



**HAL**  
open science

# A Pontryagin Maximum Principle in Wasserstein Spaces for Constrained Optimal Control Problems

Benoît Bonnet

► **To cite this version:**

Benoît Bonnet. A Pontryagin Maximum Principle in Wasserstein Spaces for Constrained Optimal Control Problems. 2019. hal-01937106v4

**HAL Id: hal-01937106**

**<https://hal.science/hal-01937106v4>**

Preprint submitted on 9 Apr 2019 (v4), last revised 21 Oct 2019 (v6)

**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.

# A Pontryagin Maximum Principle in Wasserstein Spaces for Constrained Optimal Control Problems

Benoît Bonnet\*

April 9, 2019

## Abstract

In this paper, we prove a Pontryagin Maximum Principle for constrained optimal control problems in the Wasserstein space of probability measures. The dynamics, is described by a transport equation with non-local velocities and is subject to end-point and running state constraints. Building on our previous work, we combine the classical method of needle-variations from geometric control theory and the metric differential structure of the Wasserstein spaces to obtain a maximum principle stated in the so-called Gamkrelidze form.

## 1 Introduction

Transport equations with non-local velocities have drawn a great amount of attention from several scientific communities for almost a century. They were first introduced in statistical physics to describe averaged Coulomb interactions within large assemblies of particles (see e.g. [43]), and are still to this day a widely studied topic in mathematical physics. More recently, a growing interest in the mathematical modelling of multi-agent systems has opened a whole new panel of problems in which these equations play a central role. Starting from the seminal paper of Cucker and Smale [20] dealing with emergent behaviour in animal flocks, a large literature has been devoted to the fine mathematical analysis of kinetic cooperative systems, i.e. systems described by non-local transport equations with attractive velocity fields [4, 12, 27, 33].

Later on, the focus shifted partly to include control-theoretic problems such as reachability analysis, optimal control, explicit design of sparse control strategies. For these purposes, the vast majority of the existing contributions have taken advantage of the recent developments in the theory of optimal transport. We refer the reader to [41, 39] for a comprehensive introduction to this ever-expanding topic. In particular, the emergence of powerful tools of analysis in the so-called Wasserstein spaces has allowed for the establishment of a comprehensive existence theory for these equations (see e.g. [9, 8]), which incorporates natural Lipschitz and metric estimates in the smooth cases (see [37]).

Apart from a few controllability results as in [24], most of the attention of the community has been devoted to optimal control problems in Wasserstein spaces. The existence of optimal solutions has been investigated with various degrees of generality in [2, 1, 29, 30, 28], mostly by means of  $\Gamma$ -convergence arguments. Besides, a few papers have been dealing with numerical methods either in the presence of diffusion terms, which considerably simplify the convergence analysis of the corresponding schemes (see e.g. [25]), or in the purely metric setting through Boltzmann-based methods [5]. However to the best of our knowledge, there currently are no general results describing shooting methods or PDE-flavoured gradient schemes for optimal control problems in Wasserstein spaces.

On the contrary, the derivation of Hamilton-Jacobi and Pontryagin optimality conditions - which are customary and well-known in the classical theory of optimal control - has been an active topic in

---

\*Aix Marseille Université, CNRS, ENSAM, Université de Toulon, LIS, Marseille, France. *benoit.bonnet@lis-lab.fr*

the recent years. Starting from the seminal paper [31] on Hamilton-Jacobi equations in the Wasserstein space, several contributions such as [17, 18] aim at refining a dynamic-programming principle for mean-field optimal control problems. Pontryagin-type optimality conditions, on the other hand, have received less interest. The first result derived in [13] focuses on a multi-scale ODE-PDE system in which the control only acts on the ODE part. In this setting, the Pontryagin Maximum Principle ("PMP" for short) is derived by combining  $\Gamma$ -convergence and mean-field limit arguments. Another approach, introduced in our previous work [14] studies the infinite-dimensional problem by means of the classical technique of needle-variations (see e.g. [3, 15]) and makes an extensive use of the theory of Wasserstein subdifferential calculus [9]. The corresponding maximum principle is formulated as a Hamiltonian flow in the space of measures in the spirit of [8], and is in a sense the most natural generalization to be expected of the usual finite-dimensional Pontryagin-type optimality conditions.

### 1.1 Statement of the main result

In this paper, we further this line of research by extending our previous result to the setting of constrained optimal control problems given in following the general form

$$(\mathcal{P}) \quad \begin{cases} \min_{u \in \mathcal{U}} \left[ \int_0^T L(t, \mu(t), u(t)) dt + \varphi(\mu(T)) \right], \\ \text{s.t.} \quad \begin{cases} \partial_t \mu(t) + \nabla \cdot (u(t, \cdot)) \mu(t) = 0, \\ \mu(0) = \mu^0 \in \mathcal{P}_c(\mathbb{R}^d). \end{cases} \\ \text{and} \quad \begin{cases} \Psi^I(\mu(T)) \leq 0, \quad \Psi^E(\mu(T)) = 0, \\ \Lambda(t, \mu(t)) \leq 0 \quad \text{for all } t \in [0, T]. \end{cases} \end{cases}$$

The methodology that we follow relies on the technique of *packages of needle-variations*, combined with a *Lagrange multiplier rule*. In essence, this method allows to derive the maximum principle from a family of finite dimensional first-order optimality conditions combined with the introduction of a suitable costate. Even though classical in the unconstrained case, this direct approach requires some care to be translated to constrained problem. Indeed, the presence of constraints induces an unwanted dependency between the Lagrange multipliers and the needle-parameters. This extra difficulty can be circumvented by considering  $N$ -dimensional perturbations of the optimal trajectory instead of a single one, and by performing a limiting procedure as  $N \rightarrow +\infty$ . Originally introduced in [11] for smooth optimal control problems with end-point constraints, this approach was extended in [40] to the case of non-smooth and state-constrained problems.

In the following Theorem, we state the main result of this article consisting in a Pontryagin Maximum Principle formulated in the so-called Gamkrelidze form (see e.g. [10]) for problem  $(\mathcal{P})$ . Its proof is inspired from that of [40] and from our previous contribution [14].

**Theorem 1.1** (Pontryagin Maximum Principle for  $(\mathcal{P})$ ). *Let  $(u^*(\cdot), \mu^*(\cdot))$  be an optimal pair control-trajectory for  $(\mathcal{P})$  and assume that the set of hypotheses **(H)** below holds.*

**(H)**

**(H1)** *The set of admissible controls is defined as  $\mathcal{U} = L^\infty([0, T], U)$  where  $U$  is any  $C^0$ -closed subset of  $\{\omega \in C^0(\mathbb{R}^d, \mathbb{R}^d) \text{ s.t. } \|\omega(\cdot)\|_{C^0(\mathbb{R}^d, \mathbb{R}^d)} + \text{Lip}(\omega(\cdot), \mathbb{R}^d) \leq L_U\}$  for a given constant  $L_U > 0$ .*

**(H2)** *The optimal control  $(t, x) \mapsto u^*(t, x) \in \mathbb{R}^d$  is measurable with respect to  $t \in [0, T]$  and  $C^1$ -smooth with respect to  $x \in \mathbb{R}^d$ .*

**(H3)** The non-local velocity field  $\mu \mapsto v[\mu] \in L^\infty([0, T], C^1(\mathbb{R}^d, \mathbb{R}^d))$  satisfies the classical Cauchy-Lipschitz assumptions in Wasserstein spaces, i.e. there exists positive constants  $M, L_1$  and  $L_2$  such that

$$\begin{aligned} |v[\mu](t, x)| &\leq M(1 + |x|), \quad |v[\mu](t, x) - v[\mu](t, y)| \leq L_1|x - y|, \\ |v[\mu](t, x) - v[\nu](t, x)| &\leq L_2W_1(\mu, \nu) \end{aligned}$$

for all  $(x, y) \in \mathbb{R}^d$  and  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ . Moreover, the maps  $\mu \mapsto D_x v[\mu](t, x)$  and  $(x, y) \mapsto \nabla_\mu (v^i[\cdot](t, x))(\mu)(y)$  are continuous, uniformly with respect to  $t \in [0, T]$ .

**(H4)** The final cost  $\mu \mapsto \varphi(\mu)$  and the boundary constraints maps  $\mu \mapsto (\Psi_i^I(\mu), \Psi_j^E(\mu)) \in \mathbb{R}^{n_I + n_E}$  are bounded and Lipschitz continuous in the  $W_2$ -metric over  $\mathcal{P}(K)$  for any compact set  $K \subset \mathbb{R}^d$ . They are furthermore Wasserstein-differentiable at  $\mu^*(T)$  and their Wasserstein gradients  $\nabla_\mu \varphi(\mu^*(T))(\cdot)$ ,  $\nabla_\mu \Psi_i^I(\mu^*(T))(\cdot)$  and  $\nabla_\mu \Psi_j^E(\mu^*(T))(\cdot)$  are continuous.

**(H5)** The running cost  $(t, \mu, \omega) \mapsto L(t, \mu, \omega)$  is  $\mathcal{L}^1$ -measurable with respect to  $t \in [0, T]$  and Lipschitz continuous with respect to the product  $W_2 \times C^0$ -metric defined over  $\mathcal{P}(K) \times U$  for any compact set  $K \subset \mathbb{R}^d$ . It is furthermore Wasserstein-differentiable at  $\mu^*(t)$  for any  $t \in [0, T]$  and its Wasserstein gradient  $\nabla_\mu L(t, \mu^*(t), \omega)(\cdot)$  is continuous, uniformly with respect to  $\omega \in U$ .

**(H6)** The state constraints maps  $(t, \mu) \mapsto (\Lambda_l(t, \mu))_{1 \leq l \leq r} \in \mathbb{R}^r$  are bounded and Lipschitz-continuous over  $[0, T] \times \mathcal{P}(K)$  for any compact set  $K \subset \mathbb{R}^d$ . Moreover, the maps  $(t, \mu) \mapsto \partial_t \Lambda(t, \mu)$  and  $(t, \mu) \mapsto \nabla_\mu \Lambda_l(t, \mu)(\cdot)$  are well-defined and continuously differentiable at  $(t, \mu^*(t))$  with

$$\begin{aligned} \nabla_\mu \partial_t \Lambda_l(t, \mu^*(t))(\cdot) &= \partial_t \nabla_\mu \Lambda_l(t, \mu^*(t))(\cdot) \in C^0(\mathbb{R}^d, \mathbb{R}^d), \quad D_x \nabla_\mu \Lambda_l(t, \mu^*(t))(\cdot) \in C^0(\mathbb{R}^d, \mathbb{R}^d), \\ \nabla_\mu [\nabla_\mu \Lambda(t, \mu^*(t))(\cdot)](\cdot) &\in C^0(\mathbb{R}^{2d}, \mathbb{R}^d) \end{aligned}$$

Then there exists a constant  $R'_T > 0$ , Lagrange multipliers  $(\lambda_0, \dots, \lambda_n, \eta_1, \dots, \eta_m, \varpi_1, \dots, \varpi_l) \in \{0, 1\} \times \mathbb{R}_+^n \times \mathbb{R}^m \times \mathcal{M}_+([0, T], \mathbb{R}^d)^l$  and a curve  $\nu^*(\cdot) \in \text{Lip}([0, T], \mathcal{P}(\overline{B_{2d}(0, R'_T)}))$  such that

(i) The map  $t \mapsto \nu^*(t)$  is a solution of the forward-backward system of continuity equations

$$\begin{cases} \partial_t \nu^*(t) + \nabla \cdot (\mathbb{J}_{2d} \nabla_\nu \mathcal{H}(t, \nu^*(t), \zeta^*(t), u^*(t)) \nu^*(t)) = 0 & \text{in } [0, T] \times \mathbb{R}^{2d}, \\ \pi_{\#}^1 \nu^*(0) = \mu^0, \\ \pi_{\#}^2 \nu^*(T) = (-\nabla_\mu \mathcal{S}(\mu^*(T)))_{\#} \mu^*(T), \end{cases} \quad (1.1)$$

where the augmented infinite-dimensional Hamiltonian  $\mathcal{H}(\cdot, \cdot, \cdot, \cdot)$  of the system is defined by

$$\mathcal{H}(t, \nu, \zeta, \omega) = \int_{\mathbb{R}^{2d}} \langle r, v[\pi_{\#}^1 \nu](t, x) + \omega(x) \rangle d\nu(x, r) - L(t, \pi_{\#}^1 \nu, \omega) - \mathcal{C}(t, \pi_{\#}^1 \nu, \zeta), \quad (1.2)$$

for any  $(t, \nu, \zeta, \omega) \in [0, T] \times \mathcal{P}(\overline{B_{2d}(0, R'_T)}) \times \mathbb{R}^r \times U$ . The final gradient and penalized state constraints maps are given respectively by

$$\nabla_\mu \mathcal{S}(\mu) = \lambda_0 \nabla_\mu \varphi(\mu) + \sum_{i=1}^n \lambda_i \nabla_\mu \Psi_i^I(\mu) + \sum_{j=1}^m \eta_j \nabla_\mu \Psi_j^E(\mu), \quad (1.3)$$

and

$$\mathcal{C}(t, \mu, \zeta) = \sum_{l=1}^r \zeta_l \left( \partial_t \Lambda_l(t, \mu) + \int_{\mathbb{R}^d} \langle \nabla_{\mu} \Lambda_l(t, \mu)(x), v[\mu](t, x) + u^*(t, x) \rangle d\mu(x) \right). \quad (1.4)$$

For all  $l \in \{1, \dots, r\}$ , the map  $t \in [0, T] \mapsto \zeta_l^*(t)$  denotes the cumulated state constraints multiplier, defined by

$$\zeta_l^*(t) = \mathbb{1}_{[0, T]}(t) \int_t^T d\varpi_l(s),$$

(ii) The Lagrange multipliers satisfy the Non-Degeneracy condition

$$(\lambda_0, \dots, \lambda_n, \eta_1, \dots, \eta_m, \varpi_1, \dots, \varpi_r) \neq 0. \quad (1.5)$$

as well as the Complementary Slackness condition

$$\begin{cases} \lambda_i \Psi_i^I(\mu^*(T)) = 0 & \text{for all } i \in \{1, \dots, n\}, \\ \text{supp}(\varpi_l) = \{t \in [0, T] \text{ s.t. } \Lambda_l(t, \mu(t)) = 0\} & \text{for all } l \in \{1, \dots, r\}. \end{cases}$$

(iii) The Pontryagin maximization condition

$$\mathcal{H}(t, \nu^*(t), \zeta^*(t), u^*(t)) = \max_{\omega \in U} [\mathcal{H}(t, \nu^*(t), \zeta^*(t), \omega)] \quad (1.6)$$

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

**Remark 1** (The Gamkrelidze Maximum Principle). *The so-called Gamkrelidze formulation of the PMP corresponds to the case in which one includes the derivative of the state constraints inside the Hamiltonian function. The consequence of this choice is that the costate variables are absolutely continuous in time instead of being merely BV as it is the case in the more classical formulation of the constrained PMP (see e.g. [42, Chapter 9]).*

*This fact is crucial for our purpose, since absolutely continuous curves in Wasserstein spaces are exactly those curves which solve a continuity equation (see e.g. [9, Theorem 8.3.1]). Hence, one cannot derive a Hamiltonian system such as (1.1) by sticking to the classical formulation of the PMP.*

**Remark 2** (On the regularity hypotheses **(H1)**-**(H2)**). *One of the distinctive features of continuity equations in Wasserstein spaces, compared to other families of PDEs is that they require Cauchy-Lipschitz assumptions on the driving velocity fields in order to be classically well-posed. Even though the existence theory has gone far beyond this basic setting, notably through the DiPerna-Lions-Ambrosio theory (see [21, 6] or [9, Section 8.2]), such extensions come at the price of losing the strict micro-macro correspondence of the solutions embodied by the underlying flow-structure. Therefore, from a mean-field control-theoretic viewpoint, it seemed more meaningful for the authors to work in a setting where classical well-posedness holds for the optimal trajectory.*

*Furthermore, the proof of Theorem 1.1 relies heavily on the geometric flow of diffeomorphism structure of the underlying characteristic system of ODEs, both forward and backward in time. For this reason, the Lipschitz-regularity assumption **(H1)** is instrumental in our argumentation. Let it be remarked however that there exists common examples of Wasserstein optimal control problems for which the optimal control is  $C^1$ -smooth in space. Such a situation is given e.g. by controlled vector fields of the form  $u(t, x) = \sum_{k=1}^m u_k(t) X_k(\cdot)$  where  $X_1, \dots, X_m \in C^1(\mathbb{R}^d, \mathbb{R}^d)$  and  $u_1, \dots, u_m \in L^\infty([0, T], \mathbb{R})$ .*

## 1.2 Examples of functionals satisfying hypotheses (H)

In this Section, we show that the rather long list of hypotheses (H) is not too restrictive and that a good score of relevant functionals for applications fit into the framework of Theorem 1.1.

**Proposition 1.1** (Example of non-local velocity field). *Let  $(t, x, y) \in \mapsto H(t, x, y) \in \mathbb{R}^d$  be measurable with respect to  $t \in [0, T]$ , sublinear and  $C^1$ -with respect to  $(x, y) \in \mathbb{R}^{2d}$ . Then, the map  $\mu \in \mathcal{P}_c(\mathbb{R}^d) \mapsto v[\mu](\cdot, \cdot)$  defined by*

$$v[\mu](t, x) = \int_{\mathbb{R}^d} H(t, x, y) d\mu(y)$$

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$  and any  $x \in \mathbb{R}^d$  satisfies the hypotheses (H3) of Theorem 1.1. Moreover, its first-order variations  $D_x v[\mu](t, x)$  and  $\int_{\mathbb{R}^d} \mathbb{I}_{(t,x)}^v(y) d\mu(y)$  are given by

$$D_x v[\mu](t, x) = \int_{\mathbb{R}^d} D_x H(t, x, y) d\mu(y) \quad , \quad \int_{\mathbb{R}^d} \mathbb{I}_{(t,x)}^v(y) d\mu(y) = \int_{\mathbb{R}^d} D_y H(t, x, y) d\mu(y).$$

**Proposition 1.2** (Example of cost and constraint functions). *Let  $n \geq 1$  and  $W \in C^1(\mathbb{R}^{nd}, \mathbb{R})$ . Then, the functional*

$$\varphi : \mu \in \mathcal{P}_c(\mathbb{R}^d) \mapsto \int_{\mathbb{R}^{kd}} W(x_1, \dots, x_k) d\mu^{\otimes n}(x_1, \dots, x_k)$$

with  $\mu^{\otimes n} = \mu \times \dots \times \mu$  satisfy hypotheses (H4) of Theorem 1.1 and its Wasserstein gradient at some  $\mu \in \mathcal{P}(K)$  is given by

$$\nabla_{\mu} \varphi(\mu)(x_1, \dots, x_n) = \sum_{j=1}^n \nabla_{x_j} W(x_1, \dots, x_n).$$

Let  $m \in C^1(\mathbb{R}^d, \mathbb{R}^n)$  and  $(t, x, v, r) \mapsto l(t, x, v, r) \in \mathbb{R}$  be  $\mathcal{L}^1$ -measurable with respect to  $t \in [0, T]$  and  $C^1$ -smooth with respect to  $(x, v, r) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^n$ . Then, the functional

$$L : (t, \mu, \omega) \in [0, T] \times \mathcal{P}(K) \times U \mapsto \int_{\mathbb{R}^d} l(t, x, \omega(x), \int m d\mu) d\mu(x),$$

satisfies the hypotheses (H5) of Theorem 1.1 and its Wasserstein gradient is given by

$$\begin{aligned} \nabla_{\mu} L(t, \mu, \omega)(x) &= \nabla_x l(t, x, \omega(x), \int m d\mu) + D_x \omega(x)^{\top} \nabla_v l(t, x, \omega(x), \int m d\mu) \\ &\quad + D_x m(x)^{\top} \int_{\mathbb{R}^d} \nabla_r l(t, y, \omega(y), \int m d\mu) d\mu(y). \end{aligned}$$

**Proposition 1.3** (Example of state constraints). *Let  $m \in C^2(\mathbb{R}^d, \mathbb{R}^k)$  and  $\lambda \in C^2([0, T] \times \mathbb{R}^d \times \mathbb{R}^k, \mathbb{R}^r)$ . Then for any  $l \in \{1, \dots, r\}$ , the functionals*

$$\Lambda_l(t, \mu) \in [0, T] \times \mathcal{P}(K) \mapsto \int_{\mathbb{R}^d} \lambda_l(t, x, \int m d\mu) d\mu(x)$$

satisfy the hypotheses (H6) of Theorem 1.1 and their derivatives can be computed using Propositions 1.2.

**Remark 3.** *Particular cases of functionals which are of great interest for applications are for instance the variance functional  $\mu \mapsto \int_{\mathbb{R}^d} |x - \bar{\mu}|^2 d\mu(x)$  where  $\bar{\mu} = \int y d\mu(y)$  or the target-support map to a closed set  $S \subset \mathbb{R}^d$   $\mu \mapsto \frac{1}{2} \int_{\mathbb{R}^d} d_S(x)^2 d\mu(x)$ .*

## 2 Preliminary results

In this Section, we recall several notions about analysis in the space of measures, optimal transport theory, Wasserstein spaces, continuity equations and subdifferential calculus in the space  $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ . We also introduce some elementary notions of non-smooth calculus in Banach spaces. For a complete introduction to these topics, see [9, 41] and [36, 19] respectively.

## 2.1 Analysis in measure spaces and the optimal transport problem

In this Section, we introduce some classical notations and results of measure theory, optimal transport and analysis in Wasserstein spaces. We denote by  $(\mathcal{M}_+(\mathbb{R}^d), \|\cdot\|_{TV})$  the set of real-valued non-negative Borel measures defined on  $\mathbb{R}^d$  and by  $\mathcal{L}^d$  the standard Lebesgue measure on  $\mathbb{R}^d$ . It is known by Riesz Theorem (see e.g. [26]) that this space can be identified with the topological dual of the Banach space  $(C_c^0(\mathbb{R}^d), \|\cdot\|_{C^0})$  of continuous and compactly supported functions endowed with the  $C^0$ -norm.

We denote by  $\mathcal{P}(\mathbb{R}^d) \subset \mathcal{M}_+(\mathbb{R}^d)$  the set of Borel probability measures, and for  $p \geq 1$  we define  $\mathcal{P}_p(\mathbb{R}^d)$  as the subset of  $\mathcal{P}(\mathbb{R}^d)$  of measures having finite  $p$ -th moment, i.e.

$$\mathcal{P}_p(\mathbb{R}^d) = \left\{ \mu \in \mathcal{P}(\mathbb{R}^d) \text{ s.t. } \int_{\mathbb{R}^d} |x|^p d\mu(x) < +\infty \right\}.$$

The *support* of a Borel measure  $\mu \in \mathcal{M}_+(\mathbb{R}^d)$  is defined as the closed set  $\text{supp}(\mu) = \{x \in \mathbb{R}^d \text{ s.t. } \mu(\mathcal{N}) > 0 \text{ for any neighbourhood } \mathcal{N} \text{ of } x\}$ . We denote by  $\mathcal{P}_c(\mathbb{R}^d) \subset \mathcal{P}(\mathbb{R}^d)$  the set of probability measures with compact support.

We say that a sequence  $(\mu_n) \subset \mathcal{P}(\mathbb{R}^d)$  of Borel probability measures *converges narrowly* towards  $\mu \in \mathcal{P}(\mathbb{R}^d)$ , denoted by  $\mu_n \xrightarrow{n \rightarrow +\infty} \mu$ , provided that

$$\int_{\mathbb{R}^d} \phi(x) d\mu_n(x) \xrightarrow{n \rightarrow +\infty} \int_{\mathbb{R}^d} \phi(x) d\mu(x), \quad (2.1)$$

for all  $\phi \in C_b^0(\mathbb{R}^d)$ , where  $C_b^0(\mathbb{R}^d)$  denotes the set of continuous and bounded real-valued functions over  $\mathbb{R}^d$ . This convergence coincides with the usual weak-\* topology in  $\mathcal{M}_+(\mathbb{R}^d)$ , which itself induces a convergence property on the measures of Borel sets usually referred to as the *Portmanteau Theorem* (see e.g. [7, Proposition 1.62]).

**Proposition 2.1** (Portmanteau Theorem). *Let  $(\mu_n) \subset \mathcal{M}_+(\mathbb{R}^d)$  be a sequence of measures converging in the weak-\* topology towards  $\mu \in \mathcal{M}(\mathbb{R}^d, \mathbb{R})$ . Then for any Borel set  $A \subset \mathbb{R}^d$  such that  $\mu(\partial A) = 0$ , it holds that  $\mu_n(A) \xrightarrow{n \rightarrow +\infty} \mu(A)$ .*

We recall in the following definition the notions of *pushforward* of a Borel probability measure through a Borel map and of *transport plan*.

**Definition 2.1** (Pushforward of a measure through a Borel map). *Given  $\mu \in \mathcal{P}(\mathbb{R}^d)$  and a Borel map  $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$ , the pushforward  $f_{\#}\mu$  of  $\mu$  through  $f(\cdot)$  is defined as the Borel probability measure such that  $f_{\#}\mu(B) = \mu(f^{-1}(B))$  for any Borel set  $B \subset \mathbb{R}^d$ .*

**Definition 2.2** (Transport plan). *Let  $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$ . We say that  $\gamma \in \mathcal{P}(\mathbb{R}^{2d})$  is a transport plan between  $\mu$  and  $\nu$ , denoted by  $\gamma \in \Gamma(\mu, \nu)$ , provided that  $\gamma(A \times \mathbb{R}^d) = \mu(A)$  and  $\gamma(\mathbb{R}^d \times B) = \nu(B)$  for any Borel sets  $A, B \subset \mathbb{R}^d$ . This property can be equivalently formulated in terms of pushforwards as  $\pi_{\#}^1 \gamma = \mu$  and  $\pi_{\#}^2 \gamma = \nu$ .*

In 1942, the Russian mathematician Leonid Kantorovich introduced the *optimal mass transportation problem* in its modern mathematical formulation. Given two probability measures  $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$  and a cost function  $c : \mathbb{R}^{2d} \rightarrow \mathbb{R}$ , one searches for a *transport plan*  $\gamma \in \Gamma(\mu, \nu)$  such that

$$\int_{\mathbb{R}^{2d}} c(x, y) d\gamma(x, y) = \min_{\gamma} \left\{ \int_{\mathbb{R}^{2d}} c(x, y) d\gamma'(x, y) \text{ s.t. } \gamma' \in \Gamma(\mu, \nu) \right\}.$$

This problem has been extensively studied in very broad contexts (see e.g. [9, 41]) with high levels of generality on the underlying spaces and cost functions. In the particular case where  $c(x, y) = |x - y|^p$  for some real number  $p \geq 1$ , the optimal transport problem can be used to define a distance over the subset  $\mathcal{P}_p(\mathbb{R}^d) \subset \mathcal{P}(\mathbb{R}^d)$ .

**Definition 2.3** (Wasserstein distance and Wasserstein spaces). *Given two probability measures  $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$ , the  $p$ -Wasserstein distance between  $\mu$  and  $\nu$  is defined by*

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

The set of plans  $\gamma \in \Gamma(\mu, \nu)$  achieving this optimal value is denoted<sup>1</sup> by  $\Gamma_o(\mu, \nu)$  and referred to as the set of optimal transport plans between  $\mu$  and  $\nu$ . The space  $(\mathcal{P}_p(\mathbb{R}^d), W_p)$  of probability measures with finite  $p$ -th moment endowed with the  $p$ -Wasserstein metric is called the Wasserstein space of order  $p$ .

We recall some of the interesting properties of these spaces in the following Proposition (see e.g. [9, Chapter 7] or [41, Chapter 6]).

**Proposition 2.2** (Properties of the Wasserstein spaces). *The Wasserstein spaces  $(\mathcal{P}_p(\mathbb{R}^d), W_p)$  are separable geodesic spaces. The topology generated by the  $p$ -Wasserstein metric metrizes the weak-\* topology of probability measures induced by the narrow convergence (2.1). More precisely, it holds that*

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

Given two measures  $\mu, \nu \in \mathcal{P}_c(\mathbb{R}^d)$ , the Wasserstein distances are ordered, i.e.  $W_{p_1}(\mu, \nu) \leq W_{p_2}(\mu, \nu)$  whenever  $p_1 \leq p_2$ . Moreover, when  $p = 1$ , the following Kantorovich-Rubinstein duality formula holds

$$W_1(\mu, \nu) = \sup \left\{ \int_{\mathbb{R}^d} \phi(x) d(\mu - \nu)(x) \quad \text{s.t. } \text{Lip}(\phi, \mathbb{R}^d) \leq 1 \right\}. \quad (2.2)$$

In what follows, we shall mainly restrict our considerations to the Wasserstein spaces of order 1 and 2 built over  $\mathcal{P}_c(\mathbb{R}^d)$ . We end these introductory paragraphs by recalling the concepts of *disintegration* and *barycenter* in the context of optimal transport.

**Definition 2.4** (Disintegration and barycenter). *Let  $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$  and  $\gamma \in \Gamma(\mu, \nu)$  be a transport plan between  $\mu$  and  $\nu$ . We define the disintegration  $\{\gamma_x\}_{x \in \mathbb{R}^d} \subset \mathcal{P}_p(\mathbb{R}^d)$  of  $\gamma$  on its first marginal  $\mu$ , usually denoted by  $\gamma = \int \gamma_x d\mu(x)$ , as the  $\mu$ -almost uniquely determined Borel family of probability measures such that*

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

for any Borel map  $\phi : \mathbb{R}^{2d} \rightarrow \mathbb{R}^d$ . The barycenter  $\bar{\gamma} \in L^p(\mathbb{R}^d, \mathbb{R}^d; \mu)$  of the plan  $\gamma$  is then defined by

$$\bar{\gamma} : x \in \text{supp}(\mu) \mapsto \int_{\mathbb{R}^d} y d\gamma_x(y).$$

**Proposition 2.3** (Wasserstein estimate between disintegrations). *Let  $\mu \in \mathcal{P}_c(\mathbb{R}^d)$ ,  $\gamma^1 = \int \gamma_x^1 d\mu(x) \in \mathcal{P}_c(\mathbb{R}^{2d})$  and  $\gamma^2 = \int \gamma_x^2 d\mu(x) \in \mathcal{P}_c(\mathbb{R}^{2d})$ . Then, it holds that*

$$W_1(\gamma^1, \gamma^2) \leq \int_{\mathbb{R}^d} W_1(\gamma_x^1, \gamma_x^2) d\mu(x). \quad (2.3)$$

*Proof.* Take  $\xi \in \text{Lip}(\mathbb{R}^{2d})$  with  $\text{Lip}(\xi, \mathbb{R}^{2d}) \leq 1$ . One has that

$$\int_{\mathbb{R}^{2d}} \xi(x, r) d(\gamma^1 - \gamma^2)(x, r) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \xi(x, r) d(\gamma_x^1 - \gamma_x^2)(r) d\mu(x) \leq \int_{\mathbb{R}^d} W_1(\gamma_x^1, \gamma_x^2) d\mu(x)$$

by Kantorovich duality (2.2) since the maps  $r \mapsto \xi(x, r)$  are 1-Lipschitz for all  $\mu$ -almost every  $x \in \mathbb{R}^d$ . Taking now the supremum over  $\xi \in \text{Lip}(\mathbb{R}^{2d})$  with  $\text{Lip}(\xi, \mathbb{R}^{2d}) \leq 1$  yields the desired estimate.  $\square$

<sup>1</sup>We omit the dependence on  $p$  for clarity and conciseness.

## 2.2 Subdifferential calculus in $(\mathcal{P}_2(\mathbb{R}^d), W_2)$

In this Section, we recall the definition of Wasserstein gradients for Wasserstein differentiable maps in the spirit of [32] and state a chain rule formula for Wasserstein differentiable functionals along some  $N$ -dimensional families measures. We refer the reader to [9, Chapters 9-11] for a thorough introduction to the theory of subdifferential calculus in Wasserstein spaces, as well as to [32] and [41, Chapter 15] for complementary material.

Let  $\phi : \mathcal{P}_2(\mathbb{R}^d) \rightarrow (-\infty, +\infty]$  be a lower-semicontinuous and proper functional with effective domain  $D(\phi) = \{\mu \in \mathcal{P}_2(\mathbb{R}^d) \text{ s.t. } \phi(\mu) < +\infty\}$ . We introduce in the following definition the concepts of *classical subdifferential* and *superdifferential* in  $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ .

**Definition 2.5** (Wasserstein subdifferentials and superdifferentials). *Let  $\mu \in D(\phi)$ . We say that a map  $\xi \in L^2(\mathbb{R}^d, \mathbb{R}^d; \mu)$  belongs to the classical subdifferential  $\partial^- \phi(\mu)$  of  $\phi(\cdot)$  at  $\mu$  provided that*

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

for all  $\nu \in \mathcal{P}_2(\mathbb{R}^d)$ . Similarly, we say that a map  $\xi \in L^2(\mathbb{R}^d, \mathbb{R}^d; \mu)$  belongs to the classical superdifferential  $\partial^+ \phi(\mu)$  of  $\phi(\cdot)$  at  $\mu$  if  $(-\xi) \in \partial^-(-\phi)(\mu)$ .

It has been proven recently in [32] that the definition of Wasserstein subdifferential involving a supremum taken over the set of optimal transport plans is equivalent to the usual one introduced in [9] which involves an infimum. This allows for the elaboration of a convenient notion of differentiability in Wasserstein spaces as detailed below.

**Definition 2.6** (Differentiable functionals in  $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ ). *A functional  $\phi : \mathcal{P}_2(\mathbb{R}^d) \mapsto \mathbb{R}$  is said to be Wasserstein-differentiable at some  $\mu \in D(\phi)$  if  $\partial^- \phi(\mu) \cap \partial^+ \phi(\mu) \neq \emptyset$ . In this case, there exists a unique elements  $\nabla_\mu \phi(\mu) \in \partial^- \phi(\mu) \cap \partial^+ \phi(\mu)$  called the Wasserstein gradient of  $\phi(\cdot)$  at  $\mu$ , which satisfies*

$$\phi(\nu) - \phi(\mu) = \int_{\mathbb{R}^{2d}} \langle \nabla_\mu \phi(\mu)(x), y - x \rangle d\gamma(x, y) + o(W_2(\mu, \nu)), \quad (2.4)$$

for any  $\nu \in \mathcal{P}_2(\mathbb{R}^d)$  and  $\gamma \in \Gamma_o(\mu, \nu)$ .

We already listed in the subsection 1.2 of the introduction a series of commonly encountered functionals which are differentiable in the sense of Definition 2.6, and provided their Wasserstein gradient. We conclude these recalls by stating in Proposition 2.4 below a chain rule formula for Wasserstein differentiable functionals along suitable multi-dimensional families of measures.

**Proposition 2.4** (Chain rule along multidimensional families generated by smooth vector fields). *Let  $\mu \in \mathcal{P}_c(\mathbb{R}^d)$ ,  $N \geq 1$  and  $K \subset \mathbb{R}^d$  be a compact set. Let  $\phi : \mathcal{P}(K) \rightarrow \mathbb{R}$  be a functional satisfying hypotheses **(H4)** of Theorem 1.1. Let  $\mathcal{G} \in C^0([-\epsilon, \epsilon]^N \times K, \mathbb{R}^d)$  be such that*

- (i)  $\mathcal{G}(0, \cdot) = I_d$  and  $e \mapsto \mathcal{G}(e, x)$  is Fréchet-differentiable at  $e = 0$  uniformly with respect to  $x \in K$ .
- (ii)  $\text{supp}(\mathcal{G}(e, \cdot) \# \mu) \subset K$  for all  $e \in [-\epsilon, \epsilon]^N$ .
- (iii) The directional derivative map

$$\mathcal{F}_\sigma : x \in K \mapsto D_e \mathcal{G}(0, x) \sigma = \sum_{k=1}^N \sigma_k \mathcal{F}_k(x)$$

is continuous for all  $\sigma \in [-\epsilon, \epsilon]^N$ .

Then, the map  $e \in [-\epsilon, \epsilon]^N \mapsto \phi(\mathcal{G}(e, \cdot) \# \mu)$  is Fréchet-differentiable at  $e = 0$  and

$$D_e(\phi(\mathcal{G}(e, \cdot) \# \mu))(0)\sigma = \int_{\mathbb{R}^d} \langle \nabla_\mu \phi(\mu)(x), \mathcal{F}_\sigma(x) \rangle d\mu(x) = \sum_{k=1}^N \sigma_k \int_{\mathbb{R}^d} \langle \nabla_\mu \phi(\mu)(x), \mathcal{F}_k(x) \rangle d\mu(x),$$

for any  $\sigma \in [-\epsilon, \epsilon]^N$ .

The proof of this result is essentially the same as that of the classical chainrule result along the minimal subdifferential selection that can be found e.g. in [14, Proposition 7]. However, unlike the simpler 1-dimensional case, the assumption of Wasserstein differentiability is crucial for proving the  $N$ -dimensional chainrule.

### 2.3 The continuity equation with non-local velocities on $\mathbb{R}^d$

In this Section, we introduce the continuity equation with non-local velocities in  $(\mathcal{P}_c(\mathbb{R}^d), W_1)$ . This equation is commonly written as

$$\partial_t \mu(t) + \nabla \cdot (v[\mu(t)](t, \cdot) \mu(t)) = 0, \quad (2.5)$$

where  $t \mapsto \mu(t)$  is a narrowly continuous family of probability measures on  $\mathbb{R}^d$  and  $(t, x) \mapsto v[\mu](t, x)$  is a Borel family of vector fields satisfying the condition

$$\int_0^T \int_{\mathbb{R}^d} |v[\mu(t)](t, x)| d\mu(t)(x) dt < +\infty. \quad (2.6)$$

Equation (2.5) has to be understood in duality with smooth and compactly supported functions, i.e.

$$\int_0^T \int_{\mathbb{R}^d} \left( \partial_t \xi(t, x) + \langle \nabla_x \xi(t, x), v[\mu(t)](t, x) \rangle \right) d\mu(t)(x) dt = 0 \quad (2.7)$$

for all  $\xi \in C_c^\infty([0, T] \times \mathbb{R}^d)$ . This definition can be alternatively written as

$$\frac{d}{dt} \int_{\mathbb{R}^d} \xi(x) d\mu(t)(x) = \int_{\mathbb{R}^d} \langle \nabla \xi(x), v[\mu(t)](t, x) \rangle d\mu(t)(x) \quad (2.8)$$

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

We recall in Theorem 2.1 the classical existence, uniqueness and representation formula for solutions of non-local PDEs. Although these results were first derived in [8], we state here a version explored in [37, 38] which is better suited to our control-theoretic framework.

**Theorem 2.1** (Existence, uniqueness and representation of solutions for (2.5)). *Consider a non-local velocity field  $v[\cdot](\cdot, \cdot)$  satisfying the Cauchy-Lipschitz assumptions **(H3)** of Theorem 1.1. Then, for every initial datum  $\mu^0 \in \mathcal{P}_c(\mathbb{R}^d)$ , the Cauchy problem*

$$\begin{cases} \partial_t \mu(t) + \nabla \cdot (v[\mu(t)](t, \cdot) \mu(t)) = 0 \\ \mu(0) = \mu^0, \end{cases} \quad (2.9)$$

*admits a unique solution  $\mu(\cdot) \in C^0(\mathbb{R}, \mathcal{P}_c(\mathbb{R}^d))$ . If  $\mu^0$  is absolutely continuous with respect to  $\mathcal{L}^d$ , then  $\mu(t)$  is absolutely continuous with respect to  $\mathcal{L}^d$  as well for all times  $t \geq 0$ . Furthermore for every  $T > 0$  and every  $\mu^0, \nu^0 \in \mathcal{P}_c(\mathbb{R}^d)$ , there exists positive constants  $R_T, L_T > 0$  such that*

$$\text{supp}(\mu(t)) \subset \overline{B(0, R_T)} \quad \text{and} \quad W_1(\mu(t), \nu(t)) \leq L_T W_1(\mu^0, \nu^0),$$

*for all times  $t \in [0, T]$  and any solutions  $(\mu(\cdot), \nu(\cdot))$  of (2.9) with initial conditions  $(\mu^0, \nu^0)$ .*

Let  $\mu(\cdot)$  be the unique solution of (2.9) and  $(\Phi_{(0,t)}^v[\mu^0](\cdot))_{t \geq 0}$  be the family of flows of diffeomorphisms generated by the non-autonomous velocity field  $(t, x) \mapsto v[\mu(t)](t, x)$ , i.e.

$$\begin{cases} \partial_t \Phi_{(0,t)}^v[\mu^0](x) = v[\mu(t)](t, \Phi_{(0,t)}^v[\mu^0](x)), \\ \Phi_{(0,0)}^v[\mu^0](x) = x \quad \text{for all } x \text{ in } \mathbb{R}^d. \end{cases} \quad (2.10)$$

Then, the curve  $\mu(\cdot)$  satisfies

$$\mu(t) = \Phi_{(0,t)}^v[\mu^0](\cdot) \# \mu^0, \quad (2.11)$$

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

We recall in the following Proposition a standard result which links the differential in space of the flow of diffeomorphisms of an ODE to the solution of the corresponding linearized Cauchy problem (see e.g. [15]). In our previous work [14], we extended this result to the Wasserstein differential of a flow with respect to its initial measure.

**Proposition 2.5** (Classical and Wasserstein differentials of a flow of diffeomorphisms). *Let  $\mu \in \mathcal{P}(K)$ ,  $v[\cdot](\cdot, \cdot)$  be a non-local velocity field satisfying hypotheses **(H')** and  $(\Phi_{(0,t)}^v[\mu](\cdot))_{t \in [0, T]}$  be its associated flow of diffeomorphisms. Let  $\mathcal{G} \in C^0([-\epsilon, \epsilon]^N \times K, \mathbb{R}^d)$  be a map satisfying the hypotheses of Proposition 2.4. It then holds that*

- (a) *The differential in space  $D_x \Phi_{(s,t)}^v[\mu](x)h$  of the flow  $\Phi_{(s,t)}^v[\mu](\cdot)$  is the unique solution  $w(\cdot, x)$  of the linearized Cauchy problem*

$$\partial_t w(t, x) = D_x v(t, \Phi_{(s,t)}^v[\mu](x)) w(t, x), \quad w(s, x) = h. \quad (2.12)$$

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

- (b) *The map  $e \in [-\epsilon, \epsilon]^N \mapsto \Phi_{(s,t)}^v[\mathcal{G}(e, \cdot) \# \mu](x)$  is Fréchet-differentiable at  $e = 0$  and its differential  $w_\sigma(\cdot, x)$  in an arbitrary direction  $\sigma \in [-\epsilon, \epsilon]^N$  can be expressed as*

$$w_\sigma(t, x) = \sum_{k=1}^N \sigma_k w_k(t, x),$$

where for any  $k \in \{1, \dots, N\}$  the map  $w_k(\cdot, x)$  is the unique solution of the non-local Cauchy problem

$$\begin{cases} \partial_t w_k(t, x) = D_x v(t, \Phi_{(s,t)}^v[\mu](x)) w_k(t, x) \\ \quad + \int_{\mathbb{R}^d} \mathbb{I}^v_{(t, \Phi_{(s,t)}^v[\mu](x))} \left( \Phi_{(s,t)}^v[\mu](y) \right) \left( D_x \Phi_{(s,t)}^v[\mu](y) \mathcal{F}_k(y) + w_k(t, y) \right) d\mu(y) \\ w_k(s, x) = 0. \end{cases} \quad (2.13)$$

Here,  $(t, x, y) \mapsto \mathbb{I}^v_{(t,x)}(y) = (\mathbb{I}^v_{(t,x)}(y))_{1 \leq i \leq d}$  is the matrix-valued map which columns are defined as the Wasserstein gradients of the components  $v^i[\cdot](t, x)$  of the non-local velocity field at  $\mu(t)$ , i.e.

$$\mathbb{I}^v_{(t,x)}(y) = \nabla_\mu \left( v^i[\cdot](t, x) \right) (\mu(t))(y) \quad (2.14)$$

*Proof.* The differential in space result stated in (a) is well known in the literature and its proof can be found e.g. in [15]. By Proposition 2.4, we know that the map  $e \in [-\epsilon, \epsilon]^N \mapsto \Phi_{(s,t)}^v[\mathcal{G}(e, \cdot) \# \mu](x)$  is

Fréchet-differentiable at  $e = 0$ . Therefore, the action of its differential on a given direction  $\sigma \in [-\epsilon, \epsilon]^N$  can be expressed in coordinates using partial derivatives, i.e.

$$w_\sigma(t, x) = \sum_{k=1}^N \sigma_k \partial_{e_k} \left( \Phi_{(s,t)}^v [\mathcal{G}(e, \cdot) \# \mu](x) \right) (0).$$

It has been proven in [14, Proposition 5] that such one-dimensional variations could be characterized as the unique solution of the linearized Cauchy problems (2.13).  $\square$

## 2.4 Non-smooth multiplier rule and differentiable extension of functions

In this Section, we recall some facts of non-smooth analysis as well as a non-smooth Lagrange multiplier rule which is instrumental in the proof of our main result. This multiplier rule is expressed in terms of the so-called *Michel-Penot subdifferential*, see e.g. [36, 34]. In the sequel, we denote by  $(X, \|\cdot\|_X)$  a separable Banach space and by  $X^*$  its topological dual associated with the duality bracket  $\langle \cdot, \cdot \rangle_X$ .

**Definition 2.7** (Michel-Penot subdifferential). *Given a map  $f : X \rightarrow \mathbb{R}$ , the Michel-Penot subdifferential (MP-subdifferential in the sequel) of  $f(\cdot)$  at some  $x \in D(f)$  is defined by*

$$\partial_{\text{MP}} f(x) = \{ \xi \in X^* \text{ s.t. } \langle \xi, h \rangle_X \leq d_{\text{MP}} f(x; h) \text{ for all } h \in X \},$$

where

$$d_{\text{MP}} f(x; h) = \sup_{e \in X} \limsup_{t \downarrow 0} \left[ \frac{f(x + t(e + h)) - f(x + te)}{t} \right].$$

denotes the so-called Michel-Penot derivative of  $f(\cdot)$  at  $x$  in the direction  $h$ . Moreover, if  $f : X \rightarrow \mathbb{R}$  is locally convex around  $x \in X$ , then its Michel-Penot and convex subdifferentials coincide, i.e.  $\partial_{\text{MP}} f(x) = \partial f(x)$ .

The MP-subdifferential - smaller than the Clarke subdifferential - bears the nice property of shrinking to a singleton whenever the functional  $f(\cdot)$  is merely Fréchet-differentiable. It also enjoys a summation rule and a chained-derivative formula for compositions of locally Lipschitz and Fréchet-differentiable maps. We list these properties in the following Proposition.

**Proposition 2.6** (Properties of the Michel-Penot subdifferentials). *Let  $x \in X$ ,  $f, g : X \rightarrow \mathbb{R}$  and  $\mathcal{G} : \mathbb{R}^N \rightarrow \mathbb{R}$ .*

- (a) *If  $f(\cdot)$  is Fréchet-differentiable at  $x$ , then  $\partial_{\text{MP}} f(x) = \{\nabla f(x)\}$ .*
- (b) *If  $d_{\text{MP}} f(x; h) < +\infty$  and  $d_{\text{MP}} g(x; h) < +\infty$  for any  $h \in X$ , it holds that*

$$\partial_{\text{MP}}(f + g)(x) \subseteq \partial_{\text{MP}} f(x) + \partial_{\text{MP}} g(x).$$

- (c) *If  $\mathcal{G}(\cdot)$  is Fréchet-differentiable at  $0 \in \mathbb{R}^N$  and  $f(\cdot)$  is Lipschitz in a neighbourhood of  $\mathcal{G}(0)$ , one has that*

$$d_{\text{MP}}(f \circ \mathcal{G})(0; \sigma) = \langle \xi, D\mathcal{G}(0)\sigma \rangle_X, \tag{2.15}$$

*for some  $\xi \in \partial_{\text{MP}} f(\mathcal{G}(0))$ . In other words,  $\partial_{\text{MP}}(f \circ g)(0) \subseteq Dg(0)^* \partial_{\text{MP}} f(\mathcal{G}(0))$ .*

These propositions can be verified easily by computing explicitly the Michel-Penot derivatives of the corresponding maps and using the definition of the set  $\partial_{\text{MP}}(\bullet)$ , see e.g. [40]. Another useful feature of this notion of subdifferential is that it allows to write Lagrange multiplier rules for locally Lipschitz functions. This family of optimality conditions was initially derived in [34] and refined in [40] where the author extended the result to the class of so-called *calm* functions.

**Definition 2.8** (Calm functions). A map  $f : X \rightarrow \mathbb{R}$  is said to be calm at  $x \in X$  provided that

(i) For any  $\delta \in X$  with  $\|\delta\|_X$  sufficiently small, there exists a constant  $L > 0$  such that

$$\|f(x + \delta) - f(x)\|_X \leq L \|\delta\|_X.$$

(ii)  $d_{\text{MP}}f(x; h) < +\infty$  for any  $h \in X$ .

**Theorem 2.2** (Multiplier rule for Michel-Penot subdifferentials). Let  $f_0, \dots, f_n, g_1, \dots, g_m : X \rightarrow \mathbb{R}$  and  $\Omega \subset X$  be a closed and convex set. Suppose that  $x_*$  is a local solution of the non-linear optimization problem

$$\begin{cases} \min_{x \in \Omega} [f_0(x)] \\ \text{s.t.} \begin{cases} f_i(x) \leq 0 & \text{for all } i \in \{1, \dots, n\}, \\ g_j(x) = 0 & \text{for all } j \in \{1, \dots, m\}, \end{cases} \end{cases}$$

and that the maps  $f_0(\cdot), \dots, f_n(\cdot), g_1(\cdot), \dots, g_m(\cdot)$  are calm at  $x_*$ . Then, there exists Lagrange multipliers  $(\lambda_0, \dots, \lambda_n, \eta_1, \dots, \eta_m) \in \{0, 1\} \times \mathbb{R}_+^n \times \mathbb{R}_m$  such that the following Stationarity (S), Non-Triviality (NT) and Complementary-Slackness (CS) conditions hold

$$\begin{cases} 0 \in \partial_{\text{MP}} \left( \lambda_0 f_0(x) + \sum_{i=1}^n \lambda_i f_i(x) + \sum_{j=1}^m \eta_j g_j(x) \right) (x_*) + \mathcal{N}(\Omega, x_*), & \text{(S)} \\ \lambda_0 + \sum_{i=1}^n \lambda_i + \sum_{j=1}^m |\eta_j| = 1, & \text{(NT)} \\ \lambda_i f_i(x_*) = 0 \text{ for all } i \in \{1, \dots, n\}, & \text{(CS)} \end{cases}$$

where  $\mathcal{N}(\Omega, x_*)$  denotes the normal cone of convex analysis to  $\Omega$  at  $x_*$ .

We end this introductory section by stating a Lusin-type Lemma for vector valued functions and a derivative-preserving continuous extension result that will both prove useful in the sequel.

**Lemma 2.1** (Pointwise convergence and restriction). Let  $f : [0, T] \rightarrow X$  be an  $L^1$ -function in the sense of Bochner (see e.g. [22]) and  $\mathcal{T}$  be any subset of  $[0, T]$  with full Lebesgue measure. Then, there exists  $\mathcal{A}, \mathcal{M} \subset \mathcal{T}$  respectively with null and full Lebesgue measure satisfying the property that for any  $\tau \in \mathcal{M}$ , there exists  $(\tau_k) \subset \mathcal{A}$  such that

$$\tau_k \xrightarrow[k \rightarrow +\infty]{} \tau \text{ and } \|g(\tau) - g(\tau_k)\|_X \xrightarrow[k \rightarrow +\infty]{} 0.$$

*Proof.* This result is a consequence of Lusin's Theorem applied to vector valued maps (see e.g. [22]).  $\square$

**Lemma 2.2** (A continuous extension preserving the derivative). Let  $\epsilon > 0$  and  $f : [0, \epsilon]^N \rightarrow X$  be a continuous map differentiable at 0 relatively to  $\mathbb{R}_+^N$ . Then, there exists a continuous extension  $\tilde{f} : [-\bar{\epsilon}_N, \bar{\epsilon}_N]^N \rightarrow X$  of  $f(\cdot)$  such that  $D_e \tilde{f}(0) = D_e f(0)$ .

*Proof.* We adapt here a simple proof that can be found e.g. in [40]. Define the map

$$g : e \in \mathbb{R}_+^N \mapsto \frac{1}{|e|} \left( f(e) - f(0) - D_e f(0)e \right).$$

By definition,  $g(\cdot)$  is continuous over  $\mathbb{R}_+^N \setminus \{0\}$  and can be extended to  $\mathbb{R}_+^N$  by imposing that  $g(0) = 0$ . Invoking Dugundji's extension theorem (see [23]), we can define a continuous extension  $\tilde{g}(\cdot)$  of  $g(\cdot)$  on the whole of  $\mathbb{R}^N$ .

We now define the auxiliary map  $\tilde{f} : e \in \mathbb{R}^N \mapsto f(0) + D_e f(0)e + |e|\tilde{g}(e)$ . By construction,  $\tilde{f}(\cdot)$  is continuous and coincides with  $f(\cdot)$  over  $\mathbb{R}_+^N$ . Moreover, one has for any  $e \in \mathbb{R}^N$  that

$$\tilde{f}(e) - \tilde{f}(0) = D_e f(0)e + |e|\tilde{g}(e) = D_e f(0)e + o(|e|)$$

by continuity of  $\tilde{g}(\cdot)$  at 0. Therefore, the extended map  $\tilde{f}$  is differentiable at 0 with  $D_e \tilde{f}(0) = D_e f(0)$ .  $\square$

### 3 Proof of the main result

In this Section, we prove the general constrained maximum principle stated in Theorem 1.1. We divide the proof of this result into 4 steps. In Step 1, we introduce the concept of *packages of needle-like variations* of an optimal control and compute the corresponding perturbations induced on the optimal trajectory. In Step 2, we apply Theorem 2.2 to write a Lagrange multiplier rule for the finite-dimensional optimization problem written on the length of the needle variations. We introduce in Step 3 a suitable notion of costate allowing to propagate this optimality condition backward in time, yielding the PMP with a relaxed maximization condition restricted to a countable subset of needle parameters. The full result is then recovered in Step 4 through a limiting procedure combined with density and approximation arguments.

#### Step 1 : Packages of needle-like variations :

We start by considering an optimal pair control-trajectory  $(u^*(\cdot), \mu^*(\cdot)) \in \mathcal{U} \times \text{Lip}([0, T], \overline{B(0, R_T)})$  where  $R_T > 0$  is given by Theorem 2.1. Let  $\mathcal{T} \subset [0, T]$  be the set of Lebesgue points of  $t \mapsto (u^*(t, \cdot), L(t, \mu^*(t), \cdot)) \in \mathbb{R}^{d+1}$  in the sense of Bochner's integral (see e.g. [22, Theorem 9]). This set has full Lebesgue measure in  $[0, T]$ , and Lemma 2.1 implies the existence of two subsets  $\mathcal{A}, \mathcal{M} \subset \mathcal{T}$  having respectively null and full Lebesgue measure such that for any  $\tau \in \mathcal{M}$ , there exists  $(\tau_k) \subset \mathcal{A}$  converging towards  $\tau$  and such that

$$\|u^*(\tau, \cdot) - u^*(\tau_k, \cdot)\|_{C^0(\mathbb{R}^d, \mathbb{R}^d)} \xrightarrow[k \rightarrow +\infty]{} 0, \quad \|L(\tau, \mu^*(\tau), \cdot) - L(\tau_k, \mu^*(\tau_k), \cdot)\|_{C^0(\mathbb{R}^d, \mathbb{R}^d)} \xrightarrow[k \rightarrow +\infty]{} 0.$$

**Definition 3.1** (Package of needle-like variations). *Let  $(u^*(\cdot), \mu^*(\cdot)) \in \mathcal{U} \times \text{Lip}([0, T], \mathcal{P}_c(\mathbb{R}^d))$  be an optimal pair control-trajectory and  $U_D$  be a countable and dense subset of  $U$ . Given  $N \geq 1$ ,  $\{(\omega_k, \tau_k)\}_{k=1}^N \subset U_D \times \mathcal{A}$  and  $e = (e_1, \dots, e_N) \in [0, \bar{\epsilon}_N]^N$  such that  $[\tau_i - e_i, \tau_i] \cap [\tau_j - e_j, \tau_j] = \emptyset$  for all distinct pairs  $i, j \in \{1, \dots, N\}$ , we define the  $N$ -package of needle-like variations of  $u^*(\cdot, \cdot)$  by*

$$\tilde{u}_e^N \equiv \tilde{u}_e : t \mapsto \begin{cases} \omega_k & \text{if } t \in [\tau_k - e_k, \tau_k], \\ u^*(t) & \text{otherwise.} \end{cases} \quad (3.1)$$

We denote by  $t \mapsto \tilde{\mu}_e(\cdot)$  the corresponding perturbed trajectory, i.e. the solution of (2.9) associated with  $\tilde{u}_e(\cdot, \cdot)$ .

This class of variations is known in the literature of control theory to generate admissible perturbations of the optimal control without any assumption on the structure of the control set  $\mathcal{U}$ , while allowing for an explicit and tractable computation of the relationship between the optimal and perturbed states (see e.g. [15]).

In the following Lemma, we make use the geometric structure of solutions to non-local transport equations presented in Theorem 2.1 to express  $\tilde{\mu}_e(t)$  as a function of  $\mu^*(t)$  for all times  $t \in [0, T]$ .

**Lemma 3.1** (First-order perturbation induced by a package of needle-like variations in the non-local case). *There exists a family of maps  $(\mathcal{G}_t^N(\cdot, \cdot))_{t \in [0, T]} \subset C^0([-\bar{\epsilon}_N, \bar{\epsilon}_N]^N \times \mathbb{R}^d, \mathbb{R}^d)$  such that*

(i) *For all  $(t, e) \in [0, T] \times [0, \bar{\epsilon}_N]^N$ , the perturbed measures  $\tilde{\mu}_e(t)$  satisfy*

$$\tilde{\mu}_e(t) = \mathcal{G}_t^N(e, \cdot) \# \mu^*(t). \quad (3.2)$$

(ii) *For all  $(t, e) \in [0, T] \times [0, \bar{\epsilon}_N]^N$ , the maps  $\mathcal{G}_t^N(e, \cdot)$  are  $C^1$ -diffeomorphisms over  $\overline{B(0, R_T)}$ .*

(iii) *There exists a constant  $R_T^\Phi > 0$  depending on  $R_T, L_U$  such that for all  $(t, e) \in [0, T] \times [0, \bar{\epsilon}_N]^N$  one has  $\text{supp}(\mathcal{G}_t^N(e, \cdot) \# \mu^*(T)) \subset \overline{B(0, R_T^\Phi)}$*

(iv) The map  $e \in [-\bar{\epsilon}_N, \bar{\epsilon}_N]^N \mapsto \mathcal{G}_t^N(e, \cdot)$  is Fréchet-differentiable at  $e = 0$  with respect to the  $C^0(B(0, R_T), \mathbb{R}^d)$ -norm uniformly with respect to  $t \in [0, T]$ . The corresponding Taylor expansion can be written explicitly as

$$\mathcal{G}_t^N(e, \cdot) = I_d + \sum_{k=1}^{\iota(t)} e_k \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{u^*}[\mu^*(t)](\cdot) + o(|e|), \quad (3.3)$$

where  $\iota(t) \in \{1, \dots, N\}$  is the biggest index such that  $\tau_{\iota(t)} \leq t \leq \tau_{\iota(t)+1} - e_{\iota(t)+1}$ . For all  $x \in \overline{B(0, R_T)}$  and  $k \in \{1, \dots, N\}$ , the map  $t \in [\tau_k, T] \mapsto \mathcal{F}_t^{\omega_k, \tau_k}(x)$  is the unique solution of the non-local Cauchy problem

$$\begin{cases} \partial_t \mathcal{F}_t^{\omega_k, \tau_k}(x) = \left[ D_x u^* \left( t, \Phi_{(\tau_k, t)}^{u^*}(x) \right) + D_x v[\mu^*(t)] \left( t, \Phi_{(\tau_k, t)}^{u^*}(x) \right) \right] \mathcal{F}_t^{\omega_k, \tau_k}(x) \\ \quad + \int_{\mathbb{R}^d} \Pi^v \left( t, \Phi_{(\tau_k, t)}^{v, u^*}(x) \right) \left( \Phi_{(\tau_k, t)}^{v, u^*}(y) \right) \mathcal{F}_t^{\omega_k, \tau_k}(y) d\mu^*(\tau_k)(y), \\ \mathcal{F}_{\tau_k}^{\omega_k, \tau_k}(x) = \omega_k(x) - u^*(\tau_k, x). \end{cases} \quad (3.4)$$

*Proof.* The proof of this result is similar to that of [14, Lemma 5], with some extra technicalities arising from the induction argument performed on the non-local terms. By definition of a package of needle-like variations, the perturbed controls  $\tilde{u}_e(\cdot, \cdot)$  generate well-defined flows of diffeomorphisms  $(\Phi_{(0, t)}^{\tilde{u}_e, v}[\mu^0](\cdot))_{t \in [0, T]}$  such that

$$\tilde{\mu}_e(t) = \Phi_{(0, t)}^{\tilde{u}_e, v}[\mu^0] \circ \Phi_{(t, 0)}^{u^*, v}[\mu^*(t)](\cdot) \# \mu^*(t),$$

so that items (i), (ii) and (iii) hold for any  $e \in [0, \epsilon_N]^N$  with  $\mathcal{G}_t^N(e, \cdot) = \Phi_{(0, t)}^{\tilde{u}_e, v} \circ \Phi_{(t, 0)}^{u^*, v}(\cdot)$ .

We focus our attention on the proof by induction of (iv). Let  $t \in [0, T]$  be such that  $\iota(t) = 1$ . By (3.1), one has that

$$\tilde{\mu}_e(t) = \Phi_{(\tau_1, t)}^{v, u^*}[\tilde{\mu}_e(\tau_1)] \circ \Phi_{(\tau_1 - e_1, \tau_1)}^{v, \omega_1}[\mu^*(\tau_1 - e_1)] \circ \Phi_{(\tau_1, \tau_1 - e_1)}^{v, u^*}[\mu^*(\tau_1)] \circ \Phi_{(t, \tau_1)}^{v, u^*}[\mu^*(t)](\cdot) \# \mu^*(t).$$

Invoking Lebesgue's Differentiation Theorem along with the continuity of  $e \mapsto v[\tilde{\mu}_e(t)](t, \cdot)$  in the  $C^0$ -norm topology, it holds that

$$\begin{aligned} & \Phi_{(\tau_1 - e_1, \tau_1)}^{v, \omega_1}[\mu^*(\tau_1 - e_1)](x) \\ &= x + \int_{\tau_1 - e_1}^{\tau_1} \left( v[\tilde{\mu}_e(t)] \left( t, \Phi_{(\tau_1 - e_1, t)}^{v, \omega_1}[\mu^*(\tau_1 - e_1)](x) \right) + \omega_1 \left( \Phi_{(\tau_1 - e_1, t)}^{v, \omega_1}[\mu^*(\tau_1 - e_1)](x) \right) \right) dt \\ &= x + e_1 \left( v[\mu^*(\tau_1)](\tau_1, x) \right) + \omega_1(x) + o(e_1) \end{aligned}$$

as well as

$$\begin{aligned} \Phi_{(\tau_1, \tau_1 - e_1)}^{v, u^*}[\mu^*(\tau_1)](x) &= x - \int_{\tau_1 - e_1}^{\tau_1} \left( v[\tilde{\mu}_e(t)] \left( t, \Phi_{(\tau_1, t)}^{v, u^*}[\mu^*(\tau_1)](x) \right) + u^* \left( t, \Phi_{(\tau_1, t)}^{v, u^*}[\mu^*(\tau_1)](x) \right) \right) dt \\ &= x - e_1 \left( v[\mu^*(\tau_1)](\tau_1, x) \right) + u^*(\tau_1, x) + o(e_1). \end{aligned}$$

Chaining these two expansions, we obtain that

$$\begin{aligned} \tilde{\mu}_e(\tau_1) &= \Phi_{(\tau_1 - e_1, \tau_1)}^{v, \omega_1}[\mu^*(\tau_1 - e_1)] \circ \Phi_{(\tau_1, \tau_1 - e_1)}^{v, u^*}[\mu^*(\tau_1)](\cdot) \# \mu^*(\tau_1) \\ &= (I_d + e_1 [\omega_1(\cdot) - u^*(\tau_1, \cdot)] + o(e_1)) \# \mu^*(\tau_1). \end{aligned}$$

We can now proceed to compute the induced first order expansion on the non-local flows  $\Phi_{(\tau_1, t)}^{v, u^*}[\tilde{\mu}_e(\tau_1)](\cdot)$  as follows

$$\begin{aligned} & \Phi_{(\tau_1, t)}^{v, u^*}[\tilde{\mu}_e(\tau_1)]\left(x + e_1 [\omega_1(x) - u^*(\tau_1, x)] + o(e_1)\right) \\ &= \Phi_{(\tau_1, t)}^{v, u^*}[\tilde{\mu}_e(\tau_1)](x) + e_1 D_x \Phi_{(\tau_1, t)}^{v, u^*}[\tilde{\mu}_e(\tau_1)](x) \cdot [\omega_1(x) - u^*(\tau_1, x)] + o(e_1) \\ &= \Phi_{(\tau_1, t)}^{v, u^*}[\mu^*(\tau_1)](x) + e_1 \left( D_x \Phi_{(\tau_1, t)}^{v, u^*}[\mu^*(\tau_1)](x) \cdot [\omega_1(x) - u^*(\tau_1, x)] + w_1(t, x) \right) + o(e_1) \end{aligned}$$

where  $w_1(\cdot, \cdot)$  is defined as in Proposition 2.5-(b), and where we used the fact that the  $e \mapsto D_x \Phi_{(\tau_1, \cdot)}^{v, u^*}[\tilde{\mu}_e(\cdot)](\cdot)$  is continuous as a consequence of hypothesis **(H3)**. Introducing for all times  $t \in [\tau_1, T]$  the map

$$\mathcal{F}_t^{\omega_1, \tau_1} : x \in \overline{B(0, R_T)} \mapsto D_x \Phi_{(\tau_1, t)}^{v, u^*}[\mu^*(\tau_1)](x) [\omega_1(x) - u^*(\tau_1, x)] + w_1(t, x)$$

and using again the statements of Proposition 2.5, we have that both (3.3) and (3.4) hold for any  $e_1 \in [0, \bar{\epsilon}_N]$  and all times  $t \in [0, T]^N$  such that  $\iota(t) = 1$ .

Let us now assume that (3.3) and (3.4) hold for all times  $t \in [0, T]$  such that  $\iota(t) = k - 1$ , i.e.

$$\tilde{\mu}_e(t) = \mathcal{G}_t^N(e, \cdot) \# \mu^*(t) = \left( I_d + \sum_{l=1}^{k-1} e_l \mathcal{F}_t^{\omega_l, \tau_l} \circ \Phi_{(t, \tau_l)}^{v, u^*}[\mu^*(t)](\cdot) + o(|e|) \right) \# \mu^*(t), \quad (3.5)$$

for  $e \in [0, \bar{\epsilon}_N]^N$ . Take  $t \in [0, T]$  such that  $\iota(t) = k$ . By Definition of an  $N$ -package of needle-like variations (3.1), we have that

$$\tilde{\mu}_e(\tau_k) = \Phi_{(\tau_k - e_k, \tau_k)}^{v, \omega_k}[\tilde{\mu}_e(\tau_k - e_k)] \circ \Phi_{(\tau_k, \tau_k - e_k)}^{v, u^*}[\tilde{\mu}_e(\tau_k)] \circ \Phi_{(\tau_{k-1}, \tau_k)}^{v, u^*}[\tilde{\mu}_e(\tau_{k-1})](\cdot) \# \tilde{\mu}_e(\tau_{k-1}).$$

As in the initialization step, we can write using Lebesgue's Differentiation Theorem that

$$\Phi_{(\tau_k - e_k, \tau_k)}^{v, \omega_k}[\tilde{\mu}_e(\tau_k - e_k)] \circ \Phi_{(\tau_k, \tau_k - e_k)}^{v, u^*}[\tilde{\mu}_e(\tau_k)](x) = x + e_k [\omega_k(x) - u^*(\tau_k, x)] + o(e_k). \quad (3.6)$$

Furthermore, invoking the induction hypothesis (3.5) and the results of Proposition 2.4, we obtain that

$$\begin{aligned} & \Phi_{(\tau_{k-1}, \tau_k)}^{v, u^*}[\tilde{\mu}_e(\tau_{k-1})] \left( x + \sum_{l=1}^{k-1} \left( e_l \mathcal{F}_{\tau_{k-1}}^{\omega_l, \tau_l} \circ \Phi_{(\tau_{k-1}, \tau_l)}^{v, u^*}(x) + o(e_l) \right) \right) \\ &= \Phi_{(\tau_{k-1}, \tau_k)}^{v, u^*}[\tilde{\mu}_e(\tau_{k-1})](x) + \sum_{l=1}^{k-1} \left( e_l D_x \Phi_{(\tau_{k-1}, \tau_k)}^{v, u^*}[\tilde{\mu}_e(\tau_{k-1})](x) \mathcal{F}_{\tau_{k-1}}^{\omega_l, \tau_l} \circ \Phi_{(\tau_{k-1}, \tau_l)}^{v, u^*}(x) + o(e_l) \right) \\ &= \Phi_{(\tau_{k-1}, \tau_k)}^{v, u^*}[\mu^*(\tau_{k-1})](x) + \sum_{l=1}^{k-1} e_l \left( D_x \Phi_{(\tau_{k-1}, \tau_k)}^{v, u^*}[\mu^*(\tau_{k-1})](x) \mathcal{F}_{\tau_{k-1}}^{\omega_l, \tau_l} \circ \Phi_{(\tau_{k-1}, \tau_l)}^{v, u^*}(x) + w_l(\tau_k, x) + o(e_l) \right) \end{aligned} \quad (3.7)$$

where the maps  $(w_l(\cdot, \cdot))_{1 \leq l \leq k-1}$  are defined as in Proposition 2.5-(b) with  $\mathcal{F}_l(\cdot) \equiv \mathcal{F}_{\tau_{k-1}}^{\omega_l, \tau_l} \circ \Phi_{(\tau_{k-1}, \tau_l)}^{v, u^*}[\mu^*(\tau_k)](\cdot)$ . Plugging together equations (2.12) and (2.13) of Proposition 2.5, one can see that the maps

$$t \in [\tau_{k-1}, \tau_k] \mapsto D_x \Phi_{(\tau_{k-1}, t)}^{v, u^*}[\mu^*(\tau_{k-1})] \left( \Phi_{(\tau_l, \tau_{k-1})}^{v, u^*}(x) \right) \mathcal{F}_{\tau_{k-1}}^{\omega_l, \tau_l}(x) + w_l \left( t, \Phi_{(\tau_l, \tau_{k-1})}^{v, u^*}(x) \right)$$

are solution of (3.4) on  $[\tau_{k-1}, \tau_k]$  with initial condition  $\mathcal{F}_{\tau_{k-1}}^{\omega_l, \tau_l}(\cdot)$  for any  $l \in \{1, \dots, k-1\}$ . By Cauchy-Lipschitz uniqueness, we can therefore extend the definition of the maps  $t \in [\tau_l, \tau_{k-1}] \mapsto \mathcal{F}_t^{\omega_l, \tau_l}(x)$  to  $[\tau_l, \tau_k]$  for any  $l \in \{1, \dots, k-1\}$ .

Chaining the expansions (3.6) and (3.7) along with our previous extension argument, we obtain that both (3.3) and (3.4) hold up to time  $\tau_k$ , i.e.

$$\mathcal{G}_{\tau_k}^N(e, \cdot) = I_d + \sum_{l=1}^k e_l \mathcal{F}_{\tau_k}^{\omega_l, \tau_l} \circ \Phi_{(\tau_k, \tau_l)}^{v, u^*}[\mu^*(\tau_k)](\cdot) + o(|e|)$$

for any  $e \in [0, \bar{\epsilon}_N]^N$ . Performing yet another coupled Taylor expansion of the same form on the expression

$$\tilde{\mu}_e(t) = \Phi_{(\tau_k, t)}^{v, u^*}[\tilde{\mu}_e(\tau_k)](\cdot) \# \tilde{\mu}_e(\tau_k),$$

and invoking the same prolongation argument yields the full induction step for all times  $t \in [0, T]$  such that  $\iota(t) = k$ . Hence, we have proven that item (iv) holds for all  $e \in [0, \bar{\epsilon}_N]^N$ . Using Lemma 2.2, we can now extend the map  $e \in [0, \bar{\epsilon}_N]^N \mapsto \mathcal{G}_t^N(e, \cdot)$  to the whole of  $[-\bar{\epsilon}_N, \bar{\epsilon}_N]^N$  in a continuous and bounded way, uniformly with respect to  $t \in [0, T]$ , while preserving its differential at  $e = 0$ . This achieves the proof of Lemma 3.1.  $\square$

In the sequel, we drop the explicit dependence of the flow maps on their starting measures and adopt the simplified notation  $\Phi_{(s, t)}^{v, u^*}[\mu(s)](x) = \Phi_{(s, t)}^{v, u^*}(x)$  for clarity and conciseness.

## Step 2 : First-order optimality condition

In Lemma 3.1, we derived the analytical expression of the first-order perturbation induced by a  $N$ -package of needle-like variations on the solution of a controlled non-local continuity equation. By the very definition of an  $N$ -package of needle-like variations, we know that the finite-dimensional optimization problem

$$(\mathcal{P}^N) \quad \begin{cases} \min_{e \in [0, \bar{\epsilon}_N]^N} \left[ \int_0^T L(t, \tilde{\mu}_e(t), \tilde{u}_e(t)) dt + \varphi(\tilde{\mu}_e(T)) \right] \\ \text{s.t.} \quad \begin{cases} \Psi^E(\tilde{\mu}_e(T)) = 0, \\ \Psi^I(\tilde{\mu}_e(T)) \leq 0, \\ \max_{t \in [0, T]} \Lambda(t, \tilde{\mu}_e(t)) \leq 0, \end{cases} \end{cases}$$

admits  $e = 0$  as a local solution.

In the following Lemma, we check that the functionals involved in  $(\mathcal{P}^N)$  meet the requirements of the Lagrange multiplier rule stated in Theorem 2.2. We also compute their first-order variation induced by the package of needle-like variations at  $e = 0$ .

**Lemma 3.2** (Differentiability and calmness of the functionals involved in  $(\mathcal{P}^N)$ ). *The maps  $e \in [-\bar{\epsilon}_N, \bar{\epsilon}_N]^N \mapsto \varphi(\tilde{\mu}_T(e))$ ,  $\Psi^E(\tilde{\mu}_T(e))$ ,  $\Psi^I(\tilde{\mu}_T(e))$  and  $e \in [-\bar{\epsilon}_N, \bar{\epsilon}_N]^N \mapsto \int_0^T L(t, \tilde{\mu}_e(t), \tilde{u}_e(t)) dt$  are calm and differentiable at  $e = 0$ . Their Fréchet derivative in a direction  $\sigma \in [0, \bar{\epsilon}_N]^N$  are respectively given by*

$$\begin{cases} \nabla_e(\varphi(\tilde{\mu}_T(e)))|_{e=0}(\sigma) = \sum_{k=1}^N \sigma_k \int_{\tau_k}^T \int_{\mathbb{R}^d} \langle \nabla_{\mu} \varphi(\mu^*(T))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(t)(x) dt, \\ D_e(\Psi_i^I(\tilde{\mu}_T(e)))|_{e=0}(\sigma) = \sum_{k=1}^N \sigma_k \int_{\tau_k}^T \int_{\mathbb{R}^d} \langle \nabla_{\mu} \Psi_i^I(\mu^*(T))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(t)(x) dt, \\ D_e(\Psi^E(\tilde{\mu}_T(e)))|_{e=0}(\sigma) = \sum_{k=1}^N \sigma_k \int_{\tau_k}^T \int_{\mathbb{R}^d} \langle \nabla_{\mu} \Psi_j^E(\mu^*(T))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(t)(x) dt, \end{cases} \quad (3.8)$$

for any  $(i, j) \in \{1, \dots, n_I\} \times \{1, \dots, n_E\}$  and

$$\begin{aligned} \nabla_e \left( \int_0^T L(t, \tilde{\mu}_e(t), \tilde{u}_e(t)) dt \right) \Big|_{e=0}(\sigma) &= \sum_{k=1}^N \sigma_k \int_{\tau_k}^T \int_{\mathbb{R}^d} \langle \nabla_{\mu} L(t, \mu^*(t), u^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(t)(x) dt \\ &\quad + \sum_{k=1}^N \sigma_k \left( L(\tau_k, \mu^*(\tau_k), u^*(\tau_k)) - L(\tau_k, \mu^*(\tau_k), \omega_k) \right). \end{aligned} \quad (3.9)$$

The maps  $e \in [-\bar{\epsilon}_N, \bar{\epsilon}_N]^N \mapsto \max_{t \in [0, T]} \Lambda_l(t, \tilde{\mu}_e(t)), \dots, \max_{t \in [0, T]} \Lambda_r(t, \tilde{\mu}_e(t))$  are calm and locally Lipschitz around  $e = 0$ . Their Michel-Penot derivatives in a direction  $\sigma \in [0, \bar{\epsilon}_N]^N$  are given by

$$d_{\text{MP}} \left( \max_{t \in [0, T]} \Lambda_l(t, \tilde{\mu}_e(t)) \right) (0; \sigma) = \sum_{k=1}^N \sigma_k \int_0^T \int_{\mathbb{R}^d} \left\langle \nabla_{\mu} \Lambda_l(t, \mu^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \right\rangle d\mu^*(t) d\varpi_l^N(t), \quad (3.10)$$

where the measures  $\varpi_l^N \in \mathcal{M}_+([0, T])$  are such that  $\text{supp}(\varpi_l^N) = \{t \in [0, T] \text{ s.t. } \Lambda_l(t, \mu^*(t)) = 0\}$  and  $\|\varpi_l^N\|_{TV} = 1$  for any  $l \in \{1, \dots, r\}$ .

*Proof.* The calmness property of the maps  $e \mapsto \varphi(\tilde{\mu}_T(e))$ ,  $\Psi^E(\tilde{\mu}_T(e))$ ,  $\Psi^I(\tilde{\mu}_T(e))$  and  $e \mapsto \int_0^T L(t, \tilde{\mu}_e(t), \tilde{u}_e(t)) dt$  at  $e = 0$  stems from the fact that they are compositions of Fréchet-differentiable and locally Lipschitz mappings as a by-product of hypotheses **(H4)**, **(H5)** and Lemma 3.1. The differentials of the final cost and boundary constraints can be computed with a direct application of Proposition 2.4.

We split the computation of the first-order variation at  $e = 0$  of the running cost functional into two parts. One can first derive that

$$\begin{aligned} \int_0^T \left( L(t, \tilde{\mu}_e(t), \tilde{u}_e(t)) - L(t, \tilde{\mu}_e(t), u^*(t)) \right) dt &= \sum_{k=1}^N \int_{\tau_k - e_k}^{\tau_k} \left( L(t, \tilde{\mu}_e(t), \omega_k) - L(t, \tilde{\mu}_e(t), u^*(t)) \right) dt \\ &= \sum_{k=1}^N e_k \left( L(\tau_k, \tilde{\mu}_e(\tau_k), \omega_k) - L(\tau_k, \tilde{\mu}_e(\tau_k), u^*(\tau_k)) \right) + o(|e|) \\ &= \sum_{k=1}^N e_k \left( L(\tau_k, \mu^*(\tau_k), \omega_k) - L(\tau_k, \mu^*(\tau_k), u^*(\tau_k)) \right) + o(|e|) \end{aligned}$$

by Lebesgue's Differentiation Theorem, and since the maps  $e \mapsto L(\tau_k, \tilde{\mu}_e(\tau_k), u^*(\tau_k))$  and  $e \mapsto L(\tau_k, \tilde{\mu}_e(\tau_k), \omega_k)$  are continuous for all  $k \in \{1, \dots, N\}$ . Furthermore, invoking the Wasserstein chain rule of Proposition 2.4 along with the results of Lemma 3.1, we have for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$  that

$$L(t, \tilde{\mu}_e(t), u^*(t)) - L(t, \mu^*(t), u^*(t)) = \sum_{k=1}^N e_k \int_{\mathbb{R}^d} \left\langle \nabla_{\mu} L(t, \mu^*(t), u^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \right\rangle d\mu^*(t)(x) + o(|e|).$$

Combining these expansions with an application of Lebesgue's Dominated Convergence Theorem yields (3.9)

We now turn our attention to the state constraints functionals. By hypothesis **(H6)** and Proposition 2.4, the maps  $e \mapsto \Lambda_l(t, \tilde{\mu}_e(t))$  are Fréchet-differentiable at  $e = 0$  for any  $l \in \{1, \dots, r\}$ , uniformly with respect to  $t \in [0, T]$ . Moreover, the functional

$$\gamma \in C^0([0, T], \mathbb{R}) \mapsto \|\gamma(\cdot)\|_{C^0([0, T], \mathbb{R})}$$

is locally convex and therefore locally Lipschitz over  $C^0([0, T], \mathbb{R})$ . Hence, the maps  $e \mapsto \max_{t \in [0, T]} \Lambda_l(t, \tilde{\mu}_e(t))$  are calm at  $e = 0$  as compositions of Fréchet-differentiable and locally Lipschitz mappings. By Proposition 2.6-(c), we can compute their Michel-Penot derivatives in a direction  $\sigma \in [0, \bar{\epsilon}_N]^N$  as follows

$$d_{\text{MP}} \left( \max_{t \in [0, T]} \Lambda_l(t, \tilde{\mu}_e(t)) \right) (0; \sigma) = \int_0^T \nabla_e \left( \Lambda_l(t, \tilde{\mu}_e(t)) \right) \Big|_{e=0} (\sigma) d\varpi_l^N(t)$$

where  $\varpi_l^N \in \mathcal{M}_+([0, T])$  belongs to the convex subdifferential of the  $C^0([0, T], \mathbb{R})$ -norm evaluated at  $\Lambda_l(\cdot, \mu^*(\cdot))$ . This subdifferential can be classically characterized (see e.g. [35, Section 4.5.1]) as the set of Borel regular measures such that

$$\|\varpi_l^N\|_{TV} = 1 \text{ and } \text{supp}(\varpi_l^N) = \left\{ t \in [0, T] \text{ s.t. } \Lambda_l(t, \mu^*(t)) = 0 \right\}.$$

Invoking again Proposition 2.4 and Lemma 3.1, we can write for all times  $t \in [0, T]$  the differential of  $e \mapsto \Lambda(t, \tilde{\mu}_e(t))$  at  $e = 0$  evaluated in a direction  $\sigma \in [0, \bar{\epsilon}_N]^N$  as

$$\nabla_e \left( \Lambda_l(t, \tilde{\mu}_e(t)) \right) \Big|_{e=0} (\sigma) = \sum_{k=1}^N \sigma_k \int_{\mathbb{R}^d} \left\langle \nabla_{\mu} \Lambda_l(t, \mu^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \right\rangle d\mu^*(t)(x).$$

Using the measure-theoretic version of Lebesgue's Dominated Convergence Theorem (see e.g. [7, Theorem 1.21]), we conclude that (3.10) holds as well, which ends the proof of Lemma 3.2.  $\square$

Using the results of Lemma 3.2, we can apply the Lagrange multiplier rule of Theorem 2.2 to  $(\mathcal{P}^N)$  and obtain the existence of scalar multipliers  $(\lambda_0^N, \lambda_1^N, \dots, \lambda_n^N, \eta_1^N, \dots, \eta_m^N, \theta_1^N, \dots, \theta_r^N) \in \{0, 1\} \times \mathbb{R}_+^n \times \mathbb{R}^m \times \mathbb{R}_+^r$  such that

$$\left\{ \begin{array}{l} \partial_{\text{MP}} \left( \lambda_0^N \varphi(\tilde{\mu}_e(T)) + \lambda_0^N \int_0^T L(t, \tilde{\mu}_e(t), \tilde{u}_e(t)) dt + \sum_{l=1}^r \theta_l^N \max_{t \in [0, T]} \Lambda_l(t, \tilde{\mu}_e(t)), \right. \\ \left. + \sum_{i=1}^n \lambda_i^N \Psi_i^I(\tilde{\mu}_e(T)) + \sum_{j=1}^m \eta_j^N \Psi_j^E(\tilde{\mu}_e(T)) \right) (0) \in -\mathcal{N}([0, \bar{\epsilon}_N]^N, 0) \\ \lambda_0^N + \sum_{i=1}^n \lambda_i^N + \sum_{j=1}^m |\eta_j^N| + \sum_{l=1}^r \theta_l^N = 1, \\ \theta_l^N \max_{t \in [0, T]} \Lambda_l(t, \mu^*(t)) = 0 \text{ for all } l \in \{1, \dots, r\}, \\ \lambda_i^N \Psi_i^I(\mu^*(T)) = 0 \text{ for all } i \in \{1, \dots, n\}. \end{array} \right. \quad \begin{array}{l} \text{(S)} \\ \text{(NT)} \\ \text{(CS)} \end{array}$$

Since all the functions involved in the subdifferential inclusion (S) are calm, we can use the summation rule of Proposition 2.6-(b) along with the expressions derived in Lemma 3.2 to obtain that

$$\begin{aligned} & \sum_{k=1}^N \sigma_k \int_{\mathbb{R}^d} \left\langle -\nabla_{\mu} \mathcal{S}_N(\mu^*(T))(x), \mathcal{F}_T^{\omega_k, \tau_k} \circ \Phi_{(T, \tau_k)}^{v, u^*}(x) \right\rangle d\mu^*(T)(x) - \sum_{k=1}^N \sigma_k \left( L(\tau_l, \mu^*(\tau_l), u^*(\tau_l)) - L(\tau_l, \mu^*(\tau_l), \omega_l) \right) \\ & - \sum_{k=1}^N \sigma_k \int_{\tau_k}^T \int_{\mathbb{R}^d} \left\langle \nabla_{\mu} L(t, \mu^*(t), u^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \right\rangle d\mu^*(t)(x) dt \\ & - \sum_{k=1}^N \sigma_k \sum_{l=1}^r \theta_l^N \int_{\tau_k}^T \int_{\mathbb{R}^d} \left\langle \nabla_{\mu} \Lambda_l(t, \mu^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \right\rangle d\mu^*(t)(x) d\varpi_l^N(t) \leq 0 \end{aligned}$$

for any direction  $\sigma \in [0, \bar{\epsilon}_N]^N$ . By choosing vectors  $\sigma \in [0, \bar{\epsilon}_N]^N$  which have all their components except one equal to 0, and by introducing the notation  $\varpi_l^N \equiv \theta_l^N \varpi_l^N$ , this family of inequalities can be rewritten as

$$\begin{aligned} & \int_{\mathbb{R}^d} \left\langle -\nabla_{\mu} \mathcal{S}_N(\mu^*(T))(x), \mathcal{F}_T^{\omega_k, \tau_k} \circ \Phi_{(T, \tau_k)}^{v, u^*}(x) \right\rangle d\mu^*(T)(x) - \left( L(\tau_l, \mu^*(\tau_l), u^*(\tau_l)) - L(\tau_l, \mu^*(\tau_l), \omega_l) \right) \\ & - \int_{\tau_k}^T \int_{\mathbb{R}^d} \left\langle \nabla_{\mu} L(t, \mu^*(t), u^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \right\rangle d\mu^*(t)(x) dt \\ & - \sum_{l=1}^r \int_{\tau_k}^T \int_{\mathbb{R}^d} \left\langle \nabla_{\mu} \Lambda_l(t, \mu^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \right\rangle d\mu^*(t)(x) d\varpi_l^N(t) \leq 0 \end{aligned} \quad (3.11)$$

for all  $k \in \{1, \dots, N\}$ , where  $\nabla_{\mu} \mathcal{S}_N(\mu^*(T))(\cdot)$  is defined as in (1.3).

### Step 3 : Backward dynamics and partial Pontryagin maximization condition

The next step of our proof is to introduce a suitable notion of state-costate variable transporting the family of inequalities (3.11) derived at time  $T$  to the base points  $(\tau_1, \dots, \tau_N)$  of the needle-like variations while generating a Hamiltonian dynamical structure. To this end, we build for all  $N \geq 1$  a curve  $\nu_N^*(\cdot) \in \text{Lip}([0, T], \mathcal{P}_c(\mathbb{R}^{2d}))$  solution of the forward-backward system of continuity equations

$$\begin{cases} \partial_t \nu_N^*(t) + \nabla \cdot (\mathcal{V}_N^*[\nu_N^*(t)](t, \cdot, \cdot) \nu_N^*(t)) = 0 & \text{in } [0, T] \times \mathbb{R}^{2d}, \\ \pi_{\#}^1 \nu_N^*(t) = \mu^*(t) & \text{for all } t \in [0, T], \\ \nu^*(T) = (I_d \times (-\nabla_{\mu} \mathcal{S}_N(\mu^*(T)))_{\#} \mu^*(T)), \end{cases} \quad (3.12)$$

where the non-local velocity field  $\mathcal{V}_N^*[\cdot](\cdot, \cdot, \cdot)$  is given for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$  any  $(x, r, \nu) \in \overline{B_{2d}(0, R_T)} \times \mathcal{P}(\overline{B_{2d}(0, R_T)})$  by

$$\mathcal{V}_N^*[\nu](t, x, r) = \begin{pmatrix} v[\pi_{\#}^1 \nu](t, x) + u^*(t, x) \\ \nabla_{\mu} L(t, \pi_{\#}^1 \nu, u^*(t))(x) + \nabla_{\mu} \mathcal{C}(t, \pi_{\#}^1 \nu, \zeta_N^*(t))(x) - \mathbb{I}^v[\nu](t, x) - D_x v[\pi_{\#}^1 \nu](t, x)^{\top} r - D_x u^*(t, x)^{\top} r \end{pmatrix},$$

where we introduced the notations

$$\mathbb{I}^v[\nu](t, x) = \int_{\mathbb{R}^{2d}} \left( \mathbb{I}_{(t, y)}^v(x) \right)^{\top} p \, d\nu(y, p).$$

Notice that the transport equation (3.12) does not satisfy the classical hypotheses of Theorem 2.1. Following a methodology introduced in our previous work [14], it is possible to circumvent this difficulty by building explicitly a solution of (3.12) relying on the *cascade structure* of the equations.

**Lemma 3.3** (Definition and well-posedness of solutions of (3.12)). *Let  $(u^*(\cdot), \mu^*(\cdot))$  be an optimal pair control-trajectory for  $(\mathcal{P}_1)$ . For  $\mu^*(T)$ -almost every  $x \in \mathbb{R}^d$ , we consider the family of backward flows  $(\Psi_{(T, t)}^{x, N}(\cdot))_{t \leq T}$  solution of the non-local Cauchy problems*

$$\begin{cases} \partial_t w_x(t, r) = \nabla_{\mu} L(t, \mu^*(t), u^*(t)) \left( \Phi_{(T, t)}^{v, u^*}(x) \right) + \nabla_{\mu} \mathcal{C}(t, \mu^*(t), \zeta_N^*(t)) \left( \Phi_{(T, t)}^{v, u^*}(x) \right) \\ \quad - D_x u^* \left( t, \Phi_{(T, t)}^{v, u^*}(x) \right)^{\top} w_x(t, r) - D_x v[\mu^*(t)] \left( t, \Phi_{(T, t)}^{v, u^*}(x) \right)^{\top} w_x(t, r) \\ \quad - \int_{\mathbb{R}^{2d}} \mathbb{I}_{(t, \Phi_{(T, t)}^{v, u^*}(y))}^v \left( \Phi_{(T, t)}^{v, u^*}(x) \right)^{\top} w_y(t, r) d(I_d \times (-\nabla_{\mu} \mathcal{S}_N(\mu^*(T)))_{\#} \mu^*(T))(y, p) \\ w_x(T, r) = r. \end{cases} \quad (3.13)$$

and consider the associated curves of measures

$$\sigma_{x, N}^* : t \mapsto (\Psi_{(T, t)}^{x, N}(\cdot))_{\#} \delta_{(-\nabla_{\mu} \mathcal{S}_N(\mu^*(T)))(x)}.$$

Define the map  $\nu_{T, N}^* : t \in [0, T] \mapsto \int \sigma_{x, N}^*(t) d\mu^*(T)(x) \in \mathcal{P}_c(\mathbb{R}^{2d})$ . Then, the curve  $\nu_N^* : t \in [0, T] \mapsto (\Phi_{(T, t)}^{u^*}(\cdot), I_d)_{\#} \nu_{T, N}^*(t)$  is the unique solution of (3.12). Moreover, there exists two constants  $R'_T, L'_T > 0$  such that

$$\text{supp}(\nu_N^*(t)) \subset \overline{B_{2d}(0, R'_T)} \quad \text{and} \quad W_1(\nu_N^*(t), \nu_N^*(s)) \leq L'_T |t - s|$$

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

*Proof.* Let us denote by  $\Omega \subset \mathbb{R}^{2d}$  a compact set such that

$$\bigcup_{N \geq 1} \text{supp}(I_d \times (-\nabla_{\mu} \mathcal{S}_N(\mu^*(T)))_{\#} \mu^*(T)) \subseteq \Omega.$$

Such a set exists since the maps  $(\nabla_{\mu} \mathcal{S}_N(\mu^*(T)))(\cdot)$  are continuous by **(H3)**, **(H4)** and uniformly bounded as a consequence of the non-triviality condition (NT) on the Lagrange multipliers  $(\lambda_0^N, \dots, \lambda_n^N, \eta_1^N, \dots, \eta_m^N)$ .

The existence and uniqueness of the maps  $(t, x, r) \mapsto w_x(t, r)$  solving the family of non-local Cauchy problems (3.13) can be derived from a standard fixed point argument using hypotheses **(H2)**, **(H3)**, **(H5)** and **(H6)**, see e.g. [14, Proposition 5]. In this context, the Banach space under consideration is that of all maps  $f : [0, T] \times \Omega \rightarrow \mathbb{R}^d$  endowed with the norm  $\text{ess-sup}_{\mu^*(T)}(\|w_x(\cdot, \cdot)\|_{C^0([0, T] \times \mathbb{R}^d)})$ , where we recall that  $\text{ess-sup}_{\mu}(f) = \|f\|_{L^\infty(\mu)}$  for any  $\mu$ -measurable and essentially bounded map  $f : \mathbb{R}^d \rightarrow \mathbb{R}_+$ .

By compactness of  $[0, T] \times \Omega$  and uniform boundedness of the right-hand side of (3.13), we have that the maps  $t \mapsto \Psi_{(T,t)}^{x,N}(r)$  are Lipschitz uniformly with respect to  $(x, r) \in \Omega$  and  $N$ . Furthermore, a simple application of Grönwall's Lemma yields that the flow maps  $r \mapsto \Psi_{(T,t)}^{x,N}(r)$  are uniformly Lipschitz and bounded. This allows us to conclude by applying Proposition 2.3 that the curves of measures  $t \mapsto \nu_N^*(t)$  are well defined and that there exists two constants  $R'_T, L'_T > 0$  such that

$$\text{supp}(\nu_N^*(t)) \subset \overline{B_{2d}(0, R'_T)} \quad \text{and} \quad W_1(\nu_N^*(t), \nu_N^*(s)) \leq L'_T |t - s|.$$

In order to prove that  $\nu_N^*(\cdot)$  is indeed a solution of (3.12), take  $\xi \in C_c^\infty(\mathbb{R}^{2d})$  and compute the time derivative

$$\begin{aligned} & \frac{d}{dt} \int_{\mathbb{R}^{2d}} \xi(x, r) d\nu_N^*(t)(x, r) = \frac{d}{dt} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \xi \left( \Phi_{(T,t)}^{v,u^*}(x), r \right) d\sigma_{x,N}^*(t)(r) d\mu^*(T)(x) \\ = & \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left\langle \nabla_x \xi \left( \Phi_{(T,t)}^{v,u^*}(x), r \right), v[\mu^*(t)] \left( t, \Phi_{(T,t)}^{v,u^*}(x) \right) + u^* \left( t, \Phi_{(T,t)}^{v,u^*}(x) \right) \right\rangle d\sigma_{x,N}^*(t)(r) d\mu^*(T)(x) \\ & + \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left\langle \nabla_r \xi \left( \Phi_{(T,t)}^{v,u^*}(x), r \right), \nabla_\mu L(t, \mu^*(t), u^*(t)) \left( \Phi_{(T,t)}^{v,u^*}(x) \right) \right\rangle d\sigma_{x,N}^*(t)(r) d\mu^*(T)(x) \\ & + \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left\langle \nabla_r \xi \left( \Phi_{(T,t)}^{v,u^*}(x), r \right), \nabla_\mu \mathcal{C}(t, \mu^*(t), \zeta_N^*(t)) \left( \Phi_{(T,t)}^{v,u^*}(x) \right) \right\rangle d\sigma_{x,N}^*(t)(r) d\mu^*(T)(x) \\ & - \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left\langle \nabla_r \xi \left( \Phi_{(T,t)}^{v,u^*}(x), r \right), D_x v[\mu^*(t)] \left( t, \Phi_{(T,t)}^{v,u^*}(x) \right)^\top r + D_x u^* \left( t, \Phi_{(T,t)}^{v,u^*}(x) \right)^\top r \right\rangle d\sigma_{x,N}^*(t)(r) d\mu^*(T)(x) \\ & - \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left\langle \nabla_r \xi \left( \Phi_{(T,t)}^{v,u^*}(x), r \right), \mathbb{I}^v[\nu_N^*(t)] \left( t, \Phi_{(T,t)}^{v,u^*}(x) \right) \right\rangle d\sigma_{x,N}^*(t)(r) d\mu^*(T)(x) \end{aligned}$$

where we used the fact that by Fubini's Theorem

$$\begin{aligned} \mathbb{I}^v[\nu_N^*(t)] \left( t, \Phi_{(T,t)}^{v,u^*}(x) \right) &= \int_{\mathbb{R}^{2d}} \mathbb{I}^v \left( t, \Phi_{(T,t)}^{v,u^*}(y) \right) \left( \Phi_{(T,t)}^{v,u^*}(x) \right)^\top p d\nu_{T,N}^*(t)(y, p) \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{I}^v \left( t, \Phi_{(T,t)}^{v,u^*}(y) \right) \left( \Phi_{(T,t)}^{v,u^*}(x) \right)^\top p d\sigma_{y,N}^*(t)(p) d\mu^*(T)(y) \\ &= \int_{\mathbb{R}^{2d}} \mathbb{I}^v \left( t, \Phi_{(T,t)}^{v,u^*}(y) \right) \left( \Phi_{(T,t)}^{v,u^*}(x) \right)^\top \Psi_{(T,t)}^{y,N}(p) d(I_d \times (-\nabla_\mu \mathcal{S}_N(\mu^*(T)))_{\#} \mu^*(T))(y, p). \end{aligned}$$

This can in turn be reformulated into the more concise expression

$$\begin{aligned} \frac{d}{dt} \int_{\mathbb{R}^{2d}} \xi(x, r) d\nu_N^*(t)(x, r) &= \int_{\mathbb{R}^{2d}} \langle \nabla_x \xi(x, r), v[\mu^*(t)](t, x) + u^*(t, x) \rangle d\nu_N^*(t)(x, r) \\ &+ \int_{\mathbb{R}^{2d}} \langle \nabla_r \xi(x, r), \nabla_\mu L(t, \mu^*(t), u^*(t))(x) + \nabla_\mu \mathcal{C}(t, \mu^*(t), \zeta_N^*(t))(x) \rangle d\nu_N^*(t)(x, r) \\ &- \int_{\mathbb{R}^{2d}} \langle \nabla_r \xi(x, r), D_x v[\mu^*(t)](t, x)^\top r + D_x u^*(t, x)^\top r + \mathbb{I}^v[\nu_N^*(t)](t, x) \rangle d\nu_N^*(t)(x, r), \end{aligned}$$

which is precisely the definition of  $\nu_N^*(\cdot)$  being a solution of (3.12).  $\square$

Now that we have built a suitable notion of solution for (3.12), let us prove that  $\nu_N^*(\cdot)$  is such that the PMP holds with a relaxed maximization condition formulated over the collection of needle parameters  $\{(\omega_k, \tau_k)\}_{k=1}^N \subset U_D \times \mathcal{A}$ .

**Lemma 3.4** (A partial Pontryagin Maximum Principle). *For any  $N \geq 1$ , the curve of measures  $\nu_N^*(\cdot)$  introduced in Lemma 3.3 is a solution of the Hamiltonian flow (1.1) associated to the Lagrange multipliers  $(\lambda_0^N, \dots, \lambda_n^N, \dots, \eta_1^N, \dots, \eta_m^N, \varpi_1^N, \dots, \varpi_r^N) \in \{0, 1\} \times \mathbb{R}_+^n \times \mathbb{R}^m \times \mathcal{M}_+([0, T])^r$ . Moreover, the relaxed maximization condition*

$$\begin{aligned} & \mathcal{H}(\tau_k, \nu_N^*(\tau_k), \zeta_N^*(\tau_k), \omega_k) - \sum_{l=1}^r \varpi_l^N(\{\tau_k\}) \int_{\mathbb{R}^d} \langle \nabla_{\mu} \Lambda_l(\tau_k, \mu^*(\tau_k))(x), \omega_k \rangle d\mu^*(\tau_k)(x) \\ & \leq \mathcal{H}(\tau_k, \nu_N^*(\tau_k), \zeta_N^*(\tau_k), u^*(\tau_k)) - \sum_{l=1}^r \varpi_l^N(\{\tau_k\}) \int_{\mathbb{R}^d} \langle \nabla_{\mu} \Lambda_l(\tau_k, \mu^*(\tau_k))(x), u^*(\tau_k, x) \rangle d\mu^*(\tau_k)(x) \end{aligned} \quad (3.14)$$

holds for all  $\{(\omega_k, \tau_k)\}_{k=1}^N \subset U_D \times \mathcal{A}$ .

*Proof.* Using the expression (1.2) of the augmented infinite-dimensional Hamiltonian  $\mathcal{H}(\cdot, \cdot, \cdot, \cdot)$  along with the definition of the Wasserstein gradient (2.4), one can check that

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

for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$  and every  $(x, r) \in \overline{B_{2d}(0, R_T^*)}$ .

For  $k \in \{1, \dots, N\}$ , we introduce the collection of maps  $\mathcal{K}_{\omega_k, \tau_k}^N(\cdot)$  defined for  $\mathcal{L}^1$ -almost all  $t \in [\tau_k, T]$  by

$$\begin{aligned} \mathcal{K}_{\omega_k, \tau_k}^N(t) &= \int_{\mathbb{R}^{2d}} \langle r, \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \rangle d\nu^*(t)(x, r) + [L(\tau_k, \mu^*(\tau_k), u^*(\tau_k)) - L(\tau_k, \mu^*(\tau_k), \omega_k)] \\ &\quad - \int_{\tau_k}^t \int_{\mathbb{R}^d} \langle \nabla_{\mu} L(s, \mu^*(s), u^*(s))(x), \mathcal{F}_s^{\omega_k, \tau_k} \circ \Phi_{(s, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(s)(x) ds \\ &\quad - \sum_{l=1}^r \int_{\tau_k}^t \int_{\mathbb{R}^d} \langle \nabla_{\mu} \Lambda_l(s, \mu^*(s))(x), \mathcal{F}_s^{\omega_k, \tau_k} \circ \Phi_{(s, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(s)(x) d\varpi_l^N(s) \\ &\quad - \sum_{l=1}^r \zeta_{l, N}^*(t) \int_{\mathbb{R}^d} \langle \nabla_{\mu} \Lambda_l(t, \mu^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(t)(x). \end{aligned}$$

$k \in \{1, \dots, N\}$ . By construction, the maps  $\mathcal{K}_{\omega_k, \tau_k}^N(\cdot)$  are in  $BV([0, T], \mathbb{R})$  and therefore admit a distributional derivative in the form of a finite Borel regular measure. A simple computation shows that the non-absolutely continuous parts of their derivatives cancel each other out, since  $d\zeta_{N, l}^*(t) = -\varpi_l^N(t)$  on  $[0, T]$ .

Hence, the maps  $\mathcal{K}_{\omega_k, \tau_k}^N(\cdot)$  are in fact absolutely continuous. One can therefore compute their derivative at  $\mathcal{L}^1$ -almost every  $t \in [\tau_k, T]$  as follows.

$$\begin{aligned} \frac{d}{dt} \mathcal{K}_{\omega_k, \tau_k}^N(t) &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \langle r, \partial_t \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(T, \tau_k)}^{v, u^*}(x) \rangle d\sigma_{x, N}^*(t)(r) d\mu^*(T)(x) \\ &\quad + \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \langle \mathcal{V}_N^*[\nu_N^*(t)](t, x, r), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(T, \tau_k)}^{v, u^*}(x) \rangle d\sigma_{x, N}^*(t)(r) d\mu^*(T)(x) \\ &\quad - \int_{\mathbb{R}^d} \langle \nabla_{\mu} L(t, \mu^*(t), u^*(t)) \left( \Phi_{(T, t)}^{v, u^*}(x) \right), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(T, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(T)(x) \\ &\quad - \sum_{l=1}^r \zeta_{l, N}^*(t) \frac{d}{dt} \left[ \int_{\mathbb{R}^d} \langle \nabla_{\mu} \Lambda_l(t, \mu^*(t)) \left( \Phi_{(T, t)}^{v, u^*}(x) \right), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(T, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(T)(x) \right] \end{aligned} \quad (3.15)$$

The time-derivatives of the summands of the last term can be computed using Proposition 2.4 and the

geometric structure (2.10) of solutions of (2.9) as shown below.

$$\begin{aligned}
& \frac{d}{dt} \left[ \int_{\mathbb{R}^d} \left\langle \nabla_{\mu} \Lambda_l(t, \mu^*(t)) \left( \Phi_{(T,t)}^{v,u^*}(x) \right), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(T, \tau_k)}^{v,u^*}(x) \right\rangle d\mu^*(T)(x) \right] \\
&= \int_{\mathbb{R}^d} \left\langle \nabla_{\mu} \Lambda_l(t, \mu^*(t)) \left( \Phi_{(T,t)}^{v,u^*}(x) \right), \partial_t \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(T, \tau_k)}^{v,u^*}(x) \right\rangle d\mu^*(T)(x) \\
&+ \int_{\mathbb{R}^d} \left\langle \partial_t \nabla_{\mu} \Lambda_l(t, \mu^*(t)) \left( \Phi_{(T,t)}^{v,u^*}(x) \right), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(T, \tau_k)}^{v,u^*}(x) \right\rangle d\mu^*(T)(x) \\
&+ \int_{\mathbb{R}^d} \left\langle D_x \nabla_{\mu} \Lambda_l(t, \mu^*(t)) \left( \Phi_{(T,t)}^{v,u^*}(x) \right) v[\mu^*(t)] \left( t, \Phi_{(T,t)}^{v,u^*}(x) \right), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(T, \tau_k)}^{v,u^*}(x) \right\rangle d\mu^*(T)(x) \\
&+ \int_{\mathbb{R}^d} \left\langle D_x \nabla_{\mu} \Lambda_l(t, \mu^*(t)) \left( \Phi_{(T,t)}^{v,u^*}(x) \right) u^* \left( t, \Phi_{(T,t)}^{v,u^*}(x) \right), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(T, \tau_k)}^{v,u^*}(x) \right\rangle d\mu^*(T)(x) \\
&+ \int_{\mathbb{R}^d} \left\langle \int_{\mathbb{R}^d} \Pi_{\left( t, \Phi_{(T,t)}^{v,u^*}(x) \right)}^{\nabla_{\mu} \Lambda_l} \left( \Phi_{(T,t)}^{v,u^*}(y) \right) v[\mu^*(t)] \left( t, \Phi_{(T,t)}^{v,u^*}(y) \right) d\mu^*(T)(y), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(T, \tau_k)}^{v,u^*}(x) \right\rangle d\mu^*(T)(x) \\
&+ \int_{\mathbb{R}^d} \left\langle \int_{\mathbb{R}^d} \Pi_{\left( t, \Phi_{(T,t)}^{v,u^*}(x) \right)}^{\nabla_{\mu} \Lambda_l} \left( \Phi_{(T,t)}^{v,u^*}(y) \right) u^* \left( t, \Phi_{(T,t)}^{v,u^*}(y) \right) d\mu^*(T)(y), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(T, \tau_k)}^{v,u^*}(x) \right\rangle d\mu^*(T)(x) \\
&= \int_{\mathbb{R}^d} \left\langle \nabla_{\mu} \mathcal{E}_l(t, \mu^*(t)) \left( \Phi_{(T,t)}^{v,u^*}(x) \right), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(T, \tau_k)}^{v,u^*}(x) \right\rangle d\mu^*(T)(x)
\end{aligned}$$

by applying Fubini's Theorem and identifying the analytical expressions of the summands  $\nabla_{\mu} \mathcal{E}_l(t, \mu^*(t))$  derived in (A.1). Plugging this expression into (3.15) along with the characterization of  $\partial_t \mathcal{F}_t^{\omega_k, \tau_k}(\cdot)$  derived in Lemma 3.1, we obtain that

$$\begin{aligned}
\frac{d}{dt} \mathcal{K}_{\omega_k, \tau_k}^N(t) &= \int_{\mathbb{R}^{2d}} \left\langle D_x u^*(t, x)^{\top} r + D_x v[\mu^*(t)](t, x)^{\top} r + \Pi^v[\nu_N^*(t)](t, x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v,u^*}(x) \right\rangle d\nu_N^*(t)(x, r) \\
&+ \int_{\mathbb{R}^{2d}} \left\langle \mathcal{V}_N^*[\nu_N^*(t)](t, x, r), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v,u^*}(x) \right\rangle d\nu_N^*(t)(x) \\
&- \int_{\mathbb{R}^d} \left\langle \nabla_{\mu} L(t, \mu^*(t), u^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v,u^*}(x) \right\rangle d\mu^*(t)(x) \\
&- \int_{\mathbb{R}^d} \left\langle \nabla_{\mu} \mathcal{E}_l(t, \mu^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(T,t)}^{v,u^*}(x) \right\rangle d\mu^*(t)(x)
\end{aligned}$$

since by Fubini's Theorem, it holds

$$\begin{aligned}
\int_{\mathbb{R}^{2d}} \left\langle r, \int_{\mathbb{R}^d} \Pi_{(t,x)}^v(y) \mathcal{F}_t^{\omega_k, \tau_k}(y) d\mu^*(t)(y) \right\rangle d\nu_N^*(t)(x, r) &= \int_{\mathbb{R}^d} \Pi^v[\nu_N^*(t)](t, x) d\mu^*(t)(x) \\
&= \int_{\mathbb{R}^{2d}} \Pi^v[\nu_N^*(t)](t, x) d\nu_N^*(t)(x, r).
\end{aligned}$$

Recalling the definition of the vector field  $\mathcal{V}_N^*[\cdot](\cdot, \cdot, \cdot)$  given in (3.12), we observe that  $\frac{d}{dt} \mathcal{K}_{\omega_k, \tau_k}^N(t) = 0$  for  $\mathcal{L}^1$ -almost every  $t \in [\tau_k, T)$ , so that it is constant over this time interval. In particular, it holds that

$$\mathcal{K}_{\omega_k, \tau_k}^N(\tau_k) = \mathcal{K}_{\omega_k, \tau_k}^N(T) \leq 0. \tag{3.16}$$

The evaluation of the map  $\mathcal{K}_{\omega_k, \tau_k}^N(\cdot)$  at  $\tau_k$  can be written explicitly as

$$\begin{aligned}
\mathcal{K}_{\omega_k, \tau_k}^N(\tau_k) &= \mathcal{H}(\tau_k, \nu_N^*(\tau_k), \zeta_N^*(\tau_k), \omega_k) - \mathcal{H}(\tau_k, \nu_N^*(\tau_k), \zeta_N^*(\tau_k), u^*(\tau_k)) \\
&- \sum_{l=1}^r \varpi_l^N(\{\tau_k\}) \int_{\mathbb{R}^d} \langle \nabla_{\mu} \Lambda_l(\tau_k, \mu^*(\tau_k)), \omega_k(x) - u^*(\tau_k, x) \rangle d\mu^*(\tau_k)(x).
\end{aligned} \tag{3.17}$$

Combining (3.16) and (3.17) yields the proof of our claim.  $\square$

## Step 4 : Limiting procedure

### The PMP for absolutely continuous state constraints multipliers

In Step 3, we have built for any  $N \geq 1$  a suitable state-costate curve  $\nu_N^*(\cdot)$  solution of the Hamiltonian system (1.1), and such that the relaxed Pontryagin maximization condition 3.14 holds on an  $N$ -dimensional subset of needle parameters. The last step in the proof of Theorem 1.1 is to take the limit as  $N$  goes to infinity of the previous optimality conditions in order to recover the PMP formulated on the whole set of needle parameters.

By the Non-Triviality condition (NT), the sequence of Lagrange multipliers  $(\lambda_0^N, \dots, \lambda_n^N, \eta_1^N, \dots, \eta_m^N) \subset \{0, 1\} \times \mathbb{R}_+^n \times \mathbb{R}^m$  is bounded uniformly with respect to  $N$ . Hence, we can extract a subsequence of multipliers such that  $(\lambda_0^N, \dots, \lambda_n^N, \eta_1^N, \dots, \eta_m^N) \rightarrow (\lambda_0, \dots, \lambda_n, \eta_1, \dots, \eta_m)$  as  $N \rightarrow +\infty$ . A straightforward passage to the limit shows that these limit multipliers satisfy the Complementary Slackness condition

$$\lambda_i \Psi_i^J(\mu^*(T)) = 0 \quad \text{for all } i \in \{1, \dots, n\}.$$

Similarly, the sequence of measures  $(\varpi_1^N, \dots, \varpi_r^N) \subset \mathcal{M}_+([0, T])^r$  is uniformly bounded with respect to the total variation norm. By Banach-Alaoglu's Theorem (see e.g. [16, Theorem 3.16]), it admits a weakly-\* converging subsequence to some  $(\varpi_1, \dots, \varpi_r) \in \mathcal{M}_+([0, T])^r$ . Moreover, since for any  $l \in \{1, \dots, r\}$  the measures  $(\varpi_l^N)$  are equi-supported in the sets  $\{t \in [0, T] \text{ s.t. } \Lambda_l(t, \mu^*(t)) = 0\}$ . It can be shown by standard convergence properties of the supports of converging sequences of measures (see e.g. [9, Proposition 5.1.8]) that the limit measures  $(\varpi_1, \dots, \varpi_r)$  are supported on these sets as well. Therefore, it holds that

$$\text{supp}(\varpi_l) \subseteq \left\{ t \in [0, T] \text{ s.t. } \Lambda_l(t, \mu^*(t)) = 0 \right\},$$

for any  $l \in \{1, \dots, r\}$ . Furthermore, if the scalar Lagrange multipliers  $(\lambda_0^N, \dots, \lambda_n^N, \eta_1^N, \dots, \eta_m^N)$  vanish for large  $N$ , it follows from the Non-Triviality condition (NT) that  $\|\varpi_l^N\|_{TV} > 0$  so that  $\varpi_l \neq 0$ . Hence, we recover the Non-Degeneracy condition (1.5), i.e.

$$(\lambda_0, \dots, \lambda_n, \eta_1, \dots, \eta_m, \varpi_1, \dots, \varpi_r) \neq 0.$$

In Lemma 3.3, we have shown that the curves of measures  $\nu_N^*(\cdot)$  are uniformly equi-compactly supported and equi-Lipschitz. Hence,  $(\nu^*(\cdot))$  admits converging subsequences in the  $C^0W_1$ -topology by Ascoli-Arzelà Theorem.

We now prove that there exists an accumulation point  $\nu^*(\cdot)$  of  $(\nu_N^*(\cdot))$  which solves the system of equations (1.1) associated to the limit multipliers  $(\lambda_0, \dots, \lambda_n, \eta_1, \dots, \eta_m, \varpi_1, \dots, \varpi_r)$ . To this end, we start by making an extra simplifying assumption on the state constraints multipliers.

**(H7)** : The measures  $(\varpi_1, \dots, \varpi_r)$  are absolutely continuous with respect to  $\mathcal{L}^1$  on  $[0, T]$ .

We shall see in the sequel how this extra assumption can be lifted at the price of an extra approximation argument by absolutely continuous measures. Let  $\nu^*(\cdot) \in \text{Lip}([0, T], \mathcal{P}(\overline{B_{2d}(0, R'_T)}))$  be an accumulation point of  $(\nu_N^*(\cdot))$  along a suitable subsequence. As a direct consequence of the convergence of the scalar Lagrange multipliers, one recovers the uniform convergence of the final gradient map

$$\nabla_{\mu} \mathcal{S}_N(\mu^*(T)) \xrightarrow[N \rightarrow +\infty]{C^0} \nabla_{\mu} \mathcal{S}(\mu^*(T)).$$

This implies by standard convergence results for pushforwards of measures (see e.g. [9, Lemma 5.2.1]) that  $\nu^*(\cdot)$  satisfies the boundary condition

$$\pi_{\#}^2 \nu^*(T) = (-\nabla_{\mu} \mathcal{S}(\mu^*(T)))_{\#} \mu^*(T).$$

Moreover, the weak-\* convergence of  $(\varpi_1^N, \dots, \varpi_r^N)$  towards  $(\varpi_1, \dots, \varpi_r)$  along with **(H7)** implies by Proposition 2.1 that

$$\zeta_{l,N}^*(t) = \mathbb{1}_{[0,T)}(t) \int_t^T d\varpi_l^N(s) \xrightarrow{N \rightarrow +\infty} \mathbb{1}_{[0,T)}(t) \int_t^T d\varpi_l(s) = \zeta_l^*(t)$$

for all times  $t \in [0, T]$ . By definition (2.7) of distributional solutions to transport equations, the fact that  $\nu_N^*(\cdot)$  is a solution of (3.12) can be written as

$$\int_0^T \int_{\mathbb{R}^{2d}} \left( \partial_t \xi(t, x, r) + \left\langle \nabla_{(x,r)} \xi(t, x, r), \mathbb{J}_{2d} \nabla_{\nu} \mathcal{H}(t, \nu_N^*(t), \zeta_N^*(t), u^*(t))(x, r) \right\rangle \right) d\nu_N^*(t)(x, r) dt = 0 \quad (3.18)$$

for any  $\xi \in C_c^\infty([0, T] \times \mathbb{R}^{2d})$ . Since all the functionals involved in the definition of the Wasserstein gradient of the augmented infinite-dimensional Hamiltonian are continuous and bounded, we have that

$$\nabla_{\nu} \mathcal{H}(t, \nu_N^*(t), \zeta_N^*(t), u^*(t))(\cdot, \cdot) \xrightarrow[N \rightarrow +\infty]{C^0} \nabla_{\nu} \mathcal{H}(t, \nu^*(t), \zeta^*(t), u^*(t))(\cdot, \cdot)$$

uniformly with respect to  $t \in [0, T]$ . Therefore, passing to the limit in (3.18) as  $N \rightarrow +\infty$  along with an application of Lebesgue's Dominated Convergence Theorem, we recover that

$$\int_0^T \int_{\mathbb{R}^{2d}} \left( \partial_t \xi(t, x, r) + \left\langle \nabla_{(x,r)} \xi(t, x, r), \nabla_{\nu} \mathcal{H}(t, \nu^*(t), \zeta^*(t), u^*(t))(x, r) \right\rangle \right) d\nu^*(t)(x, r) dt = 0$$

for any  $\xi \in C_c^\infty([0, T] \times \mathbb{R}^{2d})$ .

In order to complete our proof of Theorem 1.1, there remains to show that the limit curve  $\nu^*(\cdot)$  is such that the maximization condition (1.6) holds for  $\mathcal{L}^1$ -almost every  $t \in [0, T]$ . We know that for any  $(\omega_k, \tau_k) \in U_D \times \mathcal{A}$ , the modified maximization condition (3.14) holds. By the preliminary assumptions that the limit measures  $(\varpi_1, \dots, \varpi_r)$  are absolutely continuous with respect to  $\mathcal{L}^1$ , we can apply Proposition 2.1 to recover that

$$\varpi_l^N(\{\tau_k\}) \xrightarrow[N \rightarrow +\infty]{} \varpi_l(\{\tau_k\}) = 0,$$

for any  $l \in \{1, \dots, N\}$ . Since the infinite-dimensional Hamiltonian is continuous with respect to its second argument in the  $W_1$ -metric and linear in its third argument, it holds that

$$\mathcal{H}(\tau_k, \nu_N^*(\tau_k), \zeta_N^*(\tau_k), \omega_k) \xrightarrow[N \rightarrow +\infty]{} \mathcal{H}(\tau_k, \nu^*(\tau_k), \zeta^*(\tau_k), \omega_k)$$

and

$$\mathcal{H}(\tau_k, \nu_N^*(\tau_k), \zeta_N^*(\tau_k), u^*(\tau_k)) \xrightarrow[N \rightarrow +\infty]{} \mathcal{H}(\tau_k, \nu^*(\tau_k), \zeta^*(\tau_k), u^*(\tau_k))$$

uniformly with respect to  $k$ . We can therefore pass to the limit as  $N \rightarrow +\infty$  in the partial maximization condition (3.14) to obtain that

$$\mathcal{H}(\tau_k, \nu^*(\tau_k), \zeta^*(\tau_k), \omega_k) \leq \mathcal{H}(\tau_k, \nu^*(\tau_k), \zeta^*(\tau_k), u^*(\tau_k)) \quad (3.19)$$

for any  $(\omega_k, \tau_k) \in U_D \times \mathcal{A}$ .

Given an arbitrary pair  $(\omega, \tau) \in U \times \mathcal{M}$ , it is possible to choose a sequence of elements  $\{(\omega_k, \tau_k)\}_k \subset U_D \times \mathcal{A}$  such that

$$\tau_k \xrightarrow[k \rightarrow +\infty]{} \tau, \quad \omega_k \xrightarrow[k \rightarrow +\infty]{C^0} \omega,$$

and

$$\|u^*(\tau, \cdot) - u^*(\tau_k, \cdot)\|_{C^0(\overline{B(0, R_T)}, \mathbb{R}^d)} \xrightarrow{k \rightarrow +\infty} 0, \quad \|L(\tau, \mu^*(\tau), \cdot) - L(\tau_k, \mu^*(\tau_k), \cdot)\|_{C^0(\overline{B(0, R_T)}, \mathbb{R})} \xrightarrow{k \rightarrow +\infty} 0. \quad (3.20)$$

Remark first that under assumption **(H7)**, the maps  $t \rightarrow \zeta_l^*(t)$  are continuous on  $[0, T]$ . By (3.20) along with the continuity of the augmented infinite-dimensional Hamiltonian in the  $C^0$ -norm topology with respect to its fourth argument, we can pass to the limit as  $k \rightarrow +\infty$  in (3.19). This yields the Pontryagin Maximization condition

$$\mathcal{H}(\tau, \nu^*(\tau), \zeta^*(\tau), \omega) \leq \mathcal{H}(\tau, \nu^*(\tau), \zeta^*(\tau), u^*(\tau))$$

for any pair  $(\omega, \tau) \in U \times \mathcal{M}$ .

### Lifting the absolute continuity hypothesis (H7)

In order to recover the full proof of Theorem 1.1, we show how to relax the absolute continuity assumption **(H7)** made on the state constraints multipliers. For a given small parameter  $\epsilon > 0$ , we consider a sequence of mollifiers  $\rho_\epsilon : t \in [0, T] \mapsto \epsilon^{-1} \rho(t/\epsilon)$  where  $\rho \in C_c^\infty([0, T])$  and  $\int_0^T \rho(t) dt = 1$ . Given  $N \geq 1$ , we define the mollified measure  $(\varpi_1^{N, \epsilon}, \dots, \varpi_r^{N, \epsilon})$  by

$$\varpi_l^{N, \epsilon} = (\rho_\epsilon * \varpi_l^N)(t) \cdot \mathcal{L}^1$$

where for any  $l \in \{1, \dots, r\}$  the convolution maps are defined by  $\rho_\epsilon * \varpi_l^N : t \in [0, T] \mapsto \int_0^T \rho_\epsilon(t-s) d\varpi_l^N(s)$ . Using the fact that the functions

$$t \in [\tau_k, T] \mapsto \int_{\mathbb{R}^d} \langle \nabla_\mu \Lambda_l(t, \mu^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(t)(x)$$

are Lipschitz and bounded, one can assert using the definition of the dual norm in the Banach space  $\mathcal{M}_+([0, T])$  that

$$\begin{aligned} & - \int_{\tau_k}^T \int_{\mathbb{R}^d} \langle \nabla_\mu \Lambda_l(t, \mu^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(t)(x) d\varpi_l^N(t) \\ & \geq - \int_{\tau_k}^T \int_{\mathbb{R}^d} \langle \nabla_\mu \Lambda_l(t, \mu^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(t)(x) d\varpi_l^{N, \epsilon}(t) - C\epsilon \end{aligned}$$

for some uniform constant  $C > 0$ . This allows us to rewrite the optimality conditions (3.11) derived at time  $T$  as

$$\begin{aligned} & \int_{\mathbb{R}^d} \langle -\nabla_\mu \mathcal{S}_N(\mu^*(T))(x), \mathcal{F}_T^{\omega_k, \tau_k} \circ \Phi_{(T, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(T)(x) - \left( L(\tau_k, \mu^*(\tau_k), u^*(\tau_k)) - L(\tau_k, \mu^*(\tau_k), \omega_k) \right) \\ & - \int_{\tau_k}^T \int_{\mathbb{R}^d} \langle \nabla_\mu L(t, \mu^*(t), u^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(t)(x) dt \\ & - \sum_{l=1}^r \int_{\tau_k}^T \int_{\mathbb{R}^d} \langle \nabla_\mu \Lambda_l(t, \mu^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(t)(x) d\varpi_l^{N, \epsilon}(t) \leq C\epsilon \end{aligned} \quad (3.21)$$

By defining the family of measure curves  $(\nu_{N, \epsilon}^*(\cdot))$  as in Lemma 3.3, we can prove that the corresponding maps  $\mathcal{K}_{\omega_k, \tau_k}^{N, \epsilon}(\cdot)$  defined as in (3.15) are constant over  $[\tau_k, T]$  and that the partial maximization conditions

$$\mathcal{H}(\tau_k, \nu_{N, \epsilon}^*(\tau_k), \zeta_{N, \epsilon}^*(\tau_k), \omega_k) \leq \mathcal{H}(\tau_k, \nu_{N, \epsilon}^*(\tau_k), \zeta_{N, \epsilon}^*(\tau_k), u^*(\tau_k)) + C\epsilon$$

hold for any  $\epsilon > 0$ . Performing the same limiting arguments for  $N \rightarrow +\infty$  as in Section 3 - Step 4 and remarking that

$$\varpi_l^{N,\epsilon} \xrightarrow{N \rightarrow +\infty} \varpi_l^\epsilon = (\rho_\epsilon * \varpi_l) \cdot \mathcal{L}^1,$$

we recover the statement of the PMP with a measure curve  $\nu_\epsilon^*(\cdot)$  associated to the Lagrange multipliers  $(\lambda_0, \dots, \lambda_n, \eta_1, \dots, \eta_m, \varpi_1^\epsilon, \dots, \varpi_r^\epsilon)$ . This limit curve is such that the relaxed maximization condition

$$\mathcal{H}(\tau, \nu_\epsilon^*(\tau), \zeta_\epsilon^*(\tau), \omega) \leq \mathcal{H}(\tau_k, \nu_\epsilon^*(\tau), \zeta_\epsilon^*(\tau), u^*(\tau)) + C\epsilon \quad (3.22)$$

holds for any  $(\omega, \tau) \in U \times \mathcal{M}$ . There now remains to perform a last limiting argument as  $\epsilon \downarrow 0$  to recover the full maximum principle.

By Lebesgue's Decomposition Theorem for finite Borel measures on the real line (see e.g. [7, Remark 3.32, Corollary 3.33]), the sets  $\{t \in [0, T] \text{ s.t. } \varpi_l(\{t\}) > 0\}$  are at most countable for all  $l \in \{1, \dots, r\}$ . Therefore, the set  $\mathcal{M}^\circ = \mathcal{M} \setminus \bigcup_{l=1}^r \{t \in [0, T] \text{ s.t. } \varpi_l(\{t\}) > 0\}$  has full Lebesgue measure in  $[0, T]$ , and by Proposition 2.1 it is such that

$$\zeta_\epsilon^*(\tau) = \int_\tau^T d\varpi_N^\epsilon(t) \xrightarrow{\epsilon \downarrow 0} \int_t^T d\varpi_N^\epsilon(\tau) = \zeta^*(\tau).$$

for any  $\tau \in \mathcal{M}^\circ$ . By Ascoli-Arzelà Theorem, it holds that  $\nu_\epsilon^*(\cdot) \rightarrow \nu^*(\cdot)$  uniformly in the  $W_1$ -metric and it can be checked that this limit curve solves the forward-backward system of continuity equations (1.1) associated with the multipliers  $(\lambda_0, \dots, \lambda_n, \eta_1, \dots, \eta_m, \varpi_1, \dots, \varpi_r)$ . Moreover, letting  $\epsilon \downarrow 0$  in (3.22) implies that the Pontryagin maximization condition (1.6) holds on  $U \times \mathcal{M}^\circ$ .

## A Wasserstein differential of the running constraint penalization

In this Section, we give the analytical expression of the Wasserstein differential of the running constraint penalization map  $(t, \mu, \zeta) \mapsto \mathcal{C}(t, \mu, \zeta)$  defined in (1.4).

**Proposition A.1** (Wasserstein differential of the state constraints penalization map). *Let  $K \subset \mathbb{R}^d$  be a compact set. Under hypotheses (H6), the map*

$$\mu \in \mathcal{P}(K) \mapsto \mathcal{C}(t, \mu, \zeta^*(t)) = \sum_{l=1}^r \zeta_l^*(t) \left( \partial_t \Lambda_l(t, \mu) + \int_{\mathbb{R}^d} \langle \nabla_\mu \Lambda_l(t, \mu)(x), v[\mu](t, x) + u^*(t, x) \rangle d\mu(x) \right)$$

*is Wasserstein-differentiable at any  $\mu \in \mathcal{P}(K)$  and the Wasserstein gradients of its summands can be computed explicitly as*

$$\begin{aligned} \nabla_\mu \mathcal{C}_l(t, \mu)(x) &= \partial_t \nabla_\mu \Lambda_l(t, \mu)(x) + D_x \nabla_\mu \Lambda_l(t, \mu)(x)^\top (v[\mu](t, x) + u^*(t, x)) \\ &\quad + \left( D_x v[\mu](t, x)^\top + D_x u^*(t, x)^\top \right) \nabla_\mu \Lambda_l(t, \mu)(x) + \int_{\mathbb{R}^d} \mathbb{I}_{(t,y)}^v(x)^\top \nabla_\mu \Lambda_l(t, \mu)(y) d\mu(y) \\ &\quad + \int_{\mathbb{R}^d} \mathbb{I}_{(t,y)}^{\nabla_\mu \Lambda_l}(x)^\top (v[\mu](t, y) + u^*(t, y)) d\mu(y) \end{aligned} \quad (\text{A.1})$$

where  $(t, x, y) \mapsto \mathbb{I}_{(t,y)}^{\nabla_\mu \Lambda_l}(x) = \left( \mathbb{I}_{(t,y)}^{\nabla_\mu \Lambda_l, i}(x) \right)_{1 \leq i \leq d}$  are the matrix-valued maps which columns are the Wasserstein gradients of the components of  $\nabla_\mu \Lambda_l^i(t, \mu)(x)$ , i.e.

$$\mathbb{I}_{(t,y)}^{\nabla_\mu \Lambda_l, i}(x) = \nabla_\mu \left[ \nabla_\mu \Lambda_l^i(t, \cdot)(y) \right](\mu)(x).$$

*Proof.* In order to lighten the coming computations, we introduce the auxiliary functions

$$\mathcal{C}_l^1(t, \mu) = \partial_t \Lambda_l(t, \mu) \text{ and } \mathcal{C}_l^2(t, \mu) = \int_{\mathbb{R}^d} \langle \nabla_\mu \Lambda_l(t, \mu)(x), v[\mu](t, x) + u^*(t, x) \rangle d\mu(x).$$

Let  $t \in [0, T]$  and  $\mu \in \mathcal{P}(K)$ . The Wasserstein gradient of  $\mathcal{E}_l^1(t, \cdot)$  at  $\mu$  is given by

$$\nabla_{\mu} \mathcal{E}_l^1(t, \mu) = \nabla_{\mu} (\partial_t \Lambda_l(t, \cdot)) (\mu) = \partial_t \nabla_{\mu} \Lambda_l(t, \mu) \quad (\text{A.2})$$

We turn our attention to  $\mathcal{E}_l^2(t, \cdot)$ . For any  $\nu \in \mathcal{P}(K)$  and  $\gamma \in \Gamma_o(\mu, \nu)$ , it holds that

$$\begin{aligned} & \mathcal{E}_l^2(t, \nu) - \mathcal{E}_l^2(t, \mu) \\ &= \int_{\mathbb{R}^{2d}} \left( \langle \nabla_{\mu} \Lambda_l(t, \nu)(y), v[\nu](t, y) + u^*(t, y) \rangle - \langle \nabla_{\mu} \Lambda_l(t, \mu)(x), v[\mu](t, x) + u^*(t, x) \rangle \right) d\gamma(x, y) \\ &= \int_{\mathbb{R}^d} \left( \langle \nabla_{\mu} \Lambda_l(t, \nu)(x), v[\nu](t, x) + u^*(t, x) \rangle - \langle \nabla_{\mu} \Lambda_l(t, \mu)(x), v[\mu](t, x) + u^*(t, x) \rangle \right) d\mu(x) \quad (\text{A.3}) \\ &+ \int_{\mathbb{R}^{2d}} \left\langle \mathbb{D}_x \nabla_{\mu} \Lambda_l(t, \mu)(x)^{\top} (v[\mu](x) + u^*(t, x)), y - x \right\rangle d\gamma(x, y) \\ &+ \int_{\mathbb{R}^{2d}} \left\langle (\mathbb{D}_x v[\mu](t, x)^{\top} + \mathbb{D}_x u^*(t, x)^{\top}) \nabla_{\mu} \Lambda_l(t, \mu)(x), y - x \right\rangle d\gamma(x, y) + \int_{\mathbb{R}^{2d}} o(|x - y|) d\gamma(x, y) \end{aligned}$$

By definition of the Landau notation  $o(\cdot)$ , for any  $\epsilon > 0$  there exists  $\eta > 0$  such that whenever  $|x - y| \leq \eta$ , one has that  $o(|x - y|) \leq \epsilon|x - y|$ . Therefore,

$$\begin{aligned} \int_{\mathbb{R}^{2d}} o(|x - y|) d\gamma(x, y) &\leq \epsilon \int_{\{|x-y| \leq \eta\}} |x - y| d\gamma(x, y) + C \int_{\{|x-y| > \eta\}} |x - y| d\gamma(x, y) \\ &\leq \epsilon W_2(\mu, \nu) + 2C \text{diam}(K) \gamma \left( \{(x, y) \in \mathbb{R}^{2d} \text{ s.t. } |x - y| > \eta\} \right) \\ &\leq \epsilon W_2(\mu, \nu) + \frac{2C}{\eta^2} \text{diam}(K) W_2^2(\mu, \nu) \end{aligned}$$

by Chebyshev's inequality and where the constant  $C > 0$  exists because  $o(|x - y|)$  is in particular a  $O(|x - y|)$  on the compact set  $\text{supp}(\gamma) \subset \mathbb{R}^{2d}$  for  $|x - y| > \eta$ . Upon choosing  $\eta' = \eta^2 \epsilon / (2C \text{diam}(K))$ , we recover that

$$\int_{\mathbb{R}^{2d}} o(|x - y|) d\gamma(x, y) \leq 2\epsilon W_2(\mu, \nu).$$

whenever  $W_2(\mu, \nu) \leq \eta'$ . By definition, this estimate precisely amounts to the fact that  $\int_{\mathbb{R}^{2d}} o(|x - y|) d\gamma(x, y) = o(W_2(\mu, \nu))$  as  $W_2(\mu, \nu) \rightarrow 0$ .

We further compute the first-order variations arising from the remaining measure terms as follows

$$\begin{aligned} & \int_{\mathbb{R}^d} \left( \langle \nabla_{\mu} \Lambda_l(t, \nu)(x), v[\nu](t, x) + u^*(t, x) \rangle - \langle \nabla_{\mu} \Lambda_l(t, \mu)(x), v[\mu](t, x) + u^*(t, x) \rangle \right) d\mu(x) \\ &= \int_{\mathbb{R}^d} \left\langle \int_{\mathbb{R}^{2d}} \mathbb{I}_{(t,x)}^{\nabla_{\mu} \Lambda_l}(x')(y' - x') d\gamma'(x', y'), v[\mu](t, x) + u^*(t, x) \right\rangle d\mu(x) \\ &+ \int_{\mathbb{R}^d} \left\langle \int_{\mathbb{R}^{2d}} \mathbb{I}_{(t,x)}^v(x')(y' - x') d\gamma'(x', y'), \nabla_{\mu} \Lambda_l(t, \mu)(x) \right\rangle d\mu(x) + o(W_2(\mu, \nu)) \\ &= \int_{\mathbb{R}^{2d}} \left\langle \int_{\mathbb{R}^d} \left( \mathbb{I}_{(t,y')}^v(x)^{\top} \nabla_{\mu} \Lambda_l(t, \mu)(y') + \mathbb{I}_{(t,y')}^{\nabla_{\mu} \Lambda_l}(x)^{\top} \nabla_{\mu} \Lambda_l(t, \mu)(y') \right) d\mu(y'), y - x \right\rangle d\gamma(x, y) + o(W_2(\mu, \nu)) \quad (\text{A.4}) \end{aligned}$$

by a standard application of Fubini's Theorem. Merging equations (A.2), (A.3) and (A.4), we recover the characterization (2.4) of the Wasserstein gradient  $\nabla_{\mu} \mathcal{E}_l(t, \mu)$  of  $\mathcal{E}_l(t, \cdot)$  at  $\mu$  given by (A.1).  $\square$

## References

- [1] Y. Achdou and M. Laurière. On the System of Partial Differential Equations Arising in Mean Field type Control. *Discrete and Continuous Dynamical Systems*, 35(9):3879–3900, 2015.
- [2] Y. Achdou and M. Laurière. Mean Field Type Control with Congestion. *Applied Mathematics & Optimization*, 73(3):393–418, Jun 2016.

- [3] A. Agrachev and Y. Sachkov. *Control Theory from the Geometric Viewpoint*, volume 87 of *Encyclopaedia of Mathematical Sciences*. Springer, 2004.
- [4] G. Albi, M. Bongini, E. Cristiani, and D. Kalise. Invisible control of self-organizing agents leaving unknown environments. *SIAM J. Appl. Math.*, 76(4):1683–1710, July 2016.
- [5] G. Albi, L. Pareschi, and M. Zanella. Boltzmann type Control of Opinion Consensus through Leaders. *Proc. of the Roy. Soc. A.*, 2014.
- [6] L. Ambrosio. Transport Equation and Cauchy Problem for bv Vector Fields. *Inventiones Mathematicae*, 158, 2004.
- [7] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variations and Free Discontinuity Problems*. Oxford Mathematical Monographs, 2000.
- [8] L. Ambrosio and W. Gangbo. Hamiltonian ODEs in the Wasserstein space of probability measures. *Comm. Pure Appl. Math.*, 61(1):18–53, 2008.
- [9] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.
- [10] A.V. Arutyunov, D. Y. Karamzin, and F. L. Pereira. The Maximum Principle for Optimal Control Problems with State Constraints by R.V. Gamkrelidze: Revisited. *Journal of Optimization Theory and Applications*, 149(3), 2011.
- [11] A.V. Arutyunov and R. Vinter. A Simple "Finite Approximations" Proof of the Pontryagin Maximum Principle under Reduced Differentiability Hypotheses. *Set-Valued Analysis*, 12(1), 2004.
- [12] 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.
- [13] M. Bongini, M. Fornasier, F. Rossi, and F. Solombrino. Mean Field Pontryagin Maximum Principle. *Journal of Optimization Theory and Applications*, 175:1–38, 2017.
- [14] B. Bonnet and F. Rossi. The Pontryagin Maximum Principle in the Wasserstein Space. *Calculus of Variations and Partial Differential Equations*, 58:11, 2019.
- [15] 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.
- [16] H. Brézis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Universitext. Springer, 2010.
- [17] G. Cavagnari, A. Marigonda, K.T. Nguyen, and F.S. Priuli. Generalized Control Systems in the Space of Probability Measures. *Set-Valued and Variational Analysis*, 26(3):663–691, 2018.
- [18] G. Cavagnari, A. Marigonda, and B. Piccoli. Averaged Time-Optimal Control Problem in the Space of Positive Borel Measures. *ESAIM COCV*, 24, 2017.
- [19] F. Clarke. *Functional Analysis, Calculus of Variations and Optimal Control*. Springer, 2013.
- [20] F. Cucker and S. Smale. Emergent Behavior in Flocks. *IEEE Trans. Automat. Control*, 52(5):852–862, 2007.
- [21] R.L. Di Perna and Lions P.L. Ordinary Differential Equations, Transport Theory and Sobolev Spaces. *Inventiones Mathematicae*, 98, 1989.
- [22] J. Diestel and J.J.Jr Uhl. *Vector Measures*, volume 15. American Mathematical Society, 1977.
- [23] J. Dugundji. An Extension of Tietze’s Theorem. *Pacific Journal of Mathematics*, 1(3), 1951.
- [24] M. Duprez, M. Morancey, and F. Rossi. Approximate and Exact Controllability of the Continuity Equation with a Localized Vector Field. *Submitted*.
- [25] K. Elamvazhuthi and S. Berman. Optimal Control of Stochastic Coverage Strategies for Robotic Swarms. In *IEEE Int. Conf. Rob. Aut.*, 2015.
- [26] L.C. Evans and R.F. Gariepy. *Measure Theory and Fine Properties of Functions*. CRC Press, 1992.
- [27] A. Ferscha and K. Zia. Lifebelt: Crowd evacuation based on vibro-tactile guidance. *IEEE Pervasive Computing*, 9(4):33–42, 2010.
- [28] M. Fornasier, S. Lisini, C. Orrieri, and G. Savaré. Mean-Field Optimal Control as Gamma-Limit of Finite Agent Controls. *preprint*, 2018. <https://arxiv.org/pdf/1803.04689.pdf>.
- [29] M. Fornasier, B. Piccoli, and F. Rossi. Mean-Field Sparse Optimal Control. *Phil. Trans. R. Soc. A*, 372(2028):20130400, 2014.
- [30] M. Fornasier and F. Solombrino. Mean field optimal control. *Esaim COCV*, 20(4):1123–1152, 2014.
- [31] W Gangbo, T Nguyen, , and A. Tudorascu. Hamilton-Jacobi Equations in the Wasserstein Space. *Methods and Applications of Analysis*, 15(2):155–184, 2008.
- [32] W. Gangbo and A. Tudorascu. On Differentiability in Wasserstein Spaces and Well-Posedness for Hamilton-Jacobi Equations. *Technical Report*, 2017.

- [33] R. Hegselmann and U. Krause. Opinion Dynamics and Bounded Confidence Models, Analysis, and Simulation. *Journal of artificial societies and social simulation*, 5(3), 2002.
- [34] A.D. Ioffe. A Lagrange Multiplier Rule with Small Convex-Valued Subdifferentials for Nonsmooth Problems of Mathematical Programming Involving Equality and Nonfunctional Constraints. *Math. Program.*, 58:137–145, 1993.
- [35] A.D. Ioffe and V.M. Tihomirov. *Theory of Extremal Problems*. North Holland Publishing Company, Elsevier, 1979.
- [36] J.P. Penot and P. Michel. Calcul sous-différentiel pour les fonction Lipschitziennes et non-Lipschitziennes. *C. R. Acad. Sci. Paris Sér.I*, 298:269–272, 1984.
- [37] 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.
- [38] B. Piccoli, F. Rossi, and E. Trélat. Control of the kinetic Cucker-Smale model. *SIAM J. Math. Anal.*, 47(6):4685–4719, 2015.
- [39] F. Santambrogio. *Optimal Transpor for Applied Mathematicians*, volume 87. Birkhauser Basel, 2015.
- [40] I.A. Shvartsman. New approximation method in the proof of the maximum principle for nonsmooth optimal control problems with state constraints. *J. Math. Anal. Appl*, (326):974–1000, 2006.
- [41] C. Villani. *Optimal Transport : Old and New*. Springer-Verlag, Berlin, 2009.
- [42] R. Vinter. *Optimal Control*. Modern Birkhauser Classics. Birkhauser Basel, 2000.
- [43] A.A. Vlasov. *Many-Particle Theory and its Application to Plasma*. New York, Gordon and Breach, 1961.