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. 2023. hal-03969118

HAL Id: hal-03969118

https://hal.science/hal-03969118v1

Preprint submitted on 2 Feb 2023

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† and Hélène Frankowska‡

February 2, 2023

Abstract

In this article, we extend the foundations of the theory of differential inclusions in the space of probability measures recently laid down in one of our previous work to the setting of general Wasserstein spaces. Anchoring our analysis on novel estimates for solutions of continuity equations, we prove new variants of the Filippov theorem, compactness of solution set and relaxation theorem for continuity inclusions studied in the Cauchy-Lipschitz framework. We also propose an existence result "à la Peano" for this class of dynamics, under Carathéodory-type regularity assumptions. The latter is based on a set-valued generalisation of the semi-discrete Euler scheme originally proposed by Filippov to study ordinary differential equations with measurable right-hand sides.

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

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

Contents

1 Introduction **2**

2 Preliminaries **4**
2.1 Measure theory and optimal transport in Banach spaces 4
2.2 Set-valued analysis and topological properties of continuous functions 6
2.3 Continuity equations and inclusions in the space of probability measures 11

3 Well-posedness in the Cauchy-Lipschitz setting **15**
3.1 Existence of solutions and Filippov estimates 16
3.2 Compactness of the solution set for convex velocities 22
3.3 Relaxation property for non-convex velocities 24

4 Existence "à la Peano" in the Carathéodory setting **28**

Appendices **34**
A Proof of Lemma 2.10 35
B Proof of Lemma 2.21 36
C Proof of Proposition 2.22 38
D Proof of Proposition 2.24 39

* *This material is based upon work supported by the Air Force Office of Scientific Research under award number FA9550-18-1-0254, by the FMJH Program PGMO and the support to this program from EDF-THALES-ORANGE-CRITEO. This article was written in equal parts by both authors and is not subject to any conflict of interest with respect to Springer.*

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

‡ 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 a blooming interest in several mathematical communities. Even though 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 originating in equal parts from pure and applied mathematics motivated various kinds of explorations outside this paradigm.

Amongst these research endeavours, one of the most influential was certainly the development of the modern theory of optimal transport – rendered in the monographs [9, 80, 84] – along with that of gradient flows in Wasserstein spaces. The concepts introduced in the seminal papers [66, 71] and later formalised in the reference treatise [9] expounded the fact that, by applying a continuous-time steepest descent schemes to energy functionals defined over the space of probability measures, one could construct solutions to transport equations with irregular driving fields enjoyed. This simple yet 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 possessed a variational structure, see e.g. [9, 28, 30, 46, 61, 66, 69, 71, 81] and references therein. The ensuing infatuation for this innovative viewpoint stemmed both from its theoretical merits, and from its practical efficiency and adaptedness for designing numerical methods [13, 27, 73]. In a similar vein, the theory of mean-field games [26, 64, 68] – which is located halfway between control theory and the calculus of variations – largely contributed to popularising the study of dynamical problems in measure spaces. Incidentally, some of the core concepts of these emerging research schools found relevant application outlets in a wealth of multiscale models aiming at efficiently describe pedestrian dynamics [40, 70, 77], opinion propagation [2, 3, 62, 78], flocks and swarms [1, 29, 31] or macroscopic approximations of biological systems [15, 82]. Another active field of research that put this corpus of results to good use is that of mean-field control [17, 20, 32, 33, 49, 50, 51, 65], which saw the birth of several relevant extensions of the classical theory as further elaborated hereinbelow.

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

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

Therein, the initial datum $\mu^0 \in \mathscr{P}(\mathbb{R}^d)$ is a probability measure while $v : [0,T] \times \mathscr{P}(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ is a Lebesgue-Borel velocity field 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. [9, Chapter 11]), where $v(t,\mu(t)) \in L^p(\mathbb{R}^d, \mathbb{R}^d; \mu(t))$ happens to be the subgradients 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 agent cost satisfies a suitable convexity condition [68]. It is also worth noting that general existence results are available for very irregular driving fields when the dynamics exhibits a Hamiltonian structure [8].

As amply highlighted by the discussions in [9, Chapter 8], the dynamics in (1.1) admits a natural interpretation as an ordinary differential equation in the Wasserstein "manifolds" – or bundles – $(\mathscr{P}_p(\mathbb{R}^d), W_p)$. Thus, in the absence of an underlying variational structure, one should expect that the well-posedness of said Cauchy problem would 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 depends on the measure variable or not. When $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ is independent of the measure variable, the 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 Sobolev [42] or BV [4] regularity in the space variable, combined with integral bounds on the divergence or incompressibility assumptions [16], 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 [6]. We also mention the article [5] in which a local existence theory akin to the classical Peano one for ODEs is developed, as well as the older paper [38] wherein sharp stability estimates are derived for the DiPerna-Lions flow. On the other hand when $v : [0,T] \times \mathscr{P}(\mathbb{R}^d) \times \mathbb{R}^d \mapsto \mathbb{R}^d$ 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. [17, 39, 65, 76] 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 adverting to a recent body of work initiated in [75] and furthered in [22, 35, 74], in which well-posedness results for relatives of (1.1) and (1.2) in which the driving fields are replaced by probability measures on the tangent bundle are investigated. This line of study – which is highly reminiscent of [14] – bears strong resemblance with the theory of Young measures [10, Section 4.3], and has already produced very promising results shedding light on the interplay between contraction semigroups in measure spaces and various kinds of explicit Euler schemes.

The aim of this paper is to provide several important and far-reaching refinements of the theory of *continuity inclusions* in Wasserstein spaces, whose elaboration started in our previous work [17]. Therein, given a compactly supported datum $\mu^0 \in \mathscr{P}_c(\mathbb{R}^d)$ and a *set-valued map* $(t,\mu) \in [0,T] \times \mathscr{P}_c(\mathbb{R}^d) \rightrightarrows V(t,\mu) \subset L^p(\mathbb{R}^d, \mathbb{R}^d; \mu)$, we define the solution set of the Cauchy problem

$$\begin{cases} \partial_t \mu(t) \in -\mathrm{div}_x\Big(V(t,\mu(t))\mu(t)\Big), \\ \mu(0) = \mu^0, \end{cases} \tag{1.2}$$

as the collection of absolutely continuous curves $\mu(\cdot) \in \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d))$ for which there exists an $\mathscr{L}^1$*-measurable selection* $t \in [0,T] \mapsto v(t) \in V(t,\mu(t))$ such that (1.1) holds. Compared with other notions that were put forth to define differential inclusions in measure spaces, such as those of [34, 65], our approach presents the advantage of being coherent with the modern theory of differential inclusions in vector spaces, surveyed e.g. in [11, 12], and conceptually compatible with the geometric structure of Wasserstein spaces, as well as with the interpretation of continuity equations as generalised ODEs following [9, 71]. 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 dynamics and of their set-valued counterparts.

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 monographs [11, 12, 37, 85] and [25, 36]. They provide a convenient setting that allows to prove existence results [48] for variational problems, to derive first- and second-order optimality conditions – both in the form of a Pontryagin Maximum Principle [52, 57, 58, 60] or Hamilton-Jacobi-Bellman equations [53, 59] –, and to investigate qualitative properties of optimal trajectories [24]. Owing to their mathematical versatility, these schemes have recently started to percolate in the communities of mean-field control and mean-field games [18, 19, 23, 32, 65, 74]. With the goal of transposing these powerful and tried concepts to the setting of mean-field control, we generalised in [17] the cornerstones of the theory of Cauchy-Lipschitz differential inclusions to dynamics formulated in the space $(\mathscr{P}_c(\mathbb{R}^d), W_p)$ of compactly supported measures endowed with a Wasserstein metric. The first of these key results are the Filippov estimates, which provide the existence of a solution to (1.2) whose distance to an a priori fixed measure curve is controlled, and that is extremely useful to produce admissible trajectories when conducting perturbative and linearisation arguments. The second one is the compactness of the set of trajectories when the right-hand side of the dynamics is convex, which is naturally needed in virtually every existence proof based on weak compactness arguments, both in optimal control theory and in the calculus of variations. The third one 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. In addition to its topological relevance, this result is widely used for the derivation of sharp first- and second-order optimality conditions for optimal control problems in extremely varied contexts, see e.g. [54, 56, 57] for ordinary differential equations, [55] for infinite-dimensional control problems, and our previous work [18] concerned with mean-field optimal control problems. 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.

As underlined in the previous paragraph, the analyses and results of [17] were confined to curves of measures whose supports are contained in some compact set. While this modelling assumption does permit to handle most of the commonly encountered applications in population dynamics, mean-field control and mean-field games, this limitation is not satisfactory with regard to the preexisting literature in optimal transport and in the calculus of variations, or to their more recent offsprings such as [35]. Furthermore, while it is easier from a purely technical standpoint to work with compactly supported measures, one could argue that this constraint may somewhat obscure the exact role of some of the objects involved in the definition of continuity inclusions, as well as in the proofs of the corresponding structure theorems. For the aforedescribed reasons, it is crucial to extend the theory of continuity inclusions to the more general and streamlined setting of measure dynamics formulated in $(\mathscr{P}_p(\mathbb{R}^d), W_p)$. We 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 measures may have unbounded supports –, the corresponding results strictly contain those derived in [17] up to minor technical adjustments. Besides achieving greater generality and making the theory of continuity inclusions more synthetic, revisiting our previous contributions was the occasion to derive new quantitative stability estimates for continuity equations, which should constitute a fine addition to the tooling used to investigate dynamics in measure spaces. Moreover, we also propose a completely novel existence result "à la Peano" for continuity inclusions whose velocities are merely continuous instead of Lipschitz, by astutely generalising to the set-valued framework a variant of the explicit Euler scheme due to Filippov for ordinary differential equations, see [47, Chapter 1]

The contributions and organisation of the article can be summarised as follows. In Section 2, after recalling a broad range of 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 2.22 and Proposition 2.24 respectively. We then define solutions of (1.2) for set-valued maps $V : [0,T] \times \mathscr{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d, \mathbb{R}^d)$ in terms of velocity selections which are measurable for the standard Fréchet topology of the space of continuous functions (see Definition 2.12 below), and subsequently move on in Section 3 to the three main results of the theory of differential inclusions studied in the Cauchy-Lipschitz setting. In Section 3.1, we start by establishing two far-reaching versions – a local one in Theorem 3.3 and a global one in Corollary 3.5 – of the Filippov estimates. Then, in Theorem 3.6 of Section 3.2, we show that the set of solution to (1.2) is compact for the topology of uniform convergence when its right-hand side is convex-valued. In Section 3.3, we proceed to relax this latter assumption, and prove in Theorem 3.7 that in this case, the closure of the solution set coincides with that of the convexified Cauchy problem. Finally in Section 4, we abandon the Cauchy-Lipschitz framework, and provide in Theorem 4.1 an existence result for solutions of (1.2) when the right-hand side is Caratheodory and convex-valued. To the best of our knowledge, the proof strategy developed therein, based on a variation of the delayed semi-discrete scheme of [47, Theorem 1], is completely new even in the context of classical differential inclusions with measurable right-hand sides. We then close the paper by an appendix containing the proofs of several technical results and estimates. Even though some of these partly rely on ideas similar to those of [17, Appendices A and B], we expose them here for the sake of completeness and self-containedness.

2 Preliminaries

In the coming sections, we expose a 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 the monographs [7, 44] and [9] respectively.

Elements of measure theory and integration. 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 bracket. In the sequel, the notation $C^0_b(X, \mathbb{R}^d)$ 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 $C^\infty_c(\mathbb{R}^m, \mathbb{R})$ the vector space of infinitely differentiable functions with compact support from $\mathbb{R}^m$ into $\mathbb{R}$. 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 $\mathrm{AC}(I, Y)$ for that of absolutely continuous arcs defined over an interval $I \subset \mathbb{R}$ with values in Y. In addition, $\mathrm{Lip}(X, Y)$ will stand for the set of Lipschitz maps from X into Y, and we shall write $\mathrm{Lip}(\phi\,; X)$ for the Lipschitz constant of an element $\phi \in \mathrm{Lip}(X, Y)$.

In what follows, we will consider the vector space $\mathscr{M}(K, \mathbb{R}^d)$ of $\mathbb{R}^d$-valued Radon measures defined over a compact set $K \subset \mathbb{R}^d$. By the Riesz representation theorem (see e.g. [7, Theorem 1.54]), it is known that the latter is isomorphic to the topological dual of the Banach space $(C^0(K, \mathbb{R}^d), \|\cdot\|_{C^0(K,\mathbb{R}^d)})$ under the action of the duality pairing

$$\langle \boldsymbol{\nu}, \phi \rangle_{C^0(K,\mathbb{R}^d)} := \sum_{i=1}^d \int_K \phi_i(x) \mathrm{d}\boldsymbol{\nu}_i(x), \tag{2.1}$$

defined for all $\boldsymbol{\nu} \in \mathscr{M}(K, \mathbb{R}^d)$ and $\phi \in C^0(K, \mathbb{R}^d)$. Throughout the article, we denote by $\mathscr{P}(X)$ the space of Borel probability measures over X endowed with the *narrow topology*, i.e. the coarsest topology for which the applications

$$\mu \in \mathscr{P}(X) \mapsto \int_X \phi(x) \mathrm{d}\mu(x) \tag{2.2}$$

are continuous for every element $\phi \in C^0_b(X, \mathbb{R})$. It is a standard fact in measure theory (see e.g. [9, Remark 5.1.2]) that $\mathscr{P}(X)$ is a Polish space, and we will write

$$\mu_n \underset{n \to +\infty}{\rightharpoonup^*} \mu,$$

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

Given two separable Banach spaces $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ along with some $p \in [1, +\infty)$, we will write $(L^p(X, Y; \mu), \|\cdot\|_{L^p(X,Y;\mu)})$ for the space of p-integrable maps with respect to a measure $\mu \in \mathscr{P}(X)$, defined in the sense of Bochner (see e.g. [44, Chapter II]). Analogously, we will let $L^\infty(X, Y; \mu)$ be the space of μ-essentially bounded maps from X into Y, and use the denser notation $(L^p(I, Y), \|\cdot\|_{L^p(I)})$ when $X = I$ is an interval and $\mu = \mathscr{L}^1$ is the standard 1-dimensional Lebesgue measure. We recall below a powerful sufficient condition for weak compactness in $L^1(I, X)$, whose statement can be found in [41, Corollary 2.6] (see also the earlier contributions of [43, 83]).

Theorem 2.1 (A weak compactness criterion for Bochner integrable maps)**.** *Let $(X, \|\cdot\|_X)$ be a 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*

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

for $\mathscr{L}^1$-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

$$\int_I \langle \boldsymbol{\phi}(t), v(t) - v_{n_k}(t) \rangle_X \mathrm{d}t \underset{k \to +\infty}{\longrightarrow} 0,$$

for every $\boldsymbol{\phi}(\cdot) \in L^\infty(I, X^*) \subset L^1(I, X)^*$.

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

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

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

Definition 2.2 (Image measures and transport plans)**.** *The* pushforward *of a measure $\mu \in \mathscr{P}(X)$ through a Borel map $f : X \to Y$ – denoted by $f_\sharp \mu \in \mathscr{P}(Y)$ – is defined by*

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

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

$$\pi^1_\sharp \gamma = \mu \qquad \text{and} \qquad \pi^2_\sharp \gamma = \nu,$$

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

Throughout the remainder of this subsection, we will assume that $(X, \|\cdot\|_X) := (\mathbb{R}^d, |\cdot|)$ is a d-dimensional real vector space endowed with its usual Euclidean structure. For a real number $p \in [1, +\infty)$, it is a standard result in optimal transport theory that the quantity defined by

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

for each $\mu, \nu \in \mathscr{P}_p(\mathbb{R}^d)$ is a distance over $\mathscr{P}_p(\mathbb{R}^d)$. 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 so-called *Wasserstein spaces*, along with handy distance estimates.

Proposition 2.3 (Properties of Wasserstein spaces)**.** *The spaces $(\mathscr{P}_p(\mathbb{R}^d), W_p)$ are complete separable metric spaces on which the W_p-distance metrises the narrow topology (2.2), in the sense that*

$$W_p(\mu_n, \mu) \underset{n\to+\infty}{\longrightarrow} 0 \qquad \text{if and only if} \qquad \begin{cases} \mu_n \underset{n\to+\infty}{\rightharpoonup^*} \mu, \\ \displaystyle\int_{\mathbb{R}^d} |x|^p \mathrm{d}\mu_n(x) \underset{n\to+\infty}{\longrightarrow} \int_{\mathbb{R}^d} |x|^p \mathrm{d}\mu(x), \end{cases}$$

for any sequence $(\mu_n) \subset \mathscr{P}_p(\mathbb{R}^d)$ and each $\mu \in \mathscr{P}_p(\mathbb{R}^d)$. Moreover, a subset $\mathscr{K} \subset \mathscr{P}_p(\mathbb{R}^d)$ is relatively compact with respect to the W_p-metric if and only if *it satisfies*

$$\sup_{\mu \in \mathscr{K}} \int_{\{x \text{ s.t. } |x| \geq k\}} |x|^p \mathrm{d}\mu(x) \underset{k\to+\infty}{\longrightarrow} 0.$$

Proof. We point the interested reader to [9, Chapter 7] or [84, Chapter 6]. □

Proposition 2.4 (Classical optimal transport estimate)**.** *Let $\mu, \nu \in \mathscr{P}_p(\mathbb{R}^d)$ and consider a map $\phi \in \mathrm{Lip}(\mathbb{R}^d, \mathbb{R})$. Then, it holds that*

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

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 reference monographs [11, 12] for the former and [63, 79] for the latter.

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

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

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

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

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 a *multifunction* – from X into Y. A set-valued map $\mathcal{F}(\cdot)$ is said to have closed (respectively convex) images if the sets $\mathcal{F}(x) \subset Y$ are closed (respectively convex) for each $x \in X$. In addition, we define the *graph* of $\mathcal{F} : X \rightrightarrows Y$ by

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

Below, we recall a standard measurability concept for set-valued mappings defined over subintervals of the real line, endowed with the complete Lebesgue σ-algebra.

Definition 2.5 (Measurable set-valued maps and measurable selections)**.** *A multifunction $\mathcal{F} : [0, T] \rightrightarrows Y$ is said to be* $\mathscr{L}^1$-measurable *provided that the sets*

$$\mathcal{F}^{-1}(\mathcal{O}) := \Big\{t \in [0, T] \ \text{ s.t. } \mathcal{F}(t) \cap \mathcal{O} \neq \emptyset\Big\}$$

are $\mathscr{L}^1$-measurable for every open set $\mathcal{O} \subset Y$. A mapping $f : [0, T] \to Y$ is called a measurable selection *of $\mathcal{F}(\cdot)$ if it is $\mathscr{L}^1$-measurable and such that $f(t) \in \mathcal{F}(t)$ for $\mathscr{L}^1$-almost every $t \in [0, T]$.*

In the following theorem, we recollect a deep result of set-valued analysis which provides the existence of measurable selections for measurable multifunctions with closed nonempty images in complete separable metric spaces.

Theorem 2.6 (Existence of measurable selections)**.** *If $\mathcal{F} : [0, T] \rightrightarrows Y$ is an $\mathscr{L}^1$-measurable set-valued map with closed and nonempty images, then it admits an $\mathscr{L}^1$-measurable selection.*

In our subsequent developments, we will resort to 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 metric 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 $\Omega \subset X$.

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*

$$\mathcal{F}(x') \subset \mathbb{B}_Y\Big(\mathcal{F}(x)\,,\, L d_X(x, x')\Big),$$

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 $\mathcal{G} : [0, T] \times X \rightrightarrows Y$ is said to be* Carathéodory *if the application $t \in [0, T] \rightrightarrows \mathcal{G}(t, x)$ is $\mathscr{L}^1$-measurable for all $x \in X$, and the map $x \in X \rightrightarrows \mathcal{G}(t, x)$ is continuous for $\mathscr{L}^1$-almost every $t \in [0, T]$.*

In the following lemma – whose proof is outlined in Appendix A –, we state measurable selections principles adapted from [12, Section 8.1] and [86, Section 9] that we shall extensively use in the sequel. Therein, we let $\varphi : [0,T] \times Y \to \mathbb{R}_+ \cup \{+\infty\}$ be an extended real-valued function satisfying

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

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and all $y \in Y$, where $(\varphi_n(\cdot,\cdot))$ is a pointwisely non-decreasing sequence of Carathéodory maps. This implies in particular that $t \in [0,T] \mapsto \varphi(t,y)$ is $\mathscr{L}^1$-measurable for each $y \in Y$, whereas $y \in Y \mapsto \varphi(t,y)$ is lower-semicontinuous for $\mathscr{L}^1$-almost every $t \in [0,T]$.

Lemma 2.10 (Measurable selections principles)**.** *Let $\mathcal{F} : [0,T] \rightrightarrows Y$ and $\mathcal{G} : [0,T] \times X \rightrightarrows Y$ be two set-valued maps with nonempty images. In addition, fix two $\mathscr{L}^1$-measurable functions $x : [0,T] \to X$ and $L : [0,T] \to \mathbb{R}_+$.*

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

$$t \in [0,T] \rightrightarrows \mathcal{F}(t) \cap \Big\{ y \in Y \ \text{ s.t. } \varphi(t,y) \leq L(t) \Big\}$$

has nonempty images. Then the latter is $\mathscr{L}^1$-measurable, and there exists a measurable selection $t \in [0,T] \mapsto f(t) \in \mathcal{F}(t)$ such that $\varphi(t, f(t)) \leq L(t)$ for $\mathscr{L}^1$-almost every $t \in [0,T]$.

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

$$t \in [0,T] \rightrightarrows \mathcal{F}(t) \cap \Big\{ y \in Y \ \text{ s.t. } \varphi(t,y) = \inf_{z \in \mathcal{F}(t)} \varphi(t,z) \Big\}$$

is $\mathscr{L}^1$-measurable with closed nonempty 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 $t \in [0,T] \rightrightarrows \mathcal{G}(t, x(t))$ is $\mathscr{L}^1$-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 [12, Theorem 8.6.4].

Theorem 2.11 (Convexity of the Aumann integral)**.** *Suppose that $(X, \|\cdot\|_X)$ is a separable Banach space and let $\mathcal{F} : [0,T] \to X$ be a set-valued map with closed nonempty images that is* integrably bounded*, in the sense that there exists $k(\cdot) \in L^1([0,T], \mathbb{R}_+)$ such that*

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

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

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

Topologies and metrics over $C^0(\mathbb{R}^d, \mathbb{R}^d)$. Throughout the remainder of this manuscript, we will almost exclusively work with multifunctions valued in the power set of $C^0(\mathbb{R}^d, \mathbb{R}^d)$. Following the pioneering work [87], 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*

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

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 metrised by the complete, separable and translation-invariant metric

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

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

Remark 2.13 (Link with the compact-open topology)**.** *It is a standard – albeit nontrivial – fact in functional analysis (see e.g. [67, Chapter 7 – Theorem 11]) that the topology of uniform convergence on compact sets defined over $C^0(\mathbb{R}^d, \mathbb{R}^d)$ coincides with the so-called* compact-open *topology. While the former is usually considered when studying maps between metric spaces, the latter can be defined more generally over the space of continuous functions between topological spaces via bases of neighbourhoods.*

In our subsequent developments, we will always consider $C^0(\mathbb{R}^d, \mathbb{R}^d)$ as a separable Fréchet space equipped with the metric $\mathsf{d}_{\mathrm{cc}}(\cdot,\cdot)$. One of the interesting features of this latter is that it provides an equivalent functional characterisation of Carathéodory vector fields, as illustrated by the following result whose proof can be found e.g. in [72, page 511].

Lemma 2.14 (Carathéodory vector fields as measurable functions in $C^0(\mathbb{R}^d, \mathbb{R}^d)$)**.** *A vector field $(t, x) \in [0, T] \times \mathbb{R}^d \mapsto v(t, x) \in \mathbb{R}^d$ is Carathéodory* if and only if *its functional lift $t \in [0, T] \mapsto v(t) \in C^0(\mathbb{R}^d, \mathbb{R}^d)$ is $\mathscr{L}^1$-measurable for the topology of uniform convergence on compact sets.*

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

$$\mathscr{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d)) := \left\{v : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d \text{ s.t. } v(\cdot,\cdot) \text{ is a Carathéodory vector field}\right\}.$$

In addition to its amenable topological properties, the notion of uniform convergence on compact sets is particularly well-tailored to formulate compactness results, as illustrated by the following adaptation of the Ascoli-Arzelà theorem from [79, Theorem 11.28].

Theorem 2.15 (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*

$$\mathsf{d}_{\mathrm{cc}}(v_{n_k}, v) \underset{k\to+\infty}{\longrightarrow} 0,$$

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

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

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 [79, Theorem 11.28], while that of the second one simply follows from the fact that the topology induced by $\mathsf{d}_{\mathrm{cc}}(\cdot,\cdot)$ on $C^0(K, \mathbb{R}^d)$ coincides with the standard topology of uniform convergence. □

In order to formulate the well-posedness results of Section 3, we will also need a global notion of vicinity for continuous maps, that will inherently be stronger than the local one described in Definition 2.12. For this reason, we will also consider in the sequel the extended supremum metric defined by

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

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 $\mathsf{d}_{\sup}(\cdot,\cdot)$ to the topological vector space $\mathscr{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$ is complete, in the sense that its Cauchy sequences converge.

Lemma 2.16 (Completeness of the integral of the extended metric)**.** *Consider a sequence of maps* $(v_n(\cdot)) \subset \mathscr{L}([0,T], C^0(\mathbb{R}^d, \mathbb{R}^d))$ *satisfying the Cauchy condition*

$$\int_0^T \mathsf{d}_{\sup}(v_n(t), v_m(t))\mathrm{d}t \underset{n,m\to+\infty}{\longrightarrow} 0. \tag{2.8}$$

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

$$\int_0^T \mathsf{d}_{\sup}(v_n(t), v(t))\mathrm{d}t \underset{t\to+\infty}{\longrightarrow} 0.$$

Proof. The proof of this result is a transposition of standard arguments used to show that the Lebesgue spaces are complete, and for which we refer e.g. to [79, Theorem 3.11]. Let $(v_n(\cdot)) \subset \mathscr{L}([0,T], C^0(\mathbb{R}^d, \mathbb{R}^d))$ be a Cauchy sequence in the sense of (2.8), which can be chosen so that

$$\int_0^T \mathsf{d}_{\sup}(v_n(t), v_{n+1}(t))\mathrm{d}t \leq \frac{1}{2^n} \tag{2.9}$$

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

$$d_n(t) := \sum_{k=1}^{n} \mathsf{d}_{\sup}(v_k(t), v_{k+1}(t)), \tag{2.10}$$

defined for $\mathscr{L}^1$-almost every $t \in [0,T]$ and each $n \geq 1$, and observe that by (2.9), one has that

$$\sup_{n\geq 1} \int_0^T d_n(t)\mathrm{d}t \leq \sum_{k=1}^{+\infty} \int_0^T \mathsf{d}_{\sup}(v_k(t), v_{k+1}(t))\mathrm{d}t \leq 1. \tag{2.11}$$

By [79, Theorem 1.38], the partial sums of series defined in (2.10) converge for $\mathscr{L}^1$-almost every $t \in [0,T]$ towards a map $d(\cdot) \in L^1([0,T], \mathbb{R}_+)$, which must then satisfy

$$\int_0^T d(t)\mathrm{d}t \leq 1.$$

Thus, it necessarily follows that $d(t) < +\infty$ for $\mathscr{L}^1$-almost every $t \in [0,T]$, which by (2.10) together with the definition (2.7) of the extended metric $\mathsf{d}_{\sup}(\cdot,\cdot)$ implies that the series of functions defined by

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

is normally convergent for $\mathscr{L}^1$-almost every $t \in [0,T]$. Consequently, for every $(t,x) \in [0,T] \times \mathbb{R}^d$, we can define the mapping $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ 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 $\mathscr{L}^1$-measurable with respect to $t \in [0,T]$ and continuous with respect to $x \in \mathbb{R}^d$. Moreover, the map $v(\cdot) \in \mathscr{L}([0,T], C^0(\mathbb{R}^d, \mathbb{R}^d))$ is the pointwise limit of the sequence $(v_n(\cdot))$, since

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

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. [79, Statement 1.28]), one can finally show that

$$\int_0^T \mathsf{d}_{\sup}(v_n(t), v(t))\mathrm{d}t \underset{n\to+\infty}{\longrightarrow} 0,$$

which concludes the proof of our claim. ∎

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 [6, 9], as well as their set-valued counterpart introduced by the authors of the present manuscript in [17]. Given a real number $p \in [1, +\infty)$, a measure $\mu^0 \in \mathscr{P}_p(\mathbb{R}^d)$ and a velocity field $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, we will focus our attention on the well-posedness of the Cauchy problem

$$\begin{cases} \partial_t \mu(t) + \mathrm{div}_x(v(t)\mu(t)) = 0, \\ \mu(0) = \mu^0, \end{cases} \tag{2.12}$$

wherein $\mu(\cdot) \in C^0([0,T], \mathscr{P}(\mathbb{R}^d))$ is a narrowly continuous curve of measure. The dynamics appearing in the first line of (2.12) is a *continuity equation*, that is understood in the sense of distributions as

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

for every test function $\phi \in C_c^\infty((0,T) \times \mathbb{R}^d, \mathbb{R})$. Throughout this article, we will mainly work with velocity fields satisfying the following set of assumptions, which are classical when studying continuity equations in the Cauchy-Lipschitz framework, and provide standard well-posedness results for (2.12).

Hypotheses (CE).

(*i*) *The vector field* $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ *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 $\mathscr{L}^1$*-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*

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

for $\mathscr{L}^1$*-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 [17] 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 $(\mathscr{P}_p(\mathbb{R}^d), W_p)$. In the following definition, we recall the notion of *flows of diffeomorphisms* generated by a Carathéodory vector field.

Definition 2.17 (Flows of diffeomorphisms)**.** *Let* $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ *be a velocity field satisfying hypotheses* **(CE)**. *Then, we denote by* $(\Phi^v_{(0,t)}(\cdot))_{t\in[0,T]} \subset C^0(\mathbb{R}^d, \mathbb{R}^d)$ *the unique* semigroup of diffeomorphisms *that solve the Cauchy problems*

$$\Phi^v_{(0,0)}(x) = x \qquad \text{and} \qquad \partial_t \Phi^v_{(0,t)}(x) = v(t, \Phi^v_{(0,t)}(x)),$$

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

Under the Cauchy-Lipschitz assumptions stated in hypotheses **(CE)**, the following strong well-posedness result holds for solutions of continuity equations in $(\mathscr{P}_p(\mathbb{R}^d), W_p)$. Therein and 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 2.18 (Classical well-posedness of continuity equations)**.** *Let* $\mu^0 \in \mathscr{P}_p(\mathbb{R}^d)$ *and* $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ *be a velocity field satisfying hypotheses* **(CE)**-(*i*). *Then, the Cauchy problem* (2.12) *admits solutions* $\mu(\cdot) \in \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d))$, *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*

$$W_p(\mu(\tau), \mu(t)) \leq c_p \int_\tau^t m(s)\mathrm{d}s, \tag{2.13}$$

for all times $0 \leq \tau \leq t \leq T$*. If in addition the velocity field* $v(\cdot,\cdot)$ *satisfies hypothesis* **(CE)**-(ii)*, then the solution* $\mu(\cdot) \in \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d))$ *of* (2.12) *is unique and can be represented explicitly as*

$$\mu(t) = \Phi^v_{(0,t)}(\cdot)_\sharp \mu^0, \tag{2.14}$$

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

Proof. These statements follow e.g. from a combination of several standard results from [9, Section 8.1] along with the momentum estimates displayed in Proposition 2.22 below. □

Definition 2.17 and Theorem 2.18 together inform us 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 using the concept of *superposition measure*. In what follows, we will write $\Sigma_T := C^0([0,T], \mathbb{R}^d)$ for the space of continuous arcs from $[0,T]$ into $\mathbb{R}^d$.

Definition 2.19 (Superposition measures)**.** *An element* $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Sigma_T)$ *is called a* superposition measure *associated with a Lebesgue-Borel velocity field* $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ *if it is concentrated on the set of pairs* $(x,\sigma) \in \mathbb{R}^d \times \mathrm{AC}([0,T], \mathbb{R}^d)$ *satisfying*

$$\sigma(0) = x \qquad \text{and} \qquad \dot{\sigma}(t) = v(t, \sigma(t)), \tag{2.15}$$

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

A direct link can be provided between superposition measures and solutions of (2.12) under the action of the so-called *evaluation map*, which is defined for all times $t \in [0,T]$ by

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

More precisely, it can be checked that if $\boldsymbol{\eta} \in \mathscr{P}(\mathbb{R}^d \times \Sigma_T)$ is a superposition measure associated with a velocity field $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ such that $(\pi_{\mathbb{R}^d})_\sharp \boldsymbol{\eta} = \mu^0$, 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 which satisfies the local integrability condition

$$\int_0^T \int_{\mathbb{R}^d \times \Sigma_T} \mathbb{1}_K(\sigma(t)) |v(t,\sigma(t))| \mathrm{d}\boldsymbol{\eta}(x,\sigma) \mathrm{d}t < +\infty$$

for each compact set $K \subset \mathbb{R}^d$, then the curve defined by $\mu(t) := (\mathsf{e}_t)_\sharp \boldsymbol{\eta}$ for all times $t \in [0,T]$ is a solution of (2.12). 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 [6, Theorem 3.4].

Theorem 2.20 (Superposition principle)**.** *Let* $\mu^0 \in \mathscr{P}_p(\mathbb{R}^d)$ *and* $\mu(\cdot) \in \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d))$ *be a solution of* (2.12) *driven by a Lebesgue-Borel velocity field* $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ *satisfying*

$$\int_0^T \int_{\mathbb{R}^d} \frac{|v(t,x)|}{1+|x|} \mathrm{d}\mu(t)(x) \mathrm{d}t < +\infty. \tag{2.16}$$

Then, there exists a superposition measure $\boldsymbol{\eta}_\mu \in \mathscr{P}(\mathbb{R}^d \times \Sigma_T)$ *in the sense of Definition 2.19 such that* $(\mathsf{e}_t)_\sharp \boldsymbol{\eta}_\mu = \mu(t)$ *for all times* $t \in [0,T]$*.*

Throughout the remainder of this section, we derive a series of useful momentum and stability estimates for solutions of (2.12) driven by Lebesgue-Borel velocity fields $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, under the additional assumption that there exists a map $m(\cdot) \in L^1([0,T], \mathbb{R}_+)$ such that

$$|v(t,x)| \leq m(t)(1+|x|) \tag{2.17}$$

holds for $\mathscr{L}^1$-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 [17], the results that we discuss here are both new and proven under less restrictive regularity assumptions. We start by stating an adaptation of a structural result established in [17, Lemma 1], whose proof is detailed in Appendix B.

Lemma 2.21 (Superposition plans inducing optimal transports)**.** *Consider two curves of measures $\mu(\cdot), \nu(\cdot) \in \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d))$ driven by the Lebesgue-Borel velocity fields $v, w : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ complying with the pointwise sublinearity estimates*

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

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and $\mu(t) \times \nu(t)$-almost every $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$. Moreover, let $\boldsymbol{\eta}_\mu, \boldsymbol{\eta}_\nu \in \mathscr{P}(\mathbb{R}^d \times \Sigma_T)$ be two superposition measures given by Theorem 2.20, such that

$$\mu(t) = (\mathsf{e}_t)_\sharp \boldsymbol{\eta}_\mu \qquad \text{and} \qquad \nu(t) = (\mathsf{e}_t)_\sharp \boldsymbol{\eta}_\nu$$

for all times $t \in [0,T]$. Then, for every $\gamma^0 \in \Gamma_o(\mu^0, \nu^0)$, there exists a plan $\hat{\boldsymbol{\eta}}_{\mu,\nu} \in \Gamma(\boldsymbol{\eta}_\mu, \boldsymbol{\eta}_\nu)$ such that

$$(\pi_{\mathbb{R}^d}, \pi_{\mathbb{R}^d})_\sharp \hat{\boldsymbol{\eta}}_{\mu,\nu} = \gamma^0 \qquad \text{and} \qquad (\mathsf{e}_t, \mathsf{e}_t)_\sharp \hat{\boldsymbol{\eta}}_{\mu,\nu} \in \Gamma_o(\mu(t), \nu(t))$$

for all times $t \in [0,T]$, where Γ_o stands here for the set of p-optimal transport plans.

In what follows, we leverage the general superposition results of Theorem 2.20 and Lemma 2.21 to prove momentum and equi-integrability inequalities for solutions of (2.12) in Proposition 2.22, as well as two stability estimates with respect to initial data and driving fields in Proposition 2.24. For the sake of readability, we postpone the proof of these results to Appendix C and Appendix D respectively.

Proposition 2.22 (Momentum and equi-integrability estimates)**.** *Let $\mu^0 \in \mathscr{P}_p(\mathbb{R}^d)$ be given and $\mu(\cdot) \in \mathrm{AC}([0,T], \mathscr{P}(\mathbb{R}^d))$ be a solution of (2.12) driven by a Lebesgue-Borel velocity field $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ satisfying the sublinearity estimate (2.17). Then, the following momentum bound*

$$\mathcal{M}_p(\mu(t)) \leq C_p \Big(\mathcal{M}_p(\mu^0) + \int_0^t m(s) \mathrm{d}s \Big) \exp\Big(C'_p \, \|m(\cdot)\|_1^p \Big), \tag{2.18}$$

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

$$C_p := 2^{(p-1)/p} \qquad \text{and} \qquad C'_p := \tfrac{2^{p-1}}{p}. \tag{2.19}$$

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

$$\sup_{t \in [0,T]} \int_{\{x \text{ s.t. } |x| \geq R\}} |x|^p \, \mathrm{d}\mu(t)(x) \;\leq\; C_T^p \int_{\{x \text{ s.t. } |x| \geq R/C_T - 1\}} (1+|x|)^p \, \mathrm{d}\mu^0(x), \tag{2.20}$$

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

Remark 2.23 (A refined momentum inequality)**.** *In the sequel, we will often use the fact that, if a velocity field $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ satisfies the slightly more general sublinearity inequality*

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

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

$$\mathcal{M}_p(\mu(t)) \leq C_p \Big(\mathcal{M}_p(\mu^0) + \int_0^t m(s)(1+M(s)) \mathrm{d}s \Big) \exp\Big(C'_p \, \|m(\cdot)\|_{L^1([0,t]}^p \Big),$$

for all times $t \in [0,T]$. The latter inequality can be established by repeating verbatim the arguments detailed in Appendix C.

Proposition 2.24 (Two general stability estimates for continuity equations)**.** *Let $\mu^0, \nu^0 \in \mathscr{P}_p(\mathbb{R}^d)$ and $\mu(\cdot), \nu(\cdot) \in \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d))$ be two solutions of (2.12), driven respectively by a Carathéodory velocity field $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ satisfying hypotheses* **(CE)** *and a Lebesgue-Borel velocity field $w : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ complying with the sublinearity inequality*

$$|w(t,y)| \leq m(t)(1+|y|), \tag{2.21}$$

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

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

is Lebesgue integrable, the following global stability estimate

$$W_p(\mu(t),\nu(t)) \leq C_p\Big(W_p(\mu^0,\nu^0) + \int_0^t \|v(s) - w(s)\|_{L^\infty(\mathbb{R}^d,\mathbb{R}^d;\,\nu(s))}\,\mathrm{d}s\Big)\exp\Big(C'_p\,\|l(\cdot)\|^p_{L^1([0,t])}\Big), \tag{2.22}$$

holds for all times $t \in [0,T]$, wherein $C_p, C'_p > 0$ are defined in (2.19). More generally, under our assumptions, the application

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

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

$$\begin{aligned} W_p(\mu(t),\nu(t)) \leq C_p\Big(W_p(\mu^0,\nu^0) + \int_0^t \|v(s) - w(s)\|_{L^\infty(B(0,R),\mathbb{R}^d;\,\nu(s)))}\,\mathrm{d}s + \mathcal{E}_\nu(t,R)\Big) \\ \times \exp\Big(C'_p\,\|l(\cdot)\|^p_{L^1([0,t])}\Big), \end{aligned} \tag{2.23}$$

holds for all times $t \in [0,T]$, where the error term is defined by

$$\mathcal{E}_\nu(t,R) := 2\,\|m(\cdot)\|_{L^1([0,t])}\,(1 + C_T)\Big(\int_{\{y \text{ s.t. } |y| \geq R/C_T - 1\}} (1+|y|)^p \mathrm{d}\nu^0(y)\Big)^{1/p}. \tag{2.24}$$

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

Remark 2.25 (Comparison with the estimates of [17])**.** *The stability estimates displayed in Proposition 2.24 above improve on those of [17] in the two following ways. Firstly, the global inequality (2.22) now holds for a general Lebesgue-Borel velocity field $w(\cdot,\cdot)$ satisfying the sublinearity inequality (2.17), without requiring the latter to be Carathéodory. Secondly, the inequality (2.23) involving the quantitative error term $\mathcal{E}_\nu(\cdot,\cdot)$ is completely new to the best of our knowledge, and will prove crucial to palliate the fact that the metric of compact convergence $\mathsf{d}_{\mathrm{cc}}(\cdot,\cdot)$ only grants access to local and non-uniform estimates between functions.*

We end this preliminary section by recalling the definition of *continuity inclusions* in Wasserstein spaces. The latter, introduced by the authors of the present manuscript in [17], and constitutes a natural set-valued generalisation of continuity equations, as its definition mimics the well-established viewpoint that solutions of differential inclusions should be absolutely continuous curves whose derivative are measurable selections of admissible velocities (see e.g. [12, Chapter 10]).

Definition 2.26 (Continuity inclusions in $(\mathscr{P}_p(\mathbb{R}^d), W_p)$)**.** *Let $V : [0,T] \times \mathscr{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d,\mathbb{R}^d)$ be a set-valued map. We say that a curve $\mu(\cdot) \in \mathrm{AC}([0,T],\mathscr{P}_p(\mathbb{R}^d))$ is a solution of the* continuity inclusion

$$\partial_t\mu(t) \in -\mathrm{div}_x\Big(V(t,\mu(t))\mu(t)\Big), \tag{2.25}$$

if there exists an $\mathscr{L}^1$-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 $(\mu(\cdot), v(\cdot))$ *solves*

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

in the sense of distributions.

3 Well-posedness in the Cauchy-Lipschitz setting

In this section, we derive general well-posedness results for the set-valued Cauchy problem

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

posed in the Wasserstein space $(\mathscr{P}_p(\mathbb{R}^d), W_p)$. We start by establishing in Section 3.1 a general version of the renowned Filippov estimates (see e.g. [85, Section 2.4]) for this class of problems, which combine an existence result along with a powerful a priori estimate on trajectories. We then proceed by showing in Section 3.2 that the solution set of (3.1) is compact for the topology of uniform convergence when the sets of admissible velocities are convex, and then conclude in Section 3.3 by proving that when this latter property does not hold, the underlying solution set remains dense in that of the relaxed dynamics whose right-hand side has been convexified.

Throughout the remainder of this article, we will assume that $T > 0$ and $p \in [1, +\infty)$ are given real numbers, and we shall systematically consider that the space $C^0(\mathbb{R}^d, \mathbb{R}^d)$ is endowed with the topology induced by the metric of uniform convergence on compact sets $\mathsf{d}_{\mathrm{cc}}(\cdot,\cdot)$ introduced in Definition 2.12.

Hypotheses (CI).

(i) *For any $\mu \in \mathscr{P}_p(\mathbb{R}^d)$, the set-valued map $t \in [0,T] \rightrightarrows V(t,\mu) \subset C^0(\mathbb{R}^d,\mathbb{R}^d)$ is $\mathscr{L}^1$-measurable and has closed nonempty images.*

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

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

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

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

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

$$\mathsf{d}_{\sup}(v,w) \leq L(t) W_p(\mu,\nu).$$

Remark 3.1 (Comparison with the main assumptions of [17])**.** *It is worth noting that our working assumptions are essentially the same as those of [17], except for the fact that the Lipschitz constants appearing in hypothesis* **(CI)**-*(iii), (iv) are required to be global. This turns out to be necessary in order to establish stability estimates for curves of measures whose supports may be unbounded. That being said, the study of curves of compactly supported measures generated by locally Lipschitz set-valued maps 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 allowed to vary from one compact sets of $\mathscr{P}_p(\mathbb{R}^d)$ to the other.*

Remark 3.2 (Example of set-valued mapping satisfying our assumptions)**.** *A relevant example of set-valued map $V : [0,T] \times \mathscr{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d,\mathbb{R}^d)$ satisfying hypotheses* **(CI)** *is given by the set of admissible velocities of a controlled system, which is of the form*

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

Therein, $(U, d_U(\cdot,\cdot))$ is a compact metric space representing the collection of control signals, while $v : [0,T] \times \mathscr{P}_p(\mathbb{R}^d) \times U \times \mathbb{R}^d \to \mathbb{R}^d$ is a vector field that is Carathéodory in $(t,u) \in [0,T] \times U$ as well as globally Lipschitz in $(\mu,x) \in \mathscr{P}_p(\mathbb{R}^d) \times \mathbb{R}^d$ with constants that are Lebesgue integrable functions of time. In the more general case in which $v : [0,T] \times \mathscr{P}_p(\mathbb{R}^d) \times U \times \mathbb{R}^d \to \mathbb{R}^d$ is simply Carathéodory with local uniform continuity constants in $(\mu,u,x) \in \mathscr{P}_p(\mathbb{R}^d) \times U \times \mathbb{R}^d$ given by Lebesgue integrable functions of time, this control theoretic model also fits in hypotheses **(P)** *of Section 4.*

3.1 Existence of solutions and Filippov estimates

In this section, we prove a general version of the well-posedness result of Filippov for solutions of the set-valued Cauchy problem (3.1). We stress that, while the underlying strategy is greatly inspired by the methodology we previously developed in [17], the corresponding results are much more general and the details of the proof fairly different from a technical standpoint.

Theorem 3.3 (Filippov estimates)**.** *Let $V : [0,T] \times \mathscr{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d, \mathbb{R}^d)$ be a set-valued map satisfying hypotheses* **(CI)** *and $\nu(\cdot) \in \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d))$ be a solution of the continuity equation*

$$\partial_t \nu(t) + \mathrm{div}_x(w(t)\nu(t)) = 0$$

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

$$|w(t,y)| \leq m(t)(1+|y|), \tag{3.2}$$

for $\mathscr{L}^1$-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*

$$\eta_R(t) := \mathrm{dist}_{L^\infty(B(0,R),\mathbb{R}^d;\,\nu(t))}\Big(w(t)\,;V(t,\nu(t))\Big) \tag{3.3}$$

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

Then for every initial datum $\mu^0 \in \mathscr{P}_p(\mathbb{R}^d)$ and each $R > 0$, there exists a trajectory-selection pair $(\mu(\cdot), v(\cdot)) \in \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d)) \times \mathscr{L}([0,T], C^0(\mathbb{R}^d, \mathbb{R}^d))$ solution of the Cauchy problem (3.1) which satisfies the distance estimate

$$W_p(\mu(t), \nu(t)) \leq \mathcal{D}_p(t), \tag{3.4}$$

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

$$\mathcal{D}_p(t) := C_p\Big(W_p(\mu^0, \nu(0)) + \int_0^t \eta_R(s)\mathrm{d}s + \mathcal{E}_\nu(t,R)\Big) \exp\Big(C'_p\, \|l(\cdot)\|^p_{L^1([0,t])} + \chi_p(t)\Big). \tag{3.5}$$

Therein, the constants $C_p, C'_p > 0$ are as in (2.19), the error term $\mathcal{E}_\nu(t,R)$ is given by

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

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 as

$$\chi_p(t) := C_p\, \|L(\cdot)\|_{L^1([0,t])} \exp\Big(C'_p\, \|l(\cdot)\|^p_{L^1([0,t])}\Big). \tag{3.6}$$

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

$$\|v(t) - w(t)\|_{L^\infty(B(0,R),\mathbb{R}^d;\,\nu(t))} \leq \eta_R(t) + L(t)\mathcal{D}_p(t) \tag{3.7}$$

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

Before moving on to the proof of Theorem 3.3, we need to state a technical lemma dealing with chained integral estimates, whose proof is the matter of a straightforward induction argument.

Lemma 3.4 (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 $(f_n(\cdot)) \subset C^0([0,T], \mathbb{R}_+)$ such that $\|f_0(\cdot)\|_{C^0([0,T],\mathbb{R}_+)} \leq f^0$ and*

$$f_{n+1}(t) \leq \alpha\Big(1 + \int_0^t m(s) f_n(s)\mathrm{d}s\Big),$$

for all times $t \in [0,T]$ and each $n \geq 0$, complies with the following inequality

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

The proof of Theorem 3.3 is based on a constructive scheme in which one builds a sequence of trajectory-selection pairs, 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 to build 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 (3.1) which satisfies the estimates (3.4) and (3.7).

Proof of Theorem 3.3. Our goal in what follows is to build a sequence of trajectory-selection pairs $(\mu_n(\cdot), v_n(\cdot)) \subset \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d)) \times \mathscr{L}([0,T], C^0(\mathbb{R}^d, \mathbb{R}^d))$ solving the Cauchy problems

$$\begin{cases} \partial_t \mu_n(t) + \mathrm{div}_x(v_n(t)\mu_n(t)) = 0, \\ \mu_n(0) = \mu^0, \end{cases} \tag{3.8}$$

and satisfying the conditions

$$v_{n+1}(t) \in V(t, \mu_n(t)) \qquad \text{and} \qquad \mathsf{d}_{\sup}(v_n(t), v_{n+1}(t)) \leq L(t) W_p(\mu_{n-1}(t), \mu_n(t)), \tag{3.9}$$

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

$$\mathcal{M}_p(\mu_n(t)) \leq \mathcal{C} \qquad \text{and} \qquad W_p(\mu_n(\tau), \mu_n(t)) \leq c_p \int_\tau^t m(s) \mathrm{d}s \tag{3.10}$$

for all times $0 \leq \tau \leq t \leq T$ and each $n \geq 1$. Therein, the constants $\mathcal{C}, c_p > 0$ will 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 2.15, 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 $\mathscr{L}^1$-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 $\mathscr{L}^1$-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 $\mathscr{L}^1$-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 $\mathsf{d}_{\mathrm{cc}}(\cdot,\cdot)$. Thus, recalling the definition (3.3) of the mismatch function $\eta_R(\cdot)$ and applying Lemma 2.10-(b) to the correspondence

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

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

$$\|v_1(t) - w(t)\|_{L^\infty(B(0,R), \mathbb{R}^d;\, \nu(t))} = \eta_R(t) \tag{3.11}$$

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

Remark now that as a by-product of hypotheses **(CI)**-(ii) and (iii) along with Lemma 2.14, the velocity field $v_1 : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ is Carathéodory and such that

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

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and all $x \in \mathbb{R}^d$. In particular, it satisfies hypotheses **(CE)**. Hence by Theorem 2.18, there exists a unique curve $\mu_1(\cdot) \in \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d))$ solution of the Cauchy problem

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

Owing to the momentum bound of Remark 2.23, the curve $\mu_1(\cdot)$ complies with the estimate

$$\mathcal{M}_p(\mu_1(t)) \leq C_p\Big(\mathcal{M}_p(\mu^0) + \int_0^t m(s)\big(1+\mathcal{M}_p(\nu(s))\big)\mathrm{d}s\Big)\exp\left(C_p'\,\|m(\cdot)\|_1^p\right), \tag{3.12}$$

for all times $t \in [0,T]$. Consequently, by Lemma 3.4 combined and (2.18) of Proposition 2.22 applied to the curve $\nu(\cdot) \in \mathrm{AC}([0,T],\mathscr{P}_p(\mathbb{R}^d))$ – which is valid owing to (3.2) –, 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

$$\max\Big\{\mathcal{M}_p(\nu(t))\,,\,\mathcal{M}_p(\mu_1(t))\Big\} \leq \mathcal{C} \tag{3.13}$$

for all times $t \in [0,T]$. In particular, the velocity field $v_1(\cdot,\cdot)$ satisfies the uniform sublinearity estimate

$$|v_1(t,x)| \leq (1+\mathcal{C})m(t)(1+|x|), \tag{3.14}$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and all $x \in \mathbb{R}^d$, which combined with (2.13) of Theorem 2.18 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

$$W_p(\mu_1(\tau),\mu_1(t)) \leq c_p\int_\tau^t m(s)\mathrm{d}s$$

for all times $0 \leq \tau \leq t \leq T$. Moreover, by applying the approximate stability inequality (2.23) of Proposition 2.24 while taking into account the sublinearity estimate (3.14), it further holds that

$$W_p(\mu_1(t),\nu(t)) \leq C_p\Big(W_p(\mu^0,\nu(0)) + \int_0^t \eta_R(s)\mathrm{d}s + \mathcal{E}_\nu(t,R)\Big)\exp\left(C_p'\,\|l(\cdot)\|_1^p\right) \tag{3.15}$$

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

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

with $\mathcal{C}_T := \max\{1,(1+\mathcal{C})\|m(\cdot)\|_1\}\exp\big((1+\mathcal{C})\|m(\cdot)\|_1\big)$. This together with the momentum estimate of (3.12) concludes the initialisation step of our induction argument.

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

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

for $\mathscr{L}^1$-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 $\mathscr{L}^1$-measurable map $t \in [0,T] \mapsto v_2(t) \in V(t,\mu_1(t))$ which satisfies

$$\mathsf{d}_{\sup}(v_1(t),v_2(t)) \leq L(t)W_p(\mu_1(t),\nu(t)),$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$. From hypotheses **(CI)**-(ii) and (iii) combined with (3.13) and Lemma 2.14, we may then infer that the velocity field $v_2 : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ is Carathéodory, and such that

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

for $\mathscr{L}^1$-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 $\mu_2(\cdot) \in \mathrm{AC}([0,T],\mathscr{P}_p(\mathbb{R}^d))$ of the corresponding Cauchy problem

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

By the refined momentum bound of Remark 2.23 and the estimate (2.22) of Proposition 2.24, it can again be checked that

$$\begin{cases} \mathcal{M}_p(\mu_2(t)) \leq C_p\Big(\mathcal{M}_p(\mu^0) + \int_0^t m(s)\big(1+\mathcal{M}_p(\mu_1(s))\big)\mathrm{d}s\Big)\exp\Big(C'_p\,\|m(\cdot)\|_1^p\Big), \\ W_p(\mu_1(t),\mu_2(t)) \leq C_p\Big(\int_0^t \mathsf{d}_{\sup}(v_1(s),v_2(s))\mathrm{d}s\Big)\exp\Big(C'_p\,\|l(\cdot)\|_1^p\Big). \end{cases}$$

As a consequence of Lemma 3.4, this implies in particular that

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

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 iteratively build a sequence $(\mu_n(\cdot), v_n(\cdot)) \subset \mathrm{AC}([0,T],\mathscr{P}_p(\mathbb{R}^d)) \times \mathscr{L}([0,T], C^0(\mathbb{R}^d,\mathbb{R}^d))$ of trajectory-selection pairs solutions of the Cauchy problems (3.8), and which satisfy the conditions of (3.9) as well as the bounds of (3.10). In addition, the elements of the sequence comply for each $n \geq 1$ with the uniform sublinearity estimates

$$|v_n(t,x)| \leq (1+\mathcal{C})m(t)(1+|x|), \tag{3.16}$$

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

$$W_p(\mu_n(t),\mu_{n+1}(t)) \leq C_p\Big(\int_0^t \mathsf{d}_{\sup}(v_n(s),v_{n+1}(s))\mathrm{d}s\Big)\exp\Big(C'_p\,\|l(\cdot)\|_1^p\Big), \tag{3.17}$$

for all times $t \in [0,T]$. In particular by (3.16) and Theorem 2.18, 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

$$W_p(\mu_n(\tau),\mu_n(t)) \leq c_p\int_\tau^t m(s)\mathrm{d}s,$$

for all times $0 \leq \tau \leq t \leq T$ and each $n \geq 1$.

Step 3 – Convergence of the sequence of trajectory-selection pairs. In what follows, we prove that the sequence of pairs $(\mu_n(\cdot), v_n(\cdot))$ built in Step 1 and Step 2 is Cauchy in a suitable sense. First, observe that by merging the distance estimates of (3.9), (3.15) and (3.17), one has that

$$\begin{aligned} &W_p(\mu_n(t),\mu_{n+1}(t)) \\ &\leq C_p\Big(\int_0^t \mathsf{d}_{\sup}(v_n(s_n),v_{n+1}(s_n))\mathrm{d}s_n\Big)\exp\Big(C'_p\,\|l(\cdot)\|^p_{L^1([0,t])}\Big) \\ &\leq C_p\Big(\int_0^t L(s_n)W_p(\mu_{n-1}(s_n),\mu_n(s_n))\mathrm{d}s_n\Big)\exp\Big(C'_p\,\|l(\cdot)\|^p_{L^1([0,t])}\Big) \\ &\leq C_p^2\Big(\int_0^t L(s_n)\int_0^{s_n}\mathsf{d}_{\sup}(v_{n-1}(s_{n-1}),v_n(s_{n-1}))\mathrm{d}s_{n-1}\mathrm{d}s_n\Big)\exp\Big(2C'_p\,\|l(\cdot)\|^p_{L^1([0,t])}\Big) \\ &\vdots \\ &\leq C_p^n\Big(\int_0^t L(s_n)\int_0^{s_n}L(s_{n-1})\ldots\int_0^{s_2}L(s_1)W_p(\mu_1(s_1),\nu(s_1))\mathrm{d}s_1\ldots\mathrm{d}s_{n-1}\mathrm{d}s_n\Big)\exp\Big(n\,C'_p\,\|l(\cdot)\|^p_{L^1([0,t])}\Big) \\ &\leq \frac{C_p^{n+1}}{n!}\Big(\int_0^t L(s)\mathrm{d}s\Big)^n\Big(W_p(\mu^0,\nu(0)) + \int_0^t \eta_R(s)\mathrm{d}s + \mathcal{E}_\nu(t,R)\Big)\exp\Big((n+1)\,C'_p\,\|l(\cdot)\|^p_{L^1([0,t])}\Big), \\ &= \frac{\chi_p(t)^n}{n!}\,C_p\Big(W_p(\mu^0,\nu(0)) + \int_0^t \eta_R(s)\mathrm{d}s + \mathcal{E}_\nu(t,R)\Big)\exp\Big(C'_p\,\|l(\cdot)\|^p_{L^1([0,t])}\Big) \end{aligned} \tag{3.18}$$

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 (3.6). Whence, for any pair of integers $m, n \geq 1$, one can deduce from (3.18) that

$$\begin{aligned}
&\sup_{t\in[0,T]} W_p(\mu_n(t), \mu_{n+m}(t)) \\
&\leq \sup_{t\in[0,T]} \sum_{k=n}^{n+m-1} W_p(\mu_k(t), \mu_{k+1}(t)) \\
&\leq \left(\sum_{k=n}^{n+m-1} \frac{\|\chi_p(\cdot)\|_\infty^k}{k!} \right) C_p \left(W_p(\mu^0, \nu(0)) + \int_0^T \eta_R(s)\mathrm{d}s + \mathcal{E}_\nu(T,R) \right) \exp\left(C_p' \, \|l(\cdot)\|_1^p \right) \underset{m,n\to+\infty}{\longrightarrow} 0,
\end{aligned} \tag{3.19}$$

which means that $(\mu_n(\cdot)) \subset \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d))$ is a Cauchy sequence in $C^0([0,T], \mathscr{P}_p(\mathbb{R}^d))$. Noting that the latter is a complete metric space as a consequence e.g. of [67, Chapter 7 – Theorem 12], there exists some $\mu(\cdot) \in C^0([0,T], \mathscr{P}_p(\mathbb{R}^d))$ such that

$$\sup_{t\in[0,T]} W_p(\mu_n(t), \mu(t)) \underset{n\to+\infty}{\longrightarrow} 0. \tag{3.20}$$

Besides, as a consequence of the equi-absolute continuity estimate established in (3.10) for the sequence $(\mu_n(\cdot)) \subset \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d))$, the limit curve $\mu(\cdot)$ is absolutely continuous as well, with

$$W_p(\mu(\tau), \mu(t)) \leq c_p \int_\tau^t m(s)\mathrm{d}s$$

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

Concerning the sequence of Carathéodory velocity fields $(v_n(\cdot)) \subset \mathscr{L}([0,T], C^0(\mathbb{R}^d, \mathbb{R}^d))$, it directly follows from (3.9) together with the estimates in (3.19) that

$$\begin{aligned}
\int_0^T \mathsf{d}_{\sup}(v_n(t), v_{n+m}(t))\mathrm{d}t &\leq \sum_{k=n}^{m+n-1} \int_0^T \mathsf{d}_{\sup}(v_k(t), v_{k+1}(t))\mathrm{d}t \\
&\leq \|L(\cdot)\|_1 \sup_{t\in[0,T]} \sum_{k=n}^{m+n-1} W_p(\mu_{k-1}(t), \mu_k(t)) \underset{m,n\to+\infty}{\longrightarrow} 0,
\end{aligned} \tag{3.21}$$

for each $m, n \geq 1$. Hence, by Lemma 2.16, there exists a map $v(\cdot) \in \mathscr{L}([0,T], C^0(\mathbb{R}^d, \mathbb{R}^d))$ such that

$$\int_0^T \mathsf{d}_{\sup}(v_n(t), v(t))\mathrm{d}t \underset{n\to+\infty}{\longrightarrow} 0. \tag{3.22}$$

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

$$\mathsf{d}_{\sup}(v_{n_k}(t), v(t)) \underset{n\to+\infty}{\longrightarrow} 0$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$, which combined with (3.16) directly implies that the limit velocity field $v : [0,T] \times \mathbb{R}^d \mapsto \mathbb{R}^d$ satisfies the sublinearity and regularity estimates

$$|v(t,x)| \leq (1+\mathcal{C})m(t)(1+|x|) \qquad \text{and} \qquad \mathrm{Lip}(v(t)\,;\mathbb{R}^d) \leq l(t), \tag{3.23}$$

for $\mathscr{L}^1$-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 (2.23) in Proposition 2.24 as well as (3.23), it can be straightforwardly deduced from the convergence results of (3.20) and (3.22) that the trajectory-selection pair $(\mu(\cdot), v(\cdot))$ is a distributional solution of the Cauchy problem

$$\begin{cases} \partial_t \mu(t) + \mathrm{div}_x(v(t)\mu(t)) = 0, \\ \mu(0) = \mu^0. \end{cases} \tag{3.24}$$

In addition, it follows from (3.23) and Theorem 2.18 that the curve $\mu(\cdot) \in \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d))$ is in fact the unique solution of (3.24).

The next step in our argument consists in showing that $(\mu(\cdot), v(\cdot))$ is a solution of the continuity inclusion (3.1). By (3.9), the sequence of trajectory-selection pairs is such that

$$(\mu_n(t), v_{n+1}(t)) \in \mathrm{Graph}(V(t)), \tag{3.25}$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and each $n \geq 1$. By (3.20) and (3.22), it further holds that

$$W_p(\mu_n(t), \mu(t)) \underset{n\to+\infty}{\longrightarrow} 0, \tag{3.26}$$

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

$$\mathsf{d}_{\sup}(v_n(t), v(t)) \underset{n\to+\infty}{\longrightarrow} 0, \tag{3.27}$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$, along a subsequence that we do not relabel. Whence, upon combining (3.25) with the pointwise convergence results (3.26) and (3.27) and hypothesis **(CI)**-(iv), we obtain

$$(\mu(t), v(t)) \in \mathrm{Graph}(V(t)),$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$. In light of Definition 2.26, this latter fact together with (3.24) implies that the limit curve $\mu(\cdot) \in \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d))$ is a solution of the continuity inclusion (3.1), driven by the $\mathscr{L}^1$-measurable 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 (3.4) and (3.7). To this end, note that the chain of pointwise inequalities of (3.18) implies

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

for all times $t \in [0,T]$ and each $n \geq 1$, with $\mathcal{D}_p(\cdot)$ being defined as in (3.5). By letting $n \to +\infty$ in the previous expression, we directly recover (3.4). Concerning the velocity estimates, it holds as a consequence of (3.9), (3.11) and (3.15) that

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

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and each $n \geq 1$. By taking the limit as $n \to +\infty$ along a suitable subsequence in the previous expression, we recover (3.7), which concludes the proof of Theorem 3.3. □

In the following corollary, we state a global version of the Filippov estimate which can be obtained by a simple adaptation of the proof of Theorem 3.3.

Corollary 3.5 (Global version of Filippov's estimates)**.** *Suppose that the assumptions of Theorem 3.3 hold, and in addition that the* global mismatch function*, defined by*

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

for $\mathscr{L}^1$-almost every $t \in [0,T]$, is Lebesgue integrable. Then for every $\mu^0 \in \mathscr{P}_p(\mathbb{R}^d)$, there exists a trajectory-selection pair $(\mu(\cdot), v(\cdot)) \in \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d)) \times \mathscr{L}([0,T], C^0(\mathbb{R}^d, \mathbb{R}^d))$ solution of the Cauchy problem (3.1), which satisfies the distance estimate

$$W_p(\mu(t), \nu(t)) \leq C_p \Big(W_p(\mu^0, \nu(0)) + \int_0^t \eta(s) \mathrm{d}s \Big) \exp\Big(C'_p \, \|l(\cdot)\|^p_{L^1([0,t])} + \chi_p(t) \Big),$$

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

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

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

Proof. One can simply repeat the proof strategy of Theorem 3.3 with $R = +\infty$ and $\mathcal{E}(t,R) = 0$. □

3.2 Compactness of the solution set for convex velocities

In this section, we study the compactness properties of the solution set $\mathcal{S}_{[0,T]}(\mu^0)$ associated with the Cauchy problem (3.1), defined by

$$\mathcal{S}_{[0,T]}(\mu^0) := \Big\{ \mu(\cdot) \in \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d)) \ \text{ s.t. } \mu(\cdot) \text{ is a solution of (3.1) with } \mu(0) = \mu^0 \Big\} \tag{3.28}$$

for each $\mu^0 \in \mathscr{P}_p(\mathbb{R}^d)$. More specifically, we show in the following theorem that $\mathcal{S}_{[0,T]}(\mu^0)$ is compact for the topology of uniform convergence as soon as the set-valued map $V : [0,T] \times \mathscr{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d, \mathbb{R}^d)$ has convex images.

Theorem 3.6 (Compactness of the solution set)**.** *Let $\mu^0 \in \mathscr{P}_p(\mathbb{R}^d)$ and $V : [0,T] \times \mathscr{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d, \mathbb{R}^d)$ be a set-valued map satisfying hypotheses **(CI)** and whose images are convex. Then, the solution set $\mathcal{S}_{[0,T]}(\mu^0) \subset C^0([0,T], \mathscr{P}_p(\mathbb{R}^d))$ associated with the Cauchy problem (3.1) is compact for the topology of uniform convergence.*

Proof. Let $(\mu_n(\cdot), v_n(\cdot)) \subset \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d)) \times \mathscr{L}([0,T], C^0(\mathbb{R}^d, \mathbb{R}^d))$ be a sequence of trajectory selection pairs for the Cauchy problem (3.1). Observe first that by the momentum bound of Remark 2.23, there exists $\mathcal{C} > 0$ depending only on the magnitudes of $p, \mathcal{M}(\mu^0)$ and $\|m(\cdot)\|_1$ such that

$$\sup_{n \geq 1} \mathcal{M}(\mu_n(t)) \leq \mathcal{C},$$

for all times $t \in [0,T]$. Thus, by hypothesis **(CI)**-(ii) together with the equi-integrability estimate of Proposition 2.22 and the characterisation of relatively compact subsets of $\mathscr{P}_p(\mathbb{R}^d)$ given in Proposition 2.3, there exists a compact set $\mathscr{K} \subset \mathscr{P}_p(\mathbb{R}^d)$ depending only on μ^0 and $\|m(\cdot)\|_1$, such that

$$\mu_n(t) \in \mathscr{K}$$

for all times $t \in [0,T]$ and each $n \geq 1$. Besides, recalling that the velocity fields $v_n : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ are Carathéodory and uniformly equi-sublinear for each $n \geq 1$, one has by Theorem 2.18 that

$$W_p(\mu_n(\tau), \mu_n(t)) \leq c_p \int_\tau^t m(s) \mathrm{d}s,$$

for all times $0 \leq \tau \leq t \leq T$ and each $n \geq 1$, where $c_p > 0$ only depends on the magnitudes of $p, \mathcal{M}_p(\mu^0)$ and $\|m(\cdot)\|_1$. Whence, by the Ascoli-Arzelà theorem for complete separable metric spaces (see e.g. [67, Chapter 7 – Theorem 18]), there exists a curve $\mu(\cdot) \in \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d))$ such that

$$\sup_{t \in [0,T]} W_p(\mu_n(t), \mu(t)) \underset{n \to +\infty}{\longrightarrow} 0, \tag{3.29}$$

along a subsequence that we do not relabel.

Observe now that for each $m \geq 1$, the restrictions to $B(0,m)$ of the sequence of maps $(v_n(\cdot)) \subset \mathscr{L}([0,T], C^0(\mathbb{R}^d,\mathbb{R}^d))$ is uniformly bounded and equi-integrable in $L^1([0,T], C^0(B(0,m),\mathbb{R}^d))$ as a consequence of the sublinearity estimate of hypothesis **(CI)**-(ii) and Proposition 2.22. This, together with hypothesis **(CI)**-(iii) and the Ascoli-Arzelà theorem, yields the existence of family of compact sets $(\mathcal{K}_t^m)_{t\in[0,T]} \subset C^0(B(0,m),\mathbb{R}^d)$ such that

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

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

$$v_n(\cdot) \underset{n\to+\infty}{\rightharpoonup} v(\cdot) \tag{3.30}$$

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

$$\int_0^T \langle \boldsymbol{\nu}(t), v(t) - v_n(t)\rangle_{C^0(B(0,R),\mathbb{R}^d)} \mathrm{d}t \underset{n\to+\infty}{\longrightarrow} 0, \tag{3.31}$$

whenever $\boldsymbol{\nu}(\cdot) \in L^\infty([0,T], \mathscr{M}(B(0,R),\mathbb{R}^d))$, again as a consequence of Theorem 2.1. Furthermore, by combining Mazur's lemma (see e.g. [21, Corollary 3.8]) with (3.29), one can show that

$$|v(t,x)| \leq m(t)\Big(1 + |x| + \mathcal{M}_p(\mu(t))\Big) \qquad \text{and} \qquad \mathrm{Lip}(v(t)\,;\mathbb{R}^d) \leq l(t), \tag{3.32}$$

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

Our goal is now to prove that the limit pair $(\mu(\cdot), v(\cdot)) \in \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d)) \times \mathscr{L}([0,T], C^0(\mathbb{R}^d,\mathbb{R}^d))$ is a solution of (3.1). Upon combining the convergence result of (3.29) and (3.32) together with the Wasserstein estimate (2.4) of Proposition 2.4, it is clear that

$$\int_0^T \int_{\mathbb{R}^d} \partial_t\phi(t,x)\mathrm{d}\mu_n(t)(x)\mathrm{d}t \underset{n\to+\infty}{\longrightarrow} \int_0^T \int_{\mathbb{R}^d} \partial_t\phi(t,x)\mathrm{d}\mu(t)(x)\mathrm{d}t, \tag{3.33}$$

and

$$\int_0^T \int_{\mathbb{R}^d} \langle \nabla_x\phi(t,x), v_n(t,x)\rangle \mathrm{d}(\mu(t) - \mu_n(t))(x)\mathrm{d}t \underset{n\to+\infty}{\longrightarrow} 0, \tag{3.34}$$

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

$$\phi(t,x) := \zeta(t)\psi(x), \tag{3.35}$$

for all $(t,x) \in [0,T]\times\mathbb{R}^d$, with $(\zeta,\psi) \in C_c^\infty((0,T),\mathbb{R})\times \in C_c^\infty(\mathbb{R}^d,\mathbb{R})$, and then setting $\boldsymbol{\nu}(t) := \zeta(t)\nabla\psi\cdot\mu(t) \in \mathscr{M}(\mathrm{supp}(\psi),\mathbb{R}^d)$ for all times $t \in [0,T]$ in (3.31), it also holds that

$$\int_0^T \zeta(t) \int_{\mathbb{R}^d} \langle \nabla\psi(x), v(t,x) - v_n(t,x)\rangle \mathrm{d}\mu(t)(x)\mathrm{d}t \underset{n\to+\infty}{\longrightarrow} 0, \tag{3.36}$$

along a subsequence that depends on $\mathrm{supp}(\psi) \subset \mathbb{R}^d$. Therefore, by merging the convergence results of (3.33), (3.34) and (3.36) while recalling that the linear span of test functions of the form (3.35) is dense in $C_c^\infty((0,T)\times\mathbb{R}^d,\mathbb{R})$ (see e.g. [9, Chapter 8]), we finally recover that

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

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 (3.29) equivalently means that the pair $(\mu(\cdot),v(\cdot))$ is a solution of the Cauchy problem

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

To conclude the proof of Theorem 3.6, there now remains to show that $v(t)\in V(t,\mu(t))$ for $\mathscr{L}^1$-almost every $t\in[0,T]$. To this end, observe first 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

$$t\in[0,T]\mapsto \tilde{v}_n(t)\in V(t,\mu(t)), \tag{3.37}$$

which satisfies

$$\mathsf{d}_{\sup}(\tilde{v}_n(t),v_n(t))\leq L(t)W_p(\mu_n(t),\mu(t)) \underset{n\to+\infty}{\longrightarrow} 0$$

for $\mathscr{L}^1$-almost every $t\in[0,T]$, where we also used (3.29). Whence, by a direct application of Lebesgue's dominated convergence theorem, it necessarily follows that

$$\int_0^T \|\tilde{v}_n(t)-v_n(t)\|_{C^0(B(0,R),\mathbb{R}^d)}\,\mathrm{d}t \underset{n\to+\infty}{\longrightarrow} 0$$

for each $R>0$, which combined with (3.30) implies in particular that

$$\tilde{v}_n(\cdot) \underset{n\to+\infty}{\rightharpoonup} v(\cdot) \tag{3.38}$$

weakly in $L^1([0,T],C^0(B(0,R),\mathbb{R}^d))$, along adequate subsequences which depend on $R>0$. At this point, one can easily check that under hypotheses **(CI)**, the sets

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

are closed for the strong topology of $L^1([0,T],C^0(B(0,R),\mathbb{R}^d))$. The latter are also convex under our working assumptions, and thus weakly closed in $L^1([0,T],C^0(B(0,R),\mathbb{R}^d))$ by Mazur's lemma (see e.g. [21, Theorem 3.7]). Thus, from (3.37) and (3.38), 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 $\mathscr{V}_R$ for each $R>0$, which equivalently means that $v(t)\in V(t,\mu(t))$ for $\mathscr{L}^1$-almost every $t\in[0,T]$ and concludes the proof of Theorem 3.6. □

3.3 Relaxation property for non-convex velocities

In this section, we investigate the topological properties of the solution set $\mathcal{S}_{[0,T]}(\mu^0)$ defined in (3.28) when the admissible velocities are not convex. While the compactness result of Theorem 3.6 does not hold anymore in this context, we show in what follows that one can still characterise the closure of the underlying solution set by convexifying the right-hand side of the dynamics.

Theorem 3.7 (Relaxation theorem for continuity inclusions)**.** *Let $\mu^0\in\mathscr{P}_p(\mathbb{R}^d)$ and $V:[0,T]\times\mathscr{P}_p(\mathbb{R}^d)\rightrightarrows C^0(\mathbb{R}^d,\mathbb{R}^d)$ be a set-valued map satisfying hypotheses* **(CI)**. *Then, for any solution $\mu(\cdot)\in\mathrm{AC}([0,T],\mathscr{P}_p(\mathbb{R}^d))$ of the* relaxed Cauchy problem

$$\begin{cases}\partial_t\mu(t)\in-\mathrm{div}_x\Big(\overline{\mathrm{co}}\,V(t,\mu(t))\mu(t)\Big),\\ \mu(0)=\mu^0,\end{cases} \tag{3.39}$$

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

$$\begin{cases}\partial_t\mu_\delta(t)\in-\mathrm{div}_x\Big(V(t,\mu_\delta(t))\mu_\delta(t)\Big),\\ \mu(0)=\mu^0,\end{cases} \tag{3.40}$$

which satisfies

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

In particular, the solution set of the Cauchy problem (3.40) *is dense in that of the relaxed Cauchy problem* (3.39) *for the topology of uniform convergence over* $C^0([0,T],\mathscr{P}_p(\mathbb{R}^d))$.

The proof of Theorem 3.7 will be split into three steps. In Step 1, we start by choosing an adequate subdivision of the interval $[0,T]$ that we shall use to build an intermediate measure curve. In Step 2, we subsequently prove that the latter construction based on Aumann's theorem produces a curve which is close to $\mu(\cdot)$ in $C^0([0,T],\mathscr{P}_p(\mathbb{R}^d))$. We then conclude in Step 3 by applying the generalised Filippov estimates of Theorem 3.3, which allows us to recover a solution of the Cauchy problem (3.40) that remains sufficiently close to our initial measure curve.

Proof of Theorem 3.7. We start by observing that, if $V : [0,T]\times\mathscr{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d,\mathbb{R}^d)$ is a set-valued map satisfying hypotheses **(CI)** and $\mu(\cdot)\in \mathrm{AC}([0,T],\mathscr{P}_p(\mathbb{R}^d))$ is a solution of either (3.39) or (3.40), then by Proposition 2.22 every measurable selection $v(\cdot)\in\mathscr{L}([0,T],C^0(\mathbb{R}^d,\mathbb{R}^d))$ in either $t\in[0,T]\mapsto \overline{\mathrm{co}}\,V(t,\mu(t))$ or $t\in[0,T]\mapsto V(t,\mu(t))$ necessarily satisfies

$$|v(t,x)| \leq (1+\mathcal{C})m(t)(1+|x|), \tag{3.41}$$

for $\mathscr{L}^1$-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 $\boldsymbol{\eta}\in\mathscr{P}(\mathbb{R}^d\times\Sigma_T)$ the superposition measure associated with such a velocity field via Definition 2.19, it directly follows from Grönwall's lemma that

$$\|\sigma(\cdot)\|_{C^0([0,T],\mathbb{R}^d)} \leq \mathcal{C}_T(1+|x|), \tag{3.42}$$

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

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

$$\left(\int_{\{x \text{ s.t. } |x|\geq R_\delta/\mathcal{C}_T-1\}}(1+|x|)^p\mathrm{d}\mu^0(x)\right)^{1/p} \leq \frac{\delta}{2(1+\mathcal{C})(1+\mathcal{C}_T)(1+\|m(\cdot)\|_1)}. \tag{3.43}$$

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

$$\int_{t_i}^{t_{i+1}} m(s)\mathrm{d}s \leq \frac{\delta}{2(1+\mathcal{C})(1+R_\delta)} \tag{3.44}$$

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

$$\Big(\overline{\mathrm{co}}\,V(t,\mu(t))\Big)_{|B(0,R_\delta)} = \overline{\mathrm{co}}\Big(V(t,\mu(t))_{|B(0,R_\delta)}\Big), \tag{3.45}$$

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 $t\in[0,T]\mapsto v(t)\in\overline{\mathrm{co}}\,V(t,\mu(t))$ the velocity selection associated with the solution $\mu(\cdot)\in\mathrm{AC}([0,T],\mathscr{P}_p(\mathbb{R}^d))$ of (3.39), 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

$$\left\|\int_{t_i}^{t_{i+1}} v_{|B(0,R_\delta)}(s)\mathrm{d}s - \int_{t_i}^{t_{i+1}} v_i^\delta(s)\mathrm{d}s\right\|_{C^0(B(0,R_\delta),\mathbb{R}^d)} \leq \frac{\delta}{N} \tag{3.46}$$

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 the maps $(t,w) \in [t_i, t_{i+1}] \times C^0(\mathbb{R}^d, \mathbb{R}^d) \mapsto \|w - v_i^\delta(t)\|_{C^0(B(0,R_\delta),\mathbb{R}^d)}$ are Carathéodory, whereas

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

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 $\mathscr{L}^1$-almost every $t \in [0,T]$. Therefore, the velocity field $w : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ defined by

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

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

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

as a consequence of Theorem 2.18.

Step 2 – Estimation of the W_p-distance between $\mu(\cdot)$ and $\nu(\cdot)$. Let $\boldsymbol{\eta}_\mu, \boldsymbol{\eta}_\nu \in \mathscr{P}(\mathbb{R}^d \times \Sigma_T)$ be the superposition measures associated with $\mu(\cdot)$ and $\nu(\cdot)$ respectively, and $\hat{\boldsymbol{\eta}}_{\mu,\nu} \in \Gamma(\boldsymbol{\eta}_\mu, \boldsymbol{\eta}_\nu)$ be a transport plan given by Lemma 2.21, which satisfies

$$(\pi_{\mathbb{R}^d}, \pi_{\mathbb{R}^d})_\sharp \hat{\boldsymbol{\eta}}_{\mu,\nu} = (\mathrm{Id}, \mathrm{Id})_\sharp \mu^0 \qquad \text{and} \qquad \gamma(t) := (\mathsf{e}_t, \mathsf{e}_t)_\sharp \hat{\boldsymbol{\eta}}_{\mu,\nu} \in \Gamma_o(\mu(t), \nu(t))$$

for all times $t \in [0,T]$. Then, one has that

$$\begin{aligned} W_p(\mu(t), \nu(t)) &= \left(\int_{\mathbb{R}^{2d}} |x-y|^p \mathrm{d}\gamma(t)(x,y) \right)^{1/p} \\ &= \left(\int_{(\mathbb{R}^d \times \Sigma_T)^2} |\sigma_\mu(t) - \sigma_\nu(t)|^p \mathrm{d}\hat{\boldsymbol{\eta}}_{\mu,\nu}(x, \sigma_\mu, y, \sigma_\nu) \right)^{1/p} \\ &\leq \left(\int_{(\mathbb{R}^d \times \Sigma_T)^2} \left| \int_0^t \Big(v(s, \sigma_\mu(s)) - v(s, \sigma_\nu(s)) \Big) \mathrm{d}s \right|^p \mathrm{d}\hat{\boldsymbol{\eta}}_{\mu,\nu}(x, \sigma_\mu, y, \sigma_\nu) \right)^{1/p} \\ &\qquad + \left(\int_{\mathbb{R}^d \times \Sigma_T} \left| \int_0^t \Big(v(s, \sigma_\nu(s)) - w(s, \sigma_\nu(s)) \Big) \mathrm{d}s \right|^p \mathrm{d}\boldsymbol{\eta}_\nu(y, \sigma_\nu) \right)^{1/p}, \end{aligned} \tag{3.47}$$

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 2.19. By repeating the computations detailed in Appendix C and D, it can be shown that

$$\begin{aligned} &\left(\int_{(\mathbb{R}^d \times \Sigma_T)^2} \left| \int_0^t \Big(v(s, \sigma_\mu(s)) - v(s, \sigma_\nu(s)) \Big) \mathrm{d}s \right|^p \mathrm{d}\hat{\boldsymbol{\eta}}_{\mu,\nu}(x, \sigma_\mu, y, \sigma_\nu) \right)^{1/p} \\ &\qquad \leq \|l(\cdot)\|_{L^1([0,t])}^{(p-1)/p} \left(\int_0^t l(s) W_p^p(\mu(s), \nu(s)) \mathrm{d}s \right)^{1/p}, \end{aligned} \tag{3.48}$$

for all times $t \in [0,T]$. Concerning the second term in the right-hand side of (3.47), it can be bounded from above by the sum of two integrals as

$$\begin{aligned} &\left(\int_{\mathbb{R}^d \times \Sigma_T} \left| \int_0^t \Big(v(s, \sigma_\nu(s)) - w(s, \sigma_\nu(s)) \Big) \mathrm{d}s \right|^p \mathrm{d}\boldsymbol{\eta}_\nu(y, \sigma_\nu) \right)^{1/p} \\ &\leq \left(\int_{\{(y,\sigma_\nu) \text{ s.t. } \|\sigma_\nu(\cdot)\|_{C^0([0,T],\mathbb{R}^d)} \leq R_\delta\}} \left| \int_0^t \Big(v(s, \sigma_\nu(s)) - w(s, \sigma_\nu(s)) \Big) \mathrm{d}s \right|^p \mathrm{d}\boldsymbol{\eta}_\nu(y, \sigma_\nu) \right)^{1/p} \\ &\quad + \left(\int_{\{(y,\sigma_\nu) \text{ s.t. } \|\sigma_\nu(\cdot)\|_{C^0([0,T],\mathbb{R}^d)} > R_\delta\}} \left| \int_0^t \Big(v(s, \sigma_\nu(s)) - w(s, \sigma_\nu(s)) \Big) \mathrm{d}s \right|^p \mathrm{d}\boldsymbol{\eta}_\nu(y, \sigma_\nu) \right)^{1/p}. \end{aligned} \tag{3.49}$$

Recalling that $(\pi_{\mathbb{R}^d})_\sharp \boldsymbol{\eta}_\nu = \mu^0$ and invoking the sublinearity estimates of (3.41) and (3.42), the second term in the right-hand side of (3.49) can be estimated as

$$\begin{aligned}
&\Big(\int_{\{(y,\sigma_\nu)\text{ s.t. }\|\sigma_\nu(\cdot)\|_{C^0([0,T],\mathbb{R}^d)}>R_\delta\}} \Big|\int_0^t \Big(v(s,\sigma_\nu(s)) - w(s,\sigma_\nu(s))\Big)\mathrm{d}s\Big|^p \mathrm{d}\boldsymbol{\eta}_\nu(y,\sigma_\nu)\Big)^{1/p} \\
&\le 2(1+\mathcal{C})\,\|m(\cdot)\|_1 \Big(\int_{\{(y,\sigma_\nu)\text{ s.t. }\|\sigma_\nu(\cdot)\|_{C^0([0,T],\mathbb{R}^d)}>R_\delta\}} \Big(1+\|\sigma_\nu(\cdot)\|_{C^0([0,T],\mathbb{R}^d)}\Big)^p \mathrm{d}\boldsymbol{\eta}_\nu(y,\sigma_\nu)\Big)^{1/p} \qquad (3.50)\\
&\le 2(1+\mathcal{C})(1+\mathcal{C}_T)\,\|m(\cdot)\|_1 \Big(\int_{\{x\text{ s.t. }|x|>R_\delta/\mathcal{C}_T-1\}} (1+|x|)^p \mathrm{d}\mu^0(x)\Big)^{1/p},
\end{aligned}$$

which together with our choice of $R_\delta > 0$ in (3.43) further yields that

$$\Big(\int_{\{(y,\sigma_\nu)\text{ s.t. }\|\sigma_\nu(\cdot)\|_{C^0([0,T],\mathbb{R}^d)}>R_\delta\}} \Big|\int_0^t \Big(v(s,\sigma_\nu(s)) - w(s,\sigma_\nu(s))\Big)\mathrm{d}s\Big|^p \mathrm{d}\boldsymbol{\eta}_\nu(y,\sigma_\nu)\Big)^{1/p} \le \delta. \qquad (3.51)$$

We now focus our attention on the first term in the right-hand side of (3.49). By construction, there exists an integer $j \in \{0,\dots,N-1\}$ such that $t \in [t_j, t_{j+1}]$, which means that

$$\begin{aligned}
&\Big(\int_{\{(y,\sigma_\nu)\text{ s.t. }\|\sigma_\nu(\cdot)\|_{C^0([0,T],\mathbb{R}^d)}\le R_\delta\}} \Big|\int_0^t \Big(v(s,\sigma_\nu(s)) - w(s,\sigma_\nu(s))\Big)\mathrm{d}s\Big|^p \mathrm{d}\boldsymbol{\eta}_\nu(y,\sigma_\nu)\Big)^{1/p} \\
&\le \sum_{i=0}^{j-1} \Big(\int_{\{(y,\sigma_\nu)\text{ s.t. }\|\sigma_\nu(\cdot)\|_{C^0([0,T],\mathbb{R}^d)}\le R_\delta\}} \Big|\int_{t_i}^{t_{i+1}} \Big(v(s,\sigma_\nu(s)) - v_i(s,\sigma_\nu(s))\Big)\mathrm{d}s\Big|^p \mathrm{d}\boldsymbol{\eta}_\nu(y,\sigma_\nu)\Big)^{1/p} \\
&\quad + \Big(\int_{\{(y,\sigma_\nu)\text{ s.t. }\|\sigma_\nu(\cdot)\|_{C^0([0,T],\mathbb{R}^d)}\le R_\delta\}} \Big|\int_{t_j}^{t} \Big(v(s,\sigma_\nu(s)) - v_j(s,\sigma_\nu(s))\Big)\mathrm{d}s\Big|^p \mathrm{d}\boldsymbol{\eta}_\nu(y,\sigma_\nu)\Big)^{1/p} \\
&\le \sum_{i=0}^{j-1} \Big(\int_{\{(y,\sigma_\nu)\text{ s.t. }\|\sigma_\nu(\cdot)\|_{C^0([0,T],\mathbb{R}^d)}\le R_\delta\}} \Big|\int_{t_i}^{t_{i+1}} \Big(v(s,\sigma_\nu(s)) - v_i(s,\sigma_\nu(s))\Big)\mathrm{d}s\Big|^p \mathrm{d}\boldsymbol{\eta}_\nu(y,\sigma_\nu)\Big)^{1/p} + \delta,
\end{aligned} \qquad (3.52)$$

where we used the estimates of (3.41) and (3.44) in the second inequality. Now, given an arbitrary integer $i \in \{0,\dots,j-1\}$, the facts that $v(\cdot), v_i(\cdot) \in \mathscr{L}([0,T], C^0(\mathbb{R}^d,\mathbb{R}^d))$ both satisfy the sublinearity estimate (3.41) along with the regularity assumption of hypothesis **(CI)**-(iii) allow us to write

$$\begin{aligned}
&\Big(\int_{\{(y,\sigma_\nu)\text{ s.t. }\|\sigma_\nu(\cdot)\|_{C^0([0,T],\mathbb{R}^d)}\le R_\delta\}} \Big|\int_{t_i}^{t_{i+1}} \Big(v(s,\sigma_\nu(s)) - v_i(s,\sigma_\nu(s))\Big)\mathrm{d}s\Big|^p \mathrm{d}\boldsymbol{\eta}_\nu(y,\sigma_\nu)\Big)^{1/p} \\
&\le \Big(\int_{\{(y,\sigma_\nu)\text{ s.t. }\|\sigma_\nu(\cdot)\|_{C^0([0,T],\mathbb{R}^d)}\le R_\delta\}} \Big|\int_{t_i}^{t_{i+1}} \Big(v(s,\sigma_\nu(t_i)) - v_i(s,\sigma_\nu(t_i))\Big)\mathrm{d}s\Big|^p \mathrm{d}\boldsymbol{\eta}_\nu(y,\sigma_\nu)\Big)^{1/p} \\
&\quad + \Big(\int_{\{(y,\sigma_\nu)\text{ s.t. }\|\sigma_\nu(\cdot)\|_{C^0([0,T],\mathbb{R}^d)}\le R_\delta\}} \Big|\int_{t_i}^{t_{i+1}} \Big(v(s,\sigma_\nu(s)) - v(s,\sigma_\nu(t_i))\Big)\mathrm{d}s\Big|^p \mathrm{d}\boldsymbol{\eta}_\nu(y,\sigma_\nu)\Big)^{1/p} \\
&\quad + \Big(\int_{\{(y,\sigma_\nu)\text{ s.t. }\|\sigma_\nu(\cdot)\|_{C^0([0,T],\mathbb{R}^d)}\le R_\delta\}} \Big|\int_{t_i}^{t_{i+1}} \Big(v_i(s,\sigma_\nu(s)) - v_i(s,\sigma_\nu(t_i))\Big)\mathrm{d}s\Big|^p \mathrm{d}\boldsymbol{\eta}_\nu(y,\sigma_\nu)\Big)^{1/p} \\
&\le \Big\|\int_{t_i}^{t_{i+1}} v(s)_{|B(0,R_\delta)}\mathrm{d}s - \int_{t_i}^{t_{i+1}} v_i^\delta(s)\mathrm{d}s\Big\|_{C^0(B(0,R_\delta),\mathbb{R}^d)} \\
&\quad + 2\Big(\int_{\{(y,\sigma_\nu)\text{ s.t. }\|\sigma_\nu(\cdot)\|_{C^0([0,T],\mathbb{R}^d)}\le R_\delta\}} \Big|\int_{t_i}^{t_{i+1}} l(s)\int_{t_i}^{s} m(\zeta)\Big(1+|\sigma_\nu(\zeta)|+\mathcal{C}\Big)\mathrm{d}\zeta\mathrm{d}s\Big|^p \mathrm{d}\boldsymbol{\eta}_\nu(y,\sigma_\nu)\Big)^{1/p} \\
&\le \Big\|\int_{t_i}^{t_{i+1}} v(s)_{|B(0,R_\delta)}\mathrm{d}s - \int_{t_i}^{t_{i+1}} v_i^\delta(s)\mathrm{d}s\Big\|_{C^0(B(0,R_\delta,\mathbb{R}^d))} + 2(1+R_\delta+\mathcal{C})\int_{t_i}^{t_{i+1}} l(s)\int_{t_i}^{s} m(\zeta)\mathrm{d}\zeta\mathrm{d}s \\
&\le \frac{\delta}{N} + \delta\int_{t_i}^{t_{i+1}} l(s)\mathrm{d}s,
\end{aligned} \qquad (3.53)$$

where we leveraged the preliminary estimates (3.41)-(3.42) in the second and third to last inequalities, as well as (3.44)-(3.46) in the last one. Hence, by inserting the estimates of (3.53) for $i \in \{0, \dots, j-1\}$ into (3.52), combining the resulting expression with (3.51) and plugging the latter in (3.49) then yields

$$\left(\int_{\mathbb{R}^d \times \Sigma_T} \left| \int_0^t \Big(v(s, \sigma_\nu(s)) - w(s, \sigma_\nu(s)) \Big) \mathrm{d}s \right|^p \mathrm{d}\boldsymbol{\eta}_\nu(y, \sigma_\nu) \right)^{1/p} \leq \delta \big(3 + \|l(\cdot)\|_1 \big). \tag{3.54}$$

By merging (3.48) and (3.54) into (3.47), raising the resulting inequality to the power p, applying Grönwall's lemma and then taking the power $1/p$ of the corresponding expression, we finally recover the estimate

$$W_p(\mu(t), \nu(t)) \leq \delta\, C_p (3 + \|l(\cdot)\|_1) \exp\Big(C'_p \, \|l(\cdot)\|_1^p \Big), \tag{3.55}$$

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

Step 3 – Back to the unrelaxed problem via Filippov's estimates. At this stage, one should note that $\nu(\cdot) \in \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d))$ is not a solution of (3.40), since $w(t) \in V(t, \mu(t))$ for $\mathscr{L}^1$-almost every $t \in [0,T]$. Nevertheless, observe that the mismatch function $\eta_\delta(\cdot) \in L^1([0,T], \mathbb{R}_+)$, defined by

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

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

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

as a consequence of hypothesis **(CI)**-(iv). Thus, by leveraging the results of Theorem 3.3 while recollecting the a priori bound of (3.43), there exists a solution $\mu_\delta(\cdot) \in \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d))$ of the unrelaxed Cauchy problem (3.40) which satisfies

$$\begin{aligned} W_p(\mu_\delta(t), \nu(t)) &\leq C_p \Big(\delta C_p \, \|L(\cdot)\|_1 \, (3 + \|l(\cdot)\|_1) \exp\Big(C'_p \, \|l(\cdot)\|_1^p \Big) + \delta \Big) \exp\Big(C'_p \, \|l(\cdot)\|^p_{L^1([0,t])} + \chi_p(t) \Big) \\ &\leq \delta C_p \Big(1 + (3 + \|l(\cdot)\|_1) \, \|\chi_p(\cdot)\|_\infty \Big) \exp\Big(C'_p \, \|l(\cdot)\|_1^p + \|\chi_p(\cdot)\|_\infty \Big), \end{aligned} \tag{3.56}$$

for all times $t \in [0,T]$, where the map $\chi_p(\cdot)$ is defined in (3.6). Thus, by combining (3.55) and (3.56), applying the triangle inequality and redefining the constant $\delta > 0$ as

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

we can finally conclude that the solution $\mu_\delta(\cdot)$ of the unrelaxed Cauchy problem (3.40) is such that

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

which ends the proof of Theorem 3.7. □

4 Existence "à la Peano" in the Carathéodory setting

In this section, we establish a general existence result for continuity inclusions whose right-hand sides are continuous with respect to the measure and space variables. In that case, it is well-known even that one must impose a convexity hypothesis on the admissible velocities of the inclusion (see e.g. [11, Chapter 2]) in order to prove the existence of solutions, even in finite-dimensional euclidean spaces.

Hypotheses (P).

(*i*) *The set-valued map* $(t,\mu) \in [0,T] \times \mathscr{P}_p(\mathbb{R}^d) \rightrightarrows V(t,\mu) \subset C^0(\mathbb{R}^d,\mathbb{R}^d)$ *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* $\mathscr{L}^1$*-almost every* $t \in [0,T]$*, any* $\mu \in \mathscr{P}_p(\mathbb{R}^d)$ *and every* $v \in V(t,\mu)$*, it holds*

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

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

(*iii*) *For every compact set* $\mathscr{K} \subset \mathscr{P}_p(\mathbb{R}^d)$*, any* $\varepsilon > 0$ *and each* $R > 0$*, there exists a map* $\delta_R(\cdot) \in L^1([0,T],\mathbb{R}_+^*)$ *such that for* $\mathscr{L}^1$*-almost every* $t \in [0,T]$*, any* $\mu \in \mathscr{K}$ *and every* $v \in V(t,\mu)$*, it holds*

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

for all $x, y \in B(0,R)$ *satisfying* $|x-y| \leq \delta_R(t)$*.*

In the sequel, we adapt to the setting of continuity inclusions an argument outlined in [47, Chapter 1 – Theorem 1] for Carathéodory differential equations, which is based on a variant of the classical explicit Euler scheme. To the best of our knowledge, this approach is fairly new even for classical differential inclusions, as it does not seem to have been thoroughly investigated in the context of finite-dimensional vector spaces. In the context of continuity equations with nonlocal velocities, the latter can also be seen as a close relative to the methodologies developped e.g. in [35, 75, 76].

Theorem 4.1 (A Peano existence theorem for continuity inclusions)**.** *Let* $V : [0,T] \times \mathscr{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d,\mathbb{R}^d)$ *be a set-valued map satisfying hypotheses* **(P)***. Then, for every initial datum* $\mu^0 \in \mathscr{P}_p(\mathbb{R}^d)$*, there exists a solution* $\mu(\cdot) \in \mathrm{AC}([0,T],\mathscr{P}_p(\mathbb{R}^d))$ *to the Cauchy problem*

$$\begin{cases} \partial_t \mu(t) \in -\mathrm{div}_x\Big(V(t,\mu(t))\mu(t)\Big), \\ \mu(0) = \mu^0. \end{cases} \tag{4.1}$$

The proof of Theorem 4.1 follows a constructive scheme – in a fashion that is somewhat similar to that of Theorem 3.3 –, and is split into three steps. We start in Step 1 by building a sequence of trajectory-selection pairs solutions of continuity equations with delayed velocity inclusions, and proceed by showing in Step 2 that it is weakly compact by carefully adapting some of the arguments subtending Theorem 3.6. We then conclude in Step 3 by proving that the underlying cluster points are indeed solution of the Cauchy problem (4.1).

Proof of Theorem 4.1. In what follows, our goal is to build a sequence of trajectory-selection pairs $(\mu_n(\cdot), v_n(\cdot)) \in \mathrm{AC}([0,T],\mathscr{P}_p(\mathbb{R}^d)) \times \mathscr{L}([0,T],C^0(\mathbb{R}^d,\mathbb{R}^d))$ solutions of the Cauchy problems

$$\begin{cases} \partial_t \mu_n(t) + \mathrm{div}_x(v_n(t)\mu_n(t)) = 0, \\ \mu_n(0) = \mu^0. \end{cases} \tag{4.2}$$

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

$$v_n(t) \in V\Big(t, \mu_n(t - \tfrac{T}{n})\Big), \tag{4.3}$$

for $\mathscr{L}^1$-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 momentum and regularity bounds

$$\mathcal{M}_p(\mu_n(t)) \leq \mathcal{C} \qquad \text{and} \qquad W_p(\mu_n(\tau), \mu_n(t)) \leq c_p \int_\tau^t m(s)\mathrm{d}s \tag{4.4}$$

for all times $0 \leq \tau \leq t \leq T$ and each $n \geq 1$. Therein, the constants $\mathcal{C}, c_p > 0$ shall be uniform with respect to $n \geq 1$, and 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 start by building the pair $(\mu_n(\cdot), v_n(\cdot))$ satisfying the aforedescribed conditions by performing an induction on $k \in \{0, \dots, n-1\}$. First, observe that as a consequence of hypotheses **(P)**-(i), the set-valued map

$$t \in [0, \tfrac{T}{n}] \rightrightarrows V(t, \mu^0) \subset C^0(\mathbb{R}^d, \mathbb{R}^d)$$

is $\mathscr{L}^1$-measurable with closed nonempty images. Thus, it admits a measurable selection $t \in [0, \frac{T}{n}] \mapsto v_n^0(t) \in V(t, \mu^0)$ by Theorem 2.6, and the underlying Carathéodory velocity field $v_n^0 : [0, \frac{T}{n}] \times \mathbb{R}^d \mapsto \mathbb{R}^d$ satisfies the sublinearity estimate

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

for $\mathscr{L}^1$-almost every $t \in [0, \frac{T}{n}]$ and all $x \in \mathbb{R}^d$, as consequence of hypothesis **(P)**-(ii) along with Lemma 2.14. In particular it complies with hypothesis **(CE)**-(i), and by Theorem 2.18 applied on the time interval $[0, \frac{T}{n}]$ with $v(\cdot, \cdot) := v_n^0(\cdot, \cdot)$, there exists a curve of measures $\mu_n^0 \in \mathrm{AC}([0, \frac{T}{n}], \mathscr{P}_p(\mathbb{R}^d))$ solution of the Cauchy problem

$$\begin{cases} \partial_t \mu_n^0(t) + \mathrm{div}_x(v_n^0(t)\mu_n^0(t)) = 0 & \text{in } [0, \frac{T}{n}] \times \mathbb{R}^d, \\ \mu_n^0(0) = \mu^0. \end{cases}$$

By repeating this process for $k \in \{1, \dots, n-1\}$, we show how to iteratively build a family of trajectory-selection pairs $(\mu_n^k(\cdot), v_n^k(\cdot)) \in \mathrm{AC}([\frac{kT}{n}, \frac{(k+1)T}{n}], \mathscr{P}_p(\mathbb{R}^d)) \times \mathscr{L}([\frac{kT}{n}, \frac{(k+1)T}{n}], C^0(\mathbb{R}^d, \mathbb{R}^d))$ as follows.

Given a pair $(\mu_n^{k-1}(\cdot), v_n^{k-1}(\cdot))$, it stems from hypotheses **(P)** along with the Ascoli-Arzelà theorem that $t \in [\frac{(k-1)T}{n}, \frac{kT}{n}] \rightrightarrows V(t, \mu_n^{k-1}(t)) \subset C^0(\mathbb{R}^d, \mathbb{R}^d)$ is $\mathscr{L}^1$-measurable with nonempty compact images. Whence by Lemma 2.10-(c), there exists a measurable selection

$$t \in [\tfrac{kT}{n}, \tfrac{(k+1)T}{n}] \mapsto v_n^k(t) \in V\Big(t, \mu_n^{k-1}(t - \tfrac{T}{n})\Big). \tag{4.5}$$

Besides by hypothesis **(P)**-(ii) and Lemma 2.14, the Carathéodory vector field $v_n^k : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ satisfies the sublinearity estimate

$$|v_n^k(t, x)| \leq m(t)\Big(1 + |x| + \mathcal{M}_p(\mu_n^{k-1}(t - \tfrac{T}{n}))\Big) \tag{4.6}$$

for $\mathscr{L}^1$-almost every $t \in [\frac{kT}{n}, \frac{(k+1)T}{n}]$, all $x \in \mathbb{R}^d$ and each $k \in \{1, \dots, n-1\}$. Thus by applying Theorem 2.18 on the time interval $[\frac{kT}{n}, \frac{(k+1)T}{n}]$ with $v(\cdot, \cdot) := v_n^k(\cdot, \cdot)$, we can subsequently define $\mu_n^k(\cdot) \in \mathrm{AC}([\frac{kT}{n}, \frac{(k+1)T}{n}], \mathscr{P}_p(\mathbb{R}^d))$ as being one of the solutions of the Cauchy problem

$$\begin{cases} \partial_t \mu_n^k(t) + \mathrm{div}_x(v_n^k(t)\mu_n^k(t)) = 0 & \text{in } [\frac{kT}{n}, \frac{(k+1)T}{n}] \times \mathbb{R}^d, \\ \mu_n^k(\frac{kT}{n}) = \mu_n^{k-1}(\frac{kT}{n}). \end{cases}$$

By classical concatenation properties for solutions of continuity equations (see e.g. [45, Lemma 4.4]), the trajectory-selection pair $(\mu_n(\cdot), v_n(\cdot)) \in \mathrm{AC}([0, T], \mathscr{P}_p(\mathbb{R}^d)) \times \mathscr{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$ defined by

$$\mu_n(t) := \mu_n^k(t) \qquad \text{and} \qquad v_n(t) := v_n^k(t), \tag{4.7}$$

for $t \in [\frac{kT}{n}, \frac{(k+1)T}{n}]$ and $k \in \{0, \dots, n-1\}$ is a solution of the Cauchy problem (4.2). Moreover, it can be checked using (4.5) that it satisfies the shifted pointwise velocity inclusion (4.3).

Step 2 – Momentum estimates and compactness. Our next goal is to establish the uniform regularity and momentum bounds of (4.4), and to show that they yield the compactness of the sequence of pairs $(\mu_n(\cdot), v_n(\cdot)) \subset \mathrm{AC}([0, T], \mathscr{P}_p(\mathbb{R}^d)) \times \mathscr{L}([0, T], C^0(\mathbb{R}^d, \mathbb{R}^d))$. First, notice that as a consequence of the construction detailed in Step 1, the curves $\mu_n(\cdot)$ satisfy

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

for each $n \geq 1$. This allows us to resort to the momentum estimate of Remark 2.23, which writes as

$$\mathcal{M}_p(\mu_n(t)) \leq C_p\Big(\mathcal{M}_p(\mu^0) + \int_0^t m(s)\Big(1 + \mathcal{M}_p(\mu_n(s - \tfrac{T}{n}))\Big)\mathrm{d}s\Big) \exp\Big(C_p' \, \|m(\cdot)\|^p_{L^1([0,t])}\Big), \tag{4.8}$$

for all times $t \in [0, T]$, again with 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 the periodisation of $m(\cdot)$ on $[T, 2T]$, namely

$$\hat{m}(t) := \begin{cases} m(t) & \text{if } t \in [0, T], \\ m(t - T) & \text{if } t \in [T, 2T], \end{cases} \tag{4.9}$$

one can rewrite the estimate displayed in (4.8) as

$$\begin{aligned}
\mathcal{M}_p(\mu_n(t)) \leq C_p\Bigg(&\mathcal{M}_p(\mu^0) + \int_0^t m(s)\mathrm{d}s + \mathcal{M}_p(\mu^0)\int_0^{T/n} m(s)\mathrm{d}s \\
&+ \int_0^{\max\{0, t-T/n\}} m(s + \tfrac{T}{n})\mathcal{M}_p(\mu_n(s))\mathrm{d}s\Bigg) \exp\Big(C_p' \, \|m(\cdot)\|^p_{L^1([0,t])}\Big) \\
\leq C_p(1 + \mathcal{M}_p(\mu^0))&\Big(1 + \|m(\cdot)\|_{L^1([0,t])}\Big) \exp\Big(C_p' \, \|m(\cdot)\|^p_{L^1([0,t])}\Big) \\
&+ C_p\Big(\int_0^t \hat{m}(s + \tfrac{T}{n})\mathcal{M}_p(\mu_n(s))\mathrm{d}s\Big) \exp\Big(C_p' \, \|m(\cdot)\|^p_{L^1([0,t])}\Big),
\end{aligned}$$

where the first inequality follows from a simple change of variable along with the fact that $\mu_n(t) = \mu^0$ for $t \in [-\frac{T}{n}, 0]$, while the second one stems from the fact that both functions $\hat{m}(\cdot)$ and $\mathcal{M}_p(\mu_n(\cdot))$ are non-negative. Thence, a standard application of Grönwall's lemma yields

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

In turn, upon noticing that

$$\|\hat{m}(\cdot)\|_{L^1([0,t+T/n])} \leq 2 \, \|m(\cdot)\|_1$$

for all times $t \in [0, t]$ and each $n \geq 1$ as a consequence of (4.9), 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 $n \geq 1$. Thus, the velocity fields $v_n : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ satisfy the uniform sublinearity bounds

$$|v_n(t, x)| \leq (1 + \mathcal{C})m(t)(1 + |x|) \tag{4.10}$$

for $\mathscr{L}^1$-almost every $t \in [0, T]$, all $x \in \mathbb{R}^d$ and each $n \geq 1$. In particular by Theorem 2.18, 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

$$W_p(\mu_n(\tau), \mu_n(t)) \leq c_p \int_\tau^t m(s)\mathrm{d}s$$

for all times $0 \leq \tau \leq t \leq T$. To summarise, we have shown that the sequence of curves $(\mu_n(\cdot)) \subset \mathrm{AC}([0, T], \mathscr{P}_p(\mathbb{R}^d))$ satisfies the uniform momentum and absolute continuity estimates (4.4).

By invoking the momentum inequality (2.18) and the equi-integrability bounds (2.20) of Proposition 2.22 with possibly different constants $C_T > 0$, we can deduce from Proposition 2.3 the existence of a compact set $\mathscr{K} \subset \mathscr{P}_p(\mathbb{R}^d)$, depending only on μ^0 and $\|m(\cdot)\|_1$, such that

$$\mu_n(t) \in \mathscr{K}$$

for all times $t \in [0,T]$. Analogously, it follows from hypothesis **(P)**-(iii) and the Ascoli-Arzelà theorem that for $\mathscr{L}^1$-almost every $t \in [0,T]$ and each $m \geq 1$, there exists a compact set $\mathcal{K}_t^m \subset C^0(B(0,m),\mathbb{R}^d)$ such that

$$v_n(t)_{|B(0,m)} \in \mathcal{K}_t^m$$

for each $n \geq 1$. By repeating the compactness argument detailed in the proof of Theorem 3.6, there exists a trajectory-selection pair $(\mu(\cdot), v(\cdot)) \in \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d)) \times \mathscr{L}([0,T], C^0(\mathbb{R}^d,\mathbb{R}^d))$ for which

$$\sup_{t\in[0,T]} W_p(\mu_n(t),\mu(t)) \underset{n\to+\infty}{\longrightarrow} 0 \tag{4.11}$$

along a subsequence that we do not relabel, as well as

$$v_n(\cdot) \underset{n\to+\infty}{\rightharpoonup} v(\cdot) \tag{4.12}$$

weakly in $L^1([0,T], C^0(B(0,R),\mathbb{R}^d))$, along adequate subsequences which depend on $R > 0$. In particular, the latter convergence property implies that

$$\int_0^T \langle \boldsymbol{\nu}(t), v(t) - v_n(t)\rangle_{C^0(B(0,R),\mathbb{R}^d)} \mathrm{d}t \underset{n\to+\infty}{\longrightarrow} 0, \tag{4.13}$$

for each $\boldsymbol{\nu}(\cdot) \in L^\infty([0,T], \mathscr{M}(B(0,R),\mathbb{R}^d))$, again up to a subsequence.

Step 3 – Properties of the limit trajectory-selection pair. To conclude the proof of Theorem 4.1, there remains to show that the limit pair $(\mu(\cdot), v(\cdot))$ solves the Cauchy problem (4.1). First, it can be straightforwardly verified that

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

for all times $t \in [0,T]$ as a consequence of (4.11), while (4.10) and (4.12) yield by Mazur's lemma that

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

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and all $x \in \mathbb{R}^d$. Moreover, one can check that the limits (3.33) and (3.36) derived within the proof of Theorem 3.6 are still valid in the present context for the sequence $(\mu_n(\cdot), v_n(\cdot))$. Thus, the limit pair $(\mu(\cdot), v(\cdot)) \in \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d)) \times \mathscr{L}([0,T], C^0(\mathbb{R}^d,\mathbb{R}^d))$ will solve

$$\begin{cases} \partial_t \mu(t) + \mathrm{div}_x(v(t)\mu(t)) = 0, \\ \mu(0) = \mu^0, \end{cases} \tag{4.14}$$

provided that

$$\int_0^T \int_{\mathbb{R}^d} \langle \nabla_x \phi(t,x), v(t,x) - v_n(t,x)\rangle \mathrm{d}\mu_n(t)(x)\mathrm{d}t \underset{n\to+\infty}{\longrightarrow} 0 \tag{4.15}$$

for each $\phi \in C_c^\infty((0,T)\times\mathbb{R}^d,\mathbb{R})$, along adequate subsequences. Proving so is however much trickier than in Theorem 3.6, as we traded the strong global Lipschitz continuity assumption of **(CI)**-(iii) for the much weaker local uniform equi-continuity of **(P)**-(iii).

To this end, we shall consider test functions $\phi \in C_c^\infty((0,T)\times\mathbb{R}^d,\mathbb{R})$ of the form

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

for all $(t,x) \in [0,T]\times\mathbb{R}^d$ with $(\zeta,\psi) \in C_c^\infty((0,T),\mathbb{R}) \times C_c^\infty(\mathbb{R}^d,\mathbb{R})$, whose linear span is dense in $C_c^\infty((0,T)\times\mathbb{R}^d,\mathbb{R})$. We fix such a pair (ζ,ψ) and let $R_\psi > 0$ be such that $\mathrm{supp}(\psi) \subset B(0,R_\psi)$. By Step 2, there exists a subsequence of $(v_n(\cdot))$ that we do not relabel for which (4.12) holds in $L^1([0,T], C^0(B(0,R),\mathbb{R}^d))$, with the radius

$$R := (1+R_\psi)\max\left\{1, (1+\mathcal{C})\,\|m(\cdot)\|_1\right\}\exp\Big((1+\mathcal{C})\,\|m(\cdot)\|_1\Big), \tag{4.16}$$

and we claim that in that case, one has

$$\sup_{\sigma(\cdot)\in \mathrm{AC}([0,T],\mathbb{R}^d)} \left\{ \left| \int_0^T \zeta(t) \left\langle \nabla\psi(\sigma(t)), v(t,\sigma(t)) - v_n(t,\sigma(t)) \right\rangle \mathrm{d}t \right| \text{ s.t. } \dot\sigma(t) = v_n(t,\sigma(t)) \right.$$
$$\left. \text{for } \mathscr{L}^1\text{-almost every } t\in[0,T] \right\} \underset{n\to+\infty}{\longrightarrow} 0. \tag{4.17}$$

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

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

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

$$\left| \int_0^T \zeta(t) \left\langle \nabla\psi(\sigma_k(t)), v(t,\sigma_k(t)) - v_{n_k}(t,\sigma_k(t)) \right\rangle \mathrm{d}t \right| \geq \varepsilon \tag{4.18}$$

for each $k\geq 1$. Then for every $k\geq 1$, there must exist some $\tau_k \in \mathrm{supp}(\zeta)$ such that $\sigma_k(\tau_k)\in B(0,R_\psi)$, and it follows from (4.10) and Grönwall's lemma that

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

for all times $0\leq\tau\leq t\leq T$ and each $k\geq 1$, where $R>0$ is defined in (4.16). Thence by the Ascoli-Arzelà theorem, there exists a curve $\sigma(\cdot)\in\mathrm{AC}([0,T],\mathbb{R}^d)$ such that

$$\sup_{t\in[0,T]} |\sigma(t)-\sigma_k(t)| \underset{k\to+\infty}{\longrightarrow} 0, \tag{4.19}$$

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

$$\sup_{n\geq 1}\int_0^T |v_n(t,\sigma_k(t)) - v_n(t,\sigma(t))|\mathrm{d}t \underset{k\to+\infty}{\longrightarrow} 0, \tag{4.20}$$

which together with (4.18) and (4.19) further implies that

$$\left| \int_0^T \zeta(t) \left\langle \nabla\psi(\sigma(t)), v(t,\sigma(t)) - v_{n_k}(t,\sigma(t)) \right\rangle \mathrm{d}t \right| \geq \frac{\varepsilon}{2}, \tag{4.21}$$

whenever $k\geq 1$ is sufficiently large. However, by applying the weak-compactness property of (4.13) with the measures $\boldsymbol{\nu}(t) := \zeta(t)\nabla\psi\cdot\delta_{\sigma(t)} \in \mathscr{M}(B(0,R),\mathbb{R}^d)$ defined for all times $t\in[0,T]$, one has that

$$\int_0^T \zeta(t) \left\langle \nabla\psi(\sigma(t)), v(t,\sigma(t)) - v_{n_k}(t,\sigma(t)) \right\rangle \mathrm{d}t \underset{k\to+\infty}{\longrightarrow} 0 \tag{4.22}$$

along a subsequence that we do not relabel, which produces a contradiction with (4.21). To conclude, we observe that by Theorem 2.20, there exists a sequence of superposition measures $(\boldsymbol{\eta}^n)\subset\mathscr{P}(\mathbb{R}^d\times\Sigma_T)$ associated with the vector fields $v_n : [0,T]\times\mathbb{R}^d\to\mathbb{R}^d$ in the sense of Definition 2.19, which satisfy

$$(\pi_{\mathbb{R}^d})_\sharp\boldsymbol{\eta}^n = \mu^0 \qquad \text{and} \qquad (\mathsf{e}_t)_\sharp\boldsymbol{\eta}^n = \mu_n(t),$$

for all times $t\in[0,T]$ and each $n\geq 1$. Then, by (4.17) and Fubini's theorem, one can check that

$$\begin{aligned}
&\left| \int_0^T \zeta(t) \int_{\mathbb{R}^d} \langle \nabla\psi(x), v(t,x) - v_n(t,x)\rangle \mathrm{d}\mu_n(t)(x)\mathrm{d}t \right| \\
&\qquad = \left| \int_0^T \int_{\mathbb{R}^d\times\Sigma_T} \zeta(t) \left\langle \nabla\psi(\sigma(t)), v(t,\sigma(t)) - v_n(t,\sigma(t)) \right\rangle \mathrm{d}\boldsymbol{\eta}^n(x,\sigma)\mathrm{d}t \right| \\
&\qquad \leq \left\| \int_0^T \zeta(t) \left\langle \nabla\psi(\sigma(t)), v(t,\sigma(t)) - v_n(t,\sigma(t)) \right\rangle \mathrm{d}t \right\|_{L^\infty(\mathbb{R}^d\times\Sigma_T,\mathbb{R};\boldsymbol{\eta}^n)} \underset{n\to+\infty}{\longrightarrow} 0,
\end{aligned}$$

which implies (4.15) and thus that $(\mu(\cdot), v(\cdot))$ solves (4.14) by what precedes.

Our goal is now to show that $v(t) \in V(t, \mu(t))$ for $\mathscr{L}^1$-almost every $t \in [0, T]$. To this end, we recall that the curves $\mu_n(\cdot) \in \mathrm{AC}([0, T], \mathscr{P}_p(\mathbb{R}^d))$ are uniformly equi-continuous by (4.4), which implies that for each $\delta > 0$, there exists an integer $N_\delta \geq 1$ such that

$$W_p\Big(\mu_n(t - \tfrac{T}{n}), \mu_n(t)\Big) \leq \delta,$$

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

$$\sup_{t \in [0,T]} W_p(\tilde{\mu}_n(t), \mu(t)) \underset{n \to +\infty}{\longrightarrow} 0. \tag{4.23}$$

as well as

$$v_n(t) \in V(t, \tilde{\mu}_n(t)), \tag{4.24}$$

for $\mathscr{L}^1$-almost every $t \in [0, T]$. Besides, observe that the set-valued map $t \in [0, T] \rightrightarrows V(t, \mu(t))$ has nonempty and compact images by hypotheses **(P)**-(ii) and (iii), and is $\mathscr{L}^1$-measurable by Lemma 2.10-(c). Whence by Lemma 2.10-(a), there exists for each $n \geq 1$ a measurable selection

$$t \in [0, T] \mapsto \tilde{v}_n(t) \in V(t, \mu(t)) \tag{4.25}$$

which satisfies

$$\mathsf{d}_{\mathrm{cc}}(\tilde{v}_n(t), v_n(t)) = \mathrm{dist}_{C^0(\mathbb{R}^d, \mathbb{R}^d)}\Big(v_n(t)\,; V(t, \mu(t))\Big) \underset{n \to +\infty}{\longrightarrow} 0,$$

for $\mathscr{L}^1$-almost every $t \in [0, T]$, as a consequence of (4.23) and (4.24) together with hypothesis **(P)**-(i). By Lebesgue's dominated convergence theorem combined with (4.12), this further implies that

$$\tilde{v}_n(\cdot) \underset{n \to +\infty}{\rightharpoonup} v(\cdot) \tag{4.26}$$

weakly in $L^1([0, T], C^0(B(0, R), \mathbb{R}^d))$, along subsequences that depend on $R > 0$. Following the reasoning outlined in the proof of Theorem 3.6, we consider the functional set defined by

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

which are closed in the strong topology of $L^1([0, T], C^0(B(0, R), \mathbb{R}^d))$ under our working assumptions. These latter are also convex by hypothesis **(P)**-(i), and therefore weakly closed by Mazur's lemma, which combined with (4.25) and the convergence results of (4.26) finally yields that $v(t) \in V(t, \mu(t))$ for $\mathscr{L}^1$-almost every $t \in [0, T]$. □

Remark 4.2 (On the choice of proving Theorem 4.1 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 [11, Section 2.1 – Theorem 3] or [59, Theorem 2.9] – could be to consider first a sufficiently regular exact or approximate velocity selection $(t, \mu) \in [0, T] \times \mathscr{P}_p(\mathbb{R}^d) \mapsto v(t, \mu) \in C^0(\mathbb{R}^d, \mathbb{R}^d)$, 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 known results such as [12, Theorem 9.5.2] ensuring the existence of Carathéodory selections for Carathéodory set-valued maps do not hold 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 via an Euler scheme.via Secondly, even though results such as [12, Theorem 9.2.1] which provide families of regular approximate selections for upper-semicontinuous set-valued maps may still hold for Fréchet spaces instead of Banach ones, establishing the compactness of the underlying sequences may prove challenging, owing to the lack of uniformity with regard to their regularity.

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 for the sake of completeness parts of the proof of the measurable selection principles of Lemma 2.10, as the latter relies on general and 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 [12, 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 indeed measurable. To this end, we consider the multifunctions

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

defined for each $n \geq 1$, and which are $\mathscr{L}^1$-measurable by [12, Theorem 8.2.9]. For each closed set $\mathcal{C} \subset Y$, this implies in particular by [12, Theorem 8.1.4] that

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

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

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

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

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

as a consequence of (2.5). Conversely for any $\tau \in \mathscr{D}_{\mathcal{C}}$, 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. [79, 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

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

which 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

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

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 $\mathscr{L}^1$-measurable. It follows from classical measurability results (see e.g. [12, 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 $\mathscr{L}^1$-measurable. Since the sequence $(\varpi_n(\cdot))$ is pointwisely non-decreasing and bounded, the limits

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

exist for $\mathscr{L}^1$-almost every $t \in [0,T]$, and the map $\varpi(\cdot)$ is $\mathscr{L}^1$-measurable by [12, Theorem 8.2.5]. $\square$

B Proof of Lemma 2.21

In this second appendix, we provide a proof of Lemma 2.21. The overall strategy is identical to that developed in [17, Appendix A], up to some technical details in the compactness arguments.

Proof of Lemma 2.21. Let $\{t_k\}_{k=1}^{+\infty} \subset [0,T]$ be an arbitrary countable and dense family, $\gamma^0 \in \Gamma_o(\mu^0,\nu^0)$ be given and $\boldsymbol{\eta}_\mu, \boldsymbol{\eta}_\nu \in \mathscr{P}(\mathbb{R}^d \times \Sigma_T)$ be the superposition measures associated with $\mu(\cdot), \nu(\cdot) \in \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d))$ according to Definition 2.19 respectively. By repeating the argument detailed in [17, Appendix A] – which is based on iterative applications of the disintegration theorem and gluing lemma (see e.g. [9, Theorem 5.3.1 and Lemma 5.3.2]) –, there exists for each $n \geq 1$ a plan $\hat{\boldsymbol{\eta}}^n_{\mu,\nu} \in \Gamma(\boldsymbol{\eta}_\mu, \boldsymbol{\eta}_\nu)$ such that

$$(\pi_{\mathbb{R}^d}, \pi_{\mathbb{R}^d})_\sharp \hat{\boldsymbol{\eta}}^n_{\mu,\nu} = \gamma^0 \qquad \text{and} \qquad (\mathsf{e}_{t_k}, \mathsf{e}_{t_k})_\sharp \hat{\boldsymbol{\eta}}^n_{\mu,\nu} \in \Gamma_o(\mu(t_k), \nu(t_k)),$$

for every $k \in \{1, \dots, n\}$.

Step 1 – Tightness of the sequence $(\hat{\boldsymbol{\eta}}^n_{\mu,\nu})$. Our first goal is to show that the sequence $(\hat{\boldsymbol{\eta}}^n_{\mu,\nu}) \subset \mathscr{P}((\Sigma_T \times \mathbb{R}^d)^2)$ relatively compact for the narrow topology (2.2). Observing that $\mathbb{R}^d \times \Sigma_T$ is a complete separable Banach space and invoking [9, Theorem 5.1.3 and Remark 5.1.5], this latter property is tantamount to the existence of map $\Psi : (\mathbb{R}^d \times \Sigma_T)^2 \to [0, +\infty]$ whose sublevels are compact, such that

$$\sup_{n \geq 1} \int_{(\mathbb{R}^d \times \Sigma_T)^2} \Psi(x, \sigma_\mu, y, \sigma_\nu) \mathrm{d}\hat{\boldsymbol{\eta}}^n_{\mu,\nu}(x, \sigma_\mu, y, \sigma_\nu) < +\infty.$$

A classical candidate functional[1] when working with superposition measures is provided by

$$\psi : (x, \sigma) \mapsto \begin{cases} |x| + \displaystyle\int_0^T \frac{|\dot{\sigma}(t)|}{1 + |\sigma(t)|} \mathrm{d}t & \text{if } \sigma(\cdot) \in \mathrm{AC}([0,T], \mathbb{R}^d) \text{ and } |\dot{\sigma}(t)| \leq m(t)(1 + |\sigma(t)|), \\ +\infty & \text{otherwise,} \end{cases} \tag{B.1}$$

whose sublevels are compact in $\mathbb{R}^d \times \Sigma_T$ as an easy consequence of the Ascoli-Arzelà compactness theorem. Thus, by choosing

$$\Psi(x, \sigma_\mu, y, \sigma_\nu) := \psi(x, \sigma_\mu) + \psi(y, \sigma_\nu) < +\infty, \tag{B.2}$$

for $\boldsymbol{\eta}_\mu$-almost every $(x, \sigma_\mu) \in \mathbb{R}^d \times \Sigma_T$ and $\boldsymbol{\eta}_\nu$-almost every $(y, \sigma_\nu) \in \mathbb{R}^d \times \Sigma_T$ while recalling that $\boldsymbol{\eta}_\mu, \boldsymbol{\eta}_\nu \in \mathscr{P}(\mathbb{R}^d \times \Sigma_T)$ are concentrated on the characteristic curves generated by $v, w : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ which satisfy the sublinearity bound (2.17), one can infer from (B.1) and (B.2) that

$$\begin{aligned} &\sup_{n \geq 1} \int_{(\mathbb{R}^d \times \Sigma_T)^2} \Psi(x, \sigma_\mu, y, \sigma_\nu) \mathrm{d}\hat{\boldsymbol{\eta}}^n_{\mu,\nu}(x, \sigma_\mu, y, \sigma_\nu) \\ &\quad \leq \int_{\mathbb{R}^d \times \Sigma_T} \psi(x, \sigma_\mu) \mathrm{d}\boldsymbol{\eta}_\mu(x, \sigma_\mu) + \int_{\mathbb{R}^d \times \Sigma_T} \psi(y, \sigma_\nu) \mathrm{d}\boldsymbol{\eta}_\nu(y, \sigma_\nu) \\ &\quad \leq \mathcal{M}_p(\mu^0) + \int_0^T \int_{\mathbb{R}^d} \frac{|v(t,x)|}{1 + |x|} \mathrm{d}\mu(t)(x) \mathrm{d}t + \mathcal{M}_p(\nu^0) + \int_0^T \int_{\mathbb{R}^d} \frac{|w(t,y)|}{1 + |y|} \mathrm{d}\nu(t)(y) \mathrm{d}t < +\infty, \end{aligned}$$

where we used Fubini's theorem along with the fact that $(\mathsf{e}_t)_\sharp \boldsymbol{\eta}_\mu = \mu(t)$ and $(\mathsf{e}_t)_\sharp \boldsymbol{\eta}_\nu = \nu(t)$ for all times $t \in [0,T]$. Whence, it follows that the sequence $(\hat{\boldsymbol{\eta}}^n_{\mu,\nu}) \subset \mathscr{P}((\mathbb{R}^d \times \Sigma_T)^2)$ admits cluster points for the narrow topology.

[1] In [17], the proof of this compactness argument contains a small caveat in the definition of the functional $\psi(\cdot,\cdot)$, which is not coercive as it is written therein. The result remains however correct, up to redefining the latter as in (B.1)

Step 2 – Optimality of the limit plans. Let $\hat{\boldsymbol{\eta}}_{\mu,\nu} \in \mathscr{P}((\mathbb{R}^d \times \Sigma_T)^2)$ be a cluster point of the sequence $(\hat{\boldsymbol{\eta}}^n_{\mu,\nu})$. We recall that by construction, it holds

$$(\mathsf{e}_{t_k}, \mathsf{e}_{t_k})_\sharp \hat{\boldsymbol{\eta}}^n_{\mu,\nu} \in \Gamma_o(\mu(t_k), \nu(t_k)) \tag{B.3}$$

for each $n \geq 1$ and every $k \in \{1, \dots, n\}$. It then follows from the continuity of the evaluation maps $\mathsf{e}_{t_k} : \mathbb{R}^d \times \Sigma_T \to \mathbb{R}^d$ together with standard convergence results for sequences of image measures (see e.g. [9, Lemma 5.2.1]) that

$$(\mathsf{e}_{t_k}, \mathsf{e}_{t_k})_\sharp \hat{\boldsymbol{\eta}}^n_{\mu,\nu} \underset{n\to+\infty}{\rightharpoonup^*} (\mathsf{e}_{t_k}, \mathsf{e}_{t_k})_\sharp \hat{\boldsymbol{\eta}}_{\mu,\nu}, \tag{B.4}$$

along a subsequence that we do not relabel. We recall that as a consequence of the pointwise sublinearity bound (2.17) satisfied by $v(\cdot,\cdot), w(\cdot,\cdot)$, it follows from Grönwall's lemma that

$$\|\sigma_\mu(\cdot)\|_{C^0([0,T],\mathbb{R}^d)} \leq C_T(1+|x|) \qquad \text{and} \qquad \|\sigma_\nu(\cdot)\|_{C^0([0,T],\mathbb{R}^d)} \leq C_T(1+|y|) \tag{B.5}$$

for $\boldsymbol{\eta}_\mu$-almost every $(x, \sigma_\mu) \in \mathbb{R}^d \times \Sigma_T$ and $\boldsymbol{\eta}_\nu$-almost every $(y, \sigma_\nu) \in \mathbb{R}^d \times \Sigma_T$, where we recall that $C_T = \max\{1, \|m(\cdot)\|_1\} \exp(\|m(\cdot)\|_1)$. From this pointwise estimate, one can easily deduce that

$$\begin{aligned}
&\Big(\int_{\mathbb{R}^{2d}} |x-y|^p \mathrm{d}\Big((\mathsf{e}_{t_k}, \mathsf{e}_{t_k})_\sharp \hat{\boldsymbol{\eta}}^n_{\mu,\nu}\Big)(x,y)\Big)^{1/p} \\
&\qquad \leq \Big(\int_{\mathbb{R}^d\times\Sigma_T} |\sigma_\mu(t_k)|^p \mathrm{d}\boldsymbol{\eta}_\mu(x,\sigma_\mu)\Big)^{1/p} + \Big(\int_{\mathbb{R}^d\times\Sigma_T} |\sigma_\nu(t_k)|^p \mathrm{d}\boldsymbol{\eta}_\nu(x,\sigma_\nu)\Big)^{1/p} \\
&\qquad \leq C_T\Big(2 + \mathcal{M}_p(\mu^0) + \mathcal{M}_p(\nu^0)\Big),
\end{aligned} \tag{B.6}$$

for all $n \geq 1$. Thus, by invoking classical stability results under narrow convergence for sets of optimal transport plans (see e.g. [9, Proposition 7.1.3]), it follows from (B.3), (B.4) and (B.6) that

$$(\mathsf{e}_{t_k}, \mathsf{e}_{t_k})_\sharp \hat{\boldsymbol{\eta}}_{\mu,\nu} \in \Gamma_o(\mu(t_k), \nu(t_k)), \tag{B.7}$$

for every $k \geq 1$.

We fix now $t \in [0,T]$ and let $(t_{k_m}) \subset [0,T]$ be a subsequence satisfying $t_{k_m} \leq t$ for each $m \geq 1$ and such that $t_{k_m} \to t$ as $m \to +\infty$. Owing to the narrow continuity of the curves $\mu(\cdot), \nu(\cdot) \in \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d))$, one has that

$$\mu(t_{k_m}) \underset{m\to+\infty}{\rightharpoonup^*} \mu(t) \qquad \text{and} \qquad \nu(t_{k_m}) \underset{m\to+\infty}{\rightharpoonup^*} \nu(t). \tag{B.8}$$

Moreover, by combining the sublinearity bound (2.17) and norm estimate of (B.5), it holds that

$$|\sigma_\mu(t) - \sigma_\mu(t_{k_m})| \leq \int_{t_{k_m}}^t m(s)(1+|\sigma_\mu(s)|)\mathrm{d}s \leq (1+C_T)(1+|x|)\int_{t_{k_m}}^t m(s)\mathrm{d}s \tag{B.9}$$

for $\boldsymbol{\eta}_\mu$-almost every $(x, \sigma_\mu) \in \mathbb{R}^d \times \Sigma_T$ and each $m \geq 1$, with a similar estimate for $\boldsymbol{\eta}_\nu$-almost every $(y, \sigma_\nu) \in \mathbb{R}^d \times \Sigma_T$. Thus, given a parameter $\varepsilon > 0$, it follows from Chebyshev's inequality (see e.g. [7, Remark 1.18]) combined with (B.9) that

$$\begin{aligned}
&\hat{\boldsymbol{\eta}}_{\mu,\nu}\Big(\Big\{(x,\sigma_\mu,y,\sigma_\nu) \text{ s.t. } \Big|(\mathsf{e}_t - \mathsf{e}_{t_{k_m}}, \mathsf{e}_t - \mathsf{e}_{t_{k_m}})(x,\sigma_\mu,y,\sigma_\nu) > \varepsilon\Big|\Big\}\Big)^{1/p} \\
&\qquad \leq \frac{1}{\varepsilon}\Big(\int_{\mathbb{R}^d\times\Sigma_T} |\sigma_\mu(t) - \sigma_\mu(t_{k_m})|^p \mathrm{d}\boldsymbol{\eta}_\mu(x,\sigma_\mu)\Big)^{1/p} + \frac{1}{\varepsilon}\Big(\int_{\mathbb{R}^d\times\Sigma_T} |\sigma_\nu(t) - \sigma_\nu(t_{k_m})|^p \mathrm{d}\boldsymbol{\eta}_\nu(y,\sigma_\nu)\Big)^{1/p} \\
&\qquad \leq \frac{1}{\varepsilon} C_T\Big(2 + \mathcal{M}_p(\mu^0) + \mathcal{M}_p(\nu^0)\Big)\int_{t_{k_m}}^t m(s)\mathrm{d}s \underset{m\to+\infty}{\longrightarrow} 0,
\end{aligned}$$

which implies that the sequence $((\mathsf{e}_{t_{k_m}}, \mathsf{e}_{t_{k_m}})) \subset C^0((\mathbb{R}^d\times\Sigma_T)^2, \mathbb{R}^{2d})$ converges in $\hat{\boldsymbol{\eta}}_{\mu,\nu}$-measure towards $(\mathsf{e}_t, \mathsf{e}_t) \in C^0((\mathbb{R}^d\times\Sigma_T)^2, \mathbb{R}^{2d})$. By classical convergence results on image measures (see e.g. [9, Lemma 5.4.1]), this further implies that

$$(\mathsf{e}_{t_{k_m}}, \mathsf{e}_{t_{k_m}})_\sharp \hat{\boldsymbol{\eta}}_{\mu,\nu} \underset{m\to+\infty}{\rightharpoonup^*} (\mathsf{e}_t, \mathsf{e}_t)_\sharp \hat{\boldsymbol{\eta}}_{\mu,\nu}.$$

Again, by leveraging the estimate of (B.9), it can be shown straightforwardly that

$$\sup_{m\geq 1}\int_{\mathbb{R}^{2d}}|x-y|^p\mathrm{d}\Big((\mathsf{e}_{t_{k_m}},\mathsf{e}_{t_{k_m}})_\sharp\hat{\boldsymbol{\eta}}_{\mu,\nu}\Big)(x,y)<+\infty,$$

which, combined with yet another application of [9, Proposition 7.1.3] along with (B.7) and (B.8), finally yields that

$$(\mathsf{e}_t,\mathsf{e}_t)_\sharp\hat{\boldsymbol{\eta}}_{\mu,\nu}\in\Gamma_o(\mu(t),\nu(t)).$$

for all times $t\in[0,T]$. By following the same procedure while observing that $(\pi_{\mathbb{R}^d},\pi_{\mathbb{R}^d})_\sharp\hat{\boldsymbol{\eta}}^n_{\mu,\nu}=\gamma^0$ for each $n\geq 1$, one can also conclude that $(\pi_{\mathbb{R}^d},\pi_{\mathbb{R}^d})_\sharp\hat{\boldsymbol{\eta}}_{\mu,\nu}=\gamma^0$, which ends the proof of Lemma 2.21. □

C Proof of Proposition 2.22

In this third appendix, we detail the proof of the momentum and equi-integrability bounds displayed in Proposition 2.22, which both heavily rely on the superposition principle recalled in Theorem 2.20.

Proof of Proposition 2.22. Observe first that because $v:[0,T]\times\mathbb{R}^d\to\mathbb{R}^d$ satisfies the sublinearity estimate (2.17), it directly follows that it complies with the integral bound (2.16). Hence, by Theorem 2.20, there exists a superposition measure $\boldsymbol{\eta}_\mu\in\mathscr{P}(\mathbb{R}^d\times\Sigma_T)$ such that $(\mathsf{e}_t)_\sharp\boldsymbol{\eta}_\mu=\mu(t)$ for all times $t\in[0,T]$, which implies in particular that

$$\begin{aligned}\mathcal{M}_p^p(\mu(t))&=\int_{\mathbb{R}^d}|x|^p\mathrm{d}\mu(t)(x)\\&=\int_{\mathbb{R}^d\times\Sigma_T}|\mathsf{e}_t(x,\sigma_\mu)|^p\mathrm{d}\boldsymbol{\eta}_\mu(x,\sigma_\mu)\\&\leq\int_{\mathbb{R}^d\times\Sigma_T}\left(|x|+\int_0^t|v(s,\sigma_\mu(s))|\mathrm{d}s\right)^p\mathrm{d}\boldsymbol{\eta}_\mu(x,\sigma_\mu),\end{aligned}\tag{C.1}$$

for all times $t\in[0,T]$. In addition, by the very definition of a superposition measure, one can verify that (2.17) in fact yields the following lifted estimate on the space of curves

$$|v(t,\sigma_\mu(t))|\leq m(t)(1+|\sigma_\mu(t)|),\tag{C.2}$$

which holds for $\mathscr{L}^1$-almost every $t\in[0,T]$ and $\boldsymbol{\eta}_\mu$-almost every $(x,\sigma_\mu)\in\mathbb{R}^d\times\Sigma_T$. Thus, by combining (C.1) and (C.2), one further has that

$$\mathcal{M}_p^p(\mu(t))\leq\int_{\mathbb{R}^d\times\Sigma_T}\left(|x|+\int_0^t m(s)(1+|\sigma_\mu(s)|)\mathrm{d}s\right)^p\mathrm{d}\boldsymbol{\eta}_\mu(x,\sigma_\mu),\tag{C.3}$$

for all times $t\in[0,T]$. Observing that $\sigma_\mu(\cdot)\in C^0([0,T],\mathbb{R}^d)$ for $\boldsymbol{\eta}_\mu$-almost every $(x,\sigma_\mu)\in\mathbb{R}^d\times\Sigma_T$, one further has that

$$m(\cdot)^{1/p}|\sigma_\mu(\cdot)|\in L^p([0,T],\mathbb{R}_+).$$

Hence, by applying Hölder's inequality in (C.3) while recalling that $(\pi_{\mathbb{R}^d})_\sharp\boldsymbol{\eta}_\mu=\mu^0$, we obtain that

$$\begin{aligned}\mathcal{M}_p^p(\mu(t))&\leq 2^{p-1}\int_{\mathbb{R}^d}\left(|x|+\int_0^t m(s)\mathrm{d}s\right)^p\mathrm{d}\mu^0(x)\\&\quad+2^{p-1}\,\|m(\cdot)\|_1^{p/q}\int_{\mathbb{R}^d\times\Sigma_T}\left(\int_0^t m(s)|\sigma_\mu(s)|^p\mathrm{d}s\right)\mathrm{d}\boldsymbol{\eta}_\mu(x,\sigma_\mu)\\&=2^{p-1}\int_{\mathbb{R}^d}\left(|x|+\int_0^t m(s)\mathrm{d}s\right)^p\mathrm{d}\mu^0(x)+2^{p-1}\,\|m(\cdot)\|_1^{p/q}\int_0^t m(s)\mathcal{M}_p^p(\mu(s))\mathrm{d}s,\end{aligned}\tag{C.4}$$

for all times $t\in[0,T]$, wherein $q\in(1,+\infty]$ is the conjugate exponent of $p\in[1,+\infty)$ and where we used Fubini's theorem to obtain the last equality. Thus, by applying Grönwall's lemma in (C.4) and the triangle inequality for the $L^p(\mathbb{R}^d,\mathbb{R};\mu^0)$-norm, it finally holds that

$$\mathcal{M}_p(\mu(t))\leq C_p\left(\mathcal{M}_p(\mu^0)+\int_0^t m(s)\mathrm{d}s\right)\exp\left(C_p'\,\|m(\cdot)\|_1^p\right),$$

for all times $t \in [0,T]$, which conclude the proof of the momentum bound (2.18).

We now turn our attention to the proof of the equi-integrability estimate (2.20). Start by observing that as a consequence of the sublinearity estimate (C.2), it holds by Grönwall's lemma that

$$|\sigma_\mu(t)| \leq \left(|x| + \int_0^t m(s)\mathrm{d}s\right) \exp\left(\int_0^t m(s)\mathrm{d}s\right) \leq C_T(1+|x|), \tag{C.5}$$

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

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

for all times $t \in [0,T]$, which ends the proof of our claim. □

D Proof of Proposition 2.24

In this fourth and last appendix, we detail the proof of the stability estimates of Proposition 2.24. This time, we will also need to leverage the general result of Lemma 2.21 which provides the existence of superposition plans inducing optimal transport plans.

Proof of Proposition 2.24. Observe first that, owing to the hypotheses satisfied by both $v(\cdot,\cdot)$ and $w(\cdot,\cdot)$, there exists by Theorem 2.20 two superposition measures $\boldsymbol{\eta}_\mu, \boldsymbol{\eta}_\nu \in \mathscr{P}(\mathbb{R}^d \times \Sigma_T)$ such that $(\mathsf{e}_t)_\sharp \boldsymbol{\eta}_\mu = \mu(t)$ and $(\mathsf{e}_t)_\sharp \boldsymbol{\eta}_\nu = \nu(t)$ for all times $t \in [0,T]$. Choosing an element $\gamma^0 \in \Gamma_o(\mu^0,\nu^0)$ and a transport plan $\hat{\boldsymbol{\eta}}_{\mu,\nu} \in \Gamma(\boldsymbol{\eta}_\mu, \boldsymbol{\eta}_\nu)$ provided by Lemma 2.21, one has that

$$\begin{aligned}
W_p^p(\mu(t),\nu(t)) &= \int_{(\mathbb{R}^d\times\Sigma_T)^2} |\mathsf{e}_t(x,\sigma_\mu) - \mathsf{e}_t(y,\sigma_\nu)|^p \mathrm{d}\hat{\boldsymbol{\eta}}_{\mu,\nu}(x,\sigma_\mu,y,\sigma_\nu) \\
&\leq \int_{(\mathbb{R}^d\times\Sigma_T)^2} \left(|x-y| + \int_0^t |v(s,\sigma_\mu(s)) - w(s,\sigma_\nu(s))|\mathrm{d}s\right)^p \mathrm{d}\hat{\boldsymbol{\eta}}_{\mu,\nu}(x,\sigma_\mu,y,\sigma_\nu) \\
&\leq \int_{(\mathbb{R}^d\times\Sigma_T)^2} \Bigg(|x-y| + \int_0^t l(s)|\sigma_\mu(s) - \sigma_\nu(s)|\mathrm{d}s \\
&\qquad\qquad + \int_0^t |v(s,\sigma_\nu(s)) - w(s,\sigma_\nu(s))|\mathrm{d}s\Bigg)^p \mathrm{d}\hat{\boldsymbol{\eta}}_{\mu,\nu}(x,\sigma_\mu,y,\sigma_\nu),
\end{aligned} \tag{D.1}$$

for all times $t \in [0,T]$. Notice now that since $\nu(s) = (\mathsf{e}_s)_\sharp \boldsymbol{\eta}_\nu$ for all times $s \in [0,t]$, one has

$$\begin{aligned}
&\nu(s)\Big(\Big\{y \in \mathbb{R}^d \text{ s.t. } |v(s,y) - w(s,y)| > M\Big\}\Big) \\
&\qquad = \boldsymbol{\eta}_\nu\Big(\Big\{(y,\sigma_\nu) \in \mathbb{R}^d \times \Sigma_T \text{ s.t. } |v(s,\sigma_\nu(s)) - w(s,\sigma_\nu(s))| > M\Big\}\Big),
\end{aligned}$$

for every $M > 0$. In particular, the latter identity implies that

$$|v(s,\sigma_\nu(s)) - w(s,\sigma_\nu(s))| \leq \|v(s) - w(s)\|_{L^\infty(\mathbb{R}^d,\mathbb{R}^d;\nu(s))} \tag{D.2}$$

for $\mathscr{L}^1$-almost every $s \in [0,t]$ and $\boldsymbol{\eta}_\nu$-almost every $(y,\sigma_\nu) \in \mathbb{R}^d \times \Sigma_T$. Whence, by combining (D.1) and (D.2) and applying Hölder's inequality in a similar fashion to what we did in the proof of Proposition

2.22 in Appendix C above, it further holds that

$$
\begin{aligned}
W_p^p(\mu(t),\nu(t)) &\leq 2^{p-1}\int_{\mathbb{R}^{2d}}\left(|x-y|+\int_0^t \|v(s)-w(s)\|_{L^\infty(\mathbb{R}^d,\mathbb{R}^d;\,\nu(s))}\,\mathrm{d}s\right)^p \mathrm{d}\gamma^0(x,y)\\
&\quad + 2^{p-1}\,\|l(\cdot)\|^{p/q}_{L^1([0,t])}\int_{(\mathbb{R}^d\times\Sigma_T)^2}\left(\int_0^t l(s)|\sigma_\mu(s)-\sigma_\nu(s)|^p\mathrm{d}s\right)\mathrm{d}\hat{\boldsymbol{\eta}}(x,\sigma_\mu,y,\sigma_\nu)\\
&= 2^{p-1}\int_{\mathbb{R}^{2d}}\left(|x-y|+\int_0^t \|v(s)-w(s)\|_{L^\infty(\mathbb{R}^d,\mathbb{R}^d;\,\nu(s))}\,\mathrm{d}s\right)^p \mathrm{d}\gamma^0(x,y)\\
&\quad + 2^{p-1}\,\|l(\cdot)\|^{p/q}_{L^1([0,t])}\int_0^t l(s)W_p^p(\mu(s),\nu(s))\mathrm{d}s,
\end{aligned}
\tag{D.3}
$$

where the last inequality stems from an application of Fubini's theorem. Thus, by applying Grönwall's lemma to (D.3) along with the triangle inequality of the $L^p(\mathbb{R}^d,\mathbb{R};\gamma^0)$-norm, we can conclude that

$$
W_p(\mu(t),\nu(t)) \leq C_p\left(W_p(\mu^0,\nu^0)+\int_0^t \|v(s)-w(s)\|_{L^\infty(\mathbb{R}^d,\mathbb{R}^d;\,\nu(s))}\,\mathrm{d}s\right)\exp\left(C_p'\,\|l(\cdot)\|^p_{L^1([0,t])}\right) \tag{D.4}
$$

for all times $t\in[0,T]$, which yields the first stability estimate (2.22).

Fix next an element $R>0$ and observe that under our working assumptions, the map

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

is Lebesgue integrable. In that case, one can prove by repeating the series of computations leading to (D.4) that

$$
\begin{aligned}
W_p(\mu(t),\nu(t)) \leq C_p\Bigg(W_p(\mu^0,\nu^0)+\left(\int_{\mathbb{R}^d\times\Sigma_T}\left(\int_0^t |v(s,\sigma_\nu(s))-w(s,\sigma_\nu(s))|\mathrm{d}s\right)^p \mathrm{d}\boldsymbol{\eta}_\nu(y,\sigma_\nu)\right)^{1/p}\Bigg)\\
\times\exp\left(C_p'\,\|l(\cdot)\|^p_{L^1([0,t])}\right),
\end{aligned}
\tag{D.5}
$$

for all times $t\in[0,T]$. We focus our attention on the integral term appearing in the right-hand side of (D.5), and observe that the latter can be split into two parts as

$$
\begin{aligned}
&\int_{\mathbb{R}^d\times\Sigma_T}\left(\int_0^t |v(s,\sigma_\nu(s))-w(s,\sigma_\nu(s))|\mathrm{d}s\right)^p \mathrm{d}\boldsymbol{\eta}_\nu(y,\sigma_\nu)\\
&=\int_{\{(y,\sigma_\nu)\text{ s.t. }\|\sigma_\nu(\cdot)\|_{C^0([0,T],\mathbb{R}^d)}\leq R\}}\left(\int_0^t |v(s,\sigma_\nu(s))-w(s,\sigma_\nu(s))|\mathrm{d}s\right)^p \mathrm{d}\boldsymbol{\eta}_\nu(y,\sigma_\nu)\\
&\quad+\int_{\{(y,\sigma_\nu)\text{ s.t. }\|\sigma_\nu(\cdot)\|_{C^0([0,T],\mathbb{R}^d)}>R\}}\left(\int_0^t |v(s,\sigma_\nu(s))-w(s,\sigma_\nu(s))|\mathrm{d}s\right)^p \mathrm{d}\boldsymbol{\eta}_\nu(y,\sigma_\nu),
\end{aligned}
\tag{D.6}
$$

for all $t\in[0,T]$. The first term in the right-hand side of (D.6) can be estimated straightforwardly as

$$
\begin{aligned}
&\int_{\{(y,\sigma_\nu)\text{ s.t. }\|\sigma_\nu(\cdot)\|_{C^0([0,T],\mathbb{R}^d)}\leq R\}}\left(\int_0^t |v(s,\sigma_\nu(s))-w(s,\sigma_\nu(s))|\mathrm{d}s\right)^p \mathrm{d}\boldsymbol{\eta}_\nu(y,\sigma_\nu)\\
&\qquad\leq\left(\int_0^t \|v(s)-w(s)\|_{L^\infty(B(0,R),\mathbb{R}^d;\,\nu(s))}\,\mathrm{d}s\right)^p,
\end{aligned}
\tag{D.7}
$$

for all times $t\in[0,T]$. Concerning the second term in the right-hand side of (D.6), one has as a consequence of the sublinearity assumptions made on $v(\cdot,\cdot)$ and $w(\cdot,\cdot)$ and Hölder's inequality that

$$
\begin{aligned}
&\int_{\{(y,\sigma_\nu)\text{ s.t. }\|\sigma_\nu(\cdot)\|_{C^0([0,T],\mathbb{R}^d)}>R\}}\left(\int_0^t |v(s,\sigma_\nu(s))-w(s,\sigma_\nu(s))|\mathrm{d}s\right)^p \mathrm{d}\boldsymbol{\eta}_\nu(y,\sigma_\nu)\\
&\leq 2^p\int_{\{(y,\sigma_\nu)\text{ s.t. }\|\sigma_\nu(\cdot)\|_{C^0([0,T],\mathbb{R}^d)}>R\}}\left(\int_0^t m(s)(1+|\sigma_\nu(s)|)\mathrm{d}s\right)^p \mathrm{d}\boldsymbol{\eta}_\nu(y,\sigma_\nu)\\
&\leq 2^p\,\|m(\cdot)\|^{p/q}_{L^1([0,t])}\int_0^t m(s)\left(\int_{\{(y,\sigma_\nu)\text{ s.t. }\|\sigma_\nu(\cdot)\|_{C^0([0,T],\mathbb{R}^d)}>R\}}(1+|\sigma_\nu(s)|)^p\mathrm{d}\boldsymbol{\eta}_\nu(y,\sigma_\nu)\right)\mathrm{d}s.
\end{aligned}
\tag{D.8}
$$

Recall now that by (C.5) above, it also holds that

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

for $\boldsymbol{\eta}_\nu$-almost every (y,σ_ν), which together with (D.8) then yields

$$\begin{aligned}\int_{\{(y,\sigma_\nu)\text{ s.t. }\|\sigma_\nu(\cdot)\|_{C^0([0,T],\mathbb{R}^d)}>R\}} &\left(\int_0^t |v(s,\sigma_\nu(s))-w(s,\sigma_\nu(s))|\mathrm{d}s\right)^p \mathrm{d}\boldsymbol{\eta}_\nu(y,\sigma_\nu) \\ &\leq 2^p\,\|m(\cdot)\|^p_{L^1([0,t])}\int_{\{y\text{ s.t. }|y|\geq R/C_T-1\}}\Big(1+C_T(1+|y|)\Big)^p\mathrm{d}\nu^0(y),\end{aligned} \tag{D.9}$$

for all times $t\in[0,T]$. Therefore, by merging (D.7) and (D.9) in (D.5) while applying Hölder's inequality, we finally obtain that

$$\begin{aligned}W_p(\mu(t),\nu(t)) \leq C_p\Bigg(&W_p(\mu^0,\nu^0)+\int_0^t \|v(s)-w(s)\|_{L^\infty(B(0,R),\mathbb{R}^d;\,\nu(s))}\,\mathrm{d}s \\ &+2(1+C_T)\,\|m(\cdot)\|_{L^1([0,t])}\left(\int_{\{y\text{ s.t. }|y|\geq R/C_T-1\}}(1+|y|)^p\mathrm{d}\nu^0(y)\right)^{1/p}\Bigg) \\ &\times\exp\left(C'_p\,\|l(\cdot)\|^p_{L^1([0,t])}\right),\end{aligned}$$

for all times $t\in[0,T]$, which concludes the proof of our claim. □

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] G. Albi, L. Pareschi, and M. Zanella. Boltzmann type Control of Opinion Consensus through Leaders. *Proceedings of the Royal Society A.*, 372, 2014.

[4] L. Ambrosio. Transport Equation and Cauchy Problem for BV Vector Fields. *Inventiones Mathematicae*, 158(2):227–260, 2004.

[5] L. Ambrosio, M. Colombo, and A. Figalli. Existence and Uniqueness of Maximal Regular Flows with Non-smooth Vector Fields. *Archive for Rational Mechanics and Analysis*, 218(2):1043–1081, 2015.

[6] 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.

[7] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variations and Free Discontinuity Problems*. Oxford Mathematical Monographs, 2000.

[8] 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.

[9] 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.

[10] H. Attouch, G. Buttazzo, and G. Michaille. *Variational Analysis in Sobolev and BV Spaces*. Series on Optimization. SIAM-MPS, 2005.

[11] J.-P. Aubin and A. Cellina. *Differential Inclusions*. Springer-Verlag, 1984.

[12] J.-P. Aubin and H. Frankowska. *Set-Valued Analysis*. Modern Birkhäuser Classics. Birkhäuser Basel, 1990.

[13] J.-D. Benamou, G. Carlier, and M. Laborde. An Augmented Lagragian Approach to Wasserstein Gradient Flows and Applications. *ESAIM: Proceedings and Surveys*, 54:1–17, 2016.

[14] P. Bernard. Young Measures, Superposition and Transport. *Indiana Univ. Math. Journal*, 57(1):247–275, 2008.

[15] A.L. Bertozzi, T. Laurent, and J. Rosado. L^p Theory for the Multidimensional Aggregation Equation. *Communications on Pure and Applied Mathematics*, 64(1):45–83, 2011.

[16] S. Bianchini and P. Bonicatto. A Uniqueness Result for the Decomposition of Vector Fields in $\mathbb{R}^d$. *Inventiones Matematicae*, 220:255–393, 2020.

[17] B. Bonnet and H. Frankowska. Differential Inclusions in Wasserstein Spaces: The Cauchy-Lipschitz Framework. *Journal of Differential Equations*, 271:594–637, 2021.

[18] B. Bonnet and H. Frankowska. Necessary Optimality Conditions for Optimal Control Problems in Wasserstein Spaces. *Applied Mathematics and Optimization*, 84:1281–1330, 2021.

[19] 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.

[20] B. Bonnet and F. Rossi. The Pontryagin Maximum Principle in the Wasserstein Space. *Calculus of Variations and Partial Differential Equations*, 58:11, 2019.

[21] H. Brézis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Universitext. Springer, 2010.

[22] F. Camilli, G. Cavagnari, R. De Maio, and B. Piccoli. Superposition Principle and SChemes for Measure Differential Equations. 14(1):89–113, 2021.

[23] 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.

[24] P. Cannarsa and H. Frankowska. Some Characterizations of Optimal Trajectories in Control Theory. *SIAM Journal on Control and Optimization*, 29(6):1322–1347, 1991.

[25] P. Cannarsa and C. Sinestrari. *Semiconcave Functions, Hamilton-Jacobi Equations and Optimal Control*. Progresses in Nonlinear Differential Equations and Their Applications. Birkhaäuser Basel, 2004.

[26] 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 Univesrity Press, 2019.

[27] 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.

[28] 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.

[29] 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.

[30] 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.

[31] J.A. Carrillo, S. Martin, and V. Panferov. A New Interaction Potential for Swarming Models. *Physica D:Nonlinear Phenomena*, 260:112–126, 2013.

[32] 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.

[33] G. Cavagnari, A. Marigonda, K.T. Nguyen, and F.S. Priuli. Generalized Control Systems in the Space of Probability Measures. *Set-Valued and Var. Analysis*, 26(3):663–691, 2018.

[34] 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.

[35] 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.

[36] L. Cesari. *Optimization Theory and Applications. Problems with Ordinary Differential Equation*, volume 17 of *Science and Business Media*. Springer, 2012.

[37] F Clarke. *Functional Analysis, Calculus of Variations and Optimal Control*, volume 264. Springer, 2013.

[38] G. Crippa and C. De Lellis. Estimates and Regularity Results for the DiPerna-Lions Flow. *Journal Für die Reine und Angewandte Mathematik*, 2008(616), 2008.

[39] 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.

[40] E. Cristiani, B. Piccoli, and A. Tosin. *Multiscale Modeling of Pedestrian Dynamics*, volume 12. Springer, 2014.

[41] J. Destiel, W.M. Ruess, and W. Schachermayer. On Weak Compactness in $L^1(\mu, X)$. *Proceedings of the American Mathematical Society*, 118:447–453, 1993.

[42] R.L. Di Perna and Lions P.-L. Ordinary Differential Equations, Transport Theory and Sobolev Spaces. *Inventiones Mathematicae*, 98(3):511–548, 1989.

[43] J. Diestel. Remarks on Weak Compactness in $L^1(\mu, X)$. *Glasgow Mathematical Journal*, 18(1):87–91, 1977.

[44] J. Diestel and J.J.Jr Uhl. *Vector Measures*. Number 15 in Mathematical Surveys. American Mathematical Society, 1977.

[45] 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.

[46] 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.

[47] A.F. Filippov. *Differential Equations with Discontinuous Righthand Sides: Control Systems*, volume 18 of *Science and Business Media*. Springer, 2013.

[48] 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.

[49] 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.

[50] M. Fornasier, B. Piccoli, and F. Rossi. Mean-Field Sparse Optimal Control. *Philosophical Transactions of the Royal Society A.*, 372(20130400), 2014.

[51] M. Fornasier and F. Solombrino. Mean Field Optimal Control. *ESAIM COCV*, 20(4):1123–1152, 2014.

[52] 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.

[53] H. Frankowska. Optimal Trajectories Associated with a Solution of the Contingent Hamilton-Jacobi Equation. *Applied Mathematics and Optimization*, 19(1):291–311, 1989.

[54] H. Frankowska. A Priori Estimates for Operational Differential Inclusions. *Journal of Differential Equations*, 84:100–128, 1990.

[55] 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.

[56] H. Frankowska and N.P. Osmolovskii. Second-Order Necessary Conditions for a Strong Local Minimum in a Control Problem with General Control Constraints. *Applied Mathematics and Optimization*, 80(1):135–164, 2017.

[57] 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.

[58] H. Frankowska and N.P. Osmolovskii. Distance Estimates to Feasible Controls for Systems with Final Point Constraints and Second Order Necessary Optimality Conditions. *System and Control Letters*, 144:104770, 2020.

[59] 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.

[60] 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.

[61] 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.

[62] 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.

[63] John Horváth. *Topological Vector Spaces and Distributions*. Courier Corporation, 2012.

[64] 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.

[65] 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.

[66] 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.

[67] J.L. Kelley. *General Topology*, volume 27 of *Graduate Texts in Mathematics*. Springer, 1975.

[68] J-M. Lasry and P.-L. Lions. Mean Field Games. *Japanese Journal of Mathematics*, 2(1):229–260, 2007.

[69] 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.

[70] 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.

[71] F. Otto. The Geometry of Dissipative Equations : The Porous Medium Equation. *Communications in Partial Differential Equations*, 26:101–174, 2001.

[72] N.S. Papageorgiou. Random Fixed Point Theorems for Measurable Multifunctions in Banach Spaces. *Proceedings of the American Mathematical Society*, 97(3):507–514, 1986.

[73] G. Peyré. Entropic Approximation of Wasserstein Gradient Flows. *SIAM Journal on Imaging Sciences*, 8(4):2323–2351, 2015.

[74] B. Piccoli. Measure Differential Inclusions. In *2018 IEEE Conference on Decision and Control (CDC), Miami Beach, FL*, pages 1323–1328, 2018.

[75] B. Piccoli. Measure Differential Equations. *Archive for Rational Mechanics and Analysis*, 233:1289–1317, 2019.

[76] 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.

[77] 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.

[78] 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.

[79] W. Rudin. *Real and Complex Analysis.* Mathematical Series. McGraw-Hill International Editions, 1987.

[80] F. Santambrogio. *Optimal Transport for Applied Mathematicians*, volume 87. Birkhauser Basel, 2015.

[81] F. Santambrogio. {Euclidean, metric and Wasserstein} Gradient Flows : An Overview. *Bulletin of Mathematical Sciences*, 7:87–154, 2017.

[82] C.M. Topaz, A.L. Bertozzi, and M.A. Lewis. A Nonlocal Continuum Model for Biological Aggregation. *Bulletins of Mathematical Biology*, 68(7):1601–1623, 2006.

[83] A. Ülger. Weak Compactness in $L^1(\mu, X)$. *Proceedings of the American Mathematical Society*, 113:143–149, 1991.

[84] C. Villani. *Optimal Transport : Old and New.* Springer-Verlag, Berlin, 2009.

[85] R.B. Vinter. *Optimal Control.* Systems and Control: Foundations and Applications. Birkhauser Basel, 2000.

[86] D.H. Wagner. Survey of Measurable Selection Theorems. *SIAM Journal on Control and Optimization*, 15(5):859–903, 1977.

[87] S. Warner. The Topology of Compact Convergence on Continuous Function Spaces. *Duke Mathematical Journal*, 25(2):265–282, 1958.