FA9550-18-1-0254.

<sup>†</sup>LAAS-CNRS, Université de Toulouse, CNRS, 7 avenue du colonel Roche, F-31400 Toulouse, France. *E-mail:* benoit.bonnet@laas.fr (Corresponding author)

<sup>‡</sup>CNRS, IMJ-PRG, UMR 7586, Sorbonne Université, 4 place Jussieu, 75252 Paris, France. *E-mail:* helene.frankowska@imj-prg.fr

# 1 Introduction

In recent years, the study of continuity equations in the space of measures has been the object of an extensive interest in several mathematical communities. While the analysis of such partial differential equations was more commonly conveyed in Lebesgue or Sobolev spaces – in which one could establish classical well-posedness results –, several research currents in pure and applied mathematics prompted the investigation of more general and weaker notions of solutions to these dynamics.

Amongst these research endeavours, one of the most influential was the development of the modern theory of optimal transport – rendered in the treatises [7, 74, 77] – along with that of the theory of gradient flows in measure spaces. Indeed, the concepts introduced in the seminal works [60, 65] and later formalised in [7] expounded the fact that one could construct weak solutions to transport equations with very irregular driving fields by applying continuous-time steepest descent schemes to energy functionals defined over the space of probability measures. This far-reaching observation allowed to derive general well-posedness results for a wide variety of evolution equations encountered in physics, rational mechanics and biology, on the basis that they exhibit a variational structure, see e.g. [7, 25, 27, 41, 55, 60, 63, 65, 75] and references therein. The infatuation for this innovative point of view stemmed both from its theoretical merits, and from its practical efficiency and adaptedness for designing numerical methods [10, 24, 67]. In a similar vein, the theory of mean-field games [23, 58, 62] – located halfway between control theory and the calculus of variations – largely contributed to the popularisation of dynamical problems in measure spaces. Incidentally, some of the core concepts of these emerging research trends found relevant application outlets in a wealth of multiscale models aiming towards efficient descriptions of pedestrian flows [35, 64, 71] as well as opinion propagation models [2, 56, 72] and swarming dynamics [1, 26, 28]. Another active field of research that put this corpus of results to good use is that of mean-field control [14, 17, 29, 44, 45, 46, 59], which saw the birth of several relevant extensions of the classical theory as further elaborated upon below.

In the aforescribed context, a growing body of literature at the intersection between PDE analysis, dynamical systems theory and optimal transport has been concerned with the derivation of general well-posedness results for Cauchy problems of the form

$$\begin{cases} \partial_t \mu(t) + \operatorname{div}_x(v(t, \mu(t))\mu(t)) = 0, \\ \mu(0) = \mu^0. \end{cases} \quad (1.1)$$

Therein, the initial datum  $\mu^0 \in \mathcal{P}(\mathbb{R}^d)$  is a Borel probability measure, while the velocity field  $v : [0, T] \times \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  is a Lebesgue-Borel map which may be *nonlocal*, in the sense that it is allowed to depend on the measure variable itself. As alluded to in the previous paragraph, one of the main frameworks in which one can meaningfully derive well-posedness results for (1.1) is that of Wasserstein gradient flows (see e.g. [7, Chapter 11]), wherein  $v(t, \mu(t)) \in L^p(\mathbb{R}^d, \mathbb{R}^d; \mu(t))$  is given as the (sub)gradient of a functional defined over the space of measures. Analogously, in the theory of mean-field games, the well-posedness of the forward measure dynamics frequently originates from a variational principle, which requires that the common cost of the agents satisfies a suitable convexity condition [62]. It is worth noting that general existence results are also available for irregular driving fields when the dynamics exhibits a Hamiltonian structure, see for instance [6]. On the other hand, as amply highlighted by the discussions in [7, Chapter 8], the dynamics in (1.1) admits a natural interpretation as an ordinary differential equation in the Wasserstein manifold – or rather bundle –  $(\mathcal{P}_p(\mathbb{R}^d), W_p(\cdot, \cdot))$  for  $p \in (1, +\infty)$ . Thus, in the absence of an underlying variational structure, one should expect that the well-posedness of said Cauchy problems would rather stem from the regularity properties of the driving vector field. In that case, the results available in the literature can be split into two categories, depending on whether the velocity field is nonlocal or not.

- ◊ When  $v : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  is independent of the measure variable, the known optimal well-posedness settings are, on the one hand, Carathéodory assumptions (or small variations thereof) when the initial measure is arbitrary, and on the other hand Sobolev [36] or BV [3] regularity in the space variable, combined with integral bounds on the divergence or more general incompressibility assumptions [12], when the initial measure is absolutely continuous with respect to the Lebesgue measures. The structuring concepts of this latter class of solutions – which

are usually referred to as *regular Lagrangian flows* –, are surveyed together with the classical Cauchy-Lipschitz theory in [4].

- ◊ When  $v : [0, T] \times \mathcal{P}(\mathbb{R}^d) \times \mathbb{R}^d \mapsto \mathbb{R}^d$  also depends on the measure variable, general sufficient conditions for the well-posedness of (1.1) are only available in the Cauchy-Lipschitz and Carathéodory regularity frameworks (see e.g. [14, 34, 70] and references therein). In particular, there are currently no known generalisations of the concept of Lagrangian flow to the setting of nonlocal continuity equations.

We end this literature overview by advertising a recent body of work initiated in [69] and furthered in [19, 31, 68], in which the authors investigate relaxations of (1.1) wherein the driving fields are replaced by probability measures over the tangent bundle. This line of study – which is highly reminiscent of the earlier work [11] – bears strong resemblance with the theory of Young measures, and has already produced very promising results shedding light on the interplay between contraction semigroups in measure spaces and various kinds of Euler schemes.

In the lineage of this body of work, the aim of this paper is to introduce several far-reaching refinements of the theory of *continuity inclusions* in Wasserstein spaces, whose elaboration started in our previous work [14]. Therein, given a compactly supported measure  $\mu^0 \in \mathcal{P}_c(\mathbb{R}^d)$  and a *set-valued map*  $(t, \mu) \in [0, T] \times \mathcal{P}_c(\mathbb{R}^d) \rightrightarrows V(t, \mu) \subset L^p(\mathbb{R}^d, \mathbb{R}^d; \mu)$ , we defined solutions of the Cauchy problem

$$\begin{cases} \partial_t \mu(t) \in -\operatorname{div}_x \left( V(t, \mu(t)) \mu(t) \right), \\ \mu(0) = \mu^0, \end{cases} \quad (1.2)$$

as the collection of all absolutely continuous curves  $\mu(\cdot) \in \operatorname{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d))$  for which there exists an  $\mathcal{L}^1$ -measurable selection  $t \in [0, T] \mapsto v(t) \in V(t, \mu(t))$  such that (1.1) holds. The main motivation for considering such objects is that differential inclusions, and set-valued analysis at large, play an instrumental role in control theory and in the calculus of variations, as evidenced by the reference works [8, 9, 33, 78] and [22, 32]. They provide convenient tools to establish existence results for variational problems [43], to derive first- and second-order optimality conditions – both in the form of a Pontryagin Maximum Principle [47, 52, 51, 54] or Hamilton-Jacobi-Bellman equations [48, 53] –, and to investigate qualitative properties of optimal trajectories [21]. Owing to their mathematical versatility, these schemes have recently started to percolate in the communities of mean-field control and mean-field games [15, 16, 20, 29, 59, 68]. Compared with other notions that were put forth to define solutions of differential inclusions in measure spaces, such as those of [30, 59], our approach presents the advantage of being coherent with the modern theory of differential inclusions in vector spaces, surveyed e.g. in [8, 9], as well as conceptually compatible with the geometric structure of Wasserstein spaces and the interpretation of continuity equations as generalised ODEs following [7, 65]. Besides, our construct is well-adapted to the study of control problems, as it ensures that there is a one-to-one correspondence between solutions of controlled measure dynamics and their set-valued counterparts.

The first main contribution of this article is an existence result à la Peano for (1.2), presented in Theorem 3.2, and derived under mere Carathéodory regularity assumptions. To be more precise, we show therein that if  $V : [0, T] \times \mathcal{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d, \mathbb{R}^d)$  is  $\mathcal{L}^1$ -measurable with respect to  $t \in [0, T]$ , continuous with respect to  $\mu \in \mathcal{P}_p(\mathbb{R}^d)$  and has convex images, then (1.2) admits a solution from every initial datum. The requirement that the admissible velocities are convex is commonly known to be unavoidable to prove the existence of solutions to differential inclusions whose right-hand sides are not Lipschitz (see e.g. [8, Chapter 2]), even in the familiar context of finite-dimensional vector spaces. Our strategy for proving this result revolves around an astute set-valued adaptation of the semi-discrete Euler scheme due to Filippov for Carathéodory ODEs, see e.g. [42, Chapter 1]. To the best of our knowledge, this approach was not previously investigated even for classical differential inclusions. It also relies on rather new ways of studying the compactness of solutions to continuity equations, which we believe are of independent interest.

The second contribution of this article lies in the transposition of the findings of [14], derived for curves of compactly supported measures, to the setting of general Wasserstein spaces. The first of these

key results are the so-called Filippov estimates, presented in Theorem 4.2, which provide the existence of a solution to (1.2) whose distance to an a priori fixed measure curve is precisely controlled. Such estimates are extremely useful to produce admissible trajectories for control systems when conducting perturbative arguments and performing linearisations. The second of these results is the compactness of the solution set when the right-hand side of the dynamics is convex, which is needed in virtually every existence proof based on weak compactness arguments, both in optimal control theory and in the calculus of variations. The third and last is the so-called relaxation theorem, which asserts that in the absence of convexity, the closure of the solution set of (1.2) coincides with that of the Cauchy problem in which the dynamics has been convexified. These three pivotal properties have been widely used in conjunction with one another to derive sharp first- and second-order optimality conditions for optimal control problems in various contexts, see e.g. [49, 52] for ordinary differential equations, [49, 50] for general evolution equations in infinite-dimensional Banach spaces, and our previous work [15] concerned with mean-field optimal control problems. It should be noted that the relaxation theorem is also essential when investigating the fine properties of solutions to Hamilton-Jacobi-Bellman equations, as it allows to posit without loss of generality the existence of optimal trajectories associated with the corresponding value function.

From a more technical standpoint, the extension of these results required new insights on the definition of continuity inclusions, as one needed a new functional setting that was amenable to working with velocities having unbounded support, but still enjoyed good topological properties. These reflections lead us to study continuity inclusions driven by velocity selections valued in  $C^0(\mathbb{R}^d, \mathbb{R}^d)$  endowed with the topology of local uniform convergence, which happens to be a separable Fréchet space whose compact subsets are precisely characterised by the Ascoli-Arzelà theorem. It is our personal opinion that adopting this fresh viewpoint allows for a cleaner, more general, and perhaps more streamlined angle to study measure dynamics. We finally stress that while some of the working assumptions used throughout this manuscript are more stringent than those in our previous work, mainly to palliate the fact that the measures under consideration may have unbounded supports, the results presented below strictly contain those of [14], up to minor technical adjustments. It should finally be noted that some of our developments are based on new quantitative stability estimates for continuity equations under low regularity requirements, which should again be of independent interest for our fellow practitioners.

The contributions and organisation of the article can be summarised as follows. In Section 2, after recalling some concepts pertaining to measure theory, optimal transport, set-valued and functional analysis, we derive new compactness and stability estimates for solutions of continuity equations in Proposition 2.21 and Proposition 2.23 respectively. We then define solutions of (1.2) for set-valued maps  $V : [0, T] \times \mathcal{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d, \mathbb{R}^d)$  in terms of velocity selections which are measurable for the standard Fréchet topology of the space of continuous functions (see Definition 2.12 below). Then, in Section 3, we prove a novel existence result for solutions of (1.2) in Theorem 3.2, under the mere assumption that the right-hand side of the dynamics is Caratheodory and convex-valued. We subsequently move on in Section 4 to the Cauchy-Lipschitz setting, and discuss the main results of the theory of differential inclusions that may be derived therein. In Section 4.1, we start by establishing two far-reaching versions of the Filippov estimates, starting with a local one in Theorem 4.2 and proceeding with a global one in Corollary 4.4. Then, we turn our attention to the topological properties of the solution sets of (1.2) in Section 4.2. In Theorem 4.5, we start by showing that the latter are compact for the topology of uniform convergence when the right-hand side of the dynamics has convex images. In Theorem 4.6, we then relax the convexity requirement and prove that in this case, the closure of the solution set coincides with that of the convexified Cauchy problem. We finally close the paper by an appendix containing the proofs of several technical results and estimates.

## 2 Preliminaries

In the coming sections, we expose preliminary results pertaining to measure theory, set-valued analysis, optimal transport and measure dynamics in general.

## 2.1 Measure theory and optimal transport in Banach spaces

In this first preliminary section, we recollect basic notions of measure theory and optimal transport in Banach spaces, for which we largely refer to [5, 13, 39] and [7] respectively.

**Elements of measure and integration theory.** Given a separable Banach space  $(X, \|\cdot\|_X)$ , we will denote by  $X^*$  its topological dual and write  $\langle \cdot, \cdot \rangle_X$  for the underlying duality pairing. In the sequel, the notation  $C_b^0(X, \mathbb{R}^d)$  will refer to the vector space of continuous bounded maps from  $X$  into  $\mathbb{R}^d$ , and in the particular case where  $X = \mathbb{R}^m$  for some  $m \geq 1$ , we will denote by  $C_c^\infty(\mathbb{R}^m, \mathbb{R})$  the vector space of infinitely differentiable functions with compact support. Letting  $(Y, d_Y(\cdot, \cdot))$  be a complete separable metric space, we shall more generally write  $C^0(X, Y)$  for the set of continuous maps from  $X$  into  $Y$ , as well as  $\text{AC}(I, Y)$  for that of absolutely continuous arcs defined over an interval  $I \subset \mathbb{R}$  with values in  $Y$ . In addition,  $\text{Lip}(\Omega, Y)$  will stand for the space of Lipschitz maps from a subset  $\Omega \subset X$  into  $Y$ , and we shall write  $\text{Lip}(\phi; \Omega)$  for the Lipschitz constant of an element  $\phi \in \text{Lip}(\Omega, Y)$ .

In what follows, given closed set  $\Omega \subset X$ , we will consider the vector space  $\mathcal{M}(\Omega, \mathbb{R}^d)$  of  $\mathbb{R}^d$ -valued finite Radon measures. By the Riesz representation theorem (see e.g. [5, Theorem 1.54]), it is known that when  $\Omega$  is compact, the latter is isomorphic to the topological dual of the Banach space  $(C^0(\Omega, \mathbb{R}^d), \|\cdot\|_{C^0(\Omega, \mathbb{R}^d)})$  under the action of the duality pairing

$$\langle \nu, \phi \rangle_{C^0(\Omega, \mathbb{R}^d)} := \sum_{i=1}^d \int_{\Omega} \phi_i(x) d\nu_i(x), \quad (2.1)$$

defined for all  $\nu \in \mathcal{M}(\Omega, \mathbb{R}^d)$  and  $\phi \in C^0(\Omega, \mathbb{R}^d)$ . Throughout the article, we denote by  $\mathcal{P}(\Omega)$  the space of Borel probability measures over  $\Omega$  endowed with the *narrow topology*, that is the coarsest topology for which the mappings

$$\mu \in \mathcal{P}(\Omega) \mapsto \int_X \phi(x) d\mu(x) \in \mathbb{R} \quad (2.2)$$

are continuous for every element  $\phi \in C_b^0(\Omega, \mathbb{R})$ . It is a standard fact in measure theory (see e.g. [7, Remark 5.1.2]) that  $\mathcal{P}(\Omega)$  endowed with the narrow topology is a separable space, and we will write

$$\mu_n \xrightarrow[n \rightarrow +\infty]{*} \mu,$$

for the notion of convergence induced by (2.2) over  $\mathcal{P}(\Omega)$ .

Given two separable Banach spaces  $(X, \|\cdot\|_X)$  and  $(Y, \|\cdot\|_Y)$  along with some  $p \in [1, +\infty)$ , we will write  $(L^p(\Omega, Y; \mu), \|\cdot\|_{L^p(\Omega, Y; \mu)})$  for the space of maps from a subset  $\Omega \subset X$  into  $Y$  which are  $p$ -integrable in the sense of Bochner (see e.g. [39, Chapter II]) with respect to a measure  $\mu \in \mathcal{M}(\Omega, \mathbb{R}_+)$ . Analogously, we let  $L^\infty(\Omega, Y; \mu)$  be the space of  $\mu$ -essentially bounded maps from  $X$  into  $Y$ , and use the denser notation  $(L^p(I, Y), \|\cdot\|_{L^p(I)})$  for  $p \in [1, +\infty]$  when  $X := I$  is an interval and  $\mu := \mathcal{L}^1$  is the standard 1-dimensional Lebesgue measure. We recall below a powerful criterion for weak compactness in  $L^1(I, X)$ , whose statement can be found in [38, Corollary 2.6] (see also the earlier versions [37, 76]).

**Theorem 2.1** (A weak compactness criterion for Bochner integrable maps). *Let  $(X, \|\cdot\|_X)$  be a separable Banach space,  $I \subset \mathbb{R}$  be an interval and  $(v_n(\cdot)) \subset L^1(I, X)$ . Suppose that there exists a map  $m(\cdot) \in L^1(I, \mathbb{R}_+)$  and a family  $(K_t)_{t \in I}$  of weakly compact subsets of  $X$  such that*

$$\|v_n(t)\|_X \leq m(t) \quad \text{and} \quad v_n(t) \in K_t$$

*for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$  and each  $n \geq 1$ . Then, there exists a subsequence  $(v_{n_k}(\cdot)) \subset L^1(I, X)$  that converges weakly to an element  $v(\cdot) \in L^1(I, X)$ . In particular*

$$\int_I \langle \nu(t), v(t) - v_{n_k}(t) \rangle_X dt \xrightarrow[k \rightarrow +\infty]{} 0,$$

*for every  $\nu(\cdot) \in L^\infty(I, X^*) \subset L^1(I, X)^*$ .*

**Optimal transport and Wasserstein spaces.** Throughout this article, the notation  $\mathcal{P}_p(X)$  will refer to the subset of probability measures whose *moment of order*  $p \in [1, +\infty)$  is finite, that is

$$\mathcal{M}_p^p(\mu) := \int_X |x|^p d\mu(x) < +\infty.$$

In the following definition, we recall the known concepts of *image measure* through a Borel map, as well as that of *transport plan*.

**Definition 2.2** (Image measures and transport plans). *The image of a measure  $\mu \in \mathcal{P}(X)$  through a Borel map  $f : X \rightarrow Y$ , denoted by  $f_{\#}\mu \in \mathcal{P}(Y)$ , is defined as*

$$f_{\#}\mu(B) = \mu(f^{-1}(B)),$$

for every Borel set  $B \subset Y$ . Given two probability measures  $\mu, \nu \in \mathcal{P}(X)$ , we say that an element  $\gamma \in \mathcal{P}(X \times X)$  is a transport plan between  $\mu$  and  $\nu$ , denoted by  $\gamma \in \Gamma(\mu, \nu)$ , provided that

$$\pi_{\#}^1 \gamma = \mu \quad \text{and} \quad \pi_{\#}^2 \gamma = \nu,$$

where  $\pi^1, \pi^2 : X \times X \rightarrow X$  stand for the projections onto the first and second factors.

When  $(X, \|\cdot\|_X) := (\mathbb{R}^d, |\cdot|)$  is a  $d$ -dimensional real vector space endowed with its usual Euclidean structure, it is a standard result in optimal transport theory that for any  $p \in [1, +\infty)$ , the quantity

$$W_p(\mu, \nu) := \inf_{\gamma \in \Gamma(\mu, \nu)} \left( \int_{\mathbb{R}^{2d}} |x - y|^p d\gamma(x, y) \right)^{1/p} \quad (2.3)$$

defined for each  $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$  is a distance over  $\mathcal{P}_p(\mathbb{R}^d)$ . Moreover, it comes as an easy consequence of the direct method of the calculus of variations that the infimum in (2.3) is always attained, and we denote by  $\Gamma_o(\mu, \nu)$  the corresponding set of  *$p$ -optimal transport plans*. In the following propositions, we recall some of the main properties of the *Wasserstein spaces*, along with a handy distance estimate.

**Proposition 2.3** (Topology of Wasserstein spaces). *The metric spaces  $(\mathcal{P}_p(\mathbb{R}^d), W_p(\cdot, \cdot))$  are complete and separable, and the Wasserstein distances metrize the narrow topology (2.2), in the sense that*

$$W_p(\mu_n, \mu) \xrightarrow{n \rightarrow +\infty} 0 \quad \text{if and only if} \quad \begin{cases} \mu_n \xrightarrow{n \rightarrow +\infty}^* \mu, \\ \int_{\mathbb{R}^d} |x|^p d\mu_n(x) \xrightarrow{n \rightarrow +\infty} \int_{\mathbb{R}^d} |x|^p d\mu(x), \end{cases}$$

for any sequence  $(\mu_n) \subset \mathcal{P}_p(\mathbb{R}^d)$  and each  $\mu \in \mathcal{P}_p(\mathbb{R}^d)$ . Moreover, a subset  $\mathcal{K} \subset \mathcal{P}_p(\mathbb{R}^d)$  is relatively compact for the topology induced by the metric  $W_p(\cdot, \cdot)$  if and only if it satisfies

$$\sup_{\mu \in \mathcal{K}} \int_{\{x \text{ s.t. } |x| \geq k\}} |x|^p d\mu(x) \xrightarrow{k \rightarrow +\infty} 0,$$

namely if and only if it is uniformly  $p$ -integrable.

*Proof.* We point the interested reader to [7, Chapter 7] or [77, Chapter 6]. □

**Proposition 2.4** (Classical optimal transport estimate). *Given some  $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$ , it holds that*

$$\int_{\mathbb{R}^d} \phi(x) d(\mu - \nu)(x) \leq \text{Lip}(\phi; \mathbb{R}^d) W_1(\mu, \nu) \leq \text{Lip}(\phi; \mathbb{R}^d) W_p(\mu, \nu) \quad (2.4)$$

for every  $\phi \in \text{Lip}(\mathbb{R}^d, \mathbb{R})$ .

## 2.2 Set-valued analysis and topological properties of continuous functions

In this section, we recall some notations and results of set-valued and functional analysis. We shall mostly rely on the treatises [8, 9] for the former and reference monographs [57, 73] for the latter.

In what follows, we write  $(E, d_E(\cdot, \cdot))$  to denote a separable *Fréchet space*, i.e. a complete separable locally convex topological vector space  $E$  whose topology is induced by a translation-invariant metric  $d_E(\cdot, \cdot)$ . In this context, we define the *closed convex hull* of a set  $B \subset E$  as

$$\overline{\text{co}}(B) := \overline{\bigcup_{N \geq 1} \left\{ \sum_{i=1}^N \alpha_i b_i \text{ s.t. } b_i \in B, \alpha_i \in [0, 1] \text{ for } i \in \{1, \dots, N\} \text{ and } \sum_{i=1}^N \alpha_i = 1 \right\}}^E,$$

where “ $\overline{\bullet}^E$ ” stands for the closure with respect to  $d_E(\cdot, \cdot)$ . We will also use the generic notation

$$\text{dist}_E(x; \mathcal{Q}) := \inf_{y \in \mathcal{Q}} d_E(x, y)$$

for the distance between an element  $x \in E$  and a closed set  $\mathcal{Q} \subset E$ .

**Set-valued analysis.** Given two complete separable metric spaces  $(X, d_X(\cdot, \cdot))$  and  $(Y, d_Y(\cdot, \cdot))$ , we write  $\mathcal{F} : X \rightrightarrows Y$  to mean that  $\mathcal{F}(\cdot)$  is a *set-valued map* – or *multifunction* – from  $X$  into  $Y$ . A set-valued map is said to have closed images (respectively convex, when the notion makes sense) if the sets  $\mathcal{F}(x) \subset Y$  are closed (respectively convex) for each  $x \in X$ . In addition, we define its *graph* by

$$\text{Graph}(\mathcal{F}) := \left\{ (x, y) \in X \times Y \text{ s.t. } y \in \mathcal{F}(x) \right\}.$$

Below, we recall the standard concept of measurability for set-valued mappings defined over some subinterval  $I \subset \mathbb{R}$  of the real line, endowed with the complete  $\sigma$ -algebra of Lebesgue-measurable sets.

**Definition 2.5** (Measurable set-valued maps and measurable selections). *A multifunction  $\mathcal{F} : I \rightrightarrows Y$  is said to be  $\mathcal{L}^1$ -measurable provided that*

$$\mathcal{F}^{-1}(\mathcal{O}) := \left\{ t \in I \text{ s.t. } \mathcal{F}(t) \cap \mathcal{O} \neq \emptyset \right\}$$

*is  $\mathcal{L}^1$ -measurable for every open set  $\mathcal{O} \subset Y$ . A mapping  $f : I \rightarrow Y$  is called a measurable selection of  $\mathcal{F}(\cdot)$  if it is  $\mathcal{L}^1$ -measurable and such that  $f(t) \in \mathcal{F}(t)$  for  $\mathcal{L}^1$ -almost every  $t \in I$ .*

In the following theorem, we recollect a deep result of set-valued analysis which provides the existence of measurable selections for measurable multifunctions whose images are closed subsets of a complete separable metric space.

**Theorem 2.6** (Existence of measurable selections). *Every  $\mathcal{L}^1$ -measurable set-valued map  $\mathcal{F} : I \rightrightarrows Y$  with nonempty closed image admits a measurable selection.*

In our subsequent developments, we will use the notions of *continuity* and *Lipschitz regularity* for set-valued mappings, both of which are recalled in the following definitions. Therein, we denote by  $\mathbb{B}_X(x, r)$  and  $\mathbb{B}_Y(y, r)$  the closed ball of radius  $r > 0$  centered at  $x \in X$  and  $y \in Y$  respectively, and use the condensed notation  $\mathbb{B}_X(\Omega, r) := \{x \in X \text{ s.t. } d_X(x, x') \leq r \text{ for some } x' \in \Omega\}$  for any  $\Omega \subset X$ .

**Definition 2.7** (Continuous set-valued maps). *A multifunction  $\mathcal{F} : X \rightrightarrows Y$  is said to be continuous at  $x \in X$  if both the following conditions hold:*

- (i)  $\mathcal{F}(\cdot)$  is upper-semicontinuous at  $x \in X$ , i.e. for every  $\varepsilon > 0$ , there exists  $\delta > 0$  such that

$$\mathcal{F}(x') \subset \mathbb{B}_Y(\mathcal{F}(x), \varepsilon)$$

for all  $x' \in \mathbb{B}_X(x, \delta)$ .



(ii)  $\mathcal{F}(\cdot)$  is lower-semicontinuous at  $x \in X$ , i.e. for every  $\varepsilon > 0$  and each  $y \in \mathcal{F}(x)$ , there exists  $\delta > 0$  such that

$$\mathcal{F}(x') \cap \mathbb{B}_Y(y, \varepsilon) \neq \emptyset$$

for all  $x' \in \mathbb{B}_X(x, \delta)$ .

**Definition 2.8** (Lipschitz continuous set-valued maps). *A multifunction  $\mathcal{F} : X \rightrightarrows Y$  is said to be Lipschitz continuous if there exists a constant  $L > 0$  such that*

$$\mathcal{F}(x') \subset \mathbb{B}_Y(\mathcal{F}(x), Ld_X(x, x')),$$

for every  $x, x' \in X$ .

In the sequel, we will frequently resort to the general notion of *Carathéodory* set-valued map between metric spaces, which is defined as follows.

**Definition 2.9** (Carathéodory set-valued maps). *A set-valued map  $\mathcal{G} : I \times X \rightrightarrows Y$  is said to be Carathéodory if  $t \in I \rightrightarrows \mathcal{G}(t, x)$  is  $\mathcal{L}^1$ -measurable for all  $x \in X$  and  $x \in X \rightrightarrows \mathcal{G}(t, x)$  is continuous for  $\mathcal{L}^1$ -almost every  $t \in I$ .*

In the following lemma – whose proof is outlined in Appendix A –, we state measurable selection principles adapted from [9, Section 8.1] and [79, Section 9] that we shall extensively use in the sequel. Therein, we let  $\varphi : I \times Y \rightarrow \mathbb{R}_+ \cup \{+\infty\}$  be an extended real-valued function satisfying

$$\varphi(t, y) = \sup_{n \geq 1} \varphi_n(t, y) \tag{2.5}$$

for  $\mathcal{L}^1$ -almost every  $t \in I$  and all  $y \in Y$ , with  $(\varphi_n(\cdot, \cdot))$  being a pointwisely non-decreasing sequence of Carathéodory maps. This implies in particular that  $t \in I \mapsto \varphi(t, y)$  is  $\mathcal{L}^1$ -measurable for each  $y \in Y$ , whereas  $y \in Y \mapsto \varphi(t, y)$  is lower-semicontinuous for  $\mathcal{L}^1$ -almost every  $t \in I$ .

**Lemma 2.10** (Measurable selections principles). *Let  $\mathcal{F} : I \rightrightarrows Y$  and  $\mathcal{G} : I \times X \rightrightarrows Y$  be two set-valued maps with nonempty images and  $\varphi : I \times Y \rightarrow \mathbb{R}_+$  be as above. In addition, fix two  $\mathcal{L}^1$ -measurable functions  $x : I \rightarrow X$  and  $L : I \rightarrow \mathbb{R}_+$ .*

(a) *Suppose that  $\mathcal{F}(\cdot)$  is  $\mathcal{L}^1$ -measurable with compact images, and that the set-valued map*

$$t \in I \rightrightarrows \mathcal{F}(t) \cap \left\{ y \in Y \text{ s.t. } \varphi(t, y) \leq L(t) \right\}$$

*has nonempty images. Then the latter is  $\mathcal{L}^1$ -measurable, and there exists a measurable selection  $t \in I \mapsto f(t) \in \mathcal{F}(t)$  such that  $\varphi(t, f(t)) \leq L(t)$  for  $\mathcal{L}^1$ -almost every  $t \in I$ .*

(b) *Suppose that  $\mathcal{F}(\cdot)$  is  $\mathcal{L}^1$ -measurable with compact images. Then, the set-valued map*

$$t \in I \rightrightarrows \mathcal{F}(t) \cap \left\{ y \in Y \text{ s.t. } \varphi(t, y) = \inf_{z \in \mathcal{F}(t)} \varphi(t, z) \right\}$$

*is  $\mathcal{L}^1$ -measurable with nonempty closed images, and as such admits a measurable selection.*

(c) *Suppose that  $\mathcal{G}(\cdot, \cdot)$  is Carathéodory with closed images. Then, the set-valued map  $t \in I \rightrightarrows \mathcal{G}(t, x(t))$  is  $\mathcal{L}^1$ -measurable, and as such admits a measurable selection.*

We end this primer in set-valued analysis by stating an adaptation of Aumann's integral convexity theorem, for which we point the interested reader to [9, Theorem 8.6.4].

**Theorem 2.11** (Convexity of the Aumann integral). *Suppose that  $(X, \|\cdot\|_X)$  is a separable Banach space and let  $\mathcal{F} : I \rightrightarrows X$  be a set-valued map with closed nonempty images which is integrably bounded, in the sense that there exists some  $k(\cdot) \in L^1(I, \mathbb{R}_+)$  such that*

$$\mathcal{F}(t) \subset k(t)\mathbb{B}_X,$$

for  $\mathcal{L}^1$ -almost every  $t \in I$ . Then, for every measurable set  $\Omega \subset I$ , any measurable selection  $t \in I \mapsto f(t) \in \overline{\text{co}} \mathcal{F}(t)$  and all  $\delta > 0$ , there exists a measurable selection  $t \in I \mapsto f_\delta(t) \in \mathcal{F}(t)$  such that

$$\left\| \int_{\Omega} f(t) dt - \int_{\Omega} f_\delta(t) dt \right\|_X \leq \delta.$$

**Topologies and metrics over  $C^0(\mathbb{R}^d, \mathbb{R}^d)$ .** Throughout the manuscript, we will almost exclusively work with multifunctions whose values are subsets of  $C^0(\mathbb{R}^d, \mathbb{R}^d)$ . Following the pioneering work [80], it is known that the most natural topology to endow this space with is that of *uniform convergence on compact sets* – or *compact convergence* –, whose definition and distinctive features are recalled below.

**Definition 2.12** (The topology of compact convergence). *A sequence of continuous maps  $(v_n) \subset C^0(\mathbb{R}^d, \mathbb{R}^d)$  is said to converge uniformly on compact sets to some  $v \in C^0(\mathbb{R}^d, \mathbb{R}^d)$  provided that*

$$\|v - v_n\|_{C^0(K, \mathbb{R}^d)} \xrightarrow{n \rightarrow +\infty} 0,$$

for every compact set  $K \subset \mathbb{R}^d$ . The topology that this notion of convergence induces on  $C^0(\mathbb{R}^d, \mathbb{R}^d)$  is complete, separable, and metrised by the translation-invariant metric

$$d_{cc}(v, w) := \sum_{k=1}^{+\infty} 2^{-k} \min \left\{ 1, \|v - w\|_{C^0(B(0, k), \mathbb{R}^d)} \right\} \quad (2.6)$$

that is defined for each  $v, w \in C^0(\mathbb{R}^d, \mathbb{R}^d)$ . As such, the latter endows  $(C^0(\mathbb{R}^d, \mathbb{R}^d), d_{cc}(\cdot, \cdot))$  with the structure of a separable Fréchet space.

In our subsequent developments, we will always consider  $C^0(\mathbb{R}^d, \mathbb{R}^d)$  endowed with the separable Fréchet topology induced by  $d_{cc}(\cdot, \cdot)$ , which interestingly carries a functional characterisation of Carathéodory vector fields, as illustrated by the following result borrowed from [66, page 511].

**Lemma 2.13** (Carathéodory vector fields as measurable functions in  $C^0(\mathbb{R}^d, \mathbb{R}^d)$ ). *A vector field  $(t, x) \in I \times \mathbb{R}^d \mapsto v(t, x) \in \mathbb{R}^d$  is Carathéodory if and only if its functional lift  $t \in I \mapsto v(t) \in C^0(\mathbb{R}^d, \mathbb{R}^d)$  is  $\mathcal{L}^1$ -measurable for the topology of uniform convergence on compact sets.*

In light of this result, we will systematically identify  $\mathcal{L}^1$ -measurable maps  $t \in I \mapsto v(t) \in C^0(\mathbb{R}^d, \mathbb{R}^d)$  with Carathéodory vector fields  $v : I \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ , and thus work within the vector space

$$\mathcal{L}(I, C^0(\mathbb{R}^d, \mathbb{R}^d)) := \left\{ v : I \times \mathbb{R}^d \rightarrow \mathbb{R}^d \text{ s.t. } v(\cdot, \cdot) \text{ is a Carathéodory vector field} \right\}.$$

In addition to its amenable topological properties, the notion of uniform convergence on compact sets is particularly well-tailored to the formulation of compactness results, as illustrated by the following general version of the Ascoli-Arzelà theorem from [73, Theorem 11.28].

**Theorem 2.14** (Ascoli-Arzelà compactness theorem). *Let  $(v_n) \subset C^0(\mathbb{R}^d, \mathbb{R}^d)$  be a sequence of maps which are uniformly bounded and equi-continuous on compact sets. Then, there exists an element  $v \in C^0(\mathbb{R}^d, \mathbb{R}^d)$  for which*

$$d_{cc}(v_{n_k}, v) \xrightarrow{k \rightarrow +\infty} 0,$$

along a subsequence  $(v_{n_k}) \subset C^0(\mathbb{R}^d, \mathbb{R}^d)$ . Similarly if  $K \subset \mathbb{R}^d$  is compact and  $(v_n) \subset C^0(K, \mathbb{R}^d)$  is a sequence of uniformly bounded and equi-continuous maps, then there exists  $v \in C^0(K, \mathbb{R}^d)$  such that

$$\|v - v_{n_k}\|_{C^0(K, \mathbb{R}^d)} \xrightarrow{k \rightarrow +\infty} 0,$$

along a subsequence  $(v_{n_k}) \subset C^0(K, \mathbb{R}^d)$ .

*Proof.* The proof of the first compactness statement can be found in [73, Theorem 11.28], while that of the second one is simply the standard Ascoli-Arzelà theorem.  $\square$

In order to formulate the well-posedness results of Section 4, we will also need a global notion of vicinity for continuous maps, that will be inherently stronger than the local one provided by the metric  $d_{cc}(\cdot, \cdot)$ . For this reason, we will also consider the extended supremum metric defined by

$$d_{\text{sup}}(v, w) := \sup_{x \in \mathbb{R}^d} |v(x) - w(x)| \in \mathbb{R} \cup \{+\infty\}, \quad (2.7)$$

for every  $v, w \in C^0(\mathbb{R}^d, \mathbb{R}^d)$ . In the following lemma, we recollect for the sake of completeness a classical result which shows that the natural lift of  $d_{\text{sup}}(\cdot, \cdot)$  to the topological vector space  $\mathcal{L}(I, C^0(\mathbb{R}^d, \mathbb{R}^d))$  is complete, in the sense that its Cauchy sequences converge.

**Lemma 2.15** (Completeness of the integral of the extended metric). *Consider a sequence of maps  $(v_n(\cdot)) \subset \mathcal{L}(I, C^0(\mathbb{R}^d, \mathbb{R}^d))$  satisfying the Cauchy condition*

$$\int_0^T \mathbf{d}_{\text{sup}}(v_n(t), v_m(t)) dt \xrightarrow{n, m \rightarrow +\infty} 0. \quad (2.8)$$

*Then, there exists an element  $v(\cdot) \in \mathcal{L}(I, C^0(\mathbb{R}^d, \mathbb{R}^d))$  such that*

$$\int_0^T \mathbf{d}_{\text{sup}}(v_n(t), v(t)) dt \xrightarrow{t \rightarrow +\infty} 0.$$

*Proof.* The proof of this result being somewhat standard, it is deferred to Appendix B below.  $\square$

### 2.3 Continuity equations and inclusions in the space of probability measures

In this section, we recollect known results about continuity equations in measures spaces, following the usual Cauchy-Lipschitz and superposition-type theories surveyed e.g. in [4, 7], as well as their set-valued counterpart introduced by the authors of the present manuscript in [14].

Given a real number  $p \in [1, +\infty)$ , a time horizon  $T > 0$  and a velocity field  $v : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ , we will focus our attention on the well-posedness and qualitative properties of the Cauchy problem

$$\begin{cases} \partial_t \mu(t) + \text{div}_x(v(t)\mu(t)) = 0, \\ \mu(\tau) = \mu_\tau, \end{cases} \quad (2.9)$$

wherein  $(\tau, \mu_\tau) \in [0, T] \times \mathcal{P}_p(\mathbb{R}^d)$  is a fixed datum and  $\mu(\cdot) \in C^0(I, \mathcal{P}(\mathbb{R}^d))$  is a narrowly continuous curve of measure. The dynamics appearing in the first line of (2.9) is referred to as a *continuity equation*, and understood in the sense of distributions

$$\int_0^T \int_{\mathbb{R}^d} \left( \partial_t \phi(t, x) + \langle \nabla_x \phi(t, x), v(t, x) \rangle \right) d\mu(t)(x) dt = 0,$$

for every test function  $\phi \in C_c^\infty((0, T) \times \mathbb{R}^d, \mathbb{R})$ . Throughout this article, we will mostly work with velocity fields satisfying all or parts of the following assumptions which are standard when studying continuity equations in the Cauchy-Lipschitz framework.

#### Hypotheses (CE).

(i) *The vector field  $v : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  is Carathéodory, and there exists a map  $m(\cdot) \in L^1([0, T], \mathbb{R}_+)$  such that*

$$|v(t, x)| \leq m(t)(1 + |x|),$$

*for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$  and all  $x \in \mathbb{R}^d$ .*

(ii) *There exists a map  $l(\cdot) \in L^1([0, T], \mathbb{R}_+)$  such that*

$$\text{Lip}(v(t); \mathbb{R}^d) \leq l(t),$$

*for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ .*

Let it be noted that, while the velocity fields are generally assumed to be locally Lipschitz with respect to  $x \in \mathbb{R}^d$  in [14] and other references in the literature dealing with compactly supported measures, the global condition written in (CE)-(ii) is necessary for the derivation of quantitative stability estimates in  $(\mathcal{P}_p(\mathbb{R}^d), W_p(\cdot, \cdot))$ . In the following definition, we recall the notion of *flows of homeomorphisms* generated by a Carathéodory vector field.

**Definition 2.16** (Flows of homeomorphisms). *Let  $v : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  be a velocity field satisfying hypotheses (CE). Then, we denote by  $(\Phi_{(\tau, t)}^v(\cdot))_{\tau, t \in [0, T]} \subset C^0(\mathbb{R}^d, \mathbb{R}^d)$  the unique semigroup of homeomorphisms that solve the Cauchy problems*

$$\Phi_{(\tau, t)}^v(x) = x + \int_\tau^t v(s, \Phi_{(\tau, s)}^v(x)) ds,$$

*for all times  $\tau, t \in [0, T]$  and every  $x \in \mathbb{R}^d$ .*

Under the Cauchy-Lipschitz assumptions of Hypotheses **(CE)**, the following strong well-posedness result holds for solutions of continuity equations in  $(\mathcal{P}_p(\mathbb{R}^d), W_p(\cdot, \cdot))$ . In what follows given some  $q \in [1, +\infty]$ , we will frequently write  $\|\cdot\|_q := \|\cdot\|_{L^q([0, T])}$  for the sake of conciseness.

**Theorem 2.17** (Classical well-posedness of continuity equations). *Let  $(\tau, \mu_\tau) \in [0, T] \times \mathcal{P}_p(\mathbb{R}^d)$  and  $v : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  be a velocity field satisfying Hypotheses **(CE)**-(i). Then, the Cauchy problem (2.9) admits forward solutions  $\mu(\cdot) \in \text{AC}([\tau, T], \mathcal{P}_p(\mathbb{R}^d))$ , and there exists a constant  $c_p > 0$  which only depends on the magnitudes of  $p, \mathcal{M}_p(\mu^0)$  and  $\|m(\cdot)\|_1$ , such that*

$$W_p(\mu(t_1), \mu(t_2)) \leq c_p \int_{t_1}^{t_2} m(s) ds, \quad (2.10)$$

for all times  $0 \leq t_1 \leq t_2 \leq T$ . If in addition Hypothesis **(CE)**-(ii) holds, then the solution of (2.9) is unique, globally defined on  $[0, T]$ , and can be represented explicitly as

$$\mu(t) = \Phi_{(\tau, t)}^v(\cdot) \# \mu_\tau, \quad (2.11)$$

for all times  $t \in [0, T]$ .

*Proof.* These statements follow e.g. from a combination of several standard results from [7, Section 8.1] along with the moment estimates displayed in Proposition 2.21 below.  $\square$

Definition 2.16 and Theorem 2.17 inform us together that, in the Cauchy-Lipschitz setting, solutions of continuity equations can be explicitly written as the transports of the initial datum along the characteristic curves generated by the velocity field. When the latter is less regular, it is still possible to give a rigorous meaning to this intuition by leveraging the concept of *superposition measure*. In what follows, we denote by  $\Sigma_T := C^0([0, T], \mathbb{R}^d)$  the space of continuous arcs from  $[0, T]$  into  $\mathbb{R}^d$ .

**Definition 2.18** (Superposition measures). *An element  $\boldsymbol{\eta} \in \mathcal{P}(\mathbb{R}^d \times \Sigma_T)$  is called a superposition measure associated with a Lebesgue-Borel velocity field  $v : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  from time  $\tau \in [0, T]$  if it is concentrated on the set of pairs  $(x, \sigma) \in \mathbb{R}^d \times \text{AC}([0, T], \mathbb{R}^d)$  satisfying*

$$\sigma(t) = x + \int_\tau^t v(s, \sigma(s)) ds \quad (2.12)$$

for all times  $t \in [\tau, T]$ .

A direct link can be provided between superposition measures and solutions of (2.9) under the action of the so-called *evaluation map*, which is defined for all times  $t \in [0, T]$  by

$$\mathbf{e}_t : (x, \sigma) \in \mathbb{R}^d \times \Sigma_T \mapsto \sigma(t) \in \mathbb{R}^d.$$

More precisely, it can be checked that if  $\boldsymbol{\eta} \in \mathcal{P}(\mathbb{R}^d \times \Sigma_T)$  is a superposition measure associated with a velocity field  $v : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  such that  $(\pi_{\mathbb{R}^d})_\# \boldsymbol{\eta} = \mu_\tau$  – where  $\pi_{\mathbb{R}^d} : (x, \sigma) \in \mathbb{R}^d \times \Sigma_T \mapsto x \in \mathbb{R}^d$  stands for the projection onto the space component –, and if the following local integrability condition

$$\int_\tau^T \int_{\mathbb{R}^d \times \Sigma_T} \mathbf{1}_K(\sigma(t)) |v(t, \sigma(t))| d\boldsymbol{\eta}(x, \sigma) dt < +\infty$$

holds for each compact set  $K \subset \mathbb{R}^d$ , then the curve defined by  $\mu(t) := (\mathbf{e}_t)_\# \boldsymbol{\eta}$  for all times  $t \in [\tau, T]$  is a solution of (2.9). We recall in the following theorem the converse of this statement – colloquially known in the literature as the *superposition principle* –, for which we refer e.g. to [4, Theorem 3.4].

**Theorem 2.19** (Superposition principle). *Let  $(\tau, \mu_\tau) \in [0, T] \times \mathcal{P}_p(\mathbb{R}^d)$  and  $\mu(\cdot) \in \text{AC}([\tau, T], \mathcal{P}_p(\mathbb{R}^d))$  be a solution of (2.9) driven by a Lebesgue-Borel velocity field  $v : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  satisfying*

$$\int_\tau^T \int_{\mathbb{R}^d} \frac{|v(t, x)|}{1 + |x|} d\mu(t)(x) dt < +\infty. \quad (2.13)$$

Then, there exists a superposition measure  $\boldsymbol{\eta}_\mu \in \mathcal{P}(\mathbb{R}^d \times \Sigma_T)$  in the sense of Definition 2.18 such that  $(\mathbf{e}_t)_\# \boldsymbol{\eta}_\mu = \mu(t)$  for all times  $t \in [\tau, T]$ .

Throughout the remainder of this section, we derive a series of useful moment and stability estimates for solutions of (2.9) driven by Lebesgue-Borel velocity fields  $v : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ , under the additional assumption that there exists a map  $m(\cdot) \in L^1([0, T], \mathbb{R}_+)$  such that

$$|v(t, x)| \leq m(t)(1 + |x|) \quad (2.14)$$

holds for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$  and  $\mu(t)$ -almost every  $x \in \mathbb{R}^d$ . While some of the underlying techniques were already explored in our previous work [14], the results below are new, proven under less restrictive regularity assumptions, and requested several additional arguments. We start by stating an adaptation of a result established in [14, Lemma 1], whose proof<sup>1</sup> follows straightforwardly from the latter up to some minor technical adjustments.

**Lemma 2.20** (Superposition plans inducing optimal transports). *Consider two curves of measures  $\mu(\cdot), \nu(\cdot) \in \text{AC}([\tau, T], \mathcal{P}_p(\mathbb{R}^d))$  driven by the Lebesgue-Borel velocity fields  $v, w : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  complying with the pointwise sublinearity estimates*

$$|v(t, x)| \leq m(t)(1 + |x|) \quad \text{and} \quad |w(t, y)| \leq m(t)(1 + |y|)$$

for  $\mathcal{L}^1$ -almost every  $t \in [\tau, T]$  and  $\mu(t) \times \nu(t)$ -almost every  $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$ . Moreover, let  $\eta_\mu, \eta_\nu \in \mathcal{P}(\mathbb{R}^d \times \Sigma_T)$  be two superposition measures given by Theorem 2.19, such that

$$\mu(t) = (\mathbf{e}_t)_\# \eta_\mu \quad \text{and} \quad \nu(t) = (\mathbf{e}_t)_\# \eta_\nu$$

for all times  $t \in [\tau, T]$ . Then, for every  $\gamma_\tau \in \Gamma_o(\mu(\tau), \nu(\tau))$ , there exists  $\hat{\eta}_{\mu, \nu} \in \Gamma(\eta_\mu, \eta_\nu)$  such that

$$(\pi_{\mathbb{R}^d}, \pi_{\mathbb{R}^d})_\# \hat{\eta}_{\mu, \nu} = \gamma_\tau \quad \text{and} \quad (\mathbf{e}_t, \mathbf{e}_t)_\# \hat{\eta}_{\mu, \nu} \in \Gamma_o(\mu(t), \nu(t))$$

for all times  $t \in [\tau, T]$ .

In what follows, we leverage the general superposition results of Theorem 2.19 and Lemma 2.20 to prove moment and equi-integrability inequalities for solutions of (2.9) in Proposition 2.21, as well as two stability estimates with respect to initial data and driving fields in Proposition 2.23. For the sake of readability, we postpone the proof of these results to Appendices C and D respectively.

**Proposition 2.21** (Moment and equi-integrability estimates). *Let  $(\tau, \mu_\tau) \in \mathcal{P}_p(\mathbb{R}^d)$  be given and  $\mu(\cdot) \in \text{AC}([\tau, T], \mathcal{P}_p(\mathbb{R}^d))$  be a solution of (2.9) driven by a Lebesgue-Borel velocity field  $v : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  satisfying the sublinearity estimate (2.14). Then, the following moment bound*

$$\mathcal{M}_p(\mu(t)) \leq C_p \left( \mathcal{M}_p(\mu_\tau) + \int_\tau^t m(s) ds \right) \exp \left( C'_p \|m(\cdot)\|_{L^1([\tau, t])}^p \right), \quad (2.15)$$

holds for all times  $t \in [\tau, T]$ , where the constants  $C_p, C'_p > 0$  are given explicitly by

$$C_p := 2^{(p-1)/p} \quad \text{and} \quad C'_p := \frac{2^{p-1}}{p}. \quad (2.16)$$

In addition for every  $R > 0$ , the following uniform equi-integrability estimate

$$\sup_{t \in [\tau, T]} \int_{\{x \text{ s.t. } |x| \geq R\}} |x|^p d\mu(t)(x) \leq C_T^p \int_{\{x \text{ s.t. } |x| \geq R/C_T - 1\}} (1 + |x|)^p d\mu_\tau(x), \quad (2.17)$$

holds with  $C_T := \max \{1, \|m(\cdot)\|_1\} \exp(\|m(\cdot)\|_1)$ .

**Remark 2.22** (A refined moment inequality). *In the sequel, we will often use the fact that, if a velocity field  $v : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  satisfies a slightly more general sublinearity inequality of the form*

$$|v(t, x)| \leq m(t) \left( 1 + |x| + M(t) \right),$$

<sup>1</sup>In [14], the proof of Lemma 1 contains a small caveat in the definition of the coercive functional used to establish the relative narrow compactness of a relevant sequence of superposition measures. Nevertheless, the conclusion remains correct up to setting the latter to  $+\infty$  outside the closed set of pairs  $(x, \sigma) \in \mathbb{R}^d \times \Sigma_T$  satisfying  $|\dot{\sigma}(t)| \leq m(t)(1 + |\sigma(t)|)$ .

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$  and all  $x \in \mathbb{R}^d$ , where  $M(\cdot) \in L^\infty([0, T], \mathbb{R}_+)$  is a priori given, then the corresponding curve of measures  $\mu(\cdot) \in \text{AC}([\tau, T], \mathcal{P}_p(\mathbb{R}^d))$  is such that

$$\mathcal{M}_p(\mu(t)) \leq C_p \left( \mathcal{M}_p(\mu_\tau) + \int_\tau^t m(s)(1 + M(s)) ds \right) \exp \left( C'_p \|m(\cdot)\|_{L^1([\tau, t])}^p \right),$$

for all times  $t \in [\tau, T]$ . The latter inequality can be established by repeating verbatim the arguments detailed in Appendix C while replacing  $m(\cdot)$  by  $m(\cdot)(1 + M(\cdot))$ .

**Proposition 2.23** (Two general stability estimates for continuity equations). *Let  $\mu_\tau, \nu_\tau \in \mathcal{P}_p(\mathbb{R}^d)$  and  $\mu(\cdot), \nu(\cdot) \in \text{AC}([\tau, T], \mathcal{P}_p(\mathbb{R}^d))$  be two solutions of (2.9), driven respectively by a Carathéodory velocity field  $v : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  satisfying Hypotheses (CE), and by a Lebesgue-Borel velocity field  $w : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  complying with the pointwise sublinearity estimate*

$$|w(t, y)| \leq m(t)(1 + |y|), \quad (2.18)$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$  and  $\nu(t)$ -almost every  $y \in \mathbb{R}^d$ . Then, if the map defined by

$$t \in [\tau, T] \mapsto \|v(t) - w(t)\|_{L^\infty(\mathbb{R}^d, \mathbb{R}^d; \nu(t))}$$

is Lebesgue integrable, the following global stability estimate

$$W_p(\mu(t), \nu(t)) \leq C_p \left( W_p(\mu_\tau, \nu_\tau) + \int_\tau^t \|v(s) - w(s)\|_{L^\infty(\mathbb{R}^d, \mathbb{R}^d; \nu(s))} ds \right) \exp \left( C'_p \|l(\cdot)\|_{L^1([\tau, t])}^p \right), \quad (2.19)$$

holds for all times  $t \in [\tau, T]$ , wherein  $C_p, C'_p > 0$  are defined in (2.16). More generally, under our assumptions, the application

$$t \in [\tau, T] \mapsto \|v(t) - w(t)\|_{L^\infty(B(0, R), \mathbb{R}^d; \nu(t))},$$

is Lebesgue integrable for every  $R > 0$ , and the following localised stability estimate

$$\begin{aligned} W_p(\mu(t), \nu(t)) \leq C_p \left( W_p(\mu_\tau, \nu_\tau) + \int_\tau^t \|v(s) - w(s)\|_{L^\infty(B(0, R), \mathbb{R}^d; \nu(s))} ds + \mathcal{E}_\nu(\tau, t, R) \right) \\ \times \exp \left( C'_p \|l(\cdot)\|_{L^1([\tau, t])}^p \right), \end{aligned} \quad (2.20)$$

holds for all times  $t \in [\tau, T]$ , where the additional error term is given by

$$\mathcal{E}_\nu(\tau, t, R) := 2 \|m(\cdot)\|_{L^1([\tau, t])} (1 + C_T) \left( \int_{\{y \text{ s.t. } |y| \geq R/C_T - 1\}} (1 + |y|)^p d\nu_\tau(y) \right)^{1/p}. \quad (2.21)$$

with  $C_T := \max\{1, \|m(\cdot)\|_1\} \exp(\|m(\cdot)\|_1)$ .

**Remark 2.24** (Comparison with the estimates of [14]). *The stability estimates displayed in Proposition 2.23 above improve on those of [14] in the two following ways. Firstly, the global inequality (2.19) now holds for a general Lebesgue-Borel velocity field  $w : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  satisfying the sublinearity inequality (2.14), without requiring the latter to be Carathéodory. Secondly, the inequality (2.20) involving the quantitative error term  $\mathcal{E}_\nu(\tau, t, R)$  is completely new to the best of our knowledge, and will prove crucial to palliate the fact that the metric of compact convergence  $d_{cc}(\cdot, \cdot)$  only grants access to local and non-uniform discrepancy estimates between functions.*

We end this preliminary section by recalling the definition of *continuity inclusions* in Wasserstein spaces. The latter was introduced by the authors of the present manuscript in [14] as a natural set-valued generalisation of continuity equations. Indeed, its definition adopts the well-established viewpoint that solutions of differential inclusions should be understood as absolutely continuous curves whose derivatives are measurable selections of admissible velocities, see e.g. [9, Chapter 10].

**Definition 2.25** (Continuity inclusions in  $(\mathcal{P}_p(\mathbb{R}^d), W_p(\cdot, \cdot))$ ). Let  $V : [0, T] \times \mathcal{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d, \mathbb{R}^d)$  be a set-valued map. We say that a curve  $\mu(\cdot) \in \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d))$  solves the continuity inclusion

$$\partial_t \mu(t) \in -\text{div}_x(V(t, \mu(t))\mu(t)) \quad (2.22)$$

if there exists an  $\mathcal{L}^1$ -measurable selection  $t \in [0, T] \mapsto v(t) \in V(t, \mu(t)) \subset C^0(\mathbb{R}^d, \mathbb{R}^d)$  such that the trajectory-selection pair  $(\mu(\cdot), v(\cdot))$  is a solution of the continuity equation

$$\partial_t \mu(t) + \text{div}_x(v(t)\mu(t)) = 0,$$

in the sense of distributions.

### 3 Existence à la Peano in the Carathéodory framework

In this section, we establish a general existence result for set-valued Cauchy problems of the form

$$\begin{cases} \partial_t \mu(t) \in -\text{div}_x(V(t, \mu(t))\mu(t)), \\ \mu(0) = \mu^0, \end{cases} \quad (3.1)$$

formulated in the Wasserstein space  $(\mathcal{P}_p(\mathbb{R}^d), W_p(\cdot, \cdot))$ , under the Carathéodory regularity assumptions listed below. Our strategy is based on a careful adaptation of the semi-discrete Euler scheme due to Filippov for differential equations, see e.g. [42, Chapter 1 – Theorem 1]. To the best of our knowledge, this approach is completely new, as it does not seem to have been investigated for differential inclusions even in finite-dimensional vector spaces. In the context of measure dynamics, the latter can also be seen as a kind of relative to the methods developed e.g. in [31, 69, 70]. We would also like to stress that our proof strategy greatly improves on several related compactness arguments explored in the literature of meanfield control, see e.g. [14, 46], thanks to the weak  $L^1$ -compactness criterion of Theorem 2.1.

#### Hypotheses (P).

- (i) The set-valued map  $(t, \mu) \in [0, T] \times \mathcal{P}_p(\mathbb{R}^d) \rightrightarrows V(t, \mu) \subset C^0(\mathbb{R}^d, \mathbb{R}^d)$  is Carathéodory with nonempty closed and convex images.
- (ii) There exists a map  $m(\cdot) \in L^1([0, T], \mathbb{R}_+)$  such that for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ , any  $\mu \in \mathcal{P}_p(\mathbb{R}^d)$  and each  $v \in V(t, \mu)$ , there holds

$$|v(x)| \leq m(t)(1 + |x| + \mathcal{M}_p(\mu)).$$

for all  $x \in \mathbb{R}^d$ .

- (iii) For all compact sets  $\mathcal{K} \subset \mathcal{P}_p(\mathbb{R}^d)$  and  $K \subset \mathbb{R}^d$ , there exist a map  $l(\cdot) \in L^1([0, T], \mathbb{R}_+)$  and a continuous modulus of continuity  $\omega : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  such that for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ , any  $\mu \in \mathcal{K}$  and each  $v \in V(t, \mu)$ , there holds

$$|v(t, x) - v(t, y)| \leq l(t)\omega(|x - y|)$$

for all  $x, y \in K$ .

- (iv) For every compact set  $\mathcal{K} \subset \mathcal{P}_p(\mathbb{R}^d)$ , there exists a map  $L(\cdot) \in L^1([0, T], \mathbb{R}_+)$  and a continuous modulus of continuity  $\omega : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  such that for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ , any  $\mu, \nu \in \mathcal{K}$  and each  $v \in V(t, \mu)$ , there exists  $w \in V(t, \nu)$  such that

$$d_{\text{cc}}(v, w) \leq L(t)\omega(W_p(\mu, \nu)).$$

Two remarks are in order concerning the previous set of assumptions. First, we stress that the maps  $l(\cdot), L(\cdot) \in L^1([0, T], \mathbb{R}_+)$  and the local continuity moduli  $\omega : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  appearing in Hypotheses **(P)**-(iii) and (iv) are all localised on compact sets. That being said, we did not make this dependence transparent so as to keep the notations as light as possible. Second, as mentioned in the introduction, it is a well-known fact in the theory of differential inclusions that outside of the Cauchy-Lipschitz framework, one must in general impose a convexity assumption on the admissible velocities to establish the existence of solutions, even in finite-dimensional euclidean spaces (see e.g. [8, Chapter 2]).

**Remark 3.1** (Examples of set-valued mapping satisfying our assumptions). *A relevant example of set-valued map  $V : [0, T] \times \mathcal{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d, \mathbb{R}^d)$  satisfying Hypotheses **(P)** is given by the set of admissible velocities of a controlled system of the form*

$$V(t, \mu) := \left\{ v(t, \mu, u) \in C^0(\mathbb{R}^d, \mathbb{R}^d) \text{ s.t. } u \in U \right\}.$$

Therein,  $(U, d_U(\cdot, \cdot))$  is a compact metric space representing admissible control inputs, while  $v : [0, T] \times \mathcal{P}_p(\mathbb{R}^d) \times U \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  is a vector field that is Carathéodory in  $(t, u) \in [0, T] \times U$  as well as locally uniformly continuous in  $(\mu, x) \in \mathcal{P}_p(\mathbb{R}^d) \times \mathbb{R}^d$  with constants that are Lebesgue integrable functions of time. In the simpler case in which the maps  $(\mu, x) \in \mathcal{P}_p(\mathbb{R}^d) \times \mathbb{R}^d \mapsto v(t, \mu, u, x) \in \mathbb{R}^d$  are globally Lipschitz with constants given by Lebesgue integrable functions of time, this control theoretic model also fits Hypotheses **(CI)** of Section 4.

In what follows, we state and prove the main result of this section which provides the existence of solutions to (3.1) under Hypotheses **(P)**.

**Theorem 3.2** (Existence à la Peano for continuity inclusions). *Let  $V : [0, T] \times \mathcal{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d, \mathbb{R}^d)$  be a set-valued map satisfying Hypotheses **(P)**. Then for every  $\mu^0 \in \mathcal{P}_p(\mathbb{R}^d)$ , the Cauchy problem*

$$\begin{cases} \partial_t \mu(t) \in -\operatorname{div}_x(V(t, \mu(t))\mu(t)), \\ \mu(0) = \mu^0, \end{cases}$$

*admits at least one solution  $\mu(\cdot) \in \operatorname{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d))$ .*

The proof of Theorem 3.2 follows a constructive scheme which is split into five steps. We start in Step 1 by constructing a sequence of trajectory-selection pairs solving continuity equations with delayed velocity inclusions, and proceed by showing in Step 2 that the elements of the latter comply with uniform moment, regularity and equi-integrability bounds. In Step 3, we then prove that such estimates imply the existence of suitable weak cluster points for the sequence of trajectory-selection pairs. To conclude the proof, we further establish in Step 4 that the limit curve of measures solves a continuity equation driven by the corresponding velocity selection, and finally prove in Step 5 that the latter is in turn a measurable selection in the set of admissible velocities.

*Proof of Theorem 3.2.* In what follows, our goal is to build a sequence of trajectory-selection pairs  $(\mu_n(\cdot), v_n(\cdot)) \in \operatorname{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d)) \times \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$  solutions of the Cauchy problems

$$\begin{cases} \partial_t \mu_n(t) + \operatorname{div}_x(v_n(t)\mu_n(t)) = 0, \\ \mu_n(0) = \mu^0. \end{cases} \quad (3.2)$$

These latter will be chosen so as to satisfy the delayed pointwise velocity inclusion

$$v_n(t) \in V\left(t, \mu_n\left(t - \frac{T}{n}\right)\right), \quad (3.3)$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$  – where here and in what follows we set  $\mu_n(t) := \mu^0$  for  $t \in [-\frac{T}{n}, 0]$  by convention –, along with the uniform moment and regularity bounds

$$\mathcal{M}_p(\mu_n(t)) \leq \mathcal{C} \quad \text{and} \quad W_p(\mu_n(\tau), \mu_n(t)) \leq c_p \int_\tau^t m(s) ds \quad (3.4)$$

for all times  $0 \leq \tau \leq t \leq T$  and each  $n \geq 1$ . Therein, the constants  $\mathcal{C}, c_p > 0$  only depend on the magnitudes of  $p, \mathcal{M}_p(\mu^0)$  and  $\|m(\cdot)\|_1$ .



**Step 1 – Construction of the sequence.** Given an integer  $n \geq 1$ , we explicitly build the pair  $(\mu_n(\cdot), v_n(\cdot))$  satisfying the aforescribed conditions by performing an induction on  $k \in \{0, \dots, n-1\}$ . First, let  $I_n^0 := [0, \frac{T}{n}]$  and observe that under Hypotheses **(P)**-(i) and (ii), the set-valued map

$$t \in I_n^0 \rightrightarrows V(t, \mu^0) \subset C^0(\mathbb{R}^d, \mathbb{R}^d)$$

is  $\mathcal{L}^1$ -measurable with nonempty and closed images for the topology of local uniform convergence. Hence, it admits a measurable selection  $t \in I_n^0 \mapsto v_n^0(t) \in V(t, \mu^0)$  by Theorem 2.6, and the underlying Carathéodory velocity field  $v_n^0 : I_n^0 \times \mathbb{R}^d \mapsto \mathbb{R}^d$  satisfies the sublinearity estimate

$$|v_n^0(t, x)| \leq m(t) \left( 1 + |x| + \mathcal{M}_p(\mu^0) \right),$$

for  $\mathcal{L}^1$ -almost every  $t \in I_n^0$  and all  $x \in \mathbb{R}^d$ , as a consequence of Hypothesis **(P)**-(ii) along with Lemma 2.13. It complies in particular with Hypothesis **(CE)**-(i), and by Theorem 2.17 applied on the interval  $I := I_n^0$  with  $v(t, x) := v_n^0(t, x)$ , there exists a solution  $\mu_n^0(\cdot) \in \text{AC}(I_n^0, \mathcal{P}_p(\mathbb{R}^d))$  to the Cauchy problem

$$\begin{cases} \partial_t \mu_n^0(t) + \text{div}_x(v_n^0(t) \mu_n^0(t)) = 0, \\ \mu_n^0(0) = \mu^0. \end{cases}$$

By repeating this process for  $k \in \{1, \dots, n-1\}$ , we detail below how one can inductively build a family of trajectory-selection pairs  $(\mu_n^k(\cdot), v_n^k(\cdot)) \in \text{AC}(I_k, \mathcal{P}_p(\mathbb{R}^d)) \times \mathcal{L}(I_k, C^0(\mathbb{R}^d, \mathbb{R}^d))$  defined over the family of time intervals  $I_n^k := [\frac{kT}{n}, \frac{(k+1)T}{n}]$ .

First, note that for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ , the sets  $V(t, \mu_n^{k-1}(t)) \subset C^0(\mathbb{R}^d, \mathbb{R}^d)$  are compact for the topology induced by  $d_{cc}(\cdot, \cdot)$  as a consequence of Hypotheses **(P)**-(ii) and (iii) combined with the Ascoli-Arzelà theorem. Second, under Hypothesis **(P)**-(i), it follows from Lemma 2.10-(c) that

$$t \in I_n^k \rightrightarrows V\left(t, \mu_n^{k-1}\left(t - \frac{T}{n}\right)\right) \subset C^0(\mathbb{R}^d, \mathbb{R}^d)$$

is an  $\mathcal{L}^1$ -measurable set-valued map, whose images are nonempty. Thence, by what precedes, it admits a measurable selection

$$t \in I_n^k \mapsto v_n^k(t) \in V\left(t, \mu_n^{k-1}\left(t - \frac{T}{n}\right)\right). \quad (3.5)$$

Besides by Lemma 2.13 and Hypothesis **(P)**-(ii), the vector field  $v_n^k : I_n^k \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  is Carathéodory and satisfies the sublinearity estimate

$$|v_n^k(t, x)| \leq m(t) \left( 1 + |x| + \mathcal{M}_p(\mu_n^{k-1}(t - \frac{T}{n})) \right) \quad (3.6)$$

for  $\mathcal{L}^1$ -almost every  $t \in I_n^k$ , all  $x \in \mathbb{R}^d$  and each  $k \in \{1, \dots, n-1\}$ . Thus by applying Theorem 2.17 on the time interval  $I := I_n^k$  with  $v(t, x) := v_n^k(t, x)$ , one can subsequently define  $\mu_n^k(\cdot) \in \text{AC}(I_n^k, \mathcal{P}_p(\mathbb{R}^d))$  as being one of the solutions of the Cauchy problem

$$\begin{cases} \partial_t \mu_n^k(t) + \text{div}_x(v_n^k(t) \mu_n^k(t)) = 0, \\ \mu_n^k(\frac{kT}{n}) = \mu_n^{k-1}(\frac{kT}{n}). \end{cases}$$

By classical concatenation properties for solutions of continuity equations (see e.g. [40, Lemma 4.4]), the trajectory-selection pair  $(\mu_n(\cdot), v_n(\cdot)) \in \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d)) \times \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$  defined by

$$\mu_n(t) := \mu_n^k(t) \quad \text{and} \quad v_n(t) := v_n^k(t), \quad (3.7)$$

for  $t \in I_n^k$  and  $k \in \{0, \dots, n-1\}$  solves the Cauchy problem (3.2). Moreover, it can be checked using (3.5) that this trajectory-selection pair satisfies the shifted pointwise inclusion (3.3).

**Step 2 – Uniform moment and regularity estimates.** Our goal now is to establish the uniform regularity and moment bounds posited in (3.4). First, notice that as a consequence of the construction detailed in Step 1 and Proposition 2.21, the curves  $(\mu_n(\cdot)) \subset \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d))$  satisfy

$$\sup_{t \in [0, T]} \mathcal{M}_p(\mu_n(t)) < +\infty$$

for each  $n \geq 1$ . This allows us to apply the moment estimate of Remark 2.22, which in context writes

$$\mathcal{M}_p(\mu_n(t)) \leq C_p \left( \mathcal{M}_p(\mu^0) + \int_0^t m(s) \left( 1 + \mathcal{M}_p(\mu_n(s - \frac{T}{n})) \right) ds \right) \exp \left( C'_p \|m(\cdot)\|_{L^1([0, t])}^p \right), \quad (3.8)$$

for all times  $t \in [0, T]$ , using again the convention that  $\mu_n(t) = \mu^0$  when  $t \in [-\frac{T}{n}, 0]$ . Defining the map  $\hat{m}(\cdot) \in L^1([0, 2T], \mathbb{R}_+)$  as

$$\hat{m}(t) := \begin{cases} m(t) & \text{if } t \in [0, T], \\ 0 & \text{if } t \in [T, 2T], \end{cases} \quad (3.9)$$

one may rewrite the estimate displayed in (3.8) as

$$\begin{aligned} \mathcal{M}_p(\mu_n(t)) &\leq C_p \left( \mathcal{M}_p(\mu^0) + \int_0^t m(s) ds + \int_0^{T/n} m(s) \mathcal{M}_p(\mu^0) ds \right. \\ &\quad \left. + \int_0^{\max\{0, t - T/n\}} m(s + \frac{T}{n}) \mathcal{M}_p(\mu_n(s)) ds \right) \exp \left( C'_p \|m(\cdot)\|_{L^1([0, t])}^p \right) \\ &\leq C_p (1 + \mathcal{M}_p(\mu^0)) \left( 1 + \|m(\cdot)\|_{L^1([0, t])} \right) \exp \left( C'_p \|m(\cdot)\|_{L^1([0, t])}^p \right) \\ &\quad + C_p \left( \int_0^t \hat{m}(s + \frac{T}{n}) \mathcal{M}_p(\mu_n(s)) ds \right) \exp \left( C'_p \|m(\cdot)\|_{L^1([0, t])}^p \right), \end{aligned}$$

where the first inequality follows from a simple change of variable, while the second one stems from the non-negativity of  $\hat{m}(\cdot)$  and  $\mathcal{M}_p(\mu_n(\cdot))$ . Thence, a standard application of Grönwall's lemma yields

$$\begin{aligned} \mathcal{M}_p(\mu_n(t)) &\leq C_p (1 + \mathcal{M}_p(\mu^0)) \left( 1 + \|m(\cdot)\|_{L^1([0, t])} \right) \\ &\quad \times \exp \left( C'_p \|m(\cdot)\|_{L^1([0, t])}^p + C_p \|\hat{m}(\cdot)\|_{L^1([0, t + T/n])} \exp \left( C'_p \|m(\cdot)\|_{L^1([0, t])}^p \right) \right) \end{aligned}$$

for all times  $t \in [0, T]$ . In turn, upon noticing that

$$\|\hat{m}(\cdot)\|_{L^1([0, t + T/n])} \leq \|m(\cdot)\|_1$$

for all times  $t \in [0, T]$  and each  $n \geq 1$  as a consequence of (3.9), there further exists a constant  $\mathcal{C} > 0$  which only depends on the magnitudes of  $p, \mathcal{M}_p(\mu^0)$  and  $\|m(\cdot)\|_1$ , such that

$$\sup_{t \in [0, T]} \mathcal{M}_p(\mu_n(t)) \leq \mathcal{C}$$

for each  $n \geq 1$ . In particular, the velocity fields  $(v_n(\cdot)) \subset \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$  satisfy the uniform sublinearity bounds

$$|v_n(t, x)| \leq (1 + \mathcal{C})m(t)(1 + |x|) \quad (3.10)$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ , all  $x \in \mathbb{R}^d$  and each  $n \geq 1$ . Therefore, by Theorem 2.17, there exists a constant  $c_p > 0$  depending only on the magnitudes of  $p, \mathcal{M}_p(\mu^0)$  and  $\|m(\cdot)\|_1$  such that

$$W_p(\mu_n(\tau), \mu_n(t)) \leq c_p \int_\tau^t m(s) ds$$

for all times  $0 \leq \tau \leq t \leq T$ . To summarise, we have shown that the sequence of curves  $(\mu_n(\cdot)) \subset \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d))$  satisfies the uniform moment and absolute continuity estimates (3.4).

**Step 3 – Relative compactness of the sequence of trajectory-selection pairs.** Next, we show that the regularity estimates of Step 2 yield the relative compactness of the sequence of pairs  $(\mu_n(\cdot), v_n(\cdot)) \subset \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d)) \times \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$  built in Step 1. Indeed, notice first that by invoking the moment inequality (2.15) and the equi-integrability bounds (2.17) of Proposition 2.21, possibly with different constants  $C_T > 0$ , one may infer from Proposition 2.3 the existence of a compact set  $\mathcal{K} \subset \mathcal{P}_p(\mathbb{R}^d)$ , depending only on  $\mu^0$  and  $\|m(\cdot)\|_1$ , such that

$$\mu_n(t) \in \mathcal{K} \quad (3.11)$$

for all times  $t \in [0, T]$ . Whence, taking into account the uniform regularity estimate (3.4) satisfied by the sequence  $(\mu_n(\cdot)) \subset \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d))$ , it follows from the Ascoli-Arzelà theorem for complete separable metric spaces (see e.g. [61, Chapter 7 – Theorem 18]) that

$$\sup_{t \in [0, T]} W_p(\mu_n(t), \mu(t)) \xrightarrow{n \rightarrow +\infty} 0, \quad (3.12)$$

for some curve of measures  $\mu(\cdot) \in \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d))$ , along a subsequence that we do not relabel.

Concerning the sequence of velocity selections, observe that for each  $m \geq 1$ , the space restrictions to  $B(0, m)$  of  $(v_n(\cdot)) \subset \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$  form an integrably bounded sequence owing to the sublinearity estimate of (3.10). Besides, by Hypothesis (P)-(iii) combined with (3.11) and the classical Ascoli-Arzelà theorem, there exists a family of compact sets  $(\mathcal{K}_t^m)_{t \in [0, T]} \subset C^0(B(0, m), \mathbb{R}^d)$  such that

$$V\left(t, \mu_n\left(t - \frac{T}{n}\right)\right)_{|B(0, m)} := \left\{ v_{|B(0, m)} \in C^0(B(0, m), \mathbb{R}^d) \text{ s.t. } v \in V\left(t, \mu_n\left(t - \frac{T}{n}\right)\right) \right\} \subset \mathcal{K}_t^m$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$  and each  $n, m \geq 1$ , wherein  $v_{|B(0, m)}$  stands for the restriction of the function  $v \in C^0(\mathbb{R}^d, \mathbb{R}^d)$  to  $B(0, m)$ . Thus, by iteratively applying Theorem 2.1 while performing a standard diagonal argument, we obtain the existence of a Carathéodory vector field  $v(\cdot) \in \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$  such that for every  $R > 0$ , one has that

$$v_n(\cdot) \xrightarrow{n \rightarrow +\infty} v(\cdot) \quad (3.13)$$

weakly in  $L^1([0, T], C^0(B(0, R), \mathbb{R}^d))$ , along a subsequence that depends on  $R > 0$ . In particular, this convergence property implies that

$$\int_0^T \langle \boldsymbol{\nu}(t), v(t) - v_n(t) \rangle_{C^0(B(0, R), \mathbb{R}^d)} dt \xrightarrow{n \rightarrow +\infty} 0 \quad (3.14)$$

whenever  $\boldsymbol{\nu}(\cdot) \in L^\infty([0, T], \mathcal{M}(B(0, R), \mathbb{R}^d))$ , again as a consequence of Theorem 2.1. Besides, it can be straightforwardly verified that

$$\mathcal{M}_p(\mu(t)) \leq \mathcal{C} \quad \text{and} \quad W_p(\mu(\tau), \mu(t)) \leq c_p \int_\tau^t m(s) ds$$

for all times  $t \in [0, T]$  as a consequence of (3.12), while (3.10) and (3.13) yield up to an application of Mazur's lemma (see e.g. [18, Corollary 3.8]) that

$$|v(t, x)| \leq (1 + \mathcal{C})m(t)(1 + |x|), \quad (3.15)$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$  and all  $x \in \mathbb{R}^d$ .

**Step 4 – Dynamics of the limit trajectory-selection pair.** At this stage, we need to show that the limit trajectory-selection pair  $(\mu(\cdot), v(\cdot)) \in \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d)) \times \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$  stemming from the compactness argument of Step 3 is a distributional solution of the Cauchy problem

$$\begin{cases} \partial_t \mu(t) + \text{div}_x(v(t)\mu(t)) = 0, \\ \mu(0) = \mu^0. \end{cases} \quad (3.16)$$

To begin with, note that it trivially follows from (3.12) along with the construction detailed in Step 1 that  $\mu(0) = \mu^0$ . Thus, there remains to show that one may pass to the limit as  $n \rightarrow +\infty$  in the weak formulation of (3.2), of which we remind that it is given by

$$\int_0^T \int_{\mathbb{R}^d} \left( \partial_t \phi(t, x) + \langle \nabla_x \phi(t, x), v_n(t, x) \rangle \right) d\mu_n(t)(x) dt = 0 \quad (3.17)$$

for all  $\phi \in C_c^\infty((0, T) \times \mathbb{R}^d)$  and each  $n \geq 1$ . By leveraging the uniform convergence property of (3.12) together with Proposition 2.3, it can be checked straightforwardly that

$$\int_0^T \int_{\mathbb{R}^d} \partial_t \phi(t, x) d\mu_n(t)(x) dt \xrightarrow{n \rightarrow +\infty} \int_0^T \int_{\mathbb{R}^d} \partial_t \phi(t, x) d\mu(t)(x) dt. \quad (3.18)$$

In addition, upon noting that the convergence result provided in [7, Theorem 12.1.1] for bounded Carathéodory vector fields is still valid for integrably bounded ones – up to a relevant adaptation in the spirit of [7, Lemma 5.1.7] –, it follows from (3.15) combined with (3.12) and Proposition 2.3 that

$$\int_0^T \int_{\mathbb{R}^d} \langle \nabla_x \phi(t, x), v(t, x) \rangle d\mu_n(t)(x) dt \xrightarrow{n \rightarrow +\infty} \int_0^T \int_{\mathbb{R}^d} \langle \nabla_x \phi(t, x), v(t, x) \rangle d\mu(t)(x) dt. \quad (3.19)$$

Thus, by merging (3.18) and (3.19), one can check that in order to recover the weak form of the dynamics in (3.16) by letting  $n \rightarrow +\infty$  in (3.17), there only remains to show that

$$\int_0^T \int_{\mathbb{R}^d} \langle \nabla_x \phi(t, x), v(t, x) - v_n(t, x) \rangle d\mu_n(t)(x) dt \xrightarrow{n \rightarrow +\infty} 0 \quad (3.20)$$

for each  $\phi \in C_c^\infty((0, T) \times \mathbb{R}^d, \mathbb{R})$ , possibly up to a subsequence. To do so, we consider test functions of the form

$$\phi(t, x) := \zeta(t)\psi(x)$$

for all  $(t, x) \in [0, T] \times \mathbb{R}^d$  and some  $(\zeta, \psi) \in C_c^\infty((0, T), \mathbb{R}) \times C_c^\infty(\mathbb{R}^d, \mathbb{R})$ , whose linear span is dense in  $C_c^\infty((0, T) \times \mathbb{R}^d, \mathbb{R})$  (see e.g. [7, Chapter 8]). We then let  $R_\psi > 0$  be such that  $\text{supp}(\psi) \subset B(0, R_\psi)$ , and note that by Step 3, there exists a subsequence of  $(v_n(\cdot)) \subset \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$  that we do not relabel for which (3.13) holds in  $L^1([0, T], C^0(B(0, R), \mathbb{R}^d))$ , with

$$R := (1 + R_\psi) \max \left\{ 1, (1 + \mathcal{C}) \|m(\cdot)\|_1 \right\} \exp \left( (1 + \mathcal{C}) \|m(\cdot)\|_1 \right). \quad (3.21)$$

In that case, we claim that

$$\sup_{\sigma \in \text{AC}([0, T], \mathbb{R}^d)} \left\{ \left| \int_0^T \zeta(t) \langle \nabla \psi(\sigma(t)), v(t, \sigma(t)) - v_n(t, \sigma(t)) \rangle dt \right| \text{ s.t. } \dot{\sigma}(t) = v_n(t, \sigma(t)) \right. \\ \left. \text{for } \mathcal{L}^1\text{-almost every } t \in [0, T] \right\} \xrightarrow{n \rightarrow +\infty} 0. \quad (3.22)$$

Indeed, suppose by contradiction that there exists some  $\varepsilon > 0$ , a subsequence of Carathéodory vector fields  $(v_{n_k}(\cdot)) \subset \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$  and a sequence of curves  $(\sigma_k(\cdot)) \subset \text{AC}([0, T], \mathbb{R}^d)$  such that

$$\dot{\sigma}_k(t) = v_{n_k}(t, \sigma_k(t))$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ , and

$$\left| \int_0^T \zeta(t) \langle \nabla \psi(\sigma_k(t)), v(t, \sigma_k(t)) - v_{n_k}(t, \sigma_k(t)) \rangle dt \right| \geq \varepsilon \quad (3.23)$$

for each  $k \geq 1$ . Then, there must exist a sequence  $(\tau_k) \subset \text{supp}(\zeta)$  such that  $\sigma_k(\tau_k) \in B(0, R_\psi)$  for every  $k \geq 1$ , which together with the estimates of (3.10) and Grönwall's lemma implies that

$$|\sigma_k(t)| \leq R \quad \text{and} \quad |\sigma_k(t) - \sigma_k(\tau)| \leq (1 + \mathcal{C})(1 + R) \int_\tau^t m(s) ds$$

for all times  $0 \leq \tau \leq t \leq T$  and each  $k \geq 1$ , where  $R > 0$  is given as in (3.21). Thence, by the Ascoli-Arzelà theorem, there exists a curve  $\sigma(\cdot) \in \text{AC}([0, T], \mathbb{R}^d)$  such that

$$\sup_{t \in [0, T]} |\sigma(t) - \sigma_k(t)| \xrightarrow{k \rightarrow +\infty} 0, \quad (3.24)$$

along a subsequence which we do not relabel. Besides, it follows from Hypothesis **(P)**-(iii) and Lebesgue's dominated convergence theorem that

$$\sup_{n \geq 1} \int_0^T |v_n(t, \sigma_k(t)) - v_n(t, \sigma(t))| dt \xrightarrow{k \rightarrow +\infty} 0, \quad (3.25)$$

which together with (3.23) further implies that

$$\left| \int_0^T \zeta(t) \langle \nabla \psi(\sigma(t)), v(t, \sigma(t)) - v_{n_k}(t, \sigma(t)) \rangle dt \right| \geq \frac{\varepsilon}{2}, \quad (3.26)$$

whenever  $k \geq 1$  is sufficiently large. However, by the weak-compactness property of (3.14) applied to the mapping  $t \in [0, T] \mapsto \nu(t) := \zeta(t) \nabla \psi \cdot \delta_{\sigma(t)} \in \mathcal{M}(B(0, R), \mathbb{R}^d)$ , it necessarily hold that

$$\int_0^T \zeta(t) \langle \nabla \psi(\sigma(t)), v(t, \sigma(t)) - v_{n_k}(t, \sigma(t)) \rangle dt \xrightarrow{k \rightarrow +\infty} 0 \quad (3.27)$$

along a subsequence that we do not relabel, which contradicts (3.26).

To conclude, we observe that by Theorem 2.19, there exists a sequence of superposition measures  $(\eta_n) \subset \mathcal{P}(\mathbb{R}^d \times \Sigma_T)$  associated with  $(v_n(\cdot)) \subset \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$  which satisfy

$$(\pi_{\mathbb{R}^d})_{\#} \eta_n = \mu^0 \quad \text{and} \quad (\mathbf{e}_t)_{\#} \eta_n = \mu_n(t),$$

for all times  $t \in [0, T]$  and each  $n \geq 1$ . Then, by (3.22) and Fubini's theorem, one can check that

$$\begin{aligned} & \left| \int_0^T \zeta(t) \int_{\mathbb{R}^d} \langle \nabla \psi(x), v(t, x) - v_n(t, x) \rangle d\mu_n(t)(x) dt \right| \\ &= \left| \int_0^T \int_{\mathbb{R}^d \times \Sigma_T} \zeta(t) \langle \nabla \psi(\sigma(t)), v(t, \sigma(t)) - v_n(t, \sigma(t)) \rangle d\eta_n(x, \sigma) dt \right| \\ &\leq \sup_{(x, \sigma) \in \text{supp}(\eta_n)} \left| \int_0^T \zeta(t) \langle \nabla \psi(\sigma(t)), v(t, \sigma(t)) - v_n(t, \sigma(t)) \rangle dt \right| \\ &\leq \sup_{\sigma \in \text{AC}([0, T], \mathbb{R}^d)} \left\{ \left| \int_0^T \zeta(t) \langle \nabla \psi(\sigma(t)), v(t, \sigma(t)) - v_n(t, \sigma(t)) \rangle dt \right| \text{ s.t.} \right. \\ &\quad \left. \dot{\sigma}(t) = v_n(t, \sigma(t)) \text{ for } \mathcal{L}^1\text{-almost every } t \in [0, T] \right\} \xrightarrow{n \rightarrow +\infty} 0 \end{aligned}$$

as a consequence of (3.22), which then yields (3.20). Thence, we have established the trajectory-selection pair  $(\mu(\cdot), v(\cdot)) \in \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d)) \times \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$  is a solution of (3.16).

**Step 5 – Admissibility of the limit velocity selection.** In this last step, we conclude the proof of Theorem 3.2 by showing that the limit Carathéodory vector field  $v(\cdot) \in \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$  is an admissible velocity selection for (3.2), namely

$$v(t) \in V(t, \mu(t)) \quad (3.28)$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ . To this end, we recall that the elements of the sequence  $(\mu_n(\cdot)) \subset \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d))$  are uniformly equi-continuous by (3.4), which implies that for every  $\delta > 0$ , there exists an integer  $N_\delta \geq 1$  such that

$$W_p\left(\mu_n\left(t - \frac{T}{n}\right), \mu_n(t)\right) \leq \delta,$$

for all times  $t \in [0, T]$ , whenever  $n \geq N_\delta$ . Consider now the curves  $\tilde{\mu}_n(\cdot) \in AC([0, T], \mathcal{P}_p(\mathbb{R}^d))$  defined by  $\tilde{\mu}_n(t) := \mu_n(t - \frac{T}{n})$  for all times  $t \in [0, T]$ , and notice that

$$\sup_{t \in [0, T]} W_p(\tilde{\mu}_n(t), \mu(t)) \xrightarrow{n \rightarrow +\infty} 0 \quad (3.29)$$

owing to what precedes, and also

$$v_n(t) \in V(t, \tilde{\mu}_n(t)), \quad (3.30)$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ . At this point, we recall that the set-valued map  $t \in [0, T] \rightrightarrows V(t, \mu(t)) \subset C^0(\mathbb{R}^d, \mathbb{R}^d)$  has nonempty and compact images by Hypotheses **(P)**-(ii) and (iii), and that it is  $\mathcal{L}^1$ -measurable by Lemma 2.10-(c). Moreover, observe that as a consequence of Hypothesis **(P)**-(iv) together with (3.11), there exists a map  $l(\cdot) \in L^1([0, T], \mathbb{R}_+)$  and a continuous modulus  $\omega : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ , both possibly depending on  $\mathcal{K} \subset \mathcal{P}_p(\mathbb{R}^d)$ , such that

$$V(t, \mu(t)) \cap \left\{ w \in C^0(\mathbb{R}^d, \mathbb{R}^d) \text{ s.t. } d_{cc}(v_n(t), w) \leq l(t) \omega\left(W_p(\tilde{\mu}_n(t), \mu(t))\right) \right\}$$

is nonempty for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ . Hence, it follows from Lemma 2.10-(a) that for each  $n \geq 1$ , there exists a measurable selection

$$t \in [0, T] \mapsto \tilde{v}_n(t) \in V(t, \mu(t)) \quad (3.31)$$

which is such that

$$d_{cc}(\tilde{v}_n(t), v_n(t)) \leq l(t) \omega\left(W_p(\tilde{\mu}_n(t), \mu(t))\right)$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ . By Lebesgue's dominated convergence theorem combined with (3.12) and (3.13), this further implies that

$$\tilde{v}_n(\cdot) \xrightarrow{n \rightarrow +\infty} v(\cdot) \quad (3.32)$$

weakly in  $L^1([0, T], C^0(B(0, R), \mathbb{R}^d))$ , along possibly different subsequences depending on  $R > 0$ . To conclude, we finally consider the function sets defined by

$$\mathcal{V}_R := \left\{ w(\cdot) \in L^1([0, T], C^0(B(0, R), \mathbb{R}^d)) \text{ s.t. } w(t) \in V(t, \mu(t))|_{B(0, R)} \text{ for } \mathcal{L}^1\text{-almost every } t \in [0, T] \right\}$$

for each  $R > 0$ , and note that the latter are closed for the strong topology of  $L^1([0, T], C^0(B(0, R), \mathbb{R}^d))$ . As these are convex by Hypothesis **(P)**-(i), it follows e.g. from [18, Theorem 3.7] that they are also weakly closed in  $L^1([0, T], C^0(B(0, R), \mathbb{R}^d))$ . Thus, by (3.31) and (3.32), we may infer that the maps

$$v(\cdot)|_{B(0, R)} : t \in [0, T] \mapsto v(t)|_{B(0, R)} \in C^0(B(0, R), \mathbb{R}^d)$$

belong to  $\mathcal{V}_R$  for each  $R > 0$ , which equivalently means that  $v(t) \in V(t, \mu(t))$  for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$  and ends the proof of Theorem 3.2.  $\square$

**Remark 3.3** (On the choice of proving Theorem 3.2 by means of an Euler scheme). *To prove the existence of solutions to a continuity inclusion with a Carathéodory right-hand side, a tempting strategy, implemented e.g. in [8, Section 2.1 – Theorem 3] or [53, Theorem 2.9], could be to consider first a sufficiently regular exact or approximate velocity selection  $(t, \mu) \in [0, T] \times \mathcal{P}_p(\mathbb{R}^d) \mapsto v(t, \mu) \in C^0(\mathbb{R}^d, \mathbb{R}^d)$ , and then to show that a continuity equation driven by this latter admits at least a solution.*

*While such a program may work in practice, carrying it out seemed difficult – and perhaps suboptimal – for the following reasons. Firstly, the results ensuring the existence of Carathéodory selections for Carathéodory set-valued maps such as [9, Theorem 9.5.2] do not exist for multifunctions valued in infinite-dimensional spaces. Even if an adequate counterpart were to be found in our context, one would still then need to prove that the corresponding nonlocal continuity equations admit solutions, most likely by means of an Euler scheme. Secondly, even though results providing families of regular approximate selections for upper-semicontinuous set-valued mappings such as [9, Theorem 9.2.1] may still hold for Fréchet instead of Banach spaces, establishing the compactness of the underlying sequences may prove to be challenging, owing to the lack of uniformity in the regularity of said selections.*

## 4 A priori estimates, compactness and relaxation in the Cauchy-Lipschitz framework

In this section, we derive quantitative well-posedness results and discuss some of the fine topological properties of the reachable sets of the set-valued Cauchy problem

$$\begin{cases} \partial_t \mu(t) \in -\operatorname{div}_x \left( V(t, \mu(t)) \mu(t) \right), \\ \mu(0) = \mu^0. \end{cases} \quad (4.1)$$

To this end, we depart from the Carathéodory framework investigated in Section 3, and work under the following stronger Cauchy-Lipschitz assumptions.

### Hypotheses (CI).

(i) *The set-valued map  $(t, \mu) \in [0, T] \times \mathcal{P}_p(\mathbb{R}^d) \rightrightarrows V(t, \mu) \subset C^0(\mathbb{R}^d, \mathbb{R}^d)$  is Carathéodory with nonempty closed images.*

(ii) *There exists a map  $m(\cdot) \in L^1([0, T], \mathbb{R}_+)$  such that for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ , any  $\mu \in \mathcal{P}_p(\mathbb{R}^d)$ , every  $v \in V(t, \mu)$  and all  $x \in \mathbb{R}^d$ , there holds*

$$|v(x)| \leq m(t) \left( 1 + |x| + \mathcal{M}_p(\mu) \right)$$

*for all  $x \in \mathbb{R}^d$ .*

(iii) *There exists a map  $l(\cdot) \in L^1([0, T], \mathbb{R}_+)$  such that for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ , any  $\mu \in \mathcal{P}_p(\mathbb{R}^d)$  and every  $v \in V(t, \mu)$ , there holds*

$$\operatorname{Lip}(v; \mathbb{R}^d) \leq l(t).$$

(iv) *There exists a map  $L(\cdot) \in L^1([0, T], \mathbb{R}_+)$  such that for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ , any  $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$  and each  $v \in V(t, \mu)$ , there exists an element  $w \in V(t, \nu)$  for which*

$$d_{\text{sup}}(v, w) \leq L(t) W_p(\mu, \nu).$$

**Remark 4.1** (Comparison with the main assumptions of [14]). *The study of curves of compactly supported measures generated by locally Lipschitz set-valued maps initiated in [14] is completely included in our subsequent developments, up to some minor technical adjustments. Likewise, we could have opted in the present manuscript for assumptions in which the Lipschitz constants of the dynamics are localised on compact subsets of  $\mathcal{P}_p(\mathbb{R}^d)$ .*

### 4.1 Existence of solutions and Filippov estimates

In this section, we prove a new variant of the Filippov estimates for solutions of the Cauchy problem (4.1). Given an arbitrary measure curve, these latter involve the sum of a local discrepancy term between the velocity of said curve and that of the solution of the inclusion on an arbitrary ball, to which one adds a remainder controlled by the tail of the initial datum of said auxiliary curve.

**Theorem 4.2** (Filippov estimates). *Let  $V : [0, T] \times \mathcal{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d, \mathbb{R}^d)$  be a set-valued map satisfying Hypotheses (CI) and  $\nu(\cdot) \in \operatorname{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d))$  be a solution of the continuity equation*

$$\partial_t \nu(t) + \operatorname{div}_y(w(t)\nu(t)) = 0$$

*driven by a Lebesgue-Borel velocity field  $w : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  satisfying the sublinearity estimate*

$$|w(t, y)| \leq m(t)(1 + |y|), \quad (4.2)$$

*for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$  and  $\nu(t)$ -almost every  $y \in \mathbb{R}^d$ . For every  $R > 0$ , define the localised mismatch function  $\eta_R(\cdot) \in L^1([0, T], \mathbb{R}_+)$  by*

$$\eta_R(t) := \operatorname{dist}_{L^\infty(B(0, R), \mathbb{R}^d; \nu(t))} \left( w(t); V(t, \nu(t)) \right) \quad (4.3)$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ .

Then for every  $\mu^0 \in \mathcal{P}_p(\mathbb{R}^d)$  and each  $R > 0$ , there exists a trajectory-selection pair  $(\mu(\cdot), v(\cdot)) \in \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d)) \times \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$  solution of the Cauchy problem (4.1) which satisfies the distance estimate

$$W_p(\mu(t), \nu(t)) \leq \mathcal{D}_p(t, R) \quad (4.4)$$

for all times  $t \in [0, T]$ , where

$$\mathcal{D}_p(t, R) := C_p \left( W_p(\mu^0, \nu(0)) + \int_0^t \eta_R(s) ds + \mathcal{E}_\nu(t, R) \right) \exp \left( C'_p \|l(\cdot)\|_{L^1([0, t])}^p + \chi_p(t) \right). \quad (4.5)$$

Therein, the constants  $C_p, C'_p > 0$  are as in (2.16), the error term  $\mathcal{E}_\nu(t, R)$  is given explicitly by

$$\mathcal{E}_\nu(t, R) := 2 \|m(\cdot)\|_{L^1([0, t])} (1 + \mathcal{C}_T) \left( \int_{\{y \text{ s.t. } |y| \geq R/\mathcal{C}_T - 1\}} (1 + |y|)^p d\nu(0)(y) \right)^{1/p}$$

for some constant  $\mathcal{C}_T > 0$  that only depends on the magnitudes of  $p, \mathcal{M}_p(\mu^0)$  and  $\|m(\cdot)\|_1$ , and the map  $\chi_p(\cdot) \in L^\infty([0, T], \mathbb{R}_+)$  writes for all  $t \in [0, T]$  as

$$\chi_p(t) := C_p \|L(\cdot)\|_{L^1([0, t])} \exp \left( C'_p \|l(\cdot)\|_{L^1([0, t])}^p \right). \quad (4.6)$$

Moreover, the velocity selection  $t \in [0, T] \mapsto v(t) \in V(t, \mu(t))$  complies with the pointwise estimate

$$\|v(t) - w(t)\|_{L^\infty(B(0, R), \mathbb{R}^d; \nu(t))} \leq \eta_R(t) + L(t) \mathcal{D}_p(t, R) \quad (4.7)$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ .

Before moving on to the proof of Theorem 4.2, we state a technical lemma dealing with chained integral estimates, whose proof is the matter of a straightforward induction argument.

**Lemma 4.3** (A uniform bound on families of functions satisfying recurrent integral estimates). *Let  $m(\cdot) \in L^1([0, T], \mathbb{R}_+)$  and  $f^0, \alpha > 0$  be two given constants. Then, every at most countable family of functions  $(f_n(\cdot)) \subset C^0([0, T], \mathbb{R}_+)$  satisfying  $\|f_0(\cdot)\|_{C^0([0, T], \mathbb{R}_+)} \leq f^0$  as well as the recurrence estimate*

$$f_{n+1}(t) \leq \alpha \left( 1 + \int_0^t m(s) f_n(s) ds \right)$$

for all times  $t \in [0, T]$  and each  $n \geq 0$  is uniformly bounded, with

$$\sup_{n \geq 0} \|f_n(\cdot)\|_{C^0([0, T], \mathbb{R}_+)} \leq (\alpha + f^0) \exp(\alpha \|m(\cdot)\|_1).$$

The proof of Theorem 4.2 is based on a constructive scheme, much like that of Theorem 3.2, and is split into four steps. In Step 1, we detail the initialisation of the underlying induction argument by means of a selection principle applied along  $\nu(\cdot)$ , and proceed in Step 2 by building the whole sequence of trajectory-selection pairs. We then show in Step 3 that the latter is a Cauchy sequence for a suitable extended metric, and finally prove in Step 4 that the corresponding limit pair is a solution of (4.1) which satisfies the stability estimates (4.4) and (4.7).

*Proof of Theorem 4.2.* Our goal in what follows is to build a sequence of trajectory-selection pairs  $(\mu_n(\cdot), v_n(\cdot)) \subset \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d)) \times \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$  whose elements solve the Cauchy problems

$$\begin{cases} \partial_t \mu_n(t) + \text{div}_x(v_n(t) \mu_n(t)) = 0, \\ \mu_n(0) = \mu^0, \end{cases} \quad (4.8)$$

while also satisfying the conditions

$$v_{n+1}(t) \in V(t, \mu_n(t)) \quad \text{and} \quad \mathbf{d}_{\text{sup}}(v_n(t), v_{n+1}(t)) \leq L(t) W_p(\mu_{n-1}(t), \mu_n(t)) \quad (4.9)$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ , and complying with the uniform moment and regularity bounds

$$\mathcal{M}_p(\mu_n(t)) \leq \mathcal{C} \quad \text{and} \quad W_p(\mu_n(\tau), \mu_n(t)) \leq c_p \int_\tau^t m(s) ds \quad (4.10)$$

for all times  $0 \leq \tau \leq t \leq T$  and each  $n \geq 1$ . Therein, the constants  $\mathcal{C}, c_p > 0$  only depend on the magnitudes of  $p, \mathcal{M}_p(\mu^0)$  and  $\|m(\cdot)\|_1$ .



**Step 1 – Initialisation of the sequence.** Observe first that by combining Hypotheses **(CI)**-(*ii*) and (*iii*) along with Theorem 2.14, the admissible velocity sets  $V(t, \nu(t)) \subset C^0(\mathbb{R}^d, \mathbb{R}^d)$  are compact in the topology of uniform convergence on compact sets for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ . Besides under Hypotheses **(CI)**-(*i*) and (*iv*), the set-valued map  $t \in [0, T] \rightrightarrows V(t, \nu(t)) \subset C^0(\mathbb{R}, \mathbb{R}^d)$  is  $\mathcal{L}^1$ -measurable as a consequence of Lemma 2.10-(*c*). Furthermore, it can be checked that

$$(t, v) \in [0, T] \times C^0(\mathbb{R}^d, \mathbb{R}^d) \mapsto \|v - w(t)\|_{L^\infty(B(0, R), \mathbb{R}^d; \nu(t))} \in \mathbb{R}_+$$

is a Carathéodory map for each  $R > 0$ , in the sense that it is  $\mathcal{L}^1$ -measurable with respect to  $t \in [0, T]$  and continuous with respect to  $v \in C^0(\mathbb{R}^d, \mathbb{R}^d)$  for the topology induced by  $d_{cc}(\cdot, \cdot)$ . Thus, recalling the definition (4.3) of the mismatch function  $\eta_R(\cdot)$  and applying Lemma 2.10-(*b*) to the multifunction

$$t \in [0, T] \rightrightarrows V(t, \nu(t)) \cap \left\{ v \in C^0(\mathbb{R}^d, \mathbb{R}^d) \text{ s.t. } \|v - w(t)\|_{L^\infty(B(0, R), \mathbb{R}^d; \nu(t))} = \eta_R(t) \right\},$$

we obtain the existence of an  $\mathcal{L}^1$ -measurable map  $t \in [0, T] \mapsto v_1(t) \in V(t, \nu(t))$  such that

$$\|v_1(t) - w(t)\|_{L^\infty(B(0, R), \mathbb{R}^d; \nu(t))} = \eta_R(t) \quad (4.11)$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ .

Remark now that as a by-product of Hypotheses **(CI)**-(*ii*) and (*iii*) along with Lemma 2.13, the velocity field  $v_1 : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  is Carathéodory and such that

$$|v_1(t, x)| \leq m(t) \left( 1 + |x| + \mathcal{M}_p(\nu(t)) \right) \quad \text{and} \quad \text{Lip}(v_1(t); \mathbb{R}^d) \leq l(t)$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$  and all  $x \in \mathbb{R}^d$ . In particular, it satisfies Hypotheses **(CE)**, and by Theorem 2.17 there exists a unique curve  $\mu_1(\cdot) \in \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d))$  solution of the Cauchy problem

$$\begin{cases} \partial_t \mu_1(t) + \text{div}_x(v_1(t) \mu_1(t)) = 0, \\ \mu_1(0) = \mu^0. \end{cases}$$

Owing to the moment bound of Remark 2.22, the curve  $\mu_1(\cdot)$  further complies with the estimate

$$\mathcal{M}_p(\mu_1(t)) \leq C_p \left( \mathcal{M}_p(\mu^0) + \int_0^t m(s) \left( 1 + \mathcal{M}_p(\nu(s)) \right) ds \right) \exp \left( C'_p \|m(\cdot)\|_{L^1([0, t])}^p \right), \quad (4.12)$$

for all times  $t \in [0, T]$ . Consequently, by Lemma 4.3 and (2.15) of Proposition 2.21 applied to the curve  $\nu(\cdot) \in \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d))$ , there exists a constant  $\mathcal{C} > 0$  that only depends on the magnitudes of  $p, \mathcal{M}_p(\mu^0), \mathcal{M}_p(\nu(0))$  and  $\|m(\cdot)\|_1$ , such that

$$\max \left\{ \mathcal{M}_p(\nu(t)), \mathcal{M}_p(\mu_1(t)) \right\} \leq \mathcal{C} \quad (4.13)$$

for all times  $t \in [0, T]$ . In particular, the velocity field  $v_1 : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  satisfies

$$|v_1(t, x)| \leq (1 + \mathcal{C})m(t)(1 + |x|), \quad (4.14)$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$  and all  $x \in \mathbb{R}^d$ , which combined with (2.10) of Theorem 2.17 yields the existence of a constant  $c_p > 0$  depending only on the magnitudes of  $p, \mathcal{M}_p(\mu^0), \mathcal{M}_p(\nu(0))$  and  $\|m(\cdot)\|_1$ , such that

$$W_p(\mu_1(\tau), \mu_1(t)) \leq c_p \int_\tau^t m(s) ds$$

for all times  $0 \leq \tau \leq t \leq T$ . Moreover, by applying the approximate stability inequality (2.20) of Proposition 2.23 while taking into account the sublinearity estimate (4.14), it further holds that

$$W_p(\mu_1(t), \nu(t)) \leq C_p \left( W_p(\mu^0, \nu(0)) + \int_0^t \eta_R(s) ds + \mathcal{E}_\nu(t, R) \right) \exp \left( C'_p \|l(\cdot)\|_{L^1([0, t])}^p \right) \quad (4.15)$$

for all times  $t \in [0, T]$ , where

$$\mathcal{E}_\nu(t, R) := 2 \|m(\cdot)\|_{L^1([0, t])} (1 + \mathcal{C}_T) \left( \int_{\{s \text{ s.t. } |y| \geq R/\mathcal{C}_T - 1\}} (1 + |y|)^p d\nu(0)(y) \right)^{1/p}$$

with  $\mathcal{C}_T := \max \{ 1, (1 + \mathcal{C}) \|m(\cdot)\|_1 \} \exp \left( (1 + \mathcal{C}) \|m(\cdot)\|_1 \right)$ . This together with the moment estimate of (4.12) concludes the initialisation step of our induction argument.

**Step 2 – Construction of the sequence of trajectory-selection pairs.** At this stage, note that

$$d_{\text{sup}}(v_1(t), w) = \sup_{m \geq 1} \|v_1(t) - w\|_{C^0(B(0,m), \mathbb{R}^d)}$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$  and all  $w \in C^0(\mathbb{R}^d, \mathbb{R}^d)$ , and that by construction, the maps  $(t, w) \in [0, T] \times C^0(\mathbb{R}^d, \mathbb{R}^d) \mapsto \|v_1(t) - w\|_{C^0(B(0,m), \mathbb{R}^d)} \in \mathbb{R}_+$  are Carathéodory for each  $m \geq 1$ . Thence, as a consequence of Lemma 2.10-(a) combined with Hypothesis (CI)-(iv), there exists an  $\mathcal{L}^1$ -measurable map  $t \in [0, T] \mapsto v_2(t) \in V(t, \mu_1(t))$  satisfying

$$d_{\text{sup}}(v_1(t), v_2(t)) \leq L(t)W_p(\mu_1(t), \nu(t))$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ . From Hypotheses (CI)-(ii) and (iii) combined with Lemma 2.13, we may then infer that the velocity field  $v_2 : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  is Carathéodory, and such that

$$|v_2(t, x)| \leq m(t) \left(1 + |x| + \mathcal{M}_p(\mu_1(t))\right) \quad \text{and} \quad \text{Lip}(v_2(t); \mathbb{R}^d) \leq l(t),$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$  and all  $x \in \mathbb{R}^d$ . In particular, it satisfies Hypotheses (CE), and thus generates a unique solution  $\mu_2(\cdot) \in \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d))$  of the corresponding Cauchy problem

$$\begin{cases} \partial_t \mu_2(t) + \text{div}_x(v_2(t)\mu_2(t)) = 0, \\ \mu_2(0) = \mu^0. \end{cases}$$

By the refined moment bound of Remark 2.22 and the estimate (2.19) of Proposition 2.23, it may again be checked that

$$\begin{cases} \mathcal{M}_p(\mu_2(t)) \leq C_p \left( \mathcal{M}_p(\mu^0) + \int_0^t m(s) \left(1 + \mathcal{M}_p(\mu_1(s))\right) ds \right) \exp \left( C'_p \|m(\cdot)\|_{L^1([0,t])}^p \right), \\ W_p(\mu_1(t), \mu_2(t)) \leq C_p \left( \int_0^t d_{\text{sup}}(v_1(s), v_2(s)) ds \right) \exp \left( C'_p \|l(\cdot)\|_{L^1([0,t])}^p \right). \end{cases}$$

As a consequence of Lemma 4.3, this implies in particular that

$$\max \left\{ \mathcal{M}_p(\nu(t)), \mathcal{M}_p(\mu_1(t)), \mathcal{M}_p(\mu_2(t)) \right\} \leq \mathcal{C}$$

for all times  $t \in [0, T]$ , where  $\mathcal{C} > 0$  is the same constant as in Step 1 which only depends on the magnitudes of  $p, \mathcal{M}_p(\mu^0), \mathcal{M}_p(\nu(0))$  and  $\|m(\cdot)\|_1$ .

By repeating this process, we can inductively build a sequence  $(\mu_n(\cdot), v_n(\cdot)) \subset \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d)) \times \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$  of trajectory-selection pairs solution of the Cauchy problems (4.8), whose elements satisfy the conditions of (4.9) as well as the bounds in (4.10). In addition, these latter comply for each  $n \geq 1$  with the uniform sublinearity estimates

$$|v_n(t, x)| \leq (1 + \mathcal{C})m(t)(1 + |x|), \quad (4.16)$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$  and all  $x \in \mathbb{R}^d$ , as well as the distance estimates

$$W_p(\mu_n(t), \mu_{n+1}(t)) \leq C_p \left( \int_0^t d_{\text{sup}}(v_n(s), v_{n+1}(s)) ds \right) \exp \left( C'_p \|l(\cdot)\|_{L^1([0,t])}^p \right), \quad (4.17)$$

for all times  $t \in [0, T]$ .

**Step 3 – Convergence of the sequence of trajectory-selection pairs.** In what follows, we prove that the sequence of pairs  $(\mu_n(\cdot), v_n(\cdot)) \subset \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d)) \times \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$  built in

Step 1 and Step 2 is Cauchy in a suitable sense. First, observe that by merging the distance estimates of (4.9), (4.15) and (4.17), one has that

$$\begin{aligned}
& W_p(\mu_n(t), \mu_{n+1}(t)) \\
& \leq C_p \left( \int_0^t \mathbf{d}_{\text{sup}}(v_n(s_n), v_{n+1}(s_n)) ds_n \right) \exp \left( C'_p \|l(\cdot)\|_{L^1([0,t])}^p \right) \\
& \leq C_p \left( \int_0^t L(s_n) W_p(\mu_{n-1}(s_n), \mu_n(s_n)) ds_n \right) \exp \left( C'_p \|l(\cdot)\|_{L^1([0,t])}^p \right) \\
& \leq C_p^2 \left( \int_0^t L(s_n) \int_0^{s_n} \mathbf{d}_{\text{sup}}(v_{n-1}(s_{n-1}), v_n(s_{n-1})) ds_{n-1} ds_n \right) \exp \left( 2C'_p \|l(\cdot)\|_{L^1([0,t])}^p \right) \\
& \vdots \\
& \leq C_p^n \left( \int_0^t L(s_n) \int_0^{s_n} L(s_{n-1}) \dots \int_0^{s_2} L(s_1) W_p(\mu_1(s_1), \nu(s_1)) ds_1 \dots ds_{n-1} ds_n \right) \exp \left( n C'_p \|l(\cdot)\|_{L^1([0,t])}^p \right) \\
& \leq \frac{C_p^{n+1}}{n!} \left( \int_0^t L(s) ds \right)^n \left( W_p(\mu^0, \nu(0)) + \int_0^t \eta_R(s) ds + \mathcal{E}_\nu(t, R) \right) \exp \left( (n+1) C'_p \|l(\cdot)\|_{L^1([0,t])}^p \right), \\
& = \frac{\chi_p(t)^n}{n!} C_p \left( W_p(\mu^0, \nu(0)) + \int_0^t \eta_R(s) ds + \mathcal{E}_\nu(t, R) \right) \exp \left( C'_p \|l(\cdot)\|_{L^1([0,t])}^p \right)
\end{aligned} \tag{4.18}$$

for all times  $t \in [0, T]$  and each  $n \geq 1$ , where we recall that  $\chi_p(\cdot) \in L^\infty([0, T], \mathbb{R}_+)$  is defined as in (4.6). Whence, for any pair of integers  $m, n \geq 1$ , one may deduce from (4.18) that

$$\begin{aligned}
& \sup_{t \in [0, T]} W_p(\mu_n(t), \mu_{n+m}(t)) \\
& \leq \sup_{t \in [0, T]} \sum_{k=n}^{n+m-1} W_p(\mu_k(t), \mu_{k+1}(t)) \\
& \leq \left( \sum_{k=n}^{n+m-1} \frac{\|\chi_p(\cdot)\|_\infty^k}{k!} \right) C_p \left( W_p(\mu^0, \nu(0)) + \int_0^T \eta_R(s) ds + \mathcal{E}_\nu(T, R) \right) \exp \left( C'_p \|l(\cdot)\|_{L^1([0,t])}^p \right) \xrightarrow{m, n \rightarrow +\infty} 0,
\end{aligned} \tag{4.19}$$

which means that  $(\mu_n(\cdot)) \subset \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d))$  is a Cauchy sequence in  $C^0([0, T], \mathcal{P}_p(\mathbb{R}^d))$ . Noting that the latter is a complete metric space as a consequence e.g. of [61, Chapter 7 – Theorem 12], there exists some  $\mu(\cdot) \in C^0([0, T], \mathcal{P}_p(\mathbb{R}^d))$  such that

$$\sup_{t \in [0, T]} W_p(\mu_n(t), \mu(t)) \xrightarrow{n \rightarrow +\infty} 0. \tag{4.20}$$

Besides, as a consequence of the equi-absolute continuity estimate established in (4.10) for the sequence  $(\mu_n(\cdot)) \subset \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d))$ , the limit curve is absolutely continuous as well, with

$$W_p(\mu(\tau), \mu(t)) \leq c_p \int_\tau^t m(s) ds$$

for all times  $0 \leq \tau \leq t \leq T$ .

Concerning the sequence of Carathéodory vector fields  $(v_n(\cdot)) \subset \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$ , it directly follows from (4.9) together with the estimates in (4.19) that for each  $m, n \geq 1$ , there holds

$$\begin{aligned}
\int_0^T \mathbf{d}_{\text{sup}}(v_n(t), v_{n+m}(t)) dt & \leq \sum_{k=n}^{m+n-1} \int_0^T \mathbf{d}_{\text{sup}}(v_k(t), v_{k+1}(t)) dt \\
& \leq \|L(\cdot)\|_1 \sup_{t \in [0, T]} \sum_{k=n}^{m+n-1} W_p(\mu_{k-1}(t), \mu_k(t)) \xrightarrow{m, n \rightarrow +\infty} 0.
\end{aligned} \tag{4.21}$$

Hence, by Lemma 2.15, there exists a map  $v(\cdot) \in \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$  such that

$$\int_0^T \mathbf{d}_{\text{sup}}(v_n(t), v(t)) dt \xrightarrow{n \rightarrow +\infty} 0. \tag{4.22}$$

This yields in particular the existence of a subsequence  $(v_{n_k}(\cdot)) \subset \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$  satisfying

$$d_{\text{sup}}(v_{n_k}(t), v(t)) \xrightarrow{n \rightarrow +\infty} 0$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ , which combined with (4.16) directly implies that the limit velocity field  $v : [0, T] \times \mathbb{R}^d \mapsto \mathbb{R}^d$  satisfies the sublinearity and regularity estimates

$$|v(t, x)| \leq (1 + \mathcal{C})m(t)(1 + |x|) \quad \text{and} \quad \text{Lip}(v(t); \mathbb{R}^d) \leq l(t), \quad (4.23)$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$  and all  $x \in \mathbb{R}^d$ .

**Step 4 – Properties of the limit trajectory-selection pair.** By combining the estimate (2.4) of Proposition 2.4 with the stability inequality (2.20) in Proposition 2.23 as well as (4.23), it can be straightforwardly deduced from the convergence results of (4.20) and (4.22) that the trajectory-selection pair  $(\mu(\cdot), v(\cdot)) \in \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d)) \times \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$  solves the Cauchy problem

$$\begin{cases} \partial_t \mu(t) + \text{div}_x(v(t)\mu(t)) = 0, \\ \mu(0) = \mu^0. \end{cases} \quad (4.24)$$

The next step in our argument consists in showing that the latter actually solves the continuity inclusion (4.1). By (4.9), the sequence of trajectory-selection pairs is such that

$$v_{n+1}(t) \in V(t, \mu_n(t)) \quad (4.25)$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$  and each  $n \geq 1$ . By (4.20) and (4.22), it further holds that

$$W_p(\mu_n(t), \mu(t)) \xrightarrow{n \rightarrow +\infty} 0, \quad (4.26)$$

for all times  $t \in [0, T]$ , as well as

$$d_{\text{sup}}(v_n(t), v(t)) \xrightarrow{n \rightarrow +\infty} 0, \quad (4.27)$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ , along a subsequence that we do not relabel. Whence, upon combining (4.25) with the pointwise convergence results (4.26) and (4.27) and Hypothesis (CI)-(iv), we obtain

$$v(t) \in V(t, \mu(t))$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ , which in light of Definition 2.25 implies together with (4.24) that  $\mu(\cdot) \in \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d))$  is a solution (4.1), driven by the selection  $t \in [0, T] \mapsto v(t) \in V(t, \mu(t))$ .

Now, there only remains to derive the distance and velocity estimates displayed in (4.4) and (4.7). To this end, note that the chain of pointwise inequalities of (4.18) implies

$$\begin{aligned} W_p(\mu_n(t), \nu(t)) &\leq W_p(\mu_1(t), \nu(t)) + \sum_{k=1}^{n-1} W_p(\mu_k(t), \mu_{k+1}(t)) \\ &\leq \left(1 + \sum_{k=1}^{n-1} \frac{\chi_p(t)^k}{k!}\right) C_p \left(W_p(\mu^0, \nu(0)) + \int_0^t \eta_R(s) ds + \mathcal{E}_\nu(t, R)\right) \exp(C'_p \|\iota(\cdot)\|_1^p) \\ &\leq C_p \left(W_p(\mu^0, \nu(0)) + \int_0^t \eta_R(s) ds + \mathcal{E}_\nu(t, R)\right) \exp\left(C'_p \|\iota(\cdot)\|_1^p + \chi_p(t)\right) \\ &= \mathcal{D}_p(t, R), \end{aligned}$$

for all times  $t \in [0, T]$  and each  $n \geq 1$ , with  $\mathcal{D}_p(\cdot)$  being defined as in (4.5). By letting  $n \rightarrow +\infty$  in the previous expression, we directly obtain (4.4). Concerning the velocity estimates, it holds as a consequence of (4.9), (4.11) and (4.15) that

$$\begin{aligned} \|v_n(t) - w(t)\|_{L^\infty(B(0, R), \mathbb{R}^d; \nu(t))} &\leq \|v_1(t) - w(t)\|_{L^\infty(B(0, R), \mathbb{R}^d; \nu(t))} + \sum_{k=1}^{n-1} d_{\text{sup}}(v_{k+1}(t), v_k(t)) \\ &\leq \eta_R(t) + L(t) \sum_{k=1}^{n-1} W_p(\mu_{k-1}(t), \mu_k(t)) \\ &\leq \eta_R(t) + L(t) \mathcal{D}_p(t, R), \end{aligned}$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$  and each  $n \geq 1$ . We obtain (4.7) by taking the limit as  $n \rightarrow +\infty$  along a suitable subsequence in the previous expression, which concludes the proof of Theorem 4.2.  $\square$

In the following corollary, we state a global version of the Filippov estimates which can be obtained by a simple adaptation of the proof of Theorem 4.2.

**Corollary 4.4** (Global version of Filippov estimates). *Suppose that the assumptions of Theorem 4.2 hold, and in addition that the global mismatch function, defined by*

$$\eta(t) := \text{dist}_{L^\infty(\mathbb{R}^d, \mathbb{R}^d; \nu(t))}(w(t); V(t, \nu(t)))$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ , is Lebesgue integrable. Then for every  $\mu^0 \in \mathcal{P}_p(\mathbb{R}^d)$ , there exists a trajectory-selection pair  $(\mu(\cdot), v(\cdot)) \in \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d)) \times \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$  solution of the Cauchy problem (4.1) which satisfies the distance estimate

$$W_p(\mu(t), \nu(t)) \leq C_p \left( W_p(\mu^0, \nu(0)) + \int_0^t \eta(s) ds \right) \exp \left( C'_p \|l(\cdot)\|_{L^1([0, t])}^p + \chi_p(t) \right),$$

for all times  $t \in [0, T]$ , as well as the velocity estimate

$$\begin{aligned} & \|v(t) - w(t)\|_{L^\infty(\mathbb{R}^d, \mathbb{R}^d; \nu(t))} \\ & \leq \eta(t) + L(t)C_p \left( W_p(\mu^0, \nu(0)) + \int_0^t \eta(s) ds \right) \exp \left( C'_p \|l(\cdot)\|_{L^1([0, t])}^p + \chi_p(t) \right), \end{aligned}$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ .

*Proof.* One can simply repeat the proof strategy of Theorem 4.2 with  $R = +\infty$  and  $\mathcal{E}_\nu(t, R) = 0$ .  $\square$

## 4.2 Compactness of the solution set and relaxation property

In this section, we investigate the topological properties of the solution set associated with the set-valued Cauchy problem (4.1), which is defined by

$$\mathcal{S}_{[0, T]}(\mu^0) := \left\{ \mu(\cdot) \in \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d)) \text{ s.t. } \mu(\cdot) \text{ is a solution of (4.1) with } \mu(0) = \mu^0 \right\} \quad (4.28)$$

for each  $\mu^0 \in \mathcal{P}_p(\mathbb{R}^d)$ . Our first result establishes that  $\mathcal{S}_{[0, T]}(\mu^0)$  is compact for the topology of uniform convergence when  $V : [0, T] \times \mathcal{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d, \mathbb{R}^d)$  has convex images. Its proof is not written in full details, as it is based on simpler variants of the arguments subtending Theorem 3.2.

**Theorem 4.5** (Compactness of the solution set). *Let  $\mu^0 \in \mathcal{P}_p(\mathbb{R}^d)$  and  $V : [0, T] \times \mathcal{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d, \mathbb{R}^d)$  be a set-valued map satisfying Hypotheses (CI) and whose images are convex. Then, the solution set  $\mathcal{S}_{[0, T]}(\mu^0) \subset C^0([0, T], \mathcal{P}_p(\mathbb{R}^d))$  associated with the Cauchy problem (4.1) is compact for the topology of uniform convergence.*

*Proof.* Given a sequence  $(\mu_n(\cdot), v_n(\cdot)) \subset \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d)) \times \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$  of trajectory-selection pairs for the Cauchy problem (4.1), one may repeat the arguments of Step 3 in the proof of Theorem 3.2 while using (CI)-(i) along with Remark 2.22 to show the existence of a compact set  $\mathcal{K} \subset \mathcal{P}_p(\mathbb{R}^d)$  and a constant  $c_p > 0$ , both depending only on  $\mu^0, p$  and  $\|m(\cdot)\|_1$ , such that

$$\mu_n(t) \in \mathcal{K} \quad \text{and} \quad W_p(\mu_n(\tau), \mu_n(t)) \leq c_p \int_\tau^t m(s) ds$$

for all times  $0 \leq \tau \leq t \leq T$  and each  $n \geq 1$ . Hence, by the Ascoli-Arzelà theorem, there exists a curve  $\mu(\cdot) \in \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d))$  such that

$$\sup_{t \in [0, T]} W_p(\mu_n(t), \mu(t)) \xrightarrow{n \rightarrow +\infty} 0, \quad (4.29)$$

along a subsequence that we do not relabel. By following again the compactness and diagonal argument of Step 3 in the proof of Theorem 3.2, one can show that there exists a Carathéodory vector field  $v(\cdot) \in \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$  such that

$$v_n(\cdot) \xrightarrow{n \rightarrow +\infty} v(\cdot) \quad (4.30)$$

weakly in  $L^1([0, T], C^0(B(0, R), \mathbb{R}^d))$  for every  $R > 0$ , along possibly different subsequences. In particular, it follows from a direct application of Mazur's lemma that

$$|v(t, x)| \leq m(t) \left(1 + |x| + \mathcal{M}_p(\mu(t))\right) \quad \text{and} \quad \text{Lip}(v(t); \mathbb{R}^d) \leq l(t), \quad (4.31)$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$  and all  $x \in \mathbb{R}^d$ . We now prove that  $(\mu(\cdot), v(\cdot)) \in \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d)) \times \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$  is a solution of (4.1). Upon combining the convergence result of (4.29) together with (4.31) and Proposition 2.4, it is clear that

$$\int_0^T \int_{\mathbb{R}^d} \partial_t \phi(t, x) d\mu_n(t)(x) dt \xrightarrow{n \rightarrow +\infty} \int_0^T \int_{\mathbb{R}^d} \partial_t \phi(t, x) d\mu(t)(x) dt \quad (4.32)$$

and

$$\int_0^T \int_{\mathbb{R}^d} \langle \nabla_x \phi(t, x), v_n(t, x) \rangle d(\mu(t) - \mu_n(t))(x) dt \xrightarrow{n \rightarrow +\infty} 0 \quad (4.33)$$

for each  $\phi \in C_c^\infty((0, T) \times \mathbb{R}^d, \mathbb{R})$ . Moreover, by choosing test functions of the form

$$\phi(t, x) := \zeta(t) \psi(x), \quad (4.34)$$

for all  $(t, x) \in [0, T] \times \mathbb{R}^d$ , with  $(\zeta, \psi) \in C_c^\infty((0, T), \mathbb{R}) \times C_c^\infty(\mathbb{R}^d, \mathbb{R})$ , and then setting  $\nu(t) := \zeta(t) \nabla \psi \cdot \mu(t) \in \mathcal{M}(\text{supp}(\psi), \mathbb{R}^d)$  for all times  $t \in [0, T]$ , it follows from (4.30) that

$$\int_0^T \zeta(t) \int_{\mathbb{R}^d} \langle \nabla \psi(x), v(t, x) - v_n(t, x) \rangle d\mu(t)(x) dt \xrightarrow{n \rightarrow +\infty} 0, \quad (4.35)$$

along a subsequence that depends on  $\text{supp}(\psi) \subset \mathbb{R}^d$ . Therefore, by merging the convergence results of (4.32), (4.33) and (4.35) while recalling that the linear span of test functions of the form (4.34) is dense in  $C_c^\infty((0, T) \times \mathbb{R}^d, \mathbb{R})$ , we finally obtain that

$$\int_0^T \int_{\mathbb{R}^d} \left( \partial_t \phi(t, x) + \langle \nabla_x \phi(t, x), v(t, x) \rangle \right) d\mu(t)(x) dt = 0$$

for all  $\phi \in C_c^\infty((0, T) \times \mathbb{R}^d, \mathbb{R}^d)$ , which together with the fact that  $\mu(0) = \mu^0$  as a consequence of (4.29) equivalently means that the pair  $(\mu(\cdot), v(\cdot)) \in \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d)) \times \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$  solves

$$\begin{cases} \partial_t \mu(t) + \text{div}_x(v(t) \mu(t)) = 0, \\ \mu(0) = \mu^0. \end{cases} \quad (4.36)$$

To conclude, there simply remains to show that  $v(t) \in V(t, \mu(t))$  for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ . To this end, observe that as a consequence of Hypothesis (CI)-(iv) combined with Lemma 2.10-(a) and (c), there exists for each  $n \geq 1$  a measurable selection

$$t \in [0, T] \mapsto \tilde{v}_n(t) \in V(t, \mu(t)), \quad (4.37)$$

which satisfies

$$\mathbf{d}_{\text{sup}}(\tilde{v}_n(t), v_n(t)) \leq L(t) W_p(\mu_n(t), \mu(t)) \xrightarrow{n \rightarrow +\infty} 0$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ , where we also used (4.29). Whence, by a direct application of Lebesgue's dominated convergence theorem, it necessarily follows that

$$\int_0^T \|\tilde{v}_n(t) - v_n(t)\|_{C^0(B(0, R), \mathbb{R}^d)} dt \xrightarrow{n \rightarrow +\infty} 0$$

for each  $R > 0$ , which combined with (4.30) implies in particular that

$$\tilde{v}_n(\cdot) \xrightarrow{n \rightarrow +\infty} v(\cdot) \quad (4.38)$$

weakly in  $L^1([0, T], C^0(B(0, R), \mathbb{R}^d))$ , along adequate subsequences which depend on  $R > 0$ . Repeating the reasoning at the end of Step 5 in the proof of Theorem 3.2, one can finally prove  $v(t) \in V(t, \mu(t))$  for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ , which together with (4.36) concludes the proof of Theorem 4.5.  $\square$

The compactness result established in Theorem 4.5 crucially relies on the convexity of the admissible velocities, and does not hold anymore without this assumption. In the following theorem, we show that it is nonetheless possible to precisely characterise the closure of the solution of the solution set in the absence of convexity, and that the latter coincides with that of the convexified dynamics.

**Theorem 4.6** (Relaxation property for continuity inclusions). *Let  $\mu^0 \in \mathcal{P}_p(\mathbb{R}^d)$  and  $V : [0, T] \times \mathcal{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d, \mathbb{R}^d)$  be a set-valued map satisfying Hypotheses (CI). Then, for every solution  $\mu(\cdot) \in AC([0, T], \mathcal{P}_p(\mathbb{R}^d))$  of the relaxed Cauchy problem*

$$\begin{cases} \partial_t \mu(t) \in -\operatorname{div}_x(\overline{\operatorname{co}}V(t, \mu(t))\mu(t)), \\ \mu(0) = \mu^0, \end{cases} \quad (4.39)$$

and any  $\delta > 0$ , there exists a solution  $\mu_\delta(\cdot) \in AC([0, T], \mathcal{P}_p(\mathbb{R}^d))$  of the Cauchy problem

$$\begin{cases} \partial_t \mu_\delta(t) \in -\operatorname{div}_x(V(t, \mu_\delta(t))\mu_\delta(t)), \\ \mu_\delta(0) = \mu^0, \end{cases} \quad (4.40)$$

which satisfies

$$\sup_{t \in [0, T]} W_p(\mu(t), \mu_\delta(t)) \leq \delta.$$

In particular, the solution set of (4.40) is dense in that of (4.39) for the topology of uniform convergence over  $C^0([0, T], \mathcal{P}_p(\mathbb{R}^d))$ .

The proof of Theorem 4.6 will be split into three steps. In Step 1, we start by choosing a solution of (4.39) and an adequate subdivision of the interval  $[0, T]$  that we use along with Aumann's theorem to define an intermediate curve, of which it is shown in Step 2 that it is close to the original one in  $C^0([0, T], \mathcal{P}_p(\mathbb{R}^d))$ . We then conclude in Step 3 by applying the generalised Filippov estimates of Theorem 4.2, which provides us with a solution of the Cauchy problem (4.40) that remains sufficiently close to the intermediate curve, and thus to the original one.

*Proof of Theorem 4.6.* We start by observing that, if  $V : [0, T] \times \mathcal{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d, \mathbb{R}^d)$  is a set-valued map satisfying Hypotheses (CI) and  $\mu(\cdot) \in AC([0, T], \mathcal{P}_p(\mathbb{R}^d))$  is a solution of either (4.39) or (4.40), then by Proposition 2.21 every measurable selection  $v(\cdot) \in \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$  of either  $\overline{\operatorname{co}}V(\cdot, \mu(\cdot))$  or  $V(\cdot, \mu(\cdot))$  necessarily satisfies

$$|v(t, x)| \leq (1 + \mathcal{C})m(t)(1 + |x|), \quad (4.41)$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$  and all  $x \in \mathbb{R}^d$ , where  $\mathcal{C} > 0$  is a constant that only depends on the magnitudes of  $p, \mathcal{M}_p(\mu^0)$  and  $\|m(\cdot)\|_1$ . Thus, denoting by  $\boldsymbol{\eta} \in \mathcal{P}(\mathbb{R}^d \times \Sigma_T)$  the superposition measure associated with such a velocity field via Definition 2.18, it directly follows from Grönwall's lemma that

$$\|\sigma(\cdot)\|_{C^0([0, T], \mathbb{R}^d)} \leq \mathcal{C}_T(1 + |x|), \quad (4.42)$$

for  $\boldsymbol{\eta}$ -almost every  $(x, \sigma) \in \mathbb{R}^d \times \Sigma_T$ , where  $\mathcal{C}_T := \max\{1, (1 + \mathcal{C})\|m(\cdot)\|_1\} \exp((1 + \mathcal{C})\|m(\cdot)\|_1)$ .

**Step 1 – Construction of an intermediate curve via Aumann’s theorem.** Given an arbitrary real number  $\delta > 0$ , the fact that  $\mu^0 \in \mathcal{P}_p(\mathbb{R}^d)$  implies by standard results in measure theory that there exists a positive radius  $R_\delta > 0$  for which

$$\left( \int_{\{x \text{ s.t. } |x| \geq R_\delta / C_T - 1\}} (1 + |x|)^p d\mu^0(x) \right)^{1/p} \leq \frac{\delta}{2(1 + C)(1 + C_T)(1 + \|m(\cdot)\|_1)}. \quad (4.43)$$

Besides, remarking that  $m(\cdot) \in L^1([0, T], \mathbb{R}_+)$ , there exists a subdivision  $0 = t_0 < t_1 < \dots < t_N = T$  of the interval  $[0, T]$  such that

$$\int_{t_i}^{t_{i+1}} m(s) ds \leq \frac{\delta}{2(1 + C)(1 + R_\delta)} \quad (4.44)$$

for each  $i \in \{0, \dots, N - 1\}$ . From Hypotheses **(CI)** and Lemma 2.10, one can check that the restricted set-valued maps  $t \in [0, T] \rightrightarrows V(t, \mu(t))|_{B(0, R_\delta)}$  and  $t \in [0, T] \rightrightarrows \overline{\text{co}}V(t, \mu(t))|_{B(0, R_\delta)}$  are  $\mathcal{L}^1$ -measurable. Furthermore, they have compact images and are integrably bounded, and using the fact that the topology induced by  $d_{cc}(\cdot, \cdot)$  on  $C^0(B(0, R_\delta), \mathbb{R}^d)$  coincides with the usual norm topology, one has that

$$\left( \overline{\text{co}}V(t, \mu(t)) \right) |_{B(0, R_\delta)} = \overline{\text{co}} \left( V(t, \mu(t)) |_{B(0, R_\delta)} \right), \quad (4.45)$$

wherein the second convex hull is taken in the Banach space  $(C^0(B(0, R_\delta), \mathbb{R}^d), \|\cdot\|_{C^0(B(0, R_\delta), \mathbb{R}^d)})$ . Thus, denoting by  $t \in [0, T] \mapsto v(t) \in \overline{\text{co}}V(t, \mu(t))$  the velocity selection associated with a given solution  $\mu(\cdot) \in \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d))$  of (4.39), it follows from Theorem 2.11 that there exist measurable selections  $t \in [t_i, t_{i+1}] \mapsto v_i^\delta(t) \in V(t, \mu(t))|_{B(0, R_\delta)}$  such that

$$\left\| \int_{t_i}^{t_{i+1}} v|_{B(0, R_\delta)}(s) ds - \int_{t_i}^{t_{i+1}} v_i^\delta(s) ds \right\|_{C^0(B(0, R_\delta), \mathbb{R}^d)} \leq \frac{\delta}{N} \quad (4.46)$$

for every  $i \in \{0, \dots, N - 1\}$ , where the maps  $v_i^\delta(\cdot), v|_{B(0, R_\delta)}(\cdot)$  are elements of the separable Banach space  $L^1([0, T], C^0(B(0, R_\delta), \mathbb{R}^d))$  and the integrals are understood in the sense of Bochner. Notice now that  $(t, w) \in [t_i, t_{i+1}] \times C^0(\mathbb{R}^d, \mathbb{R}^d) \mapsto \|w - v_i^\delta(t)\|_{C^0(B(0, R_\delta), \mathbb{R}^d)}$  are Carathéodory, and also that

$$t \in [t_i, t_{i+1}] \rightrightarrows V(t, \mu(t)) \cap \left\{ w \in C^0(\mathbb{R}^d, \mathbb{R}^d) \text{ s.t. } \|w - v_i^\delta(t)\|_{C^0(B(0, R_\delta), \mathbb{R}^d)} = 0 \right\}$$

have nonempty images for each  $i \in \{0, \dots, N - 1\}$  by construction. Thence, by Lemma 2.10-(a), there exist measurable selections  $t \in [t_i, t_{i+1}] \mapsto v_i(t) \in V(t, \mu(t))$  such that  $v_i(t)|_{B(0, R_\delta)} = v_i^\delta(t)$  for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ . Therefore, the velocity field  $w : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  defined by

$$w(t, x) := \sum_{i=0}^{N-1} \mathbf{1}_{[t_i, t_{i+1})}(t) v_i(t, x)$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$  and all  $x \in \mathbb{R}^d$  is Carathéodory, and satisfies Hypotheses **(CE)**. As such, by Theorem 2.17, it generates a unique solution  $\nu(\cdot) \in \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d))$  to the Cauchy problem

$$\begin{cases} \partial_t \nu(t) + \text{div}_x(w(t)\nu(t)) = 0, \\ \nu(0) = \mu^0. \end{cases}$$

**Step 2 – Estimating of the Wasserstein distance between  $\mu(\cdot)$  and  $\nu(\cdot)$ .** Let  $\eta_\mu, \eta_\nu \in \mathcal{P}(\mathbb{R}^d \times \Sigma_T)$  be two superposition measures given by Theorem 2.19 which satisfy

$$(\mathbf{e}_t)_\# \eta_\mu = \mu(t) \quad \text{and} \quad (\mathbf{e}_t)_\# \eta_\nu = \nu(t)$$

for all times  $t \in [0, T]$ , and  $\hat{\eta}_{\mu, \nu} \in \Gamma(\eta_\mu, \eta_\nu)$  be a transport plan given by Lemma 2.20, for which

$$(\pi_{\mathbb{R}^d}, \pi_{\mathbb{R}^d})_\# \hat{\eta}_{\mu, \nu} = (\text{Id}, \text{Id})_\# \mu^0 \quad \text{and} \quad \gamma(t) := (\mathbf{e}_t, \mathbf{e}_t)_\# \hat{\eta}_{\mu, \nu} \in \Gamma_o(\mu(t), \nu(t)).$$



Then

$$\begin{aligned}
W_p(\mu(t), \nu(t)) &= \left( \int_{\mathbb{R}^{2d}} |x - y|^p d\gamma(t)(x, y) \right)^{1/p} \\
&= \left( \int_{(\mathbb{R}^d \times \Sigma_T)^2} |\sigma_\mu(t) - \sigma_\nu(t)|^p d\hat{\boldsymbol{\eta}}_{\mu, \nu}(x, \sigma_\mu, y, \sigma_\nu) \right)^{1/p} \\
&\leq \left( \int_{(\mathbb{R}^d \times \Sigma_T)^2} \left| \int_0^t (v(s, \sigma_\mu(s)) - v(s, \sigma_\nu(s))) ds \right|^p d\hat{\boldsymbol{\eta}}_{\mu, \nu}(x, \sigma_\mu, y, \sigma_\nu) \right)^{1/p} \\
&\quad + \left( \int_{\mathbb{R}^d \times \Sigma_T} \left| \int_0^t (v(s, \sigma_\nu(s)) - w(s, \sigma_\nu(s))) ds \right|^p d\boldsymbol{\eta}_\nu(y, \sigma_\nu) \right)^{1/p},
\end{aligned} \tag{4.47}$$

for all times  $t \in [0, T]$ , where we used the fact that  $\boldsymbol{\eta}_\mu, \boldsymbol{\eta}_\nu$  are superposition measures in the sense of Definition 2.18. By repeating the computations detailed in Appendices C and D, it can be shown that

$$\begin{aligned}
&\left( \int_{(\mathbb{R}^d \times \Sigma_T)^2} \left| \int_0^t (v(s, \sigma_\mu(s)) - v(s, \sigma_\nu(s))) ds \right|^p d\hat{\boldsymbol{\eta}}_{\mu, \nu}(x, \sigma_\mu, y, \sigma_\nu) \right)^{1/p} \\
&\leq \|l(\cdot)\|_{L^1([0, t])}^{(p-1)/p} \left( \int_0^t l(s) W_p^p(\mu(s), \nu(s)) ds \right)^{1/p},
\end{aligned} \tag{4.48}$$

for all times  $t \in [0, T]$ . Concerning the second term in the last inequality of the right-hand side of (4.47), it can be bounded from above by the sum of two integrals as

$$\begin{aligned}
&\left( \int_{\mathbb{R}^d \times \Sigma_T} \left| \int_0^t (v(s, \sigma_\nu(s)) - w(s, \sigma_\nu(s))) ds \right|^p d\boldsymbol{\eta}_\nu(y, \sigma_\nu) \right)^{1/p} \\
&\leq \left( \int_{\{(y, \sigma_\nu) \text{ s.t. } \|\sigma_\nu(\cdot)\|_{C^0([0, T], \mathbb{R}^d)} \leq R_\delta\}} \left| \int_0^t (v(s, \sigma_\nu(s)) - w(s, \sigma_\nu(s))) ds \right|^p d\boldsymbol{\eta}_\nu(y, \sigma_\nu) \right)^{1/p} \\
&\quad + \left( \int_{\{(y, \sigma_\nu) \text{ s.t. } \|\sigma_\nu(\cdot)\|_{C^0([0, T], \mathbb{R}^d)} > R_\delta\}} \left| \int_0^t (v(s, \sigma_\nu(s)) - w(s, \sigma_\nu(s))) ds \right|^p d\boldsymbol{\eta}_\nu(y, \sigma_\nu) \right)^{1/p}.
\end{aligned} \tag{4.49}$$

Recalling that  $(\pi_{\mathbb{R}^d})_\# \boldsymbol{\eta}_\nu = \mu^0$  and invoking the sublinearity estimates of (4.41) and (4.42), the second term in the right-hand side of (4.49) can be estimated as

$$\begin{aligned}
&\left( \int_{\{(y, \sigma_\nu) \text{ s.t. } \|\sigma_\nu(\cdot)\|_{C^0([0, T], \mathbb{R}^d)} > R_\delta\}} \left| \int_0^t (v(s, \sigma_\nu(s)) - w(s, \sigma_\nu(s))) ds \right|^p d\boldsymbol{\eta}_\nu(y, \sigma_\nu) \right)^{1/p} \\
&\leq 2(1 + \mathcal{C}) \|m(\cdot)\|_1 \left( \int_{\{(y, \sigma_\nu) \text{ s.t. } \|\sigma_\nu(\cdot)\|_{C^0([0, T], \mathbb{R}^d)} > R_\delta\}} (1 + \|\sigma_\nu(\cdot)\|_{C^0([0, T], \mathbb{R}^d)})^p d\boldsymbol{\eta}_\nu(y, \sigma_\nu) \right)^{1/p} \\
&\leq 2(1 + \mathcal{C})(1 + \mathcal{C}_T) \|m(\cdot)\|_1 \left( \int_{\{x \text{ s.t. } |x| > R_\delta / \mathcal{C}_T - 1\}} (1 + |x|)^p d\mu^0(x) \right)^{1/p},
\end{aligned} \tag{4.50}$$

which together with our choice of  $R_\delta > 0$  in (4.43) further yields

$$\left( \int_{\{(y, \sigma_\nu) \text{ s.t. } \|\sigma_\nu(\cdot)\|_{C^0([0, T], \mathbb{R}^d)} > R_\delta\}} \left| \int_0^t (v(s, \sigma_\nu(s)) - w(s, \sigma_\nu(s))) ds \right|^p d\boldsymbol{\eta}_\nu(y, \sigma_\nu) \right)^{1/p} \leq \delta. \tag{4.51}$$

We now focus our attention on the first term in the right-hand side of (4.49). By construction, there exists an integer  $j \in \{0, \dots, N-1\}$  such that  $t \in [t_j, t_{j+1}]$ , which means that

$$\begin{aligned}
& \left( \int_{\{(y, \sigma_\nu) \text{ s.t. } \|\sigma_\nu(\cdot)\|_{C^0([0, T], \mathbb{R}^d)} \leq R_\delta\}} \left| \int_0^t (v(s, \sigma_\nu(s)) - w(s, \sigma_\nu(s))) ds \right|^p d\boldsymbol{\eta}_\nu(y, \sigma_\nu) \right)^{1/p} \\
& \leq \sum_{i=0}^{j-1} \left( \int_{\{(y, \sigma_\nu) \text{ s.t. } \|\sigma_\nu(\cdot)\|_{C^0([0, T], \mathbb{R}^d)} \leq R_\delta\}} \left| \int_{t_i}^{t_{i+1}} (v(s, \sigma_\nu(s)) - v_i(s, \sigma_\nu(s))) ds \right|^p d\boldsymbol{\eta}_\nu(y, \sigma_\nu) \right)^{1/p} \\
& \quad + \left( \int_{\{(y, \sigma_\nu) \text{ s.t. } \|\sigma_\nu(\cdot)\|_{C^0([0, T], \mathbb{R}^d)} \leq R_\delta\}} \left| \int_{t_j}^t (v(s, \sigma_\nu(s)) - v_j(s, \sigma_\nu(s))) ds \right|^p d\boldsymbol{\eta}_\nu(y, \sigma_\nu) \right)^{1/p} \\
& \leq \sum_{i=0}^{j-1} \left( \int_{\{(y, \sigma_\nu) \text{ s.t. } \|\sigma_\nu(\cdot)\|_{C^0([0, T], \mathbb{R}^d)} \leq R_\delta\}} \left| \int_{t_i}^{t_{i+1}} (v(s, \sigma_\nu(s)) - v_i(s, \sigma_\nu(s))) ds \right|^p d\boldsymbol{\eta}_\nu(y, \sigma_\nu) \right)^{1/p} + \delta,
\end{aligned} \tag{4.52}$$

where we used the estimates of (4.41) and (4.44) in the second inequality. Now, given an arbitrary integer  $i \in \{0, \dots, j-1\}$ , the facts that  $v(\cdot), v_i(\cdot) \in \mathcal{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$  both satisfy the sublinearity estimate (4.41) along with the regularity assumption of Hypothesis (CI)-(iii) allow us to write

$$\begin{aligned}
& \left( \int_{\{(y, \sigma_\nu) \text{ s.t. } \|\sigma_\nu(\cdot)\|_{C^0([0, T], \mathbb{R}^d)} \leq R_\delta\}} \left| \int_{t_i}^{t_{i+1}} (v(s, \sigma_\nu(s)) - v_i(s, \sigma_\nu(s))) ds \right|^p d\boldsymbol{\eta}_\nu(y, \sigma_\nu) \right)^{1/p} \\
& \leq \left( \int_{\{(y, \sigma_\nu) \text{ s.t. } \|\sigma_\nu(\cdot)\|_{C^0([0, T], \mathbb{R}^d)} \leq R_\delta\}} \left| \int_{t_i}^{t_{i+1}} (v(s, \sigma_\nu(t_i)) - v_i(s, \sigma_\nu(t_i))) ds \right|^p d\boldsymbol{\eta}_\nu(y, \sigma_\nu) \right)^{1/p} \\
& \quad + \left( \int_{\{(y, \sigma_\nu) \text{ s.t. } \|\sigma_\nu(\cdot)\|_{C^0([0, T], \mathbb{R}^d)} \leq R_\delta\}} \left| \int_{t_i}^{t_{i+1}} (v(s, \sigma_\nu(s)) - v(s, \sigma_\nu(t_i))) ds \right|^p d\boldsymbol{\eta}_\nu(y, \sigma_\nu) \right)^{1/p} \\
& \quad + \left( \int_{\{(y, \sigma_\nu) \text{ s.t. } \|\sigma_\nu(\cdot)\|_{C^0([0, T], \mathbb{R}^d)} \leq R_\delta\}} \left| \int_{t_i}^{t_{i+1}} (v_i(s, \sigma_\nu(s)) - v_i(s, \sigma_\nu(t_i))) ds \right|^p d\boldsymbol{\eta}_\nu(y, \sigma_\nu) \right)^{1/p} \\
& \leq \left\| \int_{t_i}^{t_{i+1}} v(s)|_{B(0, R_\delta)} ds - \int_{t_i}^{t_{i+1}} v_i^\delta(s) ds \right\|_{C^0(B(0, R_\delta), \mathbb{R}^d)} \\
& \quad + 2 \left( \int_{\{(y, \sigma_\nu) \text{ s.t. } \|\sigma_\nu(\cdot)\|_{C^0([0, T], \mathbb{R}^d)} \leq R_\delta\}} \left| \int_{t_i}^{t_{i+1}} l(s) \int_{t_i}^s m(\zeta) (1 + |\sigma_\nu(\zeta)| + \mathcal{C}) d\zeta ds \right|^p d\boldsymbol{\eta}_\nu(y, \sigma_\nu) \right)^{1/p} \\
& \leq \left\| \int_{t_i}^{t_{i+1}} v(s)|_{B(0, R_\delta)} ds - \int_{t_i}^{t_{i+1}} v_i^\delta(s) ds \right\|_{C^0(B(0, R_\delta), \mathbb{R}^d)} + 2(1 + R_\delta + \mathcal{C}) \int_{t_i}^{t_{i+1}} l(s) \int_{t_i}^s m(\zeta) d\zeta ds \\
& \leq \frac{\delta}{N} + \delta \int_{t_i}^{t_{i+1}} l(s) ds,
\end{aligned} \tag{4.53}$$

where we leveraged the preliminary estimate (4.41) in the second and third inequalities, as well as (4.44) and (4.46) in the last one. Hence, by inserting the estimates of (4.53) for  $i \in \{0, \dots, j-1\}$  into (4.52), combining the resulting expression with (4.51) and plugging the latter in (4.49), we then get

$$\left( \int_{\mathbb{R}^d \times \Sigma_T} \left| \int_0^t (v(s, \sigma_\nu(s)) - w(s, \sigma_\nu(s))) ds \right|^p d\boldsymbol{\eta}_\nu(y, \sigma_\nu) \right)^{1/p} \leq \delta(3 + \|l(\cdot)\|_1). \tag{4.54}$$

By merging (4.48) and (4.54) into (4.47), raising the resulting inequality to the power  $p$ , applying Grönwall's lemma and then raising the corresponding expression to the power  $1/p$ , we finally obtain

$$W_p(\mu(t), \nu(t)) \leq \delta C_p (3 + \|l(\cdot)\|_1) \exp\left(C'_p \|l(\cdot)\|_1^p\right), \tag{4.55}$$

which holds for all times  $t \in [0, T]$ .

**Step 3 – Back to the unrelaxed problem using Filippov’s estimates.** At this stage, one should note that  $\nu(\cdot) \in \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d))$  is not a solution of (4.40), since  $w(t) \in V(t, \mu(t))$  for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ . Nevertheless, one has that the mismatch function defined by

$$\eta_\delta(t) := \text{dist}_{C^0(B(0, R_\delta), \mathbb{R}^d)}(w(t); V(t, \nu(t)))$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$  is Lebesgue integrable and satisfies

$$\begin{aligned} \eta_\delta(t) &\leq L(t)W_p(\mu(t), \nu(t)) \\ &\leq \delta C_p L(t)(3 + \|l(\cdot)\|_1) \exp\left(C'_p \|l(\cdot)\|_1^p\right), \end{aligned}$$

as a consequence of Hypothesis (CI)-(iv). Thus, by leveraging the results of Theorem 4.2 while recollecting the a priori bound of (4.43), there exists a solution  $\mu_\delta(\cdot) \in \text{AC}([0, T], \mathcal{P}_p(\mathbb{R}^d))$  of the unrelaxed Cauchy problem (4.40) which satisfies

$$\begin{aligned} W_p(\mu_\delta(t), \nu(t)) &\leq C_p \left( \delta C_p \|L(\cdot)\|_1 (3 + \|l(\cdot)\|_1) \exp\left(C'_p \|l(\cdot)\|_1^p\right) + \delta \right) \exp\left(C'_p \|l(\cdot)\|_{L^1([0, t])}^p + \chi_p(t)\right) \\ &\leq \delta C_p \left( 1 + (3 + \|l(\cdot)\|_1) \|\chi_p(\cdot)\|_\infty \right) \exp\left(C'_p \|l(\cdot)\|_1^p + \|\chi_p(\cdot)\|_\infty\right), \end{aligned} \tag{4.56}$$

for all times  $t \in [0, T]$ , where the map  $\chi_p(\cdot)$  is defined in (4.6). Thus, by combining (4.55) and (4.56), applying the triangle inequality and redefining the constant  $\delta > 0$  as

$$\delta := \frac{\delta}{C_p \left( (3 + \|l(\cdot)\|_1) (1 + \|\chi_p(\cdot)\|_\infty \exp(\|\chi_p(\cdot)\|_\infty)) + \exp(\|\chi_p(\cdot)\|_\infty) \right) \exp\left(C'_p \|l(\cdot)\|_1^p\right)}$$

we can finally conclude that the solution  $\mu_\delta(\cdot)$  of the unrelaxed Cauchy problem (4.40) is such that

$$\sup_{t \in [0, T]} W_p(\mu(t), \mu_\delta(t)) \leq \delta,$$

which ends the proof of Theorem 4.6. □

## Appendices

In this auxiliary section, we detail the proofs of several technical results appearing in the manuscript.

### A Proof of Lemma 2.10

In this first appendix, we detail parts of the proof of the measurable selection principles of Lemma 2.10, as these latter rely on somewhat non-standard assumptions.

*Proof of Lemma 2.10.* In what follows, we start with the proof of item (a), and proceed with that of item (b). The proof of item (c) is completely standard and can be found e.g. in [9, Theorem 8.2.8]. The only delicate thing that needs proving in the statement of Lemma 2.10-(a) is the fact that the set-valued map under consideration is measurable. To this end, we consider the multifunctions

$$t \in [0, T] \rightrightarrows \mathcal{A}_n(t) := \left\{ y \in Y \text{ s.t. } \varphi_n(t, y) \leq L(t) \right\}$$

defined for each  $n \geq 1$ , which are  $\mathcal{L}^1$ -measurable by [9, Theorem 8.2.9]. For each closed set  $\mathcal{C} \subset Y$ , this implies in particular by [9, Theorem 8.1.4] that

$$\mathcal{D}_\mathcal{C} := \bigcap_{n \geq 1} \left\{ t \in [0, T] \text{ s.t. } \mathcal{F}(t) \cap \mathcal{A}_n(t) \cap \mathcal{C} \neq \emptyset \right\}$$

is an  $\mathcal{L}^1$ -measurable set. To conclude the proof of our claim, there remains to show that

$$\mathcal{D}_{\mathcal{C}} = \left\{ t \in [0, T] \text{ s.t. } \mathcal{F}(t) \cap \mathcal{A}(t) \cap \mathcal{C} \neq \emptyset \right\},$$

where  $\mathcal{A}(t) := \{y \in Y \text{ s.t. } \varphi(t, y) \leq L(t)\}$  for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ . By construction, one has

$$\left\{ t \in [0, T] \text{ s.t. } \mathcal{F}(t) \cap \mathcal{A}(t) \cap \mathcal{C} \neq \emptyset \right\} \subset \mathcal{D}_{\mathcal{C}},$$

as a consequence of (2.5). Conversely for any  $\tau \in \mathcal{D}_{\mathcal{C}}$ , remark that the sets defined for each  $n \geq 1$  by

$$\mathcal{B}_n(\tau) := \mathcal{F}(\tau) \cap \mathcal{A}_n(\tau) \cap \mathcal{C}$$

form a non-increasing sequence since maps  $(\varphi_n(\cdot, \cdot))$  are pointwisely non-decreasing. Moreover under our standing assumptions, the sets  $\mathcal{B}_n(\tau) \subset Y$  are compact and nonempty for each  $n \geq 1$ . Whence, it follows from Cantor's intersection theorem (see e.g. [73, Theorem 2.6]) that

$$\mathcal{B}(\tau) := \bigcap_{n \geq 1} \mathcal{B}_n(\tau) \neq \emptyset,$$

which together with the fact that  $\mathcal{B}(\tau) \subset \mathcal{F}(\tau) \cap \mathcal{A}(\tau) \cap \mathcal{C}$  finally yields that

$$\tau \in \left\{ t \in [0, T] \text{ s.t. } \mathcal{F}(t) \cap \mathcal{A}(t) \cap \mathcal{C} \neq \emptyset \right\}$$

and concludes the proof of our claim.

We now shift our focus to the statements of Lemma 2.10-(b), and start by observing that the set-valued map appearing therein can be rewritten as

$$t \in [0, T] \Rightarrow \mathcal{F}(t) \cap \left\{ y \in Y \text{ s.t. } \varphi(t, y) \leq \inf_{z \in \mathcal{F}(t)} \varphi(t, z) \right\}.$$

Thus by what precedes, it is sufficient for our purpose to show that the map

$$t \in [0, T] \mapsto \inf_{z \in \mathcal{F}(t)} \varphi(t, z) \in \mathbb{R}_+$$

is  $\mathcal{L}^1$ -measurable. It follows from classical measurability results (see e.g. [9, Theorem 8.2.11]) that for each  $n \geq 1$ , the map

$$\varpi_n : t \in [0, T] \mapsto \inf_{z \in \mathcal{F}(t)} \varphi_n(t, z)$$

is  $\mathcal{L}^1$ -measurable. Since the sequence  $(\varpi_n(\cdot))$  is pointwisely non-decreasing and bounded, the limits

$$\varpi(t) := \lim_{n \rightarrow +\infty} \varpi_n(t) = \inf_{z \in \mathcal{F}(t)} \varphi(t, z)$$

exist for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ , and the map  $\varpi(\cdot)$  is  $\mathcal{L}^1$ -measurable by [9, Theorem 8.2.5].  $\square$

## B Proof of Lemma 2.15

In this second appendix, we detail for the sake of completeness the proof of Lemma 2.15, which mimics standard arguments used to show that the  $L^p$ -spaces are complete, see e.g. to [73, Theorem 3.11]

*Proof.* Let  $(v_n(\cdot)) \subset \mathcal{L}(I, C^0(\mathbb{R}^d, \mathbb{R}^d))$  be a Cauchy sequence in the sense of (2.8), and observe that without loss of generality, the latter may be chosen so that

$$\int_0^T \mathbf{d}_{\text{sup}}(v_n(t), v_{n+1}(t)) dt \leq \frac{1}{2^n} \tag{B.1}$$

for all  $n \geq 1$ , up to extracting a subsequence. Consider now the sequence of real-valued mappings

$$d_n(t) := \sum_{k=1}^n \mathbf{d}_{\text{sup}}(v_k(t), v_{k+1}(t)), \tag{B.2}$$

defined for  $\mathcal{L}^1$ -almost every  $t \in I$  and each  $n \geq 1$ , and observe that by (B.1), one has that

$$\sup_{n \geq 1} \int_0^T d_n(t) dt \leq \sum_{k=1}^{+\infty} \int_0^T \mathbf{d}_{\text{sup}}(v_k(t), v_{k+1}(t)) dt \leq 1. \quad (\text{B.3})$$

By [73, Theorem 1.38], the partial sums of series defined in (B.2) converge for  $\mathcal{L}^1$ -almost every  $t \in I$  towards a map  $d(\cdot) \in L^1(I, \mathbb{R}_+)$ , which must then satisfy

$$\int_0^T d(t) dt \leq 1.$$

Thus, it necessarily follows that  $d(t) < +\infty$  for  $\mathcal{L}^1$ -almost every  $t \in I$ , which by (B.2) together with the definition (2.7) of the extended metric  $\mathbf{d}_{\text{sup}}(\cdot, \cdot)$  implies that the series of functions defined by

$$w(t, x) := \sum_{k=1}^{+\infty} (v_{k+1}(t, x) - v_k(t, x)) \quad \text{for all } x \in \mathbb{R}^d$$

is normally convergent for  $\mathcal{L}^1$ -almost every  $t \in I$ . Consequently, for every  $(t, x) \in I \times \mathbb{R}^d$ , we can define the mapping  $v : I \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  by

$$v(t, x) := \begin{cases} v_1(t, x) + w(t, x) & \text{if } (d_n(t)) \text{ converges,} \\ 0 & \text{otherwise,} \end{cases}$$

and observe that it is  $\mathcal{L}^1$ -measurable with respect to  $t \in I$  and continuous with respect to  $x \in \mathbb{R}^d$ . Moreover, the map  $v(\cdot) \in \mathcal{L}(I, C^0(\mathbb{R}^d, \mathbb{R}^d))$  is the pointwise limit of the sequence  $(v_n(\cdot))$ , since

$$\begin{aligned} \mathbf{d}_{\text{sup}}(v_n(t), v(t)) &= \mathbf{d}_{\text{sup}}\left(\sum_{k=1}^{n-1} (v_{k+1}(t) - v_k(t)), w(t)\right) \\ &= \mathbf{d}_{\text{sup}}\left(\sum_{k=1}^{n-1} (v_{k+1}(t) - v_k(t)), \sum_{k=1}^{+\infty} (v_{k+1}(t) - v_k(t))\right) \xrightarrow{n \rightarrow +\infty} 0, \end{aligned}$$

where we used the fact that a series that is normally convergent over  $\mathbb{R}^d$  is also uniformly convergent. By a simple application of Fatou's lemma (see e.g. [73, Statement 1.28]), one can finally show that

$$\int_0^T \mathbf{d}_{\text{sup}}(v_n(t), v(t)) dt \xrightarrow{n \rightarrow +\infty} 0,$$

which concludes the proof of our claim.  $\square$

## C Proof of Proposition 2.21

In this third appendix, we detail the proof of the moment and equi-integrability bounds displayed in Proposition 2.21, which both rely on the superposition principle recalled in Theorem 2.19.

*Proof of Proposition 2.21.* We do not detail the proof of (2.15) as it is almost identical to that of [14, Proposition 2]. Concerning the equi-integrability estimate (2.17), let  $\boldsymbol{\eta}_\mu \in \mathbb{R}^d \times \Sigma_T$  be a superposition measure given by Theorem 2.19 such that  $\mu(t) = (\mathbf{e}_t)_\# \boldsymbol{\eta}_\mu$  for all times  $t \in [\tau, T]$ , and observe then that

$$|v(t, \sigma_\mu(t))| \leq m(t)(1 + \sigma_\mu(t))$$

holds for  $\mathcal{L}^1$ -almost every  $t \in [\tau, T]$  and  $\boldsymbol{\eta}_\mu$ -almost every  $(x, \sigma_\mu) \in \mathbb{R}^d \times \Sigma_T$  under our working assumptions. Thence, it follows from a simple application of Grönwall's lemma that

$$|\sigma_\mu(t)| \leq \left(|x| + \int_\tau^t m(s) ds\right) \exp\left(\int_\tau^t m(s) ds\right) \leq C_T(1 + |x|), \quad (\text{C.1})$$

for all times  $t \in [\tau, T]$  and  $\boldsymbol{\eta}_\mu$ -almost every  $(x, \sigma) \in \mathbb{R}^d \times \Sigma_T$ , where we introduced the constant  $C_T = \max\{1, \|m(\cdot)\|_1\} \exp(\|m(\cdot)\|_1)$ . In particular, for every  $R > 0$ , the latter inequality implies

$$\begin{aligned} \left( \int_{\{x \text{ s.t. } |x| \geq R\}} |x|^p d\mu(t)(x) \right)^{1/p} &= \left( \int_{\{(x, \sigma_\mu) \text{ s.t. } |e_t(x, \sigma_\mu)| \geq R\}} |e_t(x, \sigma_\mu)|^p d\boldsymbol{\eta}_\mu(x, \sigma_\mu) \right)^{1/p} \\ &\leq C_T \left( \int_{\{(x, \sigma_\mu) \text{ s.t. } C_T(1+|x|) \geq R\}} (1+|x|)^p d\boldsymbol{\eta}_\mu(x, \sigma_\mu) \right)^{1/p} \\ &= C_T \left( \int_{\{x \text{ s.t. } |x| \geq R/C_T - 1\}} (1+|x|)^p d\mu^0(x) \right)^{1/p} \end{aligned}$$

for all times  $t \in [\tau, T]$ , which ends the proof of our claim.  $\square$

## D Proof of Proposition 2.23

In this fourth and last appendix, we detail the proof of the stability estimates of Proposition 2.23.

*Proof of Proposition 2.23.* The demonstration of the global stability estimate (2.19) is omitted, as it is almost identical to that of [14, Proposition 2]. Regarding the local stability estimate, fix some  $R > 0$  and observe that under our working assumptions, the map

$$t \in [\tau, T] \mapsto \|v(t) - w(t)\|_{L^\infty(B(0, R), \mathbb{R}^d; \nu(t))},$$

is Lebesgue integrable. In that case, one can prove by repeating the series of computations much like those of [14, Appendix B] that

$$\begin{aligned} W_p(\mu(t), \nu(t)) &\leq C_p \left( W_p(\mu_\tau, \nu_\tau) + \left( \int_{\mathbb{R}^d \times \Sigma_T} \left( \int_\tau^t |v(s, \sigma_\nu(s)) - w(s, \sigma_\nu(s))| ds \right)^p d\boldsymbol{\eta}_\nu(y, \sigma_\nu) \right)^{1/p} \right) \\ &\quad \times \exp \left( C'_p \|l(\cdot)\|_{L^1([\tau, t])}^p \right), \end{aligned} \tag{D.1}$$

for all times  $t \in [\tau, T]$ , wherein  $\boldsymbol{\eta}_\nu \in \mathcal{P}(\mathbb{R}^d \times \Sigma_T)$  is a superposition measure generating  $\nu(\cdot) \in \text{AC}([\tau, T], \mathcal{P}_p(\mathbb{R}^d))$ , whose existence follows from Theorem 2.19. We focus our attention on the integral term appearing in the right-hand side of (D.1), and observe that the latter can be split into two parts

$$\begin{aligned} &\int_{\mathbb{R}^d \times \Sigma_T} \left( \int_\tau^t |v(s, \sigma_\nu(s)) - w(s, \sigma_\nu(s))| ds \right)^p d\boldsymbol{\eta}_\nu(y, \sigma_\nu) \\ &= \int_{\{(y, \sigma_\nu) \text{ s.t. } \|\sigma_\nu(\cdot)\|_{C^0([\tau, T], \mathbb{R}^d)} \leq R\}} \left( \int_\tau^t |v(s, \sigma_\nu(s)) - w(s, \sigma_\nu(s))| ds \right)^p d\boldsymbol{\eta}_\nu(y, \sigma_\nu) \\ &\quad + \int_{\{(y, \sigma_\nu) \text{ s.t. } \|\sigma_\nu(\cdot)\|_{C^0([\tau, T], \mathbb{R}^d)} > R\}} \left( \int_\tau^t |v(s, \sigma_\nu(s)) - w(s, \sigma_\nu(s))| ds \right)^p d\boldsymbol{\eta}_\nu(y, \sigma_\nu), \end{aligned} \tag{D.2}$$

for all  $t \in [0, T]$ . The first term in the right-hand side of (D.2) can be estimated straightforwardly as

$$\begin{aligned} &\int_{\{(y, \sigma_\nu) \text{ s.t. } \|\sigma_\nu(\cdot)\|_{C^0([\tau, T], \mathbb{R}^d)} \leq R\}} \left( \int_\tau^t |v(s, \sigma_\nu(s)) - w(s, \sigma_\nu(s))| ds \right)^p d\boldsymbol{\eta}_\nu(y, \sigma_\nu) \\ &\leq \left( \int_\tau^t \|v(s) - w(s)\|_{L^\infty(B(0, R), \mathbb{R}^d; \nu(s))} ds \right)^p, \end{aligned} \tag{D.3}$$

for all times  $t \in [\tau, T]$ . Concerning the second term in the right-hand side of (D.2), it follows from the sublinearity assumptions made on  $v, w : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  and Hölder's inequality that

$$\begin{aligned} &\int_{\{(y, \sigma_\nu) \text{ s.t. } \|\sigma_\nu(\cdot)\|_{C^0([\tau, T], \mathbb{R}^d)} > R\}} \left( \int_\tau^t |v(s, \sigma_\nu(s)) - w(s, \sigma_\nu(s))| ds \right)^p d\boldsymbol{\eta}_\nu(y, \sigma_\nu) \\ &\leq 2^p \int_{\{(y, \sigma_\nu) \text{ s.t. } \|\sigma_\nu(\cdot)\|_{C^0([\tau, T], \mathbb{R}^d)} > R\}} \left( \int_\tau^t m(s)(1 + |\sigma_\nu(s)|) ds \right)^p d\boldsymbol{\eta}_\nu(y, \sigma_\nu) \\ &\leq 2^p \|m(\cdot)\|_{L^1([\tau, t])}^{p/q} \int_\tau^t m(s) \left( \int_{\{(y, \sigma_\nu) \text{ s.t. } \|\sigma_\nu(\cdot)\|_{C^0([\tau, T], \mathbb{R}^d)} > R\}} (1 + |\sigma_\nu(s)|)^p d\boldsymbol{\eta}_\nu(y, \sigma_\nu) \right) ds. \end{aligned} \tag{D.4}$$

Recall now that by (C.1) above, it also holds that

$$\|\sigma_\nu(\cdot)\|_{C^0([\tau, T], \mathbb{R}^d)} \leq C_T(1 + |y|),$$

for  $\eta_\nu$ -almost every  $(y, \sigma_\nu)$ , which together with (D.4) then yields

$$\begin{aligned} & \int_{\{(y, \sigma_\nu) \text{ s.t. } \|\sigma_\nu(\cdot)\|_{C^0([\tau, T], \mathbb{R}^d)} > R\}} \left( \int_\tau^t |v(s, \sigma_\nu(s)) - w(s, \sigma_\nu(s))| ds \right)^p d\eta_\nu(y, \sigma_\nu) \\ & \leq 2^p \|m(\cdot)\|_{L^1([\tau, t])}^p \int_{\{y \text{ s.t. } |y| \geq R/C_T - 1\}} (1 + C_T(1 + |y|))^p d\nu_\tau(y), \end{aligned} \quad (\text{D.5})$$

for all times  $t \in [\tau, T]$ . Therefore, by merging (D.3) and (D.5) in (D.1) while applying Hölder's inequality, we finally obtain that

$$\begin{aligned} W_p(\mu(t), \nu(t)) & \leq C_p \left( W_p(\mu_\tau, \nu_\tau) + \int_\tau^t \|v(s) - w(s)\|_{L^\infty(B(0, R), \mathbb{R}^d, \nu(s))} ds \right. \\ & \quad \left. + 2(1 + C_T) \|m(\cdot)\|_{L^1([\tau, t])} \left( \int_{\{y \text{ s.t. } |y| \geq R/C_T - 1\}} (1 + |y|)^p d\nu_\tau(y) \right)^{1/p} \right) \\ & \quad \times \exp \left( C'_p \|l(\cdot)\|_{L^1([\tau, t])}^p \right), \end{aligned}$$

for all times  $t \in [\tau, T]$ , which concludes the proof of our claim.  $\square$

## References

- [1] G. Albi, D. Balagué, J.A. Carrillo, and J. von Brecht. Stability Analysis of Flock and Mill Rings for Second Order Models in Swarming. *SIAM Journal of Applied Mathematics*, 74(3):794–818, 2014.
- [2] G. Albi, L. Pareschi, G. Toscani, and M. Zanella. Recent Advances in Opinion Modeling: Control and Social Influence. *Active Particles, Volume 1*, pages 49–98, 2017.
- [3] L. Ambrosio. Transport Equation and Cauchy Problem for BV Vector Fields. *Inventiones Mathematicae*, 158(2):227–260, 2004.
- [4] L. Ambrosio and G. Crippa. Continuity Equations and ODE Flows with Non-Smooth Velocities. *Proceedings of the Royal Society of Edinburgh*, 144(6):1191–1244, 2014.
- [5] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variations and Free Discontinuity Problems*. Oxford Mathematical Monographs, 2000.
- [6] L. Ambrosio and W. Gangbo. Hamiltonian ODEs in the Wasserstein Space of Probability Measures. *Communications on Pure and Applied Mathematics*, 61(1):18–53, 2008.
- [7] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, 2008.
- [8] J.-P. Aubin and A. Cellina. *Differential Inclusions*. Springer-Verlag, 1984.
- [9] J.-P. Aubin and H. Frankowska. *Set-Valued Analysis*. Modern Birkhäuser Classics. Birkhäuser Basel, 1990.
- [10] J.-D. Benamou, G. Carlier, and M. Laborde. An Augmented Lagrangian Approach to Wasserstein Gradient Flows and Applications. *ESAIM: Proceedings and Surveys*, 54:1–17, 2016.
- [11] P. Bernard. Young Measures, Superposition and Transport. *Indiana Univ. Math. Journal*, 57(1):247–275, 2008.
- [12] S. Bianchini and P. Bonicatto. A Uniqueness Result for the Decomposition of Vector Fields in  $\mathbb{R}^d$ . *Inventiones Mathematicae*, 220:255–393, 2020.
- [13] V.I. Bogachev. *Measure Theory*, volume 1. Berlin: Springer, 2007.
- [14] B. Bonnet and H. Frankowska. Differential Inclusions in Wasserstein Spaces: The Cauchy-Lipschitz Framework. *Journal of Differential Equations*, 271:594–637, 2021.
- [15] B. Bonnet and H. Frankowska. Necessary Optimality Conditions for Optimal Control Problems in Wasserstein Spaces. *Applied Mathematics and Optimization*, 84:1281–1330, 2021.
- [16] B. Bonnet and H. Frankowska. Semiconcavity and Sensitivity Analysis in Mean-Field Optimal Control and Applications. *Journal de Mathématiques Pures et Appliquées*, 157:282–345, 2022.
- [17] B. Bonnet and F. Rossi. The Pontryagin Maximum Principle in the Wasserstein Space. *Calculus of Variations and Partial Differential Equations*, 58:11, 2019.

- [18] H. Brézis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Universitext. Springer, 2010.
- [19] F. Camilli, G. Cavagnari, R. De Maio, and B. Piccoli. Superposition Principle and Schemes for Measure Differential Equations. *Kinetic and Related Models*, 14(1):89–113, 2021.
- [20] P. Cannarsa, R. Capuani, and P. Cardaliaguet. Mean-Field Games with State-Constraints: From Mild to Pointwise Solutions to the PDE System. *Calculus of Variations and Partial Differential Equations*, 60:108, 2021.
- [21] P. Cannarsa and H. Frankowska. Some Characterizations of Optimal Trajectories in Control Theory. *SIAM Journal on Control and Optimization*, 29(6):1322–1347, 1991.
- [22] P. Cannarsa and C. Sinestrari. *Semiconcave Functions, Hamilton-Jacobi Equations and Optimal Control*. Progresses in Nonlinear Differential Equations and Their Applications. Birkhäuser Basel, 2004.
- [23] P. Cardaliaguet, F. Delarue, J.-M. Lasry, and P.-L. Lions. *The Master Equation and the Convergence Problem in Mean Field Games*. Annals of Mathematics Studies. Princeton University Press, 2019.
- [24] G. Carlier, V. Duval, G. Peyré, and B. Schmitzer. Convergence of Entropic Schemes for Optimal Transport and Gradient Flows. *SIAM Journal on Mathematical Analysis*, 49(2):1385–1418, 2017.
- [25] J.A. Carrillo, M. Di Francesco, A. Figalli, T. Laurent, and D. Slepčev. Global-in-Time Weak Measure Solutions and Finite-Time Aggregation for Nonlocal Interaction Equations. *Duke Mathematical Journal*, 156(2):229–271, 2011.
- [26] J.A. Carrillo, M. Fornasier, J. Rosado, and G. Toscani. Asymptotic Flocking for the Kinetic Cucker-Smale Model. *SIAM Journal on Mathematical Analysis*, 42(1):218–236, 2010.
- [27] J.A. Carrillo, S. Lisini, and E. Mainini. Gradient Flows for Non-Smooth Interaction Potentials. *Nonlinear Analysis; Theory, Methods and Applications*, 100:122–147, 2014.
- [28] J.A. Carrillo, S. Martin, and V. Panferov. A New Interaction Potential for Swarming Models. *Physica D: Nonlinear Phenomena*, 260:112–126, 2013.
- [29] G. Cavagnari, S. Lisini, C. Orrieri, and G. Savaré. Lagrangian, Eulerian and Kantorovich Formulations of Multi-Agent Optimal Control Problems: Equivalence and Gamma-Convergence. *Journal of Differential Equations*, 322:268–364, 2022.
- [30] G. Cavagnari, A. Marigonda, and M. Quincampoix. Compatibility of State Constraints and Dynamics for Multiagent Control Systems. *Journal of Evolution Equations*, 21(4):4491–4537, 2021.
- [31] 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*, pages 1–96, 2022.
- [32] L. Cesari. *Optimization Theory and Applications. Problems with Ordinary Differential Equation*, volume 17 of *Science and Business Media*. Springer, 2012.
- [33] F. Clarke. *Functional Analysis, Calculus of Variations and Optimal Control*, volume 264. Springer, 2013.
- [34] G. Crippa and M. Lécureux-Mercier. Existence and Uniqueness of Measure Solutions for a System of Continuity Equations with Non-Local Flow. *Nonlinear Differential Equations and Applications*, 20:523–537, 2013.
- [35] E. Cristiani, B. Piccoli, and A. Tosin. *Multiscale Modeling of Pedestrian Dynamics*, volume 12. Springer, 2014.
- [36] R.L. Di Perna and P.-L. Lions. Ordinary Differential Equations, Transport Theory and Sobolev Spaces. *Inventiones Mathematicae*, 98(3):511–548, 1989.
- [37] J. Diestel. Remarks on Weak Compactness in  $L^1(\mu, X)$ . *Glasgow Mathematical Journal*, 18(1):87–91, 1977.
- [38] J. Diestel, W.M. Ruess, and W. Schachermayer. On Weak Compactness in  $L^1(\mu, X)$ . *Proceedings of the American Mathematical Society*, 118:447–453, 1993.
- [39] J. Diestel and J.J.Jr Uhl. *Vector Measures*. Number 15 in Mathematical Surveys. American Mathematical Society, 1977.
- [40] J. Dolbeault, B. Nazareth, and G. Savaré. A New Class of Transport Distances Between Measures. *Calculus of Variations and Partial Differential Equations*, 34(2):193–231, 2009.
- [41] M. Erbar. The Heat Equation on Manifolds as a Gradient Flow in the Wasserstein Space. *Annales de l’I.H.P. Probabilités et statistiques*, 46(1):1–23, 2010.
- [42] A.F. Filippov. *Differential Equations with Discontinuous Righthand Sides: Control Systems*, volume 18 of *Science and Business Media*. Springer, 2013.
- [43] A.L. Filippov. On Certain Questions in the Theory of Optimal Control. *Journal of the Society for Industrial and Applied Mathematics, Series A: Control*, 1(1):76–84, 1962.
- [44] M. Fornasier, S. Lisini, C. Orrieri, and G. Savaré. Mean-Field Optimal Control as Gamma-Limit of Finite Agent Controls. *European Journal of Applied Mathematics*, 30(6):1153–1186, 2019.
- [45] M. Fornasier, B. Piccoli, and F. Rossi. Mean-Field Sparse Optimal Control. *Philosophical Transactions of the Royal Society A.*, 372(20130400), 2014.
- [46] M. Fornasier and F. Solombrino. Mean Field Optimal Control. *ESAIM COCV*, 20(4):1123–1152, 2014.
- [47] H. Frankowska. The Maximum Principle for an Optimal Solution to a Differential Inclusion with End-Point Constraints. *SIAM Journal on Control and Optimization*, 25(1):145–157, 1987.



- [48] H. Frankowska. Optimal Trajectories Associated with a Solution of the Contingent Hamilton-Jacobi Equation. *Applied Mathematics and Optimization*, 19(1):291–311, 1989.
- [49] H. Frankowska. A Priori Estimates for Operational Differential Inclusions. *Journal of Differential Equations*, 84:100–128, 1990.
- [50] H. Frankowska, E.M. Marchini, and M. Mazzola. Necessary Optimality Conditions for Infinite Dimensional State Constrained Control Problems. *Journal of Differential Equations*, 264(12):7294–7327, 2018.
- [51] H. Frankowska and N.P. Osmolovski. Second-Order Sufficient Conditions, Part 2: Weak Local Minima in an Optimal Control Problem with a General Control Constraint. *Systems and Control Letters*, 182:105673, 2023.
- [52] H. Frankowska and N.P. Osmolovskii. Strong Local Minimizers in Optimal Control Problems with State Constraints: Second Order Necessary Conditions. *SIAM Journal on Control and Optimization*, 58(3):2353–2376, 2018.
- [53] H. Frankowska, S. Plaskacz, and T. Rzezuchowski. Measurable Viability Theorems and the Hamilton-Jacobi-Bellman Equation. *Journal of Differential Equations*, 116(2):265–305, 1995.
- [54] H. Frankowska, H. Zhang, and X. Zhang. Stochastic Optimal Control Problems with Control and Initial-Final States Constraints. *SIAM Journal on Control and Optimization*, 56(3):1823–1855, 2018.
- [55] Ugo Gianazza, Giuseppe Savaré, and Giuseppe Toscani. The Wasserstein Gradient Flow of the Fisher Information and the Quantum Drift-Diffusion Equation. *Archives for Rational Mechanics and Analysis*, 1(194):133–220, 2009.
- [56] S.-Y. Ha, K. Lee, and D. Levy. Emergence of Time-Asymptotic Flocking in a Stochastic Cucker-Smale System. *Comm. Math. Sci.*, 7(2):453–469, 2009.
- [57] John Horváth. *Topological Vector Spaces and Distributions*. Courier Corporation, 2012.
- [58] M.Y. Huang, R. Malhamé, and P.E. Caines. Large Population Stochastic Dynamic Games : Closed-Loop McKean-Vlasov Systems and the Nash Certainty Equivalence Principle. *Communications in Information and Systems*, 6(3):221–252, 2006.
- [59] C. Jimenez, A. Marigonda, and M. Quincampoix. Optimal Control of Multiagent Systems in the Wasserstein Space. *Calculus of Variations and Partial Differential Equations*, 59:58, 2020.
- [60] R. Jordan, D. Kinderlehrer, and F. Otto. The Variational Formulation of the Fokker-Planck Equation. *SIAM Journal of Mathematical Analysis*, 29(1):1–17, 1998.
- [61] J.L. Kelley. *General Topology*, volume 27 of *Graduate Texts in Mathematics*. Springer, 1975.
- [62] J-M. Lasry and P.-L. Lions. Mean Field Games. *Japanese Journal of Mathematics*, 2(1):229–260, 2007.
- [63] B. Maury, A. Roudneff-Chupin, and F. Santambrogio. A Macroscopic Crowd Motion Model of Gradient Flow Type. *Mathematical Models and Methods in Applied Sciences*, 20(10):1787–1821, 2010.
- [64] B. Maury, A. Roudneff-Chupin, F. Santambrogio, and J. Venel. Handling Congestion in Crowd Motion Models. *Networks and Heterogeneous Media*, 6(3):485 – 519, 2011.
- [65] F. Otto. The Geometry of Dissipative Equations : The Porous Medium Equation. *Communications in Partial Differential Equations*, 26:101–174, 2001.
- [66] N.S. Papageorgiou. Random Fixed Point Theorems for Measurable Multifunctions in Banach Spaces. *Proceedings of the American Mathematical Society*, 97(3):507–514, 1986.
- [67] G. Peyré. Entropic Approximation of Wasserstein Gradient Flows. *SIAM Journal on Imaging Sciences*, 8(4):2323–2351, 2015.
- [68] B. Piccoli. Measure Differential Inclusions. In *2018 IEEE Conference on Decision and Control (CDC), Miami Beach, FL*, pages 1323–1328, 2018.
- [69] B. Piccoli. Measure Differential Equations. *Archive for Rational Mechanics and Analysis*, 233:1289–1317, 2019.
- [70] B. Piccoli and F. Rossi. Transport Equation with Nonlocal Velocity in Wasserstein Spaces : Convergence of Numerical Schemes. *Acta Applicandae Mathematicae*, 124(1):73–105, 2013.
- [71] B. Piccoli and F. Rossi. Measure Theoretic Models for Crowd Dynamics. In *Crowd Dynamics Volume 1 - Theory, Models, and Safety Problems*, pages 137–165. Birkhauser, 2018.
- [72] B. Piccoli, F. Rossi, and E. Trélat. Control to Flocking of the Kinetic Cucker-Smale model. *SIAM Journal on Mathematical Analysis*, 47(6):4685–4719, 2015.
- [73] W. Rudin. *Real and Complex Analysis*. Mathematical Series. McGraw-Hill International Editions, 1987.
- [74] F. Santambrogio. *Optimal Transport for Applied Mathematicians*, volume 87. Birkhauser Basel, 2015.
- [75] F. Santambrogio. {Euclidean, metric and Wasserstein} Gradient Flows : An Overview. *Bulletin of Mathematical Sciences*, 7:87–154, 2017.
- [76] A. Ülger. Weak Compactness in  $L^1(\mu, X)$ . *Proceedings of the American Mathematical Society*, 113:143–149, 1991.
- [77] C. Villani. *Optimal Transport : Old and New*. Springer-Verlag, Berlin, 2009.
- [78] R.B. Vinter. *Optimal Control*. Systems and Control: Foundations and Applications. Birkhauser Basel, 2000.
- [79] D.H. Wagner. Survey of Measurable Selection Theorems. *SIAM Journal on Control and Optimization*, 15(5):859–903, 1977.
- [80] S. Warner. The Topology of Compact Convergence on Continuous Function Spaces. *Duke Mathematical Journal*, 25(2):265–282, 1958.