HAL
open science

Necessary Optimality Conditions for Optimal Control Problems in Wasserstein Spaces

Benoît Bonnet, Hélène Frankowska

To cite this version:

Benoît Bonnet, Hélène Frankowska. Necessary Optimality Conditions for Optimal Control Problems in Wasserstein Spaces. Applied Mathematics and Optimization, 2021, 10.1007/s00245-021-09772-w . hal-03279264

HAL Id: hal-03279264

https://hal.science/hal-03279264

Submitted on 8 Jul 2021

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.

Necessary Optimality Conditions for Optimal Control Problems in Wasserstein Spaces

Benoît Bonnet*, Hélène Frankowska†

July 6, 2021

Abstract

In this article, we derive first-order necessary optimality conditions for a constrained optimal control problem formulated in the Wasserstein space of probability measures. To this end, we introduce a new notion of localised metric subdifferential for compactly supported probability measures, and investigate the intrinsic linearised Cauchy problems associated to non-local continuity equations. In particular, we show that when the velocity perturbations belong to the tangent cone to the convexification of the set of admissible velocities, the solutions of these linearised problems are tangent to the solution set of the corresponding continuity inclusion. We then make use of these novel concepts to provide a synthetic and geometric proof of the celebrated Pontryagin Maximum Principle for an optimal control problem with inequality final-point constraints. In addition, we propose sufficient conditions ensuring the normality of the maximum principle.

Keywords : Mean-Field Optimal Control, Wasserstein Spaces, Pontryagin Maximum Principle, Differential Inclusions, Inner-Approximations of Optimal Trajectories

MSC2020 Subject Classification : 30L99, 34K09, 49J53, 49K21, 49Q22, 58E25

1 Introduction

In recent years, the mathematical analysis of collective behaviours in large systems of interacting agents has been the object of a growing attention from several mathematical communities. The corresponding class of so-called *multi-agent systems* appears in a wide number of applications, including e.g. the modelling of autonomous vehicle ensembles [15], swarms and flocking structures in the animal kingdom [23], opinion dynamics on networks [8] and pedestrian flows [22]. While these structures are frequently represented first-handedly as large dynamical systems formulated in finite-dimensional spaces – e.g. by means of families of ordinary differential equations or time-dependent graphs –, such discrete formulations often give rise to intractable problems. Indeed, the main difficulty inherent to the study of multi-agent systems is their very large dimensionality, which causes severe challenges when trying to compute solutions using numerical methods or to describe agent-to-agent interactions in a meaningful way. These aspects are particularly relevant in *control-theoretic* settings, where a policy maker aims at synthesising a control law which is as generic and global as possible in order to stir a system towards a desired goal. This type of control problems arise very frequently in the applied fields which were previously mentioned, see e.g. [1, 2, 44] and references therein.

In this context, the first step in order to circumvent these high-dimension related issues is to identify adequate infinite-dimensional approximations of the class of multi-agent systems at hand, usually in the form of a family of partial differential equations. One subsequently aims at synthesising a control law directly at the level of the PDE, and finally at returning to the original problem by using the "infinite-dimensional" control signal to stir the original finite-dimensional system (see e.g. the introduction of [13] for ampler details on this general scheme). To this day, the best identified

*CNRS, IMJ-PRG, UMR 7586, Sorbonne Université, 4 place Jussieu, 75252 Paris, France. *E-mail*: benoit.bonnet@imj-prg.fr (Corresponding author)

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

framework to conduct such a program is that of *mean-field approximations*. In this setting, the discrete agent trajectories are replaced by a curve of measures $\mu(\cdot)$, whose time-evolution is described by means of a *non-local continuity equation* (see Section 2.3 below) of the form

$$\partial_t \mu(t) + \operatorname{div}\big(v(t, \mu(t))\mu(t)\big) = 0. \tag{1}$$

These equations appear very naturally when performing infinite-dimensional approximations of deterministic particle systems driven by ODEs, and are usually studied in the space of probability measures endowed with the *Wasserstein metrics* of optimal transport (see Section 2.1 below). For a modern introduction to this topic, we point to [41] as well as to the survey [21].

In this article, we investigate necessary optimality conditions for the *mean-field optimal control problems* of Bolza type

$$(\mathcal{P}) \quad \begin{cases} \min\limits_{u(\cdot)\in\mathcal{U}} \left[\displaystyle\int_0^T L(t, \mu(t), u(t))\mathrm{d}t + \varphi(\mu(T)) \right] \\ \text{s.t.} \begin{cases} \partial_t \mu(t) + \operatorname{div}\big(v(t, \mu(t), u(t))\mu(t)\big) = 0, \\ \mu(0) = \mu^0, \\ \mu(T) \in \mathcal{Q}_T. \end{cases} \end{cases}$$

The minimisation in problem $(\mathcal{P})$ is taken over the set of admissible *open-loop* controls $\mathcal{U} := \{t \in [0,T] \mapsto u(t) \in U \text{ s.t. } u(\cdot) \text{ is } \mathscr{L}^1\text{-measurable}\}$, where (U, d_U) is a compact metric space and $\mathscr{L}^1$ denotes the standard one-dimensional Lebesgue measure. The state $\mu(\cdot)$ of the system is stirred by the controlled non-local velocity field $v : [0,T] \times \mathscr{P}_c(\mathbb{R}^d) \times U \times \mathbb{R}^d \to \mathbb{R}^d$, while $\mu^0 \in \mathscr{P}_c(\mathbb{R}^d)$ denotes a fixed initial datum and $\mathcal{Q}_T \subset \mathscr{P}_2(\mathbb{R}^d)$ is a closed set which represents final-point constraints. As customary in the modelling of optimal control problems, the cost is split into a final cost $\varphi : \mathscr{P}_c(\mathbb{R}^d) \to \mathbb{R}$ and a distributed running cost $L : [0,T] \times \mathscr{P}_c(\mathbb{R}^d) \times U \to \mathbb{R}$. Observe that in $(\mathcal{P})$, the state variable $\mu(\cdot)$ is modelled as a curve of *compactly supported* measures. This comes from the fact that the afore-described applications of multi-agent systems usually involve compactly supported initial data, and that classical stability estimates (see e.g. Theorem 2.11 below) ensure that curves of measure solving non-local continuity equations starting from compactly supported data remain compactly supported at all subsequent times.

Optimal control problems in Wasserstein spaces of the form $(\mathcal{P})$ have been extensively studied in the past few years, mostly in the absence of final-point or running constraints. The first contributions on this topic [26, 27, 28] were concerned with the rigorous convergence of discrete multi-agent optimal controls towards their infinite-dimensional counterparts (see also the more recent contributions [19, 26]), which provides a constructive way of recovering the existence of optimal trajectories. In parallel to these investigations, several attempts were made at deriving *necessary optimality conditions* for mean-field optimal control problems. The first result of this kind was proposed in [9], where the authors proved a generalisation of the celebrated *Pontryagin Maximum Principle* ("PMP" in the sequel) for a multi-scale version of $(\mathcal{P})$. Therein, a family of open-loop controls acts on a finite-dimensional subset of *leaders*, which allows to recover a PMP in Wasserstein spaces by passing to the limit in the finite-dimensional maximum principles as the number of agents goes to infinity. The first general derivation of the PMP for problem $(\mathcal{P})$ was obtained by the first author in [12], and is based on a careful adaptation of the strategy of *needle-variations* to the abstract geometric structure of Wasserstein spaces. This result was further extended in [10] to problems with general final-point and running state constraints. In the latter contribution, the proof strategy combines a finite-dimensional non-smooth multipliers rule and *outer-approximations* of optimal trajectories by countable families of curves generated using needle-variations. We also mention the earlier work [45] in which the author independently derived an alternative version of the maximum principle for a very particular instance of problem $(\mathcal{P})$. This result was subsequently extended in the recent paper [46], in which a full PMP in the spirit of [12] is derived for unconstrained *impulsive* control problems, by means of discrete approximations combined with Ekeland's principle. We likewise point out that a completely different approach to Pontryagin optimality conditions for mean-field optimal control problems – formulated in terms of McKean-Vlasov dynamics and stochastic differential equations – was independently developed in [17] (see also the recent monograph [18, Vol.I Chapter 6]).

In this manuscript, we revisit the necessary optimality conditions for problem $(\mathcal{P})$ by proposing a geometric proof of the maximum principle (see Theorem 4.4 and Theorem 4.9 below). The latter is based on a general methodology which was first developed by the second author in [29] for Mayer problems, and that can be summarised as follows. When the running cost $L(\cdot,\cdot,\cdot)$ is equal to zero, the generalised Fermat rule states that the directional derivatives of the final cost $\varphi(\cdot)$ taken at points of local minimum are non-negative when evaluated along directions that are tangent to the set of admissible trajectories. Usually, due to the non-linearity of the dynamics, it is not possible to exhaustively describe the set of all such tangents. However in order to derive the PMP, it is sufficient to identify subsets of tangent directions which are informative enough. As it can be expected from the classical theory of optimisation, the set of trajectories of the linearised control system satisfying linearised constraints are tangent to the set of all admissible trajectories, provided some additional constraint qualifications are satisfied. While classical linearisation methods for control systems usually involve derivatives of the dynamics with respect to both state and control, these latter only describe a small subset of tangent directions and lead to the so-called *weak maximum principle*. On the other hand, replacing the control system by a *relaxed differential inclusion* does not change the set of tangents to its trajectories, even though it drastically enlarges the dynamics itself (since the right-hand side becomes the convex hull of the admissible velocities of the original control system). Then, linearisations of this "larger" set of dynamics obtained by taking adequate derivatives of the resulting set-valued velocity mapping allow to describe richer subsets of tangents. Note next that the Fermat rule also implies that the set of curves which are tangent to admissible trajectories has an empty intersection with the the set of all directions in which the cost is strictly decreasing. At this point, one can recover an explicit dual description of this fact in the form of an inequality by applying a separation theorem. The final step in this procedure is then to introduce a suitably chosen *costate variable* satisfying the adjoint dynamics, which is such that the maximisation condition of the PMP holds for almost-every time. Besides, it is well known that Bolza problems can be equivalently written in Mayer form – i.e. with $L = 0$ in $(\mathcal{P})$ –, which means that the approach we just described is also valid for optimal control problems with non-trivial running costs. When the final-point constraints are described by families of functional inequalities and the set of tangents to the latter has empty interior, it follows that the abnormal version of the PMP holds true, i.e. the adjoint state does not depend on the cost of the problem. On the other hand when the tangent set to the final-point constraints has non-empty interior and when the constraints are qualified in some sense, then the normal version of the PMP holds true, and the investigation of optimal curves is greatly simplified.

The approach that we described in the previous paragraph is general enough to continue with second- (and higher-) order optimality conditions, by taking second-order linearisations of control systems and constraints, as well as second-order derivatives of the costs, even in the presence of additional running state constraints. Another interesting feature of this strategy is that it allows to precisely discriminate between the normal and abnormal versions of the PMP, which is generically not true for proofs relying on outer-approximations of optimal trajectories. This general methodology has been successfully applied to optimal control problems formulated on very broad classes of dynamics, including e.g. classical ODEs in finite-dimensional spaces [29], semi-linear evolution equations in general Banach spaces [30, 32] and stochastic differential equations [36]. In addition to its holistic nature, this procedure also presents the advantage of being readily transposable to the study of *second-order* necessary optimality conditions for optimal control problems, see e.g. [31, 33, 34] and references therein. We point more specifically to the recent publication [35] on this topic, in which an optimal control problem involving final-point equality and inequality constraint was investigated using metric inverse mapping theorems on the space of controls.

However, in the context of mean-field optimal control problems written in the general form $(\mathcal{P})$, the implementation of the afore-described program is much more delicate. Indeed, the space of probability measures is at best a metric space when endowed with a Wasserstein metric, and a number of tools and concepts of set-valued and non-smooth analysis need to be carefully adapted to this setting in order to study necessary optimality conditions. Concerning differential-theoretic aspects, it has been hinted in the pioneering works [40, 42] and further installed in [5, 49] that the Wasserstein space $(\mathscr{P}_2(\mathbb{R}^d), W_2)$ can be endowed with a *pseudo-Riemannian* structure. More specifically, the latter

provides an explicit characterisation of the exponential map and of the corresponding *analytical tangent space* $\mathrm{Tan}_\mu \mathscr{P}_2(\mathbb{R}^d)$ (see Section 2.1 below), and a full theory of first- and second-order differential calculus has been developed in this framework, see [5, Chapter 10] and [4, 37]. Regarding differential inclusions, the authors of the present manuscript have recently proposed in [11] a sound generalisation of these objects for continuity equations formulated in Wasserstein spaces, in keeping with the classical theory that is described e.g. in [7, Chapter 10]. In particular contrary to other attempts aiming at generalising differential inclusions in this context (see e.g. [20, 38]), the notion of *continuity inclusion* introduced in [11] ensures a one-to-one correspondence between the trajectories of control systems and the canonically associated set-valued dynamics (see Theorem 2.15 below).

The contributions of this manuscript can be summarised as follows. In Section 3.1, we start by proposing a new definition of *localised subdifferentials* and *differentiability* (see Definition 3.1 and Definition 3.3 below) for functionals defined over Wasserstein spaces, which are well-suited to the study of dynamical and variational problems involving compactly supported measures. In Proposition 3.5, we adapt to this new setting a result already known for absolutely continuous measures in the classical case (see e.g. [24, Chapter 1, Proposition 4.2]), which asserts that subdifferentials which belong to $\mathrm{Tan}_\mu \mathscr{P}_2(\mathbb{R}^d)$ are in fact strong (see Definition 3.1 below). We subsequently derive the expression of the intrinsic linearised Cauchy problem associated to a non-local continuity equation in Proposition 3.10 of Section 3.2, and study in Theorem 3.12 the *variational linearisation* of continuity inclusions. Thereafter in Section 4.1, we apply these differential and set-theoretic concepts to prove a Pontryagin Maximum Principle for a *Mayer* version of problem $(\mathcal{P})$ – i.e. without running cost – in Theorem 4.4, where the set of end-point constraints $\mathcal{Q}_T$ is defined by *functional inequalities*. We also point out that the proof of this result has been slightly improved compared to earlier implementations of this program e.g. in [32, 35], as the separation arguments are performed in adequate *finite-dimensional* spaces, which circumvents a lot of potential technicalities related to the separation of subsets of infinite-dimensional vector spaces. A full PMP for the corresponding Bolza problem is then recovered in Theorem 4.9 of Section 4.2 by means of a standard procedure that directly builds on the PMP for the Mayer problem.

Before moving on to the core of the manuscript, we would like to make some final remarks.

1. While our methodology would also allow to deal with end-point constraints involving *functional equalities*, the corresponding proof strategy relies on a different class of inner-approximations, combined with technical metric inverse mapping arguments in the spirit of [35]. For the sake of clarity, we shall present these results separately in a future publication.

2. Even though the constrained PMP derived in [10] incorporates state constraints and takes into account more general final-point constraints, its proof suffers from the drawbacks of being less geometrically meaningful and highly technical, while relying on outer-approximations. In this regard, the approach described in this manuscript is more synthetic, and should allow for the derivation of the maximum principle in the presence of general constraints (see e.g. [34, 35]).

3. The study of mean-field control problems of the form $(\mathcal{P})$ is also sensible when the admissible controls $u : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ have a *closed-loop* structure with respect to the space variable, see e.g. [10, 12, 16, 44] and [18, 39] in connection with the theory of mean-field games. However in this case, derivatives with respect to the space variable of the closed-loop controls also appear in the linearised systems and in the adjoint dynamics of the PMP, which requires more regularity assumptions and leads to heavier expressions while using the same geometric ideas as in the open-loop case.

The structure of the article is the following. In Section 2, we expose a list of preliminary notions of optimal transport theory and set-valued analysis. In particular, we recall some of the recent results of [11] about continuity inclusions. In Section 3, we present our new definition of localised subdifferentials along with the intrinsic and variational linearisations of the set-valued counterparts of continuity equations. In Section 4, we subsequently prove a PMP in Wasserstein spaces, first for a Mayer problem, and then for a Bolza problem. We finally present in Section 5 examples of functionals which are locally differentiable in the sense of Section 3, and provide the expressions of their gradients.

2 Preliminaries

In this section, we recall several notions pertaining to optimal transport theory, set-valued analysis and continuity inclusions in Wasserstein spaces, all of which will be needed to state and prove the main results of Section 3 and Section 4.

2.1 Analysis in measures spaces and optimal transport

We recollect here some preliminary tools of analysis in measure spaces and optimal transport theory, for which we refer the reader to the monographs [3] and [5, 48, 49] respectively.

Given two complete separable metric spaces $(\mathscr{S}, d_{\mathscr{S}})$ and $(\mathscr{T}, d_{\mathscr{T}})$, we denote by $\overline{\Omega}$ the closure of a subset $\Omega \subset \mathscr{S}$ and by $\partial\Omega$ its topological boundary. A mapping $\phi : \mathscr{S} \to \mathscr{T}$ is said to be bounded if for some (and thus all) $\tau_0 \in \mathscr{T}$, it holds that $\sup_{s\in\mathscr{S}} d_{\mathscr{T}}(\phi(s), \tau_0) < +\infty$. We will use the notation $C^0(\mathscr{S}, \mathscr{T})$ (resp. $C^0_b(\mathscr{S}, \mathscr{T})$) for the space of continuous (resp. continuous and bounded) functions from $\mathscr{S}$ into $\mathscr{T}$, as well as $\mathrm{AC}([0,T], \mathscr{S})$ for the space of absolutely continuous curves from $[0,T] \subset \mathbb{R}_+$ into $\mathscr{S}$. In the sequel, $\mathrm{Lip}(\phi(\cdot)\,;\Omega)$ stands for the Lipschitz constant of a map $\phi : \mathscr{S} \to \mathscr{T}$ over $\Omega \subset \mathscr{S}$. Given a separable Banach space $(X, \|\cdot\|_X)$ and $p \in [1, +\infty)$, we denote by $L^p([0,T], X)$ the Banach space of *p-integrable maps* from $[0,T] \subset \mathbb{R}_+$ into X in the sense of Bochner (see e.g. [25]), where $[0,T]$ is endowed with the standard one-dimensional Lebesgue measure $\mathscr{L}^1$.

Let $\mathscr{P}(\mathbb{R}^d)$ be the space of *Borel probability measures* defined over $\mathbb{R}^d$ endowed with the *narrow topology*, that is the topology induced by the weak-* convergence of measures defined by

$$\mu_n \underset{n\to+\infty}{\rightharpoonup^*} \mu \quad \text{if and only if} \quad \int_{\mathbb{R}^d} \phi(x)\mathrm{d}\mu_n(x) \underset{n\to+\infty}{\longrightarrow} \int_{\mathbb{R}^d} \phi(x)\mathrm{d}\mu(x) \quad \text{for every } \phi \in C^0_b(\mathbb{R}^d, \mathbb{R}). \tag{2}$$

Given $p \in [1, +\infty)$, we define the *momentum of order p* of an element $\mu \in \mathscr{P}(\mathbb{R}^d)$ by

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

as well as its *support* as the closed subset of $\mathbb{R}^d$

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

We will henceforth denote by $\mathscr{P}_p(\mathbb{R}^d)$ and $\mathscr{P}_c(\mathbb{R}^d)$ the subsets of Borel probability measures with finite momentum of order p and compact support respectively. In what follows, we will frequently use the *R-fattening* of the support of a measure $\mu \in \mathscr{P}_c(\mathbb{R}^d)$, defined by

$$B_\mu(R) \ := \bigcup_{x\,\in\,\mathrm{supp}(\mu)} B(x, R), \tag{3}$$

where $B(x, R) \subset \mathbb{R}^d$ denotes the closed ball of center $x \in \mathbb{R}^d$ with radius $R > 0$.

Definition 2.1 (Pushforward of a measure and transport plans)**.** *Given a Borel map $f : \mathbb{R}^d \to \mathbb{R}^d$ and an element $\mu \in \mathscr{P}(\mathbb{R}^d)$, the* pushforward *$f_\#\mu \in \mathscr{P}(\mathbb{R}^d)$ of μ through $f(\cdot)$ is defined as*

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

for any Borel set $B \subset \mathbb{R}^d$. Given two elements $\mu, \nu \in \mathscr{P}(\mathbb{R}^d)$, the set $\Gamma(\mu, \nu)$ of transport plans *between μ and ν is defined as the subset of all $\gamma \in \mathscr{P}(\mathbb{R}^{2d})$ such that $\pi^1_\#\gamma = \mu$ and $\pi^2_\#\gamma = \nu$, where the maps $\pi^1, \pi^2 : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ stand for the projection operations onto the first and second factors.*

We recall below the definition and some of the known properties of the *Wasserstein spaces* of optimal transport theory (see e.g. [5, 48, 49]).

Definition 2.2 (Wasserstein spaces)**.** *Given $p \in [1, +\infty)$ and $\mu, \nu \in \mathscr{P}_p(\mathbb{R}^d)$, the* Wasserstein distance of order p *between μ and ν is defined by*

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

and we denote by $\Gamma_o(\mu, \nu)$ the set of optimal transport plans *for which this minimum is attained. The* Wasserstein space of order p *is then defined as the metric space $(\mathscr{P}_p(\mathbb{R}^d), W_p)$ of probability measures with finite momentum of order p endowed with the W_p-distance.*

It is a well-known fact in optimal transport theory that the sets $\Gamma_o(\mu, \nu)$ are non-empty for every $p \in [1, +\infty)$ and any $\mu, \nu \in \mathscr{P}_p(\mathbb{R}^d)$. This is a direct consequence of the narrow closedness of $\Gamma(\mu, \nu)$, together with the coercivity of the map

$$\gamma \in \mathscr{P}_p(\mathbb{R}^{2d}) \mapsto \int_{\mathbb{R}^{2d}} |x - y|^p \mathrm{d}\gamma(x, y) \in \mathbb{R}_+.$$

Moreover, the space $(\mathscr{P}_p(\mathbb{R}^d), W_p)$ is a complete separable metric space, and the topology induced by the Wasserstein distance metrises the narrow topology (2) restricted to $\mathscr{P}_p(\mathbb{R}^d)$, in the sense that

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

for every sequence $(\mu_n) \subset \mathscr{P}_p(\mathbb{R}^d)$ and $\mu \in \mathscr{P}_p(\mathbb{R}^d)$. In the sequel, we will frequently consider the (non-complete) metric space $(\mathscr{P}_c(\mathbb{R}^d), W_p)$ of compactly supported measures, seen as a subset of the Wasserstein space $(\mathscr{P}_p(\mathbb{R}^d), W_p)$, where $p \in [1, +\infty)$ will depend on the context.

In the particular case where $p = 2$, it is known that $(\mathscr{P}_2(\mathbb{R}^d), W_2)$ can also be endowed with a *pseudo-Riemannian structure.* Given an element $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, the *analytical tangent space* to $\mathscr{P}_2(\mathbb{R}^d)$ at μ is defined in this context as

$$\mathrm{Tan}_\mu \mathscr{P}_2(\mathbb{R}^d) := \overline{\{\nabla \xi(\cdot) \text{ s.t. } \xi \in C_c^\infty(\mathbb{R}^d)\}}^{L^2(\mu)}.$$

We point the reader to [5, Chapter 8] for a contextual introduction of this geometric object, as well as for a thorough analysis of its interplay with absolutely continuous curves of measures and continuity equations (see also Section 2.3 below).

We end this series of prerequisites by recalling a variant of the *disintegration theorem* in the context of optimal transport, for which we refer e.g. to [5, Theorem 5.3.1].

Theorem 2.3 (Disintegration theorem)**.** *Given two elements $\mu, \nu \in \mathscr{P}_p(\mathbb{R}^d)$ and $\gamma \in \Gamma(\mu, \nu)$, there exists a μ-almost uniquely determined family of measures $\{\gamma_x\}_{x \in \mathbb{R}^d} \subset \mathscr{P}_p(\mathbb{R}^d)$ called the* disintegration *of γ onto its first marginal $\pi^1_\# \gamma = \mu$, such that*

$$\int_{\mathbb{R}^{2d}} \xi(x, y) \mathrm{d}\gamma(x, y) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \xi(x, y) \mathrm{d}\gamma_x(y) \mathrm{d}\mu(x),$$

for every map $\xi \in L^1(\mathbb{R}^{2d}, \mathbb{R}; \gamma)$.

Reciprocally given $\mu \in \mathscr{P}_c(\mathbb{R}^d)$ and a μ-measurable family of measures $\{\gamma_x\}_{x \in \mathbb{R}^d} \subset \mathscr{P}_c(\mathbb{R}^d)$, we shall sometimes write $\gamma := \int_{\mathbb{R}^d} \gamma_x \mathrm{d}\mu(x)$ to denote the unique element of $\mathscr{P}_c(\mathbb{R}^{2d})$ satisfying

$$\int_{\mathbb{R}^{2d}} \xi(x, y) \mathrm{d}\gamma(x, y) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \xi(x, y) \mathrm{d}\gamma_x(y) \mathrm{d}\mu(x)$$

for every $\xi \in C_b^0(\mathbb{R}^{2d}, \mathbb{R})$.

2.2 Elements of set-valued analysis

In this section, we recall several notions and classical results of set-valued and non-smooth analysis, for which we refer to the comprehensive monographs [6, 7].

Let $(\mathscr{S}, d_{\mathscr{S}})$ be a complete separable metric space and $(X, \|\cdot\|_X)$ be a separable Banach space. We denote by $\mathbb{B}_{\mathscr{S}}(s,\beta)$ the closed ball of radius $\beta > 0$ centred at $s \in \mathscr{S}$, and by $\mathbb{B}_X$ the closed unit ball in X centred at 0. Given a subset $B \subset X$, we define its *closed convex hull* $\overline{\mathrm{co}}(B)$ as the closure of

$$\mathrm{co}(B) := \left\{ \textstyle\sum_{i=1}^N \alpha_i b_i \text{ s.t. } N \geq 1,\ b_i \in B,\ \alpha_i \geq 0 \text{ for } i \in \{1,\dots,N\} \text{ and } \sum_{i=1}^N \alpha_i = 1 \right\} \subset X,$$

and consider the *distance function* to B defined as $x \in X \mapsto \mathrm{dist}_X(x\,;B) := \inf_{b\in B} \|x-b\|_X$. An application $\mathcal{F} : \mathscr{S} \rightrightarrows X$ is called a *set-valued map* – or a *multifunction* – from $\mathscr{S}$ into X if $\mathcal{F}(s) \subset X$ for any $s \in \mathscr{S}$, and we denote by $D(\mathcal{F}) := \{s \in \mathscr{S} \text{ s.t. } \mathcal{F}(s) \neq \emptyset\}$ its *effective domain*. A set-valued map $\mathcal{F}(\cdot)$ is said to have closed (resp. convex) images if the sets $\mathcal{F}(s)$ are closed (resp. convex) for any $s \in D(\mathcal{F})$. In the following definitions, we recall the concepts of *measurability* and *Lipschitz regularity* for set-valued maps.

Definition 2.4 (Measurable multifunctions)**.** *Let $(\mathscr{S}, \Omega, \varpi)$ be a measure space defined over $\mathscr{S}$ and $\mathcal{F} : \mathscr{S} \rightrightarrows X$ be a set-valued map. Then, $\mathcal{F}(\cdot)$ is said to be* ϖ-measurable *if*

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

for any open set $\mathcal{O} \subset X$.

Definition 2.5 (Lipschitz continuity of multifunctions)**.** *A set-valued map $\mathcal{F} : \mathscr{S} \rightrightarrows X$ is* Lipschitz regular *around $s \in D(\mathcal{F})$ with constant $L > 0$ if there exists a neighbourhood $\mathcal{N}_s \subset \mathscr{S}$ of s such that*

$$\mathcal{F}(s_1) \subset \mathcal{F}(s_2) + L\, d_{\mathscr{S}}(s_1, s_2)\mathbb{B}_X,$$

for any $s_1, s_2 \in \mathcal{N}_s$.

We end this primer in set-valued analysis by recalling the definition of *tangent cone* to a closed convex subset of a Banach space, and provide one of its equivalent characterisations.

Definition 2.6 (Tangent cone to a convex set)**.** *Let $K \subset X$ be a closed convex set and $x \in K$. The* tangent cone *$T_K(x)$ to K at x is then defined by*

$$T_K(x) := \overline{\bigcup_{\lambda>0} \lambda(K-x)}.$$

Moreover, the set $T_K(x)$ can be alternatively characterised as

$$\begin{aligned} T_K(x) &= \left\{ w \in X \text{ s.t. } \mathrm{dist}_X(x+\epsilon w; K) = o(\epsilon) \ \text{ for any } \epsilon > 0 \right\} \\ &= \Big\{ w \in X \text{ s.t. for every } \epsilon_n \underset{n\to+\infty}{\longrightarrow} 0^+, \text{ there exist } w_n \underset{n\to+\infty}{\longrightarrow} w \text{ such that} \\ &\qquad\qquad x + \epsilon_n w_n \in K \ \text{ for all } n \geq 1 \Big\}. \end{aligned} \tag{4}$$

as a consequence e.g. of [7, Proposition 4.2.1].

2.3 Continuity equations and inclusions in Wasserstein spaces

In this section, we recollect notions pertaining to *continuity equations* and *continuity inclusions* in the space of measures. Given a time horizon $T > 0$, a curve $\mu : [0,T] \to \mathscr{P}(\mathbb{R}^d)$ and a velocity-field $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, we say that $(\mu(\cdot), v(\cdot))$ is a trajectory-velocity pair of the continuity equation

$$\partial_t \mu(t) + \mathrm{div}(v(t)\mu(t)) = 0, \tag{5}$$

if the following distributional equality

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

holds for any test function $\xi \in C_c^\infty((0,T)\times\mathbb{R}^d)$. In what follows, we will mainly consider continuity equations driven by *Carathéodory* velocity-fields, namely maps such that $t \in [0,T] \mapsto v(t,x) \in \mathbb{R}^d$ is $\mathscr{L}^1$-measurable for any $x \in \mathbb{R}^d$ and $x \in \mathbb{R}^d \mapsto v(t,x) \in \mathbb{R}^d$ is continuous for $\mathscr{L}^1$-almost every $t \in [0,T]$.

In the recent work [11], we introduced a *set-valued* extension of (5), in the case where the velocity-field is *non-local* i.e. when it also depends on the whole measure $\mu(t)$ at all times $t \in [0,T]$. Based on the identification of non-local velocity-fields

$$(t,\mu,x) \in [0,T]\times\mathscr{P}_c(\mathbb{R}^d)\times\mathbb{R}^d \mapsto v(t,\mu,x) \in \mathbb{R}^d,$$

with vector-field valued maps

$$(t,\mu) \in [0,T]\times\mathscr{P}_c(\mathbb{R}^d) \mapsto v(t,\mu,\cdot) \in C^0(\mathbb{R}^d,\mathbb{R}^d),$$

we proposed the following notion of differential inclusion for continuity equations.

Definition 2.7 (Continuity inclusions in measure spaces)**.** *Let $V : [0,T]\times\mathscr{P}_c(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d,\mathbb{R}^d)$ be a set-valued map. We say that a curve of measures $\mu(\cdot)$ solves the* continuity inclusion

$$\partial_t\mu(t) \in -\mathrm{div}\Big(V(t,\mu(t))\mu(t)\Big), \tag{6}$$

if there exists an $\mathscr{L}^1$-measurable selection $t \in [0,T] \mapsto \boldsymbol{v}(t) \in V(t,\mu(t))$ such that the trajectory-selection pair $(\mu(\cdot),\boldsymbol{v}(\cdot))$ *solves the continuity equation*

$$\partial_t\mu(t) + \mathrm{div}\big(\boldsymbol{v}(t)\mu(t)\big) = 0, \tag{7}$$

in the sense of distributions.

In what follows, we introduce technical notions pertaining to *compact restrictions* of multifunctions with values in $C^0(\mathbb{R}^d,\mathbb{R}^d)$, and proceed by stating our working assumptions. We will henceforth denote by $\|\cdot\|_1 := \|\cdot\|_{L^1([0,T],\mathbb{R})}$ the L^1-norm of a real-valued functional defined over $[0,T]$.

Definition 2.8 (Restriction of multifunctions to balls)**.** *Given a set-valued map $V : [0,T]\times\mathscr{P}_c(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d,\mathbb{R}^d)$ and $R > 0$, we define its* compact restriction *to $K := B(0,R)$ by*

$$V_K(t,\mu) := \Big\{ v_{|K} \in C^0(K,\mathbb{R}^d) \text{ s.t. } v \in V(t,\mu)\Big\},$$

for all $(t,\mu) \in [0,T]\times\mathscr{P}_c(\mathbb{R}^d)$. Here, $v_{|K}$ denotes the restriction to K of the map $v \in C^0(\mathbb{R}^d,\mathbb{R}^d)$. Conversely, observe that any function $v \in C^0(K,\mathbb{R}^d)$ can be extended to $\tilde{v} \in C^0(\mathbb{R}^d,\mathbb{R}^d)$ by defining $\tilde{v}(x) := v(\pi_K(x))$ for all $x \in \mathbb{R}^d$, where $\pi_K : \mathbb{R}^d \to K$ denotes the projection onto K.

We also define the closed convex hull $\overline{\mathrm{co}}\, V(t,\mu)$ *of $V(t,\mu) \subset C^0(\mathbb{R}^d,\mathbb{R}^d)$ as*

$$\overline{\mathrm{co}}\, V(t,\mu) := \Big\{ v \in C^0(\mathbb{R}^d,\mathbb{R}^d) \text{ s.t. } v_{|K} \in \overline{\mathrm{co}}\, V_K(t,\mu) \text{ for any } K := B(0,R) \text{ with } R > 0\Big\}, \tag{8}$$

where the closed convex hull $\overline{\mathrm{co}}\, V_K(t,\mu)$ is taken in the Banach space $(C^0(K,\mathbb{R}^d), \|\cdot\|_{C^0(K,\mathbb{R}^d)})$.

Throughout the remainder of this section, we fix a time horizon $T > 0$ and a real $p \in [1,+\infty)$.

Hypotheses (DI). *For every $R > 0$, assume that the following holds with $K := B(0,R)$.*

(*i*) *For any $\mu \in \mathscr{P}_c(\mathbb{R}^d)$, the set-valued map $t \in [0,T] \rightrightarrows V_K(t,\mu)$ is $\mathscr{L}^1$-measurable with closed non-empty images in $C^0(K,\mathbb{R}^d)$.*

(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* $\mu \in \mathscr{P}_c(\mathbb{R}^d)$*, every* $v \in V(t,\mu)$ *and all* $x \in \mathbb{R}^d$*, it holds*

$$|v(x)| \leq m(t)\Big(1 + |x| + \mathcal{M}_1(\mu)\Big).$$

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

$$\mathrm{Lip}\big(v(\cdot)\,;K\big) \leq l_K(t).$$

(iv) *There exists a map* $L_K(\cdot) \in L^1([0,T],\mathbb{R}_+)$ *such that for* $\mathscr{L}^1$*-almost every* $t \in [0,T]$ *and any* $\mu,\nu \in \mathscr{P}(K)$*, it holds*

$$V_K(t,\nu) \subset V_K(t,\mu) + L_K(t)W_p(\mu,\nu)\mathbb{B}_{C^0(K,\mathbb{R}^d)}.$$

Remark 2.9 (Non-local continuity equations as a particular case of continuity inclusions)**.** *Suppose that the multifunction* $V : [0,T]\times\mathscr{P}_c(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d,\mathbb{R}^d)$ *is* single-valued *for all* $(t,\mu) \in [0,T]\times\mathscr{P}_c(\mathbb{R}^d)$*. Then, solutions of* (6) *coincide with those of the non-local continuity equation*

$$\partial_t\mu(t) + \mathrm{div}\big(v(t,\mu(t))\mu(t)\big) = 0, \tag{9}$$

where $v(t,\mu,x) := V(t,\mu)(x)$ *for all* $(t,\mu,x) \in [0,T] \times \mathscr{P}_c(\mathbb{R}^d) \times \mathbb{R}^d$*.*

In the scenario described in Remark 2.9, hypotheses **(DI)** become a localised and time-dependent generalisation of that of [43, Section 1] in the spirit of Carathéodory ODEs. We state these assumptions separately for the sake of clarity, as they will be frequently used in the sequel.

Hypotheses (CE). *For every* $R > 0$*, suppose that the following holds with* $K := B(0,R)$*.*

(i) *The application* $t \in [0,T] \mapsto v(t,\mu,x) \in \mathbb{R}^d$ *is* $\mathscr{L}^1$*-measurable for any* $(\mu,x) \in \mathscr{P}_c(\mathbb{R}^d) \times \mathbb{R}^d$*. Moreover, there exists a map* $m(\cdot) \in L^1([0,T],\mathbb{R}_+)$ *such that*

$$|v(t,\mu,x)| \leq m(t)\Big(1 + |x| + \mathcal{M}_1(\mu)\Big),$$

for $\mathscr{L}^1$*-almost every* $t \in [0,T]$ *and any* $(\mu,x) \in \mathscr{P}_c(\mathbb{R}^d) \times \mathbb{R}^d$*.*

(ii) *There exist two maps* $l_K(\cdot), L_K(\cdot) \in L^1([0,T],\mathbb{R}_+)$ *such that for* $\mathscr{L}^1$*-almost every* $t \in [0,T]$*, any* $\mu,\nu \in \mathscr{P}(K)$ *and all* $x,y \in K$*, it holds*

$$\big|v(t,\mu,x) - v(t,\mu,y)\big| \leq l_K(t)|x-y| \qquad \textit{and} \qquad \big|v(t,\mu,x) - v(t,\nu,x)\big| \leq L_K(t)W_p(\mu,\nu).$$

Remark 2.10 (Local variant of hypotheses **(CE)**)**.** *In the sequel, we say that a Carathéodory vector-field* $w : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ *which is* independent from $\mu(\cdot)$ *satisfies hypotheses* **(CE)** *if it is sub-linear and locally Lipschitz with respect to the space variable, with constants* $m(\cdot), l_K(\cdot) \in L^1([0,T],\mathbb{R}_+)$*.*

In what follows, we recall structural results for continuity inclusions in Wasserstein spaces which were derived in [11]. From now on, we suppose that $V : [0,T] \times \mathscr{P}_c(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d,\mathbb{R}^d)$ is a set-valued map satisfying hypotheses **(DI)**, and consider $\mathscr{P}_c(\mathbb{R}^d)$ as a subset of the metric space $(\mathscr{P}_p(\mathbb{R}^d), W_p)$.

Theorem 2.11 (Existence and estimates on solutions of continuity inclusions)**.** *For every* $\mu^0 \in \mathscr{P}_c(\mathbb{R}^d)$*, there exists a curve of measures* $\mu(\cdot) \in \mathrm{AC}([0,T],\mathscr{P}_c(\mathbb{R}^d))$ *solution of the continuity inclusion* (6) *such that* $\mu(0) = \mu^0$*. Moreover if* $\mu^0 \in \mathscr{P}(B(0,r))$ *for some* $r > 0$*, then there exist* $R_r > 0$ *and* $m_r(\cdot) \in L^1([0,T],\mathbb{R}_+)$ *depending only on the magnitudes of* $r, \|m(\cdot)\|_1$ *such that*

$$\mathrm{supp}(\mu(t)) \subset B(0,R_r) \qquad \textit{and} \qquad W_p(\mu(t),\mu(s)) \leq \int_s^t m_r(\tau)\mathrm{d}\tau, \tag{10}$$

for every solution $\mu(\cdot)$ *of* (6) *and all times* $0 \leq s \leq t \leq T$*.*

In the following corollary, we state elementary consequences of Theorem 2.11 in the case where $V(\cdot,\cdot)$ is single-valued. We present this result separately for the sake of clarity, as we shall frequently use this variant in the sequel. Its last statement follows e.g. from [5, Chapter 8] and [43].

Corollary 2.12 (Non-local continuity equations)**.** *In Theorem 2.11 suppose that* $V : [0,T]\times\mathscr{P}_c(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d,\mathbb{R}^d)$ *is single-valued, and let* $\mu^0 \in \mathscr{P}(B(0,r))$ *for some* $r > 0$*. Then the non-local velocity-field* $v : [0,T] \times \mathscr{P}_c(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ *defined for all* $(t,\mu,x) \in [0,T] \times \mathscr{P}_c(\mathbb{R}^d) \times \mathbb{R}^d$ *by*

$$v(t,\mu,x) := V(t,\mu)(x),$$

satisfies hypotheses **(CE)***, and the non-local Cauchy problem* (9) *admits a unique solution* $\mu(\cdot) \in \mathrm{AC}([0,T],\mathscr{P}_c(\mathbb{R}^d))$*. Moreover, the curve* $\mu(\cdot)$ *satisfies the estimates of* (10) *and can be represented as*

$$\mu(t) = \Phi_{(0,t)}[\mu^0](\cdot)_{\#}\mu^0, \tag{11}$$

for all times $t \in [0,T]$*. Here,* $(\Phi_{(s,t)}[\mu(s)](\cdot))_{s,t\in[0,T]}$ *denotes the family of* non-local flows *defined by*

$$\Phi_{(s,t)}[\mu(s)](x) = x + \int_s^t v\Big(\sigma,\mu(\sigma),\Phi_{(s,\sigma)}[\mu(s)](x)\Big)\mathrm{d}\sigma, \tag{12}$$

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

We end this section by recalling three of the main results of [11], which are generalisations to the setting of continuity inclusions in Wasserstein spaces of Filippov's estimates, the Relaxation theorem, and the one-to-one correspondence between control systems and continuity inclusions.

Theorem 2.13 (Filippov's theorem)**.** *Let* $\nu(\cdot) \in \mathrm{AC}([0,T],\mathscr{P}(K_\nu))$ *be a solution of* (5) *generated by a Carathéodory velocity-field* $w : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ *with* $K_\nu := B(0,R_\nu)$ *for some* $R_\nu > 0$*. Furthermore, suppose that the* mismatch function $\eta_\nu : [0,T] \to \mathbb{R}_+$ *defined by*

$$\eta_\nu(t) := \mathrm{dist}_{C^0(K_\nu,\mathbb{R}^d)}\Big(w_{|K_\nu}(t); V_{K_\nu}(t,\nu(t))\Big),$$

is Lebesgue integrable over $[0,T]$*. Then for any* $\mu^0 \in \mathscr{P}(B(0,r))$ *with* $r > 0$*, there exists a solution* $\mu(\cdot) \in \mathrm{AC}([0,T],\mathscr{P}_c(\mathbb{R}^d))$ *of the continuity inclusion* (6) *such that for all times* $t \in [0,T]$*, it holds*

$$W_p(\mu(t),\nu(t)) \le C_p\left(W_p(\mu^0,\nu(0)) + \int_0^t \eta_\nu(s)\mathrm{d}s\right),$$

where $C_p > 0$ *is a constant which depends only on the magnitudes of* p, r, R_ν, T *and* $\|m(\cdot)\|_1$*.*

Theorem 2.14 (Relaxation theorem)**.** *Let* $\mu^0 \in \mathscr{P}_c(\mathbb{R}^d)$ *and* $\mu(\cdot) \in \mathrm{AC}([0,T],\mathscr{P}_c(\mathbb{R}^d))$ *be a solution of the* relaxed continuity inclusion

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

where $\overline{\mathrm{co}}\,V(t,\mu(t))$ *denotes the closed convex hull of* $V(t,\mu(t)) \subset C^0(\mathbb{R}^d,\mathbb{R}^d)$ *defined as in* (8)*. Then for any* $\delta > 0$*, there exists a solution* $\mu_\delta(\cdot) \in \mathrm{AC}([0,T],\mathscr{P}_c(\mathbb{R}^d))$ *of the continuity inclusion*

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

such that

$$W_p(\mu(t),\mu_\delta(t)) \le \delta,$$

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

Theorem 2.15 (Correspondence between control systems and continuity inclusions)**.** *Let (U, d_U) be a compact metric space and $v : [0,T] \times \mathscr{P}_c(\mathbb{R}^d) \times U \times \mathbb{R}^d \to \mathbb{R}^d$ be a controlled non-local velocity-field such that $(t,\mu,x) \mapsto v(t,\mu,u,x) \in \mathbb{R}^d$ satisfies hypotheses* **(CE)** *for every $u \in U$ with constants $m(\cdot), l_k(\cdot), L_K(\cdot) \in L^1([0,T],\mathbb{R}_+)$ independent of $u \in U$, and also $u \in U \mapsto v(t,\mu,u,x) \in \mathbb{R}^d$ is continuous for $\mathscr{L}^1$-almost every $t \in [0,T]$ and any $(\mu,x) \in \mathscr{P}_c(\mathbb{R}^d) \times \mathbb{R}^d$. Moreover, define the set-valued map $V : [0,T] \times \mathscr{P}_c(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d,\mathbb{R}^d)$ by*

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

for all $(t,\mu) \in [0,T] \times \mathscr{P}_c(\mathbb{R}^d)$.

Then, $V(\cdot,\cdot)$ satisfies hypotheses **(DI)***, and a curve of measures $\mu(\cdot)$ is a solution of the continuity inclusion* (6) if and only if *it solves the controlled non-local continuity equation*

$$\partial_t \mu(t) + \operatorname{div}\big(v(t,\mu(t),u(t))\mu(t)\big) = 0,$$

for some $\mathscr{L}^1$-measurable selection $t \in [0,T] \mapsto u(t) \in U$.

3 Differential calculus and linearised dynamics in $\mathscr{P}_c(\mathbb{R}^d)$

In this section, we investigate novel differential calculus tools for functionals and dynamical systems defined over $\mathscr{P}_c(\mathbb{R}^d)$ seen as a subset of $(\mathscr{P}_2(\mathbb{R}^d), W_2)$. In Section 3.1, we introduce a localised notion of $\mathscr{P}_2(\mathbb{R}^d)$-differentiability for compactly supported measures, and prove a new chain rule formula along arbitrary transport plans. In Section 3.2, we derive the general expression for the linearised Cauchy problem associated to a non-local continuity equation, and focus on the particular case in which the velocity perturbations are tangent to the set of admissible velocities of a continuity inclusion.

3.1 Localised theory of Wasserstein calculus

In this section, we study the properties of a new notion of *localised* subdifferential defined in the spirit of [5, 37] for extended real-valued functionals $\phi : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R} \cup \{\pm\infty\}$ such that $\mathscr{P}_c(\mathbb{R}^d) \subset D(\phi) := \{\mu \in \mathscr{P}_2(\mathbb{R}^d) \text{ s.t. } \phi(\mu) \neq \pm\infty\}$. In the sequel, we will use the condensed notation $\phi : \mathscr{P}_c(\mathbb{R}^d) \to \mathbb{R}$ to refer to any such functional.

Definition 3.1 (Localised sub and superdifferential)**.** *Let $\phi : \mathscr{P}_c(\mathbb{R}^d) \to \mathbb{R}$ and $\mu \in \mathscr{P}_c(\mathbb{R}^d)$. We say that a map $\xi \in L^2(\mathbb{R}^d,\mathbb{R}^d;\mu)$ belongs to the* localised subdifferential $\partial^-_{\mathrm{loc}}\phi(\mu)$ *of $\phi(\cdot)$ at μ provided that*

$$\phi(\nu) - \phi(\mu) \geq \inf_{\gamma \in \Gamma_o(\mu,\nu)} \int_{\mathbb{R}^{2d}} \langle \xi(x), y - x \rangle \mathrm{d}\gamma(x,y) + o_R(W_2(\mu,\nu)), \tag{13}$$

for every $R > 0$ and any $\nu \in \mathscr{P}(B_\mu(R))$, with $B_\mu(R)$ being defined as in (3)*. Furthermore, we say that a map $\xi \in L^2(\mathbb{R}^d,\mathbb{R}^d;\mu)$ belongs to the* strong localised subdifferential $\partial^-_{S,\mathrm{loc}}\phi(\mu)$ *if it satisfies*

$$\phi(\nu) - \phi(\mu) \geq \int_{\mathbb{R}^{2d}} \langle \xi(x), y - x \rangle \mathrm{d}\boldsymbol{\mu}(x,y) + o_R(W_{2,\boldsymbol{\mu}}(\mu,\nu)), \tag{14}$$

for every $R > 0$, any $\nu \in \mathscr{P}(B_\mu(R))$ and each $\boldsymbol{\mu} \in \Gamma(\mu,\nu)$, where we introduced the notation

$$W_{2,\boldsymbol{\mu}}(\mu,\nu) = \left(\int_{\mathbb{R}^{2d}} |x-y|^2 \mathrm{d}\boldsymbol{\mu}(x,y) \right)^{1/2}. \tag{15}$$

Analogously, we say that $\xi \in L^2(\mathbb{R}^d,\mathbb{R}^d;\mu)$ belongs to the localised superdifferential $\partial^+_{\mathrm{loc}}\phi(\mu)$ *of $\phi(\cdot)$ at μ if $(-\xi) \in \partial^-_{\mathrm{loc}}(-\phi)(\mu)$, with a similar definition for strong localised superdifferentials.*

Remark 3.2 (On the difference between (13) and (14))**.** *In line with the theory of subdifferential calculus on manifolds, localised subdifferentials are defined via their action on optimal displacement directions. As illustrated first in [40] (see also [5, Chapter 7]), such optimal directions are given by displacement interpolations of optimal plans, which leads to the geometric definition* (13)*.*

On the other hand, defining localised subdifferentials via their action on any transport direction as in (14) *produces a class of subdgradients that is much smaller, and which does not enjoy several of the desirable features of classical subdifferentials (such as being weakly closed, see e.g. [5, Remark 10.3.2]). Yet, strong subdifferentials are extremely useful in practice as they allow to differentiate functionals along directions that are not necessarily optimal. This property is of crucial interest in control theory, as the corresponding classes of admissible variations usually belong to this category.*

By a direct adaptation of the proofs of [37, Section 3] with finite radii $R > 0$, it can be shown that $\partial^-_{\mathrm{loc}}\phi(\mu) \cap \partial^+_{\mathrm{loc}}\phi(\mu)$ contains at most one element, which also belongs to $\mathrm{Tan}_\mu \mathscr{P}_2(\mathbb{R}^d)$. This motivates the following definition of *local differentiability* for functionals defined over $\mathscr{P}_c(\mathbb{R}^d)$.

Definition 3.3 (Locally differentiable functional)**.** *A functional $\phi : \mathscr{P}_c(\mathbb{R}^d) \to \mathbb{R}$ is* locally differentiable at $\mu \in \mathscr{P}_c(\mathbb{R}^d)$ *if there exists a map $\nabla\phi(\mu) \in \mathrm{Tan}_\mu \mathscr{P}_2(\mathbb{R}^d)$ – called the* Wasserstein gradient *of $\phi(\cdot)$ at μ –, such that $\partial^-_{\mathrm{loc}}\phi(\mu) \cap \partial^+_{\mathrm{loc}}\phi(\mu) = \{\nabla\phi(\mu)\}$. Similarly, we shall say that $\phi(\cdot)$ is locally differentiable over $\mathscr{P}_c(\mathbb{R}^d)$ if it is locally differentiable at every $\mu \in \mathscr{P}_c(\mathbb{R}^d)$.*

Remark 3.4 (Uniformity of the gradient)**.** *Observe that in Definition 3.3 above, we impose on the gradient $\nabla\phi(\mu) \in \partial^-_{\mathrm{loc}}\phi(\mu) \cap \partial^+_{\mathrm{loc}}\phi(\mu)$ to be independent from $R > 0$. This is due to the fact that $\nabla\phi(\mu)$ is an element of $\mathrm{Tan}_\mu \mathscr{P}_2(\mathbb{R}^d)$ – thus defined μ-almost everywhere over $\mathbb{R}^d$ –, so that it should not depend on the fattening parameter of the support of $\nu \in \mathscr{P}(B_\mu(R))$. The latter is solely used to control the small-o of the W_2-distance in* (13)*, as illustrated in Proposition 5.1 below.*

We provide examples of locally differentiable functionals in Section 5. In the next proposition, we prove that localised subdifferentials which belong to the analytical tangent space $\mathrm{Tan}_\mu \mathscr{P}_2(\mathbb{R}^d)$ are in fact strong localised subdifferentials. The proof of this result is inspired from that of [24, Chapter 1, Proposition 4.2], where a similar property is established for classical subdifferentials and measures $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ which are *absolutely continuous* with respect to the standard Lebesgue measure $\mathscr{L}^d$.

Proposition 3.5 (Tangent localised subdifferentials are strong)**.** *Let $\phi : \mathscr{P}_c(\mathbb{R}^d) \to \mathbb{R}$ and $\mu \in \mathscr{P}_c(\mathbb{R}^d)$ be such that $\partial^-_{\mathrm{loc}}\phi(\mu) \neq \emptyset$. Then*

$$\partial^-_{\mathrm{loc}}\phi(\mu) \cap \mathrm{Tan}_\mu \mathscr{P}_2(\mathbb{R}^d) \subset \partial^-_{S,\mathrm{loc}}\phi(\mu),$$

i.e. any localised subdifferential which belongs to the analytical tangent space $\mathrm{Tan}_\mu \mathscr{P}_2(\mathbb{R}^d)$ is also a strong localised subdifferential.

Proof. By contradiction, suppose that there exists $\xi \in \partial^-_{\mathrm{loc}}\phi(\mu) \cap \mathrm{Tan}_\mu \mathscr{P}_2(\mathbb{R}^d)$ which is not a strong localised subdifferential. Then for some $R > 0$ and $\delta > 0$, we can find a sequence of measures $(\mu_n) \subset \mathscr{P}(B_\mu(R))$ which converges towards μ in the W_2-metric along with a sequence of plans $(\boldsymbol{\mu}_n) \subset \Gamma(\mu, \mu_n)$ such that for any $n \geq 1$ large enough, it holds

$$\phi(\mu_n) - \phi(\mu) - \int_{\mathbb{R}^{2d}} \langle \xi(x), y - x \rangle \mathrm{d}\boldsymbol{\mu}_n(x, y) \leq -\delta\epsilon_n, \tag{16}$$

where $\epsilon_n := W_{2,\boldsymbol{\mu}_n}(\mu, \mu_n)$ is given by (15) and satisfies $\epsilon_n \to 0^+$ as $n \to +\infty$. Choose $\gamma_n \in \Gamma_o(\mu, \mu_n)$ for any $n \geq 1$, and observe that since $\xi \in \partial^-_{\mathrm{loc}}\phi(\mu)$, it holds for $n \geq 1$ large enough

$$\phi(\mu_n) - \phi(\mu) - \int_{\mathbb{R}^{2d}} \langle \xi(x), y - x \rangle \mathrm{d}\gamma_n(x, y) \geq -\tfrac{\delta}{2}\epsilon_n, \tag{17}$$

where we used the fact that $W_2(\mu, \mu_n) \leq W_{2,\boldsymbol{\mu}_n}(\mu, \mu_n)$. By merging (16) and (17), we obtain

$$\int_{\mathbb{R}^{2d}} \langle \xi(x), \tfrac{y-x}{\epsilon_n} \rangle \mathrm{d}(\gamma_n - \boldsymbol{\mu}_n)(x, y) \leq -\tfrac{\delta}{2}. \tag{18}$$

We now introduce the sequences of rescaled plans $(\hat{\gamma}_n)$ and $(\hat{\boldsymbol{\mu}}_n)$, defined for all $n \geq 1$ by

$$\hat{\gamma}_n := (\pi^1, \tfrac{\pi^2 - \pi^1}{\epsilon_n})_\# \gamma_n \qquad \text{and} \qquad \hat{\boldsymbol{\mu}}_n := (\pi^1, \tfrac{\pi^2 - \pi^1}{\epsilon_n})_\# \boldsymbol{\mu}_n,$$

so that (18) can be rewritten as

$$\int_{\mathbb{R}^{2d}} \langle \xi(x), r\rangle \mathrm{d}(\hat{\gamma}_n - \hat{\boldsymbol{\mu}}_n)(x,r) \leq -\tfrac{\delta}{2}. \tag{19}$$

We claim that both sequences $(\hat{\gamma}_n)$ and $(\hat{\boldsymbol{\mu}}_n)$ have uniformly bounded second order momentum. Indeed

$$\sup_{n\geq 1} \mathcal{M}_2^2(\hat{\gamma}_n) = \sup_{n\geq 1} \int_{\mathbb{R}^{2d}} \Big(|r|^2 + |x|^2\Big) \mathrm{d}\hat{\gamma}_n(x,r) = \sup_{n\geq 1} \int_{\mathbb{R}^{2d}} \Big(\big|\tfrac{x-y}{\epsilon_n}\big|^2 + |x|^2\Big) \mathrm{d}\gamma_n(x,y) \leq 1 + \mathcal{M}_2^2(\mu),$$

and the same estimate also holds for $(\hat{\boldsymbol{\mu}}_n)$. Whence, $(\hat{\gamma}_n)$ and $(\hat{\boldsymbol{\mu}}_n)$ are tight (see e.g. [5, Remark 5.1.5]), and thus narrowly compact by Prokhorov's theorem (see e.g. [5, Theorem 5.1.3]). In the sequel, we denote by $\hat{\gamma}$ and $\hat{\boldsymbol{\mu}}$ two of their cluster points that can be reached along the same subsequence.

Our goal now is to pass to the limit as $n \to +\infty$ in (19). Let $(\xi_k) \subset C_c^\infty(\mathbb{R}^d, \mathbb{R}^d)$ be a sequence such that $\xi_k \to \xi$ in $L^2(\mathbb{R}^d, \mathbb{R}^d; \mu)$ as $k \to +\infty$, and observe that

$$\begin{aligned}\int_{\mathbb{R}^{2d}} \langle \xi_k(x), r\rangle \mathrm{d}\hat{\gamma}_n(x,r) &\leq \int_{\mathbb{R}^{2d}} \langle \xi(x), r\rangle \mathrm{d}\hat{\gamma}_n(x,r) + \Big(\int_{\mathbb{R}^d} |\xi(x) - \xi_k(x)|^2 \mathrm{d}\mu(x)\Big)^{1/2} \Big(\int_{\mathbb{R}^{2d}} |r|^2 \mathrm{d}\hat{\gamma}_n(x,r)\Big)^{1/2} \\ &\leq \int_{\mathbb{R}^{2d}} \langle \xi(x), r\rangle \mathrm{d}\hat{\gamma}_n(x,r) + \|\xi - \xi_k\|_{L^2(\mu)},\end{aligned} \tag{20}$$

where we used Cauchy-Schwarz's inequality and the definition of $(\hat{\gamma}_n)$. Similarly, one also has

$$\int_{\mathbb{R}^{2d}} \langle \xi_k(x), r\rangle \mathrm{d}\hat{\boldsymbol{\mu}}_n(x,r) \geq \int_{\mathbb{R}^{2d}} \langle \xi(x), r\rangle \mathrm{d}\hat{\boldsymbol{\mu}}_n(x,r) - \|\xi - \xi_k\|_{L^2(\mu)} \tag{21}$$

for all $n, k \geq 1$. Whence by plugging (20) and (21) into (19), one deduces for all $n, k \geq 1$ large enough

$$\int_{\mathbb{R}^{2d}} \langle \xi_k(x), r\rangle \mathrm{d}(\hat{\gamma}_n - \hat{\boldsymbol{\mu}}_n)(x,r) \leq -\tfrac{\delta}{4}, \tag{22}$$

We now claim that given $k \geq 1$, the map $(x,r) \in \mathbb{R}^{2d} \mapsto |\langle \xi_k(x), r\rangle| \in \mathbb{R}_+$ is uniformly integrable with respect to both $(\hat{\gamma}_n)$ and $(\hat{\boldsymbol{\mu}}_n)$. Indeed setting $C_k := \|\xi_k(\cdot)\|_{C^0(\mathbb{R}^d, \mathbb{R}^d)}$, it holds

$$\int_{\{(x,r) \text{ s.t. } |\langle \xi_k(x), r\rangle| \geq i\}} |\langle \xi_k(x), r\rangle| \mathrm{d}\hat{\gamma}_n(x,r) \leq C_k \int_{\{(x,r) \text{ s.t. } |r| \geq i/C_k\}} |r| \mathrm{d}\hat{\gamma}_n(x,r), \tag{23}$$

for every $i \geq 1$, while on the other hand it can be checked by definition of $(\hat{\gamma}_n)$ that

$$\frac{i}{C_k} \int_{\{(x,r) \text{ s.t. } |r| \geq i/C_k\}} |r| \mathrm{d}\hat{\gamma}_n(x,r) \leq \int_{\mathbb{R}^{2d}} |r|^2 \mathrm{d}\hat{\gamma}_n(x,r) \leq 1, \tag{24}$$

for any $n, k, i \geq 1$. Merging (23) and (24), we therefore obtain

$$\int_{\{(x,r) \text{ s.t. } |\langle \xi_k(x), r\rangle| \geq i\}} |\langle \xi_k(x), r\rangle| \mathrm{d}\hat{\gamma}_n(x,r) \leq \tfrac{C_k^2}{i} \underset{i \to +\infty}{\longrightarrow} 0,$$

so that the $(x,r) \in \mathbb{R}^{2d} \mapsto |\langle \xi_k(x), r\rangle| \in \mathbb{R}_+$ is uniformly integrable with respect to $(\hat{\gamma}_n)$ for all $k \geq 1$. By repeating the same arguments for $(\hat{\boldsymbol{\mu}}_n)$ and recalling that the latter map is continuous, we recover

$$\int_{\mathbb{R}^{2d}} \langle \xi_k(x), r\rangle \mathrm{d}(\hat{\gamma}_n - \hat{\boldsymbol{\mu}}_n)(x,r) \underset{n \to +\infty}{\longrightarrow} \int_{\mathbb{R}^{2d}} \langle \xi_k(x), r\rangle \mathrm{d}(\hat{\gamma} - \hat{\boldsymbol{\mu}})(x,r), \tag{25}$$

for all $k \geq 1$ by applying the convergence results of [5, Lemma 5.1.7]. Upon combining (22) and (25) while letting $k \to +\infty$, we can thus conclude

$$\int_{\mathbb{R}^{2d}} \langle \xi(x), r\rangle \mathrm{d}(\hat{\gamma} - \hat{\boldsymbol{\mu}})(x,r) \leq -\tfrac{\delta}{4}, \tag{26}$$

where we used the fact that $\pi^1_\# \hat{\gamma} = \pi^1_\# \hat{\boldsymbol{\mu}} = \mu$ as a consequence of classical stability results under narrow convergence for transport plans (see e.g. [5, Chapter 5]).

Let us now choose an arbitrary test function $\zeta \in C_c^\infty(\mathbb{R}^d)$. First, notice that there exists a constant $C_\zeta > 0$ such that the global estimates

$$\begin{cases} \zeta(y) - \zeta(x) \leq \langle \nabla\zeta(x), y - x\rangle + C_\zeta |x-y|^2, \\ \zeta(x) - \zeta(y) \leq \langle \nabla\zeta(x), x - y\rangle + C_\zeta |x-y|^2, \end{cases} \tag{27}$$

hold for any $x, y \in \mathbb{R}^d$. By using the first estimates of (27), we further obtain for any $n \geq 1$

$$\int_{\mathbb{R}^{2d}} \zeta(y) \mathrm{d}\gamma_n(x,y) \leq \int_{\mathbb{R}^{2d}} \Big(\zeta(x) + \langle \nabla\zeta(x), y - x\rangle\Big) \mathrm{d}\gamma_n(x,y) + C_\zeta W_2^2(\mu, \mu_n), \tag{28}$$

where we used the fact that $\gamma_n \in \Gamma_o(\mu, \mu_n)$, while the second estimate of (27) further yields

$$-\int_{\mathbb{R}^{2d}} \zeta(y) \mathrm{d}\boldsymbol{\mu}_n(x,y) \leq -\int_{\mathbb{R}^{2d}} \Big(\zeta(x) + \langle \nabla\zeta(x), y - x\rangle\Big) \mathrm{d}\boldsymbol{\mu}_n(x,y) + C_\zeta W_{2,\boldsymbol{\mu}_n}^2(\mu, \mu_n), \tag{29}$$

by the definition (15) of $W_{2,\boldsymbol{\mu}_n}(\mu, \mu_n)$. In addition, observe that since $\gamma_n, \boldsymbol{\mu}_n \in \Gamma(\mu, \mu_n)$, one also has

$$\int_{\mathbb{R}^{2d}} \zeta(x) \mathrm{d}(\gamma_n - \boldsymbol{\mu}_n)(x,y) = 0 \qquad \text{and} \qquad \int_{\mathbb{R}^{2d}} \zeta(y) \mathrm{d}(\gamma_n - \boldsymbol{\mu}_n)(x,y) = 0, \tag{30}$$

for all $n \geq 1$. Therefore, combining the estimates of (28) and (29) while using (30) and recalling that $W_2(\mu, \mu_n) \leq W_{2,\boldsymbol{\mu}_n}(\mu, \mu_n) = \epsilon_n$, we obtain

$$0 \leq \int_{\mathbb{R}^{2d}} \langle \nabla\zeta(x), y - x\rangle \mathrm{d}(\gamma_n - \boldsymbol{\mu}_n)(x,y) + 2C_\zeta \epsilon_n^2.$$

By dividing both sides of the previous inequality by $\epsilon_n > 0$, considering again the sequences of rescaled plans $(\hat{\gamma}_n)$ and $(\hat{\boldsymbol{\mu}}_n)$ and letting $n \to +\infty$ along an adequate subsequence, we finally recover

$$0 \leq \int_{\mathbb{R}^{2d}} \langle \nabla\zeta(x), r\rangle \mathrm{d}(\hat{\gamma} - \hat{\boldsymbol{\mu}})(x,r), \tag{31}$$

for any $\zeta \in C_c^\infty(\mathbb{R}^d)$. Since $\xi \in \mathrm{Tan}_\mu \mathscr{P}_2(\mathbb{R}^d)$, there exists a sequence $(\zeta_n) \subset C_c^\infty(\mathbb{R}^d)$ such that $\nabla\zeta_n \to \xi$ in $L^2(\mathbb{R}^d, \mathbb{R}^d; \mu)$ as $n \to +\infty$, which leads to a contradiction between (26) and (31). □

By combining Definition 3.3 and Proposition 3.5, we obtain a general chain rule along arbitrary transport plans for locally differentiable functionals.

Proposition 3.6 (Chain rule along arbitrary transport plans)**.** *Let $\phi : \mathscr{P}_c(\mathbb{R}^d) \to \mathbb{R}$, $\mu \in \mathscr{P}_c(\mathbb{R}^d)$ and suppose that $\phi(\cdot)$ is locally differentiable at μ. Then for every $R > 0$ and $\nu \in \mathscr{P}(B_\mu(R))$, it holds*

$$\phi(\nu) = \phi(\mu) + \int_{\mathbb{R}^{2d}} \langle \nabla\phi(\mu)(x), y - x\rangle \mathrm{d}\boldsymbol{\mu}(x,y) + o_R(W_{2,\boldsymbol{\mu}}(\mu, \nu)), \tag{32}$$

for any $\boldsymbol{\mu} \in \Gamma(\mu, \nu)$. Conversely, if there exists a map $\nabla\phi(\mu) \in \mathrm{Tan}_\mu \mathscr{P}_2(\mathbb{R}^d)$ such that for every $R > 0$ and $\nu \in \mathscr{P}(B_\mu(R))$ the first-order expansion (32) *holds along any optimal plan $\boldsymbol{\mu} \in \Gamma_o(\mu, \nu)$, then $\phi(\cdot)$ is locally differentiable at μ in the sense of Definition 3.3.*

Proof. By Definition 3.3, we have that $\nabla\phi(\mu) \in \partial_{\mathrm{loc}}^- \phi(\mu) \cap \partial_{\mathrm{loc}}^+ \phi(\mu) \cap \mathrm{Tan}_\mu \mathscr{P}_2(\mathbb{R}^d)$. Whence by Proposition 3.5, it further holds that $\nabla\phi(\mu) \in \partial_{S,\mathrm{loc}}^- \phi(\mu) \cap \partial_{S,\mathrm{loc}}^+ \phi(\mu)$, so that (32) is satisfied for every $\nu \in \mathscr{P}(B_\mu(R))$ and any $\boldsymbol{\mu} \in \Gamma(\mu, \nu)$. Conversely, let $\nu \in \mathscr{P}(B_\mu(R))$ for some $R > 0$, and observe that for every optimal transport plan $\boldsymbol{\mu} \in \Gamma_o(\mu, \nu)$, it holds

$$\int_{\mathbb{R}^{2d}} \langle \nabla\phi(\mu)(x), y - x\rangle \mathrm{d}\boldsymbol{\mu}(x,y) \geq \inf_{\gamma \in \Gamma_o(\mu,\nu)} \int_{\mathbb{R}^{2d}} \langle \nabla\phi(\mu), y - x\rangle \mathrm{d}\gamma(x,y),$$

so that (32) directly yields that $\nabla\phi(\mu) \in \partial_{\mathrm{loc}}^- \phi(\mu)$ since $W_{2,\boldsymbol{\mu}}(\mu, \nu) = W_2(\mu, \nu)$ when $\boldsymbol{\mu} \in \Gamma_o(\mu, \nu)$ and both $R > 0$ and $\nu \in \mathscr{P}(B_\mu(R))$ are arbitrary. Similarly, one can show that $\nabla\phi(\mu) \in \partial_{\mathrm{loc}}^+ \phi(\mu)$. Hence, $\nabla\phi(\mu) \in \partial_{\mathrm{loc}}^- \phi(\mu) \cap \partial_{\mathrm{loc}}^+ \phi(\mu) \cap \mathrm{Tan}_\mu \mathscr{P}_2(\mathbb{R}^d)$, and the functional $\phi(\cdot)$ is locally differentiable at μ in the sense of Definition 3.3. □

Corollary 3.7 (Chain rule along perturbations in $L^\infty(\mathbb{R}^d,\mathbb{R}^d;\mu)$)**.** *Let $\phi:\mathscr{P}_c(\mathbb{R}^d)\to\mathbb{R}$, $\mu\in\mathscr{P}_c(\mathbb{R}^d)$ and suppose that $\phi(\cdot)$ is locally differentiable at μ. Then for every $\mathcal{F}\in L^\infty(\mathbb{R}^d,\mathbb{R}^d;\mu)$, it holds*

$$\phi((\mathrm{Id}+\epsilon\mathcal{F})_\#\mu)=\phi(\mu)+\epsilon\int_{\mathbb{R}^d}\langle\nabla\phi(\mu)(x),\mathcal{F}(x)\rangle\mathrm{d}\mu(x)+o_{R_\mathcal{F}}(\epsilon),$$

for all $\epsilon\in[-1,1]$, where $R_\mathcal{F}:=\|\mathcal{F}\|_{L^\infty(\mu)}$.

Proof. Observe first that for every $\mathcal{F}\in L^\infty(\mathbb{R}^d,\mathbb{R}^d;\mu)$, the measures $(\mathrm{Id}+\epsilon\mathcal{F})_\#\mu$ have compact support in $B_{R_\mathcal{F}}(\mu)$ for all $\epsilon\in[-1,1]$. By applying Proposition 3.6 with the transport plan $\boldsymbol{\mu}_\epsilon:=(\mathrm{Id},\mathrm{Id}+\epsilon\mathcal{F})_\#\mu$, we therefore obtain

$$\begin{aligned}\phi((\mathrm{Id}+\epsilon\mathcal{F})_\#\mu)&=\phi(\mu)+\int_{\mathbb{R}^{2d}}\langle\nabla\phi(\mu)(x),y-x\rangle\mathrm{d}\boldsymbol{\mu}_\epsilon(x,y)+o_{R_\mathcal{F}}\big(W_{2,\boldsymbol{\mu}_\epsilon}(\mu,(\mathrm{Id}+\epsilon\mathcal{F})_\#\mu)\big)\\&=\phi(\mu)+\epsilon\int_{\mathbb{R}^d}\langle\nabla\phi(\mu)(x),\mathcal{F}(x)\rangle\mathrm{d}\mu(x)+o_{R_\mathcal{F}}(\epsilon),\end{aligned}$$

where we used the fact that $W_{2,\boldsymbol{\mu}_\epsilon}(\mu,(\mathrm{Id}+\epsilon\mathcal{F})_\#\mu)=\epsilon\,\|\mathcal{F}\|_{L^2(\mu)}$ as a direct consequence of (15). ☐

We end this section by adapting the notion of local differentiability to vector-valued maps.

Definition 3.8 (Local differentiability of vector-valued maps)**.** *For $m\geq1$, a map $\phi:\mathscr{P}_c(\mathbb{R}^d)\to\mathbb{R}^m$ is said to be* locally differentiable *at $\mu\in\mathscr{P}_c(\mathbb{R}^d)$ if its components $\{\phi_i(\cdot)\}_{1\leq i\leq m}$ are locally differentiable.*

Remark 3.9 (Chainrule for vector-valued maps)**.** *When $m=d$ and $\phi:\mathscr{P}_c(\mathbb{R}^d)\to\mathbb{R}^d$ is locally differentiable at $\mu\in\mathscr{P}_c(\mathbb{R}^d)$, we can apply Proposition 3.6 componentwisely to obtain*

$$\phi(\nu)=\phi(\mu)+\int_{\mathbb{R}^{2d}}\mathrm{D}_\mu\phi(\mu)(x)(y-x)\mathrm{d}\boldsymbol{\mu}(x,y)+o_R(W_{2,\boldsymbol{\mu}}(\mu,\nu)),\tag{33}$$

for every $\nu\in\mathscr{P}(B_\mu(R))$ with $R>0$, where $\mathrm{D}_\mu\phi(\mu)(\cdot):=(\nabla\phi_i(\mu)(\cdot))_{1\leq i\leq d}$ is the matrix-valued map which rows are the Wasserstein gradients of the components $\{\phi_i(\cdot)\}_{1\leq i\leq d}$. Similarly, by applying componentwisely the statements of Corollary 3.7, it also holds

$$\phi((\mathrm{Id}+\epsilon\mathcal{F})_\#\mu)=\phi(\mu)+\epsilon\int_{\mathbb{R}^d}\mathrm{D}_\mu\phi(\mu)(x)\mathcal{F}(x)\mathrm{d}\mu(x)+o_{R_\mathcal{F}}(\epsilon),\tag{34}$$

for all $\epsilon\in[-1,1]$ and any $\mathcal{F}\in L^\infty(\mathbb{R}^d,\mathbb{R}^d;\mu)$.

3.2 Linearisations of non-local continuity equations

In this section, we derive the expression of the canonical linearised Cauchy problem associated to a non-local continuity equation with smooth driving velocity-field. In what follows we fix $p\in[1,+\infty)$ and for any compact set $K\subset\mathbb{R}^d$, we shall consider $\mathscr{P}(K)$ as a subset of the metric space $(\mathscr{P}_p(\mathbb{R}^d),W_p)$. We henceforth focus our attention on the Cauchy problem

$$\begin{cases}\partial_t\mu(t)+\mathrm{div}\big(v(t,\mu(t))\mu(t)\big)=0,\\ \mu(0)=\mu^0\in\mathscr{P}_c(\mathbb{R}^d),\end{cases}\tag{35}$$

driven by a non-local velocity-field $v:[0,T]\times\mathscr{P}_c(\mathbb{R}^d)\times\mathbb{R}^d\mapsto\mathbb{R}^d$ satisfying the following assumptions.

Hypotheses (H). *For every $R>0$, assume that the following holds with $K:=B(0,R)$.*

(i) The non-local velocity field $v:[0,T]\times\mathscr{P}_c(\mathbb{R}^d)\times\mathbb{R}^d\to\mathbb{R}^d$ satisfies hypotheses **(CE)**.

(ii) For $\mathscr{L}^1$-almost every $t\in[0,T]$ and any $\mu\in\mathscr{P}_c(\mathbb{R}^d)$, the map $x\in\mathbb{R}^d\mapsto v(t,\mu,x)\in\mathbb{R}^d$ is Fréchet differentiable, and the application $(\mu,x)\in\mathscr{P}(K)\times K\mapsto\mathrm{D}_xv(t,\mu,x)\in\mathbb{R}^{d\times d}$ is continuous.

(iii) For $\mathscr{L}^1$-almost every $t\in[0,T]$ and any $x\in\mathbb{R}^d$, the map $\mu\in\mathscr{P}_c(\mathbb{R}^d)\mapsto v(t,\mu,x)\in\mathbb{R}^d$ is locally differentiable in the sense of Definition 3.8, and the application $(\mu,x,y)\in\mathscr{P}(K)\times K\times K\mapsto\mathrm{D}_\mu v(t,\mu,x)(y)\in\mathbb{R}^{d\times d}$ is continuous.

In the following proposition, we derive a general first-order linearisation formula for solutions of (35). Similar techniques were applied in [10, 12] in the particular case of perturbations induced by *needle variations* (see e.g. [47]).

Proposition 3.10 (Linearisation of non-local continuity equations)**.** *Let $\mu^0 \in \mathscr{P}(B(0,r))$ for some $r > 0$, $v : [0,T] \times \mathscr{P}_c(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ be a non-local velocity-field satisfying hypotheses* **(H)***, and $\mu(\cdot) \in \mathrm{AC}([0,T], \mathscr{P}(K))$ be the corresponding solution of* (35) *where $K := B(0,R_r)$ is given by* (10)*. Let $\mathcal{F}^0 \in C^0(\mathbb{R}^d, \mathbb{R}^d)$, $w : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ be a Carathéodory vector-field satisfying* **(CE)** *and consider the perturbed Cauchy problems*

$$\begin{cases} \partial_t \mu_\epsilon(t) + \mathrm{div}\Big(\big(v(t,\mu_\epsilon(t)) + \epsilon w(t)\big)\mu_\epsilon(t)\Big) = 0, \\ \mu_\epsilon(0) = (\mathrm{Id} + \epsilon \mathcal{F}^0)_\# \mu^0. \end{cases} \tag{36}$$

for $\epsilon \in [-1,1]$.

Then, there exists a family of maps $(\mathcal{G}_\epsilon(\cdot,\cdot)) \subset C^0([0,T] \times K, \mathbb{R}^d)$ such that for every $\epsilon \in [-1,1]$, the solution $\mu_\epsilon(\cdot) \in \mathrm{AC}([0,T], \mathscr{P}_c(\mathbb{R}^d))$ of (36) *can be expressed explicitly as*

$$\mu_\epsilon(t) = \mathcal{G}_\epsilon(t,\cdot)_\# \mu(t), \tag{37}$$

for all times $t \in [0,T]$. Moreover, the application $\epsilon \in [-1,1] \mapsto \mathcal{G}_\epsilon(t,\cdot) \in C^0(K, \mathbb{R}^d)$ admits a Taylor expansion at $\epsilon = 0$, of the form

$$\mathcal{G}_\epsilon(t,\cdot) = \mathrm{Id} + \epsilon \mathcal{F}\Big(t, \Phi_{(t,0)}[\mu(t)](\cdot)\Big) + o_t(\epsilon), \tag{38}$$

where $\sup_{t\in[0,T]} \|o_t(\epsilon)\|_{C^0(K,\mathbb{R}^d)} = o(\epsilon)$ and $\mathcal{F} \in C^0([0,T] \times K, \mathbb{R}^d)$ is the unique solution of the linearised non-local Cauchy problem

$$\begin{cases} \partial_t \mathcal{F}(t,x) = \mathrm{D}_x v\Big(t, \mu(t), \Phi_{(0,t)}[\mu^0](x)\Big)\mathcal{F}(t,x) + w\Big(t, \Phi_{(0,t)}[\mu^0](x)\Big) \\ \qquad\qquad + \displaystyle\int_{\mathbb{R}^d} \mathrm{D}_\mu v\Big(t, \mu(t), \Phi_{(0,t)}[\mu^0](x)\Big)\big(\Phi_{(0,t)}[\mu^0](y)\big)\mathcal{F}(t,y)\mathrm{d}\mu^0(y), \\ \mathcal{F}(0,x) = \mathcal{F}^0(x). \end{cases} \tag{39}$$

Proof. First, remark that for any $\epsilon \in [-1,1]$, the non-local velocity-field $v_\epsilon : [0,T] \times \mathscr{P}_c(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ defined for any $(t,\mu,x) \in [0,T] \times \mathscr{P}_c(\mathbb{R}^d) \times \mathbb{R}^d$ by

$$v_\epsilon(t,\mu,x) := v(t,\mu,x) + \epsilon w(t,x), \tag{40}$$

satisfies hypotheses **(CE)**. Whence, denoting $\mu^0_\epsilon := (\mathrm{Id} + \epsilon \mathcal{F}^0)_\# \mu^0$, there exists by Corollary 2.12 a unique curve of measures $\mu_\epsilon(\cdot)$ solution of (36) which can be represented as

$$\mu_\epsilon(t) = \Phi^\epsilon_{(0,t)}[\mu^0_\epsilon](\cdot)_\# \mu^0_\epsilon, \tag{41}$$

where $(\Phi^\epsilon_{(0,t)}[\mu^0_\epsilon](\cdot))_{t\in[0,T]}$ stands for the non-local flow generated by $v_\epsilon : [0,T] \times \mathscr{P}_c(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ starting from μ^0_ϵ, defined as in (12). For the sake of readability, we shall use the condensed notations

$$\Phi_{(s,t)}(x) := \Phi_{(s,t)}[\mu(s)](x) \qquad \text{and} \qquad \Phi^\epsilon_{(s,t)}(x) := \Phi^\epsilon_{(s,t)}[\mu_\epsilon(s)](x),$$

throughout the remainder of the proof. Recalling that $\mu^0_\epsilon = (\mathrm{Id} + \epsilon \mathcal{F}^0)_\# \mu^0$ and that the curve $\mu(\cdot)$ also admits the flow representation (11), the expression in (41) can be rewritten as

$$\mu_\epsilon(t) = \Big(\Phi^\epsilon_{(0,t)} \circ (\mathrm{Id} + \epsilon \mathcal{F}^0) \circ \Phi_{(t,0)}(\cdot)\Big)_\# \mu(t), \tag{42}$$

for all times $t \in [0,T]$, where "$\circ$" is the standard composition operation between functions. Hence, (37) holds with the family of maps $(\mathcal{G}_\epsilon(\cdot,\cdot))$ defined by

$$\mathcal{G}_\epsilon : (t,x) \in [0,T] \times K \mapsto \Phi^\epsilon_{(0,t)} \circ (\mathrm{Id} + \epsilon \mathcal{F}^0) \circ \Phi_{(t,0)}(x), \tag{43}$$

for every $\epsilon \in [-1,1]$. Our goal now is to establish the first-order expansion of (38).

By adapting the argument of [12, Proposition 5] to the case of non-local velocity-fields satisfying the localised differentiability assumptions **(H)**-(ii) and **(H)**-(iii), it can be shown that the map

$$\epsilon \in [-1,1] \mapsto \Phi^{\epsilon}_{(0,\cdot)}(\cdot) \in C^0([0,T] \times K', \mathbb{R}^d),$$

is Fréchet differentiable at $\epsilon = 0$, where $K' := B(0, R'_r)$ and $R'_r > 0$ are such that $(\mathrm{Id} + \epsilon \mathcal{F}^0)(K) \subset K'$ for all $\epsilon \in [-1,1]$. Denoting by $\Psi \in C^0([0,T] \times K', \mathbb{R}^d)$ the corresponding Fréchet derivative, one has

$$\Phi^{\epsilon}_{(0,t)}(y) = \Phi_{(0,t)}(y) + \epsilon \Psi(t,y) + o_{t,y}(\epsilon), \tag{44}$$

for any $(t,y) \in [0,T] \times K'$, where $\sup_{(t,y)\in[0,T]\times K'} |o_{t,y}(\epsilon)| = o(\epsilon)$. Under hypotheses **(H)**-(i) and **(H)**-(ii), it is a standard result in the theory of Carathéodory ODEs (see e.g. [14, Theorem 2.3.2]) that the flow map $x \in \mathbb{R}^d \mapsto \Phi_{(0,t)}(x) \in \mathbb{R}^d$ is Fréchet-differentiable with

$$\Phi_{(0,t)}(x + \epsilon h) = \Phi_{(0,t)}(x) + \epsilon \mathrm{D}_x \Phi_{(0,t)}(x) h + o_{t,x,h}(\epsilon), \tag{45}$$

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

$$\sup_{(t,x,h)\in[0,T]\times K\times B(0,r')} |o_{t,x,h}(\epsilon)| = o_{r'}(\epsilon),$$

for every $r' > 0$. By merging (44) and (45), we obtain for all $(t,x) \in [0,T] \times K$ that

$$\Phi^{\epsilon}_{(0,t)}(x + \epsilon \mathcal{F}^0(x)) = \Phi_{(0,t)}(x) + \epsilon \mathrm{D}_x \Phi_{(0,t)}(x) \mathcal{F}^0(x) + \epsilon \Psi(t,x) + o_{t,x}(\epsilon), \tag{46}$$

where $\sup_{(t,x)\in[0,T]\times K} |o_{t,x}(\epsilon)| = o(\epsilon)$. We now define the map $\mathcal{F} \in C^0([0,T] \times K, \mathbb{R}^d)$ by

$$\mathcal{F}(t,x) := \mathrm{D}_x \Phi_{(0,t)}(x) \mathcal{F}^0(x) + \Psi(t,x), \tag{47}$$

for any $(t,x) \in [0,T] \times K$. By merging (43) and (46), we obtain that $\epsilon \in [-1,1] \mapsto \mathcal{G}_\epsilon(t,\cdot) \in C^0(K, \mathbb{R}^d)$ satisfies for all times $t \in [0,T]$ the following Taylor expansion at $\epsilon = 0$

$$\mathcal{G}_\epsilon(t,\cdot) = \mathrm{Id} + \epsilon \mathcal{F}(t, \Phi_{(t,0)}(\cdot)) + o_t(\epsilon), \tag{48}$$

where $\sup_{t\in[0,T]} \|o_t(\epsilon)\|_{C^0(K,\mathbb{R}^d)} = o(\epsilon)$ for every $\epsilon \in [-1,1]$.

We now prove that the map $\mathcal{F} \in C^0([0,T] \times K, \mathbb{R}^d)$ defined in (47) is the unique solution of the linearised Cauchy problem (39). First, recall that by classical results on differentials of flow maps (see e.g. [14, Theorem 2.3.2]), the application $t \in [0,T] \mapsto \mathrm{D}_x \Phi_{(0,t)}(x) \mathcal{F}^0(x) \in \mathbb{R}^d$ is the unique solution of

$$\mathrm{D}_x \Phi_{(0,t)}(x) \mathcal{F}^0(x) = \mathcal{F}^0(x) + \int_0^t \mathrm{D}_x v\Big(s, \mu(s), \Phi_{(0,s)}(x)\Big) \Big(\mathrm{D}_x \Phi_{(0,s)}(x) \mathcal{F}^0(x)\Big) \mathrm{d}s, \tag{49}$$

for any $x \in K$. By (44), it further holds

$$\begin{aligned} v\Big(s, \mu_\epsilon(s), \Phi^{\epsilon}_{(0,s)}(x)\Big) &= v\Big(s, \mu_\epsilon(s), \Phi_{(0,s)}(x) + \epsilon \Psi(s,x) + o_{s,x}(\epsilon)\Big) \\ &= v\Big(s, \mu_\epsilon(s), \Phi_{(0,s)}(x)\Big) + \epsilon \mathrm{D}_x v\Big(s, \mu_\epsilon(s), \Phi_{(0,s)}(x)\Big) \Psi(s,x) + o_{s,x}(\epsilon) \\ &= v\Big(s, \mu_\epsilon(s), \Phi_{(0,s)}(x)\Big) + \epsilon \mathrm{D}_x v\Big(s, \mu(s), \Phi_{(0,s)}(x)\Big) \Psi(s,x) + o_{s,x}(\epsilon), \end{aligned} \tag{50}$$

for $\mathscr{L}^1$-almost every $s \in [0,T]$ and any $(\epsilon, x) \in [-1,1] \times K'$, with

$$\int_0^T \sup_{x\in K'} |o_{s,x}(\epsilon)| \mathrm{d}s = o(\epsilon),$$

and where we also used that $\nu \in \mathscr{P}(K') \mapsto \mathrm{D}_x v(s,\nu,\Phi_{(0,s)}(x))$ is continuous by **(H)**-(ii). Using **(H)**-(iii), we can in turn apply the vector-valued chain rule of Remark 3.9 together with (42), (43) and (48) to obtain

$$\begin{aligned}
&v\Big(s,\mu_\epsilon(s),\Phi_{(0,s)}(x)\Big)\\
&= v\Big(s,\Big(\mathrm{Id}+\epsilon\mathcal{F}(s,\Phi_{(s,0)}(\cdot))+o_s(\epsilon)\Big)_{\#}\mu(s),\Phi_{(0,s)}(x)\Big)\\
&= v\Big(s,\mu(s),\Phi_{(0,s)}(x)\Big)+\epsilon\int_{\mathbb{R}^d}\mathrm{D}_\mu v\Big(s,\mu(s),\Phi_{(0,s)}(x)\Big)(y)\mathcal{F}(s,\Phi_{(s,0)}(y))\mathrm{d}\mu(s)(y)+o_s(\epsilon)\\
&= v\Big(s,\mu(s),\Phi_{(0,s)}(x)\Big)+\epsilon\int_{\mathbb{R}^d}\mathrm{D}_\mu v\Big(s,\mu(s),\Phi_{(0,s)}(x)\Big)(\Phi_{(0,s)}(y))\,\mathcal{F}(s,y)\mathrm{d}\mu^0(y)+o_s(\epsilon),
\end{aligned}\tag{51}$$

for $\mathscr{L}^1$-almost every $s\in[0,T]$ and all $y\in K'$. Observe now that for every $\epsilon\in[-1,1]$, the family of perturbed flows $(\Phi^\epsilon_{(0,t)}(\cdot))_{t\in[0,T]}$ can be characterised as the unique solution of the Carathéodory ODE

$$\begin{aligned}
\Phi^\epsilon_{(0,t)}(y) &= y+\int_0^t v_\epsilon\Big(s,\Phi^\epsilon_{(0,s)}(\cdot)_{\#}\mu^0_\epsilon,\Phi^\epsilon_{(0,s)}(y)\Big)\mathrm{d}s\\
&= y+\int_0^t\Big(v\Big(s,\mu_\epsilon(s),\Phi^\epsilon_{(0,s)}(y)\Big)+\epsilon w\Big(s,\Phi^\epsilon_{(0,s)}(y)\Big)\Big)\mathrm{d}s,
\end{aligned}\tag{52}$$

for any $(t,y)\in[0,T]\times K'$. Therefore by plugging (44), (50) and (51) into (52), we further obtain

$$\begin{aligned}
\Phi_{(0,t)}(x)+\epsilon\Psi(t,x) = x &+\int_0^t v\Big(s,\mu(s),\Phi_{(0,s)}(x)\Big)\mathrm{d}s+\epsilon\int_0^t\mathrm{D}_x v\Big(s,\mu(s),\Phi_{(0,s)}(x)\Big)\Psi(s,x)\mathrm{d}s\\
&+\epsilon\int_0^t\int_{\mathbb{R}^d}\mathrm{D}_\mu v\Big(s,\mu(s),\Phi_{(0,s)}(x)\Big)(\Phi_{(0,s)}(y))\,\mathcal{F}(s,y)\mathrm{d}\mu^0(y)\mathrm{d}s\\
&+\epsilon\int_0^t w\Big(s,\Phi_{(0,s)}(x)\Big)\mathrm{d}s+o_{t,x}(\epsilon),
\end{aligned}$$

for any $(t,x)\in[0,T]\times K$ and all $\epsilon\in[-1,1]$, where $\sup_{(t,x)\in[0,T]\times K}|o_{t,x}(\epsilon)|=o(\epsilon)$. By identifying terms and observing that the flow maps $(\Phi_{(0,t)}(\cdot))_{t\in[0,T]}$ satisfy (12), we recover the ODE characterisation

$$\begin{aligned}
\Psi(t,x) = &\int_0^t\Big(\mathrm{D}_x v\Big(s,\mu(s),\Phi_{(0,s)}(x)\Big)\Psi(s,x)+w\Big(s,\Phi_{(0,s)}(x)\Big)\Big)\mathrm{d}s\\
&+\int_0^t\int_{\mathbb{R}^d}\mathrm{D}_\mu v\Big(s,\mu(s),\Phi_{(0,s)}(x)\Big)(\Phi_{(0,s)}(y))\,\mathcal{F}(s,y)\mathrm{d}\mu^0(y)\mathrm{d}s,
\end{aligned}\tag{53}$$

for all $(t,x)\in[0,T]\times K$. Upon combining (49) and (53), we conclude that the map $\mathcal{F}\in C^0([0,T]\times K,\mathbb{R}^d)$ defined in (47) solves (39). The uniqueness of solutions to this linearised problem can then be recovered under hypotheses **(H)** from Grönwall's Lemma. □

In the following theorem, we show that when the velocity perturbation $w:[0,T]\times\mathbb{R}^d\to\mathbb{R}^d$ belongs to the tangent cone to the admissible velocities of a continuity inclusion of the form (6) with a convexified right-hand side, there exist solutions of the non-convexified inclusion whose distance to the measure defined in (37) is infinitesimally small compared to $\epsilon>0$.

Definition 3.11 (Tangent cones to convex hulls of $C^0(\mathbb{R}^d,\mathbb{R}^d)$-valued multifunctions)**.** *Let $V:[0,T]\times\mathscr{P}_c(\mathbb{R}^d)\rightrightarrows C^0(\mathbb{R}^d,\mathbb{R}^d)$ be a set-valued map and $v\in\overline{\mathrm{co}}\,V(t,\mu)$ for some $t\in[0,T]$ and $\mu\in\mathscr{P}_c(\mathbb{R}^d)$, where $\overline{\mathrm{co}}\,V(t,\mu)$ is given in the sense of* (8)*. The* tangent cone *to $\overline{\mathrm{co}}\,V(t,\mu)$ at v is defined by*

$$T_{\overline{\mathrm{co}}\,V(t,\mu)}(v):=\Big\{w\in C^0(\mathbb{R}^d,\mathbb{R}^d)\ \textit{s.t.}\ w_{|K}\in T_{\overline{\mathrm{co}}\,V_K(t,\mu)}(v_{|K})\ \textit{for any}\ K:=B(0,R)\ \textit{with}\ R>0\Big\},$$

where $T_{\overline{\mathrm{co}}\,V_K(t,\mu)}(v_{|K})$ is taken in the Banach space $\big(C^0(K,\mathbb{R}^d),\|\cdot\|_{C^0(K,\mathbb{R}^d)}\big)$.

Theorem 3.12 (Variational linearisation)**.** *Let $\mu^0 \in \mathscr{P}(B(0,r))$ for some $r > 0$ and $V : [0,T] \times \mathscr{P}_c(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d, \mathbb{R}^d)$ be a set-valued map satisfying hypotheses* **(DI)** *with $p \in [1, +\infty)$. Moreover, let $(\mu(\cdot), \boldsymbol{v}(\cdot))$ be a trajectory-selection pair solution of the inclusion* (6)*, and suppose that for $\mathscr{L}^1$-almost every $t \in [0,T]$ the following representation*

$$\boldsymbol{v}(t) := v(t, \mu(t)) \in V(t, \mu(t)),$$

holds with $v : [0,T] \times \mathscr{P}_c(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ satisfying hypotheses **(H)** *and $v(t,\nu) \in V(t,\nu)$ for $\mathscr{L}^1$-almost every $t \in [0,T]$ and any $\nu \in \mathscr{P}_c(\mathbb{R}^d)$. Let $\mathcal{F}^0 \in C^0(\mathbb{R}^d, \mathbb{R}^d)$, $w : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ be a Carathéodory vector-field satisfying hypotheses* **(CE)** *along with the pointwise inclusions*

$$w(t) \in T_{\overline{\mathrm{co}}\, V(t,\mu(t))}\big(v(t,\mu(t))\big) \qquad \text{for } \mathscr{L}^1\text{-almost every } t \in [0,T], \tag{54}$$

and $\mathcal{F} : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ be the corresponding solution of the linearised Cauchy problem (39)*.*

Then for any $\epsilon \in [0,1]$, there exists a solution $\tilde{\mu}_\epsilon(\cdot)$ of the continuity inclusion

$$\begin{cases} \partial_t \tilde{\mu}_\epsilon(t) \in -\mathrm{div}\Big(V(t, \tilde{\mu}_\epsilon(t))\tilde{\mu}_\epsilon(t)\Big), \\ \tilde{\mu}_\epsilon(0) = (\mathrm{Id} + \epsilon \mathcal{F}^0)_\# \mu^0, \end{cases} \tag{55}$$

such that

$$\sup_{t \in [0,T]} W_p\Big(\tilde{\mu}_\epsilon(t), \mathcal{G}_\epsilon(t,\cdot)_\# \mu(t)\Big) = o(\epsilon), \tag{56}$$

where the family of maps $(\mathcal{G}_\epsilon(\cdot,\cdot)) \subset C^0([0,T] \times \mathbb{R}^d, \mathbb{R}^d)$ is given as in Proposition 3.10.

Proof. In the sequel, we shall interchangeably use the notations $v(t,\mu(t))$ and $w(t)$ to denote the vector-fields defined on $\mathbb{R}^d$ or their restrictions to some closed ball $K := B(0,R)$.

First, remark that $v : [0,T] \times \mathscr{P}_c(\mathbb{R}^d) \times \mathbb{R}^d \mapsto \mathbb{R}^d$ satisfies hypotheses **(CE)** as a direct consequence of **(H)**-(i). Hence, by Corollary 2.12, there exists $R_r > 0$ such that $\mathrm{supp}(\mu(t)) \subset K := B(0,R_r)$ for all times $t \in [0,T]$. Furthermore observe that since $w(\cdot)$ satisfies (54), it holds by Definition 2.6 that

$$\mathrm{dist}_{C^0(K,\mathbb{R}^d)}\Big(v(t,\mu(t)) + \epsilon w(t); \overline{\mathrm{co}}\, V_K(t,\mu(t))\Big) = o_t(\epsilon), \tag{57}$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and all $\epsilon \in [0,1]$, where $o_t(\epsilon)/\epsilon \to 0$ as $\epsilon \to 0^+$. Moreover, one also has

$$\mathrm{dist}_{C^0(K,\mathbb{R}^d)}\Big(v(t,\mu(t)) + \epsilon w(t); \overline{\mathrm{co}}\, V_K(t,\mu(t))\Big) \le \epsilon \left\| w(t) \right\|_{C^0(K,\mathbb{R}^d)}, \tag{58}$$

because $v(t,\mu(t)) \in V(t,\mu(t))$, which implies in particular that the remainder defined in (57) divided by $\epsilon > 0$ is integrably bounded. In addition, observe that the non-local velocity field

$$v_\epsilon : (t,\mu,x) \in [0,T] \times \mathscr{P}_c(\mathbb{R}^d) \times \mathbb{R}^d \mapsto v(t,\mu,x) + \epsilon w(t,x),$$

also satisfies hypotheses **(CE)** for every $\epsilon \in [0,1]$ by construction.

Denote by $\mu_\epsilon(\cdot) \in \mathrm{AC}([0,T], \mathscr{P}_c(\mathbb{R}^d))$ the unique solution of the perturbed Cauchy problem

$$\begin{cases} \partial_t \mu_\epsilon(t) + \mathrm{div}\Big(\big(v(t,\mu_\epsilon(t)) + \epsilon w(t)\big)\mu_\epsilon(t)\Big) = 0, \\ \mu_\epsilon(0) = (\mathrm{Id} + \epsilon \mathcal{F}^0)_\# \mu^0, \end{cases} \tag{59}$$

and let $K' := B(0,R'_r)$ be a closed ball such that $\mathrm{supp}(\mu_\epsilon(t)) \subset K'$ for any $(t,\epsilon) \in [0,T] \times [0,1]$. By Theorem 2.13 applied to $V(\cdot,\cdot) := \{v(\cdot,\cdot)\}$ together with Corollary 2.12, there exists $C > 0$ such that

$$\sup_{t \in [0,T]} W_p(\mu(t), \mu_\epsilon(t)) \le C\epsilon, \tag{60}$$

for every $\epsilon \in [0,1]$. Since $w(\cdot)$ satisfies (57) and (58), we can apply [7, Theorem 8.1.3] to recover the existence of a family $(w_\epsilon(\cdot))$ of $\mathscr{L}^1$-measurable maps such that

$$\begin{cases} \left\| w(t) - w_\epsilon(t) \right\|_{C^0(K',\mathbb{R}^d)} \underset{\epsilon \to 0^+}{\longrightarrow} 0 \quad \text{and} \quad \left\| w_\epsilon(t) \right\|_{C^0(K',\mathbb{R}^d)} \le 2 \left\| w(t) \right\|_{C^0(K',\mathbb{R}^d)} + 1, \\ v(t,\mu(t)) + \sqrt{\epsilon} w_\epsilon(t) \in \overline{\mathrm{co}}\, V_{K'}(t,\mu(t)), \end{cases} \tag{61}$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and all $\epsilon \in [0,1]$. By hypothesis **(DI)**-(iv) together with (60) and [7, Theorem 8.1.3], one can also find measurable selections $t \in [0,T] \mapsto \boldsymbol{v}_\epsilon(t) \in \overline{\mathrm{co}}\, V_{K'}(t,\mu_\epsilon(t))$ such that

$$\left\|v(t,\mu(t)) + \sqrt{\epsilon} w_\epsilon(t) - \boldsymbol{v}_\epsilon(t)\right\|_{C^0(K',\mathbb{R}^d)} \leq L_{K'}(t) W_p(\mu(t),\mu_\epsilon(t)) \leq L_{K'}(t) C\epsilon, \tag{62}$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and any $\epsilon \in [0,1]$. Moreover observing that the sets $\overline{\mathrm{co}}\, V_{K'}(t,\mu_\epsilon(t)) \subset C^0(K',\mathbb{R}^d)$ are convex, it holds

$$(1-\sqrt{\epsilon}) v(t,\mu_\epsilon(t)) + \sqrt{\epsilon}\boldsymbol{v}_\epsilon(t) \in \overline{\mathrm{co}}\, V_{K'}(t,\mu_\epsilon(t)),$$

which along with (62) further yields

$$(1-\sqrt{\epsilon}) v(t,\mu_\epsilon(t)) + \sqrt{\epsilon}\big(v(t,\mu(t)) + \sqrt{\epsilon} w_\epsilon(t)\big) \in \overline{\mathrm{co}}\, V_{K'}(t,\mu_\epsilon(t)) + L_{K'}(t) C\epsilon^{3/2}\, \mathbb{B}_{C^0(K',\mathbb{R}^d)}, \tag{63}$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and any $\epsilon \in [0,1]$. Recalling that $v : [0,T] \times \mathscr{P}_c(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ satisfies hypotheses **(CE)**, we again have by (60)

$$\|v(t,\mu(t)) - v(t,\mu_\epsilon(t))\|_{C^0(K',\mathbb{R}^d)} \;\leq\; L_{K'}(t) W_p(\mu(t),\mu_\epsilon(t)) \;\leq\; L_{K'}(t) C\epsilon,$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$. Upon plugging this last estimate into (63), we obtain

$$\begin{aligned} v(t,\mu_\epsilon(t)) + \epsilon w(t) &= (1-\sqrt{\epsilon}) v(t,\mu_\epsilon(t)) + \sqrt{\epsilon}\big(v(t,\mu_\epsilon(t)) + \sqrt{\epsilon} w_\epsilon(t)\big) + \epsilon\big(w(t) - w_\epsilon(t)\big) \\ &\in \overline{\mathrm{co}}\, V_{K'}(t,\mu_\epsilon(t)) + \Big(2L_{K'}(t) C\epsilon^{3/2} + \epsilon\, \|w(t) - w_\epsilon(t)\|_{C^0(K',\mathbb{R}^d)}\Big)\mathbb{B}_{C^0(K',\mathbb{R}^d)}. \end{aligned}$$

This together with the properties (61) of the family of maps $(w_\epsilon(\cdot))$ and the fact that $L_{K'}(\cdot) \in L^1([0,T],\mathbb{R}_+)$ imply that for every $\epsilon \in [0,1]$, the mismatch function

$$\eta_\epsilon : t \in [0,T] \mapsto \mathrm{dist}_{C^0(K',\mathbb{R}^d)}\Big(v(t,\mu_\epsilon(t)) + \epsilon w(t); \overline{\mathrm{co}}\, V_{K'}(t,\mu_\epsilon(t))\Big), \tag{64}$$

is integrable and satisfies $r(\epsilon) := \|\eta_\epsilon(\cdot)\|_1 = o(\epsilon)$ for all $\epsilon \in [0,1]$ by Lebesgue's dominated convergence theorem. Whence, we can apply Theorem 2.13 to obtain the existence of a curve $\tilde{\mu}_\epsilon(\cdot)$ solution of

$$\begin{cases} \partial_t \tilde{\mu}_\epsilon(t) \in -\mathrm{div}\Big(\overline{\mathrm{co}}\, V(t,\tilde{\mu}_\epsilon(t))\tilde{\mu}_\epsilon(t)\Big), \\ \tilde{\mu}_\epsilon(0) = (\mathrm{Id} + \epsilon\mathcal{F}^0)_\# \mu^0, \end{cases} \tag{65}$$

such that

$$\sup_{t\in[0,T]} W_p(\tilde{\mu}_\epsilon(t),\mu_\epsilon(t)) = o(\epsilon), \tag{66}$$

for any $\epsilon \in [0,1]$. Moreover by Theorem 2.14 applied with $\delta := r(\epsilon)$, we can choose the curve $\tilde{\mu}_\epsilon(\cdot)$ satisfying (66) as a solution of (55) instead of (65). Finally, observe that by Proposition 3.10 one can express $\mu_\epsilon(t)$ for all $t \in [0,T]$ as

$$\mu_\epsilon(t) = \mathcal{G}_\epsilon(t,\cdot)_\# \mu(t), \tag{67}$$

where the family of maps $(\mathcal{G}_\epsilon(\cdot,\cdot)) \subset C^0([0,T] \times K, \mathbb{R}^d)$ satisfies (38). This together with (66) yields

$$\sup_{t\in[0,T]} W_p(\tilde{\mu}_\epsilon(t), \mathcal{G}_\epsilon(t,\cdot)_\# \mu(t)) = o(\epsilon),$$

which concludes the proof of Theorem 3.12. □

4 Pontryagin Maximum Principle in Wasserstein spaces

In this section, we apply the differential-theoretic concepts studied in Section 3 to prove a *Pontryagin Maximum Principle* ("PMP" in the sequel) for optimal control problems in Wasserstein spaces with inequality final-point constraints. This result partially improves those of [10, 12] as the strategy is more concise and allows to discriminate between the normal and abnormal scenarios of the PMP.

To lighten the exposition, we will illustrate the full proof strategy on a *Mayer* problem in Section 4.1, and we will then prove the PMP for a *Bolza* problem in Section 4.2 by adapting a standard procedure which directly builds on the PMP for the Mayer problem.

4.1 The PMP for constrained Mayer problems in Wasserstein spaces

In this section, we focus our attention on the following constrained Mayer problem

$$(\mathcal{P}_{\mathrm{MC}}) \quad \begin{cases} \min\limits_{u(\cdot)\in\mathcal{U}} \Big[\varphi(\mu(T))\Big] \\ \quad \text{s.t.} \begin{cases} \partial_t\mu(t) + \mathrm{div}\Big(v(t,\mu(t),u(t))\mu(t)\Big) = 0, \\ \mu(0) = \mu^0, \\ \mu(T) \in \mathcal{Q}_T, \end{cases} \end{cases}$$

where $\varphi : \mathscr{P}_c(\mathbb{R}^d) \to \mathbb{R}$ is a given final cost. Here, we fix an initial datum $\mu^0 \in \mathscr{P}_c(\mathbb{R}^d)$ and a controlled non-local velocity field $v : [0,T] \times \mathscr{P}_c(\mathbb{R}^d) \times U \times \mathbb{R}^d \to \mathbb{R}^d$. The minimisation in $(\mathcal{P}_{\mathrm{MC}})$ is taken over the set of admissible open-loop controls

$$\mathcal{U} := \Big\{u : [0,T] \to U \text{ s.t. } u(\cdot) \text{ is } \mathscr{L}^1\text{-measurable}\Big\},$$

where (U, d_U) is a compact metric space, and we suppose that the set of *final-point constraints* $\mathcal{Q}_T$ is defined by a collection of functional inequalities of the form

$$\mathcal{Q}_T := \Big\{\mu \in \mathscr{P}_2(\mathbb{R}^d) \text{ s.t. } \Psi_i(\mu) \le 0 \text{ for all } i \in \{1,\dots,n\}\Big\},$$

where $\Psi_i : \mathscr{P}_c(\mathbb{R}^d) \to \mathbb{R}$ for every index $i \in \{1,\dots,n\}$. From now on, we consider $\mathscr{P}_c(\mathbb{R}^d)$ as a subset of the metric space $(\mathscr{P}_1(\mathbb{R}^d), W_1)$.

Hypotheses (MCP). *For every $R > 0$, assume that the following holds with $K := B(0,R)$.*

(i) For every $u \in U$, the non-local velocity-field $(t,\mu,x) \in [0,T] \times \mathscr{P}_c(\mathbb{R}^d) \times \mathbb{R}^d \mapsto v(t,\mu,u)(x) \in \mathbb{R}^d$ satisfies hypotheses **(H)** *with $p = 1$ and constants that are independent of $u \in U$. Moreover, the map $u \in U \mapsto v(t,\mu,u)(x) \in \mathbb{R}^d$ is continuous for $\mathscr{L}^1$-almost every $t \in [0,T]$ and any $(\mu,x) \in \mathscr{P}_c(\mathbb{R}^d) \times \mathbb{R}^d$.*

(ii) The final cost $\varphi(\cdot)$ and the constraint functionals $\{\Psi_i(\cdot)\}_{1\le i\le n}$ are Lipschitz continuous in the W_1-metric over $\mathscr{P}(K)$ and locally differentiable over $\mathscr{P}_c(\mathbb{R}^d)$. Moreover, the maps

$$x \in \mathbb{R}^d \mapsto \nabla\varphi(\mu)(x) \in \mathbb{R}^d \qquad \text{and} \qquad x \in \mathbb{R}^d \mapsto \nabla\Psi_i(\mu)(x) \in \mathbb{R}^d,$$

are continuous for every $i \in \{1,\dots,n\}$.

As illustrated in [11] and recalled in Theorem 2.15 above, the set of all trajectories $\mu(\cdot)$ satisfying

$$\partial_t\mu(t) + \mathrm{div}\big(v(t,\mu(t),u(t))\mu(t)\big) = 0 \quad \text{for some } u(\cdot) \in \mathcal{U}, \tag{68}$$

coincides exactly with the solution set of the continuity inclusion

$$\partial_t\mu(t) \in -\mathrm{div}\Big(V(t,\mu(t))\mu(t)\Big), \tag{69}$$

when the set-valued map $V : [0,T] \times \mathscr{P}_c(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d,\mathbb{R}^d)$ is defined by

$$V(t,\mu) := \Big\{\boldsymbol{v} \in C^0(\mathbb{R}^d,\mathbb{R}^d) \text{ s.t. } \boldsymbol{v}(\cdot) = v(t,\mu,u,\cdot) \text{ for some } u \in U\Big\}, \tag{70}$$

for all $(t,\mu) \in [0,T] \times \mathscr{P}_c(\mathbb{R}^d)$.

Definition 4.1 (Admissible pairs and strong local minimisers for $(\mathcal{P}_{\mathrm{MC}})$)**.** *We say that $(\mu(\cdot),u(\cdot))$ is an* admissible trajectory-control pair *for $(\mathcal{P}_{\mathrm{MC}})$ if $u(\cdot) \in \mathcal{U}$ and $\mu(\cdot) \in \mathrm{AC}([0,T],\mathscr{P}_c(\mathbb{R}^d))$ is a solution of the controlled non-local continuity equation (68) satisfying $\mu(0) = \mu^0$ and $\mu(T) \in \mathcal{Q}_T$. Moreover, an admissible pair $(\mu^*(\cdot),u^*(\cdot))$ is a* strong local minimiser *for $(\mathcal{P}_{\mathrm{MC}})$ if there exists $\epsilon > 0$ such that*

$$\varphi(\mu^*(T)) \le \varphi(\mu(T)),$$

for every other admissible pair $(\mu(\cdot),u(\cdot))$ which satisfies $\sup_{t\in[0,T]} W_1(\mu^*(t),\mu(t)) \le \epsilon$.

Remark 4.2 (Strong local W_p-minimisers)**.** *Observe that if $\mu, \nu \in \mathscr{P}(B(0,R))$ for some $R > 0$, it holds that $W_1(\mu,\nu) \leq W_p(\mu,\nu) \leq (2R)^{(p-1)/p} W_1(\mu,\nu)^{1/p}$ for every $p \in [1,+\infty)$. Hence, an admissible pair is a strong local W_p-minimiser if and only if it is a strong local W_1-minimiser.*

Remark 4.3 (On the existence of optimal trajectory-control pairs)**.** *In [11, Remark 4 and Theorem 7], an existence result is provided for a more general variant of problem ($\mathcal{P}_{\mathrm{MC}}$) in the presence of running and final-point constraints, under a set of hypotheses which are a natural adaptation of* **(MCP)***. The only additional requirement needed to obtain this existence result compared to necessary optimality conditions is the convexity of the sets of admissible velocities $V(t,\mu) \subset C^0(\mathbb{R}^d, \mathbb{R}^d)$, which is instrumental in showing that the solution set of the corresponding inclusion is compact in the topology of the uniform convergence (see [11, Theorem 6]). We would like to stress that in [11, Theorem 7], it is also assumed that $u \in U \mapsto v(t,\mu,u,x)$ is Lipschitz continuous. However, this hypothesis is only needed when the controls are closed-loop, and it follows from our proof therein that the usual continuity assumption is sufficient when the controls are open-loop, see [11, Remark 5].*

Let $r > 0$ be such that $\mu^0 \in \mathscr{P}(B(0,r))$, and observe that under hypotheses **(MCP)** the non-local velocity-fields $(t,\mu,x) \in [0,T] \times \mathscr{P}_c(\mathbb{R}^d) \times \mathbb{R}^d \mapsto v(t,\mu,u,x)$ satisfy hypotheses **(CE)** with constants which are uniform with respect to $u \in U$. Therefore by Theorem 2.11, there exists a radius $R_r > 0$ depending only on the magnitudes of $r, \|m(\cdot)\|_1$ such that every admissible pair $(\mu(\cdot), u(\cdot))$ verifies

$$\mathrm{supp}(\mu(t)) \subset K := B(0,R_r),$$

for all times $t \in [0,T]$. In the sequel, we will denote by $(\Phi^{u^*}_{(s,t)}(\cdot))_{s,t\in[0,T]}$ the semigroup of non-local flows which represent the optimal curve $\mu^*(\cdot)$ via (11) and (12).

We are now ready to state and prove our main result, which is a PMP for ($\mathcal{P}_{\mathrm{MC}}$). In what follows, we denote by $\mathbb{J}_{2d}$ the *symplectic matrix* of $\mathbb{R}^{2d}$, i.e.

$$\mathbb{J}_{2d} = \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0 \end{pmatrix},$$

and by $\mathbb{H} : [0,T] \times \mathscr{P}_c(\mathbb{R}^{2d}) \times U \to \mathbb{R}$ the *Hamiltonian* of the control problem, defined by

$$\mathbb{H}(t,\nu,u) = \int_{\mathbb{R}^{2d}} \langle r, v(t,\pi^1_\#\nu,u,x)\rangle \mathrm{d}\nu(x,r), \tag{71}$$

for all $(t,\nu,u) \in [0,T] \times \mathscr{P}_c(\mathbb{R}^{2d}) \times U$. We also consider the set of *active indices* at $\mu^*(T)$, given by

$$I^\circ(\mu^*(T)) := \Big\{ i \in \{1,\dots,n\} \text{ s.t. } \mu^*(T) \in \partial \mathcal{Q}^i_T \Big\},$$

where we introduced the sets $\mathcal{Q}^i_T := \{\mu \in \mathscr{P}_2(\mathbb{R}^d) \text{ s.t. } \Psi_i(\mu) \leq 0\}$ for every $i \in \{1,\dots,n\}$.

Theorem 4.4 (Pontryagin Maximum Principle for ($\mathcal{P}_{\mathrm{MC}}$))**.** *Let $(\mu^*(\cdot), u^*(\cdot)) \in \mathrm{AC}([0,T], \mathscr{P}_c(\mathbb{R}^d)) \times \mathcal{U}$ be a strong local minimiser for ($\mathcal{P}_{\mathrm{MC}}$), and suppose that hypotheses* **(MCP)** *hold.*

Then, there exist $R'_r > 0$, non-trivial Lagrange multipliers $(\lambda_0, \lambda_1, \dots, \lambda_n) \in \{0,1\} \times \mathbb{R}^n_+$ and a curve of measures $\nu^*(\cdot) \in \mathrm{AC}([0,T], \mathscr{P}(K' \times K'))$ with $K' := B(0,R'_r)$ such that the following holds.

(*i*) *The curve $\nu^*(\cdot)$ solves the* forward-backward Hamiltonian *continuity equation*

$$\begin{cases} \partial_t \nu^*(t) + \mathrm{div}\Big(\mathbb{J}_{2d} \nabla_\nu \mathbb{H}(t,\nu^*(t),u^*(t))\nu^*(t)\Big) = 0, \\ \pi^1_\#\nu^*(t) = \mu^*(t) \qquad\qquad \textit{for all times } t \in [0,T], \\ \nu^*(T) = \Big(\mathrm{Id}\,,\, \Big(-\lambda_0 \nabla\varphi(\mu^*(T)) - \displaystyle\sum_{i=1}^n \lambda_i \nabla\Psi_i(\mu^*(T))\Big)\Big)_\# \mu^*(T), \end{cases} \tag{72}$$

where the Wasserstein gradient of the Hamiltonian is given explicitly by

$$\begin{aligned} &\nabla_\nu \mathbb{H}(t,\nu^*(t),u^*(t))(x,r) \\ &\quad = \begin{pmatrix} \mathrm{D}_x v(t,\mu^*(t),u^*(t),x)^\top r + \int_{\mathbb{R}^{2d}} \mathrm{D}_\mu v(t,\mu^*(t),u^*(t),y)(x)^\top p \,\mathrm{d}\nu^*(t)(y,p) \\ \\ v(t,\mu^*(t),u^*(t),x) \end{pmatrix}, \end{aligned} \tag{73}$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and any $(x,r) \in \mathbb{R}^{2d}$.

(*ii*) *The* complementarity slackness *conditions*

$$\lambda_i \Psi_i(\mu^*(T)) = 0, \tag{74}$$

hold for every index $i \in \{1, \dots, n\}$.

(*iii*) *The* Pontryagin maximisation condition

$$\mathbb{H}(t, \nu^*(t), u^*(t)) = \max_{u \in U} \mathbb{H}(t, \nu^*(t), u), \tag{75}$$

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

Moreover, if there exists a map $t \in [0,T] \mapsto w(t) \in T_{\overline{\mathrm{co}}\, V(t,\mu^*(t))}\big(v(t, \mu^*(t), u^*(t)\big)$ *satisfying hypotheses* **(CE)** *such that the corresponding solution* $\mathcal{F} \in C^0([0,T] \times K, \mathbb{R}^d)$ *of* (39) *with* $\mathcal{F}^0 = 0$ *satisfies*

$$\langle \nabla \Psi_i(\mu^*(T)), \mathcal{F}\big(T, \Phi^{u^*}_{(T,0)}(\cdot)\big)\rangle_{L^2(\mu^*(T))} < 0,$$

for every $i \in I^\circ(\mu^*(T))$*, then the PMP is normal, i.e.* $\lambda_0 = 1$.

We split the proof of Theorem 4.4 into four steps. In Step 1, we introduce suitable variational linearisations of the controlled non-local continuity equation inspired by Theorem 3.12. In Step 2, we focus on the situation in which the end-points of the trajectories of the linearised system do not satisfy the constraint qualification condition, which leads to the *abnormal* PMP, i.e. the case $\lambda_0 = 0$. In Step 3, we analyse the converse scenario and prove that it corresponds to the *normal* PMP, i.e. the case $\lambda_0 = 1$. Finally in Step 4, we build a state-costate curve $\nu^*(\cdot)$ solution of the Hamiltonian flow (72), along which the maximum principle (75) holds.

Step 1: Variational linearisations along $(\mu^*(\cdot), u^*(\cdot))$**.** Observe that $v(t, \mu, u^*(t)) \in \overline{\mathrm{co}}\, V(t, \mu)$ for $\mathscr{L}^1$-almost every $t \in [0,T]$ and all $\mu \in \mathscr{P}_c(\mathbb{R}^d)$, where the set-valued map $V : [0,T] \times \mathscr{P}_c(\mathbb{R}^d) \rightrightarrows C^0(\mathbb{R}^d, \mathbb{R}^d)$ is given by (70), and satisfies hypotheses **(DI)** as a consequence of **(MCP)**-(*i*). Let now

$$t \in [0,T] \mapsto w(t) \in T_{\overline{\mathrm{co}}\, V(t,\mu^*(t))}\big(v(t, \mu^*(t), u^*(t))\big), \tag{76}$$

be any $\mathscr{L}^1$-measurable selection satisfying hypotheses **(CE)**, where $T_{\overline{\mathrm{co}}\, V(t,\mu^*(t))}\big(v(t, \mu^*(t), u^*(t)\big)$ is to be understood for $\mathscr{L}^1$-almost every $t \in [0,T]$ in the sense of Definition 3.11. Since the non-local velocity-field $(t, \mu, x) \mapsto v(t, \mu, u^*(t), x)$ satisfies hypotheses **(H)**, we can apply Theorem 3.12 to obtain the existence of a solution $\tilde{\mu}_\epsilon(\cdot)$ of (69) with $\tilde{\mu}_\epsilon(0) = \mu^0$, such that

$$\sup_{t \in [0,T]} W_1\Big(\tilde{\mu}_\epsilon(t), \big(\mathrm{Id} + \epsilon \mathcal{F}\big(t, \Phi^{u^*}_{(t,0)}(\cdot)\big) + o_t(\epsilon)\big)_{\#} \mu^*(t)\Big) = o(\epsilon), \tag{77}$$

for any $\epsilon > 0$ and all times $t \in [0,T]$, where $\sup_{t \in [0,T]} \|o_t(\epsilon)\|_{C^0(K,\mathbb{R}^d)} = o(\epsilon)$. Here, the map $\mathcal{F} \in C^0([0,T] \times K, \mathbb{R}^d)$ is the unique solution of (39) with $\mathcal{F}^0 = 0$ and $w(\cdot)$ satisfying (76) and **(CE)**. From now on, we will denote by $\mathcal{R}^L_T \subset C^0(\mathbb{R}^d, \mathbb{R}^d)$ the *reachable set at time* T of (39) with $\mathcal{F}^0 = 0$, where the non-local velocity-field

$$(t, x) \in [0,T] \times \mathbb{R}^d \mapsto v(t, \mu(t), \Phi_{(0,t)}[\mu^0](x)) \in \mathbb{R}^d,$$

is replaced by the controlled non-local vector-field

$$(t, x) \in [0,T] \times \mathbb{R}^d \mapsto v(t, \mu^*(t), u^*(t), \Phi^{u^*}_{(0,t)}(x)) \in \mathbb{R}^d,$$

namely

$$\mathcal{R}^L_T := \Big\{ \mathcal{G}_T \in C^0(\mathbb{R}^d, \mathbb{R}^d) \text{ s.t. } \mathcal{G}_T(\cdot) = \mathcal{F}\big(T, \Phi^{u^*}_{(T,0)}(\cdot)\big) \text{ where } \mathcal{F}(\cdot,\cdot) \text{ solves (39) with } \mathcal{F}^0 = 0 \\ \text{and } w(\cdot) \text{ satisfying (76) and hypotheses } \textbf{(CE)} \Big\}.$$

Up to relabelling the functionals $\{\Psi_i(\cdot)\}_{i=1}^n$, we can suppose without loss of generality that $I^\circ(\mu^*(T)) = \{1,\dots,k\}$ for some $k \leq n$, whenever it is non-empty. In addition, observe that if $\nabla\Psi_i(\mu^*(T)) = 0$ for some $i \in I^\circ(\mu^*(T))$, then the statements of the PMP are verified with $\lambda_i = 1$, $\lambda_j = 0$ for every $j \in \{0,\dots,n\}\backslash\{i\}$ and $\nu^*(t) := \mu^*(t) \times \delta_0$ for all times $t \in [0,T]$. In this case, one has

$$\mathbb{H}(t,\nu^*(t),u) \equiv 0, \quad \nabla_\nu \mathbb{H}(t,\nu^*(t),u^*(t))(x,r) = \begin{pmatrix} 0 \\ v(t,\mu^*(t),u^*(t),x) \end{pmatrix} \quad \text{and} \quad \nu^*(T) = \mu^*(T) \times \delta_0,$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$, all $u \in U$ and $\nu^*(t)$-almost every $(x,r) \in \mathbb{R}^d$, and it can then be checked that the curve $\nu^*(\cdot)$ defined above trivially satisfies the statements of Theorem 4.4. Similarly if $\nabla\varphi(\mu^*(T)) = 0$, we can set $\lambda_0 = 1$ and $\lambda_i = 0$ for all $i \in \{1,\dots,n\}$, so that the PMP is again satisfied with $\nu^*(t) = \mu^*(t) \times \delta_0$. Hence, there only remains to consider the case in which $\nabla\varphi(\mu^*(T)) \neq 0$ and $\nabla\Psi_i(\mu^*(T)) \neq 0$ for all $i \in I^\circ(\mu^*(T))$.

If $I^\circ(\mu^*(T)) \neq \emptyset$, we introduce the non-empty subset of $\mathbb{R}^k$ defined by

$$\mathcal{B}_T := \Big\{ \Big(\langle \nabla\Psi_1(\mu^*(T)), \mathcal{G}_T \rangle_{L^2(\mu^*(T))}, \dots, \langle \nabla\Psi_k(\mu^*(T)), \mathcal{G}_T \rangle_{L^2(\mu^*(T))} \Big) \text{ s.t. } \mathcal{G}_T \in \mathcal{R}_T^L \Big\}, \tag{78}$$

which allows to discriminate between the two scenarios corresponding to the abnormal and normal versions of the maximum principle.

Step 2: Separation without constraint qualification. In this case, we suppose that

$$\mathcal{B}_T \cap (\mathbb{R}_-^*)^k := \mathcal{B}_T \cap (-\infty,0)^k = \emptyset.$$

Notice that $\mathcal{B}_T$ is a convex subset of $\mathbb{R}^k$ since $\mathcal{R}_T^L \subset C^0(\mathbb{R}^d,\mathbb{R}^d)$ is convex. Because $(\mathbb{R}_-^*)^k$ is convex as well, there exists by the separation theorem a non-trivial element $p \in \mathbb{R}^k$ such that

$$\sup_{d \in (\mathbb{R}_-^*)^k} \langle p, d \rangle \leq \inf_{b \in \mathcal{B}_T} \langle p, b \rangle. \tag{79}$$

Remark now that since $(\mathbb{R}_-^*)^k$ is a cone, the separation inequality (79) necessarily implies

$$\sup_{d \in \mathbb{R}_-^k} \langle p, d \rangle \leq 0,$$

which yields that $p := (\lambda_1,\dots,\lambda_k) \in \mathbb{R}_+^k$. Similarly, one can check that $\mathcal{B}_T$ is a cone upon remarking that $\mathcal{R}_L^T \subset C^0(\mathbb{R}^d,\mathbb{R}^d)$ is itself a cone. Thus, (79) also implies

$$\inf_{b \in \mathcal{B}_T} \langle p, b \rangle \geq 0,$$

which combined with the definition (78) of $\mathcal{B}_T$ further yields

$$\sum_{i=1}^k \lambda_i \langle \nabla\Psi_i(\mu^*(T)), \mathcal{G}_T \rangle_{L^2(\mu^*(T))} \ \geq \ 0,$$

for every element $\mathcal{G}_T \in \mathcal{R}_T^L$. This inequality can be in turn rewritten as

$$\big\langle -P_T, \mathcal{F}\big(T, \Phi_{(T,0)}^{u^*}(\cdot)\big) \big\rangle_{L^2(\mu^*(T))} \leq 0, \tag{80}$$

for every solution $\mathcal{F} \in C^0([0,T] \times K, \mathbb{R}^d)$ of (39) with $\mathcal{F}^0 = 0$ and $w(\cdot)$ satisfying (76) and **(CE)**, where the covector $P_T \in C^0(K,\mathbb{R}^d)$ is defined by

$$P_T := \sum_{i=1}^k \lambda_i \nabla\Psi_i(\mu^*(T)). \tag{81}$$

Step 3: Separation with constraint qualification. We now investigate the scenario in which

$$\mathcal{B}_T \cap (\mathbb{R}^*_-)^k \neq \emptyset.$$

In this context, consider the set $\mathcal{A}_T \subset \mathbb{R}^{k+1}$ defined by

$$\mathcal{A}_T := \Big\{ \Big(\langle \nabla\varphi(\mu^*(T)), \mathcal{G}_T \rangle_{L^2(\mu^*(T))}, \langle \nabla\Psi_1(\mu^*(T)), \mathcal{G}_T \rangle_{L^2(\mu^*(T))}, \dots, \\ \langle \nabla\Psi_k(\mu^*(T)), \mathcal{G}_T \rangle_{L^2(\mu^*(T))} \Big) \text{ s.t. } \mathcal{G}_T \in \mathcal{R}^L_T \Big\}.$$

In the following lemma, we prove that the set $\mathcal{A}_T$ defined above is disjoint from $(\mathbb{R}^*_-)^{k+1}$.

Lemma 4.5 (Incompatible intersection)**.** *Let $(\mu^*(\cdot), u^*(\cdot))$ be a strong local minimiser for $(\mathcal{P}_{\mathrm{MC}})$. Then, it necessarily holds that*

$$\mathcal{A}_T \cap (\mathbb{R}^*_-)^{k+1} = \emptyset. \tag{82}$$

Proof. Suppose by contradiction that the intersection in (82) is non-empty. Then, there exists $w(\cdot)$ satisfying (76) and **(CE)** such that the solution $\mathcal{F} \in C^0([0,T] \times K, \mathbb{R}^d)$ of (39) with $\mathcal{F}^0 = 0$ verifies

$$\big\langle \nabla\varphi(\mu^*(T)), \mathcal{F}\big(T, \Phi^{u^*}_{(T,0)}(\cdot)\big) \big\rangle_{L^2(\mu^*(T))} < 0, \tag{83}$$

and

$$\big\langle \nabla\Psi_i(\mu^*(T)), \mathcal{F}\big(T, \Phi^{u^*}_{(T,0)}(\cdot)\big) \big\rangle_{L^2(\mu^*(T))} < 0, \tag{84}$$

for every $i \in \{1, \dots, k\}$. By Step 1, there exists a compact set $K' \subset \mathbb{R}^d$ such that for every $\epsilon > 0$ sufficiently small, there exists a curve $\tilde{\mu}_\epsilon(\cdot) \in \mathrm{AC}([0,T], \mathscr{P}(K'))$ solution of (69) which satisfies

$$\sup_{t\in[0,T]} W_1\Big(\tilde{\mu}_\epsilon(t), \big(\mathrm{Id} + \epsilon\mathcal{F}\big(t, \Phi^{u^*}_{(t,0)}(\cdot)\big) + o_t(\epsilon)\big)_{\#}\mu^*(t)\Big) = o(\epsilon),$$

where $\sup_{t\in[0,T]} \|o_t(\epsilon)\|_{C^0(K',\mathbb{R}^d)} = o(\epsilon)$. Since the solution set of (69) coincides with that of (68), there exists $\tilde{u}_\epsilon(\cdot) \in \mathcal{U}$ such that $(\tilde{\mu}_\epsilon(\cdot), \tilde{u}_\epsilon(\cdot))$ is a trajectory-control pair for (68). Moreover by hypotheses **(MCP)**-(ii), the maps $\varphi(\cdot)$ and $\{\Psi_i(\cdot)\}_{1\le i\le n}$ are locally differentiable in the sense of Definition 3.3. Therefore, it holds as a consequence of Corollary 3.7 together with (83) that

$$\begin{aligned} \varphi(\tilde{\mu}_\epsilon(T)) &= \varphi\Big(\big(\mathrm{Id} + \epsilon\mathcal{F}\big(T, \Phi^{u^*}_{(T,0)}(\cdot)\big) + o_T(\epsilon)\big)_{\#}\mu^*(T)\Big) + o(\epsilon) \\ &= \varphi(\mu^*(T)) + \epsilon\big\langle \nabla\varphi(\mu^*(T)), \mathcal{F}\big(T, \Phi^{u^*}_{(T,0)}(\cdot)\big) \big\rangle_{L^2(\mu^*(T))} + o_T(\epsilon) \\ &< \varphi(\mu^*(T)), \end{aligned} \tag{85}$$

and analogously, by (84),

$$\Psi_i(\tilde{\mu}_\epsilon(T)) < \Psi_i(\mu^*(T)), \tag{86}$$

for every index $i \in \{1, \dots, k\}$ and any $\epsilon > 0$ small enough, where we also used that fact that the maps $\varphi(\cdot)$ and $\{\Psi_i(\cdot)\}_{i=1}^n$ are Lipschitz continuous over $\mathscr{P}(K')$ in the W_1-metric. Combining (85) and (86) while observing that $\Psi_i(\tilde{\mu}_\epsilon(T)) < 0$ as well for $i \notin I^\circ(\mu^*(T))$ and $\epsilon > 0$ small enough, we have thus built an admissible pair $(\tilde{\mu}_\epsilon(\cdot), \tilde{u}_\epsilon(\cdot))$ for $(\mathcal{P}_{\mathrm{MC}})$ that produces a cost strictly smaller than $(\mu^*(\cdot), u^*(\cdot))$, and which satisfies $\sup_{t\in[0,T]} W_1(\mu^*(t), \tilde{\mu}_\epsilon(t)) \le C\epsilon$ for some constant $C > 0$ independent of $\epsilon > 0$. This contradicts our assumption that the pair $(\mu^*(\cdot), u^*(\cdot))$ is a strong local minimiser for $(\mathcal{P}_{\mathrm{MC}})$. ☐

As a consequence of Lemma 4.5, the intersection described in (82) is necessarily empty. Since both sets in the latter are convex, the separation theorem yields the existence of a non-trivial element $p := (\lambda_0, \dots, \lambda_k) \in \mathbb{R}^{k+1}$ such that

$$\sup_{d\in(\mathbb{R}^*_-)^{k+1}} \langle p, d \rangle \;\le\; \inf_{a\in\mathcal{A}_T} \langle p, a \rangle. \tag{87}$$

By repeating the same arguments as in Step 2 above, we again have that $(\lambda_0, \dots, \lambda_k) \in \mathbb{R}_+^{k+1}$. In this case however, it necessarily holds that $\lambda_0 \neq 0$. Indeed if $\lambda_0 = 0$, then (87) becomes equivalent to $\mathcal{B}_T \cap (\mathbb{R}_-^*)^k = \emptyset$ because $(\mathbb{R}_-^*)^k$ is open, which contradicts our standing assumption. Hence up to renormalising all the multipliers by $\lambda_0 > 0$, we recover

$$\langle \nabla\varphi(\mu^*(T)), \mathcal{G}_T \rangle_{L^2(\mu^*(T))} + \sum_{i=1}^{k} \lambda_i \langle \nabla\Psi_i(\mu^*(T)), \mathcal{G}_T \rangle_{L^2(\mu^*(T))} \geq 0,$$

for every $\mathcal{G}_T \in \mathcal{R}_T^L$. The latter expression can be equivalently rewritten as

$$\langle - P_T, \mathcal{F}(T, \Phi^{u^*}_{(T,0)}(\cdot)) \rangle_{L^2(\mu^*(T))} \leq 0,$$

where $P_T \in C^0(K, \mathbb{R}^d)$ is given in this context by

$$P_T := \nabla\varphi(\mu^*(T)) + \sum_{i=1}^{k} \lambda_i \nabla\Psi_i(\mu^*(T)). \tag{88}$$

Observe now that if $I^\circ(\mu^*(T)) = \emptyset$, we can repeat the same arguments to obtain

$$\langle \nabla\varphi(\mu^*(T)), \mathcal{G}_T \rangle_{L^2(\mu^*(T))} \geq 0,$$

for any $\mathcal{G}_T \in \mathcal{R}_T^L$, and set $P_T := \nabla\varphi(\mu^*(T))$.

Step 4: Proof of the PMP. Condensing the results of Step 2 and Step 3 and setting $\lambda_i = 0$ for all $i \in \{k+1, \dots, n\}$, we have built a covector $P_T \in C^0(K, \mathbb{R}^d)$ given explicitly by

$$P_T = \lambda_0 \nabla\varphi(\mu^*(T)) + \sum_{i=1}^{n} \lambda_i \nabla\Psi_i(\mu^*(T)), \tag{89}$$

for some $(\lambda_0, \lambda_1, \dots, \lambda_n) \in \{0,1\} \times \mathbb{R}_+^n$ not all equal to 0 and satisfying $\lambda_i \Psi_i(\mu^*(T)) = 0$ for every $i \in \{1, \dots, n\}$. Moreover, the covector P_T is such that the family of end-point inequalities

$$\int_{\mathbb{R}^d} \langle - P_T(x), \mathcal{F}(T, \Phi^{u^*}_{(T,0)}(x)) \rangle \mathrm{d}\mu^*(T)(x) \leq 0, \tag{90}$$

hold for any solution $\mathcal{F} \in C^0([0,T] \times K, \mathbb{R}^d)$ of (39) with $\mathcal{F}^0 = 0$ and $w(\cdot)$ satisfying (76) and **(CE)**. Our goal now is to build a curve $\nu^*(\cdot)$ solution of the *forward-backward* continuity equation

$$\begin{cases} \partial_t \nu^*(t) + \mathrm{div}\Big(\mathcal{V}(t, \nu^*(t))\nu^*(t)\Big) = 0, \\ \pi^1_\# \nu^*(t) = \mu^*(t) \quad \text{ for all times } t \in [0,T], \\ \nu^*(T) = (\mathrm{Id}, -P_T)_\# \mu^*(T), \end{cases} \tag{91}$$

where the non-local velocity field $\mathcal{V} : [0,T] \times \mathscr{P}_c(\mathbb{R}^{2d}) \times \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ is defined by

$$\mathcal{V}(t, \nu, x, r) := \begin{pmatrix} v(t, \pi^1_\# \nu, u^*(t), x) \\ \\ -\mathrm{D}_x v(t, \pi^1_\# \nu, u^*(t), x)^\top r - \displaystyle\int_{\mathbb{R}^{2d}} \mathrm{D}_\mu v(t, \pi^1_\# \nu, u^*(t), y)(x)^\top p \, \mathrm{d}\nu(y,p) \end{pmatrix}, \tag{92}$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and any $(\nu, x, r) \in \mathscr{P}_c(\mathbb{R}^{2d}) \times \mathbb{R}^{2d}$.

Observe that $\mathcal{V}(\cdot,\cdot,\cdot)$ does not satisfy hypotheses **(CE)**, so that we cannot directly apply the existence result of Corollary 2.12 to assert that (91) admits solutions. The following lemma, originally established in [12], provides an explicit disintegration construction of such solutions by exploiting the *cascaded* structure of the system (see also [10]).

Lemma 4.6 (Definition and existence of solutions to (91))**.** *Let $(\mu^*(\cdot), u^*(\cdot))$ be a strong local minimiser for $(\mathcal{P}_{\mathrm{MC}})$ and suppose that hypotheses* **(MCP)** *hold. For $\mu^*(T)$-almost every $x \in \mathbb{R}^d$, denote by $(\Psi^x_{(T,t)}(\cdot))_{t\in[0,T]}$ the family of* backward non-local flows*, solutions of*

$$\begin{cases} \partial_t \Psi^x_{(T,t)}(r) = -\,\mathrm{D}_x v\Big(t, \mu^*(t), u^*(t), \Phi^{u^*}_{(T,t)}(x)\Big)^\top \Psi^x_{(T,t)}(r) \\ \qquad\qquad - \displaystyle\int_{\mathbb{R}^{2d}} \Big(\mathrm{D}_\mu v\Big(t, \mu^*(t), u^*(t), \Phi^{u^*}_{(T,t)}(y)\Big)\big(\Phi^{u^*}_{(T,t)}(x)\big)^\top \Psi^y_{(T,t)}(p)\Big)\mathrm{d}((\mathrm{Id}, -P_T)_\# \mu^*(T))(y,p), \\ \Psi^x_{(T,T)}(r) = r, \end{cases}$$

and define the curve of measures $\sigma^*_x(\cdot) \in \mathrm{AC}([0,T], \mathscr{P}_c(\mathbb{R}^d))$ as

$$\sigma^*_x(t) := \Psi^x_{(T,t)}(\cdot)_\# \delta_{(-P_T(x))},$$

for all times $t \in [0,T]$. Then, the curve given by $\nu^* : t \in [0,T] \mapsto \big(\Phi^{u^*}_{(T,t)} \circ \pi^1, \pi^2\big)_\# \nu^*_T(t)$ with

$$\nu^*_T(t) := \int_{\mathbb{R}^d} \sigma^*_x(t)\mathrm{d}\mu^*(T)(x),$$

is a solution of (91). Moreover, there exist $R'_r > 0$ and $m'_r(\cdot) \in L^1([0,T], \mathbb{R}_+)$ such that

$$\mathrm{supp}(\nu^*(t)) \subset K' \times K' \qquad \textit{and} \qquad W_1(\nu^*(t), \nu^*(s)) \le \int_s^t m'_r(\tau)\mathrm{d}\tau,$$

for all times $0 \le s \le t \le T$, where $K' := B(0, R'_r)$.

Proof. See [10, Lemma 3.7] and [12, Lemma 6]. □

Applying to the Hamiltonian $\nu \in \mathscr{P}_c(\mathbb{R}^d) \mapsto \mathbb{H}(t,\nu,u) \in \mathbb{R}$ defined in (71) the results of Section 5, which provide analytical expressions for the gradients of smooth integral functionals, one can check

$$\mathcal{V}(t, \nu^*(t), x, r) = \mathbb{J}_{2d} \nabla_\nu \mathbb{H}(t, \nu^*(t), u^*(t))(x,r),$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and any $(x,r) \in \mathbb{R}^{2d}$. Since $\nu^*(T) = (\mathrm{Id}, -P_T)_\# \mu^*(T)$ by the construction detailed in Lemma 4.6 with P_T being given by (89), the curve $\nu^*(\cdot)$ is a solution of (72). In the next lemma, we state an auxiliary result which will in turn yield the maximisation condition.

Lemma 4.7 (Derivative of an auxiliary functional)**.** *Given $w(\cdot)$ satisfying (76) and* **(CE)***, define*

$$\mathcal{H} : t \in [0,T] \mapsto \int_{\mathbb{R}^{2d}} \langle r, \mathcal{F}(t, \Phi^{u^*}_{(t,0)}(x))\rangle \mathrm{d}\nu^*(t)(x,r), \tag{93}$$

where $\mathcal{F} \in C^0([0,T] \times K, \mathbb{R}^d)$ is the corresponding solution of (39) with $\mathcal{F}^0 = 0$. Then, $\mathcal{H}(\cdot) \in \mathrm{AC}([0,T], \mathbb{R})$ and its pointwise derivative is given explicitly by

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}(t) = \int_{\mathbb{R}^{2d}} \langle r, w(t,x)\rangle \mathrm{d}\nu^*(t)(x,r), \tag{94}$$

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

Proof. The proof of this result is a matter of fairly long computations which mimic those already performed in [10, Lemma 3.9] and [12, Lemma 7]. □

Observe that by the definition of $\nu^*(\cdot)$ given in Lemma 4.6 together with (90), one has

$$\int_{\mathbb{R}^{2d}} \langle r, \mathcal{F}(T, \Phi^{u^*}_{(T,0)}(x))\rangle \mathrm{d}\nu^*(T)(x,r) \le 0,$$

which can in turn be reformulated as

$$\mathcal{H}(T) \le 0, \tag{95}$$

with $\mathcal{H}(\cdot)$ being given as in (93). Remark now that since $\mathcal{H}(\cdot) \in \mathrm{AC}([0,T],\mathbb{R})$ and $\mathcal{F}(0) = \mathcal{F}^0 = 0$ by construction, it further holds by Lemma 4.7

$$\mathcal{H}(T) = \int_0^T \tfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}(t)\mathrm{d}t = \int_0^T \int_{\mathbb{R}^{2d}} \langle r, w(t,x)\rangle \mathrm{d}\nu^*(t)(x,r)\mathrm{d}t, \tag{96}$$

so that upon merging (95) and (96), we obtain

$$\int_0^T \int_{\mathbb{R}^{2d}} \langle r, w(t,x)\rangle \mathrm{d}\nu^*(t)(x,r)\mathrm{d}t \leq 0, \tag{97}$$

for any $w(\cdot)$ satisfying (76) and **(CE)**.

Let us consider for some given $t \in [0,T]$ and $m \geq 1$ the closed subset of controls

$$\tilde{U}_m(t) := \Big\{ u \in U \text{ s.t. } \mathbb{H}(t,\nu^*(t),u) \geq \mathbb{H}(t,\nu^*(t),u^*(t)) + \tfrac{1}{m} \Big\},$$

and suppose that the associated $\mathscr{L}^1$-measurable subset $\Omega_m \subset [0,T]$, defined by

$$\Omega_m := \{ t \in [0,T] \text{ s.t. } \tilde{U}_m(t) \neq \emptyset \},$$

has positive measure for some $m \geq 1$. Then by Theorem [7, Theorem 8.2.9], we can find a measurable selection $t \in [0,T] \mapsto \tilde{u}(t) \in U$ such that $\tilde{u}(t) \in \tilde{U}_m(t)$ for $\mathscr{L}^1$-almost every $t \in \Omega_m$ and $\tilde{u}(t) = u^*(t)$ otherwise. Observe next that the map defined by

$$\tilde{w}(t) := v(t,\mu^*(t),\tilde{u}(t)) - v(t,\mu^*(t),u^*(t)),$$

for all times $t \in [0,T]$ satisfies (76) and hypotheses **(CE)**. Moreover, it is such that

$$\int_0^T \int_{\mathbb{R}^{2d}} \langle r, \tilde{w}(t,x)\rangle \mathrm{d}\nu^*(t)(x,r)\mathrm{d}t = \int_{\Omega_m} \Big(\mathbb{H}(t,\nu^*(t),\tilde{u}(t)) - \mathbb{H}(t,\nu^*(t),u^*(t)) \Big)\mathrm{d}t \geq \tfrac{1}{m}\mathscr{L}^1(\Omega_m),$$

which violates (97). Thus, the set $\Omega := \cup_{m\geq 1}\Omega_m \subset [0,T]$ necessarily has zero measure, which by definition of the sets $\tilde{U}_m(t)$ for $\mathscr{L}^1$-almost every $t \in [0,T]$ yields the maximisation condition (75).

Remark 4.8 (On the choice of performing separations on $\mathcal{A}_T$ and $\mathcal{B}_T$)**.** *Usually, geometric proofs of the PMP and its second-order variants tend to involve separation arguments on sets of trajectories which are not transposable to the setting of Wasserstein spaces. Thus in Step 2 and Step 3, we chose to perform separation arguments directly on $\mathcal{A}_T$ and $\mathcal{B}_T$, which are subsets of finite-dimensional euclidean spaces defined as the images of the reachable set $\mathcal{R}_T^L$ under the action of the gradients of the cost and constraint functionals. As illustrated before, this choice allows for a very simple and concise discrimination between the abnormal and normal scenarios of the PMP.*

4.2 Adaptation of the PMP to constrained Bolza problems

In this section, we make use of the results of Section 4.1 to obtain an extension of Theorem 4.4 to the setting of general constrained Bolza problems of the form

$$(\mathcal{P}_{\mathrm{BC}}) \begin{cases} \min\limits_{u(\cdot)\in\mathcal{U}} \left[\displaystyle\int_0^T L(t,\mu(t),u(t))\mathrm{d}t + \varphi(\mu(T)) \right] \\ \text{s.t.} \begin{cases} \partial_t\mu(t) + \mathrm{div}\Big(v(t,\mu(t),u(t))\mu(t)\Big) = 0, \\ \mu(0) = \mu^0, \\ \mu(T) \in \mathcal{Q}_T, \end{cases} \end{cases}$$

where $L : [0,T] \times \mathscr{P}_c(\mathbb{R}^d) \times U \mapsto \mathbb{R}$ is a given running cost which satisfies the following.

Hypotheses (L). *For every $R > 0$, assume that the following holds with $K := B(0,R)$.*

(i) *The map* $t \in [0,T] \mapsto L(t,\mu,u)$ *is* $\mathscr{L}^1$*-measurable for any* $(\mu,u) \in \mathscr{P}_c(\mathbb{R}^d) \times U$*. Moreover, there exists* $k(\cdot) \in L^1([0,T],\mathbb{R}_+)$ *such that* $\sup_{u\in U} |L(t,\delta_0,u)| \leq k(t)$ *for* $\mathscr{L}^1$*-almost every* $t \in [0,T]$*.*

(ii) *The map* $u \in U \mapsto L(t,\mu,u)$ *is continuous for* $\mathscr{L}^1$*-almost every* $t \in [0,T]$ *and any* $\mu \in \mathscr{P}_c(\mathbb{R}^d)$*.*

(iii) *The map* $\mu \in \mathscr{P}_c(\mathbb{R}^d) \mapsto L(t,\mu,u)$ *is locally differentiable in the sense of Definition 3.3 and the application* $x \in \mathbb{R}^d \mapsto \nabla_\mu L(t,\mu,u)(x) \in \mathbb{R}^d$ *is continuous for* $\mathscr{L}^1$*-almost every* $t \in [0,T]$ *and any* $(\mu,u) \in \mathscr{P}_c(\mathbb{R}^d) \times U$*. Moreover, there exists a map* $\mathcal{L}_K(\cdot) \in L^1([0,T],\mathbb{R}_+)$ *such that*

$$|L(t,\mu,u) - L(t,\nu,u)| \leq \mathcal{L}_K(t) W_1(\mu,\nu),$$

for every $\mu,\nu \in \mathscr{P}(K)$ *and any* $u \in U$*.*

In the present context, we say that an admissible pair $(\mu^*(\cdot),u^*(\cdot)) \in \mathrm{AC}([0,T],\mathscr{P}_c(\mathbb{R}^d)) \times \mathcal{U}$ is a strong local minimiser for $(\mathcal{P}_{\mathrm{BC}})$ if there exists $\epsilon > 0$ such that

$$\int_0^T L(t,\mu^*(t),u^*(t))\mathrm{d}t + \varphi(\mu^*(T)) \leq \int_0^T L(t,\mu(t),u(t))\mathrm{d}t + \varphi(\mu(T)),$$

for every admissible pair $(\mu(\cdot),u(\cdot))$ which satisfies $\sup_{t\in[0,T]} W_1(\mu^*(t),\mu(t)) \leq \epsilon$.

Theorem 4.9 (Pontryagin Maximum Principe for $(\mathcal{P}_{\mathrm{BC}})$)**.** *Let* $(\mu^*(\cdot),u^*(\cdot)) \in \mathrm{AC}([0,T],\mathscr{P}_c(\mathbb{R}^d)) \times \mathcal{U}$ *be a strong local minimiser for* $(\mathcal{P}_{\mathrm{BC}})$*, and suppose that hypotheses* **(MCP)** *and* **(L)** *hold.*

Then, the conclusions of Theorem 4.4 hold with the Hamiltonian $\mathbb{H}_{\lambda_0} : [0,T] \times \mathscr{P}_c(\mathbb{R}^{2d}) \times U \to \mathbb{R}$ *associated to* $(\mathcal{P}_{\mathrm{BC}})$*, defined by*

$$\mathbb{H}_{\lambda_0}(t,\nu,u) := \int_{\mathbb{R}^{2d}} \langle r, v(t,\pi^1_\#\nu,u,x)\rangle \mathrm{d}\nu(x,r) - \lambda_0 L(t,\pi^1_\#\nu,u), \tag{98}$$

for all $(t,\nu,u) \in [0,T] \times \mathscr{P}_c(\mathbb{R}^{2d}) \times U$*, which Wasserstein gradient writes*

$$\begin{aligned} &\nabla_\nu \mathbb{H}_{\lambda_0}(t,\nu^*(t),u^*(t))(x,r) \\ &= \begin{pmatrix} \mathrm{D}_x v(t,\mu^*(t),u^*(t),x)^\top r + \int_{\mathbb{R}^{2d}} \mathrm{D}_\mu v(t,\mu^*(t),u^*(t),y)(x)^\top p \,\mathrm{d}\nu^*(t)(y,p) \\ -\lambda_0 \nabla_\mu L(t,\mu^*(t),u^*(t))(x) \\ \\ v(t,\mu^*(t),u^*(t),x) \end{pmatrix}, \end{aligned} \tag{99}$$

for $\mathscr{L}^1$*-almost every* $t \in [0,T]$ *and any* $(x,r) \in \mathbb{R}^{2d}$*.*

In what follows, we show how the statements of Theorem 4.9 can be recovered by directly applying Theorem 4.4 to a suitable augmented Mayer problem. We shall henceforth denote by $\boldsymbol{\pi}^1 : \mathbb{R}^{d+1} \times \mathbb{R}^{d+1} \to \mathbb{R}^{d+1}$ and $\boldsymbol{\pi}^2 : \mathbb{R}^{d+1} \times \mathbb{R}^{d+1} \to \mathbb{R}^{d+1}$ the projection operators onto the first and second factor.

Given an admissible pair $(\mu(\cdot),u(\cdot))$ for $(\mathcal{P}_{\mathrm{BC}})$, let $z(\cdot) \in \mathrm{AC}([0,T],\mathbb{R})$ be the curve defined as

$$z(t) := \int_0^t L(s,\mu(s),u(s))\mathrm{d}s, \tag{100}$$

and consider the corresponding *augmented state*

$$\boldsymbol{\mu}(t) := \mu(t) \times \delta_{z(t)},$$

associated to $(\mathcal{P}_{\mathrm{BC}})$, both defined for all times $t \in [0,T]$. Below given a measure $\boldsymbol{\mu} \in \mathscr{P}_c(\mathbb{R}^{d+1})$, we will set $\mu := (\pi^1,\dots,\pi^d)_\#\boldsymbol{\mu} \in \mathscr{P}_c(\mathbb{R}^d)$. It is then clear that if $(\mu^*(\cdot),u^*(\cdot))$ is a strong local minimiser for $(\mathcal{P}_{\mathrm{BC}})$, then choosing $z^*(\cdot)$ as in (100) with $(\mu(\cdot),u(\cdot)) := (\mu^*(\cdot),u^*(\cdot))$ and defining $\boldsymbol{\mu}^*(t) := \mu^*(t) \times \delta_{z^*(t)}$, the pair $(\boldsymbol{\mu}^*(\cdot),u^*(\cdot))$ is a strong local minimiser for the Mayer problem

$$(\hat{\mathcal{P}}_{\mathrm{MC}}) \quad \begin{cases} \min\limits_{u(\cdot)\in\mathcal{U}} \left[\hat{\varphi}(\boldsymbol{\mu}(T))\right], \\ \text{s.t.} \quad \begin{cases} \partial_t \boldsymbol{\mu}(t) + \mathrm{div}\Big(\hat{v}(t,\boldsymbol{\mu}(t),u(t))\boldsymbol{\mu}(t)\Big) = 0, \\ \boldsymbol{\mu}(0) = \mu^0 \times \delta_0, \\ \boldsymbol{\mu}(T) \in \hat{\mathcal{Q}}_T, \end{cases} \end{cases}$$

where we defined

$$\hat{\varphi}(\boldsymbol{\mu}) := \varphi(\mu) + \int_{\mathbb{R}^{d+1}} z \, \mathrm{d}\boldsymbol{\mu}(x,z) \quad \text{and} \quad \hat{v}(t,\boldsymbol{\mu},u,x,z) := \begin{pmatrix} v(t,\mu,u,x) \\ L(t,\mu,u) \end{pmatrix}, \tag{101}$$

for all $(t,\boldsymbol{\mu},u,x,z) \in [0,T] \times \mathscr{P}_c(\mathbb{R}^{d+1}) \times U \times \mathbb{R}^{d+1}$, along with

$$\hat{\mathcal{Q}}_T := \Big\{ \boldsymbol{\mu} \in \mathscr{P}_2(\mathbb{R}^{d+1}) \text{ s.t. } \hat{\Psi}_i(\boldsymbol{\mu}) \leq 0 \ \text{ for every } i \in \{1,\dots,n\} \Big\},$$

where $\hat{\Psi}_i(\boldsymbol{\mu}) := \Psi_i(\mu)$. By applying the results of Section 5, the first-order derivatives of the augmented dynamics, final cost and constraint maps can be written explicitly as

$$\mathrm{D}_{(x,z)}\hat{v}(t,\boldsymbol{\mu},u,x,z) = \begin{pmatrix} \mathrm{D}_x v(t,\mu,u,x) \\ 0 \end{pmatrix}, \qquad \mathrm{D}_{\boldsymbol{\mu}}\hat{v}(t,\boldsymbol{\mu},u,x,z) = \begin{pmatrix} \mathrm{D}_\mu v(t,\mu,u,x) & 0 \\ \nabla_\mu L(t,\mu,u) & 0 \end{pmatrix}, \tag{102}$$

and

$$\nabla\hat{\varphi}(\boldsymbol{\mu}) = \begin{pmatrix} \nabla\varphi(\mu) \\ 1 \end{pmatrix}, \qquad \nabla\hat{\Psi}_i(\boldsymbol{\mu}) = \begin{pmatrix} \nabla\Psi_i(\mu) \\ 0 \end{pmatrix}, \tag{103}$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$, all $(\boldsymbol{\mu},u,x,z) \in \mathscr{P}_c(\mathbb{R}^{d+1}) \times U \times \mathbb{R}^{d+1}$ and any index $i \in \{1,\dots,n\}$.

By Corollary 2.12, there exists $R_r > 0$ such that $\mathrm{supp}(\mu^*(t)) \subset K := B(0,R_r)$ for all times $t \in [0,T]$ where $r > 0$ is such that $\mu^0 \in \mathscr{P}(B(0,r))$. Hence, observe that under hypotheses **(L)**-(iii) and up to redefining the running cost for $\mathscr{L}^1$-almost every $t \in [0,T]$ and any $u \in U$ as

$$L(t,\mu,u) := L(t,(\pi_K)_\#\mu,u),$$

whenever $\mathrm{supp}(\mu) \not\subseteq K$, the extended velocity field $\hat{v} : [0,T] \times \mathscr{P}(\mathbb{R}^{d+1}) \times U \times \mathbb{R}^{d+1} \mapsto \mathbb{R}^{d+1}$ satisfies hypotheses **(H)** with $p = 1$. This together with the definitions of $\hat{\varphi}(\cdot)$ and $\{\hat{\Psi}_i(\cdot)\}_{1\leq i\leq n}$ implies that hypotheses **(MCP)** hold for $(\hat{\mathcal{P}}_{\mathrm{MC}})$. Thus by Theorem 4.4, there exist non-trivial multipliers $(\lambda_0,\dots,\lambda_n) \in \{0,1\} \times \mathbb{R}^n_+$ and a curve of measures $\boldsymbol{\nu}^*(\cdot) \in \mathrm{AC}([0,T],\mathscr{P}_c((\mathbb{R}^{d+1})^2)$ solution of the forward-backward Hamiltonian continuity equation

$$\begin{cases} \partial_t\boldsymbol{\nu}^*(t) + \mathrm{div}\Big(\mathbb{J}_{2(d+1)}\nabla_{\boldsymbol{\nu}}\hat{\mathbb{H}}(t,\boldsymbol{\nu}^*(t),u^*(t))\boldsymbol{\nu}^*(t)\Big) = 0, \\ \pi^1_\#\boldsymbol{\nu}^*(t) = \mu^*(t) \times \delta_{z^*(t)} \qquad \text{for all times } t \in [0,T], \\ \boldsymbol{\nu}^*(T) = \Big(\mathrm{Id}\,,\, -\lambda_0\nabla\hat{\varphi}(\boldsymbol{\mu}^*(T)) - \displaystyle\sum_{i=1}^n \lambda_i\nabla\hat{\Psi}_i(\boldsymbol{\mu}^*(T))\Big)_\#\boldsymbol{\mu}^*(T), \end{cases} \tag{104}$$

such that the complementarity slackness condition

$$\lambda_i\hat{\Psi}_i(\boldsymbol{\mu}^*(T)) = 0, \tag{105}$$

holds for every index $i \in \{1,\dots,n\}$, and the maximisation condition

$$\hat{\mathbb{H}}(t,\boldsymbol{\nu}^*(t),u^*(t)) = \max_{u\in U} \hat{\mathbb{H}}(t,\boldsymbol{\nu}^*(t),u). \tag{106}$$

is satisfied for $\mathscr{L}^1$-almost every $t \in [0,T]$. Here, the *augmented Hamiltonian* $\hat{\mathbb{H}} : [0,T] \times \mathscr{P}_c((\mathbb{R}^{d+1})^2) \times U \to \mathbb{R}$ associated to $(\hat{\mathcal{P}}_{\mathrm{MC}})$ is defined by

$$\hat{\mathbb{H}}(t,\boldsymbol{\nu},u) := \int_{\mathbb{R}^{2d+2}} \langle (r,q), \hat{v}(t,\boldsymbol{\pi}^1_\#\boldsymbol{\nu},u,x,z)\rangle \mathrm{d}\boldsymbol{\nu}(x,z,r,q),$$

for all $(t,\boldsymbol{\nu},u) \in [0,T] \times \mathscr{P}_c((\mathbb{R}^{d+1})^2) \times U$.

Observe now that by (101), the augmented Hamiltonian can be further expressed as

$$\begin{aligned} \hat{\mathbb{H}}(t,\boldsymbol{\nu}^*(t),u^*(t)) &= \int_{\mathbb{R}^{2d+2}} \Big\langle (r,q), \hat{v}(t,\boldsymbol{\pi}^1_\#\boldsymbol{\nu}^*(t),u^*(t),x,z)\Big\rangle \mathrm{d}\boldsymbol{\nu}^*(t)(x,r,z,q) \\ &= \int_{\mathbb{R}^{2d+2}} \Big(\langle r, v(t,\mu^*(t),u^*(t),x)\rangle + qL(t,\mu^*(t),u^*(t))\Big)\mathrm{d}\boldsymbol{\nu}^*(t)(x,z,r,q) \end{aligned} \tag{107}$$

which along with (102) allows us to derive the following analytical expression for its gradient

$$\nabla_{\boldsymbol{\nu}}\hat{\mathbb{H}}(t,\boldsymbol{\nu}^*(t),u^*(t))(x,z,r,q) = \begin{pmatrix} \mathrm{D}_{(x,z)}\hat{v}(t,\boldsymbol{\mu}^*(t),u^*(t),x,z)^\top(r,q) \\ +\int_{\mathbb{R}^{2d+2}} \mathrm{D}_{\boldsymbol{\mu}}\hat{v}(t,\boldsymbol{\mu}^*(t),u^*(t),x',z')(x,z)^\top(r',q')\,\mathrm{d}\boldsymbol{\nu}^*(t)(x',z',r',q') \\ \\ \hat{v}(t,\boldsymbol{\mu}^*(t),u^*(t),x) \end{pmatrix}$$
$$= \begin{pmatrix} \mathrm{D}_x v(t,\mu^*(t),u^*(t),x)^\top r + q\nabla_\mu L(t,\mu^*(t),u^*(t))(x) \\ +\int_{\mathbb{R}^{2d}} \mathrm{D}_\mu v(t,\mu^*(t),u^*(t),y)(x)^\top p\,\mathrm{d}\nu^*(t)(y,p) \\ 0 \\ \\ v(t,\mu^*(t),u^*(t),x) \\ L(t,\mu^*(t),u^*(t)) \end{pmatrix}. \tag{108}$$

It can now be verified straightforwardly as a consequence of (103) that

$$\pi^{2d+2}_{\#}\boldsymbol{\nu}^*(T) = \delta_{(-\lambda_0)},$$

and also that the right-hand side of the dynamics driving this extra covector variable is identically equal to 0 by plugging (108) into (104). Hence up to a permutation of the coordinates, it is possible to express the extended state-costate curve $\boldsymbol{\nu}^*(\cdot)$ for all times $t \in [0,T]$ as

$$\boldsymbol{\nu}^*(t) = \nu^*(t) \times \delta_{(z^*(t),-\lambda_0)}, \tag{109}$$

where $\nu^*(\cdot) \in \mathrm{AC}([0,T],\mathscr{P}_c(\mathbb{R}^{2d}))$. Plugging this last expression into (107) and (108) then yields

$$\hat{\mathbb{H}}(t,\boldsymbol{\nu}^*(t),u^*(t)) = \mathbb{H}_{\lambda_0}(t,\nu^*(t),u^*(t)), \tag{110}$$

and by repeating the same coordinate permutation, we recover the following expression for the gradient

$$\nabla_{\boldsymbol{\nu}}\hat{\mathbb{H}}(t,\boldsymbol{\nu}^*(t),u^*(t))(x,z,r,q) = \begin{pmatrix} \mathrm{D}_x v(t,\mu^*(t),u^*(t),x)^\top r - \lambda_0\nabla_\mu L(t,\mu^*(t),u^*(t))(x) \\ +\int_{\mathbb{R}^{2d}} \mathrm{D}_\mu v(t,\mu^*(t),u^*(t),y)(x)^\top p\,\mathrm{d}\nu^*(t)(y,p) \\ v(t,\mu^*(t),u^*(t),x) \\ \\ L(t,\mu^*(t),u^*(t)) \\ 0 \end{pmatrix},$$

for $\mathscr{L}^1$-almost every $t \in [0,T]$ and any $(x,z,r,q) \in \mathrm{supp}(\boldsymbol{\nu}^*(t))$. By applying again the differentiation results of Proposition 5.1, it can finally be checked that

$$\nabla_\nu \mathbb{H}_{\lambda_0}(t,\nu^*(t),u^*(t))(x,r) = \begin{pmatrix} \mathrm{D}_x v(t,\mu^*(t),u^*(t),x)^\top r - \lambda_0\nabla_\mu L(t,\mu^*(t),u^*(t))(x) \\ +\int_{\mathbb{R}^{2d}} \mathrm{D}_\mu v(t,\mu^*(t),u^*(t),y)(x)^\top p\,\mathrm{d}\nu^*(t)(y,p) \\ v(t,\mu^*(t),u^*(t),x) \end{pmatrix},$$

and the gradient of the augmented Hamiltonian can be rewritten as

$$\nabla_{\boldsymbol{\nu}}\hat{\mathbb{H}}(t,\boldsymbol{\nu}^*(t),u^*(t))(x,z,r,q) = \begin{pmatrix} \nabla_\nu \mathbb{H}_{\lambda_0}(t,\nu^*(t),u^*(t))(x,r) \\ L(t,\mu^*(t),u^*(t)) \\ 0 \end{pmatrix}. \tag{111}$$

Thus because $\boldsymbol{\nu}^*(\cdot)$ solves (104) and owing to the decomposition (109) combined with the identity (111), it follows that $\nu^*(\cdot)$ is a solution of (72) driven by gradient of the Hamiltonian defined in (98). Moreover, the complementarity slackness and maximisation condition can be directly recovered by using the definition of $\hat{\Psi}_i(\boldsymbol{\mu}^*(T))$ in (105) for the former and plugging (110) into (106) for the latter.

5 Examples of locally differentiable functionals

In this auxiliary section, we provide two examples of extended real-valued functionals defined over $\mathscr{P}_2(\mathbb{R}^d)$ which are locally differentiable in the sense of Definition 3.3, and compute their gradients.

Proposition 5.1 (Locally differentiable integral functionals)**.** *Let $V : \mathbb{R}^d \to \mathbb{R}$ be a continuously differentiable mapping. Then the functional*

$$\mathcal{V} : \mu \in \mathscr{P}_c(\mathbb{R}^d) \mapsto \int_{\mathbb{R}^d} V(x)\mathrm{d}\mu(x) \in \mathbb{R}, \tag{112}$$

is locally differentiable in the sense of Definition 3.3, and its Wasserstein gradient is given by

$$\nabla\mathcal{V}(\mu)(x) = \nabla V(x),$$

for every $\mu \in \mathscr{P}_c(\mathbb{R}^d)$ and any $x \in \mathrm{supp}(\mu)$.

Similarly, let $L : \mathbb{R}^d \times \mathscr{P}_c(\mathbb{R}^d) \to \mathbb{R}$ be such that $x \in \mathbb{R}^d \to L(x,\mu)$ is continuously differentiable for every $\mu \in \mathscr{P}_c(\mathbb{R}^d)$ and $\mu \in \mathscr{P}_c(\mathbb{R}^d) \mapsto \nabla_x L(x,\mu)$ is continuous for every $x \in \mathbb{R}^d$. Moreover, suppose that $\mu \in \mathscr{P}_c(\mathbb{R}^d) \to L(x,\mu)$ is locally differentiable for every $x \in \mathbb{R}^d$ and that $(x,y) \in \mathbb{R}^d \times \mathrm{supp}(\mu) \mapsto \nabla_\mu L(x,\mu)(y) \in \mathbb{R}^d$ is continuous for every $\mu \in \mathscr{P}_c(\mathbb{R}^d)$. Then, the map

$$\mathcal{L} : \mu \in \mathscr{P}_c(\mathbb{R}^d) \mapsto \int_{\mathbb{R}^d} L(x,\mu)\mathrm{d}\mu(x), \tag{113}$$

is locally differentiable, and its Wasserstein gradient writes

$$\nabla\mathcal{L}(\mu)(x) = \nabla_x L(x,\mu) + \int_{\mathbb{R}^d} \nabla_\mu L(z,\mu)(x)\mathrm{d}\mu(z),$$

for any $\mu \in \mathscr{P}_c(\mathbb{R}^d)$ and $x \in \mathrm{supp}(\mu)$.

Proof. We first prove that the functional $\mathcal{V}(\cdot)$ defined in (112) is locally differentiable. Let $\mu \in \mathscr{P}_c(\mathbb{R}^d)$, $R > 0$ and $\nu \in \mathscr{P}(B_\mu(R))$. For every $\gamma \in \Gamma_o(\mu,\nu)$, it holds

$$\begin{aligned}
\mathcal{V}(\nu) - \mathcal{V}(\mu) &= \int_{\mathbb{R}^{2d}} \Big(V(y) - V(x)\Big)\mathrm{d}\gamma(x,y) \\
&= \int_{\mathbb{R}^{2d}} \int_0^1 \langle \nabla V\big(x + s(y-x)\big), y - x\rangle \mathrm{d}s\mathrm{d}\gamma(x,y), \\
&= \int_{\mathbb{R}^{2d}} \langle \nabla V(x), y - x\rangle \mathrm{d}\gamma(x,y) + \int_0^1 \int_{\mathbb{R}^{2d}} \langle \nabla V\big(x + s(y-x)\big) - \nabla V(x), y - x\rangle \mathrm{d}\gamma(x,y)\mathrm{d}s, \\
&= \int_{\mathbb{R}^{2d}} \langle \nabla V(x), y - x\rangle \mathrm{d}\gamma(x,y) + I_R,
\end{aligned} \tag{114}$$

where we used Fubini's theorem. Our goal is now to prove that $I_R = o_R(W_2(\mu,\nu))$. Since we assumed that $\nabla V(\cdot)$ is continuous, we can find for every $\epsilon > 0$ another real $\eta > 0$ such that $|\nabla V(x+z) - \nabla V(x)| \leq \frac{\epsilon}{2}$ for any $z \in B(0,\eta)$. Whence, we can split I into two terms and estimate them as

$$\begin{aligned}
|I_R| &\leq \int_0^1 \int_{|x-y|\leq\eta} \left|\langle \nabla V\big(x + s(y-x)\big) - \nabla V(x), y - x\rangle\right| \mathrm{d}\gamma(x,y)\mathrm{d}s \\
&\quad + \int_0^1 \int_{|x-y|>\eta} \left|\langle \nabla V\big(x + s(y-x)\big) - \nabla V(x), y - x\rangle\right| \mathrm{d}\gamma(x,y)\mathrm{d}s \\
&\leq \tfrac{\epsilon}{2}\int_{\mathbb{R}^{2d}} |x-y|\mathrm{d}\gamma(x,y) + 2 \sup_{z\in\overline{\mathrm{co}}\, B_\mu(R)} |\nabla V(z)| \sup_{x,y\in B_\mu(R)} |x-y|\gamma\Big(\Big\{(x,y) \in \mathbb{R}^{2d} \text{ s.t. } |x-y| > \eta\Big\}\Big) \\
&\leq \tfrac{\epsilon}{2}\int_{\mathbb{R}^{2d}} |x-y|\mathrm{d}\gamma(x,y) + \tfrac{C_R}{\eta^2}\int_{\mathbb{R}^{2d}} |x-y|^2\mathrm{d}\gamma(x,y) \\
&\leq \tfrac{\epsilon}{2} W_2(\mu,\nu) + \tfrac{C_R}{\eta^2} W_2(\mu,\nu)^2,
\end{aligned}$$

for some constant $C_R > 0$ depending only on $\mathrm{supp}(\mu)$, $R > 0$ and $V(\cdot)$, and where we used Hölder's and Chebyshev's inequalities together with the fact that $\gamma \in \Gamma_o(\mu,\nu)$. Thus whenever $W_2(\mu,\nu) \leq$

$\eta' := \frac{\eta^2\epsilon}{2C_R}$, it holds that $|I_R| \leq \epsilon W_2(\mu,\nu)$ which precisely amounts to saying that $I_R = o_R(W_2(\mu,\nu))$. Since $\nabla V(\cdot) \in \mathrm{Tan}_\mu \mathscr{P}_2(\mathbb{R}^d)$, (114) implies by Proposition 3.6 that $\mathcal{V}(\cdot)$ is locally differentiable at μ.

Let us now consider the functional $\mathcal{L}(\cdot)$ defined in (113). By adapting the same arguments as above, one has for every $\gamma \in \Gamma_o(\mu,\nu)$

$$\begin{aligned}\mathcal{L}(\nu) - \mathcal{L}(\mu) &= \int_{\mathbb{R}^{2d}} \Big(L(y,\nu) - L(x,\mu)\Big)\mathrm{d}\gamma(x,y) \\ &= \int_{\mathbb{R}^{2d}} \Big(\langle \nabla_x L(x,\mu), y-x\rangle + \big(L(x,\nu) - L(x,\mu)\big)\Big)\mathrm{d}\gamma(x,y) + o_R(W_2(\mu,\nu)),\end{aligned} \tag{115}$$

where we used the fact that $\langle \nabla_x L(x,\nu), x-y\rangle = \langle \nabla_x L(x,\mu), x-y\rangle + o(|x-y|)$ as $W_2(\mu,\nu) \to 0$. Since $\mu \in \mathscr{P}_c(\mathbb{R}^d) \mapsto L(x,\mu)$ is locally differentiable for μ-almost every $x \in \mathbb{R}^d$, it further holds

$$L(x,\nu) - L(x,\mu) = \int_{\mathbb{R}^{2d}} \langle \nabla_\mu L(x,\mu)(z), w-z\rangle \mathrm{d}\gamma(z,w) + o_R(W_2(\mu,\nu)), \tag{116}$$

so that by plugging (116) into (115) and applying Fubini's Theorem, we can derive

$$\mathcal{L}(\nu) - \mathcal{L}(\mu) = \int_{\mathbb{R}^{2d}} \Big\langle \nabla_x L(x,\mu) + \int_{\mathbb{R}^d} \nabla_\mu L(z,\mu)(x)\mathrm{d}\mu(z), y-x\Big\rangle \mathrm{d}\gamma(x,y) + o_R(W_2(\mu,\nu)). \tag{117}$$

Finally observe that $\nabla_x L(\cdot,\mu) \in \mathrm{Tan}_\mu \mathscr{P}_2(\mathbb{R}^d)$ by definition, and also

$$\int_{\mathbb{R}^d} \nabla_\mu L(z,\mu)(\cdot)\mathrm{d}\mu(z) \in \mathrm{Tan}_\mu \mathscr{P}_2(\mathbb{R}^d),$$

because $\nabla_\mu L(z,\mu)(\cdot) \in \mathrm{Tan}_\mu \mathscr{P}_2(\mathbb{R}^d)$ for any $z \in \mathrm{supp}(\mu)$ and

$$\int_{\mathbb{R}^d} \Big\langle \int_{\mathbb{R}^d} \nabla_\mu L(z,\mu)(x)\mathrm{d}\mu(z), \xi(x)\Big\rangle \mathrm{d}\mu(x) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \langle \nabla_\mu L(z,\mu)(x), \xi(x)\rangle \mathrm{d}\mu(x)\mathrm{d}\mu(z) = 0,$$

whenever $\xi \in \mathrm{Tan}_\mu \mathscr{P}_2(\mathbb{R}^d)^\perp$, by linearity of the integral and Fubini's Theorem. Therefore, the identity (117) implies together with Proposition 3.6 that $\mathcal{L}(\cdot)$ is locally differentiable at μ. □

Remark 5.2 (Global definition of Wasserstein gradient)**.** *In general, the gradient $\nabla\phi(\mu)$ of a functional $\phi : \mathscr{P}_c(\mathbb{R}^d) \to \mathbb{R}$ at $\mu \in \mathscr{P}_c(\mathbb{R}^d)$ is an element of $L^2(\mathbb{R}^d, \mathbb{R}^d; \mu)$, and is thus well-defined on* $\mathrm{supp}(\mu)$*. In the particular case where $\nabla\phi(\mu)$ is a continuous map – as in Proposition 5.1 above – , it is possible to extend it as a continuous function defined over the whole of $\mathbb{R}^d$ and not only* $\mathrm{supp}(\mu)$*.*

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

References

[1] G. Albi, M. Bongini, E. Cristiani, and D. Kalise. Invisible Control of Self-Organizing Agents Leaving Unknown Environments. *SIAM Journal of Applied Mathematics*, 76(4):1683–1710, 2016.

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

[4] L. Ambrosio and N. Gigli. Construction of the Parallel Transport in the Wasserstein Space. *Methods and Applications of Analysis*, 15(1):1–30, 2008.

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

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

[8] N. Bellomo, M. A. Herrero, and A. Tosin. On the Dynamics of Social Conflicts: Looking for the Black Swan. *Kinetic & Related Models*, 6(3):459–479, 2013.

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

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

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

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

[13] B. Bonnet and F. Rossi. Intrinsic Lipschitz Regularity of Mean-Field Optimal Controls. *To appear in SIAM Journal on Control and Optimization, arXiv:1908.04183*, 2021.

[14] A. Bressan and B. Piccoli. *Introduction to the Mathematical Theory of Control*, volume 2 of *AIMS Series on Applied Mathematics*. American Institute of Mathematical Sciences (AIMS), Springfield, MO, 2007.

[15] F. Bullo, J. Cortés, and S. Martines. *Distributed Control of Robotic Networks*. Applied Mathematics. Princeton University Press, 2009.

[16] 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.

[17] R. Carmona and F. Delarue. Forward-Backward Stochastic Differential Equations and Controlled McKean Vlasov Dynamics. *Annals of Probability*, 43(5):2647–2700, 2015.

[18] R. Carmona and F. Delarue. *Probabilistic Theory of Mean-Field Games with Applications*. Springer, 2018.

[19] G. Cavagnari, S. Lisini, C. Orrieri, and G. Savaré. Lagrangian, Eulerian and Kantorovich Formulations of Multi-Agent Optimal Control Problems: Equivalence and Gamma-Convergence. *arxiv preprint arXiv:2011.07117*, 2021.

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

[21] Y.-P. Choi, J.A. Carrillo, and M. Hauray. The Derivation of Swarming Models: Mean-Field Limit and Wasserstein Distances. In *Collective Dynamics from Bacteria to Crowds, CISM International Centre for Mechanical Sciences*, volume 553, pages 1–46. Springer, 2014.

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

[23] F. Cucker and S. Smale. On the Mathematics of Emergence. *Japanese Journal of Mathematics*, 2(1):197–227, 2007.

[24] C.M. Dafermos and E. Feireisl. *Handbook of Differential Equations: Evolutionary Equations*, volume 3. North-Holland, 2006.

[25] J. Diestel and J.J.Jr Uhl. *Vector Measures*, volume 15. American Mathematical Society, 1977.

[26] 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.

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

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

[29] H. Frankowska. The Maximum Principle for an Optimal Solution to a Differential Inclusion with End-Point Constraints. *SIAM Journal on Control and Optimization*, 25(1):145–157, 1987.

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

[31] H. Frankowska and Q. Lü. First and Second Order Necessary Optimality Conditions for Controlled Stochastic Evolution Equations with Control and State Constraints. *Journal of Differential Equations*, 268(6):2949–3015, 2020.

[32] 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.

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

[34] 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.

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

[36] H. Frankowska, H. Zhang, and X. Zhang. Stochastic Optimal Control Problems with Control and Initial-Final States Constraints. *SIAM Journal on Control and Optimization*, 56(3):1823–1855, 2018.

[37] W. Gangbo and A. Tudorascu. On Differentiability in the Wasserstein Space and Well-Posedness for Hamilton-Jacobi Equations. *Journal de Mathématiques Pures et Appliquées*, 125:119–174, 2019.

[38] 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.

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

[40] R.J. McCann. A Convexity Principle for Interacting Gases. *Advances in Mathematics*, 128(1):153–179, 1997.

[41] A. Muntean, J. Rademacher, and A. Zagaris. *Macroscopic and Large Scale Phenomena: Coarse Graining, Mean Field Limits and Ergodicity.* Springer, 2016.

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

[43] 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.

[44] 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.

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

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

[47] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze, and E.F. Mischenko. *The Mathematical Theory of Optimal Processes*, volume 4. CRC Press, 1987.

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

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