HAL open science

On the Viability and Invariance of Proper Sets under Continuity Inclusions in Wasserstein Spaces

Benoît Bonnet-Weill, Hélène Frankowska

To cite this version:

Benoît Bonnet-Weill, Hélène Frankowska. On the Viability and Invariance of Proper Sets under Continuity Inclusions in Wasserstein Spaces. SIAM Journal on Mathematical Analysis, 2024, 56 (3), pp.2863-2914. 10.1137/23M1560410 . hal-04082985v2

HAL Id: hal-04082985

https://hal.science/hal-04082985v2

Submitted on 4 Jul 2024

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

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

On the Viability and Invariance of Proper Sets under Continuity Inclusions in Wasserstein Spaces

Benoît Bonnet-Weill* and Hélène Frankowska†

July 4, 2024

Abstract

In this article, we derive conditions for the existence of solutions to state-constrained continuity inclusions in Wasserstein spaces whose right-hand sides may be discontinuous in time. These latter are based on a fine investigation of the infinitesimal behaviour of the underlying reachable sets, through which we show that up to a negligible set of times, every admissible velocity of a continuity inclusion can be approximately realised as the metric derivative of a solution of the dynamics, and vice versa. Building on these results, we are able to establish necessary and sufficient geometric conditions for the viability and invariance of stationary and time-dependent constraint sets which involve a suitable notion of contingent cones in Wasserstein spaces, and presented in ascending order of generality. We then close the article by exhibiting two prototypical examples of constraints sets appearing in applications for which one can compute relevant subfamilies of contingent directions.

Keywords: Continuity Inclusions, Optimal Transport, Viability, Invariance, Wasserstein Geometry, Dynamics with Time Discontinuities.

MSC2020 Subject Classification: 28B20, 34G25, 46N20, 49Q22

Contents

1 Introduction **2**

2 Preliminaries **5**
2.1 Measure theory and optimal transport 6
2.2 Set-valued analysis and topological properties of $C^0(\mathbb{R}^d, \mathbb{R}^d)$ 8
2.3 Continuity equations and inclusions in Wasserstein spaces 12

3 Infinitesimal behaviour of the reachable sets **15**

4 Viability and invariance theorems for proper constraints sets **22**
4.1 The case of stationary constraints 23
4.2 The case of time-dependent constraints 27

5 Examples of constraints sets and computations of tangents **34**

Appendices **38**
A Proof of Proposition 2.3 38
B Proof of Proposition 2.8 39

*CNRS, LAAS, 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

Recent times have witnessed a surge of interest for the mathematical analysis of large-scale approximations of particle systems. During the past two decades, a series of seminal works concerned with the mean-field approximation of cooperative dynamics [39, 40, 64], the theory of mean-field games [36, 67, 71] and the sparse control of multiagent systems [33, 34, 56] have given rise to several research currents focusing on dynamical and variational problems whose aim is to describe the macroscopic behaviour of many-body systems. Amongst these latter, *mean-field control* is a research branch whose main object lies in designing scale-free and efficient control signals for large microscopic systems, by finely understanding the interplay that exists between discrete models and their continuous approximations. From a technical standpoint, these inquiries often boil down to studying variational problems in the space of probability measures, and are commonly approached using optimal transport techniques and Wasserstein geometry, in the spirit of the reference treatises [5, 81, 82]. Without aiming at full exhaustivity, we point the reader to the manuscripts [27, 42, 55, 57] for various existence and qualitative regularity results on deterministic mean-field optimal control problems, as well as to the following broad series of works dealing with optimality conditions, either in the form of a Pontryagin maximum principle [11, 18, 20, 21, 23, 26, 77, 78] or Hamilton-Jacobi-Bellman equations [13, 43, 45, 68]. We also mention the references [2, 35, 41, 76] that propose astute control strategies to stir collective systems towards specific asymptotic patterns, and finally [1, 21, 30, 31, 79] for various numerical methods in the context of mean-field optimal control.

Motivated by this blooming interest for variational problems in measure spaces, several research groups have been investigating relevant generalisations of the core concepts of *set-valued analysis* to the setting of mean-field control [13, 22, 23, 24, 28, 44, 46, 68], a lively trend that reached more recently other closely related topics, such as mean-field games [6, 32] and the study of sufficient conditions for the well-posedness of dynamics in measure spaces [47, 69]. It is de facto widely accepted that the language of correspondences, differential inclusions and generalised gradients provides, in many cases, a synthetic and powerful framework in which most problems stemming from the calculus of variations, games and control theory can be encompassed, as supported by the series of reference monographs [8, 9, 48, 83] tracing back to the nineteen eighties. For these reasons, the authors of the present manuscript introduced in [22, 28] a notion of differential inclusion in Wasserstein spaces that is tailored to the study of mean-field control problems. Therein, given a set-valued mapping $V : [0,T] \times \mathscr{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d, \mathbb{R}^d)$, we defined *continuity inclusions* as the set-valued dynamical systems

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

whose solutions are understood as the collection of all curves $\mu(\cdot) \in \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d))$ for which there exists a *measurable selection* $t \in [0,T] \mapsto v(t) \in V(t,\mu(t))$ such that the *continuity equation*

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

holds in the sense of distributions. While other notions of solutions for differential inclusions in measures spaces had already been proposed in some preexisting works, see e.g. [10, 44, 68], the approach we just described seemed more natural as well as necessary for several reasons. On the one hand, it remained coherent with the classical theory of set-valued dynamics, see e.g. [8] and [9, Chapter 11], as well as with the geometric interpretation of Wasserstein spaces developed in [5, 73] wherein the space $(\mathscr{P}_p(\mathbb{R}^d), W_p)$ is seen as a fiber bundle over which continuity equations essentially play the same role as ODEs for differential manifolds. On the other hand, it complied with one of the most important and desired features of differential inclusions, already formulated in the pioneering article [54], which stipulates that solutions of control systems should be in one-to-one correspondence with those of their set-valued counterparts. In [22, 28], based on this definition of differential inclusions, we proved analogues in the setting of Wasserstein spaces of the Filippov estimates and Peano existence theorem, as well as a relaxation principle and a compactness criterion for the underlying solution sets. These fundamental results are known to be extremely useful to investigate the fine properties of optimal control problems, both in the deterministic [61, 83] and stochastic [17] settings, while enjoying natural

generalisations to study e.g. evolution equations in Banach spaces [58, 60] or mutational dynamics in metric spaces [14, 59]. We also point to our recent works [23, 24], where we successfully applied such set-theoretic techniques to derive optimality conditions for optimal control problems in Wasserstein spaces, and to study certain qualitative properties of their solutions.

While permitting to handle a variety of relevant dynamical models in measure spaces, the afore-described framework does not cover systems with *state-constraints*, which appear nonetheless in a wide variety of applications ranging from game theory [37] to pedestrian dynamics [49] and traffic flows [38], and more recently to the dynamical formulations of deep neural networks [52]. Historically, the problem of ensuring that a differential inclusion admits trajectories that remain within a given set starting from any initial condition was coined *viability*, whereas that of ensuring that all such trajectories be viable was usually called *invariance*. The first results in this direction were established in [72] for differential equations, while their natural counterparts for differential inclusions with stationary right-hand sides later followed in [15] and [65]. To this day, the farthest-reaching viability theorems for classical differential inclusions can be found in [62, 63] – which inspired several of our contributions –, wherein the viability of general time-dependent constraint tubes is proven both in the Carathéodory and Cauchy-Lipschitz frameworks. Besides modelling incentives, viability and invariance results can be used to study the existence and uniqueness of viscosity solutions for several nonlinear partial differential equations, see e.g. [29, 63], as well as to investigate sufficient stability conditions for differential inclusions [7]. These viewpoints – like many others stemming from set-valued analysis – present the advantage of being readily transposable beyond the setting of finite-dimensional vector spaces, as illustrated e.g. by their recent applications to problems in metric spaces [10, 13, 14, 46].

In this article, we study the viability and invariance properties of constraints sets under the action of continuity inclusions in the Wasserstein spaces $(\mathscr{P}_p(\mathbb{R}^d), W_p)$ with $p \in (1, +\infty)$. Given a constraint tube $\mathcal{Q} : [0, T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ with proper images (see Definition 2.4 below), we provide necessary and sufficient conditions ensuring that either all or some of the solutions of the Cauchy problem

$$\begin{cases} \partial_t \mu(t) \in -\mathrm{div}_x\Big(V(t, \mu(t))\mu(t)\Big), \\ \mu(\tau) = \mu_\tau, \end{cases} \tag{1.1}$$

are such that $\mu(t) \in \mathcal{Q}(t)$ for all times $t \in [\tau, T]$, wherein $\tau \in [0, T]$ and $\mu_\tau \in \mathcal{Q}(\tau)$ are both arbitrary. These results, which are discussed in Section 4, rely on a careful analysis of the infinitesimal behaviour of the reachable sets of (1.1). This latter is the object of Section 3, and can be heuristically summarised as follows. In Theorem 3.1, we show that for each element $v_\tau \in V(\tau, \mu_\tau)$ taken at some adequate pair $(\tau, \mu_\tau) \in [0, T] \times \mathscr{P}_p(\mathbb{R}^d)$ and every $\varepsilon > 0$, there exists a solution $\mu_\varepsilon(\cdot)$ of (1.1) which satisfies

$$W_p\Big(\mu_\varepsilon(\tau + h^\varepsilon), (\mathrm{Id} + h^\varepsilon v_\tau)_\sharp \mu_\tau\Big) \leq \varepsilon h^\varepsilon$$

whenever $h^\varepsilon > 0$ is sufficiently small. In other words, each admissible velocity can be approximately realised – on a subset of times of full $\mathscr{L}^1$-measure and up to an arbitrarily small error – as the *right metric derivative* of an admissible curve. In Theorem 3.2, we complete this result by showing that for $\mathscr{L}^1$-almost every $\tau \in [0, T]$, each solution $\mu_\varepsilon(\cdot)$ of (1.1) and any sequence $h_i \to 0$, there exists $v_\tau^\varepsilon \in V(\tau, \mu_\tau)$ such that

$$W_p\Big(\mu_\varepsilon(\tau + h_{i_k}^\varepsilon), (\mathrm{Id} + h_{i_k}^\varepsilon v_\tau^\varepsilon)_\sharp \mu_\tau\Big) \leq \varepsilon |h_{i_k}^\varepsilon|,$$

along a subsequence $h_{i_k}^\varepsilon \to 0$. Stated otherwise, one can always find an admissible velocity which approximately represents – again on a subset of full $\mathscr{L}^1$-measure, up to an arbitrary small error, and a subsequence – the *metric derivative* of a solution of the Cauchy problem. It is worth noting that while these results are proven under Cauchy-Lipschitz assumptions in the present paper, we do believe that they remain valid under weaker Carathéodory-Peano hypotheses similar to those of [28, Section 3]. The main issue in proving so would be to replace the estimates on flow maps by their counterparts for measures concentrated on characteristic curves, whose existence is ensured by the famed superposition principle of Ambrosio, see e.g. [3, Theorem 3.4].

Let us now discuss more in depth the main contributions of this article, which are the necessary and sufficient conditions ensuring that a constraints tube $\mathcal{Q} : [0, T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ is either viable or

invariant under the dynamics of $V : [0,T] \times \mathscr{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d, \mathbb{R}^d)$ for $p \in (1, +\infty)$. For the sake of clarity, we start our investigation thereof by considering stationary constraints sets. In this context, the relevant geometric objects allowing to study viability and invariance properties are the *metric contingent cones* to $\mathcal{Q}$, defined by

$$T_{\mathcal{Q}}(\nu) := \left\{ \xi \in \mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d, \nu) \text{ s.t. } \liminf_{h \to 0^+} \tfrac{1}{h} \mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)} \Big((\mathrm{Id} + h\xi)_\sharp \nu \, ; \mathcal{Q} \Big) = 0 \right\}$$

for each $\nu \in \mathcal{Q}$, where $\mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d; \nu)$ stands for the seminormed space of Borel measurable and p-integrable maps against ν, and whose expression is akin to that recently introduced in [13]. Building on this notion, we show in Theorem 4.4 that if the geometric compatibility condition

$$V(t, \nu) \cap \overline{\mathrm{co}}\, T_{\mathcal{Q}}(\nu) \neq \emptyset \tag{1.2}$$

is satisfied for $\mathscr{L}^1$-almost every $t \in [0,T]$ and each $\nu \in \mathcal{Q}$, where "$\overline{\mathrm{co}}$" denotes the *closed convex hull* of a set, then $\mathcal{Q}$ is viable. That is, there exists then for every $(\tau, \mu_\tau) \in [0,T] \times \mathcal{Q}$ a solution $\mu(\cdot)$ of the Cauchy problem (1.1) satisfying $\mu(t) \in \mathcal{Q}$ for all times $t \in [\tau, T]$. Reciprocally, we prove in Theorem 4.6 that if $\mathcal{Q}$ is viable, then it necessarily holds that

$$V(t, \nu) \cap T_{\mathcal{Q}}(\nu) \neq \emptyset \tag{1.3}$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and each $\nu \in \mathcal{Q}$. We point the interested reader to our recent work [25], where we leveraged a weaker version of these viability theorems to prove the existence of exponentially stable solutions to a class of continuity inclusions via the second method of Lyapunov, see also the second example in Section 5 below. In Theorem 4.7, we subsequently show that the stronger condition

$$V(t, \nu) \subset \overline{\mathrm{co}}\, T_{\mathcal{Q}}(\nu), \tag{1.4}$$

which is assumed to hold for $\mathscr{L}^1$-almost every $t \in [0,T]$ and each $\nu \in \mathcal{Q}$, is equivalent to the invariance of $\mathcal{Q}$ under the dynamics of (1.1), namely to the fact that $\mu(t) \in \mathcal{Q}$ for all times $t \in [\tau, T]$ and every admissible curve $\mu(\cdot)$ starting from $\mu_\tau \in \mathcal{Q}$ at time $\tau \in [0,T]$.

We then turn our attention to the more involved scenario in which the constraints are allowed to be time-dependent. In Theorem 4.8, we start by showing that regardless of its regularity, if the tube $\mathcal{Q} : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ is viable for (1.1), namely if for any $\tau \in [0,T]$ and each $\mu_\tau \in \mathcal{Q}(\tau)$, there exists a solution of (1.1) satisfying $\mu(t) \in \mathcal{Q}(t)$ for all time $t \in [\tau, T]$, then necessarily

$$(\{1\} \times V(t, \nu)) \cap T_{\mathrm{Graph}(\mathcal{Q})}(t, \nu) \neq \emptyset \tag{1.5}$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and all $\nu \in \mathcal{Q}(t)$, where

$$\mathrm{Graph}(\mathcal{Q}) := \Big\{ (t, \nu) \in [0,T] \times \mathscr{P}_p(\mathbb{R}^d) \text{ s.t. } \nu \in \mathcal{Q}(t) \Big\}$$

denotes the *graph* of $\mathcal{Q} : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$. In addition, it follows in this context that $\mathcal{Q} : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ is actually *left absolutely continuous* (see Definition 2.7 below), as it inherits some of the regularity properties of the reachable sets of (1.1). Similarly, if one posits that $\mathcal{Q} : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ is invariant for (1.1), that is if $\mu(t) \in \mathcal{Q}(t)$ for all times $t \in [\tau, T]$ along every solution of (1.1) starting from some $\mu_\tau \in \mathcal{Q}(\tau)$ at time $\tau \in [0,T]$, then it must hold that

$$(\{1\} \times V(t, \nu)) \subset T_{\mathrm{Graph}(\mathcal{Q})}(t, \nu) \tag{1.6}$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and all $\nu \in \mathcal{Q}(t)$.

Unlike the aforedescribed necessary implications, sufficient viability and invariance conditions for time-dependent constraints call for separate analyses depending on the regularity of the tube $\mathcal{Q} : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$. When the latter is *absolutely continuous* (see Definition 2.7 below), we prove in Theorem 4.10 that $\mathcal{Q} : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ is viable for (1.1) whenever the geometric condition

$$(\{1\} \times V(t, \nu)) \cap \overline{\mathrm{co}}\, T_{\mathrm{Graph}(\mathcal{Q})}(t, \nu) \neq \emptyset \tag{1.7}$$

holds for $\mathscr{L}^1$-almost every $t \in [0,T]$ and all $\nu \in \mathcal{Q}(t)$. We stress that in this context, the convexification of the contingent directions requires much more care than in the stationary case, as one must combine tangent velocities corresponding to possibly different time instants. Similarly to what precedes, we then show in Theorem 4.11 below that $\mathcal{Q} : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ is invariant for (1.1) provided that

$$V(t,\nu) \subset \overline{\mathrm{co}}\, T_{\mathrm{Graph}(\mathcal{Q})}(t,\nu) \tag{1.8}$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and all $\nu \in \mathcal{Q}(t)$. Finally in Theorem 4.12, we address sufficient viability and invariance conditions when the tube $\mathcal{Q} : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ is merely left absolutely continuous. This regularity framework – which, as already illustrated above, appears naturally when studying continuity inclusions with state-constraints – is very similar to the one we previously discussed, with the added subtlety that in this context one cannot convexify the contingent directions anymore. Thence, the sufficient condition which ensures that $\mathcal{Q} : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ is viable for (1.1) becomes

$$(\{1\} \times V(t,\nu)) \cap T_{\mathrm{Graph}(\mathcal{Q})}(t,\nu) \neq \emptyset \tag{1.9}$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and all $\nu \in \mathcal{Q}(t)$. Similarly to the stationary and absolutely continuous cases, one can show that $\mathcal{Q} : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ is invariant for (1.1) if the geometric condition

$$(\{1\} \times V(t,\nu)) \subset T_{\mathrm{Graph}(\mathcal{Q})}(t,\nu) \tag{1.10}$$

holds for $\mathscr{L}^1$-almost every $t \in [0,T]$ and each $\nu \in \mathcal{Q}(t)$. At this stage, it is worth noting that while conditions (1.2) to (1.9) involve the whole contingent cone to the constraints, it is sufficient in practice to test their validity only on some nice subsets of tangent directions which are easy to compute. This fact is expounded in Section 5, in which we exhibit collections of tangents for two simple, yet frequently encountered families of constraints sets, defined respectively in terms of support inclusions or as lifted epigraphs. In the first situation, we display certain adjacent directions which are amenable to computations, while in the second one, we are able to fully characterise a relevant subset of the contingent cone.

In terms of bibliographical positioning, this work can be seen as a far-reaching extension of [13] by the second author, in which viability and invariance properties are established as a means to prove the well-posedness of general Hamilton-Jacobi-Bellman equations for optimal control problems in Wasserstein spaces, under more restrictive regularity assumptions on the data. We also point to the independent works [10, 12, 46] which focus on the study of viability properties for another class of set-valued dynamics in Wasserstein spaces, which we already alluded earlier. The main differences between the latter notion and the one considered in the present article and the related works of both authors is thoroughly discussed in [22].

The paper is organised as follows. In Section 2, we start by recalling a list of preliminary material pertaining to optimal transport, set-valued analysis, and continuity inclusions. Subsequently, we discuss in Section 3 the infinitesimal behaviour of reachable sets to continuity inclusions, which constitute novel contributions to the theory on which the main results of Section 4 crucially rely. This latter section is then split into two parts in which we discuss viability and invariance results for continuity inclusions, starting with the case of stationary constraints in Section 4.1. We then approach time-dependent constraints in Section 4.2, wherein the necessary and sufficient conditions for viability and invariance are exposed separately for absolutely continuous and left absolutely continuous constraints. Finally, we show a couple of relevant examples of constraints sets and compute some of their tangent directions in Section 5, and close the paper by an appendix that contains the proofs of two intermediate technical results.

2 Preliminaries

In this section, we fix the notations that will be used throughout the manuscript, and list a series of prerequisites of optimal transport theory, set-valued analysis and Wasserstein geometry.

2.1 Measure theory and optimal transport

In this first preliminary section, we recollect common notions of measure theory and optimal transport, for which we refer the reader to the monographs [4, 53] and [5, 81, 82] respectively.

Function spaces and measure theory. Given two complete separable metric spaces $(X, \mathsf{d}_X(\cdot,\cdot))$ and $(Y, \mathsf{d}_Y(\cdot,\cdot))$, we denote by $C^0(X,Y)$ the space of continuous functions from X into Y, and likewise by $C^0_b(X,Y)$ the subspace of continuous and bounded functions. In this context, we will use the notation $\mathrm{Lip}(\varphi\,;\Omega) \in \mathbb{R}_+ \cup \{+\infty\}$ for the Lipschitz constant of a map $\varphi : X \to Y$ over some subset $\Omega \subset X$. In the particular case where $(X, \mathsf{d}_X(\cdot,\cdot)) = (I, |\cdot|)$ for a closed interval $I \subset \mathbb{R}$, we shall write $\mathrm{AC}(I,Y)$ for the collection of absolutely continuous curves valued in Y. We shall also denote by $C^\infty_c(\mathbb{R}^d, \mathbb{R}^m)$ the space of infinitely differentiable functions with compact support from $\mathbb{R}^d$ into $\mathbb{R}^m$ for some $d, m \geq 1$.

In what follows, we let $\mathscr{P}(\mathbb{R}^d)$ be the space of Borel probability measures defined over $(\mathbb{R}^d, |\cdot|)$. Recalling that the latter is a subset of the topological dual of $C^0_b(\mathbb{R}^d, \mathbb{R})$, it can be endowed with the usual weak-* or *narrow topology*, which is the coarsest topology such that

$$\mu \in \mathscr{P}(\mathbb{R}^d) \mapsto \int_{\mathbb{R}^d} \varphi(x)\mathrm{d}\mu(x) \in \mathbb{R}$$

defines a continuous mapping for every $\varphi \in C^0_b(\mathbb{R}^d, \mathbb{R})$. In this context, given $\mu \in \mathscr{P}(\mathbb{R}^d)$ and some $p \in [1, +\infty)$, the notation $(\mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d; \mu), \|\cdot\|_{\mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\mu)})$ will refer to the seminormed vector space of Borel maps from $\mathbb{R}^d$ into itself which are p-integrable with respect to μ (see e.g. [16, Chapter 2.4]). We will also denote by $\mathscr{L}^1$ the standard 1-dimensional Lebesgue measure, and given a closed interval $I \subset \mathbb{R}$ along with a separable Banach space $(X, \|\cdot\|_X)$, we let $(L^1(I,X), \|\cdot\|_{L^1(I)})$ stand for the Banach space of (all equivalence classes of) maps which are $\mathscr{L}^1$-measurable and integrable in the sense of Bochner, see for instance [50, Chapter II].

In the following definition, given some interval $I \subset \mathbb{R}$, we recall the classical notions of *one-sided density points* of an $\mathscr{L}^1$-measurable subset $\mathscr{A} \subset I$, along with that of *one-sided Lebesgue points* for a Lebesgue integrable map defined over I. It is a well-known result in measure theory that these sets both have full $\mathscr{L}^1$-measure in $\mathscr{A}$ and I, respectively.

Definition 2.1 (Lebesgue and density points)**.** *Given a Lebesgue measurable set $\mathscr{A} \subset I$, its* one-sided density points *are defined as the elements $\tau \in \mathscr{A}$ satisfying*

$$\frac{\mathscr{L}^1([\tau, \tau+h] \cap \mathscr{A})}{|h|} \underset{h\to 0}{\longrightarrow} 1.$$

Similarly, given a map $f \in L^1(I,X)$, we denote by $\mathscr{T}_f \subset I$ the subset of its one-sided Lebesgue points*, which are the elements $\tau \in I$ at which*

$$\frac{1}{h}\int_\tau^{\tau+h} \|f(t) - f(\tau)\|_X \,\mathrm{d}t \underset{h\to 0}{\longrightarrow} 0.$$

For any real number $p \in [1, +\infty)$, we denote by $\mathscr{P}_p(\mathbb{R}^d)$ the subset of Borel probability measures whose *moment of order* p, defined by

$$\mathcal{M}_p(\mu) := \left(\int_{\mathbb{R}^d} |x|^p \mathrm{d}\mu(x)\right)^{1/p},$$

is finite. In what follows, we write $f_\sharp\mu \in \mathscr{P}(\mathbb{R}^d)$ for the *image measure* of an element $\mu \in \mathscr{P}(\mathbb{R}^d)$ through a Borel map $f : \mathbb{R}^d \to \mathbb{R}^d$, which is uniquely characterised by the identity

$$\int_{\mathbb{R}^d} \varphi(x)\mathrm{d}(f_\sharp\mu)(x) = \int_{\mathbb{R}^d} \varphi \circ f(x)\mathrm{d}\mu(x)$$

satisfied for every bounded Borel mapping $\varphi : \mathbb{R}^d \to \mathbb{R}^d$, wherein "$\circ$" stands for the composition operation between functions. The set of *transport plans* between two measures $\mu, \nu \in \mathscr{P}(\mathbb{R}^d)$ is then defined by

$$\Gamma(\mu,\nu) := \left\{\gamma \in \mathscr{P}_p(\mathbb{R}^{2d}) \text{ s.t. } \pi^1_\sharp\gamma = \mu \text{ and } \pi^2_\sharp\gamma = \nu\right\},$$

where $\pi^1, \pi^2 : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ represent the projections onto the first and second factors respectively. In this context, the *Wasserstein distance of order* p between μ and ν is the quantity given by

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

and we henceforth denote by $\Gamma_o(\mu,\nu)$ the set of p-optimal transport plans for which the minimum is reached. Owing to the definition of this distance in the form of an infimum, it can be checked straightforwardly following e.g. [5, Chapter 7] that

$$W_p(\zeta_\sharp\mu, \xi_\sharp\mu) \leq \|\xi-\zeta\|_{\mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\,\mu)} \tag{2.1}$$

for any $\mu \in \mathscr{P}_p(\mathbb{R}^d)$ and each pair of elements $\zeta, \xi \in \mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\mu)$. In the following proposition, we recall elementary facts regarding the topology of Wasserstein spaces, see e.g. [5, Proposition 7.1.5].

Proposition 2.2 (Topological properties of Wasserstein spaces)**.** *The metric spaces* $(\mathscr{P}_p(\mathbb{R}^d), W_p)$ *are complete and separable, and their topology is stronger than the narrow topology. Moreover, a set* $\mathcal{K} \subset \mathscr{P}_p(\mathbb{R}^d)$ *is* relatively compact *for the* W_p*-metric if and only if*

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

that is, if and only if it is p*-uniformly integrable.*

Wasserstein geometry. In addition to their convenient topological properties, the Wasserstein spaces can be endowed with a geometric structure that greatly resembles that of a Riemannian manifold when $p = 2$. For a general $p \in [1,+\infty)$, it is discussed in depth throughout [5, Chapter 8] that $(\mathscr{P}_p(\mathbb{R}^d), W_p)$ can be seen as a bundle, whose fibers are the closed cones

$$\mathrm{Tan}_\mu \mathscr{P}_p(\mathbb{R}^d) := \overline{\Big\{ j_q(\nabla\varphi) \text{ s.t. } \varphi \in C_c^\infty(\mathbb{R}^d,\mathbb{R}) \Big\}}^{\mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\,\mu)}$$

defined at each $\mu \in \mathscr{P}_p(\mathbb{R}^d)$. Therein, $q \in (1,+\infty]$ stands for the conjugate exponent of $p \in [1,+\infty)$, and $j_q : \mathcal{L}^q(\mathbb{R}^d,\mathbb{R}^d;\mu) \to \mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\mu)$ refers to the so-called *duality map*, which is given by

$$j_q(\xi) := \begin{cases} 0 & \text{if } \xi = 0, \\ |\xi|^{q-2}\,\xi & \text{otherwise} \end{cases} \tag{2.2}$$

for each $\xi \in \mathcal{L}^q(\mathbb{R}^d,\mathbb{R}^d;\mu)$. In the following proposition, we provide an adhoc version of the joint directional superdifferentiability inequalities satisfied by the p-Wasserstein distance whenever $p > 1$. We point the reader to [5, Theorem 10.2.2] for the general case.

Proposition 2.3 (Joint directional superdifferentiability of the Wasserstein distance)**.** *Fix an element* $p \in (1,+\infty)$ *and* $\mu,\nu \in \mathscr{P}_p(\mathbb{R}^d)$*. Then for each* $\xi \in \mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\mu)$ *and* $\zeta \in \mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\nu)$*, it holds that*

$$\tfrac{1}{p}W_p^p\Big((\mathrm{Id}+h\zeta)_\sharp\mu, (\mathrm{Id}+h\xi)_\sharp\nu\Big) - \tfrac{1}{p}W_p^p(\mu,\nu) \leq h\int_{\mathbb{R}^{2d}} \langle \zeta(x)-\xi(y), j_p(x-y)\rangle \mathrm{d}\gamma(x,y) + r_p(h,\zeta,\xi)$$

for any $h \in \mathbb{R}$ *and every* $\gamma \in \Gamma_o(\mu,\nu)$*, where the remainder term* $r_p(h,\zeta,\xi)$ *is given explicitly by*

$$\begin{aligned} r_p(h,\zeta,\xi) := (p-1)\Big(W_p(\mu,\nu) + |h|\Big(\|\zeta\|_{\mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\,\mu)} + \|\xi\|_{\mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\,\nu)}\Big)\Big)^{p-2} \\ \times \Big(\|\zeta\|^2_{\mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\,\mu)} + \|\xi\|^2_{\mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\,\nu)}\Big)|h|^2 \end{aligned} \tag{2.3}$$

when $p \in [2,+\infty)$*, and by*

$$r_p(h,\zeta,\xi) := \tfrac{2}{p-1}\Big(\|\zeta\|^p_{\mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\,\mu)} + \|\xi\|^p_{\mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\,\nu)}\Big)|h|^p \tag{2.4}$$

when $p \in (1,2]$. In particular, there exists a constant $C_p > 0$ which only depends on the magnitudes of $p, \mathcal{M}_p(\mu), \mathcal{M}_p(\nu), \|\zeta\|_{\mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\mu)}$ and $\|\xi\|_{\mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\nu)}$, such that

$$r_p(h,\zeta,\xi) \leq C_p |h|^{\min\{p,2\}}$$

whenever $h \in (0,1]$. Moreover, note that $r_p(h,\zeta,\xi) = 0$ if $\zeta = \xi = 0$.

Proof. For the sake of readability, the proof of this result is deferred to Appendix A. □

2.2 Set-valued analysis and topological properties of $C^0(\mathbb{R}^d,\mathbb{R}^d)$

In this second preliminary section, we recollect pivotal concepts of set-valued analysis and discuss some of the topological features of the space $C^0(\mathbb{R}^d,\mathbb{R}^d)$ that will prove useful in the sequel. We point the reader to the reference treatises [8, 9] for the former topics, and to [70, 80] for the latter.

Elementary notations. Given a metric space $(X, \mathsf{d}_X(\cdot,\cdot))$, we will denote the closed ball of radius $R > 0$ centered at $x \in X$ by $\mathbb{B}_X(x,R) := \{x' \in X \text{ s.t. } \mathsf{d}_X(x,x') \leq R\}$, and write

$$\mathrm{dist}_X(\mathcal{Q};\mathcal{Q}') := \inf_{(x,x')\in\mathcal{Q}\times\mathcal{Q}'} \mathsf{d}_X(x,x')$$

for the usual distance between two closed sets $\mathcal{Q}, \mathcal{Q}' \subset X$. Throughout the article, we will also work with the *Hausdorff metric*, which is defined by

$$\mathsf{d}_{\mathcal{H}}(\mathcal{K},\mathcal{K}') := \inf\Big\{\varepsilon > 0 \text{ s.t. } \mathcal{K} \subset \mathbb{B}_X(\mathcal{K}',\varepsilon) \text{ and } \mathcal{K}' \subset \mathbb{B}_X(\mathcal{K},\varepsilon)\Big\} \tag{2.5}$$

for each pair of compact sets $\mathcal{K}, \mathcal{K}' \subset X$, where we used the condensed notation

$$\mathbb{B}_X(\mathcal{K},\varepsilon) := \bigcup_{x\in\mathcal{K}} \mathbb{B}_X(x,\varepsilon).$$

In our subsequent developments, we will denote by $\mathrm{int}(\mathcal{Q})$ and $\partial\mathcal{Q} := \mathcal{Q} \setminus \mathrm{int}(\mathcal{Q})$ the interior and topological boundary of a set $\mathcal{Q}$. In the particular case in which $(X, \mathsf{d}_X(\cdot,\cdot))$ possesses a linear structure – e.g. when it is a Banach or a Fréchet space, see for instance [66] –, we define the *closed convex hull* of any subset $\mathcal{B} \subset X$ as

$$\overline{\mathrm{co}}(\mathcal{B}) = \overline{\mathrm{co}(\mathcal{B})}^X := \overline{\bigcup_{N\geq 1}\Big\{\sum_{j=1}^N \alpha_j b_j \text{ s.t. } b_j \in \mathcal{B},\ \alpha_j \geq 0 \text{ for } j \in \{1,\dots,N\} \text{ and } \sum_{j=1}^N \alpha_j = 1\Big\}}^X,$$

wherein "$\overline{\bullet}^X$" stands for the sequential closure with respect to $\mathsf{d}_X(\cdot,\cdot)$. We finally recall the definition of the so-called *proper subsets* of a metric space.

Definition 2.4 (Proper subsets of a metric space)**.** *A closed set $\mathcal{Q} \subset X$ is* proper *provided that $\mathcal{Q} \cap \mathbb{B}_X(x,R)$ is compact for each $x \in X$ and every $R > 0$.*

Note that in the previous definition, one could alternatively require that $\mathcal{Q} \cap \mathbb{B}_X(x,R)$ be compact for each $x \in \mathcal{Q}$ instead of $x \in X$.

Set-valued analysis. We recall that a *set-valued map* – or *correspondence* – between two metric spaces $(X, \mathsf{d}_X(\cdot,\cdot))$ and $(Y, \mathsf{d}_Y(\cdot,\cdot))$ is an application $\mathcal{F} : X \rightrightarrows Y$ whose images are subsets of Y, namely $\mathcal{F}(x) \subset Y$ for all $x \in X$. In this context, the *graph* of $\mathcal{F}$ is the subset of $X \times Y$ defined by

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

In the coming definitions, we recall the main regularity notions for set-valued mappings with values in metric spaces, starting with those of *continuity* and *Lipschitz regularity*.

Definition 2.5 (Continuity of set-valued maps)**.** *A correspondence $\mathcal{F} : X \rightrightarrows Y$ is said to be* continuous *at $x \in X$ if both the following conditions hold.*

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

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

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

(*ii*) *$\mathcal{F}$ is* upper-semicontinuous *at x, i.e. for any $\varepsilon > 0$, there exists $\delta > 0$ such that*

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

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

Definition 2.6 (Lipschitz continuity of set-valued maps)**.** *A correspondence $\mathcal{F} : X \rightrightarrows Y$ is said to be* Lipschitz continuous *with constant $L > 0$ provided that*

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

for all $x, x' \in X$. When $\mathcal{F} : X \rightrightarrows Y$ has compact images, this is equivalent to requiring that

$$\mathsf{d}_{\mathcal{H}}(\mathcal{F}(x), \mathcal{F}(x')) \leq L\,\mathsf{d}_X(x, x')$$

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

As for functions defined over the real line taking values in a metric space, it is possible to formulate several relevant notions of *absolute continuity* for set-valued maps. Some of them involve the quantity

$$\Delta_{x,R}(\mathcal{Q}\,;\mathcal{Q}') := \inf\Big\{\varepsilon > 0 \text{ s.t. } \mathcal{Q} \cap \mathbb{B}_X(x, R) \subset \mathbb{B}_X(\mathcal{Q}', \varepsilon)\Big\} \in \mathbb{R}_+ \cup \{+\infty\},$$

defined for each $x \in X$, $R > 0$ and every pair of nonempty closed sets $\mathcal{Q}, \mathcal{Q}' \subset X$, which can be seen as a sort of asymmetric and localised version of the Hausdorff distance recalled in (2.5).

Definition 2.7 (Notions of absolute continuity for set-valued mappings)**.** *We say that a correspondence $\mathcal{F} : I \rightrightarrows X$ with nonempty closed images is* absolutely continuous *if for every $x \in X$ and each $R > 0$, there exists a map $m_{x,R}(\cdot) \in L^1(I, \mathbb{R}_+)$ such that*

$$\max\Big\{\Delta_{x,R}\big(\mathcal{F}(\tau)\,;\mathcal{F}(t)\big)\,,\,\Delta_{x,R}\big(\mathcal{F}(t)\,;\mathcal{F}(\tau)\big)\Big\} \leq \int_\tau^t m_{x,R}(s)\mathrm{d}s$$

for all times $\tau, t \in I$ satisfying $\tau \leq t$. Analogously, we say that $\mathcal{F} : I \rightrightarrows X$ is left absolutely continuous *if only the one-sided inequality*

$$\Delta_{x,R}\big(\mathcal{F}(\tau)\,;\mathcal{F}(t)\big) \leq \int_\tau^t m_{x,R}(s)\mathrm{d}s$$

holds. In the case where $\mathcal{F} : I \rightrightarrows X$ has compact images, we say that it is absolutely continuous in the Hausdorff metric *if there exists a map $m_{\mathcal{F}}(\cdot) \in L^1(I, \mathbb{R}_+)$ such that*

$$\mathsf{d}_{\mathcal{H}}(\mathcal{F}(\tau), \mathcal{F}(t)) \leq \int_\tau^t m_{\mathcal{F}}(s)\mathrm{d}s$$

for all times $\tau, t \in I$ such that $\tau \leq t$.

These notions of absolute continuity allow us to establish the following regularity statements on the distance between set-valued maps, which will prove crucial in some of our subsequent developments.

Proposition 2.8 (Regularity of the distance between set-valued maps)**.** *Let $\mathcal{K} : I \rightrightarrows X$ be a set-valued map with nonempty compact images that is absolutely continuous in the Hausdorff metric and $\mathcal{Q} : I \rightrightarrows X$ be an absolutely continuous set-valued map with nonempty closed images. Then, the map*

$$g : t \in I \mapsto \mathrm{dist}_X(\mathcal{K}(t)\,;\mathcal{Q}(t)) \in \mathbb{R}_+$$

is absolutely continuous. In the case where $\mathcal{Q} : [0,T] \rightrightarrows X$ is only left absolutely continuous, then the set-valued mapping

$$\mathscr{E} : t \in I \rightrightarrows \Big\{\alpha \in \mathbb{R}_+ \ \text{ s.t. } \alpha = g(t) + r \text{ for some } r \geq 0\Big\}$$

is left absolutely continuous as well.

Proof. Being somewhat long and technical, the proof of this statement is deferred to Appendix B. □

Analogously to the classical notions of regularity exposed hereinabove, it is possible to generalise the concept of *measurability* to set-valued mappings, as highlighted by the following definition.

Definition 2.9 (Measurability of set-valued maps)**.** *A set-valued map $\mathcal{F} : I \rightrightarrows X$ is said to be* $\mathscr{L}^1$-measurable *if the preimages*

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

are $\mathscr{L}^1$-measurable for each open set $\mathcal{O} \subset X$. Moreover, we say that an $\mathscr{L}^1$-measurable map $f : I \to X$ is a measurable selection *of $\mathcal{F} : I \rightrightarrows X$ if $f(t) \in \mathcal{F}(t)$ for $\mathscr{L}^1$-almost every $t \in I$.*

In the following theorem, we recall an instrumental result of set-valued analysis excerpted from [9, Theorem 8.1.3], which asserts that measurable correspondences with nonempty closed images always admit measurable selections.

Theorem 2.10 (Existence of measurable selections)**.** *Suppose that $(X, \mathsf{d}_X(\cdot,\cdot))$ is a complete separable metric space. Then every $\mathscr{L}^1$-measurable set-valued map $\mathcal{F} : I \rightrightarrows X$ with nonempty closed images admits a measurable selection.*

We end this primer in set-valued analysis by recollecting a fine adaptation of the Scorza-Dragoni theorem for set-valued mappings between metric spaces, for which we refer to [19, Theorem 1].

Theorem 2.11 (Scorza-Dragoni property for set-valued mappings)**.** *Suppose that $(X, \mathsf{d}_X(\cdot,\cdot))$ and $(Y, \mathsf{d}_Y(\cdot,\cdot))$ are complete separable metric spaces, and let $\mathcal{F} : I \times X \rightrightarrows Y$ be a set-valued map with nonempty closed images, such that $t \in I \rightrightarrows \mathcal{F}(t,x)$ is $\mathscr{L}^1$-measurable for all $x \in X$ and $x \rightrightarrows \mathcal{F}(t,x)$ is continuous for $\mathscr{L}^1$-almost every $t \in I$.*

Then for every $\varepsilon > 0$, there exists a compact set $\mathscr{A}_\varepsilon \subset I$ such that $\mathscr{L}^1(I \setminus \mathscr{A}_\varepsilon) < \varepsilon$, and for which the following holds.

(i) The restricted set-valued mapping $\mathcal{F} : \mathscr{A}_\varepsilon \times X \rightrightarrows Y$ is lower-semicontinuous.

(ii) The graph of the restricted set-valued mapping

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

is closed in $\mathscr{A}_\varepsilon \times X \times Y$.

Topological structures over the space of continuous functions. Throughout the coming paragraphs, we recall some useful topological properties of the space $C^0(\mathbb{R}^d, \mathbb{R}^d)$. In what follows, we let

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

be the supremum extended-distance between a pair of elements $v, w \in C^0(\mathbb{R}^d, \mathbb{R}^d)$. While this latter is useful to control the global discrepancy between two continuous functions – which may be equal to $+\infty$ –, the topology that it induces is not separable and thus ill-adapted to the application of measurable selection theorems. For this reason, we will systematically endow the space $C^0(\mathbb{R}^d, \mathbb{R}^d)$ with the topology of *local uniform convergence*, whose definition is recalled hereinbelow.

Definition 2.12 (Topology of local uniform convergence)**.** *A sequence of maps* $(v_n) \subset C^0(\mathbb{R}^d, \mathbb{R}^d)$ *converges* locally uniformly – *or* 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 each compact set $K \subset \mathbb{R}^d$*. This notion of convergence endows* $C^0(\mathbb{R}^d, \mathbb{R}^d)$ *with the structure of a separable Fréchet space, whose topology is induced by the translation invariant metric*

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

that is well-defined and finite for any $v, w \in C^0(\mathbb{R}^d, \mathbb{R}^d)$*.*

Amongst its interesting properties, the topology of local uniform convergence enjoys a very explicit and amenable characterisation of compactness, for which we refer e.g. to [70, Chapter 7, Theorem 18].

Theorem 2.13 (Ascoli-Arzelà compactness criterion)**.** *A closed set* $V \subset C^0(\mathbb{R}^d, \mathbb{R}^d)$ *is compact for the topology induced by* $\mathsf{d}_{cc}(\cdot, \cdot)$ if and only if *its elements are locally uniformly equicontinuous and if, for every* $x \in \mathbb{R}^d$*, there exists a compact set* $K_x \subset \mathbb{R}^d$ *such that* $v(x) \in K_x$ *for each* $v \in V$*.*

We recall below a fact whose proof can be found in [74], which establishes a one-to-one correspondence between $\mathscr{L}^1$-measurable maps $t \in I \mapsto v(t) \in C^0(\mathbb{R}^d, \mathbb{R}^d)$ and Carathéodory vector fields. We recall that a map $v : I \times \mathbb{R}^d \to \mathbb{R}^d$ is Carathéodory if $t \in I \mapsto v(t, x)$ is $\mathscr{L}^1$-measurable for all $x \in \mathbb{R}^d$ and $x \in \mathbb{R}^d \mapsto v(t, x)$ is continuous for $\mathscr{L}^1$-almost every $t \in [0, T]$.

Lemma 2.14 (Measurable selections in $C^0(\mathbb{R}^d, \mathbb{R}^d)$ and Carathéodory vector fields)**.** *A vector field* $(t, x) \in I \times \mathbb{R}^d \mapsto v(t, x) \in \mathbb{R}^d$ *is Carathéodory* if and only if *its functional lift* $t \in I \mapsto v(t) \in C^0(\mathbb{R}^d, \mathbb{R}^d)$ *is* $\mathscr{L}^1$*-measurable with respect to the topology induced by* $\mathsf{d}_{cc}(\cdot, \cdot)$*.*

Lastly, we prove a technical result which states that for sequences of sublinear continuous functions, the convergence with respect to $\mathsf{d}_{cc}(\cdot, \cdot)$ implies the convergence in $\mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d; \mu)$ for every $p \in [1, +\infty)$ and each $\mu \in \mathscr{P}_p(\mathbb{R}^d)$.

Lemma 2.15 (Link between local uniform and Lebesgue convergences)**.** *Let* $(v_n) \subset C^0(\mathbb{R}^d, \mathbb{R}^d)$ *be a sequence of maps such that for some* $m > 0$*, there holds*

$$|v_n(x)| \leq m(1 + |x|)$$

for all $x \in \mathbb{R}^d$ *and each* $n \geq 1$*. Moreover, suppose that* $\mathsf{d}_{cc}(v_n, v) \to 0$ *as* $n \to +\infty$ *for some* $v \in C^0(\mathbb{R}^d, \mathbb{R}^d)$*. Then, for each* $p \in [1, +\infty)$ *and* $\mu \in \mathscr{P}_p(\mathbb{R}^d)$*, it holds that* $\{v_n\}_{n=1}^{+\infty} \subset \mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d; \mu)$ *and also*

$$\|v - v_n\|_{\mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d; \mu)} \underset{n \to +\infty}{\longrightarrow} 0.$$

Proof. The fact that $v_n \in \mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d; \mu)$ for each $n \geq 1$ simply follows from the observation that

$$\|v_n\|_{\mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\,\mu)} \leq m\bigg(\int_{\mathbb{R}^d}(1+|x|)^p \mathrm{d}\mu(x)\bigg)^{1/p} < +\infty,$$

and likewise $v \in \mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d; \mu)$. Fix now an arbitrary $\varepsilon > 0$ and consider some $R_\varepsilon > 0$ for which

$$\bigg(\int_{\{x\in\mathbb{R}^d \text{ s.t. } |x|\geq R_\varepsilon\}}(1+|x|)^p \mathrm{d}\mu(x)\bigg)^{1/p} \leq \frac{\varepsilon}{4(1+m)}.$$

Then, choose an integer $N_\varepsilon \geq 1$ such that

$$\|v - v_n\|_{C^0(B(0,R_\varepsilon),\mathbb{R}^d)} \leq \frac{\varepsilon}{2}$$

for each $n \geq N_\varepsilon$, which is always possible by the definition (2.6) of $\mathsf{d}_{cc}(\cdot,\cdot)$. Whence, it follows that

$$\|v - v_n\|_{\mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\,\mu)} \leq \|v - v_n\|_{C^0(B(0,R_\varepsilon),\mathbb{R}^d)} + 2m\bigg(\int_{\{x\in\mathbb{R}^d \text{ s.t. } |x|\geq R_\varepsilon\}}(1+|x|)^p \mathrm{d}\mu(x)\bigg)^{1/p} \leq \varepsilon$$

for each $n \geq N_\varepsilon$, which yields the desired convergence result. □

2.3 Continuity equations and inclusions in Wasserstein spaces

In this last preliminary section, we expose well-posedness results and estimates for solutions of continuity equations and inclusions in Wasserstein spaces. These latter are mostly borrowed from our previous works [22, 28], but we also point the reader to [5, Chapter 8] and [3, 75] for more standard versions thereof.

Continuity equations in the Carathéodory framework. In the ensuing paragraphs, we recall some elementary results pertaining to the qualitative properties of *continuity equations* of the form

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

defined over some time interval $[0,T]$ with $T > 0$, whose solutions are understood in the sense of distributions, namely, satisfying

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

against smooth test functions $\varphi \in C_c^\infty((0,T)\times\mathbb{R}^d, \mathbb{R})$. In this context, given a real number $p \in [1,+\infty)$ and a pair of elements $(\tau, \mu_\tau) \in [0,T]\times\mathscr{P}_p(\mathbb{R}^d)$, we will study the well-posedness of the Cauchy problem

$$\begin{cases} \partial_t \mu(t) + \mathrm{div}_x(v(t)\mu(t)) = 0, \\ \mu(\tau) = \mu_\tau, \end{cases} \tag{2.7}$$

in the case where the velocity field $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ satisfies either the following standard Carathéodory assumptions, or some of their variants.

Hypotheses (CE).

(*i*) *The velocity field $v : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ is Carathéodory, i.e. $t \in [0,T] \mapsto v(t,x)$ is $\mathscr{L}^1$-measurable for all $x \in \mathbb{R}^d$ while $x \in \mathbb{R}^d \mapsto v(t,x)$ is continuous for $\mathscr{L}^1$-almost every $t \in [0,T]$. Moreover, 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* $t \in [0,T]$.

In their strongest form, the well-posedness results stated in Theorem 2.17 below for continuity equations involve the notion of *characteristic flows* generated by a velocity field.

Definition 2.16 (Characteristic flow)**.** *Given a velocity field* $v : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ *satisfying Hypotheses* **(CE)**, *we define the* characteristic flows $(\Phi^v_{(\tau,t)})_{\tau,t\in[0,T]} \subset C^0(\mathbb{R}^d,\mathbb{R}^d)$ *as the unique collection of maps satisfying*

$$\Phi^v_{(\tau,t)}(x) = x + \int_\tau^t v(s,\Phi^v_{(\tau,s)}(x))\mathrm{d}s \tag{2.8}$$

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

Theorem 2.17 (Well-posedness in the Carathéodory framework)**.** *Let* $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ *be a velocity field satisfying Hypothesis* **(CE)**-(*i*), *and fix some* $(\tau,\mu_\tau) \in [0,T] \times \mathbb{R}^d$.

Then, the Cauchy problem (2.7) *admits solutions* $\mu(\cdot) \in \mathrm{AC}([\tau,T],\mathscr{P}_p(\mathbb{R}^d))$. *In the case where Hypothesis* **(CE)**-(*ii*) *holds as well, the latter is then uniquely defined on the whole interval* $[0,T]$, *and represented explicitly by the formula*

$$\mu(t) = \Phi^v_{(\tau,t)\,\sharp}\mu_\tau$$

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

Set-valued dynamics in Wasserstein spaces. In the next few paragraphs, we recollect for the sake of completeness the definition of *continuity inclusions* introduced in our earlier works [22, 28], along with several estimates on which our main contributions will strongly rely. In what follows, given some $p \in [1,+\infty)$, we focus on the set-valued Cauchy problems of the form

$$\begin{cases} \partial_t\mu(t) \in -\mathrm{div}_x\Big(V(t,\mu(t))\mu(t)\Big), \\ \mu(\tau) = \mu_\tau, \end{cases} \tag{2.9}$$

wherein $(\tau,\mu_\tau) \in [0,T]\times\mathscr{P}_p(\mathbb{R}^d)$ and $V : [0,T]\times\mathscr{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d,\mathbb{R}^d)$ are given, and whose solutions are understood in the following sense.

Definition 2.18 (Solutions to continuity inclusions)**.** *A curve of measures* $\mu(\cdot) \in \mathrm{AC}([0,T],\mathscr{P}_p(\mathbb{R}^d))$ *is said to be a solution of the Cauchy problem* (2.9) *if there exists an* $\mathscr{L}^1$*-measurable selection* $t \in [0,T] \mapsto v(t) \in V(t,\mu(t))$ *such that the* trajectory-selection pair $(\mu(\cdot),v(\cdot))$ *satisfies*

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

in the sense of distributions.

Based on our earlier contributions, we will assume throughout this article that the dynamics satisfies the following assumptions. Therein and in what follows, $C^0(\mathbb{R}^d,\mathbb{R}^d)$ is systematically endowed with the separable Fréchet structure induced by $\mathsf{d}_{cc}(\cdot,\cdot)$, exposed 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 with 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*

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

(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*

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

Examples of classical set-valued mappings defined in terms of control systems satisfying localised variants of **(CI)** are provided in [22, Section 4]. In our subsequent developments, we will frequently refer to solutions of (2.9) by using the terminology of *reachable* and *solutions sets*, defined as follows.

Definition 2.19 (Reachable and solution sets of continuity inclusions)**.** *Given a pair of elements* $(\tau,\mu_\tau) \in [0,T] \times \mathscr{P}_p(\mathbb{R}^d)$*, we define the (forward)* solution set *of the Cauchy problem* (2.9) *as*

$$\mathcal{S}_{[\tau,T]}(\tau,\mu_\tau) := \Big\{\mu(\cdot) \in \mathrm{AC}([\tau,T],\mathscr{P}_p(\mathbb{R}^d)) \text{ s.t. } \mu(\cdot) \textit{ is a solution of } (2.9)\Big\},$$

and denote by

$$\mathcal{R}_{(\tau,t)}(\mu_\tau) := \Big\{\mu(t) \text{ s.t. } \mu(\cdot) \in \mathcal{S}_{[\tau,T]}(\tau,\mu_\tau)\Big\}$$

the corresponding reachable sets *at time* $t \in [\tau,T]$*.*

By combining classical concatenation results for solutions of continuity equations (see e.g. [51, Lemma 4.4]) and Definition 2.18, it can be shown that the reachable sets satisfy the semigroup property

$$\mathcal{R}_{(\tau,t)}(\mu_\tau) = \mathcal{R}_{(s,t)} \circ \mathcal{R}_{(\tau,s)}(\mu_\tau) \tag{2.10}$$

for all times $\tau \leq s \leq t \leq T$. Besides, it follows from Hypotheses **(CI)** and Theorem 2.17 that solution curves are also well-defined and unique backward in time. Hence, each element of $\mathcal{S}_{[\tau,T]}(\tau,\mu_\tau)$ can be seen as a restriction to $[\tau,T]$ of some curve in $\mathcal{S}_{[0,T]}(\tau,\mu_\tau)$

In the next propositions, we recall several a priori estimates for solutions of (2.9), along with some useful topological properties for the reachable and solution sets. Therein and in what follows, given a map $m(\cdot) \in L^1([0,T],\mathbb{R}_+)$, we will frequently use the shorthand notation $\|m(\cdot)\|_1 := \|m(\cdot)\|_{L^1([0,T],\mathbb{R})}$.

Proposition 2.20 (Moment, equi-integrability and absolute continuity 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* $(\tau,\mu_\tau) \in [0,T] \times \mathscr{P}_p(\mathbb{R}^d)$*.*

Then, there exists a constant $\mathcal{C}_T > 0$ *which only depends on the magnitudes of* $p, \mathcal{M}_p(\mu_\tau)$ *and* $\|m(\cdot)\|_1$ *such that every curve* $\mu(\cdot) \in \mathcal{S}_{[0,T]}(\tau,\mu_\tau)$ *complies with the a priori moment bound*

$$\mathcal{M}_p(\mu(t)) \leq \mathcal{C}_T, \tag{2.11}$$

as well as the uniform equi-integrability estimate

$$\int_{\{x\in\mathbb{R}^d \text{ s.t. } |x|\geq R\}} |x|^p \mathrm{d}\mu(t)(x) \leq \mathcal{C}_T^p \int_{\{x\in\mathbb{R}^d \text{ s.t. } |x|\geq R/\mathcal{C}_T-1\}} (1+|x|)^p \mathrm{d}\mu_\tau(x) \tag{2.12}$$

for all times $t \in [0,T]$ *and each* $R > 0$*. Moreover, the following uniform absolute continuity inequality*

$$W_p(\mu(t_1),\mu(t_2)) \leq (1+\mathcal{C}_T)\int_{t_1}^{t_2} m(s)\mathrm{d}s \tag{2.13}$$

holds for all times $0 \leq t_1 \leq t_2 \leq T$ *and every* $\mu(\cdot) \in \mathcal{S}_{[0,T]}(\tau,\mu_\tau)$*.*

Proposition 2.21 (Topological properties of the reachable and solution sets)**.** *Assume that the hypotheses of Proposition 2.20 hold and that* $V : [0,T] \times \mathscr{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d,\mathbb{R}^d)$ *has convex images. Then the reachable sets* $\mathcal{R}_{(\tau,t)}(\mu_\tau) \subset \mathscr{P}_p(\mathbb{R}^d)$ *are compact for all times* $t \in [0,T]$*, and the solution set* $\mathcal{S}_{[0,T]}(\tau,\mu_\tau) \subset C^0([0,T],\mathscr{P}_p(\mathbb{R}^d))$ *is compact for the topology of uniform convergence.*

Proof. The fact that the solution set $\mathcal{S}_{[0,T]}(\tau,\mu_\tau) \subset C^0([\tau,T],\mathscr{P}_p(\mathbb{R}^d))$ is compact when $V : [0,T] \times \mathscr{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d,\mathbb{R}^d)$ has convex images was proven in [28, Theorem 3.5]. It is then straightforward to show that the underlying reachable sets are compact for all times $t \in [0,T]$. □

We end this preliminary section by recalling a simplified and condensed version of one of the main results of [28], which combines an existence result for (2.9) together with a powerful estimate "à la Grönwall" involving the distance to an a priori given curve of measures.

Theorem 2.22 (Local Filippov estimates 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* **(CI)** *and $\nu(\cdot) \in \mathrm{AC}([0,T],\mathscr{P}_p(\mathbb{R}^d))$ be a solution of the continuity equation*

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

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

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

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and all $y \in \mathbb{R}^d$. Given $R > 0$, denote by $\eta_R(\cdot) \in L^1([0,T],\mathbb{R}_+)$ the local mismatch function*, defined by*

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

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

Then for every $(\tau,\mu_\tau) \in [0,T] \times \mathscr{P}_p(\mathbb{R}^d)$ and each $R > 0$, there exists a curve of measures $\mu(\cdot) \in \mathcal{S}_{[0,T]}(\tau,\mu_\tau)$ which satisfies the a priori estimate

$$W_p(\mu(t),\nu(t)) \leq \mathcal{C}'_T\Big(W_p(\mu_\tau,\nu(\tau)) + \int_\tau^t \eta_R(s)\mathrm{d}s + \mathcal{E}_\nu(\tau,t,R)\Big) \tag{2.14}$$

for all times $t \in [\tau,T]$. Therein, the constant $\mathcal{C}'_T > 0$ only depends on the magnitudes of the data $p, \mathcal{M}_p(\mu_\tau), \|m(\cdot)\|_1, \|l(\cdot)\|_1$ and $\|L(\cdot)\|_1$, while the error term $\mathcal{E}_\nu(\tau,t,R)$ is given explicitly by

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

for all times $t \in [\tau,T]$, where $\mathcal{C}_T > 0$ is the constant appearing in Proposition 2.20.

Remark 2.23 (Link between continuity equations and inclusions)**.** *In the particular case in which $V : [0,T] \times \mathscr{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d,\mathbb{R}^d)$ happens to be single valued and independent of μ – that is if $V(t,\mu) = \{v(t)\}$ for $\mathscr{L}^1$-almost every $t \in [0,T]$ and each $\mu \in \mathscr{P}_p(\mathbb{R}^d)$ –, then the corresponding velocity field satisfies Hypotheses* **(CE)***. In addition, the solution of (2.9) is then unique, coincides with that of (2.7), and complies with the a priori estimates of Proposition 2.20 and Theorem 2.22.*

3 Infinitesimal behaviour of the reachable sets

In this section, we prove two fundamental results concerning the metric differentiability properties of solutions of the Cauchy problem

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

which are largely inspired by the analysis carried out in [63, Section 2]. The first one, discussed in the ensuing theorem, deals with the existence of curves with (approximately) prescribed initial velocities.

Theorem 3.1 (Existence of admissible curves with approximate initial velocities)**.** *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)***.*

Then, there exists a subset $\mathscr{T} \subset (0,T)$ of full $\mathscr{L}^1$-measure such that for every $\tau \in \mathscr{T}$, all $\mu_\tau \in \mathscr{P}_p(\mathbb{R}^d)$, each $v_\tau \in V(\tau, \mu_\tau)$ and any $\varepsilon > 0$, there exist some $h_\varepsilon > 0$ along with a curve $\mu_\varepsilon(\cdot) \in \mathcal{S}_{[0,T]}(\tau, \mu_\tau)$ such that

$$W_p\Big(\mu_\varepsilon(\tau + h), (\mathrm{Id} + hv_\tau)_\sharp \mu_\tau\Big) \leq \varepsilon h, \tag{3.2}$$

for all $h \in [0, h_\varepsilon]$.

Proof. To begin with, denote by $\mathscr{T}_m, \mathscr{T}_l, \mathscr{T}_L \subset (0,T)$ the sets of one-sided Lebesgue points of the maps $m(\cdot), l(\cdot)$, and $L(\cdot)$, respectively, and by $\mathscr{T}_{\mathrm{H}} \subset (0,T)$ the subset of full $\mathscr{L}^1$-measure over which Hypotheses **(CI)**-(ii), (iii) and (iv) hold. By Theorem 2.11, there exists for every $k \geq 1$ a compact set $\mathcal{A}^k \subset [0,T]$ satisfying $\mathscr{L}^1([0,T] \setminus \mathcal{A}^k) < \frac{1}{2^k}$, and such that $V : \mathcal{A}^k \times \mathscr{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d, \mathbb{R}^d)$ is lower-semicontinuous in the sense of Definition 2.5-(i). For each $n \geq 1$, define then $\mathscr{A}_n \subset [0,T]$ as

$$\mathscr{A}_n := \bigcap_{k>n} \mathcal{A}^k, \tag{3.3}$$

and denote by $\tilde{\mathscr{A}}_n \subset \mathscr{A}_n$ the subset of its one-sided density points understood in the sense of Definition 2.1, which can be characterised as the subset of full $\mathscr{L}^1$-measure in $\mathscr{A}_n$ such that

$$\lim_{h \to 0^+} \frac{\mathscr{L}^1([\tau, \tau + h] \setminus \mathscr{A}_n)}{h} = 0 \tag{3.4}$$

for each $\tau \in \tilde{\mathscr{A}}_n$. Upon noting that for each $n \geq 1$, one has

$$\begin{aligned} \mathscr{L}^1([0,T] \setminus \tilde{\mathscr{A}}_n) &= \mathscr{L}^1\Big(\bigcup\nolimits_{k>n} \big([0,T] \setminus \mathcal{A}^k\big)\Big) \\ &\leq \sum_{k=n+1}^{+\infty} \mathscr{L}^1([0,T] \setminus \mathcal{A}^k) \leq \frac{1}{2^n} \end{aligned}$$

while observing that the sequence of measurable sets $(\tilde{\mathscr{A}}_n)$ is increasing by construction, it holds that

$$\mathscr{L}^1\Big([0,T] \setminus \Big(\bigcup\nolimits_{n \geq 1} \tilde{\mathscr{A}}_n\Big)\Big) = \lim_{n \to +\infty} \mathscr{L}^1([0,T] \setminus \tilde{\mathscr{A}}_n) = 0.$$

Therefore, the set $\mathscr{T} \subset (0,T)$ defined by

$$\mathscr{T} := \Big(\bigcup\nolimits_{n \geq 1} \tilde{\mathscr{A}}_n\Big) \cap \mathscr{T}_m \cap \mathscr{T}_l \cap \mathscr{T}_L \cap \mathscr{T}_{\mathrm{H}}, \tag{3.5}$$

has full $\mathscr{L}^1$-measure in $[0,T]$, and in the sequel we fix an element $\tau \in \mathscr{T}$. We also pick a radius $R_\varepsilon > 0$ satisfying

$$2m(\tau)(1 + \mathcal{C}_T)\Big(\int_{\{x \in \mathbb{R}^d \text{ s.t. } |x| \geq R_\varepsilon/\mathcal{C}_T - 1\}} (1 + |x|^p) \mathrm{d}\mu_\tau(x)\Big)^{1/p} \leq \varepsilon, \tag{3.6}$$

where $\mathcal{C}_T > 0$ is given as in Proposition 2.20 and Theorem 2.22.

Step 1 – Construction of an admissible curve. Observe that under Hypotheses **(CI)**, each element $v_\tau \in V(\tau, \mu_\tau)$ is a time-independent vector field satisfying Hypotheses **(CE)** with constants $m(\tau), l(\tau) \geq 0$. Thus, by Theorem 2.17, there exists a unique solution $\nu(\cdot) \in \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d))$ of

$$\begin{cases} \partial_t \nu(t) + \mathrm{div}_x(v_\tau \nu(t)) = 0, \\ \nu(\tau) = \mu_\tau, \end{cases}$$

and the latter can be represented explicitly as

$$\nu(t) = \Phi^{v_\tau}_{(\tau,t)\,\sharp} \mu_\tau$$

for all times $t \in [0,T]$. Therein, the maps $(\Phi^{v_\tau}_{(\tau,t)})_{t\in[0,T]} \subset C^0(\mathbb{R}^d,\mathbb{R}^d)$ are the characteristic flows generated by $v_\tau \in V(\tau,\mu_\tau)$ in the sense of Definition 2.16. By standard linearisation techniques (see e.g. [24, Appendix A]), it can further be shown that

$$\Phi^{v_\tau}_{(\tau,\tau+h)}(x) = x + hv_\tau(x) + o_{\tau,x}(h)$$

for all $x \in \mathbb{R}^d$ and any sufficiently small $h > 0$, where $\int_{\mathbb{R}^d} |o_{\tau,x}(h)|^p \mathrm{d}\mu_\tau(x) = o_\tau(|h|^p)$. Thence, upon remarking that

$$\Big(\mathrm{Id} + hv_\tau, \Phi^{v_\tau}_{(\tau,\tau+h)}\Big)_\sharp \mu_\tau \in \Gamma\Big((\mathrm{Id} + hv_\tau)_\sharp \mu_\tau, \nu(\tau+h)\Big),$$

we straightforwardly deduce from (2.1) the distance estimate

$$W_p\Big(\nu(\tau+h), (\mathrm{Id} + hv_\tau)_\sharp \mu_\tau\Big) \leq o_\tau(h), \tag{3.7}$$

which holds for every sufficiently small $h > 0$. Moreover, since $\nu(\tau) = \mu_\tau$ by construction, there exists by the Filippov estimates of Theorem 2.22 a curve $\mu_\varepsilon(\cdot) \in \mathcal{S}_{[0,T]}(\tau,\mu_\tau)$ which satisfies

$$W_p(\mu_\varepsilon(t), \nu(t)) \leq \mathcal{C}'_T\Big(\int_\tau^t \mathrm{dist}_{C^0(B(0,R_\varepsilon),\mathbb{R}^d)}\Big(v_\tau\,; V(s,\nu(s))\Big)\mathrm{d}s + \mathcal{E}_\nu(\tau,t,R_\varepsilon)\Big) \tag{3.8}$$

for all times $t \in [\tau,T]$, where $\mathcal{C}'_T > 0$ only depends on the magnitudes of $p, \mathcal{M}_p(\mu_\tau), \|m(\cdot)\|_1, \|l(\cdot)\|_1$, and $\|L(\cdot)\|_1$.

Step 2 – Distance estimate in the vicinity of $\tau \in \mathscr{T}$. In order to conclude, we need to show that the right-hand side of (3.8) is bounded from above by $\varepsilon h + o_{\tau,\varepsilon}(h)$ when $t = \tau + h$ with $h > 0$ sufficiently small. Observe first that by our choice of $R_\varepsilon > 0$ via (3.6), one has that

$$\begin{aligned}\mathcal{E}_\nu(\tau,\tau+h,R_\varepsilon) &= 2(1+\mathcal{C}_T)\,\|m(\cdot)\|_{L^1([\tau,\tau+h])}\left(\int_{\{x\in\mathbb{R}^d \text{ s.t. } |x|\geq R_\varepsilon/\mathcal{C}_T-1\}} (1+|x|^p)\mathrm{d}\mu_\tau(x)\right)^{1/p} \\ &\leq \varepsilon h + o_{\tau,\varepsilon}(h)\end{aligned} \tag{3.9}$$

whenever $h > 0$ is small enough, since $\tau \in \mathscr{T}$ is a one-sided Lebesgue point of $m(\cdot) \in L^1([0,T],\mathbb{R}_+)$.

In order to derive an upper-bound on the integral of the mismatch function, recall that by the definition (3.5) of $\mathscr{T} \subset (0,T)$, there exists an integer $n \geq 1$ such that $\tau \in \tilde{\mathscr{A}}_n$ and the set-valued map $t \in \mathscr{A}_n \rightrightarrows V(t,\mu_\tau)$ is lower-semicontinuous. Hence for each $\varepsilon' > 0$ there exists some $\delta_n > 0$ for which

$$\mathrm{dist}_{C^0(\mathbb{R}^d,\mathbb{R}^d)}\Big(v_\tau\,; V(t,\mu_\tau)\Big) \leq \varepsilon'$$

for every $t \in [\tau,\tau+\delta_n] \cap \mathscr{A}_n$, where we recall that $(C^0(\mathbb{R}^d,\mathbb{R}^d), \mathsf{d}_{cc}(\cdot,\cdot))$ is equipped with the Fréchet structure described in Definition 2.12. In particular, by choosing $\varepsilon' > 0$ to be sufficiently small, it follows from the definition (2.6) of the metric $\mathsf{d}_{cc}(\cdot,\cdot)$ that

$$\mathrm{dist}_{C^0(B(0,R_\varepsilon),\mathbb{R}^d)}\Big(v_\tau\,; V(t,\mu_\tau)\Big) \leq \varepsilon, \tag{3.10}$$

for every $t \in [\tau,\tau+\delta_n] \cap \mathscr{A}_n$, where $(C^0(B(0,R_\varepsilon),\mathbb{R}^d), \|\cdot\|_{C^0(B(0,R_\varepsilon),\mathbb{R}^d)})$ is endowed with its usual Banach space structure. Besides, noting that the sets $V(t,\mu_\tau) \subset C^0(\mathbb{R}^d,\mathbb{R}^d)$ are compact under Hypotheses **(CI)**-(ii) and (iii) as a consequence of Theorem 2.13, while observing that the map

$$w \in C^0(\mathbb{R}^d,\mathbb{R}^d) \mapsto \|v_\tau - w\|_{C^0(B(0,R_\varepsilon,\mathbb{R}^d)} \in \mathbb{R}_+$$

is continuous, it follows from Theorem 2.10 applied with Hypothesis **(CI)**-(i) that there exists a measurable selection $t \in [0,T] \mapsto w_\tau(t) \in V(t,\mu_\tau)$ such that

$$\|v_\tau - w_\tau(t)\|_{C^0(B(0,R_\varepsilon),\mathbb{R}^d)} = \mathrm{dist}_{C^0(B(0,R_\varepsilon),\mathbb{R}^d)}\Big(v_\tau\,; V(t,\mu_\tau)\Big) \tag{3.11}$$

for every $t \in [\tau, \tau + \delta_n] \cap \mathscr{A}_n$. Furthermore, by Hypotheses **(CI)**-(i) and (iv) combined again with Theorem 2.10, there exists a measurable selection $t \in [0,T] \mapsto w(t) \in V(t, \nu(t))$ satisfying

$$\|w(t) - w_\tau(t)\|_{C^0(B(0,R_\varepsilon),\mathbb{R}^d)} \leq L(t) W_p(\mu_\tau, \nu(t))$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$, which together with (3.10) and (3.11) further yields

$$\mathrm{dist}_{C^0(B(0,R_\varepsilon),\mathbb{R}^d)}\Big(v_\tau\,; V(t,\nu(t))\Big) \leq \varepsilon + L(t) W_p(\mu_\tau, \nu(t)) \tag{3.12}$$

for $\mathscr{L}^1$-almost every $t \in [\tau, \tau + \delta_n] \cap \mathscr{A}_n$. There now remains to estimate the integral over $[\tau, \tau + h]$ of the local mismatch function. The latter can be decomposed into the sum of two terms as

$$\begin{aligned}\int_\tau^{\tau+h} \mathrm{dist}_{C^0(B(0,R_\varepsilon),\mathbb{R}^d)}\Big(v_\tau\,; V(t,\nu(t))\Big)\mathrm{d}t = &\int_{[\tau,\tau+h]\setminus\mathscr{A}_n} \mathrm{dist}_{C^0(B(0,R_\varepsilon),\mathbb{R}^d)}\Big(v_\tau\,; V(t,\nu(t))\Big)\mathrm{d}t \\ &+ \int_{[\tau,\tau+h]\cap\mathscr{A}_n} \mathrm{dist}_{C^0(B(0,R_\varepsilon),\mathbb{R}^d)}\Big(v_\tau\,; V(t,\nu(t))\Big)\mathrm{d}t.\end{aligned} \tag{3.13}$$

As a consequence of Hypothesis **(CI)**-(ii), the first of these two integrals can be estimated as

$$\begin{aligned}&\int_{[\tau,\tau+h]\setminus\mathscr{A}_n} \mathrm{dist}_{C^0(B(0,R_\varepsilon),\mathbb{R}^d)}\Big(v_\tau\,; V(t,\nu(t))\Big)\mathrm{d}t \\ &\qquad\leq \int_{[\tau,\tau+h]\setminus\mathscr{A}_n} \Big(\|v_\tau\|_{C^0(B(0,R_\varepsilon),\mathbb{R}^d)} + \sup_{w\in V(t,\nu(t))} \|w\|_{C^0(B(0,R_\varepsilon),\mathbb{R}^d)} \Big)\mathrm{d}t \\ &\qquad\leq (1 + R_\varepsilon + \mathcal{C}_T) \int_{[\tau,\tau+h]\setminus\mathscr{A}_n} (m(\tau) + m(t))\mathrm{d}t \\ &\qquad\leq (1 + R_\varepsilon + \mathcal{C}_T)\Big(2m(\tau)\mathscr{L}^1([\tau,\tau+h]\setminus\mathscr{A}_n) + \int_{[\tau,\tau+h]\setminus\mathscr{A}_n} (m(t) - m(\tau))\mathrm{d}t \Big) \\ &\qquad= o_{\tau,\varepsilon}(h)\end{aligned} \tag{3.14}$$

where we used the characterisation (3.4) of the one-sided density points of $\mathscr{A}_n$, along with the fact that $\tau \in \mathscr{T}$ is a one-sided Lebesgue point of $m(\cdot) \in L^1([0,T],\mathbb{R}_+)$, as well as the moment bound of Proposition 2.20. By (3.12), the second term in (3.13) can be bounded from above as

$$\begin{aligned}&\int_{[\tau,\tau+h]\cap\mathscr{A}_n} \mathrm{dist}_{C^0(B(0,R_\varepsilon),\mathbb{R}^d)}\Big(v_\tau\,; V(t,\nu(t))\Big)\mathrm{d}t \\ &\qquad\leq \int_{[\tau,\tau+h]\cap\mathscr{A}_n} \Big(\varepsilon + L(t) W_p(\mu_\tau,\nu(t))\Big)\mathrm{d}t \\ &\qquad\leq \varepsilon h + \mathcal{C}'_T \int_\tau^{\tau+h} L(t) \int_\tau^t \|v_\tau\|_{C^0(B(0,R_\varepsilon),\mathbb{R}^d)}\,\mathrm{d}s\,\mathrm{d}t \\ &\qquad\quad + \mathcal{C}'_T \int_\tau^{\tau+h} 2\,\|m(\cdot)\|_{L^1([\tau,\tau+h])}\,(1+\mathcal{C}_T)\Big(\int_{\{x \text{ s.t. } |x|\geq R_\varepsilon/\mathcal{C}_T - 1\}} (1+|x|)^p \mathrm{d}\mu_\tau(x) \Big)^{1/p} \mathrm{d}t \\ &\qquad\leq \varepsilon h + o_{\tau,\varepsilon}(h),\end{aligned} \tag{3.15}$$

where we used the single-valued version of the distance estimate of Theorem 2.22 – see Remark 2.23 – along with the fact that $\tau \in \mathscr{T}$ is a one-sided Lebesgue point of $m(\cdot), L(\cdot) \in L^1([0,T],\mathbb{R}_+)$. By plugging (3.14) and (3.15) into (3.13) and combining the resulting estimate with (3.8), one then obtains

$$W_p\Big(\mu_\varepsilon(\tau+h), \nu(\tau+h)\Big) \leq \varepsilon h + o_{\tau,\varepsilon}(h) \tag{3.16}$$

for every small $h \in (0, \delta_n]$. Upon merging (3.16) with (3.7) and taking $h_\varepsilon > 0$ so that $o_{\tau,\varepsilon}(h) \leq \varepsilon h$ for all $h \in [0, h_\varepsilon]$, one finally has up to rescaling $\varepsilon > 0$ by a constant that

$$W_p\Big(\mu_\varepsilon(\tau+h), (\mathrm{Id} + h v_\tau)_\sharp \mu_\tau\Big) \leq \varepsilon h,$$

for all $h \in [0, h_\varepsilon]$, which concludes the proof. □

In the following theorem, we establish a property that is complementary to the one we previously discussed, which ensures the existence of an admissible velocity representing (approximately) the local behaviour of any given solution of (2.9) when the admissible velocities are convex.

Theorem 3.2 (Infinitesimal behaviour of reachable sets)**.** *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 with convex images satisfying Hypotheses* **(CI)**.

Then, there exists a subset $\mathscr{T} \subset (0,T)$ of full $\mathscr{L}^1$-measure such that for every $\tau \in \mathscr{T}$, all $\mu_\tau \in \mathscr{P}_p(\mathbb{R}^d)$, each solution $\mu(\cdot) \in \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d))$ of (3.1), any $\varepsilon > 0$ and every sequence $h_i \to 0$, there exists an element $v_\tau^\varepsilon \in V(\tau, \mu_\tau)$ such that

$$W_p\Big(\mu(\tau + h_{i_k}^\varepsilon), (\mathrm{Id} + h_{i_k}^\varepsilon v_\tau^\varepsilon)_\sharp \mu_\tau\Big) \leq \varepsilon |h_{i_k}^\varepsilon| \tag{3.17}$$

along a subsequence $h_{i_k}^\varepsilon \to 0$ which depends both on $v_\tau^\varepsilon \in V(\tau, \mu_\tau)$ and $\varepsilon > 0$.

Proof. First, let $\mathscr{T}_m, \mathscr{T}_l, \mathscr{T}_L$ and $\mathscr{T}_{\mathrm{H}}$ be the subsets of full $\mathscr{L}^1$-measure in $(0,T)$ defined as in the proof of Theorem 3.1. By Theorem 2.11, there exists for each $k \geq 1$ a compact set $\mathcal{A}^k \subset [0,T]$ satisfying $\mathscr{L}^1([0,T] \setminus \mathcal{A}^k) < \frac{1}{2^k}$, and such that the graph

$$\mathrm{Graph}(V)_{|\mathcal{A}^k \times \mathscr{P}_p(\mathbb{R}^d) \times C^0(\mathbb{R}^d, \mathbb{R}^d)} := \Big\{(t, \mu, v) \in \mathcal{A}^k \times \mathscr{P}_p(\mathbb{R}^d) \times C^0(\mathbb{R}^d, \mathbb{R}^d) \text{ s.t. } v \in V(t, \mu)\Big\}$$

of the restriction of $V : [0,T] \times \mathscr{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d, \mathbb{R}^d)$ is closed in $\mathcal{A}^k \times \mathscr{P}_p(\mathbb{R}^d) \times C^0(\mathbb{R}^d, \mathbb{R}^d)$. For each $n \geq 1$, consider the increasing sequence of $\mathscr{L}^1$-measurable sets $(\hat{\mathcal{B}}_n)$ defined by

$$\hat{\mathcal{B}}_n := \Big\{t \in [0,T] \text{ s.t. } m(t) + l(t) \leq n\Big\}.$$

Then, there exists an increasing subsequence $(\mathcal{B}_n)$ of Lebesgue measurable sets that we do not relabel satisfying $\mathcal{B}_n \subset \hat{\mathcal{B}}_n$ as well as $\mathscr{L}^1(\mathcal{B}_n) > 0$ for each $n \geq 1$, and such that

$$\mathscr{L}^1([0,T] \setminus \mathcal{B}_n) \underset{n \to +\infty}{\longrightarrow} 0.$$

By the inner regularity of the Lebesgue measure, one can find an increasing sequence of closed sets $(\mathcal{C}_n) \subset [0,T]$ satisfying $\mathcal{C}_n \subset \mathcal{B}_n$ for each $n \geq 1$, as well as

$$\mathscr{L}^1(\mathcal{C}_n) \geq \Big(1 - \tfrac{1}{T}\mathscr{L}^1([0,T] \setminus \mathcal{B}_n)\Big)\mathscr{L}^1(\mathcal{B}_n).$$

Then, for each $n \geq 1$, define the closed set $\mathscr{A}_n \subset [0,T]$ by

$$\mathscr{A}_n := \Big(\bigcap\nolimits_{k>n} \mathcal{A}^k\Big) \cap \mathcal{C}_n,$$

and as in the proof of Theorem 3.1, denote by $\tilde{\mathscr{A}}_n$ the subset of all one-sided density points of $\mathscr{A}_n$. Noting that the sequence of measurable sets $(\tilde{\mathscr{A}}_n)$ is increasing by construction, and that it satisfies

$$\begin{aligned}
\mathscr{L}^1([0,T] \setminus \tilde{\mathscr{A}}_n) &= \mathscr{L}^1\Big(\bigcup\nolimits_{k>n} \Big([0,T] \setminus \mathcal{A}^k\Big) \cup \Big([0,T] \setminus \mathcal{C}_n\Big)\Big) \\
&\leq \sum_{k=n+1}^{+\infty} \mathscr{L}^1([0,T] \setminus \mathcal{A}^k) + T - \mathscr{L}^1(\mathcal{B}_n) + \mathscr{L}^1([0,T] \setminus \mathcal{B}_n) \\
&\leq \frac{1}{2^n} + 2\,\mathscr{L}^1([0,T] \setminus \mathcal{B}_n) \underset{n \to +\infty}{\longrightarrow} 0,
\end{aligned}$$

it follows that the set $\mathscr{T} \subset (0,T)$ defined by

$$\mathscr{T} := \Big(\bigcup\nolimits_{n \geq 1} \tilde{\mathscr{A}}_n\Big) \cap \mathscr{T}_m \cap \mathscr{T}_l \cap \mathscr{T}_L \cap \mathscr{T}_{\mathrm{H}}$$

has full $\mathscr{L}^1$-measure in $[0,T]$. Fix now some $\tau \in \mathscr{T}$, a measure $\mu_\tau \in \mathscr{P}_p(\mathbb{R}^d)$ as well as a solution $\mu(\cdot) \in \mathcal{S}_{[0,T]}(\tau, \mu_\tau)$ of (3.1) and a sequence $h_i \to 0^+$. Choose also $R_\varepsilon > 0$ in such a way that

$$2m(\tau)(1+\mathcal{C}_T)\Big(\int_{\{x\in\mathbb{R}^d \text{ s.t. } |x|\geq R_\varepsilon\}} (1+|x|^p)\mathrm{d}\mu_\tau(x)\Big)^{1/p} \leq \varepsilon, \tag{3.18}$$

where $\mathcal{C}_T > 0$ is the constant appearing in Proposition 2.20, which we recall only depends on the magnitudes of $p, \mathcal{M}_p(\mu_\tau)$ and $\|m(\cdot)\|_1$.

By the definition of solutions to (3.1), 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 curve $\mu(\cdot)$ solves the Cauchy problem

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

Besides, it can be deduced from Lemma 2.14 that $(t,x) \in [0,T]\times\mathbb{R}^d \mapsto v(t,x) \in \mathbb{R}^d$ is a Carathéodory vector field which satisfies Hypotheses **(CE)**, see for instance [28]. Hence by Theorem 2.17, the curve $\mu(\cdot) \in \mathrm{AC}([0,T],\mathscr{P}_p(\mathbb{R}^d))$ is given explicitly by

$$\mu(t) = \Phi^v_{(\tau,t)\,\sharp}\mu_\tau$$

for all times $t \in [0,T]$. Moreover, there exists by construction an integer $n \geq 1$ such that $\tau \in \tilde{\mathscr{A}_n}$ and the restricted set-valued map $t \in \mathscr{A}_n \rightrightarrows V(t,\mu(t))$ has closed graph. Since $m(\cdot)$ and $l(\cdot)$ are both bounded from above over $\mathscr{A}_n$, it stems from Hypotheses **(CI)**-(i), (ii) and (iii) together with Theorem 2.13 that there exists a compact set $\mathcal{K}_n \subset C^0(\mathbb{R}^d,\mathbb{R}^d)$ such that

$$V(t,\mu(t)) \subset \mathcal{K}_n$$

for all times $t \in \mathscr{A}_n$. Thence, it follows e.g. from [9, Proposition 1.4.8] that $t \in \mathscr{A}_n \rightrightarrows V(t,\mu(t))$ is upper-semicontinuous in the sense of Definition 2.5-(ii). In particular, for each $\varepsilon > 0$, there exists a measurable selection $t \in [0,T] \mapsto \tilde{v}^\varepsilon_\tau(t) \in V(\tau,\mu_\tau)$ along with some $\delta_n > 0$ such that

$$\|v(t) - \tilde{v}^\varepsilon_\tau(t)\|_{C^0(B(0,R_\varepsilon),\mathbb{R}^d)} \leq \varepsilon \tag{3.19}$$

for every $t \in \mathscr{A}_n \cap [\tau, \tau+\delta_n]$, where we leveraged the expression (2.6) of the metric $\mathsf{d}_{cc}(\cdot,\cdot)$.

In what follows, we use the selection $t \in [0,T] \mapsto \tilde{v}^\varepsilon_\tau(t) \in V(\tau,\mu_\tau)$ to build an element $v^\varepsilon_\tau \in V(\tau,\mu_\tau)$ satisfying (3.17) along some subsequence $h^\varepsilon_{i_k} \to 0$. We assume without loss of generality that $h_i \to 0^+$, the general case being similar. By the definition (2.8) of characteristic flows, one has that

$$\Phi^v_{(\tau,\tau+h_i)}(x) = x + \int_\tau^{\tau+h_i} v(t,x)\mathrm{d}t + \int_\tau^{\tau+h_i}\Big(v\big(t,\Phi^v_{(\tau,t)}(x)\big) - v(t,x)\Big)\mathrm{d}t \tag{3.20}$$

for all $x \in \mathbb{R}^d$. As a consequence of Hypothesis **(CE)**-(ii), the second integral term in the right-hand side of (3.20) can be estimated from above for all $x \in B(0,R_\varepsilon)$ as

$$\begin{aligned} \Big|\int_\tau^{\tau+h_i}\Big(v\big(t,\Phi^v_{(\tau,t)}(x)\big) - v(t,x)\Big)\mathrm{d}t\Big| &\leq \int_\tau^{\tau+h_i} l(t)|\Phi^v_{(\tau,t)}(x) - x|\mathrm{d}t \\ &\leq (1+R_\varepsilon+\mathcal{C}_T)\int_\tau^{\tau+h_i} l(t)\Big(\int_\tau^t m(s)\mathrm{d}s\Big)\mathrm{d}t \\ &= o_{\tau,\varepsilon}(h_i) \end{aligned} \tag{3.21}$$

for $h_i > 0$ sufficiently small, since $\tau \in \mathscr{T}$ is a one-sided Lebesgue point of $m(\cdot), l(\cdot) \in L^1([0,T],\mathbb{R}_+)$. At this stage, one may further decompose the first integral in the right-hand side of (3.20) into

$$\begin{aligned} \int_\tau^{\tau+h_i} v(t,x)\mathrm{d}t = \int_\tau^{\tau+h_i} \tilde{v}^\varepsilon_\tau(t,x)\mathrm{d}t &+ \int_{[\tau,\tau+h_i]\setminus\mathscr{A}_n}\Big(v(t,x) - \tilde{v}^\varepsilon_\tau(t,x)\Big)\mathrm{d}t \\ &+ \int_{[\tau,\tau+h_i]\cap\mathscr{A}_n}\Big(v(t,x) - \tilde{v}^\varepsilon_\tau(t,x)\Big)\mathrm{d}t \end{aligned} \tag{3.22}$$

and use Hypothesis **(CI)**-(ii) to estimate the second expression in (3.22) as

$$\Big|\int_{[\tau,\tau+h_i]\setminus\mathscr{A}_n}\Big(v(t,x)-\tilde{v}^\varepsilon_\tau(t,x)\Big)\mathrm{d}t\Big|\leq 2(1+R_\varepsilon+\mathcal{C}_T)\int_{[\tau,\tau+h_i]\setminus\mathscr{A}_n}m(t)\mathrm{d}t=o_{\tau,\varepsilon}(h_i) \tag{3.23}$$

for all $x\in B(0,R_\varepsilon)$, since $\tau\in\mathscr{T}$ is a one-sided Lebesgue point of $m(\cdot)$ as well as a one-sided density point of $\mathscr{A}_n$. Regarding the third term in the right-hand side of (3.22), it follows from (3.19) that

$$\Big\|\int_{[\tau,\tau+h_i]\cap\mathscr{A}_n}\Big(v(t)-\tilde{v}^\varepsilon_\tau(t)\Big)\mathrm{d}t\Big\|_{C^0(B(0,R_\varepsilon),\mathbb{R}^d)}\leq\varepsilon h_i, \tag{3.24}$$

and by merging the estimates of (3.23)-(3.24) with (3.22) while inserting the resulting expression together with (3.21) inside (3.20), one further obtains that

$$\Big|\Phi^v_{(\tau,\tau+h_i)}(x)-x-\int_\tau^{\tau+h_i}\tilde{v}^\varepsilon_\tau(t,x)\mathrm{d}t\Big|\leq\varepsilon h_i+o_{\tau,\varepsilon}(h_i) \tag{3.25}$$

for all $x\in B(0,R_\varepsilon)$. Observe now that since $V(\tau,\mu_\tau)_{|B(0,R_\varepsilon)}\subset C^0(B(0,R_\varepsilon),\mathbb{R}^d)$ is convex and closed, there exists by the separation theorem a family of elements $(v^\varepsilon_\tau(h_i))\subset V(\tau,\mu_\tau)$ such that

$$\frac{1}{h_i}\int_\tau^{\tau+h_i}\tilde{v}^\varepsilon_\tau(t)_{|B(0,R_\varepsilon)}\mathrm{d}t=v^\varepsilon_\tau(h_i)_{|B(0,R_\varepsilon)}.$$

Besides, noting that $V(\tau,\mu_\tau)\subset C^0(\mathbb{R}^d,\mathbb{R}^d)$ is a compact set as a consequence of Hypotheses **(CI)**-(ii) and (iii) and Theorem 2.13, there exists some $v^\varepsilon_\tau\in V(\tau,\mu_\tau)$ for which

$$\mathsf{d}_{cc}(v^\varepsilon_\tau(h^\varepsilon_{i_k}),v^\varepsilon_\tau)\underset{h^\varepsilon_{i_k}\to 0^+}{\longrightarrow}0$$

along a subsequence $h^\varepsilon_{i_k}\to 0^+$. In particular for $h^\varepsilon_{i_k}>0$ sufficiently small, one has that

$$\Big\|\int_\tau^{\tau+h^\varepsilon_{i_k}}\tilde{v}^\varepsilon_\tau(t)\mathrm{d}t-h^\varepsilon_{i_k}v^\varepsilon_\tau\Big\|_{C^0(B(0,R_\varepsilon),\mathbb{R}^d)}=o_{\tau,\varepsilon}(h^\varepsilon_{i_k}),$$

which together with (3.25) finally yields that

$$\Big|\Phi^v_{(\tau,\tau+h^\varepsilon_{i_k})}(x)-x-h^\varepsilon_{i_k}v^\varepsilon_\tau(x)\Big|\leq\varepsilon h^\varepsilon_{i_k}+o_{\tau,\varepsilon}(h^\varepsilon_{i_k}), \tag{3.26}$$

for all $x\in B(0,R_\varepsilon)$. To conclude the proof, there remains to observe that

$$\Big(\Phi^v_{(\tau,\tau+h^\varepsilon_{i_k})},\mathrm{Id}+h^\varepsilon_{i_k}v^\varepsilon_\tau\Big)_\sharp\mu_\tau\in\Gamma\Big(\mu(\tau+h^\varepsilon_{i_k}),(\mathrm{Id}+h^\varepsilon_{i_k}v^\varepsilon_\tau)_\sharp\mu_\tau\Big) \tag{3.27}$$

and to use the standard Wasserstein inequality (2.1) along with the estimate

$$\begin{aligned}\Big(\int_{\{x\in\mathbb{R}^d\text{ s.t. }|x|\geq R_\varepsilon\}}&|\Phi^v_{(\tau,\tau+h^\varepsilon_{i_k})}(x)-x-h^\varepsilon_{i_k}v^\varepsilon_\tau(x)|^p\mathrm{d}\mu_\tau(x)\Big)^{1/p}\\&\leq(1+\mathcal{C}_T)\Big(\|m(\cdot)\|_{L^1([\tau,\tau+h^\varepsilon_{i_k}])}+h^\varepsilon_{i_k}m(\tau)\Big)\Big(\int_{\{x\in\mathbb{R}^d\text{ s.t. }|x|\geq R_\varepsilon\}}(1+|x|)^p\mathrm{d}\mu_\tau(x)\Big)^{1/p}\\&\leq\varepsilon h^\varepsilon_{i_k}+o_{\tau,\varepsilon}(h^\varepsilon_{i_k})\end{aligned} \tag{3.28}$$

which holds as a consequence of Hypothesis **(CI)**-(ii) combined with the equi-integrability bound of Proposition 2.20, the definition (3.18) of $R_\varepsilon>0$, and the fact that $\tau\in\mathscr{T}$ is a one-sided Lebesgue point of $m(\cdot)\in L^1([0,T],\mathbb{R}_+)$. Thence, by merging (3.26), (3.27) and (3.28), one finally recovers that

$$W_p\Big(\mu(\tau+h^\varepsilon_{i_k}),(\mathrm{Id}+h^\varepsilon_{i_k}v^\varepsilon_\tau)_\sharp\mu_\tau\Big)\leq\varepsilon h^\varepsilon_{i_k}$$

whenever $h^\varepsilon_{i_k}>0$ is taken sufficiently small, in particular to ensure that $o_{\tau,\varepsilon}(h^\varepsilon_{i_k})\leq\varepsilon h^\varepsilon_{i_k}$, and up to rescaling the free parameter $\varepsilon>0$ by a fixed constant. The case of an unsigned sequence $h_i\to 0$ being completely similar, this concludes the proof. □

Remark 3.3 (More general assumptions for Theorem 3.1 and Theorem 3.2)**.** *While the two theorems displayed in this section have been proven under the Cauchy-Lipschtz regularity assumptions* **(CI)**, *we expect them to hold under less stringent requirements, where one only asks for the local uniform continuity of the dynamics with respect to the space and measure variables in the spirit of [28, Section 3]. For the sake of conciseness, we postpone these refinements to an ulterior work.*

Remark 3.4 (Exact differential formulations for compactly supported measures)**.** *When the initial data $\mu_\tau \in \mathscr{P}_p(\mathbb{R}^d)$ are compactly supported, the quantitative estimates on the difference quotients derived in Theorem 3.1 and Theorem 3.2 can be expressed more simply in terms of metric derivatives. In this case, following e.g. [22, Proposition 3], there exists for each $r > 0$ a radius $R_r > 0$ such that each $\mu(\cdot) \in \mathcal{S}_{[0,T]}(\tau, \mu_\tau)$ satisfies $\mu(t) \in \mathscr{P}(B(0, R_r))$ for all times $t \in [0, T]$, whenever $\mu_\tau \in \mathscr{P}(B(0, r))$. Owing to this uniform support bound, one can check that the error terms $o_{\tau,\varepsilon}(h)$ in the proofs of both Theorem 3.1 and Theorem 3.2 become independent of $\varepsilon > 0$, which consequently means that the statements of these results hold respectively for every small $h > 0$ for the former, and along a subsequence $h_{i_k} \to 0$ for the latter, which are both independent of $\varepsilon > 0$.*

In Theorem 3.1, upon making the additional assumption that the velocity sets $V(\tau, \mu_\tau)$ are convex, it follows from Proposition 2.21 that the solution set $\mathcal{S}_{[\tau,T]}(\mu_\tau)$ is compact. Then, one can first then let $\varepsilon \to 0^+$ up to considering a subsequence, and then take the limit as $h \to 0^+$ to obtain that for every $v_\tau \in V(\tau, \mu_\tau)$, there exists a solution $\mu(\cdot) \in \mathrm{AC}([0, T], \mathscr{P}_p(\mathbb{R}^d))$ of (3.1) such that

$$\lim_{h\to 0^+} \tfrac{1}{h} W_p\Big(\mu(\tau + h), (\mathrm{Id} + hv_\tau)_\sharp \mu_\tau\Big) = 0.$$

Similarly for Theorem 3.2, upon observing that under Hypotheses **(CI)** *the set $V(\tau, \mu_\tau)$ is compact for the topology of local uniform convergence by the Ascoli-Arzelà theorem, one can let $\varepsilon \to 0^+$ in conjunction with Lemma 2.15 to obtain that, for each solution $\mu(\cdot) \in \mathrm{AC}([0, T], \mathscr{P}_p(\mathbb{R}^d))$ of (3.1) and every sequence $h_i \to 0$, there exists an element $v_\tau \in V(\tau, \mu_\tau)$ such that*

$$\lim_{h_{i_k}\to 0} \tfrac{1}{|h_{i_k}|} W_p\Big(\mu(\tau + h_{i_k}), (\mathrm{Id} + h_{i_k} v_\tau)_\sharp \mu_\tau\Big) = 0$$

along a subsequence $h_{i_k} \to 0$.

4 Viability and invariance theorems for proper constraints sets

In this section, we discuss several necessary and sufficient conditions for the viability and invariance of general constraint tubes $\mathcal{Q} : [0, T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ under the action of the dynamics

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

In this context, we will always assume that $V : [0, T] \times \mathscr{P}_p(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d, \mathbb{R}^d)$ is a set-valued map satisfying Hypotheses **(CI)** for some $p \in (1, +\infty)$, and that $(\tau, \mu_\tau) \in [0, T] \times \mathscr{P}_p(\mathbb{R}^d)$ represents a some arbitrary initial condition complying with the admissibility constraint $\mu_\tau \in \mathcal{Q}(\tau)$.

Definition 4.1 (Viability and invariance)**.** *We say that the tube $\mathcal{Q} : [0, T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ is* viable *for (4.1) if for every $\tau \in [0, T]$ and any $\mu_\tau \in \mathcal{Q}(\tau)$, there exists a solution $\mu(\cdot) \in \mathcal{S}_{[\tau,T]}(\tau, \mu_\tau)$ of the Cauchy problem such that*

$$\mu(t) \in \mathcal{Q}(t)$$

for all times $t \in [\tau, T]$. Similarly, we say $\mathcal{Q} : [0, T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ is invariant *for (4.1) if all the solution curves $\mu(\cdot) \in \mathcal{S}_{[\tau,T]}(\tau, \mu_\tau)$ satisfy $\mu(t) \in \mathcal{Q}(t)$ for all times $t \in [\tau, T]$.*

In what follows, we split the exposition of our main results into two separate parts, starting in Section 4.1 with the simpler case in which the constraints sets are stationary, and then treating the more involved situation in which they are time-dependent in Section 4.2.

4.1 The case of stationary constraints

We start our investigation of the viability and invariance under the action of continuity inclusions by considering the case in which the constraints are represented by a fixed proper subset $\mathcal{Q} \subset \mathscr{P}_p(\mathbb{R}^d)$. In this context, the crucial geometric object which allows to characterise the viability or the invariance of the latter is its *contingent cone*, whose definition is inspired by that of [13] and presented below.

Definition 4.2 (Contingent cone in Wasserstein spaces)**.** *Given a closed set $\mathcal{Q} \subset \mathscr{P}_p(\mathbb{R}^d)$, we define its* contingent cone *at some $\mu \in \mathcal{Q}$ by*

$$T_{\mathcal{Q}}(\mu) := \left\{ \xi \in \mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d; \mu) \text{ s.t. } \liminf_{h \to 0^+} \tfrac{1}{h} \mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}\Big((\mathrm{Id} + h\xi)_\sharp \mu \, ; \mathcal{Q} \Big) = 0 \right\}.$$

The latter can also be characterised in terms of sequences as

$$T_{\mathcal{Q}}(\mu) = \Big\{ \xi \in \mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d; \mu) \text{ s.t. } \textit{there exists a sequence } h_i \to 0^+ \\ \textit{for which } \mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}\Big((\mathrm{Id} + h_i\xi)_\sharp \mu \, ; \mathcal{Q} \Big) = o(h_i) \Big\}.$$

Remark 4.3 (On the choice of defining contingent cones using Borel maps)**.** *It is worth noting that eventhough $\mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d; \mu)$ is merely a seminormed space for any $\mu \in \mathscr{P}_p(\mathbb{R}^d)$, the contingent cones introduced in Definition 4.2 are closed for the convergence induced by the pseudometric $\|\cdot\|_{\mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d; \mu)}$. Indeed, while the limit of a sequence $(\xi_n) \subset T_{\mathcal{Q}}(\mu)$ should be a μ-measurable map by construction (see e.g. [4, Definition 1.12]), the latter always coincides with a Borel function outside of a Borel set with zero μ-measure by [16, Proposition 2.1.11]. In what ensues, the closures of the convex hulls of contingent cones will therefore always be understood in with respect to the pseudometric $\|\cdot\|_{\mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d; \mu)}$.*

By leveraging this notion, we are able to prove the following sufficient viability conditions for proper time-independent constraints sets, which is one of our main contributions. We chose to start by presenting this latter separately, as it thoroughly illustrates the main ideas supporting the more general results of Section 4.2 a simpler setting.

Theorem 4.4 (Sufficient viability conditions for stationary constraints)**.** *Suppose that $p \in (1, +\infty)$, 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 with convex images satisfying Hypotheses* **(CI)** *and $\mathcal{Q} \subset \mathscr{P}_p(\mathbb{R}^d)$ be a proper set such that*

$$V(t, \nu) \cap \overline{\mathrm{co}}\, T_{\mathcal{Q}}(\nu) \neq \emptyset \tag{4.2}$$

for $\mathscr{L}^1$-almost every $t \in [0, T]$ and each $\nu \in \mathcal{Q}$. Then $\mathcal{Q}$ is viable for (4.1).

In the proof of this theorem and several others in the manuscript, we will extensively use the following regularity property of the reachable sets.

Lemma 4.5 (Regularity in time of the reachable sets)**.** *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 with convex images satisfying Hypotheses* **(CI)***, then the reachable maps $t \in [\tau, T] \rightrightarrows \mathcal{R}_{(\tau,t)}(\mu_\tau) \subset \mathscr{P}_p(\mathbb{R}^d)$ are absolutely continuous in the Hausdorff metric for all $(\tau, \mu_\tau) \in [0, T] \times \mathscr{P}_p(\mathbb{R}^d)$.*

Proof. Let $t_1, t_2 \in [\tau, T]$ be such that $\tau \leq t_1 < t_2 \leq T$, and observe that for each $\mu_{t_2} \in \mathcal{R}_{(\tau,t_2)}(\mu_\tau)$, there exists some $\mu_{t_1} \in \mathcal{R}_{(\tau,t_1)}(\mu_\tau)$ such that $\mu_{t_2} \in \mathcal{R}_{(t_1,t_2)}(\mu_{t_1})$. Thence, it follows from the regularity estimate (2.13) of Proposition 2.20 that

$$\mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}\Big(\mu_{t_2} \, ; \mathcal{R}_{(\tau,t_1)}(\mu_\tau) \Big) \leq W_p(\mu_{t_2}, \mu_{t_1}) \leq (1 + \mathcal{C}_T) \int_{t_1}^{t_2} m(s) \mathrm{d}s$$

for some constant $\mathcal{C}_T > 0$ that only depends on the magnitudes of p, $\mathcal{M}(\mu_\tau)$ and $\|m(\cdot)\|_1$. Analogously, it can be shown that

$$\mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}\Big(\mu_{t_1} \, ; \mathcal{R}_{(\tau,t_2)}(\mu_\tau) \Big) \leq (1 + \mathcal{C}_T) \int_{t_1}^{t_2} m(s) \mathrm{d}s,$$

for every $\mu_{t_1} \in \mathcal{R}_{(\tau,t_1)}(\mu_\tau)$. By combining both estimates while remarking that

$$\mathsf{d}_{\mathcal{H}}\Big(\mathcal{R}_{(\tau,t_1)}(\mu_\tau), \mathcal{R}_{(\tau,t_2)}(\mu_\tau)\Big) = \max\Big\{ \sup\Big\{\mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}\Big(\mu_{t_1}\,; \mathcal{R}_{(\tau,t_2)}(\mu_\tau)\Big) \text{ s.t. } \mu_{t_1} \in \mathcal{R}_{(\tau,t_1)}(\mu_\tau)\Big\},$$
$$\sup\Big\{\mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}\Big(\mu_{t_2}\,; \mathcal{R}_{(\tau,t_1)}(\mu_\tau)\Big) \text{ s.t. } \mu_{t_2} \in \mathcal{R}_{(\tau,t_2)}(\mu_\tau)\Big\}\Big\},$$

it then holds that

$$\mathsf{d}_{\mathcal{H}}\Big(\mathcal{R}_{(\tau,t_1)}(\mu_\tau), \mathcal{R}_{(\tau,t_2)}(\mu_\tau)\Big) \leq (1+\mathcal{C}_T)\int_{t_1}^{t_2} m(s)\mathrm{d}s,$$

which concludes the proof. □

This technical result being established, we can move on to the proof of Theorem 4.4.

Proof of Theorem 4.4. In what follows, we assume without loss of generality that $(\tau, \mu_\tau) = (0, \mu^0)$ in (4.1) for some $\mu^0 \in \mathcal{Q}$. Our goal will be to show that the function measuring the distance between $\mathcal{Q}$ and the reachable sets, which is defined for all times $t \in [0,T]$ by

$$g(t) := \mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}\Big(\mathcal{R}_{(0,t)}(\mu^0)\,; \mathcal{Q}\Big),$$

is identically equal to zero. By Proposition 2.8 and Lemma 4.5, it holds that $g(\cdot) \in \mathrm{AC}([0,T],\mathbb{R}_+)$, and we denote by $\mathscr{D} \subset (0,T)$ the set of full $\mathscr{L}^1$-measure over which it is differentiable. Note also that the set $\cup_{t\in[0,T]}\mathcal{R}_{(0,t)}(\mu^0) \subset \mathscr{P}_p(\mathbb{R}^d)$ is compact by Proposition 2.21 and Lemma 4.5. Since $\mu^0 \in \mathcal{Q}$, it is possible to choose a radius $R > 0$ such that

$$\mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}\Big(\mathcal{R}_{(0,t)}(\mu^0)\,; \partial\mathbb{B}_{\mathscr{P}_p(\mathbb{R}^d)}(\mu^0, R)\Big) \geq \mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}\Big(\mathcal{R}_{(0,t)}(\mu^0)\,; \mathcal{Q}\Big) + 1 \tag{4.3}$$

for all times $t \in [0,T]$, by following e.g. the arguments detailed in Appendix B. Therefore, the set $\mathcal{Q}_R := \mathcal{Q} \cap \mathbb{B}_{\mathscr{P}_p(\mathbb{R}^d)}(\mu^0, R)$, which is nonempty as well as compact by construction, is such that

$$g(t) = \mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}\Big(\mathcal{R}_{(0,t)}(\mu^0)\,; \mathcal{Q}_R\Big)$$

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

Step 1 – A Grönwall estimate on the distance function. In this first step, we show by contradiction that the map $g : [0,T] \to \mathbb{R}_+$ is identically equal to zero. Otherwise, since $g(0) = 0$, there should exist a time $t \in [0,T)$ and some $\delta > 0$ such that $g(t) = 0$ and $g(\tau) > 0$ for $\tau \in (t, t+\delta)$.

Let $\mathscr{T} \subset (0,T)$ be the subset of full $\mathscr{L}^1$-measure over which the statement of Theorem 3.1 and Hypotheses **(CI)**-(ii), (iii) and (iv) hold, and fix an element $\tau \in (t, t+\delta) \cap \mathscr{T} \cap \mathscr{D}$. Since $\mathcal{R}_{(0,\tau)}(\mu^0)$ and $\mathcal{Q}_R$ are both compact, one has that

$$g(\tau) = W_p(\mu_\tau, \nu_\tau) \tag{4.4}$$

for some $\mu_\tau \in \mathcal{R}_{(0,\tau)}(\mu^0)$ and $\nu_\tau \in \mathcal{Q}_R$. Moreover, recalling that $R > 0$ is defined in such a way that (4.3) is satisfied, it necessarily holds

$$\nu_\tau \in \mathcal{Q} \cap \mathrm{int}\Big(\mathbb{B}_{\mathscr{P}_p(\mathbb{R}^d)}(\mu^0, R)\Big),$$

which implies in particular that $T_{\mathcal{Q}_R}(\nu_\tau) = T_{\mathcal{Q}}(\nu)$. Thence, by using the definition of contingent cones provided in Definition 4.2, there exists for each $\xi_\tau \in T_{\mathcal{Q}}(\nu_\tau)$ a sequence $h_i \to 0^+$ such that

$$\begin{aligned} W_p\Big(\mu_\tau, (\mathrm{Id} + h_i\xi_\tau)_\sharp\nu_\tau\Big) &\geq \mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}(\mu_\tau\,; \mathcal{Q}_R) + o_\tau(h_i) \\ &= W_p(\mu_\tau, \nu_\tau) + o_\tau(h_i). \end{aligned} \tag{4.5}$$

Besides, by the directional superdifferentiability property of Proposition 2.3 above, one further has

$$\tfrac{1}{p}W_p^p\Big(\mu_\tau,(\mathrm{Id}+h_i\xi_\tau)_\sharp\nu_\tau\Big)-\tfrac{1}{p}W_p^p(\mu_\tau,\nu_\tau)\leq h_i\int_{\mathbb{R}^{2d}}\langle\xi_\tau(y),j_p(y-x)\rangle\mathrm{d}\gamma_\tau(x,y)+o_\tau(h_i) \tag{4.6}$$

for any small $h_i>0$ and every $\gamma_\tau\in\Gamma_o(\mu_\tau,\nu_\tau)$. Hence, by combining the estimates of (4.5) and (4.6), dividing the resulting expression by $h_i>0$ and letting $h_i\to 0^+$, one obtains the following inequality

$$\int_{\mathbb{R}^{2d}}\langle\xi_\tau(y),j_p(x-y)\rangle\mathrm{d}\gamma_\tau(x,y)\leq 0, \tag{4.7}$$

which holds for every $\xi_\tau\in T_{\mathcal{Q}}(\nu_\tau)$ and each $\gamma_\tau\in\Gamma_o(\mu_\tau,\nu_\tau)$.

Observe now that by Theorem 3.1, there exists for any $\varepsilon>0$ and every $v_\tau\in V(\tau,\mu_\tau)$ a curve of measures $\mu_\varepsilon(\cdot)\in\mathcal{S}_{[0,T]}(\tau,\mu_\tau)$ such that

$$W_p\Big(\mu_\varepsilon(\tau+h),(\mathrm{Id}+hv_\tau)_\sharp\mu_\tau\Big)\leq\varepsilon h, \tag{4.8}$$

whenever $h>0$ is sufficiently small. Furthermore, since $\mu_\varepsilon(\tau+h)\in\mathcal{R}_{(\tau,\tau+h)}(\mu_\tau)\subset\mathcal{R}_{(0,\tau+h)}(\mu^0)$, one can estimate from above the forward difference quotient of $\frac{1}{p}g^p(\cdot)$ at $\tau\in(t,t+\delta)\cap\mathscr{T}\cap\mathscr{D}$ as

$$\tfrac{1}{p}g^p(\tau+h)-\tfrac{1}{p}g^p(\tau)\leq\tfrac{1}{p}W_p^p(\mu_\varepsilon(\tau+h),\nu_\tau)-\tfrac{1}{p}W_p^p(\mu_\tau,\nu_\tau). \tag{4.9}$$

Besides, assuming without loss of generality that $\varepsilon,h\in(0,1]$ and noting that $v_\tau\in\mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\mu_\tau)$ by Hypothesis **(CI)**-(ii), it can be deduced from (4.8) along with the estimates of Lemma A.1 below that

$$\begin{aligned}
&\tfrac{1}{p}W_p^p(\mu_\varepsilon(\tau+h),\nu_\tau)-\tfrac{1}{p}W_p^p\Big((\mathrm{Id}+hv_\tau)_\sharp\mu_\tau,\nu_\tau\Big)\\
&\qquad\leq W_p^{p-1}\Big((\mathrm{Id}+hv_\tau)_\sharp\mu_\tau,\nu_\tau\Big)\Big(W_p(\mu_\varepsilon(\tau+h),\nu_\tau)-W_p\Big((\mathrm{Id}+hv_\tau)_\sharp\mu_\tau,\nu_\tau\Big)\Big)\\
&\qquad\quad+C_p\Big|W_p(\mu_\varepsilon(\tau+h),\nu_\tau)-W_p\Big((\mathrm{Id}+hv_\tau)_\sharp\mu_\tau,\nu_\tau\Big)\Big|^{\min\{p,2\}}\\
&\qquad\leq C_p'\Big(W_p\Big(\mu_\varepsilon(\tau+h),(\mathrm{Id}+hv_\tau)_\sharp\mu_\tau\Big)+W_p\Big(\mu_\varepsilon(\tau+h),(\mathrm{Id}+hv_\tau)_\sharp\mu_\tau\Big)^{\min\{p,2\}}\Big)\\
&\qquad\leq C_p'\varepsilon h+o_{\tau,\varepsilon}(h)
\end{aligned} \tag{4.10}$$

for $h>0$ sufficiently small, and where the constants $C_p,C_p'>0$ only depend on the magnitudes of $p,\mathcal{M}(\mu_\tau),\mathcal{M}(\nu_\tau)$, $\|m(\cdot)\|_1$ and $\|v_\tau\|_{\mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\mu_\tau)}$. Thus, by merging the estimate of (4.10) with (4.9), one further obtains up to rescaling $\varepsilon>0$ by a positive constant that

$$\tfrac{1}{p}g^p(\tau+h)-\tfrac{1}{p}g^p(\tau)\leq\tfrac{1}{p}W_p^p\Big((\mathrm{Id}+hv_\tau)_\sharp\mu_\tau,\nu_\tau\Big)-\tfrac{1}{p}W_p^p(\mu_\tau,\nu_\tau)+\varepsilon h+o_{\tau,\varepsilon}(h) \tag{4.11}$$

Besides, it follows again from the directional superdifferentiability property of Proposition 2.3 that

$$\tfrac{1}{p}W_p^p\Big((\mathrm{Id}+hv_\tau)_\sharp\mu_\tau,\nu_\tau\Big)-\tfrac{1}{p}W_p^p(\mu_\tau,\nu_\tau)\leq h\int_{\mathbb{R}^{2d}}\langle v_\tau(x),j_p(x-y)\rangle\mathrm{d}\gamma_\tau(x,y)+o_{\tau,\varepsilon}(h) \tag{4.12}$$

for each optimal transport plan $\gamma_\tau\in\Gamma_o(\mu_\tau,\nu_\tau)$. Therefore, by combining the estimates of (4.11) and (4.12), dividing the resulting expression by $h>0$ and then letting $h\to 0^+$ while recalling that $\frac{1}{p}g^p(\cdot)$ is differentiable at $\tau\in\mathscr{D}$ and that $\varepsilon>0$ is arbitrary, we obtain the differential inequality

$$g^{p-1}(\tau)\dot{g}(\tau)\leq\int_{\mathbb{R}^{2d}}\langle v_\tau(x),j_p(x-y)\rangle\mathrm{d}\gamma_\tau(x,y) \tag{4.13}$$

which holds for every time $\tau\in(t,t+\delta)\cap\mathscr{T}\cap\mathscr{D}$.

Our goal in what follows is to use differential estimate derived in (4.13) to show that $g(\cdot)$ vanishes identically on $(t, t+\delta)$. By inserting crossed terms in the latter expression, one can easily check that

$$\begin{aligned} g^{p-1}(\tau)\dot{g}(\tau) \leq & \int_{\mathbb{R}^{2d}} \langle v_\tau(x) - v_\tau(y), j_p(x-y)\rangle \mathrm{d}\gamma_\tau(x,y) \\ & + \int_{\mathbb{R}^{2d}} \langle v_\tau(y) - \xi_\tau(y), j_p(x-y)\rangle \mathrm{d}\gamma_\tau(x,y) \\ & + \int_{\mathbb{R}^{2d}} \langle \xi_\tau(y), j_p(x-y)\rangle \mathrm{d}\gamma_\tau(x,y) \\ \leq & \, l(\tau) g^p(\tau) + \int_{\mathbb{R}^{2d}} \langle v_\tau(y) - \xi_\tau(y), j_p(x-y)\rangle \mathrm{d}\gamma_\tau(x,y) \end{aligned} \tag{4.14}$$

where we resorted to (4.7) and Hypothesis **(CI)**-(iii), as well as to the elementary observation that

$$\int_{\mathbb{R}^{2d}} |x-y||j_p(x-y)| \mathrm{d}\gamma_\tau(x,y) = \int_{\mathbb{R}^{2d}} |x-y|^p \mathrm{d}\gamma_\tau(x,y) = g^p(\tau)$$

which follows from (4.4) together with the fact that $\gamma_\tau \in \Gamma_o(\mu_\tau, \nu_\tau)$. Recall now that as a consequence of Hypothesis **(CI)**-(iv), there exists for each $w_\tau \in V(\tau, \nu_\tau)$ an element $v_\tau \in V(\tau, \mu_\tau)$ such that

$$\mathsf{d}_{\sup}(v_\tau, w_\tau) \leq L(\tau) W_p(\mu_\tau, \nu_\tau).$$

This, together with the fact that the estimates in (4.14) hold for every $v_\tau \in V(\tau, \mu_\tau)$, further yields

$$g^{p-1}(\tau)\dot{g}(\tau) \leq \Big(l(\tau) + L(\tau)\Big) g^p(\tau) + \int_{\mathbb{R}^{2d}} \langle w_\tau(y) - \xi_\tau(y), j_p(x-y)\rangle \mathrm{d}\gamma_\tau(x,y) \tag{4.15}$$

for all times $\tau \in (t, t+\delta) \cap \mathscr{T} \cap \mathscr{D}$, each $\xi_\tau \in T_{\mathcal{Q}}(\nu_\tau)$ and every $w_\tau \in V(\tau, \nu_\tau)$. Noting that the right-hand side of the previous expression is linear and strongly continuous with respect to $\xi_\tau \in \mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d; \nu_\tau)$, one gets that (4.15) also holds for all $\xi_\tau \in \overline{\mathrm{co}}\, T_{\mathcal{Q}}(\nu_\tau)$. Thus, by choosing

$$\xi_\tau = w_\tau \in V(\tau, \nu_\tau) \cap \overline{\mathrm{co}}\, T_{\mathcal{Q}}(\nu_\tau),$$

where the intersection is nonempty as a consequence of our standing assumption (4.2), one finally has

$$\dot{g}(\tau) \leq \Big(l(\tau) + L(\tau)\Big) g(\tau)$$

for $\mathscr{L}^1$-almost every $\tau \in (t, t+\delta)$. As we assumed that $g(t) = 0$, a direct application of Grönwall's lemma yields that $g(\tau) = 0$ for all times $\tau \in (t, t+\delta)$, thus leading to a contradiction.

Step 2 – Existence of a viable curve. In the first step of the proof, we have shown without loss of generality that

$$\mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}\Big(\mathcal{R}_{(\tau,t)}(\mu_\tau)\,;\mathcal{Q}_R\Big) = 0, \tag{4.16}$$

which equivalently means that $\mathcal{R}_{(\tau,t)}(\mu_\tau) \cap \mathcal{Q} \neq \emptyset$ for all times $0 \leq \tau \leq t \leq T - \tau$ and each $\mu_\tau \in \mathcal{R}_{(0,\tau)}(\mu^0) \cap \mathcal{Q}_R$. Given an integer $n \geq 1$, consider the following dyadic subdivision $[0,T] := \cup_{k=0}^{2^n-1}[t_k, t_{k+1}]$ of the time interval, wherein $t_k := Tk/2^n$ for $k \in \{0, \ldots, 2^n - 1\}$. By inductively leveraging (4.16) along with the semigroup property (2.10) of the reachable sets, we can build for each $n \geq 1$ a curve $\mu_n(\cdot) \in \mathcal{S}_{[0,T]}(0, \mu^0)$ that is such that

$$\mu_n(t_k) \in \mathcal{Q} \tag{4.17}$$

for each $k \in \{0, \ldots, 2^n - 1\}$. At this stage recall that, as a consequence of Proposition 2.21, the solution set $\mathcal{S}_{[0,T]}(0, \mu^0)$ is compact for the topology of uniform convergence, so that

$$\sup_{t \in [0,T]} W_p(\mu_{n_j}(t), \mu(t)) \underset{n_j \to +\infty}{\longrightarrow} 0$$

for some $\mu(\cdot) \in \mathcal{S}_{[0,T]}(\mu^0)$ and along a subsequence $(\mu_{n_j}(\cdot)) \subset \mathrm{AC}([0,T], \mathscr{P}_p(\mathbb{R}^d))$. In particular, it then follows from (4.17) that

$$\mu(\tfrac{kT}{2^m}) \in \mathcal{Q},$$

for every integer $m \geq 1$ and each $k \in \{0, \ldots, 2^m - 1\}$. From there, we conclude by a classical density argument that $\mu(t) \in \mathcal{Q}$ for all times $t \in [0,T]$. □

To complement the sufficient viability conditions stated in Theorem 4.4, we derive below necessary viability conditions which also involve the contingent cone to the constraints.

Theorem 4.6 (Necessary viability conditions for stationary constraints)**.** *Suppose that $p \in (1, +\infty)$, 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 with convex images satisfying Hypotheses* **(CI)** *and $\mathcal{Q} \subset \mathscr{P}_p(\mathbb{R}^d)$ be a proper set. Then if $\mathcal{Q}$ is viable for (4.1), it necessarily holds that*

$$V(t,\nu) \cap T_{\mathcal{Q}}(\nu) \neq \emptyset$$

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

Proof. This result is a particular case of Theorem 4.8 whose proof is detailed in Section 4.2 below. □

Lastly, we end this section by providing necessary and sufficient conditions for the invariance of a stationary constraints set, based on a geometric condition that is stronger than that of Theorem 4.4.

Theorem 4.7 (Invariance conditions for stationary constraint sets)**.** *Under the assumptions of Theorem 4.4, the set $\mathcal{Q} \subset \mathscr{P}_p(\mathbb{R}^d)$ is invariant for (4.1)* if and only if

$$V(t,\nu) \subset \overline{\mathrm{co}}\, T_{\mathcal{Q}}(\nu)$$

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

Proof. This result is a particular case of Theorems 4.8 and Theorem 4.12, whose proofs are discussed in depth in Section 4.2 below. □

4.2 The case of time-dependent constraints

As mentioned hereinabove, the viability and invariance results exposed in Theorem 4.4, Theorem 4.6 and Theorem 4.7 can be generalised to time-dependent constraint tubes $\mathcal{Q} : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$. In this setting, the relevant geometric objects on which these statements are based are the contingent cones to the graph of the constraints, which are defined for each $(\tau, \mu) \in \mathrm{Graph}(\mathcal{Q})$ as

$$T_{\mathrm{Graph}(\mathcal{Q})}(\tau,\mu) := \Big\{ (\zeta,\xi) \in \mathbb{R} \times \mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d; \mu) \text{ s.t. } \liminf_{h \to 0^+} \tfrac{1}{h} \mathrm{dist}_{[0,T]\times\mathscr{P}_p(\mathbb{R}^d)} \Big((\tau + h\zeta, (\mathrm{Id} + h\xi)_\sharp \mu) \,;\, \mathrm{Graph}(\mathcal{Q}) \Big) = 0 \Big\}.$$

By a simple adaptation of [9, Proposition 5.1.4] following [13, Sections 2.3 and 2.4], this set can be equivalently characterised as

$$T_{\mathrm{Graph}(\mathcal{Q})}(\tau,\mu) = \Big\{ (\zeta,\xi) \in \mathbb{R} \times \mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d; \mu) \text{ s.t. there exist sequences } h_i \to 0^+ \text{ and } \zeta_i \to \zeta \text{ for which } \mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)} \Big((\mathrm{Id} + h_i\xi)_\sharp \mu \,;\, \mathcal{Q}(\tau + h_i\zeta_i) \Big) = o_\tau(h_i) \Big\}. \tag{4.18}$$

In the following theorem, we begin our investigation by a discussion centered around necessary viability and invariance conditions, as these latter do not depend on the regularity of the constraints tubes.

Theorem 4.8 (Necessary viability, invariance and regularity conditions for constraints tubes)**.** *Suppose that $p \in (1, +\infty)$, 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 with convex images satisfying Hypotheses* **(CI)**, *and $\mathcal{Q} : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ be a constraints tube with proper images.*

Then if $\mathcal{Q} : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ is viable for (4.1), it must be left absolutely continuous and satisfy

$$(\{1\} \times V(t,\nu)) \cap T_{\mathrm{Graph}(\mathcal{Q})}(t,\nu) \neq \emptyset \tag{4.19}$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and each $\nu \in \mathcal{Q}(t)$. Analogously if $\mathcal{Q} : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ is invariant for (4.1), it necessarily holds that

$$(\{1\} \times V(t,\nu)) \subset T_{\mathrm{Graph}(\mathcal{Q})}(t,\nu) \tag{4.20}$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and each $\nu \in \mathcal{Q}(t)$.

Proof. Let us start by showing the necessity of (4.19) when $\mathcal{Q} : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ is viable for (4.1). Let $\mathscr{T} \subset (0,T)$ be the set of full $\mathscr{L}^1$-measure over which the statements of Theorem 3.1 and Theorem 3.2 as well as Hypotheses **(CI)**-(ii), (iii) and (iv) hold. Fix some $\tau \in \mathscr{T}$, an element $\mu_\tau \in \mathcal{Q}(\tau)$, a sequence $h_i \to 0^+$ and a viable curve $\mu(\cdot) \in \mathcal{S}_{[\tau,T]}(\tau,\mu_\tau)$. By Theorem 3.2, there exists for each $\varepsilon > 0$ a velocity $v_\tau^\varepsilon \in V(\tau,\mu_\tau)$ such that

$$W_p\Big(\mu(\tau + h_{i_k}^\varepsilon), (\mathrm{Id} + h_{i_k}^\varepsilon v_\tau^\varepsilon)_\sharp \mu_\tau\Big) \leq \varepsilon h_{i_k}^\varepsilon$$

along an adequate subsequence $h_{i_k}^\varepsilon \to 0^+$. Observe that by Theorem 2.13 and our choice of $\tau \in \mathscr{T}$, the set $V(\tau,\mu_\tau) \subset C^0(\mathbb{R}^d,\mathbb{R}^d)$ is compact for the topology induced by $\mathsf{d}_{cc}(\cdot,\cdot)$. In particular for each sequence $\varepsilon_n \to 0^+$, there exists a subsequence that we do not relabel and some $v_\tau \in V(\tau,\mu_\tau)$ for which

$$\|v_\tau - v_\tau^{\varepsilon_n}\|_{\mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\,\mu_\tau)} \underset{\varepsilon_n \to 0^+}{\longrightarrow} 0,$$

where we used Lemma 2.15. Note also that for every $\varepsilon_n > 0$, one can choose $\delta_n := h_{i_n}^{\varepsilon_n}$ in such a way that $o_{\tau,\varepsilon_n}(h_{i_n}^{\varepsilon_n}) \leq \varepsilon_n h_{i_n}^{\varepsilon_n}$. Thus, recalling that $\mu(t) \in \mathcal{Q}(t)$ for all times $t \in [\tau,T]$, one further has that

$$\begin{aligned}
&\mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}\Big((\mathrm{Id} + \delta_n v_\tau)_\sharp \mu_\tau\,;\mathcal{Q}(\tau+\delta_n)\Big)\\
&\quad\leq \mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}\Big((\mathrm{Id} + \delta_n v_\tau^{\varepsilon_n})_\sharp \mu_\tau\,;\mathcal{Q}(\tau+\delta_n)\Big) + W_p\Big((\mathrm{Id} + \delta_n v_\tau^{\varepsilon_n})_\sharp \mu_\tau, (\mathrm{Id} + \delta_n v_\tau)_\sharp \mu_\tau\Big)\\
&\quad\leq W_p\Big(\mu(\tau+\delta_n), (\mathrm{Id} + \delta_n v_\tau^{\varepsilon_n})_\sharp \mu_\tau\Big) + \delta_n\,\|v_\tau - v_\tau^{\varepsilon_n}\|_{\mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\,\mu_\tau)}\\
&\quad\leq \delta_n\Big(\varepsilon_n + \|v_\tau - v_\tau^{\varepsilon_n}\|_{\mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\,\mu_\tau)}\Big),
\end{aligned}$$

which in turn implies

$$\liminf_{\delta_n \to 0^+} \tfrac{1}{\delta_n}\mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}\Big((\mathrm{Id} + \delta_n v_\tau)_\sharp \mu_\tau\,;\mathcal{Q}(\tau+\delta_n)\Big) = 0.$$

By (4.18), this is tantamount to the fact that $(1, v_\tau) \in T_{\mathrm{Graph}(\mathcal{Q})}(\tau,\mu_\tau)$, and thus yields (4.19).

Suppose now that $\mathcal{Q} : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ is invariant for (4.1), fix an arbitrary $v_\tau \in V(\tau,\mu_\tau)$, and observe that by Theorem 3.1, there exists for every $\varepsilon > 0$ some $h_\varepsilon > 0$ and a curve $\mu_\varepsilon(\cdot) \in \mathcal{S}_{[\tau,T]}(\tau,\mu_\tau)$ such that

$$W_p\Big(\mu_\varepsilon(\tau+h), (\mathrm{Id} + h v_\tau)_\sharp \mu_\tau\Big) \leq \varepsilon h$$

for all $h \in [0,h_\varepsilon]$. This estimate, combined with the fact that $\mu_\varepsilon(t) \in \mathcal{Q}(t)$ for all times $t \in [\tau,T]$ owing to the invariance of the tube, implies that

$$\begin{aligned}
\mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}\Big((\mathrm{Id} + h v_\tau)_\sharp \mu_\tau\,;\mathcal{Q}(\tau+h)\Big) &\leq W_p\Big(\mu_\varepsilon(\tau+h), (\mathrm{Id} + h v_\tau)_\sharp \mu_\tau\Big)\\
&\leq \varepsilon h.
\end{aligned}$$

Thus, dividing by $h > 0$ and letting $h \to 0^+$ while recalling that $\varepsilon > 0$ is arbitrary, we finally obtain

$$\liminf_{h \to 0^+} \tfrac{1}{h}\mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}\Big((\mathrm{Id} + h v_\tau)_\sharp \mu_\tau\,;\mathcal{Q}(\tau+h)\Big) = 0,$$

which equivalently means that $(1, v_\tau) \in T_{\mathrm{Graph}(\mathcal{Q})}(\tau,\mu_\tau)$ for all $v_\tau \in V(\tau,\mu_\tau)$, and thus yields (4.20).

Let us finally prove that if $\mathcal{Q} : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ is viable for (4.1), then it is left absolutely continuous. To do so, fix $\mu \in \mathscr{P}_p(\mathbb{R}^d)$ and $R > 0$ in such a way that

$$\mathcal{Q}(\tau) \cap \mathbb{B}_{\mathscr{P}_p(\mathbb{R}^d)}(\mu, R) \neq \emptyset.$$

Then, since $\mathcal{Q} : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ is viable, there exists for all $\tau \in [0,T]$ and each $\mu_\tau \in \mathcal{Q}(\tau) \cap \mathbb{B}_{\mathscr{P}_p(\mathbb{R}^d)}(\mu, R)$ a curve $\mu(\cdot) \in \mathcal{S}_{[\tau,T]}(\tau,\mu_\tau)$ such that $\mu(t) \in \mathcal{Q}(t)$ for all times $t \in [\tau,T]$. Besides by

Proposition 2.20, there exists a constant $\mathcal{C}_{\mu,R} > 0$ depending only on the magnitudes of $p, \mathcal{M}_p(\mu), R$ and $\|m(\cdot)\|_1$ such that

$$W_p(\mu_\tau, \mu(t)) \leq (1 + \mathcal{C}_{\mu,R}) \int_\tau^t m(s)\mathrm{d}s$$

for all times $t \in [\tau, T]$. Thus, noting by construction that for any $0 \leq \tau \leq t \leq T$, it holds that

$$\begin{aligned}
\Delta_{\mu,R}(\mathcal{Q}(\tau), \mathcal{Q}(t)) &\leq \sup\Big\{\mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}(\mu_\tau\,;\mathcal{Q}(t)) \text{ s.t. } \mu_\tau \in \mathcal{Q}(\tau) \cap \mathbb{B}_{\mathscr{P}_p(\mathbb{R}^d)}(\mu, R)\Big\} \\
&\leq \sup\Big\{W_p(\mu_\tau, \mu(t)) \text{ s.t. } \mu_\tau \in \mathcal{Q}(\tau) \cap \mathbb{B}_{\mathscr{P}_p(\mathbb{R}^d)}(\mu, R) \text{ and } \mu(\cdot) \in \mathcal{S}_{[\tau,T]}(\tau, \mu_\tau) \\
&\qquad\qquad \text{satisfies } \mu(t) \in \mathcal{Q}(t)\Big\} \\
&\leq (1 + \mathcal{C}_{\mu,R}) \int_\tau^t m(s)\mathrm{d}s,
\end{aligned}$$

we can conclude that $\mathcal{Q} : [0, T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ is left absolutely continuous. □

Remark 4.9 (On the role of left absolute continuity)**.** *It is worth noting that in the previous theorem, we have shown that being viable for* (4.1) *under Hypotheses* **(CI)** *entails the left absolute continuity of the constraint tube. This supports the fact that this regularity framework – for which we provide sufficient viability conditions in Theorem 4.12 – appears quite naturally when studying Cauchy-Lipschitz continuity inclusions with state-constraints.*

In the next theorem, we provide sufficient viability conditions for absolutely continuous constraint tubes, which are the natural generalisation of Theorem 4.4 to the time-dependent setting.

Theorem 4.10 (Sufficient viability conditions for absolutely continuous constraints tubes)**.** *Suppose that $p \in (1, +\infty)$, 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 with convex images satisfying Hypotheses* **(CI)** *and $\mathcal{Q} : [0, T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ be an absolutely continuous tube with proper images such that*

$$(\{1\} \times V(t, \nu)) \cap \overline{\mathrm{co}}\, T_{\mathrm{Graph}(\mathcal{Q})}(t, \nu) \neq \emptyset \tag{4.21}$$

for $\mathscr{L}^1$-almost every $t \in [0, T]$ and all $\nu \in \mathcal{Q}(t)$. Then $\mathcal{Q} : [0, T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ is viable for (4.1)*.*

Proof. As in the proof of Theorem 4.4, we assume without loss of generality that $(\tau, \mu_\tau) = (0, \mu^0)$ for some $\mu^0 \in \mathcal{Q}(0)$. From there on, the arguments will essentially follow along the same line as those of Theorem 4.4, in which one aims at showing that the distance function, defined by

$$g(t) := \mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}\Big(\mathcal{R}_{(0,t)}(\mu^0)\,;\mathcal{Q}(t)\Big)$$

for all times $t \in [0, T]$, is identically equal to 0.

In what follows, we let $\mathscr{T} \subset (0, T)$ be the set of full $\mathscr{L}^1$-measure on which the statements of Theorem 3.1 and Theorem 3.2 as well as Hypotheses **(CI)**-(ii), (iii) and (iv) hold. Observing that $g(\cdot) \in \mathrm{AC}([0, T], \mathbb{R}_+)$ by Proposition 2.8 and Lemma 4.5 , we also denote by $\mathscr{D} \subset (0, T)$ the subset of full $\mathscr{L}^1$-measure where it is differentiable. Moreover, owing to the absolute continuity of $\mathcal{Q} : [0, T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ and to the fact that $\cup_{t\in[0,T]}\mathcal{R}_{(0,t)}(\mu^0)$ is compact by Proposition 2.21 and Lemma 4.5, one can find some radius $R > 0$ satisfying

$$\mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}\Big(\mathcal{R}_{(0,t)}(\mu^0)\,;\partial\mathbb{B}_{\mathscr{P}_p(\mathbb{R}^d)}(\mu^0, R)\Big) \geq \mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}\Big(\mathcal{R}_{(0,t)}(\mu^0)\,;\mathcal{Q}(t)\Big) + 1 \tag{4.22}$$

for all times $t \in [0, T]$, by following e.g. the arguments detailed in Appendix B below. Note that by construction, the sets $\mathcal{Q}_R(t) := \mathcal{Q}(t) \cap \mathbb{B}_{\mathscr{P}_p(\mathbb{R}^d)}(\mu^0, R)$ are nonempty and such that

$$g(t) = \mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}\Big(\mathcal{R}_{(0,t)}(\mu^0)\,;\mathcal{Q}_R(t)\Big)$$

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

Step 1 – Local variations of the distance along contingent directions. Suppose by contradiction that there exist some $t \in [0,T)$ along with a small $\delta > 0$ such that $g(t) = 0$ and $g(\tau) > 0$ for all times $\tau \in (t, t+\delta)$, and fix an arbitrary element $\tau \in (t, t+\delta) \cap \mathscr{T} \cap \mathscr{D}$. Since $\mathcal{R}_{(0,\tau)}(\mu^0)$ and $\mathcal{Q}_R(\tau)$ are both compact because $\mathcal{Q}(\tau)$ is proper, one has that

$$g(\tau) = W_p(\mu_\tau, \nu_\tau)$$

for some $\mu_\tau \in \mathcal{R}_{(0,\tau)}(\mu^0)$ and $\nu_\tau \in \mathcal{Q}_R(\tau)$. By (4.18) along with the estimate (4.22) imposed on $R > 0$, there exists for every $(\zeta_\tau, \xi_\tau) \in T_{\mathrm{Graph}(\mathcal{Q})}(\tau, \nu_\tau)$ two sequences $h_i \to 0^+$ and $\zeta_\tau^i \to \zeta_\tau$ for which

$$\mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}\Big((\mathrm{Id} + h_i\xi_\tau)_\sharp \nu_\tau \,;\, \mathcal{Q}_R(\tau + h_i\zeta_\tau^i)\Big) = o_\tau(h_i).$$

This allows us to estimate from above the variation of $g^p(\cdot)$ around τ as

$$\begin{aligned}\tfrac{1}{p}g^p(\tau + h_i\zeta_\tau^i) - \tfrac{1}{p}g^p(\tau) &= \tfrac{1}{p}\mathrm{dist}^p_{\mathscr{P}_p(\mathbb{R}^d)}\Big(\mathcal{R}_{(0,\tau+h_i\zeta_\tau^i)}(\mu^0)\,;\,\mathcal{Q}_R(\tau + h_i\zeta_\tau^i)\Big) - \tfrac{1}{p}W_p^p(\mu_\tau,\nu_\tau)\\ &\le \tfrac{1}{p}\mathrm{dist}^p_{\mathscr{P}_p(\mathbb{R}^d)}\Big(\mathcal{R}_{(0,\tau+h_i\zeta_\tau^i)}(\mu^0)\,;\,(\mathrm{Id} + h_i\xi_\tau)_\sharp\nu_\tau\Big) - \tfrac{1}{p}W_p^p(\mu_\tau,\nu_\tau) + o_\tau(h_i)\end{aligned} \tag{4.23}$$

for any $h_i > 0$ that is sufficiently small.

In order to extract further information from (4.23), we need to discriminate between two possible scenarios depending on the asymptotic behaviour of the sequence $(\zeta_\tau^i) \subset \mathbb{R}$. If there exists a subsequence $i_k \to +\infty$ for which $\zeta_\tau^{i_k} \ge 0$, it follows from Theorem 3.1 that for every $\varepsilon > 0$ and any $v_\tau \in V(\tau, \mu_\tau)$, there exists a curve $\mu_\varepsilon(\cdot) \in \mathcal{S}_{[\tau,T]}(\tau, \mu_\tau)$ such that

$$W_p\Big(\mu_\varepsilon(\tau + h_{i_k}\zeta_\tau^{i_k}), (\mathrm{Id} + h_{i_k}\zeta_\tau^{i_k}v_\tau)_\sharp\mu_\tau\Big) \le \varepsilon h_{i_k}\zeta_\tau^{i_k}. \tag{4.24}$$

On the other hand, if $\zeta_\tau^i < 0$ for all large $i \ge 1$, we can apply Theorem 3.2 to obtain for each curve $\mu(\cdot) \in \mathcal{S}_{[0,T]}(\tau, \mu_\tau)$ the existence of an element $v_\tau^\varepsilon \in V(\tau, \mu_\tau)$ and of a subsequence $i_k \to +\infty$, both depending on $\varepsilon > 0$, for which

$$W_p\Big(\mu(\tau + h_{i_k}\zeta_\tau^{i_k}), (\mathrm{Id} + h_{i_k}\zeta_\tau^{i_k}v_\tau^\varepsilon)_\sharp\mu_\tau\Big) \le \varepsilon h_{i_k}|\zeta_\tau^{i_k}|. \tag{4.25}$$

Thus by combining (4.24) and (4.25), one may assert that there exist curves $\mu_\varepsilon(\cdot) \in \mathcal{S}_{[0,T]}(\tau, \mu_\tau)$ along with two subsequences $h_{i_k} \to 0^+$ and $\zeta_\tau^{i_k} \to \zeta_\tau$, all possibly depending on $\varepsilon > 0$, such that

$$W_p\Big(\mu_\varepsilon(\tau + h_{i_k}\zeta_\tau^{i_k}), (\mathrm{Id} + h_{i_k}\zeta_\tau^{i_k}v_\tau^\varepsilon)_\sharp\mu_\tau\Big) \le \varepsilon h_{i_k}|\zeta_\tau^{i_k}| \quad \text{for} \quad \begin{cases}\text{every } v_\tau^\varepsilon \in V(\tau,\mu_\tau) & \text{if } \zeta_\tau^{i_k} \ge 0,\\ \text{some } v_\tau^\varepsilon \in V(\tau,\mu_\tau) & \text{if } \zeta_\tau^{i_k} < 0,\end{cases} \tag{4.26}$$

when $i_k \ge 1$ is large enough. Furthermore, observing that

$$\mu_\varepsilon(\tau + h_{i_k}\zeta_\tau^{i_k}) \in \mathcal{R}_{(0,\tau+h_{i_k}\zeta_\tau^{i_k})}(\mu^0),$$

one may refine the estimate of (4.23) using the curves $\mu_\varepsilon(\cdot)$ as follows

$$\tfrac{1}{p}g^p(\tau + h_{i_k}\zeta_\tau^{i_k}) - \tfrac{1}{p}g^p(\tau) \le \tfrac{1}{p}W_p^p\Big(\mu_\varepsilon(\tau + h_{i_k}\zeta_\tau^{i_k}), (\mathrm{Id} + h_{i_k}\xi_\tau)_\sharp\nu_\tau\Big) - \tfrac{1}{p}W_p^p(\mu_\tau,\nu_\tau) + o_\tau(h_{i_k}). \tag{4.27}$$

At this stage, upon noting that $v_\tau \in \mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d; \mu_\tau)$ by Hypothesis **(CI)**-(ii) while $\xi_\tau \in \mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d; \nu_\tau)$ by definition, and assuming without loss of generality that $\varepsilon \in (0,1]$ and $h_{i_k} \in (0,1]$, one may reproduce the computations of (4.10) in the proof of Theorem 4.4 and combine them with (4.26) to obtain that

$$\begin{aligned}&\tfrac{1}{p}W_p^p\Big(\mu_\varepsilon(\tau + h_{i_k}\zeta_\tau^{i_k}), (\mathrm{Id} + h_{i_k}\xi_\tau)_\sharp\nu_\tau\Big)\\ &\qquad - \tfrac{1}{p}W_p^p\Big((\mathrm{Id} + h_{i_k}\zeta_\tau^{i_k}v_\tau^\varepsilon)_\sharp\mu_\tau, (\mathrm{Id} + h_{i_k}\xi_\tau)_\sharp\nu_\tau\Big) \le \varepsilon h_{i_k} + o_{\tau,\varepsilon}(h_{i_k})\end{aligned} \tag{4.28}$$

for $h_{i_k} > 0$ small enough, and up to rescaling $\varepsilon > 0$ by a constant since the sequence $(\zeta_\tau^{i_k}) \subset \mathbb{R}$ is bounded. Whence, by merging (4.27) and (4.28), it then holds that

$$\begin{aligned}\tfrac{1}{p}g^p(\tau + h_{i_k}\zeta_\tau^{i_k}) - \tfrac{1}{p}g^p(\tau) \leq \tfrac{1}{p}W_p^p\Big((\mathrm{Id} + h_{i_k}\zeta_\tau^{i_k}v_\tau^\varepsilon)_\sharp\mu_\tau, (\mathrm{Id} + h_{i_k}\xi_\tau)_\sharp\nu_\tau\Big) \\ - \tfrac{1}{p}W_p^p(\mu_\tau,\nu_\tau) + \varepsilon h_{i_k} + o_{\tau,\varepsilon}(h_{i_k}),\end{aligned} \tag{4.29}$$

for $i_k \geq 1$ sufficiently large, with $v_\tau^\varepsilon \in V(\tau,\mu_\tau)$ being either fixed or arbitrary depending on the asymptotic behaviour of $(\zeta_\tau^{i_k}) \subset \mathbb{R}$. One may then apply the joint superdifferentiability inequality of Proposition 2.3 to obtain

$$\begin{aligned}\tfrac{1}{p}W_p^p\Big((\mathrm{Id} + h_{i_k}\zeta_\tau^{i_k}v_\tau^\varepsilon)_\sharp\mu_\tau, (\mathrm{Id}+h_{i_k}\xi_\tau)_\sharp\nu_\tau\Big) - \tfrac{1}{p}W_p^p(\mu_\tau,\nu_\tau) \\ \leq h_{i_k}\int_{\mathbb{R}^{2d}}\langle\zeta_\tau^{i_k}v_\tau^\varepsilon(x) - \xi_\tau(y), j_p(x-y)\rangle\mathrm{d}\gamma_\tau(x,y) + o_{\tau,\varepsilon}(h_{i_k}),\end{aligned} \tag{4.30}$$

for each $\gamma_\tau \in \Gamma_o(\mu_\tau,\nu_\tau)$, where we used the analytical expressions (2.3)-(2.4) of the remainder term, along with the fact that $(\zeta_\tau^{i_k}) \subset \mathbb{R}$ is bounded. In turn, by combining (4.29) and (4.30), letting $i_k \to +\infty$ and recalling that $\frac{1}{p}g^p(\cdot)$ is differentiable at $\tau \in \mathscr{D}$, one finally gets

$$\zeta_\tau\, g^{p-1}(\tau)\dot{g}(\tau) \leq \int_{\mathbb{R}^{2d}}\langle\zeta_\tau v_\tau^\varepsilon(x) - \xi_\tau(y), j_p(x-y)\rangle\mathrm{d}\gamma_\tau(x,y) + \varepsilon \quad \text{for} \quad \begin{cases}\text{every } v_\tau^\varepsilon \in V(\tau,\mu_\tau) & \text{if } \zeta_\tau \geq 0,\\ \text{some } v_\tau^\varepsilon \in V(\tau,\mu_\tau) & \text{if } \zeta_\tau < 0,\end{cases} \tag{4.31}$$

where $(\zeta_\tau,\xi_\tau) \in T_{\mathrm{Graph}(\mathcal{Q})}(\tau,\nu_\tau)$ and $\gamma_\tau \in \Gamma_o(\mu_\tau,\nu_\tau)$ are arbitrary while $v_\tau^\varepsilon \in V(\tau,\mu_\tau)$ may possibly depend on $\zeta_\tau \in \mathbb{R}$ as well as on the free parameter $\varepsilon > 0$.

Step 2 – Convexification of the contingent directions and viability. In this second step, we show how one can convexify the contingent directions in (4.31) and then prove the existence of viable curves. With this goal in mind, we draw inspiration from [63, Lemma 4.9] and consider arbitrary collections of $N \geq 1$ elements

$$(\zeta_\tau^j,\xi_\tau^j) \in T_{\mathrm{Graph}(\mathcal{Q})}(\tau,\nu_\tau) \qquad \text{and} \qquad \alpha_j \in [0,1]$$

indexed by $j \in \{1,\dots,N\}$, which are chosen in such a way that

$$\sum_{j=1}^N \alpha_j = 1 \qquad \text{and} \qquad \zeta_\tau := \sum_{j=1}^N \alpha_j\zeta_\tau^j > 0. \tag{4.32}$$

Up to reordering the labels, we may posit that there exists an $m \in \{1,\dots,N\}$ such that $\zeta_\tau^j \geq 0$ if $j \geq m$ and $\zeta_\tau^j < 0$ otherwise. By applying (4.31) to each $(\zeta_\tau^j,\xi_\tau^j) \in T_{\mathrm{Graph}(\mathcal{Q})}(\tau,\nu_\tau)$ with $j \in \{1,\dots,N\}$, and then summing the resulting expressions depending on whether $j < m$ or $j \geq m$, one has that

$$\sum_{j<m}\alpha_j\zeta_\tau^j\, g^{p-1}(\tau)\dot{g}(\tau) \leq \sum_{j<m}\alpha_j\Big(\int_{\mathbb{R}^{2d}}\langle\zeta_\tau^j v_\tau^{\varepsilon,j}(x) - \xi_\tau^j(y), j_p(x-y)\rangle\mathrm{d}\gamma_\tau(x,y) + \varepsilon\Big), \tag{4.33}$$

for some fixed (and potentially empty) tuple $(v_\tau^{\varepsilon,j})_{j<m} \in V(\tau,\mu_\tau)^{m-1}$, as well as

$$\sum_{j\geq m}\alpha_j\zeta_\tau^j\, g^{p-1}(\tau)\dot{g}(\tau) \leq \sum_{j\geq m}\alpha_j\Big(\int_{\mathbb{R}^{2d}}\langle\zeta_\tau^j v_\tau(x) - \xi_\tau^j(y), j_p(x-y)\rangle\mathrm{d}\gamma_\tau(x,y) + \varepsilon\Big) \tag{4.34}$$

for any element $v_\tau \in V(\tau,\mu_\tau)$. Introducing in turn the coefficients

$$\beta_j := \frac{\alpha_j|\zeta_\tau^j|}{\sum_{j\geq m}\alpha_j\zeta_\tau^j} \in (0,1) \ \text{ for each } j < m \qquad \text{and} \qquad \beta := 1 - \sum_{j<m}\beta_j,$$

which are well-defined as a consequence of (4.32), while recalling that set $V(\tau,\mu_\tau)$ is convex by assumption, it holds for each $v_\tau \in V(\tau,\nu_\tau)$ that

$$v_\tau' := \beta v_\tau + \sum_{j<m}\beta_j v_\tau^{\varepsilon,j} \in V(\tau,\nu_\tau).$$

Whence, by merging the estimate of (4.33) and that of (4.34) evaluated at $v'_\tau \in V(\tau, \mu_\tau)$ defined via the previous expression, one eventually obtains that

$$\zeta_\tau\, g^{p-1}(\tau)\dot{g}(\tau) \leq \int_{\mathbb{R}^{2d}} \langle \zeta_\tau v_\tau(x) - \xi_\tau(y), j_p(x-y)\rangle \mathrm{d}\gamma_\tau(x,y) + \varepsilon \tag{4.35}$$

for any given $(\zeta_\tau, \xi_\tau) \in \mathrm{co}T_{\mathrm{Graph}(\mathcal{Q})}(\tau, \nu_\tau)$ satisfying $\zeta_\tau > 0$ and every $v_\tau \in V(\tau, \mu_\tau)$. Remarking that the right-hand side in (4.35) is linear and continuous with respect $(\zeta_\tau, \xi_\tau) \in \mathbb{R} \times \mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d; \nu_\tau)$, the latter expression remains valid for every $(\zeta_\tau, \xi_\tau) \in \overline{\mathrm{co}}\, T_{\mathrm{Graph}(\mathcal{Q})}(\tau, \nu_\tau)$ such that $\zeta_\tau > 0$.

At this stage, starting from (4.35), one may repeat the argument discussed at the end of Step 1 in the proof of Theorem 4.4 while using the facts that $\varepsilon > 0$ is arbitrary and $\gamma_\tau \in \Gamma_o(\mu_\tau, \nu_\tau)$ to show that the latter estimate further yields

$$\zeta_\tau\, g^{p-1}(\tau)\dot{g}(\tau) \leq \zeta_\tau\Big(l(\tau) + L(\tau)\Big)g^p(\tau) + \int_{\mathbb{R}^{2d}} \langle \zeta_\tau w_\tau(y) - \xi_\tau(y), j_p(x-y)\rangle \mathrm{d}\gamma_\tau(x,y),$$

for every $(\zeta_\tau, \xi_\tau) \in \overline{\mathrm{co}}\, T_{\mathrm{Graph}(\mathcal{Q})}(\tau, \nu_\tau)$ such that $\zeta_\tau > 0$ and each $w_\tau \in V(\tau, \nu_\tau)$. Choosing in particular

$$(\zeta_\tau, \xi_\tau) = (1, w_\tau) \in (\{1\} \times V(\tau, \nu_\tau)) \cap \overline{\mathrm{co}}\, T_{\mathrm{Graph}(\mathcal{Q})}(\tau, \nu_\tau),$$

which is licit under our standing assumption (4.21), one finally gets that

$$\dot{g}(\tau) \leq \Big(l(\tau) + L(\tau)\Big)g(\tau)$$

for all times $\tau \in (t, t+\delta) \cap \mathscr{T} \cap \mathscr{D}$. Noting that $g(t) = 0$ and $\mathscr{T}, \mathscr{D} \subset (0,T)$ both have full $\mathscr{L}^1$-measure, it follows from Grönwall's lemma that $g(\tau) = 0$ for all $\tau \in [t, t+\delta)$, which implies that $g : [0,T] \to \mathbb{R}_+$ is identically equal to zero and thus leads to a contradiction. One can then deduce the existence of a viable curve by repeating the argument detailed above in Step 2 of the proof of Theorem 4.4. □

In the following theorem, we state the natural counterpart of the sufficient implication of the invariance result of Theorem 4.7 for absolutely continuous time-dependent constraint sets.

Theorem 4.11 (Sufficient invariance conditions for absolutely continuous constraints tubes)**.** *Suppose that the assumptions of Theorem 4.10 hold and that the tube $\mathcal{Q} : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ is such that*

$$(\{1\} \times V(t, \nu)) \subset \overline{\mathrm{co}}\, T_{\mathrm{Graph}(\mathcal{Q})}(t, \nu) \tag{4.36}$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and all $\nu \in \mathcal{Q}(t)$. Then $\mathcal{Q} : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ is invariant for (4.1).

Proof. In what follows, we let $\mathscr{T} \subset (0,T)$ and $R > 0$ be given as in the proof of Theorem 4.10 above, and assume without loss of generality that $(\tau, \mu_\tau) = (0, \mu^0)$ for some $\mu^0 \in \mathcal{Q}(0)$. Given an arbitrary curve $\mu(\cdot) \in \mathcal{S}_{[0,T]}(\mu^0)$, our goal is to show that the distance function, defined by

$$g(t) := \mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}(\mu(t)\,; \mathcal{Q}(t))$$

for all times $t \in [0,T]$, is identically equal to zero. Note that since $\mathcal{Q} : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ is absolutely continuous, it can be easily verified that $g(\cdot) \in \mathrm{AC}([0,T], \mathbb{R}_+)$, and we denote by $\mathscr{D} \subset (0,T)$ the subset of full $\mathscr{L}^1$-measure over which it is differentiable. We posit by contradiction that $g(t) = 0$ for some $t \in [0,T]$ and that there exists $\delta > 0$ such that $g(\tau) > 0$ for all $\tau \in (t, t+\delta)$. Observe now that by the compactness of $\mathcal{Q}_R(\tau)$, there exists an element $\nu_\tau \in \mathcal{Q}_R(\tau)$ such that

$$g(\tau) = W_p(\mu(\tau), \nu_\tau).$$

Besides, owing to the choice of $R > 0$ made via (4.3), it holds for every pair of contingent directions $(\zeta_\tau, \xi_\tau) \in T_{\mathrm{Graph}(\mathcal{Q})}(\tau, \nu_\tau)$ that

$$\mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}\Big((\mathrm{Id} + h_i\xi_\tau)_\sharp \nu_\tau\,; \mathcal{Q}_R(\tau + h_i\zeta_\tau^i)\Big) = o_\tau(h_i)$$

along two given sequences $h_i \to 0^+$ and $\zeta_\tau^i \to \zeta_\tau$. By Theorem 3.2, there exists for every $\varepsilon > 0$ some $v_\tau^\varepsilon \in V(\tau, \mu(\tau))$ such that

$$W_p\Big(\mu(\tau + h_{i_k}^\varepsilon \zeta_\tau^{i_k}), (\mathrm{Id} + h_{i_k}^\varepsilon \zeta_\tau^{i_k} v_\tau^\varepsilon)_\sharp \mu(\tau)\Big) \leq \varepsilon h_{i_k}^\varepsilon |\zeta_\tau^{i_k}| + o_{\tau,\varepsilon}(h_{i_k}^\varepsilon)$$

along a subsequence $h_{i_k}^\varepsilon \to 0^+$. By following the arguments leading to the differential inequality (4.31) on $\frac{1}{p} g^p(\cdot)$ in the proof of Theorem 4.10, one can derive the estimate

$$\zeta_\tau \, g^{p-1}(\tau)\dot{g}(\tau) \leq \int_{\mathbb{R}^{2d}} \langle \zeta_\tau v_\tau^\varepsilon(x) - \xi_\tau(y), j_p(x-y)\rangle \mathrm{d}\gamma_\tau(x,y) + \varepsilon$$

for all $\gamma_\tau \in \Gamma_o(\mu(\tau), \nu_\tau)$, up to rescaling $\varepsilon > 0$. Since the latter expression is linear and continuous with respect to $(\zeta_\tau, \xi_\tau) \in T_{\mathrm{Graph}(\mathcal{Q})}(\tau, \nu_\tau)$, it holds more generally for elements of $\overline{\mathrm{co}}\, T_{\mathrm{Graph}(\mathcal{Q})}(\tau, \nu_\tau)$. There now remains to observe that, by Hypothesis **(CI)**-(iv), there exists $w_\tau^\varepsilon \in V(\tau, \nu_\tau)$ such that

$$\mathsf{d}_{\mathrm{sup}}(v_\tau^\varepsilon, w_\tau^\varepsilon) \leq L(\tau) W_p(\mu(\tau), \nu_\tau),$$

which together with Hypothesis **(CI)**-(iii), the definition (2.2) of the duality map $j_p : \mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d; \mu) \to \mathcal{L}^q(\mathbb{R}^d, \mathbb{R}^d; \mu)$ and the fact that $\gamma_\tau \in \Gamma_o(\mu(\tau), \nu_\tau)$ yields the differential estimate

$$\zeta_\tau \, g^{p-1}(\tau)\dot{g}(\tau) \leq \int_{\mathbb{R}^{2d}} \langle \zeta_\tau w_\tau^\varepsilon(y) - \xi_\tau(y), j_p(x-y)\rangle \mathrm{d}\gamma_\tau(x,y) + \zeta_\tau\Big(l(\tau) + L(\tau)\Big) g^p(\tau) + \varepsilon,$$

which is valid whenever $\zeta_\tau > 0$. Thence, choosing in particular

$$(\zeta_\tau, \xi_\tau) = (1, w_\tau^\varepsilon) \in (\{1\} \times V(\tau, \nu_\tau)) \subset \overline{\mathrm{co}}\, T_{\mathrm{Graph}(\mathcal{Q})}(\tau, \nu_\tau),$$

which is licit under our standing assumption (4.36), one can deduce that

$$\dot{g}(\tau) \leq \Big(l(\tau) + L(\tau)\Big) g(\tau) + \varepsilon,$$

which finally yields $g(\tau) = 0$ for each $\tau \in (t, t+\delta)$ by applying Grönwall's lemma while noting that $\varepsilon > 0$ is arbitrary. This contradicts our initial choice of $t \in [0, T]$. □

In the following theorem, we finally present sufficient viability and invariance conditions for constraints tubes which are merely left absolutely continuous. In order to treat this less regular case, we shall see that one must relinquish the convexification of the contingent directions which was available both in Theorem 4.4 and Theorem 4.10 above.

Theorem 4.12 (Sufficient viability and invariance conditions for left absolutely continuous tubes)**.** *Suppose that $p \in (1, +\infty)$, 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 with convex images satisfying Hypotheses* **(CI)***, and $\mathcal{Q} : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ be a left absolutely continuous constraints tube with proper images such that*

$$(\{1\} \times V(t,\nu)) \cap T_{\mathrm{Graph}(\mathcal{Q})}(t,\nu) \neq \emptyset \tag{4.37}$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and each $\nu \in \mathcal{Q}(t)$. Then $\mathcal{Q} : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ is viable for (4.1). Analogously, if the constraints tube satisfies the stronger condition

$$(\{1\} \times V(t,\nu)) \subset T_{\mathrm{Graph}(\mathcal{Q})}(t,\nu) \tag{4.38}$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and each $\nu \in \mathcal{Q}(t)$, then $\mathcal{Q} : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ is invariant for (4.1).

Proof. In what follows, we only prove the viability of $\mathcal{Q} : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ under the sufficient condition (4.37), as its invariance under (4.38) follows from a straightforward transposition of the method detailed previously in the proof of Theorem 4.11.

Without loss of generality, we assume that $(\tau, \mu_\tau) := (0, \mu^0)$ for some $\mu^0 \in \mathcal{Q}(0)$ and let $\mathscr{T} \subset (0,T)$ be the set of full $\mathscr{L}^1$-measure such that the statements of Theorem 3.1 as well as Hypotheses **(CI)**-(ii), (iii) and (iv) hold. Since the constraints tube $\mathcal{Q} : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ is merely left absolutely continuous, the distance function

$$g : t \in [0,T] \mapsto \mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}\Big(\mathcal{R}_{(0,t)}(\mu^0)\,;\mathcal{Q}(t)\Big)$$

is not absolutely continuous in general. To estimate its local variations, we cannot resort to the usual Grönwall approach detailed above, and need to carry out a subtler viability analysis on its epigraph. By contradiction, assume that there exist $t \in [0,T]$ and $\delta > 0$ such that $g(t) = 0$ and $g(\tau) > 0$ for $\tau \in (t, t+\delta)$. By combining the arguments detailed in the proof of Theorem 4.10 with (4.37), one can check that for each $\tau \in (t, t+\delta) \cap \mathscr{T}$, it holds that

$$\liminf_{h \to 0^+} \frac{g(\tau + h) - g(\tau)}{h} \leq \Big(l(\tau) + L(\tau)\Big) g(\tau). \tag{4.39}$$

Consider now the auxiliary real-valued tube defined by

$$\mathscr{E}(\tau) := \Big\{\alpha \in \mathbb{R}_+ \ \text{ s.t. } \alpha = g(\tau) + r \ \text{ for some } r \geq 0\Big\},$$

as well as the 1-dimensional linear vector field

$$f(\tau, \alpha) := \Big(l(\tau) + L(\tau)\Big)\alpha$$

given for all times $\tau \in [t, t+\delta]$ and each $\alpha \in \mathbb{R}$, and notice that $\mathscr{E} : [t, t+\delta] \rightrightarrows \mathbb{R}$ is left absolutely continuous by Proposition 2.8. Given some $\tau \in (t, t+\delta) \cap \mathscr{T}$, remark that if $\alpha \in \mathscr{E}(\tau)$ is such that $\alpha = g(\tau)$, it can then be deduced from (4.39) and [9, Proposition 5.1.4] that

$$f(\tau, g(\tau)) = \Big(l(\tau) + L(\tau)\Big) g(\tau) \in \Big\{\xi \in \mathbb{R} \ \text{ s.t. } (1, \xi) \in T_{\mathrm{Graph}(\mathscr{E})}(\tau, g(\tau))\Big\}. \tag{4.40}$$

On the other hand, if $\alpha \in \mathscr{E}(\tau)$ is such that $\alpha > g(\tau)$, it is then clear that

$$f(\tau, \alpha) \in \Big\{\xi \in \mathbb{R} \ \text{ s.t. } (1, \xi) \in T_{\mathrm{Graph}(\mathscr{E})}(\tau, \alpha)\Big\} = \mathbb{R}. \tag{4.41}$$

Whence, by combining (4.40) and (4.41), one then obtains

$$f(\tau, \alpha) \in \Big\{\xi \in \mathbb{R} \ \text{ s.t. } (1, \xi) \in T_{\mathrm{Graph}(\mathscr{E})}(\tau, \alpha)\Big\}$$

for $\mathscr{L}^1$-almost every $\tau \in (t, t+\delta)$ and each $\alpha \in \mathscr{E}(\tau)$. Thus, observing that $g(t) = 0 \in \mathscr{E}(t)$, the classical measurable viability theorem of [63, Theorem 4.2] yields the existence of a curve $\alpha(\cdot) \in \mathrm{AC}([t, t+\delta], \mathbb{R})$ solution of the Cauchy problem

$$\begin{cases} \dot{\alpha}(\tau) = \Big(l(\tau) + L(\tau)\Big)\alpha(\tau), \\ \alpha(t) = 0, \end{cases} \tag{4.42}$$

such that $\alpha(\tau) \in \mathscr{E}(\tau)$ for all times $\tau \in [t, t+\delta]$. Noting that the unique solution of (4.42) is identically equal to zero, we conclude that $g(\tau) = 0$ on that same interval, which contradicts our initial choice of $t \in [0,T]$. From there, the existence of a viable measure can be obtained by repeating the argument in Step 2 of the proof of Theorem 4.4 above. □

5 Examples of constraints sets and computations of tangents

In this section, we provide two examples of proper constraint sets $\mathcal{Q} \subset \mathscr{P}_p(\mathbb{R}^d)$ which frequently arise in applications, and compute in each case some relevant subsets of tangent directions.

Constraints sets defined by support inclusions. In this first example, we start by considering the prototypical case in which the constraints set is given as

$$\mathcal{Q}_K := \Big\{\mu \in \mathscr{P}_p(\mathbb{R}^d) \text{ s.t. } \mathrm{supp}(\mu) \subset K\Big\},$$

for some compact set $K \subset \mathbb{R}^d$, where $\mathrm{supp}(\mu) \subset \mathbb{R}^d$ denotes the *support* of $\mu \in \mathscr{P}(\mathbb{R}^d)$, defined by

$$\mathrm{supp}(\mu) := \Big\{x \in \mathbb{R}^d \text{ s.t. } \mu(\mathcal{N}_x) > 0 \text{ for each neighbourhood } \mathcal{N}_x \text{ of } x \text{ in } \mathbb{R}^d\Big\}.$$

By Proposition 2.2, it can easily be checked that $\mathcal{Q}_K \subset \mathscr{P}_p(\mathbb{R}^d)$ is compact and thus proper. In what ensues, we fully characterise a nice subset of the *adjacent cone* (see e.g. [9, Definition 4.1.5] for a general definition) to $\mathcal{Q}_K$ at some $\mu \in \mathcal{Q}$, defined in our context by

$$T^\flat_{\mathcal{Q}_K}(\mu) := \Big\{\xi \in \mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d; \mu) \text{ s.t. } \lim_{h\to 0^+} \tfrac{1}{h}\mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}\Big((\mathrm{Id} + h\xi)_\sharp\mu\,;\mathcal{Q}_K\Big) = 0\Big\}. \tag{5.1}$$

Notice in particular that $T^\flat_{\mathcal{Q}_K}(\mu) \subset T_{\mathcal{Q}_K}(\mu)$ by construction.

Proposition 5.1 (Computation of adjacent directions to $\mathcal{Q}_K$)**.** *For every $\mu \in \mathcal{Q}_K$, it holds that*

$$\Big\{\xi \in \mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d; \mu) \text{ s.t. } \xi(x) \in T^\flat_K(x) \text{ for } \mu\text{-almost every } x \in K\Big\} \subset T^\flat_{\mathcal{Q}_K}(\mu)$$

where $T^\flat_K(x)$ denotes the standard adjacent cone to $K \subset \mathbb{R}^d$ at $x \in K$.

Proof. Let $\xi \in \mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d; \mu)$ be such that $\xi(x) \in T^\flat_K(x)$ for μ-almost every $x \in K$ and $h > 0$ be given. In addition, denote by $\hat{\mu}$ any complete extension of the Borel measure μ given e.g. by [80, Theorem 1.36], and observe that the map $x \mapsto x + h\xi(x)$ is Borel and thus $\hat{\mu}$-measurable, see for instance [4, Definition 1.12]. Then, it follows from [9, Theorem 8.2.11] that the set-valued mapping

$$\mathcal{D}_K : x \in K \rightrightarrows \operatorname*{argmin}_{y\in K} |x + h\xi(x) - y| \subset K$$

is $\hat{\mu}$-measurable as well. Because the latter has closed images and since the σ-algebra of $\hat{\mu}$-measurable sets is complete, it follows from [9, Theorem 8.1.3] that there exists a $\hat{\mu}$-measurable selection $x \in K \mapsto \hat{d}_K(x) \in \mathcal{D}_K(x) \subset K$ which satisfies

$$|x + h\xi(x) - \hat{d}_K(x)| = \mathrm{dist}_{\mathbb{R}^d}(x + h\xi(x)\,; K) = o_x(h)$$

for $\hat{\mu}$-almost every $x \in K$ as $h \to 0^+$, where $|o_x(h)| \le h|\xi(x)|$. Moreover, by [16, Proposition 2.1.11], the latter coincides with a Borel map $d_K : K \to K$ outside of a Borel set of zero $\hat{\mu}$-measure. Therefore, noting in turn that

$$\mathrm{supp}(d_{K\,\sharp}\mu) \subset \overline{\Big\{d_K(x) \text{ s.t. } x \in \mathrm{supp}(\mu)\Big\}}^{\mathbb{R}^d} \subset K$$

since $d_K^{-1}(B(y,\varepsilon)) = \emptyset$ for every $y \in \mathbb{R}^d \setminus K$ and $\varepsilon > 0$ for which $B(y,\varepsilon) \cap K = \emptyset$, because $K \subset \mathbb{R}^d$ is closed, it further holds that

$$\begin{aligned}\mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}\Big((\mathrm{Id} + h\xi)_\sharp\mu\,;\mathcal{Q}_K\Big) &\le W_p\Big((\mathrm{Id} + h\xi)_\sharp\mu, d_{K\,\sharp}\mu\Big)\\ &\le \|\mathrm{Id} + h\xi - d_K\|_{\mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\,\mu)} = o(h)\end{aligned}$$

as $h \to 0^+$ by Lebesgue's dominated convergence theorem. This concludes the proof by definition (5.1) of the adjacent cone $T^\flat_{\mathcal{Q}_K}(\mu)$. □

It is possible to generalise this example to time-dependent tubes $K : [0,T] \rightrightarrows \mathbb{R}^d$. In this context, we define the constraints sets by

$$\mathcal{Q}_K(t) := \Big\{\mu \in \mathscr{P}_p(\mathbb{R}^d) \text{ s.t. } \mathrm{supp}(\mu) \subset K(t)\Big\}$$

for all times $t \in [0,T]$. In what follows, we treat the case in which $K : [0,T] \rightrightarrows \mathbb{R}^d$ is left absolutely continuous with nonempty compact images.

Proposition 5.2 (Regularity in time of the constraints)**.** *Under our assumptions on $K : [0,T] \rightrightarrows \mathbb{R}^d$, the constraints tube $\mathcal{Q}_K : [0,T] \rightrightarrows \mathscr{P}_p(\mathbb{R}^d)$ is left absolutely continuous.*

Proof. Fix some $\tau \in [0,T]$ along with $\mu \in \mathscr{P}_p(\mathbb{R}^d)$ and $R > 0$ for which $\mathcal{Q}_K(\tau) \cap \mathbb{B}_{\mathscr{P}_p(\mathbb{R}^d)}(\mu, R) \neq \emptyset$, and let $\mu_\tau \in \mathcal{Q}_K(\tau) \cap \mathbb{B}_{\mathscr{P}_p(\mathbb{R}^d)}(\mu, R)$. Since $K : [0,T] \rightrightarrows \mathbb{R}^d$ is left absolutely continuous, there exists for each $x_\tau \in \mathrm{supp}(\mu_\tau)$ and every $t \in [0,T]$ such that $\tau \leq t$ some point $x_t \in K(t)$ for which

$$|x_t - x_\tau| \leq \int_\tau^t m_K(s)\mathrm{d}s$$

where $m_K(\cdot) \in L^1([0,T], \mathbb{R}_+)$ only depends on $K : [0,T] \rightrightarrows \mathbb{R}^d$. Whence, up to a trivial extension argument outside $\mathrm{supp}(\mu_\tau)$, it is possible to construct a Borel map $\phi_{(\tau,t)} : K(\tau) \to K(t)$ which satisfies

$$|\phi_{(\tau,t)}(x_\tau) - x_\tau| \leq \int_\tau^t m_K(s)\mathrm{d}s$$

for each $x_\tau \in \mathrm{supp}(\mu_\tau)$. Then, one can easily show that the measure $\mu_t := \phi_{(\tau,t)\,\sharp}\mu_\tau$ satisfies $\mathrm{supp}(\mu_t) \subset K(t)$, and by (2.1) the latter also complies with the estimate

$$W_p(\mu_\tau, \mu_t) \leq \int_\tau^t m_K(s)\mathrm{d}s.$$

Repeating the arguments supporting the regularity statement of Theorem 4.8 then closes the proof. □

Proposition 5.3 (Computation of adjacent directions to $\mathrm{Graph}(\mathcal{Q}_K)$)**.** *For all times $t \in [0,T]$ and each $\mu \in \mathcal{Q}_K(t)$, it holds that*

$$\Big\{(\zeta, \xi) \in \mathbb{R} \times \mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d; \mu) \text{ s.t. } (\zeta, \xi(x)) \in T^\flat_{\mathrm{Graph}(K)}(t, x) \text{ for } \mu\text{-almost every } x \in K(t)\Big\} \subset T^\flat_{\mathrm{Graph}(\mathcal{Q}_K)}(t, \mu).$$

Proof. Given some $t \in [0,T]$, an element $\mu \in \mathcal{Q}(t)$ and some $(\zeta, \xi) \in \mathbb{R} \times \mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d; \mu)$ such that $(\zeta, \xi(x)) \in T^\flat_{\mathrm{Graph}(K)}(t, x)$ for μ-almost every $x \in \mathbb{R}^d$, one can show by repeating the arguments in the proof of Proposition 5.1 that the set-valued map

$$\mathcal{D}_{K(t)} : K(t) \rightrightarrows \operatorname*{argmin}_{y \in K(t+h\zeta)} |x + h\xi(x) - y| \subset K(t + h\zeta)$$

admits a Borel selection $x \in K(t) \mapsto d_{K(t)}(x) \in K(t + h\zeta)$ which satisfies

$$|x + h\xi(x) - d_{K(t)}(x)| = o_x(h)$$

for μ-almost every $x \in \mathbb{R}^d$ as $h \to 0^+$, where $|o_x(h)| \leq h|\xi(x)|$. Then, there simply remains to note that $d_{K(t)\,\sharp}\mu \in \mathcal{Q}_K(t + h)$ by construction, while

$$\begin{aligned}\mathrm{dist}_{\mathscr{P}_p(\mathbb{R}^d)}\Big((\mathrm{Id} + h\xi)_\sharp\mu\,;\mathcal{Q}_K(t + h)\Big) &\leq W_p\Big((\mathrm{Id} + h\xi)_\sharp\mu, d_{K(t)\,\sharp}\mu\Big) \\ &\leq \|\mathrm{Id} + h\xi - d_{K(t)}\|_{\mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\,\mu)} = o(h)\end{aligned}$$

by Lebesgue's dominated convergence theorem, which concludes the proof. □

Constraints sets defined as lifted epigraphs. In this second example, which is discussed in our earlier work [25], we consider an extended real-valued map $\mathcal{W} : \mathscr{P}_p(\mathbb{R}^d) \to \mathbb{R} \cup \{+\infty\}$ with compact sublevels that is continuous over its domain $\mathrm{dom}(\mathcal{W}) \subset \mathscr{P}_p(\mathbb{R}^d)$, and define its *lifted epigraph* by

$$\mathcal{Q}_\mathcal{W} := \Big\{\boldsymbol{\mu} \in \mathscr{P}_p(\mathbb{R}^{d+1}) \text{ s.t. } \boldsymbol{\mu} = \mu \times \delta_\alpha \text{ with } \mu \in \mathscr{P}_p(\mathbb{R}^d) \text{ and } \mathcal{W}(\mu) \leq \alpha\Big\}.$$

In what follows, we discuss the topological properties of this constraint set and provide a full characterisation of a relevant subset of its contingent cone.

Proposition 5.4 (Topological properties of $\mathcal{Q}_{\mathcal{W}}$)**.** *Under the assumptions listed hereinabove on the function $\mathcal{W} : \mathscr{P}_p(\mathbb{R}^d) \to \mathbb{R}_+ \cup \{+\infty\}$, the set $\mathcal{Q}_{\mathcal{W}} \subset \mathscr{P}_p(\mathbb{R}^d)$ is proper.*

Proof. Showing that $\mathcal{Q}_{\mathcal{W}}$ is closed under our assumptions is a matter of routine computations. Given $\boldsymbol{\mu} = \mu \times \delta_\alpha \in \mathcal{Q}_{\mathcal{W}}$ and some $\boldsymbol{\nu} \in \mathscr{P}_p(\mathbb{R}^{d+1})$, one has that

$$\begin{aligned} W_p(\boldsymbol{\mu}, \boldsymbol{\nu}) &= \Big(\int_{\mathbb{R}^{2(d+1)}} |(x,\alpha) - (y,\beta)|^p \mathrm{d}\boldsymbol{\gamma}(x,\alpha,y,\beta) \Big)^{1/p} \\ &\geq \Big(\int_{\mathbb{R}^{2(d+1)}} |\alpha - \beta|^p \mathrm{d}\boldsymbol{\gamma}(x,\alpha,y,\beta) \Big)^{1/p} \\ &\geq |\alpha| - \textstyle\int_{\mathbb{R}^{d+1}} |\beta| \mathrm{d}\boldsymbol{\nu}(y,\beta) \end{aligned}$$

for all $\boldsymbol{\gamma} \in \Gamma_o(\boldsymbol{\mu}, \boldsymbol{\nu})$, where we used Jensen's inequality. In particular given $R > 0$, it then holds for every $\boldsymbol{\mu} \in \mathcal{Q}_{\mathcal{W}} \cap \mathbb{B}_{\mathscr{P}_p(\mathbb{R}^{d+1})}(\boldsymbol{\nu}, R)$ that

$$\mathcal{W}(\mu) \leq \alpha \leq R + \textstyle\int_{\mathbb{R}^{d+1}} |\beta| \mathrm{d}\boldsymbol{\nu}(y,\beta).$$

Since $\mathcal{W} : \mathscr{P}_p(\mathbb{R}^d) \to \mathbb{R} \cup \{+\infty\}$ has compact sublevels, we deduce that $\mathcal{Q}_{\mathcal{W}} \cap \mathbb{B}_{\mathscr{P}_p(\mathbb{R}^{d+1})}(\boldsymbol{\nu}, R)$ is compact for any $\boldsymbol{\nu} \in \mathscr{P}_p(\mathbb{R}^{d+1})$ and $R > 0$, and thus that $\mathcal{Q}_{\mathcal{W}}$ is proper. ☐

Similarly to what is known in nonsmooth analysis, the contingent cone to $\mathcal{Q}_{\mathcal{W}}$ can be expressed in terms of *lower directional derivatives* of the functional. Following [14], we define these latter as

$$\mathrm{D}_\uparrow \mathcal{W}(\mu)(\xi) := \liminf_{\substack{h \to 0^+,\ \mu_h \in \mathrm{dom}(\mathcal{W}) \\ W_p((\mathrm{Id}+h\xi)_\sharp \mu, \mu_h) = o(h)}} \frac{\mathcal{W}(\mu_h) - \mathcal{W}(\mu)}{h} \tag{5.2}$$

for each $\mu \in \mathrm{dom}(\mathcal{W})$ and every $\xi \in T_{\mathrm{dom}(\mathcal{W})}(\mu) \subset \mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d; \mu)$.

Proposition 5.5 (Characterisation of contingent directions to $\mathcal{Q}_{\mathcal{W}}$)**.** *For every $\boldsymbol{\mu} \in \mathcal{Q}_{\mathcal{W}}$ and each $(\xi, \rho) \in \mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d; \mu) \times \mathbb{R}$, it holds that*

$$(\xi,\rho) \in T_{\mathcal{Q}_{\mathcal{W}}}(\boldsymbol{\mu}) \qquad \textit{if and only if} \qquad \begin{cases} \xi \in T_{\mathrm{dom}(\mathcal{W})}(\mu) \ \text{ and } \ \mathrm{D}_\uparrow \mathcal{W}(\mu)(\xi) \leq \rho & \textit{when } \alpha = \mathcal{W}(\mu), \\ \xi \in T_{\mathrm{dom}(\mathcal{W})}(\mu) & \textit{when } \alpha > \mathcal{W}(\mu). \end{cases}$$

Proof. The ensuing arguments are largely inspired by those of [9, Proposition 6.1.4]. We only treat the case $\alpha = \mathcal{W}(\mu)$, the other being similar. Assume first that $(\xi, \rho) \in T_{\mathcal{Q}_{\mathcal{W}}}(\boldsymbol{\mu})$ so that one may find a sequence $h_i \to 0^+$ and some $\boldsymbol{\mu}_{h_i} := \mu_{h_i} \times \delta_{\alpha_{h_i}} \in \mathcal{Q}_{\mathcal{W}}$ which satisfy

$$W_p\Big((\mathrm{Id} + h_i(\xi,\rho))_\sharp \boldsymbol{\mu}, \boldsymbol{\mu}_{h_i} \Big) \leq o(h_i). \tag{5.3}$$

Note also that, as a consequence of the disintegration theorem (see e.g. [5, Theorem 5.2.1]) and up to a permutation of coordinates, each optimal plan $\boldsymbol{\gamma}_{h_i} \in \Gamma_o((\mathrm{Id} + h_i(\xi,\rho))_\sharp \boldsymbol{\mu}, \boldsymbol{\mu}_{h_i})$ is of the form

$$\boldsymbol{\gamma}_{h_i} = \gamma_{h_i} \times \delta_{(\mathcal{W}(\mu) + h_i\rho, \alpha_{h_i})}$$

for some $\gamma_{h_i} \in \Gamma_o((\mathrm{Id} + h_i\xi)_\sharp \mu, \mu_{h_i})$. Thence, it necessarily holds that

$$\begin{aligned} W_p\Big((\mathrm{Id} + h_i\xi)_\sharp \mu, \mu_{h_i} \Big) &= \Big(\int_{\mathbb{R}^{2d}} |x-y|^p \mathrm{d}\gamma_{h_i}(x,y) \Big)^{1/p} \\ &\leq \Big(\int_{\mathbb{R}^{2(d+1)}} |(x,\alpha) - (y,\beta)|^p \mathrm{d}\boldsymbol{\gamma}_{h_i}(x,\alpha,y,\beta) \Big)^{1/p} \\ &= W_p\Big((\mathrm{Id} + h_i(\xi,\rho))_\sharp \boldsymbol{\mu}, \boldsymbol{\mu}_{h_i} \Big). \end{aligned}$$

Owing to (5.3), this implies in particular that $\xi \in T_{\mathrm{dom}(\mathcal{W})}(\mu)$. Similarly, one can show that

$$|\alpha_{h_i} - \mathcal{W}(\mu) - h_i\rho| \leq W_p\Big((\mathrm{Id} + h_i(\xi,\rho))_\sharp \boldsymbol{\mu}, \boldsymbol{\mu}_{h_i} \Big) = o(h_i)$$

which together with (5.2) and (5.3) finally yields

$$\mathrm{D}_{\uparrow}\mathcal{W}(\mu)(\rho) \leq \liminf_{h_i \to 0^+} \frac{\mathcal{W}(\mu_{hi}) - \mathcal{W}(\mu)}{h_i} \leq \rho.$$

Conversely, let $(\xi, \rho) \in T_{\mathrm{dom}(\mathcal{W})}(\mu) \times \mathbb{R}$ be such that $\mathrm{D}_{\uparrow}\mathcal{W}(\mu)(\xi) \leq \rho$, and observe then that there exist sequences $h_i \to 0^+$ and $(\mu_{h_i}) \subset \mathrm{dom}(\mathcal{W})$ satisfying $W_p(\mu_{h_i}, (\mathrm{Id} + h_i\xi)_\sharp\mu) = o(h_i)$, for which

$$\mathcal{W}(\mu_{h_i}) \leq \mathcal{W}(\mu) + h_i\rho + o(h_i)$$

when $h_i > 0$ is small enough. Hence, there exists $\rho_i \to \rho$ such that $\boldsymbol{\mu}_{h_i} := \mu_{h_i} \times \delta_{\mathcal{W}(\mu)+h_i\rho_i} \in \mathcal{Q}_{\mathcal{W}}$ and

$$W_p\Big((\mathrm{Id} + h_i(\xi, \rho))_\sharp(\mu \times \delta_{\mathcal{W}(\mu)}), \boldsymbol{\mu}_{h_i}\Big) \leq W_p((\mathrm{Id} + h_i\xi)_\sharp\mu, \mu_{h_i}) + h_i(\rho - \rho_i) = o(h_i)$$

as $h_i \to 0^+$, which equivalently means that $(\xi, \rho) \in T_{\mathcal{Q}_{\mathcal{W}}}(\boldsymbol{\mu})$. □

Appendices

A Proof of Proposition 2.3

In this appendix, we detail the proof of Proposition 2.3. For the sake of self-containedness, we recall first the following technical result taken from [5, Lemma 10.2.1].

Lemma A.1 (Quantitative superdifferentiability estimates on powers of the euclidean norm)**.** *Given $x, y \in \mathbb{R}^d$, one has for $p \in (1, 2]$ that*

$$\tfrac{1}{p}|y|^p - \tfrac{1}{p}|x|^p - \langle y - x, j_p(x)\rangle \leq \tfrac{2^{2-p}}{p-1}|x - y|^p,$$

whereas for $p \in [2, +\infty)$, it holds that

$$\tfrac{1}{p}|y|^p - \tfrac{1}{p}|x|^p - \langle y - x, j_p(x)\rangle \leq \tfrac{p-1}{2}|x - y|^2 \max\{|x|, |y|\}^{p-2}.$$

Therein, $j_p : \mathbb{R}^d \to \mathbb{R}^d$ is the usual duality map defined by

$$j_p(x) := \begin{cases} 0 & \text{if } x = 0, \\ |x|^{p-2}x & \text{otherwise.} \end{cases}$$

Proof of Proposition 2.3. Given an element $\gamma \in \Gamma_o(\mu, \nu)$ and some $h \in \mathbb{R}$, define the transport plan

$$\gamma_h := \Big((\mathrm{Id} + h\zeta) \circ \pi^1, (\mathrm{Id} + h\xi) \circ \pi^2\Big)_\sharp\gamma \in \Gamma\Big((\mathrm{Id} + h\zeta)_\sharp\mu, (\mathrm{Id} + h\xi)_\sharp\nu\Big),$$

and note that by construction, one has that

$$\begin{aligned} \tfrac{1}{p}W_p^p\Big((\mathrm{Id} + h\zeta)_\sharp\mu, (\mathrm{Id} + h\xi)_\sharp\nu\Big) - \tfrac{1}{p}W_p^p(\mu, \nu) &\leq \int_{\mathbb{R}^{2d}} \tfrac{1}{p}|x_h - y_h|^p \mathrm{d}\gamma_h(x_h, y_h) - \int_{\mathbb{R}^{2d}} \tfrac{1}{p}|x - y|^p \mathrm{d}\gamma(x, y) \\ &\leq \int_{\mathbb{R}^{2d}} \Big(\tfrac{1}{p}|x - y + h(\zeta(x) - \xi(y))|^p - \tfrac{1}{p}|x - y|^p\Big)\mathrm{d}\gamma(x, y). \end{aligned} \tag{A.1}$$

By leveraging the identities of Lemma A.1 above, it can be checked that for $p \in (1, 2]$, one has that

$$\begin{aligned} &\int_{\mathbb{R}^{2d}} \Big(\tfrac{1}{p}|x - y + h(\zeta(x) - \xi(y))|^p - \tfrac{1}{p}|x - y|^p - h\left\langle \zeta(x) - \xi(y), j_p(x - y)\right\rangle\Big)\mathrm{d}\gamma(x, y) \\ &\qquad \leq \tfrac{2^{2-p}}{p-1}\int_{\mathbb{R}^{2d}} |h(\zeta(x) - \xi(y))|^p \mathrm{d}\gamma(x, y) \\ &\qquad \leq \tfrac{2}{p-1}|h|^p\Big(\|\zeta\|^p_{\mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d;\, \mu)} + \|\xi\|^p_{\mathcal{L}^p(\mathbb{R}^d, \mathbb{R}^d;\, \nu)}\Big), \end{aligned} \tag{A.2}$$

whereas for $p \in [2, +\infty)$, it holds

$$\begin{aligned}
&\int_{\mathbb{R}^{2d}} \Big(\tfrac{1}{p}|x-y+h(\zeta(x)-\xi(y))|^p - \tfrac{1}{p}|x-y|^p - h\,\langle \zeta(x)-\xi(y), j_p(x-y)\rangle\Big)\mathrm{d}\gamma(x,y)\\
&\leq \tfrac{p-1}{2}\int_{\mathbb{R}^{2d}} |h(\zeta(x)-\xi(y))|^2 \max\Big\{|x-y|, |x-y+h(\zeta(x)-\xi(y))|\Big\}^{p-2}\mathrm{d}\gamma(x,y)\\
&\leq \tfrac{(p-1)}{2}|h|^2\int_{\mathbb{R}^{2d}} |\zeta(x)-\xi(y)|^2\Big(|x-y|+|x-y+h(\zeta(x)-\xi(y))|\Big)^{p-2}\mathrm{d}\gamma(x,y)\\
&\leq (p-1)|h|^2\Big(W_p(\mu,\nu)+|h|\Big(\|\zeta\|_{\mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\,\mu)}+\|\xi\|_{\mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\,\nu)}\Big)\Big)^{p-2}\Big(\|\zeta\|^2_{\mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\,\mu)}+\|\xi\|^2_{\mathcal{L}^p(\mathbb{R}^d,\mathbb{R}^d;\,\nu)}\Big),
\end{aligned} \tag{A.3}$$

where we used elementary Hölder and convexity inequalities to derive both estimates. Thence, upon combining (A.2) and (A.3) with (A.1) depending on the value of $p \in (1, +\infty)$, one finally obtains that

$$\tfrac{1}{p}W_p^p\Big((\mathrm{Id}+h\xi)_\sharp\mu, (\mathrm{Id}+h\zeta)_\sharp\nu\Big) - \tfrac{1}{p}W_p^p(\mu,\nu) \leq h\int_{\mathbb{R}^{2d}} \langle \zeta(x)-\xi(y), j_p(x-y)\rangle \mathrm{d}\gamma(x,y) + r_p(h,\xi,\zeta)$$

with $r_p(h,\zeta,\xi)$ being defined as in (2.3) or (2.4) depending on the value of $p \in (1, +\infty)$. □

B Proof of Proposition 2.8

In this appendix section, we detail the proof of Proposition 2.8.

Proof of Proposition 2.8. In what follows, we assume without loss of generality that $I := [0,T]$ for some $T > 0$, and start by showing that when $\mathcal{Q} : [0,T] \rightrightarrows X$ is absolutley continuous, the map

$$t \in [0,T] \mapsto g(t) := \mathrm{dist}_X(\mathcal{K}(t)\,;\mathcal{Q}(t))$$

is absolutely continuous as well. To do so, we first need to establish some preliminary facts. Observe that since $\mathcal{K}(0)$ is compact, there exist $x_0 \in X$ and some $r_0 > 0$ such that

$$\mathcal{K}(0) \subset \mathbb{B}_X(x_0, r_0) \tag{B.1}$$

for all times $t \in [0,T]$. In what follows, we show that the map $t \in [0,T] \mapsto \mathrm{dist}_X(x_0\,;\mathcal{Q}(t))$ is continuous. Indeed, fixing $\tau \in [0,T]$, setting $R_\tau := \mathrm{dist}_X(x_0\,;\mathcal{Q}(\tau))$ and recalling that $\mathcal{Q} : [0,T] \rightrightarrows X$ is absolutely continuous, for each $\varepsilon > 0$ there exists some $\delta > 0$ such that

$$\mathcal{Q}(\tau) \cap \mathbb{B}_X(x_0, R_\tau+\varepsilon) \subset \mathbb{B}_X\big(\mathcal{Q}(t), \varepsilon\big) \qquad \text{and} \qquad \mathcal{Q}(t) \cap \mathbb{B}_X(x_0, R_\tau+\varepsilon) \subset \mathbb{B}_X\big(\mathcal{Q}(\tau), \varepsilon\big)$$

whenever $|t-\tau| \leq \delta$. Noticing in turn that $\mathcal{Q}(\tau) \cap \mathbb{B}_X(x_0, R_\tau+\varepsilon) \neq \emptyset$ and $\mathcal{Q}(t) \cap \mathbb{B}_X(x_0, R_\tau+\varepsilon) \neq \emptyset$ by construction, the first of these inclusions implies that

$$\begin{aligned}
\mathrm{dist}_X(x_0\,;\mathcal{Q}(t)) &\leq \mathrm{dist}_X(x_0\,;\mathcal{Q}(\tau)) + \mathrm{dist}_X\Big(\mathcal{Q}(\tau) \cap \mathbb{B}_X(x_0, R_\tau+\varepsilon)\,;\mathcal{Q}(t)\Big)\\
&\leq \mathrm{dist}_X(x_0\,;\mathcal{Q}(\tau)) + \varepsilon,
\end{aligned}$$

while the second one analogously yields

$$\mathrm{dist}_X(x_0\,;\mathcal{Q}(\tau)) \leq \mathrm{dist}_X(x_0\,;\mathcal{Q}(t)) + \varepsilon,$$

from whence we can deduce that $t \mapsto \mathrm{dist}_X(x_0\,;\mathcal{Q}(t))$ is continuous at $\tau \in [0,T]$. Recalling that $\mathcal{K} : [0,T] \rightrightarrows X$ has compact images and that it satisfies

$$\mathsf{d}_{\mathcal{H}}(\mathcal{K}(\tau)\,;\mathcal{K}(t)) \leq \int_\tau^t m_{\mathcal{K}}(s)\mathrm{d}s \tag{B.2}$$

for all times $0 \leq \tau \leq t \leq T$ and some $m_{\mathcal{K}}(\cdot) \in L^1([0,T],\mathbb{R}_+)$, it follows from (B.1) that $\mathcal{K}(t) \subset \mathbb{B}_X(x_0, r_{\mathcal{K}})$ for all times $t \in [0,T]$ with $r_{\mathcal{K}} := r_0 + \|m_{\mathcal{K}}(\cdot)\|_1$. Notice then that

$$\mathrm{dist}_X(x\,;\mathcal{Q}(t)) \leq r_{\mathcal{K}} + \max_{t\in[0,T]} \mathrm{dist}_X(x_0\,;\mathcal{Q}(t))$$

for all $t \in [0,T]$ and every $x \in \mathbb{B}_X(x_0, r_{\mathcal{K}})$, so that the quantity

$$R_T := (r_{\mathcal{K}} + 1) + \sup\Big\{\mathrm{dist}_X(x\,;\mathcal{Q}(t)) \ \text{ s.t. } (t,x) \in [0,T] \times \mathbb{B}_X(x_0, r_{\mathcal{K}})\Big\} \tag{B.3}$$

is well-defined and such that

$$\mathrm{dist}_X(\mathcal{K}(t)\,;\partial\mathbb{B}_X(x_0,R_T)) \geq \sup_{\tau,s\in[0,T]} \mathrm{dist}_X(\mathcal{K}(\tau)\,;\mathcal{Q}(s)) + 1$$

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

We now prove that $g : [0,T] \to \mathbb{R}_+$ is absolutely continuous. For all times $\tau, t \in [0,T]$ satisfying $0 \leq \tau \leq t \leq T$, it holds that

$$\begin{aligned} |g(t) - g(\tau)| \leq &\big|\mathrm{dist}_X(\mathcal{K}(t)\,;\mathcal{Q}(t)) - \mathrm{dist}_X(\mathcal{K}(\tau)\,;\mathcal{Q}(t))\big| \\ &+ \big|\mathrm{dist}_X(\mathcal{K}(\tau)\,;\mathcal{Q}(t)) - \mathrm{dist}_X(\mathcal{K}(\tau)\,;\mathcal{Q}(\tau))\big|. \end{aligned} \tag{B.4}$$

In order to estimate the first term in (B.4), note that for each $\varepsilon > 0$, there exist $y_t^\varepsilon \in \mathcal{Q}(t)$ and $x_\tau^\varepsilon \in \mathcal{K}(\tau)$ such that

$$\mathsf{d}_X(x_\tau^\varepsilon, y_t^\varepsilon) \leq \mathrm{dist}_X(\mathcal{K}(\tau)\,;\mathcal{Q}(t)) + \varepsilon.$$

Furthermore, it stems from (B.2) that there exists an element $x_t^\varepsilon \in \mathcal{K}(t)$ for which

$$\mathsf{d}_X(x_\tau^\varepsilon, x_t^\varepsilon) \leq \int_\tau^t m_{\mathcal{K}}(s)\mathrm{d}s.$$

Merging both estimates, it then follows that

$$\begin{aligned} \mathrm{dist}_X(\mathcal{K}(t)\,;\mathcal{Q}(t)) - \mathrm{dist}_X(\mathcal{K}(\tau)\,;\mathcal{Q}(t)) &\leq \mathsf{d}_X(x_t^\varepsilon, y_t^\varepsilon) - \mathsf{d}_X(x_\tau^\varepsilon, y_t^\varepsilon) + \varepsilon \\ &\leq \int_\tau^t m_{\mathcal{K}}(s)\mathrm{d}s + \varepsilon \end{aligned}$$

and repeating the same argument while exchanging the roles of τ and t further yields

$$\big|\mathrm{dist}_X(\mathcal{K}(t)\,;\mathcal{Q}(t)) - \mathrm{dist}_X(\mathcal{K}(\tau)\,;\mathcal{Q}(t))\big| \leq \int_\tau^t m_{\mathcal{K}}(s)\mathrm{d}s \tag{B.5}$$

since $\varepsilon > 0$ was arbitrary. Concerning the second term in (B.4), it stems from our choice of $R_T > 0$ in (B.3) that

$$\mathrm{dist}_X(\mathcal{K}(\tau)\,;\mathcal{Q}(\tau)) = \mathrm{dist}_X\Big(\mathcal{K}(\tau)\,;\mathcal{Q}(\tau) \cap \mathbb{B}_X(x_0,R_T)\Big)$$

and

$$\mathrm{dist}_X(\mathcal{K}(\tau)\,;\mathcal{Q}(t)) = \mathrm{dist}_X\Big(\mathcal{K}(\tau)\,;\mathcal{Q}(t) \cap \mathbb{B}_X(x_0,R_T)\Big)$$

Moreover, since $\mathcal{Q} : [0,T] \rightrightarrows X$ is absolutely continuous in the sense of Definition 2.7, there exists a map $m_{x_0,R_T}(\cdot) \in L^1([0,T],\mathbb{R}_+)$ for which

$$\mathcal{Q}(\tau)\cap\mathbb{B}(x_0,R_T) \subset \mathbb{B}_X\Big(\mathcal{Q}(t), \textstyle\int_\tau^t m_{x_0,R_T}(s)\mathrm{d}s\Big) \quad \text{and} \quad \mathcal{Q}(t)\cap\mathbb{B}(x_0,R_T) \subset \mathbb{B}_X\Big(\mathcal{Q}(\tau), \textstyle\int_\tau^t m_{x_0,R_T}(s)\mathrm{d}s\Big).$$

Combining these few latter facts together, we further obtain

$$\big|\mathrm{dist}_X(\mathcal{K}(\tau)\,;\mathcal{Q}(t)) - \mathrm{dist}_X(\mathcal{K}(\tau)\,;\mathcal{Q}(\tau))\big| \leq \int_\tau^t m_{x_0,R_T}(s)\mathrm{d}s, \tag{B.6}$$

which along with (B.4) and (B.5) finally yields that

$$|g(t) - g(\tau)| \leq \int_\tau^t \Big(m_{\mathcal{K}}(s) + m_{x_0,R_T}(s)\Big)\mathrm{d}s,$$

for all times $\tau, t \in [0,T]$ satisfying $0 \leq \tau \leq t \leq T$, which equivalently means that $g(\cdot) \in \mathrm{AC}([0,T],\mathbb{R}_+)$.

We finally conclude by showing that whenever $\mathcal{Q} : [0,T] \rightrightarrows X$ is left absolutely continuous, then the set-valued map

$$\mathscr{E} : t \in [0,T] \rightrightarrows \Big\{ \alpha \in \mathbb{R}_+ \text{ s.t. } \alpha = g(t) + r \ \text{ for some } r \geq 0 \Big\}$$

is left absolutely continuous as well. To do so, let $x_0 \in \mathcal{K}(0)$ and $R_T > 0$ be as above, fix an element $\alpha_\tau \in \mathscr{E}(\tau)$, and observe that

$$g(\tau) = \mathrm{dist}_X\Big(\mathcal{K}(\tau)\,;\mathcal{Q}(\tau) \cap \mathbb{B}_X(x_0, R_T)\Big) \leq \alpha_\tau$$

by construction. It then follows from elementary applications of the triangle inequality that

$$\begin{aligned} g(t) &= \mathrm{dist}_X(\mathcal{K}(t)\,;\mathcal{Q}(t)) \\ &\leq \mathrm{dist}_X(\mathcal{K}(t)\,;\mathcal{K}(\tau)) + \mathrm{dist}_X\Big(\mathcal{K}(\tau)\,;\mathcal{Q}(\tau) \cap \mathbb{B}_X(x_0, R_T)\Big) + \mathrm{dist}_X\Big(\mathcal{Q}(\tau) \cap \mathbb{B}_X(x_0, R_T)\,;\mathcal{Q}(t)\Big) \\ &\leq \alpha_\tau + \int_\tau^t \Big(m_{\mathcal{K}}(s) + m_{x_0,R_T}(s)\Big)\mathrm{d}s, \end{aligned}$$

for all times $t \in [0,T]$ such that $\tau \leq t$. In particular, we have shown that

$$\Delta_{\alpha,R}(\mathscr{E}(\tau)\,;\mathscr{E}(t)) \leq \int_\tau^t \Big(m_{\mathcal{K}}(s) + m_{x_0,R_T}(s)\Big)\mathrm{d}s$$

for all times $0 \leq \tau \leq t \leq T$, every $\alpha \in \mathbb{R}_+$ and each $R > 0$, which yields the desired claim □

References

[1] G. Albi, Y.P. Choi, M. Fornasier, and D. Kalise. Mean-Field Control Hierarchy. *Applied Mathematics and Optimization*, 76(1):93–135, 2017.

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

[3] 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.

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

[5] 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.

[6] S.S. Arjmand and G. Mazanti. On the Characterization of Equilibria of Nonsmooth Minimal-Time Mean Field Games with State Constraints. In *2021 IEEE 60th Conference on Decision and Control (CDC)*, pages 5300–5305, 2021.

[7] J.-P. Aubin. A Survey of Viability Theory. *SIAM Journal on Control and Optimization*, 28(4):749–788, 1990.

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

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

[10] Y. Averboukh. Viability Theorem for Deterministic Mean-Field Type Control Systems. *Set-Valued and Variational Analysis*, 26(4):993–1008, 2018.

[11] Y. Averboukh and D. Khlopin. Pontryagin Maximum Principle for the Deterministic Mean Field Type Optimal Control Problem via the Lagrangian Approach. *arXiv preprint arXiv:2207.01892*, 2023.

[12] Y. Averboukh, A. Marigonda, and M. Quincampoix. Extremal Shift Rule and Viability Property for Mean Field-Type Control Systems. *Journal of Optimization Theory and Applications*, 189(1):244–270, 2021.

[13] Z. Badreddine and H. Frankowska. Solutions to Hamilton-Jacobi Equation on a Wasserstein Space. *Calculus of Variations and Partial Differential Equations*, 61(1):1–41, 2022.

[14] Z. Badreddine and H. Frankowska. Viability and Invariance on Metric Spaces. *Nonlinear Analysis*, 225:113133, 2022.

[15] J.W. Bebernes and J.D. Schuur. The Wazewski Topological Method for Contingent Equations. *Annali di Matematica Pura e Applicata*, 87(1):271–279, 1970.

[16] V.I. Bogachev. *Measure Theory*, volume 1. Berlin: Springer, 2007.

[17] R. Bonalli and B. Bonnet. First-Order Pontryagin Maximum Principle for Risk-Averse Stochastic Optimal Control Problems. *SIAM Journal on Control and Optimization*, 61(3):1881–1909, 2023.

[18] M. Bongini, M. Fornasier, F. Rossi, and F. Solombrino. Mean Field Pontryagin Maximum Principle. *Journal of Optimization Theory and Applications*, 175:1–38, 2017.

[19] G. Bonnano. Two Theorems on the Scorza-Dragoni Property for Multifunctions. *Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti*, 83(1):51–56, 1989.

[20] B. Bonnet. A Pontryagin Maximum Principle in Wasserstein Spaces for Constrained Optimal Control Problems. *ESAIM COCV*, 25(52), 2019.

[21] B. Bonnet, C. Cipriani, M. Fornasier, and H. Huang. A Measure Theoretical Approach to the Mean-field Maximum Principle for Training NeurODEs. *Nonlinear Analysis, Theory, Methods and Applications*, 227:113161, 2023.

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

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

[24] 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.

[25] B. Bonnet and H. Frankowska. Viability and Exponentially Stable Trajectories for Differential Inclusions in Wasserstein Spaces. *2022 IEEE 61st Conference on Decision and Control (CDC)*, pages 5086–5091, 2022.

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

[27] B. Bonnet and F. Rossi. Intrinsic Lipschitz Regularity of Mean-Field Optimal Controls. *SIAM Journal on Control and Optimization*, 59(3):2011–2046, 2021.

[28] B. Bonnet-Weill and H. Frankowska. Carathéodory Theory and A Priori Estimates for Continuity Inclusions in the Space of Probability Measures. *arxiv preprint arXiv:2302.00963*, 2023.

[29] R. Buckdahn, M. Quincampoix, and A. Rascanu. Viability Property for a Backward Stochastic Differential Equation and Applications to Partial Differential Equations. *Probability Theory and Related Fields*, 116:485–504, 2000.

[30] M. Burger, R. Pinnau, O. Totzeck, and O. Tse. Mean-Field Optimal Control and Optimality Conditions in the Space of Probability Measures. *SIAM Journal on Control and Optimization*, 59(2):977–1006, 2021.

[31] M. Burger, R. Pinnau, O. Totzeck, O. Tse, and A. Roth. Instantaneous Control of Interacting Particle Systems in the Mean-Field Limit. *Journal of Computational Physics*, 405:109–181, 2020.

[32] 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.

[33] M. Caponigro, M. Fornasier, B. Piccoli, and E. Trélat. Sparse Stabilization and Optimal Control of the Cucker-Smale Model. *Mathematical Control and Related Fields*, 3(4):447–466, 2013.

[34] M. Caponigro, M. Fornasier, B. Piccoli, and E. Trélat. Sparse Stabilization and Control of Alignment Models. *Mathematical Models and Methods in Applied Sciences*, 25 (3):521–564, 2015.

[35] M. Caponigro, B. Piccoli, F. Rossi, and E. Trélat. Mean-Field Sparse Jurdjevic-Quinn Control. *Mathematical Models and Methods in Applied Sciences*, 27(7):1223–1253, 2017.

[36] P. Cardaliaguet. Notes on Mean-Field Games. 2012. https://www.ceremade.dauphine.fr/ cardaliaguet/MFG20130420.pdf.

[37] P. Cardaliaguet, A. Mészarós, and F. Santambrogio. First Order Mean Field Games with Density Constraints: Pressure Equals Price. *SIAM Journal on Control and Optimization*, 54(5):2672–2709, 2016.

[38] G. Carlier, C. Jimenez, and F. Santambrogio. Optimal Transportation with Traffic Congestion and Wardrop Equilibria. *SIAM Journal on Control and Optimization*, 47(3):1330–1350, 2008.

[39] J.A. Carrillo, M.R. D'Orsogna, and V. Panferov. Double Milling in Self-Propelled Swarms from Kinetic Theory. *Kinetic & Related Models*, 2(2):363, 2009.

[40] 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.

[41] J.A. Carrillo, D. Kalise, F. Rossi, and E. Trélat. Controlling Swarms toward Flocks and Mills. *SIAM Journal on Control and Optimization*, 60(3):1863–1891, 2022.

[42] 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.

[43] 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.

[44] G. Cavagnari, A. Marigonda, and B. Piccoli. Superposition Principle for Differential Inclusions. In *Large-Scale Scientific Computing*, pages 201–209, 2018.

[45] G. Cavagnari, A. Marigonda, and B. Piccoli. Generalized Dynamic Programming Principle and Sparse Mean-Field Control Problems. *Journal of Mathematical Analysis and Applications*, 481(1):123437, 2020.

[46] 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.

[47] 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.

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

[49] E. Cristiani and D. Peri. Handling Obstacles in Pedestrian Simulations: Models and Optimization. *Applied Mathematical Modelling*, 45:285–302, 2017.

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

[51] 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.

[52] W. E. A Proposal on Machine Learning via Dynamical Systems. *Communications in Mathematics and Statistics*, 1(5):1–11, 2017.

[53] L.C. Evans and R.F. Gariepy. *Measure Theory and Fine Properties of Functions*. CRC Press, 1992.

[54] 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.

[55] 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.

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

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

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

[59] H. Frankowska and T. Lorenz. Filippov's Theorem for Mutational Inclusions in a Metic Space. *Annali della Scuola Normale Superiore di Pisa, Classe di Scienze*, XXIV(5):1053–1094, 2023.

[60] 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.

[61] 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.

[62] H. Frankowska and S. Plaskacz. A Measurable Upper Semicontinuous Viability Theorem for Tubes. *Nonlinear Analysis. Theory, Methods & Applications*, 26(3):565–582, 1996.

[63] 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.

[64] S.-Y. Ha and J.G. Liu. A Simple Proof of the Cucker-Smale Flocking Dynamics and Mean-Field Limit. *Comm. Math. Sci.*, 7(2):297–325, 2009.

[65] G. Haddad. Monotone Viable Trajectories for Functional Differential Inclusions. *Journal of Differential Equations*, 42(1):1–24, 1981.

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

[67] 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.

[68] 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.

[69] S. Karimghasemi, S. Müller, and M. Westdickenberg. Flow Solutions of Transport Equations. *Communications in Partial Differential Equations*, 46(1):98–134, 2021.

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

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

[72] M. Nagumo. Über die Lage der Integralkurven Gewöhnlicher Differentialgleichungen. In *Proceedings of the Physico-Mathematical Society of Japan, 3rd Series*, volume 24, pages 551–559, 1942.

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

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

[75] 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.

[76] 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.

[77] N. Pogodaev. Optimal Control of Continuity Equations. *Nonlinear Differential Equations and Applications*, 23:21, 2016.

[78] N. Pogodaev and M. Saritsyn. Impulsive Control of Nonlocal Transport Equation. *Journal of Differential Equations*, 269(4):3585–3623, 2020.

[79] N. Pogodaev, M. Staritsyn, and R. Chertovskih. Optimal Control of Nonlocal Continuity Equations: Numerical Solution. *arxiv preprint arXiv:2212.06608*, 2023.

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

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

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

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