



HAL
open science

Optimal Control in Wasserstein Spaces

Benoît Bonnet

► **To cite this version:**

Benoît Bonnet. Optimal Control in Wasserstein Spaces. Automatic Control Engineering. Aix-Marseille Université; Laboratoire d'Informatique et Systèmes - LIS; Università degli studi di Padova, 2019. English. NNT: . tel-02361353

HAL Id: tel-02361353

<https://hal.science/tel-02361353>

Submitted on 13 Nov 2019

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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

AIX-MARSEILLE UNIVERSITÉ

ÉCOLE DOCTORALE EN MATHÉMATIQUES ET INFORMATIQUE DE MARSEILLE

Laboratoire d'Informatique et des Systèmes

Thèse présentée pour obtenir le grade universitaire de docteur

Discipline : Mathématiques Appliquées
Spécialité : Automatique

Optimal Control in Wasserstein Spaces

Benoît Bonnet

Soutenue le 28/10/2019 devant le jury composé de :

Nicola GIGLI	SISSA Trieste	Rapporteur
Pierre CARDALIAGUET	Université Paris Dauphine	Rapporteur
Hélène FRANKOWSKA	CNRS, Université Paris 6	Examineur
Francesca CHITTARO	Université de Toulon	Examineur
Filippo SANTAMBROGIO	Université Lyon 1	Examineur
José Antonio CARRILLO	Imperial College London	Examineur
Francesco ROSSI	Università degli Studi di Padova	Directeur de thèse
Maxime HAURAY	Aix-Marseille Université	Co-directeur de thèse

Numéro national de thèse/suffixe local : 2019AIXM0442/026ED184



Cette œuvre est mise à disposition selon les termes de la [Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Résumé

Une vaste quantité d'outils mathématiques permettant la modélisation et l'analyse des problèmes multi-agents ont récemment été développés dans le cadre de la théorie du transport optimal. Dans cette thèse, nous étendons pour la première fois plusieurs de ces concepts à des problématiques issues de la théorie du contrôle.

Le premier résultat présenté dans ce manuscrit est la généralisation du principe du maximum de Pontryagin aux problèmes de contrôle optimal multi-agent étudiés dans leur approximation par limite de champs moyen. La preuve de ces résultats repose sur la généralisation de techniques du contrôle géométrique au cadre de la structure Riemannienne des espaces de Wasserstein.

Par la suite, nous investiguons des conditions suffisantes de régularité Lipschitz en espace pour les contrôles optimaux. Ces résultats sont généralement cruciaux pour assurer une correspondance stricte entre les modèles microscopiques et leurs approximations macroscopiques. Nous les obtenons en combinant une approximation par limite de champs moyens et un argument d'existence de feedback Lipschitz optimaux pour les modèles microscopiques sous-jacents.

Nous nous intéressons ensuite aux modèles d'alignement. Nous proposons une analyse de convergence de type Lyapunov pour une classe de systèmes coopératifs présentant des défauts aléatoires de communication. Par la suite, nous présentons une stratégie de contrôle parcimonieuse permettant d'assurer la convergence de systèmes faiblement coopératifs vers un état de presque-alignement.

Nous présentons enfin un résultat de géométrie sous-Riemannienne, dans lequel nous achevons la classification des singularités génériques du lieu conjugué pour les distributions de contact en dimension 3. Ce résultat se base sur des arguments de transversalité appliqués aux jets de la métrique au voisinage de l'origine.

Mots clés: contrôle optimal, systèmes multi-agents, limite de champs moyen, transport optimal, espaces de Wasserstein, calcul sous-différentiel métrique, analyse non-lisse, régularité Lipschitz, régularité métrique, problèmes d'alignement, persistance d'excitations, contrôle parcimonieux, géométrie sous-Riemannienne, théorie des singularités, transversalité.

Abstract

A wealth of mathematical tools allowing to model and analyse multi-agent systems has been brought forth as a consequence of recent developments in optimal transport theory. In this thesis, we extend for the first time several of these concepts to the framework of control theory.

The first result presented in this manuscript is the generalization of the Pontryagin Maximum Principle, both in the absence and presence of constraints, to optimal control problems of multi-agent systems studied in the so-called mean-field approximation framework. The proof of this result relies on the generalization of techniques from geometric control theory to the setting of the Riemannian structure of the Wasserstein spaces.

Subsequently, we investigate sufficient conditions for the Lipschitz-in-space regularity of mean-field optimal control. These results are generally crucial for ensuring the correspondence between the microscopic multi-agent systems and their macroscopic approximations. We obtain them by combining a mean-field approximation argument with an existence result for Lipschitz optimal feedbacks formulated for microscopic multi-agent models.

Later on, we focus our attention on alignment models. We propose a convergence analysis based on Lyapunov-type arguments for cooperative systems subject to random communication failures. We further propose a sparse control strategy which allows to stir weakly-cooperative systems towards a state of almost-alignment.

We finally present a result of sub-Riemannian geometry, in which we complete the classification of the generic singularities of the conjugate locus for contact distributions in dimension 3. This result is based on transversality arguments applied to the jets of the metric in a suitable neighbourhood of the origin.

Keywords : optimal control, multi-agent systems, mean-field limit, optimal transport, Wasserstein spaces, subdifferential calculus in metric spaces, non-smooth analysis, Lipschitz regularity, metric regularity, alignment problems, persistency of excitation, sparse controls, sub-Riemannian geometry, singularity theory, transversality.

Remerciements

Ce manuscrit représente l'aboutissement d'un cheminement personnel et professionnel dans le monde universitaire qui s'est étendu sur trois à six années (selon l'endroit où l'on marque son début), et dont le visage serait résolument différent sans un bon nombre de personnes qui m'ont accompagné durant mes études et que je souhaite remercier ici.

Je voudrais commencer par remercier Pierre Cardaliaguet et Nicola Gigli d'avoir accepté d'être les rapporteurs de ce mémoire. Ce sont deux personnes dont j'admire profondément les travaux académiques, et je leur suis infiniment reconnaissant d'avoir accepté de consacrer un peu de leur temps à la relecture de ceux que j'ai menés durant ces trois années. Cette reconnaissance s'étend aussi à Hélène Frankowska, Francesca Chittaro, José Antonio Carrillo et Filippo Santambrogio qui ont acceptés de faire partie du jury de soutenance de ma thèse.

A la tête des personnes sans qui je n'en serais pas là se situe bien évidemment Francesco. Je lui suis extrêmement reconnaissant de m'avoir accepté en temps qu'élève – avec les contraintes et les responsabilités que cela représente – durant une période pour le moins chargée de sa vie. J'ai énormément appris de sa vision dynamique et toujours en recherche de nouveauté des mathématiques, et je suis très heureux des liens professionnels et humains que nous avons tissés ensemble. Je lui suis aussi très reconnaissant de l'autonomie dont il m'a laissé faire preuve et de la confiance qu'il m'a accordé durant mes pérégrinations mathématiques au cours de ces trois dernières années. Je souhaite par la même occasion remercier Maxime, d'avoir toujours su m'accorder une oreille bienveillante et sûre.

J'adresse des pensées toutes particulières aux enseignants chercheurs que j'ai eu la chance de côtoyer durant mes années d'études, et qui m'ont donné envie par leur passion et leur pédagogie de me lancer dans le monde académique. Un grand merci donc à Michal pour m'avoir encadré en stage à Edinburgh et m'avoir donné goût à la vie de chercheur, à Frédéric pour son enseignement d'AOT et son accompagnement durant mes années à l'ENSTA, et enfin à Hasnaa, Filippo et Ugo pour leurs cours de master fascinants, leur attention et leur gentillesse. Dans cette lignée, il m'est impossible de ne pas mentionner Emmanuel qui a toujours répondu présent tant à l'appel du partage scientifique qu'à celui des sorties au bar les soirs de conférence, ainsi que Jean-Paul dont la personnalité haute en couleurs est une source d'inspiration de tous les instants.

Du côté des Marseillais passés comme présents, j'ai une pensée particulière pour Corinne et Béatrice qui ont toujours été d'une très grande aide – et d'une grande patience ! – pour mes démarches administratives au LIS et au LabEx. Je dois une gratitude toute particulière à Michel, pour m'avoir si fréquemment hébergé dans son bureau à l'I2M pour que l'on puisse travailler ensemble, ainsi qu'à Magali pour l'opportunité qu'elle m'a donné d'enseigner à ses côtés dans le cursus de mathématiques de l'École Centrale. Je remercie aussi Pierluigi pour toutes les sessions de billard que l'on a fait ensemble, ainsi que Sabine pour les soirées géniales que j'ai passées en leur compagnie. Pour ce qui est

des Padovani, je remercie Antoine, Sebastiano, Davide, Annalisa, Francesco et Roberto pour ces superbes groupes de travail de GMT où j'ai eu l'occasion de découvrir ce que "mathématiquement constructif" signifiait, ainsi que Roberto, Pietro, Guglielmo "giro della morte", Federico, Simone "so di Latina" et les autres doctorants du bureau 721 pour tous les bons moments que l'on a passés ensemble durant mes séjours en Italie.

Enfin, cette liste de remerciements ne saurait être complète sans la mention de mes amis et de ma famille, qui sont un phare rassurant, illuminant en toutes circonstances mes sorties dans les mers agitées de la vie d'adulte. Toute ma tendresse va à mes parents, dont l'éducation et l'affection m'ont permis d'arriver là où j'en suis aujourd'hui en restant heureux tout au long du chemin. À Maman, d'être et d'avoir toujours été présente à mes côtés, et pour la confiance en moi qu'elle a su me donner; et à Papa, pour toute l'aide qu'il m'a apporté – et m'apporte encore – que ce soit dans mes études ou dans beaucoup d'autres aspects de ma vie. Je dois aussi beaucoup en terme de développement personnel à mes deux frères, Pierre et Xavier, ainsi qu'à Grand-Maman pour être toujours aussi dynamique et vive d'esprit du haut des soixante années qu'elle a de plus que moi ! Du côté amical du spectre, la liste risque d'être longue... Un remerciement du coeur va à Benjamin pour la patience dont il a fait preuve en m'écoutant parler de sciences depuis toutes ces années, ainsi qu'à JB, Barthélémy, Paul et Claire pour avoir apporté un quota de pipetteurs éclairés à nos discussions. Un grand merci à Paul, le maître artisan dentistier, à Elsa la sculpteuse de Mickeys, ainsi qu'aux autres membres de l'Antre de la Folish pour être les beaux amis que l'on puisse souhaiter avoir. À Émilien, Riccardo et Adrien qui font partie des rares personnes à me supporter plus d'une heure une fois que j'ai commencé à m'oublier dans des torrents d'enthousiasme scientifique. À Florian, Katie, Jérôme et Gilbert, ainsi qu'au grand collectif ENSTA des Gens Qui Connaissent Beaucoup De Choses, majoritairement concentré sur le Slack Enstalibur.

La dernière et la plus tendre de toutes ces pensées va à Élise, qui m'accompagne tous les jours depuis déjà quatre ans, et qui a fait de moi le plus heureux des hommes en acceptant l'été dernier de partager sa vie avec la mienne.

A la mémoire de Jean Winter

Publications and presentations

Publications in referred journals

1. *Intrinsic Lipschitz Regularity of Mean-Field Optimal Controls* (with F. Rossi), Submitted (2019).
2. *Generic Singularities of the 3D Sub-Riemannian Conjugate Locus* (with J.P. Gauthier and F. Rossi), C. R. Acad. Sci. Paris, Ser. I (2019)
3. *A Pontryagin Maximum Principle in Wasserstein Spaces for Constrained Optimal Control Problems*, accepted for publication in *ESAIM COCV* (2019).
4. *The Pontryagin Maximum Principle in the Wasserstein Space* (with F. Rossi), *Calc.Var.PDEs.* 58:11 (2019)

International conference proceedings

1. *Consensus and Flocking in Cooperative Systems with Random Communication Failures* (with É. Flayac and F. Rossi), Preprint (2019).
2. *Sparse Control of Kinetic Cooperative Systems to Approximate Alignment* (with F. Rossi), Proceedings of the 20th IFAC World Congress (2017).

Presentations at international conferences and seminars

1. *Intrinsic Lipschitz Regularity of Mean-Field Optimal Controls*. Seminar of Control and Probability Theory, ENSTA Paris (10/2019).
2. *Topics in Analysis and Optimal Control of Multi-Agent Systems*. Seminar of Analysis, Dipartimento di Matematica “Tullio Levi-Civita”, Padova (03/2019).
3. *Optimal Control of Multi-Agent Systems : A Pontryagin Approach*. Winter School in Calculus of Variations and Probability Theory, Toulouse (02/2019).
4. *Optimal Control Problems in Wasserstein Spaces*. ‘Variational Analysis and Optimal Control’ Invited Session, 14th Viennese Conference on Optimal Control and Dynamic Games, Vienna (07/2018).
5. _____, 12th International Young Researcher Workshop on Geometry, Mechanics and Control, Padova (01/2018).
6. *Sparse Alignment of Kinetic Cooperative Systems*, IFAC World Conference, Toulouse (07/2017).

Poster sessions

1. *Some Problems in Modelling and Optimal Control of Multi-Agent Systems*, Poster session of the conference “Crowds : models and control”, Marseille (06/2019).

Contents

Résumé	3
Abstract	4
Remerciements	6
Introduction	10
1 Analysis in the Space of Measures and Optimal Transport Theory	26
1.1 Elements of measure theory and Wasserstein spaces	26
1.1.1 Elementary notions of measure theory	26
1.1.2 The optimal transport problem and Wasserstein spaces	28
1.1.3 Absolutely continuous curves in $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ and the analytical tangent space $\text{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$	31
1.2 First and second order differential calculus in $(\mathcal{P}_2(\mathbb{R}^d), W_2)$	32
1.2.1 Subdifferential calculus in Wasserstein spaces	32
1.2.2 Classical subdifferentials and Wasserstein gradients	36
1.2.3 Second order calculus in Wasserstein spaces	40
1.3 The continuity equation with non-local velocities in \mathbb{R}^d	41
1.4 Directional derivatives of non-local flows	43
2 The Pontryagin Maximum Principle in the Wasserstein Space	50
2.1 A Simpler Pontryagin Maximum Principle with no interaction field and no running cost	54
2.2 Proof of Theorem 2.1	61
Step 1 : Needle-like variations in the non-local case	61
Step 2 : First-order optimality condition	64
Step 3 : Backward dynamics and Pontryagin maximization condition	65
Step 4 : Proof of the Pontryagin Maximum Principle for (\mathcal{P})	68
2.3 Examples of functionals satisfying hypotheses (H) of Theorem 2.1	69
3 A Pontryagin Maximum Principle for Constrained Optimal Control Problems in Wasserstein Spaces	73
3.1 Non-smooth multiplier rule and differentiable extension of functions	76
3.2 Proof of Theorem 3.1	79
Step 1 : Packages of needle-like variations	79
Step 2 : First-order optimality condition	83
Step 3 : Backward dynamics and partial Pontryagin maximization condition	87
Step 4 : Limiting procedure	92
3.3 Wasserstein differential of the running constraint penalization	97

3.4	Examples of functionals satisfying hypotheses (H)	99
4	Intrinsic Lipschitz Regularity of Mean-Field Optimal Controls	101
4.1	Preliminary material	104
4.1.1	Mean-field adapted structures and discrete measures	104
4.1.2	Existence of mean-field optimal controls for problem (\mathcal{P})	109
4.1.3	Existence of locally optimal Lipschitz feedbacks in finite-dimensional optimal control problems	111
4.2	Proof of Theorem 4.1	116
	Step 1 : Packages of needle-like variations	117
	Step 2 : Construction of a Lipschitz-in-space optimal controls for (\mathcal{P}_N)	119
	Step 3 : Existence of Lipschitz optimal controls for problem (\mathcal{P})	123
4.3	Discussions on the coercivity assumption (\mathbf{CO}_N)	125
4.3.1	A generic sufficient condition for coercivity	125
4.3.2	Sharpness of the mean-field coercivity condition (\mathbf{CO}_N) on a 1D example	127
5	Convergence Analysis and Sparse Control of Weakly Cooperative Alignment Models	132
5.1	Convergence to consensus and flocking of randomly failed cooperative systems	132
5.1.1	Consensus under persistent excitation for first-order dynamics	133
5.1.2	Flocking for Cucker-Smale type systems with strong interactions	137
5.2	Sparse control of kinetic cooperative systems to approximate alignment	143
5.2.1	Invariance properties of kinetic cooperative systems	146
5.2.2	Proof of Theorem 5.3	147
6	Generic Singularities of the 3D-Contact Sub-Riemannian Conjugate Locus	155
6.1	3D-Contact sub-Riemannian manifolds and their conjugate locus	157
6.2	Generic singularities of the full-conjugate locus	160

Introduction

In this thesis, we study several control problems fitting into the broad topic of multi-agent systems. The denomination “multi-agent system” encompasses a wide range of mathematical models describing a large number of autonomous and undistinguishable entities interacting together. One of the fascinating aspects of multi-agent systems is their mathematical versatility: they can be studied from various perspectives and with different formalisms, depending on the modelling tools at hand and on the type of results one aims at deriving. While the simplest, direct models evolved from discrete time-space automata to many-body ensembles described by differential equations, a wide portion of the more recent literature is devoted to the study of these systems in the so-called mean-field limit. In this setting, multi-agent systems are approximated by transport equations describing the evolution of the macroscopic density of agents. The best identified setting for studying such equations is the metric space of probability measures, endowed with the so-called Wasserstein distance. The huge developments undergone by optimal transport theory during the past two decades have brought forth a wide range of powerful mathematical tools for studying such problems.

A first crucial question when it comes to multi-agent systems is the analysis of pattern formations and destructions. Such self-arising motifs come in different forms and shapes, but it is possible to isolate a few of them which have been the object of intensive studies. The so-called aggregation models describe the formation of tight local clusters in a crowd of agents. When there is a single of such clusters, one usually talks about consensus formation. Another broadly studied topic, which is relevant when describing opinion dynamics, is that of alignment models. In this scenario, one aims at understanding under which circumstances a crowd is likely to flock, i.e. to form a consensus in the velocity variable, so that all the agents go in the same direction and “agree” on something.

Most of these patterns were first studied in a strictly finite-dimensional context where the agents are modelled as interacting particles. Yet as mentioned hereabove, it was realized later that infinite-dimensional models were of high relevance for describing large multi-agent systems. A lot of attention was therefore devoted to understand which mechanisms of pattern formations were preserved between discrete and continuous models, how intermediate regimes may behave in the mean-field limit, etc... When a clear picture had emerged on whether these patterns did arise on their own or not, and in which situation, the question of controlling multi-agent systems to promote the emergence of a desired behaviour naturally arose in the literature.

The topics of interest of the community progressively shifted to incorporate classical control-theoretic questions such as reachability and stability analysis, existence of optimal controls, optimality conditions, explicit design of control laws and numerical resolution of optimal control problems. The results presented in this thesis are among the first to tackle in a general and systematic way several central problems in the field of multi-agent control, such as Pontryagin optimality conditions, regularity results for the optimal solutions, and general consensus and flocking formation under communication failures.

For almost a decade now, a lively community at the interface of control theory, computer science, calculus of variations and PDE analysis has been studying control and optimal control problems on multi-agent systems. Let it be noted that these developments were quite parallel to those of the deeply powerful mean-field games theory introduced separately by Lasry and Lions [98] as well as Caines, Huang and Malhamé [92]. Even though it was already sketched at the time that certain classes of mean-field games problems could be reformulated as multi-agent optimal control problems with a central planner, this analogy did not really (to the best of the author’s knowledge) play a major role in the development of the latter theory.

In this thesis, we bring innovative answers to several open questions in the analysis and control of multi-agent systems. Our results are exposed in distinct chapters that roughly correspond the publications prepared during this thesis, at the exception of Chapter 1 which deals with preliminary material for our subsequent developments. Throughout the following paragraphs, we provide a detailed summary of the main results presented in this manuscript. We start by providing the reader with an informal introduction to the mathematical modelling of multi-agent systems.

Mathematical modelling of multi-agent systems

As mentioned previously, when it comes to multi-agent systems one can choose between several possible definitions and formalisms. The more intuitive and direct to manipulate is the following one, which is discrete in space and continuous in time.

Microscopic multi-agent models

A microscopic multi-agent system is the datum of a family of N curves $t \in \mathbb{R}_+ \mapsto (x_1(t), \dots, x_N(t))$ with values in a state space M . This state-space can be fairly arbitrary (metric space, smooth manifold, etc...), but in this thesis we will restrict our attention to the case where $M = \mathbb{R}^d$. The agents are moreover assumed to be *indistinguishable*. This means that none of them plays a special role, and that all the agents could theoretically be interchanged without modifying the global behaviour of the system. Besides its relevance in applications, this hypothesis is crucial to perform the mean-field approximation procedures that we are going to sketch in a subsequent paragraph.

The curves $(x_1(\cdot), \dots, x_N(\cdot))$ are usually defined as solutions of a system of coupled differential equations of the form

$$\dot{x}_i(t) = \mathbf{v}_N[\mathbf{x}_N(t)](t, x_i(t)). \quad (1)$$

where $\mathbf{x}_N = (x_1, \dots, x_N) \in (\mathbb{R}^d)^N$. In this case, the indistinguishability of the agents must be encoded in the dynamics as well. Thus, the vector field $(t, x, \mathbf{x}) \mapsto \mathbf{v}_N[\mathbf{x}](t, x)$ must satisfy some invariance under permutations property in the \mathbf{x} variable, namely

$$\mathbf{v}_N[\sigma(\mathbf{x})](t, x) = \mathbf{v}_N[\mathbf{x}](t, x),$$

for any d -blockwise permutation $\sigma : (\mathbb{R}^d)^N \rightarrow (\mathbb{R}^d)^N$. A typical example, which is frequently encountered in applications ranging from theoretical physics to pedestrian

modelling, is that of general interaction kernels of the form

$$\mathbf{v}_N[\mathbf{x}](t, x) = \frac{1}{N} \sum_{j=1}^N K(x_j - x).$$

However convenient these ODE-based models may be, they do not always capture the global features of the problems at hand. Indeed, be it in social media analysis, robot swarm control or the pedestrian flows, the number N of agents (or particles) is extremely big. Similarly to fluid and gas dynamics, it is then natural to approximate this large finite-dimensional system by a simpler infinite-dimensional one. This is the concept of *mean-field limit*.

Mean-field limits and macroscopic multi-agent models

Mean-field approximations of many-particle systems were first introduced in theoretical physics to describe gas dynamics as early as the work of Vlasov [134]. The mean-field approximation of a multi-agent system such as (1) can be formally derived by introducing the curve of *empirical density of agents*

$$\mu_N(\cdot) = \frac{1}{N} \sum_{i=1}^N \delta_{x_i(\cdot)}$$

associated with $(x_1(\cdot), \dots, x_N(\cdot))$. Here, δ_x denotes the Dirac measure centered at $x \in \mathbb{R}^d$. Suppose now that as one lets N grow larger and larger, i.e. $N \rightarrow +\infty$, the sequence of density curves $(\mu_N(\cdot))$ converges¹ towards a limit curve of densities $\mu(\cdot)$ which are not necessarily empirical.

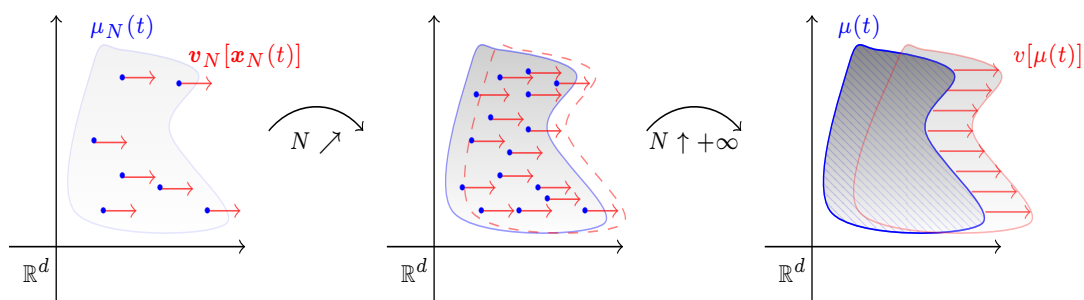


Figure 1: *Illustration of the mean-field approximation procedure as $N \rightarrow +\infty$.*

It is now natural to ask what will be the dynamics of the limit curve $\mu(\cdot)$. Under sufficiently nice assumptions on the agent-based velocity field $\mathbf{v}_N[\cdot](\cdot, \cdot)$ given in (1), it can be shown in great generality that the limit dynamics for the curve of density is given by a *transport equation in divergence form* which writes

$$\partial_t \mu(t) + \nabla \cdot (v[\mu(t)](t, \cdot) \mu(t)) = 0. \quad (2)$$

1. In a sense which will be made precise in Chapter 1.

Here, $v[\mu](\cdot, \cdot)$ is a *non-local vector field* satisfying the compatibility² condition

$$v[\mu_N](\cdot, \cdot) = \mathbf{v}_N[\mathbf{x}_N](\cdot, \cdot).$$

Transport equations with non-local interaction terms have been intensively studied for decades by various communities. As mentioned previously, they frequently appear as mean-field limits of particle systems, see e.g. [111, 129]. More recently, the study of crowd modelling has stimulated a renewed interest for these equations. Indeed, pedestrians have a long-range perception of their space, and thus choose their path based on long-range interactions. While such interactions do not enjoy action-reaction properties which are typical in physical models, methods connected to mean-field limit approaches have shown their adaptability in this setting too (see e.g. [25, 61, 104, 105]). More generally, the study of other kind of interacting agents, such as opinion dynamics on networks [26, 91], or animal flocks [23, 62], has been conveyed with similar techniques.

As it will be highlighted in Chapter 1, the most natural framework in which one can rigorously derive these mean-field convergence properties and study the resulting equations is the *space of probability measures* endowed with the *Wasserstein distance* (see e.g. [17, 125, 132]). This class of metric coming from optimal transport theory have proved very powerful in studying mean-field approximations and continuity equations in general. Indeed, such equations have an intrinsic link with the geometric structure of the space of measures endowed with the Wasserstein metric (see e.g. [15] and [17, Chapter 8]), and generalize the notion of ODE in this infinite-dimensional framework.

Controlled multi-agent dynamics

As explained previously, one may wish to stir a microscopic multi-agent system such as (1) towards a desired configuration or steady state. This can be achieved as in the usual setting of control theory via the addition of control functions to the right-hand side of the equation. In the framework of this manuscript, we focus on linearly controlled dynamics of the form

$$\begin{cases} \dot{x}_i(t) = \mathbf{v}_N[\mathbf{x}(t)](t, x_i(t)) + u_i(t), \\ x_i(0) = x_i^0, \end{cases} \quad (3)$$

where $u_1(\cdot), \dots, u_N(\cdot) \in L^\infty(\mathbb{R}_+, U)$ are control functions. The conceptual model behind (3) is the following : a central planner is designing a control law $u_i(\cdot)$ adapted to each agent $x_i(\cdot)$ in order to achieve a given goal (reaching a desired configuration, maximizing a certain pay-off, etc...).

Even though practical and useful to some extent in multi-agent modelling (see e.g. Chapter 5), systems such as (3) are not completely satisfactory from a conceptual standpoint. Indeed, when N is very large, it does not make much sense to design individual control laws for each agent. It would be better for the central planner to design a general function $u : (t, x) \in [0, T] \times \mathbb{R}^d \rightarrow U$ once and for all and to apply it to each agent.

While it may take some care to make this intuition rigorous – especially when the control laws are not smooth with respect to the space variable –, this idea is the most

2. One could more generally choose $v[\cdot](\cdot, \cdot)$ as the limit of a suitable sequence of maps $(v_N[\cdot](\cdot, \cdot))$ which are compatible with $\mathbf{v}_N[\cdot](\cdot, \cdot)$ for each N .

natural one when studying multi-agent systems. In the mean-field limit framework, the equivalent formulation of (3) is given by *controlled continuity equation* of the form

$$\partial_t \mu(t) + \nabla \cdot ((v[\mu(t)])(t, \cdot) + u(t, \cdot))\mu(t) = 0, \quad (4)$$

where $u(\cdot, \cdot)$ is a Borel velocity field satisfying some adequate integrability properties along the curve $\mu(\cdot)$. There is a large literature devoted to the well-posedness of such equations, see e.g. [11, 12, 13, 68] and references therein.

In this manuscript, we will extensively study control and optimal control problems formulated on continuity equations. Apart from a few recent results about controllability [73, 74], most of the contributions in this direction have considered optimal control problems, i.e. the minimization of a functional where the constraint is a controlled dynamics. Applications of these problems are of great interest, canonical examples being provided e.g. by the minimal escape time problem for a crowd [10, 77, 106] or the enforcement of consensus in a network by minimizing the variance of the opinions (see e.g. [43, 44, 115]). Existence of optimal controls has been investigated in [31, 80, 81, 82], mostly via Γ -convergence techniques, as well as in the setting of mean-field control [1, 2]. While Hamilton-Jacobi optimality conditions in Wasserstein spaces have received some attention, see e.g. the seminal paper [85] and recent developments in the field of control theory [50, 51], Pontryagin optimality conditions remain rather unexplored. A first result in this direction was presented in [31], where a coupled PDE-ODE system in which the control acts on the ODE part only was studied. This result was the first step towards our more refined developments [34, 35], presented in Chapter 2 and 3.

In the remainder of the introduction, we summarize the main contributions presented in the different chapters of this manuscript.

Chapter 1 : Analysis in measure spaces and optimal transport theory

In Chapter 1, we present preliminary notions pertaining to measure theory and more specifically to optimal transport. It will mostly contain state-of-the-art material borrowed from [14, 17]. Some original results dealing with the characterization of the differential of non-local flows exposed in Section 1.4.

The theory of optimal transportation was first introduced by Gaspard Monge in his “*Mémoire sur la Théorie des Déblais et des Remblais*” [109] in 1781. Written in more modern terms, the Monge problem is the following: given two probability measures $\mu, \nu \in \mathcal{P}_c(\mathbb{R}^d)$ (which represent piles of concrete) and a cost function $c : \mathbb{R}^d \times \mathbb{R}^d \rightarrow [0, +\infty]$ (which represents the cost to pay to move from point A to point B), find a Borel map $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that

$$\int_{\mathbb{R}^d} c(x, T(x)) d\mu(x) = \inf_{T'} \left\{ \int_{\mathbb{R}^d} c(x, T'(x)) d\mu(x) \text{ s.t. } T'_\# \mu = \nu \right\}, \quad (5)$$

where $T'_\# : \mathcal{P}(\mathbb{R}^d) \rightarrow \mathcal{P}(\mathbb{R}^d)$ denotes the pushforward operation through T' .

There are several issues with (5). First, the set of admissible transport maps satisfying the image-measure condition $T'_\# \mu = \nu$ may be empty, for instance if $\mu = \delta_0$ and $\nu = \mathcal{L}_{[0,1]^d}$. Second, even in the case where transport maps do exist, the feasible set is not closed or convex in any nice topology, which prevents the application of the direct method of calculus of variations to recover the existence of minimizers. Therefore, the

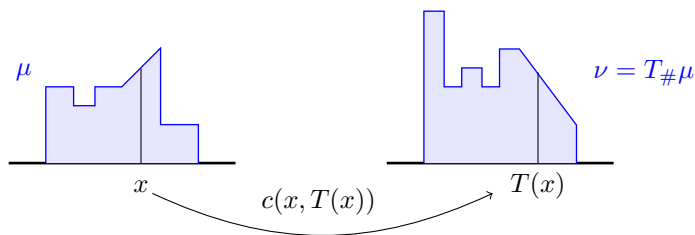


Figure 2: *Illustration of the Monge formulation of the optimal transport problem*

fascinating problem formulated by Monge remained mostly unstudied – at the exception of some attempts made during the 19th century, see e.g. [125, Chapter 1] – for a long time due to a deficit in mathematical artillery.

In 1942, the Russian mathematician Leonid Kantorovich – who was then working on the premises of linear programming – introduced a relaxation of Monge’s problem in [97]. In his new formulation, he aimed at finding an optimal transport plan $\gamma \in \mathcal{P}_c(\mathbb{R}^{2d})$ solution of

$$\int_{\mathbb{R}^{2d}} c(x, y) d\gamma(x, y) = \inf_{\gamma'} \left\{ \int_{\mathbb{R}^{2d}} c(x, y) d\gamma(x, y) \text{ s.t. } \pi_{\#}^1 \gamma = \mu \text{ and } \pi_{\#}^2 \gamma = \nu \right\}. \quad (6)$$

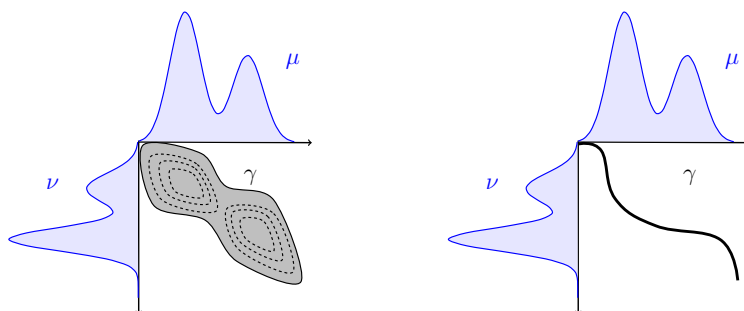


Figure 3: *Illustration of the Kantorovich formulation of the optimal transport problem: a generic transport plan (left) and an optimal one concentrated on a graph (right)*

It turns out that (6) is a linear optimization problem over a weakly-closed admissible set, so that it always admit solutions under mild assumptions on the cost function $c(\cdot, \cdot)$ (see e.g. [17, 125, 132]). By taking in particular $c(x, y) = |x - y|^p$ for a given $p \geq 1$, one can define a distance between probability measures, called the Wasserstein distance.

The development of optimal transport and the use of Wasserstein spaces in general have had enormous repercussions throughout modern analysis. It is at the center of fields as diverse as the study of dissipative PDEs [17, 87, 96, 112], metric geometry [18, 101, 130], mean-field games theory [45, 46, 98], control theory [34, 35, 50, 51, 115], imaging [55, 121, 128] and machine-learning [19, 56]. In Chapter 1, we will recall several results of this theory, with a strong emphasis on notions pertaining to the weak-Riemannian and differential structures of these spaces introduced in [112] and formalized later on in [17], as well as on the special geometric role played by continuity equations. We provide at the end of Chapter 1 a multi-dimensional chainrule formula along with characterizations of first-order variations of non-local characteristic flows which did not

exist up to now in the literature.

Chapter 2 : The Pontryagin Maximum Principle in the Wasserstein space

In Chapter 2, we establish a Pontryagin Maximum Principle for unconstrained optimal control problems of continuity equations, written in the general form

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

Here $(t, x) \mapsto v[\mu](t, x)$ is a non-local velocity field, while the maps $(t, \mu, \omega) \mapsto L(\mu, \omega)$ and $\mu \mapsto \varphi(\mu)$ represent a running and a final cost respectively. The control $(t, x) \mapsto u(t, x)$ is a velocity field which is chosen among a set $\mathcal{U} = L^\infty([0, T], U)$ of admissible controls where $U \subset C^1(\mathbb{R}^d, \mathbb{R}^d)$ is compact in the C^0 -topology. The choice of essentially bounded admissible controls is very natural in many applications, and it prevents the occurrence of Lavrentiev-type phenomena [100]. This phenomenon, which is well-known in calculus of variations, has the drawback of invalidating the first-order necessary optimality conditions unless some further structural assumptions are imposed on the problem (see e.g. [59, Chapter 22] and [133, Chapter 11] for detailed accounts in the finite-dimensional case).

The Pontryagin Maximum Principle is a fundamental tool to study optimal control problems. It allows to derive regularity and stability estimates on the optimal controls [58, 70, 83, 90, 93] and to compute explicitly optimal syntheses [8, 36]. It is also broadly used in conjunction with the so-called *shooting methods* for numerical optimal trajectory planning, see e.g. [30, 32, 33]. Let it be noted that, in essence, the only structural assumption required to write a PMP is that the ambient space is endowed with a manifold structure. As it will be shown more explicitly in Chapter 1, this structural prerequisite is available in the setting of Wasserstein spaces. By adapting several tools and ideas of geometric control to the metric spaces of probability measures, we are able to derive the following result, stated precisely in Theorem 2.1.

Theorem (Pontryagin Maximum Principle for (\mathcal{P})). *Let $(u^*(\cdot), \mu^*(\cdot))$ be an optimal pair control-trajectory for (\mathcal{P}) and assume that hypotheses **(H)** of Theorem 2.1 hold. Then, there exist a constant $R_T > 0$ and a curve $\nu^*(\cdot) \in \text{Lip}([0, T], \mathcal{P}(\overline{B_{2d}(0, R_T)}))$, Lipschitz with respect to the W_1 -metric, which satisfies the following conditions.*

(i) *It solves the forward-backward system of continuity equations*

$$\begin{cases} \partial_t \nu^*(t) + \nabla \cdot \left(\mathbb{J}_{2d} \tilde{\nabla}_\nu \mathbb{H}(t, \nu^*(t), u^*(t)) \nu^*(t) \right) = 0 \text{ in } [0, T] \times \mathbb{R}^{2d}, \\ \pi_{\#}^1 \nu^*(t) = \mu^*(t), \\ \pi_{\#}^2 \nu^*(T) = (-\bar{\gamma}_\varphi^\circ)_{\#} \mu^*(T), \end{cases}$$

where $\bar{\gamma}_\varphi^\circ(\cdot)$ is the barycenter of the minimal selection $\partial^\circ \varphi(\mu^*(T))$ of the final cost $\varphi(\cdot)$ at $\mu^*(T)$ and \mathbb{J}_{2d} is the symplectic matrix in \mathbb{R}^{2d} .

The infinite-dimensional Hamiltonian of the system is defined by

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

for any $(t, \nu, \omega) \in [0, T] \times \mathcal{P}_c(\mathbb{R}^{2d}) \times U$, and the vector field $\tilde{\nabla}_{\nu} \mathbb{H}_c(t, \nu^*(t), u^*(t))(\cdot, \cdot)$ is “almost” the Wasserstein gradient of this Hamiltonian map with respect to the differential structure of $(\mathcal{P}_2(\mathbb{R}^{2d}), W_2)$.

(ii) The Pontryagin maximization condition

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

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

The derivation of Pontryagin optimality conditions for problems of the general form (\mathcal{P}) remained open during a long time. The first contribution in the setting described in this thesis was obtained in [31], where the authors derived a Pontryagin Maximum Principle for a multi-scale problem similar to (\mathcal{P}) but where the control only acted on a finite-dimensional population of *leaders*. The proof of this result was based on a combination of mean-field limit and Γ -convergence arguments, and was not straightforwardly generalizable to the more difficult case of (\mathcal{P}) in which the control $u(\cdot, \cdot)$ is a vector field driving the measure $\mu(\cdot)$. In [118], the author proved a Pontryagin-type optimality condition, but for simpler instances of (\mathcal{P}) , and again for controls that were only finite-dimensional.

Among the several strategies to prove the PMP, the technique of *needle-like variations*, introduced as early as [119], was a promising approach to generalize this result to the setting of Wasserstein spaces where other methods based on convex duality or mean-field limit failed. A needle-like variation $\tilde{u}_{\epsilon}(\cdot)$ of an optimal control $u^*(\cdot) \in L^{\infty}([0, T], U)$ is a pointwise perturbation which is defined by

$$\tilde{u}_{\epsilon} : t \in [0, T] \mapsto \begin{cases} u^*(t) & \text{if } t \in [\tau - \epsilon, \tau], \\ \omega & \text{otherwise,} \end{cases}$$

where $\epsilon > 0$ is a small parameter, $\tau \in [0, T]$ is a Lebesgue point of $t \mapsto v[\mu^*(t)](t, \cdot) + u^*(t, \cdot)$ and $\omega \in U$ is an arbitrary admissible control value. This class of perturbation allows for tractable computation of first-order perturbations in the state and can be performed without any a priori assumption on the set of control values U .

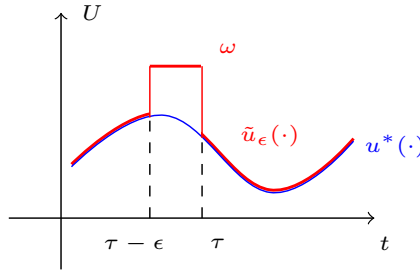


Figure 4: Illustration of a needle-like perturbation of the optimal control $u^*(\cdot)$

The outline of the proof presented in Chapter 2 is the following. We start by computing the first-order variation of the total cost generated by the needle variation using the non-linear semigroup structure of continuity equations. Afterwards, we derive a family of first-order necessary optimality conditions expressed via the differential structure of Wasserstein spaces. We then build a suitable state-costate measure curve solving the Wasserstein Hamiltonian flow, and we finally show that the family of first-order condition can be rewritten as the pointwise maximization condition along the integral curves of this flow.

Chapter 3 : A Pontryagin Maximum Principle for Constrained Optimal Control Problems in the Wasserstein Space

In Chapter 3, we present an extension of the Pontryagin maximum principle in Wasserstein spaces described hereabove to the case of constrained optimal control problems. Such problems can be written in the general form

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

In this setting, the maps $(t, \mu) \mapsto \Lambda(t, \mu)$ and $\mu \mapsto \Psi^I(\mu), \Psi^E(\mu)$ respectively represent running and end-point constraints, the latter incorporating both equality and inequality constraints.

Constraints appear very naturally in the modelling of control problems. End-point constraints are needed e.g. in shortest path and planning problems [36] or when aiming to drive the controlled system in a desired set of final configurations described by a manifold [8, Chapter 12]. On the other hand, running constraints frequently appear in classical control theory where the state variables are confined to time-varying domains, often depending on the control [30]. Constrained problems have also many applications in multi-agent modelling. For instance, end-point constraints arise in the optimal control formulation of the Wasserstein geodesic problem due to Benamou and Brenier [27] or in exit minimal-time problems for crowds [51]. Besides, several papers deal with density constraints in the framework of mean-field games theory [99, 120].

To generalize the Wasserstein PMP to $(\mathcal{P}_{\text{const}})$, we choose to follow a modern methodology proposed in [22], based again on needle-like variations. This proof strategy was further refined in [126] to incorporate state constraints and non-smooth functionals. By adapting and generalizing the ideas introduced in these papers to the setting of Wasserstein, we proved the following result which is presented in details in Theorem 3.1 in Chapter 3.

Theorem (Pontryagin Maximum Principle for $(\mathcal{P}_{\text{const}})$). *Let $(u^*(\cdot), \mu^*(\cdot))$ be an optimal pair control-trajectory for $(\mathcal{P}_{\text{const}})$ and assume that the hypotheses **(H)** of Theorem 3.1 hold. Then there exists a constant $R'_T > 0$, a family of Lagrange multipli-*

ers $(\lambda_0, \dots, \lambda_n, \eta_1, \dots, \eta_m, \varpi_1, \dots, \varpi_l) \in \{0, 1\} \times \mathbb{R}_+^n \times \mathbb{R}^m \times \mathcal{M}_+([0, T])^r$ and a curve $\nu^*(\cdot) \in \text{Lip}([0, T], \mathcal{P}(B_{2d}(0, R_T)))$ such that the following holds.

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

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

where \mathbb{J}_{2d} is the symplectic matrix of \mathbb{R}^{2d} . The augmented infinite-dimensional Hamiltonian $\mathcal{H}(\cdot, \cdot, \cdot, \cdot)$ of the system is defined by

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

The penalized state constraints and final gradient maps are given respectively by

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

and

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

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

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

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

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

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

There are two main difficulties to overcome in order to adapt the results described in Chapter 2 to constrained problems. The first one is in some sense purely technical, and

has to do with the needle-variations approach. Because of the presence of constraints, one cannot simply write the optimality condition on the length of the needle parameter as the non-negativity of the derivative total cost, which was possible for proving the unconstrained PMP. Instead, one needs to resort to a Lagrange multiplier rule, which multipliers will structurally depend on the needle parameters. To get around this difficulty, the authors proposed in [22] to consider multi-dimensional perturbations and to get rid of this dependence via a limiting process as the number of perturbations goes to infinity.

The second main issue is deeper. It is proven in [17, Theorem 8.3.1] that solutions of continuity equations coincide exactly with absolutely continuous curves of measures in Wasserstein spaces. Yet in the classical formulation of the maximum principle with state-constraints, the costate variables are merely BV with respect to time. In order to be able to write a Wasserstein Hamiltonian flow in the sense of [15, 35], we therefore chose to prove a modified PMP introduced by R.V. Gamkrelidze in [84] (we refer the reader to [21] for a modern treatise on this topic). In this formulation, the natural BV -in-time dependence induced by the dualization of the state-constraints is placed on an extra multiplier. The corresponding costates are absolutely continuous in time, which allows to express the state-costate pairs as the solutions of Wasserstein Hamiltonian flows.

Chapter 4 : Intrinsic Lipschitz Regularity of Mean-Field Optimal Controls

In Chapter 4, we present an intrinsic Lipschitz regularity result for mean-field optimal control problems. As already mentioned in the introduction, Cauchy-Lipschitz regularity is a most desirable property for the vector fields driving continuity equations. It ensures classical well-posedness and uniqueness, and provides a powerful representation formula for the solution as the pushforwards of the initial datum via the characteristic flow, see [13, Part I] and [17, Section 8.1]. Moreover, it is known in the literature that a strict micro-macro correspondence of the solutions of continuity equations – which is very convenient in applications –, is only available under Cauchy-Lipschitz regularity assumptions (see e.g. [11] and the comments in [78, Theorem 4.1.1]).

Yet, it has been well-known that solutions of Wasserstein optimal control problems are not generically Lipschitz, or even continuous or Sobolev for that matter. Indeed, an astonishingly vast and sophisticated literature has been devoted to the regularity of the solution of Monge’s optimal transport problem (see e.g. [66, 67, 79]) mostly via PDE techniques. However, few of these results can be translated into regularity properties on the optimal tangent velocity field $v^*(\cdot, \cdot)$ solving the Benamou-Brenier problem

$$(\mathcal{P}_{\text{BB}}) \quad \begin{cases} \min_{v \in L^2} \left[\int_0^T \int_{\mathbb{R}^d} \frac{1}{2} |v(t, x)|^2 d\mu(t)(x) dt \right] \\ \text{s.t.} \begin{cases} \partial_t \mu(t) + \nabla \cdot (v(t, \cdot) \mu(t)) = 0, \\ \mu(0) = \mu^0, \quad \mu(T) = \mu^1. \end{cases} \end{cases}$$

with arbitrary measures μ^0, μ^1 . For this optimal control problem in Wasserstein spaces, Caffarelli proved in [40] that $v(t, \cdot) \in C_{\text{loc}}^{k-1, \alpha}(\mathbb{R}^d, \mathbb{R}^d)$ for some $\alpha \in (0, \bar{\alpha})$ whenever $\mu^0, \mu^1 \in \mathcal{P}^{ac}(\mathbb{R}^d)$ have densities in $C_{\text{loc}}^{k, \alpha}(\mathbb{R}^d, \mathbb{R}^d)$ with respect to the d -dimensional Lebesgue measure.

Another context in which the regularity of mean-field optimal controls has been

(indirectly) investigated is that of mean-field games theory. Indeed, there is a large literature devoted to the regularity of the value function $(t, x) \mapsto u^*(t, x)$ solving forward-backward system of Hamilton-Jacobi and continuity equations

$$\begin{cases} \partial_t u(t, x) + H(t, u(t, x), D_x u(t, x)) = f(t, x), & u(T, x) = g_T(x), \\ \partial_t \mu(t) - \nabla \cdot (\nabla_p H(t, u(t, x), D_x u(t, x)) \mu(t)) = 0, & \mu(0) = \mu^0. \end{cases}$$

We refer the reader e.g. to [47] for Sobolev regularity results and to [48] for Hölderian regularity properties. In the setting of potential mean-field games, the tangent velocity field $v^*(t, x) = -\nabla_p H(t, u^*(t, x), D_x u^*(t, x))$ is the optimal control associated to a mean-field optimal control problem. Therefore, regularity properties of the optimal control can be recovered from that of the optimal value function, and are expected to have one order of differentiation fewer with respect to the latter.

In our work, we consider general smooth and unconstrained optimal control problems of the form

$$(\mathcal{P}) \quad \begin{cases} \min_{u \in \mathcal{U}} \left[\int_0^T \left(L(t, \mu(t)) + \int_{\mathbb{R}^d} \psi(u(t, x)) d\mu(t)(x) \right) dt + \varphi(\mu(T)) \right] \\ \text{s.t.} \quad \begin{cases} \partial_t \mu(t) + \nabla \cdot ((v[\mu(t)](t, \cdot) + u(t, \cdot)) \mu(t)) = 0, \\ \mu(0) = \mu^0. \end{cases} \end{cases}$$

where $u(\cdot, \cdot)$ is a priori an element of $L^\infty([0, T] \times \mathbb{R}^d, U; \tilde{\mu})$ with $\tilde{\mu} \in \mathcal{P}([0, T] \times \mathbb{R}^d)$ denoting the Young measure lift of the curve $\mu(\cdot)$. Here, $\psi(\cdot)$ is a strictly convex and $C_{\text{loc}}^{2,1}$ -regular map. All the datum of the problem are supposed to be smooth, namely at least twice continuously differentiable both with respect to space and measure variables. In this setting, we were able to derive the following result stated precisely in Theorem 4.1.

Theorem (Existence of Lipschitz-in-space optimal controls for (\mathcal{P})). *Let $\mu^0 \in \mathcal{P}_c(\mathbb{R}^d)$ and assume that hypotheses **(H)** and **(CO_N)** of Theorem 4.1 hold. Then, there exists an optimal pair control-trajectory $(u^*(\cdot, \cdot), \mu^*(\cdot)) \in \mathcal{U} \times \text{Lip}([0, T], \mathcal{P}_c(\mathbb{R}^d))$ for problem (\mathcal{P}) . Moreover, the map $x \in \mathbb{R}^d \mapsto u^*(t, x) \in U$ is \mathcal{L}_U -Lipschitz for \mathcal{L}^1 -almost every $t \in [0, T]$, where the uniform constant \mathcal{L}_U only depends on the datum of the problem (\mathcal{P}) .*

In order to obtain this result, we start by invoking a general existence result of mean-field optimal controls for (\mathcal{P}) derived in [80]. Then, we prove that for each of the approximating finite-dimensional multi-agent optimal control problem, there exists an optimal Lipschitz map. To do so, we invoke the main Theorem of [71] which is itself based on a comprehensive study of metrically regular differential generalized equations performed in [58]. These results are essentially quantitative strong metric regularity arguments, which can be carried out under a global coercivity condition on our problem (see hypothesis **(CO_N)** in Chapter 4). This assumptions, which is inspired from the seminal work [90], can be seen as a second-order sufficient optimality condition for a suitable linearized problem in a neighbourhood of an optimal pair control-trajectory. It is then possible to obtain a Lipschitz regular optimal control for (\mathcal{P}) in the limit by standard compactness arguments.

Chapter 5 : Convergence Analysis and Sparse Control of Weakly Cooperative Alignment Models

In Chapter 5, we turn our attention to the analysis and control of consensus and alignment models. In Section 5.1, we prove convergence towards consensus and flocking for systems subject to random communication failures. To do so, we require an assumption of persistence of excitations expressed on the Laplacian matrix describing the system. In Section 5.2, we build an explicit control strategy enabling to drive weakly cooperative kinetic systems towards approximate alignment.

Consensus and Flocking in Cooperative Systems with Random Communication Failures

Among the family of discrete models for multi-agents described in this introduction, a special class of interest is that of cooperative systems (see e.g. [127]). These systems can be written as

$$\dot{x}_i(t) = \frac{1}{N} \sum_{j=1}^N \phi(x_i(t) - x_j(t)), \quad x_i(0) = x_i^0, \quad (7)$$

for first-order dynamics and

$$\begin{cases} \dot{x}_i(t) = v_i(t), & x_i(0) = x_i^0, \\ \dot{v}_i(t) = \frac{1}{N} \sum_{j=1}^N \psi(x_i(t) - x_j(t), v_i(t) - v_j(t)) \end{cases} \quad (8)$$

for second-order ones. Here, the ‘‘cooperative’’ part comes from the extra assumption that in both case, $\phi(x)$ and $\psi(x, v)$ are non-positively colinear to x and v respectively. In the case where the interaction kernel is strictly positive, the system is said to be strongly cooperative, otherwise it is said to be weakly cooperative.

First-order cooperative systems are instrumental in the modelling of consensus behaviour [61, 91], while second-order cooperative models have gained a lot of steam since the seminal works of Cucker and Smale [62, 63] on flocking behaviours. They are also used to design centralized control laws for large robot swarms, see e.g. [60], and frequently appear in graph theoretic analysis of discrete automaton systems [24].

While the picture is fairly well understood for consensus and flocking behaviours in the strongly cooperative setting [49, 89], the weakly cooperative one is still largely open. In particular, little works to this day take into account possible communication failures among agents. The seminal contribution [110] in the setting of discrete-state systems has been extended over the years, but few progresses have been made on general continuous multi-agent systems and even less on their kinetic approximations.

In this context, we were able to prove a convergence towards consensus result for randomly failed first order systems, which is described in Theorem 5.1. We were also able to obtain the following more delicate flocking result for weakly-cooperative second order systems, stated in Theorem 5.2.

Theorem (Flocking for randomly failed Cucker-Smale systems). *Consider the randomly*

failed second-order Cucker-Smale system

$$\begin{cases} \dot{x}_i(t) = v_i(t), \\ \dot{v}_i(t) = \frac{1}{N} \sum_{j=1}^N \xi_{ij}(t) \phi(|x_i(t) - x_j(t)|) (v_j(t) - v_i(t)), \end{cases} \quad (\text{CSM}_2)$$

where $\xi_{ij} \in L^\infty(\mathbb{R}_+, [0, 1])$ are symmetric weights satisfying the (τ, μ) -persistence assumption

$$\frac{1}{2N^2} \sum_{i,j=1}^N \int_t^{t+\tau} \xi_{ij}(s) dt |v_i - v_j|^2 \geq \frac{\mu}{2N^2} \sum_{i,j=1}^N |v_i - v_j|^2, \quad (9)$$

for \mathcal{L}^1 -almost every $t \in \mathbb{R}_+$ and any $\mathbf{v} \in (\mathbb{R}^d)^N$, and where $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is a non-increasing, non-summable function such that

$$\int_0^{+\infty} \exp\left(-\sqrt{r} \phi \circ \Phi^{-1}(\sqrt{r})\right) dr < +\infty, \quad (10)$$

where $\Phi(X) = \int_{X(0)}^X \phi(r) dr$. Then every solution of (CSM_2) converges asymptotically to flocking, i.e.

$$\sup_{t \geq 0} |x_i(t) - \bar{\mathbf{x}}(t)| < +\infty, \quad \lim_{t \rightarrow +\infty} |v_i(t) - \bar{\mathbf{v}}(t)| = 0.$$

In the statement of this result, the maps $\xi_{ij}(\cdot)$ are weights that can vanish on possibly large time intervals, and which account for the possible communication failures between agents. Our proof of this result is based on a novel combination of the Lyapunov methods introduced in [89] with a *persistence of excitations* hypothesis (see e.g. [53, 54, 102, 103]) expressed in terms of graph Laplacian matrices [41, 44]. We also provide connexions between our hypotheses and existing sufficient convergence conditions expressed in terms of the average connectedness of the interaction graph of the system.

This result illustrates the fact that general convergence to flocking is not granted for weakly cooperative systems. Indeed, the persistence assumption (9) can be interpreted as a lower bound on the average interaction of the agents which are not yet aligned, and this for every time window of the form $[t, t + \tau]$. Moreover, this result is currently limited to the so-called “strong interaction” scenario in which $\phi \notin L^1(\mathbb{R}_+, \mathbb{R}_+)$. Our method is also limited by the extra constraint given in (10), which is satisfied e.g. when $\phi(r) \sim \frac{1}{r^\beta}$ at infinity with $\beta < \frac{1}{2}$, which is not the natural critical exponent to be expected in this setting. In the deterministic case, some sharp bounds are needed on the initial position and velocity supports for flocking to emerge in the case where $\phi \in L^1(\mathbb{R}_+, \mathbb{R}_+)$ (see e.g. [63, 89, 115]).

Sparse Control of Kinetic Cooperative Systems to Approximate Alignment

The results previously stated illustrate that consensus and flocking may likely fail to arise in weakly-cooperative systems. Whence, several papers have been aiming at designing control laws steering weakly cooperative systems towards flocking [41, 115, 116]. A lot of attention has moreover been devoted to ensure that these control strategies are sparse, i.e. that they act on few agents at each time in a sense to be specified. With this goal in mind, we would like to design a sparse control strategy steering any kinetic

weakly cooperative system of the form

$$\partial_t \mu(t) + \nabla \cdot \left(\left(v, \Psi[\mu(t)](x, v) + \chi_{\omega(t)} u(t, x, v) \right) \mu(t) \right) = 0, \quad (11)$$

towards ϵ -approximate alignment around a prescribed velocity $v^* \in \mathbb{R}^d$. Mathematically speaking, this condition can be expressed as

$$\text{supp}(\mu(t)) \subset \mathbb{R}^d \times B(v^*, \epsilon)$$

for a given $M > 0$ and for all times $t \geq T_c$ where $T_c \geq 0$ is the time horizon over which we control the system. Remark that (11) is precisely the kinetic version of (8). Expanding the methodologies developed in [115, 116], we obtained the following result, stated in Theorem 5.3.

Theorem. *Let $\mu^0 \in \mathcal{P}_c^{ac}(\mathbb{R}^d)$ be a given initial data for (11). For any constant $c > 0$, any limit velocity $v^* \in \mathbb{R}^d$ and any precision $\epsilon > 0$, there exists a time T_c and a Lipschitz-in-space control strategy $u(\cdot, \cdot, \cdot)$ which support $\omega(\cdot)$ satisfies the constraints*

$$\|u(t, \cdot, \cdot)\|_{L^\infty(\mathbb{R}^d \times \mathbb{R}^d)} \leq 1, \quad \int_{\omega(t)} d\mu(t)(x) \leq c,$$

such that the corresponding solution $\mu \in C^0([0, +\infty), \mathcal{P}_c^{ac}(\mathbb{R}^d \times \mathbb{R}^d))$ of (11) is approximately aligned around v^* with precision ϵ for all $t \geq T_c$.

The proof strategy is to show that control towards alignment can be performed component by component and around the origin, as a consequence of invariance properties of cooperative systems. Then, we build a piecewise constant control strategy which brings any 1-dimensional cooperative system to approximate alignment in a finite number of steps, and show how the latter can then be generalized to arbitrary dimensions.

Chapter 6 : Generic Singularities of the 3D-Contact Sub-Riemannian Conjugate Locus

In Chapter 6, we move to a relatively different topic. We present a result about the generic singularities of the conjugate locus for contact sub-Riemannian distributions in \mathbb{R}^3 .

A sub-Riemannian structure over a manifold M is given by the datum of a distribution $\Delta \subset TM$ such that $\dim(\Delta(q)) = m < n$ at any point $q \in M$, along with a Riemannian metric \mathbf{g} defined over Δ . The distribution $\Delta(q)$ at a point $q \in M$ represents a prescribed set of admissible directions which a curve passing through q can follow. Curves satisfying these non-holonomic constraints are called admissible, and they can be written as solutions of a control problem of the form

$$\dot{\gamma}(t) = \sum_{k=1}^m u_k(t) X_k(\gamma(t))$$

where $X_1(\cdot), \dots, X_m(\cdot)$ are smooth vector field spanning Δ and $(u_1(\cdot), \dots, u_m(\cdot)) \in L^2([0, T], \mathbb{R})$ are control functions.

Under the well-known Hörmander condition (see e.g. [5, 8]), any pair of points $q_0, q_1 \in M$ can be connected by such an admissible curve, and it is in particular possible to search for the shortest admissible path between two points. However, since Δ is a strict sub-bundle of TM , the admissible geodesics exhibit uncanny behaviours which do not appear in classical Riemannian geometry. In particular, the sub-Riemannian distance function usually fails to be smooth, even in an arbitrarily small neighbourhood of its base point. This property gives rise to many singularities along several geometric objects generated by the sub-Riemannian distribution, such as the wave fronts and balls.

In this context, we focus on the generic structure of the caustics for 3-dimensional contact distributions. The caustic is defined in the context of 3D-contact sub-Riemannian geometry as the critical value set of the exponential map starting from a given point, or in other words as the image of the *conjugate locus* through the exponential map. The classification of these singularities was initiated in [3] and furthered in [4, 52] and [6], but it was only done for the upper and lower semi-caustics. In Theorem 6.1, we complete this classification and extend it to the whole caustic. This result can be summarized as follows in a non-technical way.

Theorem (Generic singularities of the 3D-contact sub-Riemannian conjugate locus). *Let M be a 3-dimensional smooth and connected manifold. For a generic 3D-contact sub-Riemannian distribution over M , the following holds.*

- (i) *There exists a smooth curve $\mathcal{C} \subset M$ such that, outside \mathcal{C} , the intersections of the caustic with neighbouring horizontal planes are the closed curves exhibiting 4 cusp points exhibited in [3].*
- (ii) *On \mathcal{C} , the intersections of the upper and lower semi-caustics with horizontal planes are the closed curves with 6 cusp points exhibited in [6]. Moreover, the semi-caustics are independent, and the only obstructions to the pairing of two families of singularities can be expressed in terms of codimension in the corresponding space of jet coefficients.*

The proof of this result relies on three key steps. First, one uses sub-Riemannian normal coordinates, introduced in [6, 52], which allow to express any 3D-contact sub-Riemannian distribution as a smooth perturbation of the Heisenberg group. Then, one expresses the different singularities of the conjugate locus in the form of a suspension in a privileged set of coordinates. Finally, one resorts to arguments of transversality theory to classify the genericity of these singularities in terms of codimension in the space of jet coefficients of the equations that they must verify.

1 Analysis in the Space of Measures and Optimal Transport Theory

In this chapter, we recall useful notions that will be used extensively in the manuscript, and which pertain mainly to analysis in measure spaces and optimal transport theory. Most of the results stated in this chapter are state of the art which is mostly borrowed from [14, 17, 125]. Let it be noted however that the results of Proposition 1.8 and Section 1.4 are new and were first written in [34, 35].

1.1 Elements of measure theory and Wasserstein spaces

In this section, we introduce some general tools of measure theory and proceed to recall some of the main definitions of optimal transport theory and Wasserstein spaces. We refer the interested reader to [14] for the former and to [17, Chapter 5-6] for the latter.

1.1.1 Elementary notions of measure theory

We denote by $(\mathcal{M}(\mathbb{R}^d, \mathbb{R}^m), \|\cdot\|_{TV})$ the Banach space of m -dimensional vector-valued Borel measures defined on \mathbb{R}^d endowed with the total variation norm, defined by

$$\|\nu\|_{TV} = |\nu|(\mathbb{R}^d) \equiv \sup \left\{ \sum_{k=1}^{+\infty} |\nu(E_k)| \text{ s.t. } E_k \text{ are disjoint Borel sets and } \bigcup_{k=1}^{+\infty} E_k = \mathbb{R}^d \right\},$$

for any $\nu \in \mathcal{M}(\mathbb{R}^d, \mathbb{R}^m)$. By Riesz Theorem (see e.g. [14, Theorem 1.54]), this space can be identified with the topological dual of the Banach space $(C_c^0(\mathbb{R}^d, \mathbb{R}^m), \|\cdot\|_{C^0})$ of continuous and compactly supported functions endowed with the duality bracket defined by

$$\langle \nu, \phi \rangle_{C^0} = \sum_{k=1}^m \int_{\mathbb{R}^d} \phi_k(x) d\nu_k(x), \quad (1.1)$$

for any $\nu \in \mathcal{M}(\mathbb{R}^d, \mathbb{R}^m)$ and any $\phi \in C_c^0(\mathbb{R}^d, \mathbb{R}^m)$.

Given a Borel measure $\nu \in \mathcal{M}(\mathbb{R}^d, \mathbb{R}_+)$ and a real number $p \in [1, \infty]$, we denote respectively by $L^p(\Omega, \mathbb{R}^m; \nu)$ and $W^{1,p}(\Omega, \mathbb{R}^m; \nu)$ the corresponding spaces of p -integrable and p -Sobolev functions defined over a subset $\Omega \subset \mathbb{R}^d$ with values in \mathbb{R}^m . In the case where $\nu = \mathcal{L}^d$ is the standard d -dimensional Lebesgue measure, we simply denote these spaces by $L^p(\Omega, \mathbb{R}^m)$ and $W^{1,p}(\Omega, \mathbb{R}^m)$.

We denote by $\mathcal{P}(\mathbb{R}^d) \subset \mathcal{M}(\mathbb{R}^d, \mathbb{R}_+)$ the set of *Borel probability measures* and for $p \geq 1$, we define $\mathcal{P}_p(\mathbb{R}^d) \subset \mathcal{P}(\mathbb{R}^d)$ the subset 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 $\nu \in \mathcal{M}(\mathbb{R}^d, \mathbb{R}^m)$ is defined as the closed set $\text{supp}(\nu) = \{x \in \mathbb{R}^d \text{ s.t. } \nu(\mathcal{N}) \neq 0 \text{ for any neighbourhood } \mathcal{N} \text{ of } x\}$. We denote by $\mathcal{P}_c(\mathbb{R}^d) \subset \mathcal{P}(\mathbb{R}^d)$ the subset of Borel 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 $\mu_n \rightharpoonup^* \mu$ – provided that

$$\int_{\mathbb{R}^d} \phi(x) d\mu_n(x) \xrightarrow{n \rightarrow +\infty} \int_{\mathbb{R}^d} \phi(x) d\mu(x), \quad (1.2)$$

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 notion of convergence induces a topology over $\mathcal{P}(\mathbb{R}^d)$ which coincides with the restriction of the usual weak- $*$ topology of $\mathcal{M}(\mathbb{R}^d, \mathbb{R}_+)$ induced by the duality pairing (1.1).

In the following proposition, we recall a useful convergence property on the measures of Borel sets, usually referred to as the *Portmanteau Theorem* (see e.g. [14, Proposition 1.62]).

Proposition 1.1 (Portmanteau Theorem). *Let $(\mu_n) \subset \mathcal{M}_+(\mathbb{R}^d)$ be a sequence of measures converging in the weak- $*$ topology towards $\mu \in \mathcal{M}_+(\mathbb{R}^d)$. Then for any Borel set $A \subset \mathbb{R}^d$ such that $\mu(\partial A) = 0$, it holds that $\mu_n(A) \xrightarrow{n \rightarrow +\infty} \mu(A)$.*

Another key notion in measure theory is that of absolute continuity with respect to a given measure, which we recall in the following definition. We also state in Theorem 1.1 the *Radon-Nykodym Theorem*.

Definition 1.1 (Absolutely continuous measures and Radon-Nikodym derivative). *Let $\Omega \subset \mathbb{R}^m$ and $U \subset \mathbb{R}^d$ be two Borel sets. Given a pair of measures $(\nu, \mu) \in \mathcal{M}(\Omega, U) \times \mathcal{M}(\Omega, \mathbb{R}_+)$, we say that ν is absolutely continuous with respect to μ – denoted by $\nu \ll \mu$ – if $\mu(A) = 0$ implies that $|\nu|(A) = 0$ for any Borel set $A \subset \Omega$.*

Theorem 1.1 (Radon-Nykodym). *Let $(\nu, \mu) \in \mathcal{M}(\Omega, U) \times \mathcal{M}(\Omega, \mathbb{R}_+)$. Then, $\nu \ll \mu$ if and only if there exists a Borel map $u \in L^1(\Omega, U; \mu)$ such that $\nu = u(\cdot)\mu$. This map is usually referred to as the Radon-Nikodym derivative of ν with respect to μ , and denoted by $u(\cdot) = \frac{d\nu}{d\mu}(\cdot)$.*

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

Definition 1.2 (Pushforward of a measure through a Borel map). *Given a measure $\mu \in \mathcal{P}(\mathbb{R}^d)$ and a Borel map $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$, the pushforward $f_{\#}\mu$ of μ through $f(\cdot)$ is defined as the Borel probability measure such that $f_{\#}\mu(B) = \mu(f^{-1}(B))$ for any Borel set $B \subset \mathbb{R}^d$.*

We recall in the following proposition three useful convergence results for sequences of probability measures and functions (see e.g. [17, Chapter 5]).

Proposition 1.2 (Convergence results). *Let $(\mu_n) \subset \mathcal{P}(\mathbb{R}^d)$ be a sequence narrowly converging to $\mu \in \mathcal{P}(\mathbb{R}^d)$, (f_n) be a sequence of μ -measurable functions pointwisely converging to f and $g \in C^0(\mathbb{R}^d)$. Then the following facts hold.*

(i) Suppose that the map $x \mapsto |g(x)|$ is uniformly integrable with respect to the family $\{\mu_n\}_{n=1}^\infty$, i.e.

$$\lim_{k \rightarrow +\infty} \int_{\{|g(x)| \geq k\}} |g(x)| d\mu_n(x) = 0$$

for all $n \geq 1$. Then, the sequence $(\int_{\mathbb{R}^d} g(x) d\mu_n(x)) \subset \mathbb{R}$ converges to $\int_{\mathbb{R}^d} g(x) d\mu(x)$ as $n \rightarrow +\infty$.

(ii) The sequence $(g_{\#}\mu_n) \subset \mathcal{P}(\mathbb{R}^d)$ narrowly converges to $g_{\#}\mu$ as $n \rightarrow +\infty$.

(iii) (Vitali convergence theorem) Suppose that the family $x \mapsto |f_n(x)|$ is uniformly integrable with respect to the measure μ , i.e.

$$\lim_{k \rightarrow +\infty} \int_{\{|f_n(x)| \geq k\}} |f_n(x)| d\mu(x) = 0$$

for all $n \geq 1$ and also assume that $|f(x)| < +\infty$ for μ -almost every $x \in \mathbb{R}^d$. Then (f_n) converges uniformly to f in $L^1(\mathbb{R}^d; \mu)$ as $n \rightarrow +\infty$.

We end this introductory paragraph by recalling in the following theorem the concept of *disintegration* of a family of vector-valued probability measures, see e.g. [14, Theorem 2.28].

Theorem 1.2 (Disintegration). *Let $\Omega_1 \subset \mathbb{R}^{m_1}$, $\Omega_2 \subset \mathbb{R}^{m_2}$ and $U \subset \mathbb{R}^d$ be arbitrary sets. Let $\nu \in \mathcal{M}(\Omega_1 \times \Omega_2, U)$ and $\pi^1 : \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \rightarrow \mathbb{R}^{m_1}$ be the projection map on the first factor. Denoting $\mu = \pi_{\#}^1 |\nu| \in \mathcal{M}(\Omega_1, \mathbb{R}_+)$, there exists a μ -almost uniquely determined Borel family of measures $\{\nu_x\}_{x \in \Omega_1} \subset \mathcal{M}(\Omega_2, U)$ such that*

$$\int_{\Omega_1 \times \Omega_2} f(x, y) d\nu(x, y) = \int_{\Omega_1} \left(\int_{\Omega_2} f(x, y) d\nu_x(y) \right) d\mu(x) \quad (1.3)$$

for any Borel map $f \in L^1(\Omega_1 \times \Omega_2, |\nu|)$. This construction is referred to as the disintegration of ν onto μ and denoted by $\nu = \int_{\Omega_1} \nu_x d\mu(x)$.

1.1.2 The optimal transport problem and Wasserstein spaces

The optimal mass transfer problem was first introduced in 1781 by Gaspard Monge in his “*Mémoire sur la Théorie des Déblais et des Remblais*”. The seminal problem studied by Monge¹ can be formulated as follows in modern mathematical terms: given two probability measures $\mu, \nu \in \mathcal{P}_c(\mathbb{R}^d)$, find a Borel map $T : \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that $T_{\#}\mu = \nu$ and

$$\int_{\mathbb{R}^d} c(x, T(x)) d\mu(x) = \min_{T'} \left\{ \int_{\mathbb{R}^d} c(x, T(x)) d\mu(x) \text{ s.t. } T'_{\#}\mu = \nu \right\}.$$

where $c : \mathbb{R}^{2d} \rightarrow (-\infty, +\infty]$ is a suitable cost function.

As mentioned in the introduction, this problem is in fact extremely tough mathematically speaking and remained unsolved for a long time. In particular, the image-measure constraint $T_{\#}\mu = \nu$ is badly behaved in the sense that it does not define a nice closed or

1. In his manuscript, Monge only considered the case $c(x, y) = |x - y|$

convex set in any reasonable functional topology. In 1942, the Russian mathematician Leonid Kantorovich introduced in [97] a relaxation of this constraint expressed in terms of *transport plans*.

Definition 1.3 (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 μ and ν – 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 pair of Borel sets $A, B \subset \mathbb{R}^d$. This property can be equivalently formulated in terms of pushforwards as $\pi_{\#}^1 \gamma = \mu$ and $\pi_{\#}^2 \gamma = \nu$.*

Transport plans are a much more convenient to study mass transportation problems than transport maps. Indeed, the set $\Gamma(\mu, \nu)$ is never empty and closed in the narrow topology of measures. The optimal mass transportation problem can then be formulated as follows: 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. [17, 125, 132]) 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 1.4 (Wasserstein distance and Wasserstein spaces). *Given two probability measures $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$, the p -Wasserstein distance between μ and ν is defined by*

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

The set of plans $\gamma \in \Gamma(\mu, \nu)$ achieving this optimal value is denoted² by $\Gamma_o(\mu, \nu)$ and referred to as the set of optimal transport plans between μ and ν . 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. [17, Chapter 7] or [132, Chapter 6]).

Proposition 1.3 (Elementary 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 (1.2). More precisely, it holds that*

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

Given $\mu, \nu \in \mathcal{P}(\mathbb{R}^d)$, 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

2. We omit the dependence on p for clarity and conciseness.

formula holds

$$W_1(\mu, \nu) = \sup_{\phi} \left\{ \int_{\mathbb{R}^d} \phi(x) d(\mu - \nu)(x) \text{ s.t. } \text{Lip}(\phi; \mathbb{R}^d) \leq 1 \right\}, \quad (1.4)$$

where $\text{Lip}(\phi; \mathbb{R}^d)$ denotes the Lipschitz constant of the map $\phi(\cdot)$ over \mathbb{R}^d .

In what follows, we shall mainly restrict our considerations to the Wasserstein spaces of order 1 and 2 built over $\mathcal{P}_c(\mathbb{R}^d)$. In the particular case of the square Wasserstein distance, optimal transport plans enjoy a fairly explicit characterization that we recall in the following theorem which is a consequence of [17, Theorem 6.1.5].

Theorem 1.3 (Characterization of optimal plans for the square Wasserstein distance). *Let $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^{2d})$ and $\gamma \in \Gamma(\mu, \nu)$. Then, γ is a W_2 -optimal transport plan between μ and ν if and only if there exists a convex function $\psi : \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that $y \in \partial\psi(x)$ for γ -almost every $(x, y) \in \mathbb{R}^{2d}$.*

This result is based on the necessary and sufficient optimality condition for the optimal transport problem stating that γ is concentrated on a $|\cdot|^2$ -monotone set described in terms of Kantorovich potentials. From this general result, we can deduce the following easy corollary which will prove useful in studying the geometric structure of $(\mathcal{P}_2(\mathbb{R}^d), W_2)$.

Corollary 1.1 (Perturbation of the identity by gradients of test functions). *For any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $\xi \in C_c^\infty(\mathbb{R}^d)$, there exists $\bar{s} > 0$ such that the application $\text{Id} + s\nabla\xi(\cdot)$ is an W_2 -optimal transport map between μ and $(\text{Id} + s\nabla\xi)_\# \mu$ for any $s \in (-\bar{s}, \bar{s})$.*

Proof. For any $s \in \mathbb{R}$, the map $\text{Id} + s\nabla\xi(\cdot)$ is the gradient of the application $x \in \mathbb{R}^d \mapsto \frac{1}{2}|x|^2 + s\xi(x)$. This function is smooth, and for $|s|$ smaller than a given $\bar{s} > 0$, it is also convex. Whence, $(\text{Id}, \text{Id} + s\nabla\xi)_\# \mu$ is concentrated on the subdifferential of a convex function and is therefore W_2 -optimal as a consequence of Theorem 1.3. \square

We end these introductory paragraphs by recalling the concept *barycenter* in the context of optimal transport, and by providing an easy estimate for the 1-Wasserstein distance between barycenters.

Definition 1.5 (Disintegration and barycenter). *Let $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$ and $\gamma \in \Gamma(\mu, \nu)$ be a transport plan between μ and ν . The barycenter $\bar{\gamma} \in L^p(\mathbb{R}^d, \mathbb{R}^d; \mu)$ of the plan γ is then defined by*

$$\bar{\gamma}(x) = \int_{\mathbb{R}^d} y d\gamma_x(y). \quad (1.5)$$

for μ -almost every $x \in \mathbb{R}^d$, where $\{\gamma_x\}_{x \in \mathbb{R}^d}$ is the disintegration of γ onto its first marginal μ .

Proposition 1.4 (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*

$$W_1(\gamma^1, \gamma^2) \leq \int_{\mathbb{R}^d} W_1(\gamma_x^1, \gamma_x^2) d\mu(x). \quad (1.6)$$

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 (1.4) since the maps $r \mapsto \xi(x, r)$ are 1-Lipschitz for all $x \in \mathbb{R}^d$. 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, again as a consequence of (1.4). \square

1.1.3 Absolutely continuous curves in $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ and the analytical tangent space $\text{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$

In this section, we recall some structural results concerning continuity equations and absolutely continuous curves in Wasserstein spaces. We also link these concepts with the geometric structure of $(\mathcal{P}_2(\mathbb{R}^d), W_2)$, and in particular with its tangent space $\text{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$. We restrict our attention to the Wasserstein spaces of order 2 because in the subsequent chapters, we will only deal with compactly supported measures. We can therefore choose to use the squared Wasserstein distance which induces the metric structure which is the best suited to our control-theoretic purposes. Most of the notions presented in this section are borrowed from [17, Chapter 8].

Continuity equation in $\mathcal{P}_2(\mathbb{R}^d)$ are hyperbolic equations in divergence form given in the general form

$$\partial_t \mu(t) + \nabla \cdot (\mathbf{w}(t, \cdot) \mu(t)) = 0, \quad (1.7)$$

where $(t, x) \in [0, T] \times \mathbb{R}^d \mapsto \mathbf{w}(t, x) \in \mathbb{R}^d$ is a Borel vector field satisfying the integrability condition

$$\int_0^T \int_{\mathbb{R}^d} |\mathbf{w}(t, x)|^2 d\mu(t)(x) dt < +\infty.$$

Equation (1.7) has to be understood in the sense of distributions against smooth and compactly supported test functions, i.e.

$$\int_0^T \int_{\mathbb{R}^d} \left(\partial_t \xi(t, x) + \langle \nabla_x \xi(t, x), \mathbf{w}(t, x) \rangle \right) d\mu(t)(x) dt = 0, \quad (1.8)$$

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

$$\frac{d}{dt} \int_{\mathbb{R}^d} \xi(x) d\mu(t)(x) = \int_{\mathbb{R}^d} \langle \nabla \xi(x), \mathbf{w}(t, x) \rangle d\mu(t)(x), \quad (1.9)$$

for any $\xi \in C_c^\infty(\mathbb{R}^d)$.

Continuity equations play a fundamental role in the geometric structure of Wasserstein spaces. Indeed, they are the natural generalization of the concept of ODE to this setting, as supported by the following result which is a simplified statement of [17, Chapter 8.3.1].

Theorem 1.4 (Absolutely continuous curves in $(\mathcal{P}_2(\mathbb{R}^d), W_2)$). *A curve of measures $t \in [0, T] \mapsto \mu(t) \in \mathcal{P}_2(\mathbb{R}^d)$ is absolutely continuous in the W_2 -metric if and only if there exists a Borel velocity field $(t, x) \in [0, T] \times \mathbb{R}^d \mapsto \mathbf{w}(t, x)$ such that $\mathbf{w}(t, \cdot) \in L^2(\mathbb{R}^d, \mathbb{R}^d; \mu(t))$ for \mathcal{L}^1 -almost every $t \in [0, T]$ and (1.7) holds in the sense of distributions.*

This strict correspondence between absolutely continuous curves of measures and

solutions of continuity equations is what inspired the definition of the analytical tangent space to $\mathcal{P}_2(\mathbb{R}^d)$.

Definition 1.6 (Wasserstein tangent space). *We define the analytical tangent space $\text{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$ to $\mathcal{P}_2(\mathbb{R}^d)$ at a measure $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ by*

$$\text{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d) = \overline{\{\nabla \xi(\cdot) \text{ s.t. } \xi \in C_c^\infty(\mathbb{R}^d)\}}^{L^2(\mu)}. \quad (1.10)$$

This definition can be heuristically understood as follows. As illustrated by (1.8)-(1.9), the velocity field $\mathbf{w}(\cdot, \cdot)$ driving a continuity equation only acts on gradients of test functions. Besides, one can check that adding any divergence-free vector field to $\mathbf{w}(\cdot, \cdot)$ does not modify the corresponding solution of the continuity equation. It is therefore possible to select a canonical representative of minimum $L^2(\mathbb{R}^d, \mathbb{R}^d; \mu(t))$ -norm among the velocity fields generating the same absolutely continuous curve $t \in [0, T] \mapsto \mu(t)$. Such a canonical representative belongs to $\text{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$.

Remark 1.1 (On the different notions of tangent spaces to $(\mathcal{P}_2(\mathbb{R}^d), W_2)$). *As illustrated in [17, Chapters 8], it is possible to define several notions of tangent spaces to the Wasserstein space of order 2. We termed the definition given in (1.10) to be analytical in the sense that it does not stem from any geometric or metric consideration, but only from the structure of continuity equations. Let it be noted however that this definition incorporates indirectly the geometry of $(\mathcal{P}_2(\mathbb{R}^d), W_2)$, since we know that gradients of test functions generate optimal transport maps by perturbation of the identity as a consequence of Theorem 1.3 and Corollary 1.1. Whence, gradients of test functions can be thought of as being tangent vectors to $(\mathcal{P}_2(\mathbb{R}^d), W_2)$. Moreover, it is proven in [17, Chapter 12] that $\text{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$ is the image of the metric tangent space – which is made of transport plans – by the barycentric projection (1.5).*

1.2 First and second order differential calculus in $(\mathcal{P}_2(\mathbb{R}^d), W_2)$

In this section, we recall elements of subdifferential calculus in the Wasserstein space $(\mathcal{P}_2(\mathbb{R}^d), W_2)$, along with notions pertaining to second order calculus. For a thorough introduction to Wasserstein subdifferential calculus, we refer the reader [17, Chapters 9-11] where the full theory is developed and applied to the study of gradient flows. Concerning second order calculus, we refer the reader to [16, 88] for a comprehensive theoretical study of the topic, and to [57] for hands-on facts on Wasserstein Hessians.

1.2.1 Subdifferential calculus in Wasserstein spaces

Throughout this section, we denote by $\phi : \mathcal{P}_2(\mathbb{R}^d) \rightarrow (-\infty, +\infty]$ a proper and lower-semicontinuous functional. We denote the effective domain $D(\phi)$ of $\phi(\cdot)$ as the set of points where it is finite, i.e.

$$D(\phi) = \{\mu \in \mathcal{P}_2(\mathbb{R}^d) \text{ s.t. } \phi(\mu) < +\infty\}.$$

We further assume that for $\tau_* > 0$ small enough, the *Moreau-Yosida* relaxation of $\phi(\cdot)$ defined by

$$\phi_M(\mu, \tau; \cdot) : \nu \mapsto \frac{1}{2\tau} W_2^2(\mu, \nu) + \phi(\nu) \quad (1.11)$$

attains a minimum at some $\mu_\tau \in D(\phi)$ for any $\tau \in (0, \tau_*)$. This technical assumption is satisfied whenever $\phi(\cdot)$ is bounded from below and at least lower-semicontinuous and is crucial for proving the main results of the theory developed in [17, Chapter 10].

We start by introducing the concept of *extended subdifferentials* for a functional defined over the Wasserstein space $(\mathcal{P}_2(\mathbb{R}^d), W_2)$.

Definition 1.7 (Extended subdifferential). *Let $\mu^1 \in D(\phi)$. We say that a transport plan $\gamma \in \mathcal{P}_2(\mathbb{R}^{2d})$ belongs to the extended (Fréchet) subdifferential $\partial\phi(\mu^1)$ of $\phi(\cdot)$ at μ^1 provided that*

- (i) $\pi_{\#}^1 \gamma = \mu^1$,
- (ii) for all $\mu^3 \in \mathcal{P}_2(\mathbb{R}^d)$ it holds

$$\phi(\mu^3) - \phi(\mu^1) \geq \inf_{\mu \in \Gamma_o^{1,3}(\gamma, \mu^3)} \left[\int_{\mathbb{R}^{3d}} \langle x_2, x_3 - x_1 \rangle d\mu(x_1, x_2, x_3) \right] + o(W_2(\mu^1, \mu^3)),$$

where $\Gamma_o^{1,3}(\gamma, \mu^3) = \{\mu \in \Gamma(\gamma, \mu^3) \text{ s.t. } \pi_{\#}^{1,3} \mu \in \Gamma_o(\mu^1, \mu^3)\}$. Moreover, we say that an extended subdifferential γ is induced by a plan if there exists $\xi \in L^2(\mathbb{R}^d, \mathbb{R}^d; \mu^1)$ such that $\gamma = (\text{Id} \times \xi)_{\#} \mu^1$.

We say that a transport plan $\gamma \in \mathcal{P}_2(\mathbb{R}^{2d})$ belongs to the strong extended subdifferential $\partial_S \phi(\mu^1)$ of $\phi(\cdot)$ at μ^1 if the following stronger condition

$$\phi(\mu^3) - \phi(\mu^1) \geq \int_{\mathbb{R}^{3d}} \langle x_2, x_3 - x_1 \rangle d\mu(x_1, x_2, x_3) + o(W_{2,\mu}(\mu^1, \mu^3)), \quad (1.12)$$

holds for any $\mu^3 \in \mathcal{P}_2(\mathbb{R}^d)$ and $\mu \in \Gamma(\gamma, \mu^3)$. Here for $\mu \in \mathcal{P}(\mathbb{R}^{3d})$, the quantity $W_{2,\mu}(\mu^1, \mu^3)$ is defined by

$$W_{2,\mu}(\mu^1, \mu^3) = \left(\int_{\mathbb{R}^{2d}} |x_1 - x_3|^2 d\mu(x_1, x_2, x_3) \right)^{1/2}.$$

We now introduce the technical notions of *regularity* and *metric slope* that are instrumental in deriving a sufficient condition for the extended subdifferential of a functional to be non-empty. This result is stated in Theorem 1.5 and its proof can be found in [17, Theorem 10.3.10].

Definition 1.8 (Regular functionals over $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ and metric slope). *A proper and lower semicontinuous functional $\phi(\cdot)$ is said to be regular provided that whenever $(\mu_n) \subset \mathcal{P}_2(\mathbb{R}^d)$ and $(\gamma_n) \subset \mathcal{P}_2(\mathbb{R}^{2d})$ are taken such that*

$$\begin{cases} \mu_n \xrightarrow{W_2} \mu \text{ in } \mathcal{P}_2(\mathbb{R}^d) \text{ , } \phi(\mu_n) \longrightarrow \tilde{\phi} \text{ in } \mathbb{R}, \\ \gamma_n \in \partial_S \phi(\mu_n) \ \forall n \geq 1 \text{ , } \gamma_n \xrightarrow{W_2} \gamma \text{ in } \mathcal{P}_2(\mathbb{R}^{2d}), \end{cases}$$

it implies that $\gamma \in \partial\phi(\mu)$ and $\tilde{\phi} = \phi(\mu)$.

Furthermore, we define the metric slope $|\partial\phi|(\mu)$ of the functional $\phi(\cdot)$ at $\mu \in D(\phi)$ as

$$|\partial\phi|(\mu) = \limsup_{\nu \rightarrow \mu} \left[\frac{(\phi(\mu) - \phi(\nu))^+}{W_2(\mu, \nu)} \right],$$

where $(\bullet)^+$ denotes the positive part.

Theorem 1.5 (Link between extended subdifferentials and metric slopes). *Let $\phi(\cdot)$ be a proper, lower-semicontinuous, bounded from below and regular functional over $\mathcal{P}_2(\mathbb{R}^d)$. Then, the extended subdifferential $\partial\phi(\mu)$ of $\phi(\cdot)$ at some $\mu \in D(\phi)$ is non-empty if and only if its metric slope $|\partial\phi|(\mu)$ at μ is finite.*

In which case, there exists a unique minimal selection in $\partial\phi(\mu)$, denoted by $\partial^\circ\phi(\mu)$, satisfying

$$\begin{aligned} \left(\int_{\mathbb{R}^{2d}} |r|^2 d(\partial^\circ\phi(\mu))(x, r) \right)^{1/2} &= \min_{\gamma} \left\{ \left(\int_{\mathbb{R}^{2d}} |r|^2 d\gamma(x, r) \right)^{1/2} \quad \text{s.t. } \gamma \in \partial\phi(\mu) \right\} \\ &= |\partial\phi|(\mu). \end{aligned}$$

This minimal selection can be explicitly characterized as follows : let μ_τ be the minimizer of the Moreau-Yosida functional (1.11) for some $\tau \in (0, \tau_)$. Then there exists a family of strong subdifferentials $(\gamma_\tau) \subset (\partial_S\phi(\mu_\tau))$ which converges towards $\partial^\circ\phi(\mu)$ in the W_2 -metric along a suitable vanishing sequence $\tau_n \downarrow 0$.*

As in the more classical theory of subdifferential calculus in Euclidean spaces, it is possible to compute directional derivatives of sufficiently nice functionals by using subgradients. These results can be generalized to smooth Riemannian manifolds to compute directional derivatives along vectors which belong to the tangent space. This analogy still holds in the setting of Wasserstein spaces and gives rise to the following chainrule involving $\text{Tan}_\mu\mathcal{P}_2(\mathbb{R}^d)$, which is a simplified version of [17, Proposition 10.3.18].

Proposition 1.5 (Directional derivative along tangent vectors). *Let $\xi \in \text{Tan}_\mu\mathcal{P}_2(\mathbb{R}^d)$, $\bar{s} > 0$ and $\phi : \mathcal{P}_2(\mathbb{R}^d) \rightarrow (-\infty, +\infty]$ be a lower-semicontinuous and proper functional. Furthermore, suppose that*

- (i) $|\partial\phi(\mu)| < +\infty$.
- (ii) *The map $s \in (-\bar{s}, \bar{s}) \mapsto \phi((\text{Id} + s\xi)_\#\mu)$ is differentiable at $s = 0$.*

Then, it holds that

$$\frac{d}{ds} \phi((\text{Id} + s\xi)_\#\mu)|_{s=0} = \int_{\mathbb{R}^{2d}} \langle r, \xi(x) \rangle d\gamma(x, r),$$

for any $\gamma \in \partial\phi(\mu)$.

In Chapter 2, our proof of the Pontryagin Maximum Principle by perturbation methods will require a more general Wasserstein chainrule. More precisely, we will need such a chainrule for curves of measures of the form $\epsilon \mapsto \mu(\epsilon) = \mathcal{G}(\epsilon, \cdot)_\#\mu$, where $\epsilon \mapsto \mathcal{G}(\epsilon, \cdot)$ is Fréchet differentiable at $\epsilon = 0$. In this case, one would a priori need to restrict himself to strong subdifferentials, since $\frac{d}{d\epsilon}\mathcal{G}(\epsilon, \cdot)$ is not an optimal transport map in general. Yet, there is generically no reason for the strong subdifferential of a functional be non-empty, except at the points μ_{τ_n} where the minimum of the Moreau-Yosida relaxation (1.11) is reached.

In Proposition 1.6, we condense some of the results of [17, Chapter 10] in order to provide a chain rule that allows to compute derivatives along smooth vector fields using the minimal selection $\partial^\circ\phi(\mu)$. For simplicity, we state this result in the framework of the Wasserstein space $\mathcal{P}_c(\mathbb{R}^d)$.

Proposition 1.6 (Minimal selection and chain rule along smooth vector fields). *Let $K \subset \mathbb{R}^d$ be a compact set, $\mu \in \mathcal{P}(K)$ and $\phi : \mathcal{P}(K) \rightarrow (-\infty, +\infty]$ be Lipschitz in the W_2 -metric, regular in the sense of Definition 1.8, and uniformly bounded over $\mathcal{P}(K)$. Let $\mathcal{G} \in \text{Lip}((-\bar{\epsilon}, \bar{\epsilon}), C^0(\mathbb{R}^d, \mathbb{R}^d))$ be a function satisfying the following assumptions.*

- (i) $\mathcal{G}(0, \cdot) = \text{Id}$ and $\epsilon \mapsto \mathcal{G}(\epsilon, x)$ is Fréchet-differentiable at $\epsilon = 0$ uniformly with respect to $x \in K$.
- (ii) $\text{supp}(\mathcal{G}(\epsilon, \cdot) \# \mu) \subset K$ for all $\epsilon \in (-\bar{\epsilon}, \bar{\epsilon})$
- (iii) The directional derivative map $\mathcal{F} : x \mapsto \frac{d}{d\epsilon} [\mathcal{G}(\epsilon, x)]_{\epsilon=0}$ is continuous.

Furthermore, suppose that the map $\epsilon \mapsto \phi(\mathcal{G}(\epsilon, \cdot) \# \mu)$ is differentiable at $\epsilon = 0$.

Then, it holds that

$$\frac{d}{d\epsilon} \phi(\mathcal{G}(\epsilon, \cdot) \# \mu)_{\epsilon=0} = \int_{\mathbb{R}^{2d}} \langle \bar{\gamma}_\phi^\circ(x), \mathcal{F}(x) \rangle d\mu(x),$$

where $\bar{\gamma}_\phi^\circ \in L^2(\mathbb{R}^d, \mathbb{R}^d; \mu)$ is the barycenter of $\partial^\circ \phi(\mu)$.

Proof. First remark that it holds for any $\nu \in \mathcal{P}(K)$

$$(\phi(\mu) - \phi(\nu))^+ \leq \text{Lip}(\phi, \mathcal{P}(K)) W_2(\mu, \nu)$$

where $\text{Lip}(\phi, \mathcal{P}(K))$ is the Lipschitz constant of $\phi(\cdot)$ on $\mathcal{P}(K)$. Hence, $|\partial\phi|(\mu)$ is uniformly bounded by $\text{Lip}(\phi, \mathcal{P}(K))$. Moreover, the assumption that $\phi(\cdot)$ is bounded from below and Lipschitz on sets of uniformly compactly supported measures implies that for $\tau_* > 0$ small enough, the Moreau-Yosida functional $\Phi_M(\mu, \tau; \cdot)$ defined in (1.11) attains a minimum point $\mu_\tau \in D(\phi) \subset \mathcal{P}(K)$ for any $\tau \in (0, \tau_*)$. Thus, by Theorem 1.5, $\partial\phi(\mu)$ is non-empty and contains at least the minimal selection $\partial^\circ \phi(\mu)$ at any $\mu \in \mathcal{P}(K)$.

Consider a sequence $(\tau_n) \subset (0, \tau_*)$ converging to 0 and the corresponding sequence of strong subdifferentials $(\gamma_{\tau_n}) \subset (\partial_S \phi(\mu_{\tau_n}))$ converging towards $\partial^\circ \phi(\mu)$ in the W_2 -metric. Pick $\epsilon \in (0, \bar{\epsilon})$ small enough and choose $\mu_\epsilon^{\tau_n} = (\pi^1, \pi^2, \mathcal{G}(\epsilon, \cdot) \circ \pi^1) \# \gamma_{\tau_n} \in \Gamma(\gamma_{\tau_n}, \mathcal{G}(\epsilon, \cdot) \# \mu_{\tau_n})$. By the definition of strong subdifferentials given in (1.12), it holds that

$$\frac{\phi(\mathcal{G}(\epsilon, \cdot) \# \mu_{\tau_n}) - \phi(\mu_{\tau_n})}{\epsilon} \geq \int_{\mathbb{R}^{2d}} \left\langle r, \frac{\mathcal{G}(\epsilon, x) - x}{\epsilon} \right\rangle d\gamma_{\tau_n}(x, r) + o(1). \quad (1.13)$$

since

$$o(W_{2, \mu}(\mathcal{G}(\epsilon, \cdot) \# \mu_{\tau_n}, \mu_{\tau_n})) = o\left(\|\mathcal{G}(\epsilon, \cdot) - \text{Id}\|_{L^2(\mu_{\tau_n})}\right) = o(\epsilon),$$

for all $n \geq 1$. Remark that the left hand side of (1.13) is bounded over $\mathcal{P}(K)$ uniformly with respect to $n \geq 1$ and $\epsilon \in (0, \bar{\epsilon})$ by Lipschitzianity of $\phi(\cdot)$.

We recall that $\gamma_{\tau_n} \xrightarrow{W_2} \partial^\circ \phi(\mu)$ in $\mathcal{P}_2(\mathbb{R}^{2d})$. Notice that the whole sequence (μ_{τ_n}) is in $\mathcal{P}(K)$, thus for all $\epsilon \in (0, \bar{\epsilon})$ the maps $x \mapsto |(\mathcal{G}(\epsilon, x) - x)/\epsilon|^2$ are uniformly integrable with respect to $\{\pi_{\#}^1 \gamma_{\tau_n}\}_{n=1}^{+\infty}$. Hence, the maps $(x, r) \mapsto |\langle r, (\mathcal{G}(\epsilon, x) - x)/\epsilon \rangle|$ are uniformly integrable with respect to $\{\gamma_n\}_{n=1}^{+\infty}$ and the application of Proposition 1.2-(i) implies that

for all $\epsilon \in (0, \bar{\epsilon})$,

$$\begin{aligned} \int_{\mathbb{R}^{2d}} \left\langle r, \frac{\mathcal{G}(\epsilon, x) - x}{\epsilon} \right\rangle d\gamma_{\tau_n}(x, r) &\xrightarrow{\tau_n \downarrow 0} \int_{\mathbb{R}^{2d}} \left\langle r, \frac{\mathcal{G}(\epsilon, x) - x}{\epsilon} \right\rangle d(\partial^\circ \phi(\mu))(x, r) \\ &= \int_{\mathbb{R}^d} \left\langle \bar{\gamma}_\phi^\circ(x), \frac{\mathcal{G}(\epsilon, x) - x}{\epsilon} \right\rangle d\mu(x) \end{aligned} \quad (1.14)$$

where we used the notion *barycenter* of a plan introduced in Definition 1.5.

Moreover, the Lipschitz regularity in the W_2 -metric of $\phi(\cdot)$ over $\mathcal{P}(K)$ together with Proposition 1.2-(ii) imply that

$$\phi(\mathcal{G}(\epsilon, \cdot) \# \mu_{\tau_n}) \longrightarrow \phi(\mathcal{G}(\epsilon, \cdot) \# \mu). \quad (1.15)$$

Thus, merging (1.13), (1.14) and (1.15), we prove that for any $\epsilon \in (0, \bar{\epsilon})$ with $\bar{\epsilon} > 0$ small enough, it holds

$$\frac{\phi(\mathcal{G}(\epsilon, \cdot) \# \mu) - \phi(\mu)}{\epsilon} \geq \int_{\mathbb{R}^d} \left\langle \bar{\gamma}_\phi^\circ(x), \frac{\mathcal{G}(\epsilon, x) - x}{\epsilon} \right\rangle d\mu(x) + o(1).$$

The family of maps $(|\langle \bar{\gamma}_\phi^\circ(\cdot), (\mathcal{G}(\epsilon, \cdot) - \text{Id})/\epsilon \rangle|)_{\epsilon \in (0, \bar{\epsilon})}$ is uniformly integrable with respect to μ as a consequence of our assumptions made on the functions $\mathcal{G}(\cdot, \cdot)$. It also holds that $|\langle \bar{\gamma}_\phi^\circ(\cdot), \mathcal{F}(\cdot) \rangle| < +\infty$ μ -almost everywhere. Therefore, letting $\epsilon \downarrow 0$ and invoking Proposition 1.2-(iii), we recover that

$$\lim_{\epsilon \downarrow 0} \left[\frac{\phi(\mathcal{G}(\epsilon, \cdot) \# \mu) - \phi(\mu)}{\epsilon} \right] \geq \int_{\mathbb{R}^d} \langle \bar{\gamma}_\phi^\circ(x), \mathcal{F}(x) \rangle d\mu(x).$$

Following the same steps with $\epsilon \in (-\bar{\epsilon}, 0)$, we obtain the converse inequality for $\epsilon \uparrow 0$

$$\lim_{\epsilon \uparrow 0} \left[\frac{\phi(\mathcal{G}(\epsilon, \cdot) \# \mu) - \phi(\mu)}{\epsilon} \right] \leq \int_{\mathbb{R}^d} \langle \bar{\gamma}_\phi^\circ(x), \mathcal{F}(x) \rangle d\mu(x).$$

Since we assumed that $\epsilon \mapsto \phi(\mathcal{G}(\epsilon, \cdot) \# \mu)$ is differentiable at $\epsilon = 0$, both limits coincide and

$$\frac{d}{d\epsilon} [\phi(\mathcal{G}(\epsilon, \cdot) \# \mu)]_{\epsilon=0} = \int_{\mathbb{R}^d} \langle \bar{\gamma}_\phi^\circ(x), \mathcal{F}(x) \rangle d\mu(x),$$

which proves our claim. \square

Remark 1.2 (The case $\partial_S \phi(\mu) \neq \emptyset$). *When $\partial_S \phi(\mu)$ is non-empty, the previous chain rule can be applied with any strong subdifferential and for more general classes of vector fields, see e.g. [17, Remark 10.3.2].*

The interest of proving this kind of result for the minimal selection is twofold. First, as recalled in Theorem 1.5, a minimal selection always exists when the extended subdifferential is non-empty. Second, minimal selections can be computed explicitly even in very general settings for a wide range of functionals, see e.g. [17, Chapter 10.4].

1.2.2 Classical subdifferentials and Wasserstein gradients

In Proposition 1.6, we have proven a general chain rule along curves of measures generated by 1-parameter families of smooth vector fields. In Chapter 3, we will also

require such a chain rule for curves of measures generated by multi-dimensional families of vector fields $\mathcal{G} : \mathbb{R}^N \times \mathbb{R}^d \rightarrow \mathbb{R}^d$. In this context however, the subdifferentiability of the functional in the sense of Definition 1.7 is not sufficient to perform such an expansion.

With this goal in mind, we introduce the stronger and simpler notion of *classical Wasserstein subdifferential* in the following definition.

Definition 1.9 (Classical 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 μ 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 μ if $(-\xi) \in \partial^-(-\phi)(\mu)$.

It has been proven recently in [86] that the definition of classical Wasserstein subdifferential involving a supremum taken over the set of optimal transport plans is equivalent to the usual one introduced in [17] and described in Definition 1.7 which involves an infimum. This allows for the elaboration of a convenient notion of differentiability in Wasserstein spaces as detailed below.

Definition 1.10 (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 element $\nabla_\mu \phi(\mu) \in \partial^- \phi(\mu) \cap \partial^+ \phi(\mu) \cap \text{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$ called the Wasserstein gradient of $\phi(\cdot)$ at μ , which satisfies*

$$\phi(\nu) - \phi(\mu) = \int_{\mathbb{R}^{2d}} \langle \nabla_\mu \phi(\mu)(x), y - x \rangle d\gamma(x, y) + o(W_2(\mu, \nu)) \quad (1.16)$$

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

Given a Wasserstein-differentiable functional $\phi(\cdot)$, we can define its *Lie derivative* in a tangent direction $\xi \in \text{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$ by

$$\mathcal{L}_\xi \phi(\mu) = \frac{d}{ds} \phi((\text{Id} + s\xi)_\# \mu)|_{s=0} = \int_{\mathbb{R}^d} \langle \nabla_\mu \phi(\mu)(x), \xi(x) \rangle d\mu(x) \quad (1.17)$$

which exists and is uniquely determined as a consequence of Proposition 1.5 and Definition 1.10.

In the following proposition, we show that the Wasserstein gradient introduced in Definition 1.10 coincides – when it exists – with the barycenter of the minimal selection in the extended subdifferential defined in Theorem 1.5.

Proposition 1.7 (Wasserstein gradients are minimal selections). *Let $\phi : \mathcal{P}_2(\mathbb{R}^d) \rightarrow (-\infty, \infty]$ be differentiable at $\mu \in D(\phi)$ in the sense of Definition 1.10. Then the minimal selection in the extended subdifferential exists and is induced by the Wasserstein gradient, i.e. $\partial^\circ \phi(\mu) = (\text{Id} \times \nabla_\mu \phi(\mu))_\# \mu$.*

Proof. Since $\phi(\cdot)$ is differentiable at $\mu \in D(\phi)$, it holds that $(\text{Id} \times \nabla_\mu \phi(\mu))_\# \mu \in \partial \phi(\mu)$ so that the extended Fréchet subdifferential of $\phi(\cdot)$ in the sense of Definition 1.7 is non-empty. Therefore by Theorem 1.5, there exists a minimal selection $\partial^\circ \phi(\mu)$.

Let $\xi \in \text{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$ and $\bar{s} > 0$ be a small parameter. As a consequence of (1.16) and (1.17), the map $s \in (-\bar{s}, \bar{s}) \mapsto \phi((\text{Id} + s\xi)_\# \mu)$ is differentiable at $s = 0$ with derivative

$$\frac{d}{ds} \phi((\text{Id} + s\xi)_\# \mu)|_{s=0} = \int_{\mathbb{R}^d} \langle \nabla_\mu \phi(\mu)(x), \xi(x) \rangle d\mu(x). \quad (1.18)$$

Moreover, the chain rule of Proposition 1.5 also implies that

$$\frac{d}{ds} \phi((\text{Id} + s\xi)_\# \mu)|_{s=0} = \int_{\mathbb{R}^d} \langle \bar{\gamma}_\phi^\circ(x), \xi(x) \rangle d\mu(x) \quad (1.19)$$

where $\bar{\gamma}_\phi^\circ(\cdot)$ denotes the barycentric projection of $\partial^\circ \phi(\mu)$. By merging (1.18) and (1.19), we obtain that

$$\int_{\mathbb{R}^d} \langle \nabla_\mu \phi(\mu)(x) - \bar{\gamma}_\phi^\circ(x), \xi(x) \rangle d\mu(x) = 0, \quad (1.20)$$

for any $\xi \in \text{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$.

Recall now that by Definition 1.10, the Wasserstein gradient $\nabla_\mu \phi(\mu)(\cdot)$ belongs to the analytical tangent space $\text{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$. Choosing in particular $\xi(\cdot) = \nabla_\mu \phi(\mu)(\cdot)$ in (1.20) and applying Cauchy-Schwarz inequality, we obtain that

$$\left(\int_{\mathbb{R}^d} |\nabla_\mu \phi(\mu)(x)|^2 d\mu(x) \right)^{1/2} \leq \left(\int_{\mathbb{R}^d} |\bar{\gamma}_\phi^\circ(x)|^2 d\mu(x) \right)^{1/2}.$$

There now remains to observe that as a consequence of the definition of barycentric projection (1.5) and Young's inequality, it holds that

$$\begin{aligned} \int_{\mathbb{R}^d} |\bar{\gamma}_\phi^\circ(x)|^2 d\mu(x) &= \int_{\mathbb{R}^d} \left| \int_{\mathbb{R}^d} r d(\partial^\circ \phi(\mu))_x(r) \right|^2 d\mu(x) \\ &\leq \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |r|^2 d(\partial^\circ \phi(\mu))_x(r) d\mu(x) = \int_{\mathbb{R}^d} |r|^2 d(\partial^\circ \phi(\mu))(x, r). \end{aligned}$$

We can therefore conclude that $\partial^\circ \phi(\mu) = (\text{Id} \times \nabla_\mu \phi(\mu))_\# \mu$ since $(\text{Id} \times \nabla_\mu \phi(\mu))_\# \mu \in \partial \phi(\mu)$ and $\|\nabla_\mu \phi(\mu)\|_{L^2(\mu)} \leq |\partial \phi|(\mu)$. \square

Remark 1.3 (Barycenters of minimal selections are tangent). *The proof that we gave for Proposition 1.7 is pretty direct, but maybe not the most insightful from a geometric point of view. Another way to look at this is to remark that $\bar{\gamma}_\phi^\circ(\cdot)$ belongs to $\text{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$. This is due to the fact that the minimal selection belongs to the metric tangent space by construction (see the proofs of Theorems 10.3.10 and 10.3.11), which image under the barycentric projection is $\text{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$ (see [17, Theorem 12.4.4]).*

We end this primer on first order Wasserstein calculus by generalizing in Proposition 1.8 below the chain rule formula of Proposition 1.6 for Wasserstein differentiable functionals along multi-dimensional families of measures.

Proposition 1.8 (Chain rule along multi-dimensional perturbations by smooth vector fields). *Let $K \subset \mathbb{R}^d$ be a compact set and $\mu \in \mathcal{P}(K)$. Suppose that $\phi : \mathcal{P}(K) \rightarrow (-\infty, +\infty]$ is Lipschitz in the W_2 -metric, regular in the sense of Definition 1.8 and Wasserstein differentiable over $\mathcal{P}(K)$. Given $N \geq 1$ and a small parameter $\epsilon > 0$, suppose that $\mathcal{G} \in C^0([-\epsilon, \epsilon]^N \times K, \mathbb{R}^d)$ is a function such that the following holds.*

(i) $\mathcal{G}(0, \cdot) = \text{Id}$ 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

$$\begin{aligned} \nabla_e(\phi(\mathcal{G}(e, \cdot)_{\#}\mu))|_{e=0}(\sigma) &= \int_{\mathbb{R}^d} \langle \nabla_\mu \phi(\mu)(x), D_e \mathcal{G}(0, x)\sigma \rangle d\mu(x) \\ &= \sum_{k=1}^N \sigma_k \int_{\mathbb{R}^d} \langle \nabla_\mu \phi(\mu)(x), \mathcal{F}_k(x) \rangle d\mu(x), \end{aligned} \quad (1.21)$$

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

Proof. Even though the proof relies on the same arguments as that of Proposition 1.6, we include it in the manuscript to highlight where the need of higher differentiability requirements comes into play. By assumption, $\phi(\cdot)$ is Lipschitz over $\mathcal{P}(K)$ in the W_2 -metric, regular in the sense of Definition 1.8 and Wasserstein-differentiable in the sense of Definition 1.10 at μ . Hence, $|\partial\phi|(\mu) < +\infty$ and $\partial^\circ\phi(\mu) = (\text{Id} \times \nabla_\mu\phi(\mu))_{\#}\mu$ as a consequence of Proposition 1.7.

Let $(\tau_n) \subset (0, \tau_*)$ be a suitable vanishing sequence, $(\mu_{\tau_n}) \subset D(\phi)$ be a sequence of minimizers of the Moreau-Yosida functional (1.11) and $(\gamma_{\tau_n}) \subset (\partial_S\phi(\mu_{\tau_n}))$ be the corresponding sequence of strong subdifferentials converging in the W_2 -metric towards $(\text{Id} \times \nabla_\mu\phi(\mu))_{\#}\mu$. For any $\sigma \in [-\epsilon, \epsilon]^N$, define the family of 3-plans $\mu_\sigma^{\tau_n} = (\pi^1, \pi^2, \mathcal{G}(\sigma, \cdot) \circ \pi^1)_{\#}\gamma_{\tau_n}$. By (1.12), it holds that

$$\frac{\phi(\mathcal{G}(\sigma, \cdot)_{\#}\mu_{\tau_n}) - \phi(\mu_{\tau_n})}{|\sigma|} \geq \int_{\mathbb{R}^{2d}} \left\langle r, \frac{\mathcal{G}(\sigma, x) - x}{|\sigma|} \right\rangle d\gamma_{\tau_n}(x, r) + o(1),$$

since $o(W_{2, \mu_\sigma^{\tau_n}}(\mu_{\tau_n}, \mathcal{G}(\sigma, \cdot)_{\#}\mu_{\tau_n})) = o(\|\mathcal{G}(\sigma, \cdot) - \text{Id}\|_{L^2(\mu_{\tau_n})}) = o(|\sigma|)$. For all $\sigma \in [-\epsilon, \epsilon]^N$, the maps $(x, r) \mapsto \langle r, (\mathcal{G}(\sigma, x) - x)/|\sigma| \rangle$ are continuous and uniformly integrable with respect to the family of measures $\{\gamma_{\tau_n}\}_{n=1}^{+\infty}$. By a classical convergence result for sequences of measures (see e.g. [17, Lemma 5.1.7]), we recover that

$$\begin{aligned} \int_{\mathbb{R}^{2d}} \left\langle r, \frac{\mathcal{G}(\sigma, x) - x}{|\sigma|} \right\rangle d\gamma_{\tau_n}(x, r) &\xrightarrow{n \rightarrow +\infty} \int_{\mathbb{R}^{2d}} \left\langle r, \frac{\mathcal{G}(\sigma, x) - x}{|\sigma|} \right\rangle d(\partial^\circ\phi(\mu))(x, r) \\ &= \int_{\mathbb{R}^d} \left\langle \nabla_\mu\phi(\mu)(x), \frac{\mathcal{G}(\sigma, x) - x}{|\sigma|} \right\rangle d\mu(x). \end{aligned} \quad (1.22)$$

Furthermore, the continuity of $x \mapsto \mathcal{G}(\sigma, x)$ uniformly with respect to $\sigma \in [-\epsilon, \epsilon]^N$

together with the Lipschitzianity of $\phi(\cdot)$ with respect to the W_2 -topology implies that

$$\frac{\phi(\mathcal{G}(\sigma, \cdot) \# \mu_{\tau_n}) - \phi(\mu_{\tau_n})}{|\sigma|} \xrightarrow{n \rightarrow +\infty} \frac{\phi(\mathcal{G}(\sigma, \cdot) \# \mu_\tau) - \phi(\mu_\tau)}{|\sigma|}. \quad (1.23)$$

Merging together (1.22) and (1.23) and applying Vitali's Convergence Theorem to the μ -uniformly integrable family of maps $\sigma \mapsto \langle \nabla_\mu \phi(\mu)(\cdot), (\mathcal{G}(\sigma, \cdot) - \text{Id})/|\sigma| \rangle$ (see e.g. [14, Exercise 1.18]), we obtain that

$$\begin{aligned} \frac{\phi(\mathcal{G}(\sigma, \cdot) \# \mu) - \phi(\mu)}{|\sigma|} &\geq \int_{\mathbb{R}^d} \langle \nabla_\mu \phi(\mu)(x), D_e \mathcal{G}(0, x) \sigma \rangle d\mu(x) + o(1) \\ &= \sum_{k=1}^N \sigma_k \int_{\mathbb{R}^d} \langle \nabla_\mu \phi(\mu)(x), \mathcal{F}_k(x) \rangle d\mu(x) + o(1). \end{aligned}$$

Using the fact that $(\text{Id} \times (-\nabla_\mu \phi(\mu))) \# \mu$ is the minimal selection in the extended Fréchet subdifferential of $(-\phi(\cdot))$ as a consequence of Proposition 1.7, we recover the converse inequality. \square

Remark 1.4 (Comparing Proposition 1.6 and Proposition 1.8). *Notice that in order to obtain the chain rule formula of Proposition 1.6, we need to posit a priori that the map $\epsilon \mapsto \phi(\mathcal{G}(\epsilon, \cdot) \# \mu)$ is differentiable at $\epsilon = 0$. In Proposition 1.8, this differentiability is recovered a posteriori as a consequence of the stronger Wasserstein differentiability hypothesis made on $\phi(\cdot)$.*

1.2.3 Second order calculus in Wasserstein spaces

In Chapter 4, we are also going to need a notion of intrinsic second order derivative for functionals defined over $\mathcal{P}_2(\mathbb{R}^d)$. We therefore introduce in the following definition the notion of *Wasserstein Hessian bilinear form* for a sufficiently regular functional $\phi(\cdot)$ defined over $\mathcal{P}_2(\mathbb{R}^d)$.

Definition 1.11 (Hessian bilinear form in $(\mathcal{P}_2(\mathbb{R}^d), W_2)$). *Suppose that $\phi(\cdot)$ is differentiable at $\mu \in D(\phi)$ in the sense of Definition 1.10 and that for any $\xi \in \nabla C_c^\infty(\mathbb{R}^d)$, the Lie derivative map*

$$\mathcal{L}_\xi \phi : \nu \in \mathcal{P}_2(\mathbb{R}^d) \mapsto \langle \nabla_\mu \phi(\nu), \xi \rangle_{L^2(\nu)} = \int_{\mathbb{R}^d} \langle \nabla_\mu \phi(\nu)(x), \xi(x) \rangle d\nu(x)$$

is also differentiable at μ . Then, we define the partial Wasserstein Hessian of $\phi(\cdot)$ at μ as the bilinear form

$$\text{Hess } \phi[\mu](\xi_1, \xi_2) = \mathcal{L}_{\xi_2} (\mathcal{L}_{\xi_1} \phi(\mu)) - \mathcal{L}_{D\xi_1 \xi_2} \phi(\mu) \quad (1.24)$$

for any $\xi_1, \xi_2 \in \nabla C_c^\infty(\mathbb{R}^d)$. If moreover there exists a positive constant $C_\mu > 0$ such that

$$\text{Hess } \phi[\mu](\xi_1, \xi_2) \leq C_\mu \|\xi_1\|_{L^2(\mu)} \|\xi_2\|_{L^2(\mu)},$$

we denote again by $\text{Hess } \phi[\mu](\cdot, \cdot)$ its extension to $\text{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d) \times \text{Tan}_\mu \mathcal{P}_2(\mathbb{R}^d)$ and we say that $\phi(\cdot)$ is twice differentiable at μ .

By definition of the Lie derivative in $(\mathcal{P}_2(\mathbb{R}^d), W_2)$, the Wasserstein Hessian can be written more explicitly as

$$\text{Hess } \phi[\mu](\xi_1, \xi_2) = \int_{\mathbb{R}^d} \langle \nabla_{\mu} \mathcal{L}_{\xi_1} \phi(\mu)(x), \xi_2(x) \rangle d\mu(x) - \int_{\mathbb{R}^d} \langle \nabla_{\mu} \phi(\mu)(x), D\xi_1(x) \xi_2(x) \rangle d\mu(x).$$

In the following Proposition, we provide a condensed statement of several listed listed in [57, Section 3]. This allows us to derive an analytical and natural expression for the Hessian bilinear form, as well as a second-order differentiation formula for Wasserstein functionals.

Proposition 1.9 (Expression of the Wasserstein Hessian and second-order expansion). *Let $\phi : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ be a proper and lower-semicontinuous functional differentiable at $\mu \in D(\phi)$ in the sense of Definition 1.10. Furthermore, suppose that the maps $y \in \mathbb{R}^d \mapsto \nabla_{\mu} \phi(\mu)(y)$ and $\nu \in \mathcal{P}_2(\mathbb{R}^d) \mapsto \nabla_{\mu} \phi(\nu)(x)$ are differentiable at $x \in \mathbb{R}^d$ and $\mu \in D(\phi)$ respectively. Then, $\phi(\cdot)$ is twice differentiable in the sense of Definition 1.11, and its Wasserstein Hessian is given explicitly by*

$$\begin{aligned} \text{Hess } \phi[\mu](\xi_1, \xi_2) &= \int_{\mathbb{R}^d} \langle D_x \nabla_{\mu} \phi(\mu)(x) \xi_1(x), \xi_2(x) \rangle d\mu(x) \\ &+ \int_{\mathbb{R}^{2d}} \langle D_{\mu}^2 \phi(\mu)(x, y) \xi_1(x), \xi_2(y) \rangle d\mu(x) d\mu(y), \end{aligned} \quad (1.25)$$

for any $\xi_1, \xi_2 \in \text{Tan}_{\mu} \mathcal{P}_2(\mathbb{R}^d)$. Here, the map $D_x \nabla_{\mu} \phi(\mu)(x) \in \mathbb{R}^{d \times d}$ is the classical differential of $\nabla_{\mu} \phi(\mu)(\cdot)$ at $x \in \mathbb{R}^d$ while $D_{\mu}^2 \phi(\mu)(x, \cdot) : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$ denotes the matrix-valued map which columns are the Wasserstein gradients of the components of $\nabla_{\mu} \phi(\mu)(x)$ defined as in Definition 1.10. Moreover for any $\xi_1, \xi_2 \in \nabla C_c^{\infty}(\mathbb{R}^d)$, it holds that

$$\frac{d}{ds} \mathcal{L}_{\xi_1} \phi((\text{Id} + s\xi_2)_{\#} \mu) = \text{Hess } \phi[\mu](\xi_1, \xi_2) + \mathcal{L}_{D\xi_1 \xi_2} \phi(\mu). \quad (1.26)$$

Proof. The explicit expression (1.25) can be derived by following the proof of [57, Theorem 3.2] which deploys a more general argument. The second order differentiation formula (1.26) can be recovered as a direct consequence of Proposition 1.5 and of the Definition (1.24) of the Wasserstein Hessian. \square

1.3 The continuity equation with non-local velocities in \mathbb{R}^d

In this section, we introduce several facts about continuity equations with non-local velocities in $(\mathcal{P}_c(\mathbb{R}^d), W_1)$. We refer the reader to [15, 114] for a study of non-local equations and to [11, 12, 68] and [17, Chapter 8] for a detailed account on some of the underlying tools developed originally for continuity equations.

Transport equations with non-local velocities can be generally written as

$$\partial_t \mu(t) + \nabla \cdot (v[\mu(t)](t, \cdot) \mu(t)) = 0, \quad (1.27)$$

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 for any $\mu \in \mathcal{P}_c(\mathbb{R}^d)$, satisfying the

condition

$$\int_0^T \int_{\mathbb{R}^d} |v[\mu(t)](t, x)| d\mu(t)(x) dt < +\infty. \quad (1.28)$$

Equation (1.27) has to be understood in the sense of distributions, i.e.

$$\int_0^T \int_{\mathbb{R}^d} \left(\partial_t \xi(t, x) + \langle \nabla_x \xi(t, x), v[\mu(t)](t, x) \rangle \right) d\mu(t)(x) dt = 0 \quad (1.29)$$

for all $\xi \in C_c^\infty([0, T] \times \mathbb{R}^d)$, or alternatively as

$$\frac{d}{dt} \int_{\mathbb{R}^d} \xi(x) d\mu(t)(x) = \int_{\mathbb{R}^d} \langle \nabla \xi(x), v[\mu(t)](t, x) \rangle d\mu(t)(x) \quad (1.30)$$

for all $\xi \in C_c^\infty(\mathbb{R}^d)$ and \mathcal{L}^1 -almost every $t \in [0, T]$. As highlighted previously, these equations are interesting for a large number of applications in the field of multi-agent dynamics.

We state in the following theorem a general existence result for solutions of continuity equations of the form (1.27) under mere L^p -integrability of the driving velocity field with $p \geq 1$. We refer the reader to the seminal papers [11, 68] as well as to [12, Section 3] and [82, Section 6.1].

Theorem 1.6 (Superposition principle). *Let $\mu \in C^0([0, T], \mathcal{P}_p(\mathbb{R}^d))$ be a curve of measures and $v \in C^0(\mathcal{P}_p(\mathbb{R}^d), L^p([0, T] \times \mathbb{R}^d, \mathbb{R}^d; \tilde{\mu}))$ be a non-local Borel velocity field. Then, $\mu(\cdot)$ is a solution of (1.27) associated to $v[\cdot](\cdot, \cdot)$ if and only if there exists a probability measure $\boldsymbol{\eta} \in \mathcal{P}_p(\mathbb{R}^d \times \text{AC}([0, T], \mathbb{R}^d))$ such that*

- (i) $\boldsymbol{\eta}$ is concentrated on the set of pairs $(x, \gamma(\cdot)) \in \mathbb{R}^d \times \Gamma_T$ such that $\dot{\gamma}(t) = v[\mu(t)](t, \gamma(t))$ for \mathcal{L}^1 -almost every $t \in [0, T]$ and $\gamma(0) = x$.
- (ii) It holds that $\mu(t) = (e_t)_\# \boldsymbol{\eta}$ where for all times $t \in [0, T]$ we introduced the evaluation map $e_t : (x, \gamma(\cdot)) \in \mathbb{R}^d \times \text{AC}([0, T], \mathbb{R}^d) \mapsto \gamma(t) \in \mathbb{R}^d$.

In the following theorem, we state a stronger classical well-posedness result for (1.27) under additional regularity assumptions on the non-local velocity field $v[\cdot](\cdot, \cdot)$. This result, which was originally derived in [15], incorporates uniqueness properties and representation formula for solutions of (1.27). We adopt here a formalism closer to that of [114] which is more suited to our smooth control-theoretic framework.

Theorem 1.7 (Existence, uniqueness and representation of solutions for (1.27)). *Consider a non-local velocity field $v[\cdot](\cdot, \cdot)$ defined as*

$$v : \mu \in \mathcal{P}_c(\mathbb{R}^d) \mapsto v[\mu](\cdot, \cdot) \in L^1(\mathbb{R}, C^1(\mathbb{R}^d, \mathbb{R}^d)), \quad (1.31)$$

and satisfying the following Cauchy-Lipschitz assumptions.

(H')

(i) *There exists positive constants L_1 and M such that*

$$|v[\mu](t, x) - v[\mu](t, y)| \leq L_1 |x - y| \quad \text{and} \quad |v[\mu](t, x)| \leq M(1 + |x|)$$

for every $\mu \in \mathcal{P}_c(\mathbb{R}^d)$, $t \in \mathbb{R}$ and $(x, y) \in \mathbb{R}^{2d}$;
(ii) There exists a positive constant L_2 such that

$$\|v[\mu](t, \cdot) - v[\nu](t, \cdot)\|_{C^0(\mathbb{R}^d)} \leq L_2 W_1(\mu, \nu)$$

for every $\mu, \nu \in \mathcal{P}_c(\mathbb{R}^d)$ and $t \in \mathbb{R}$;

Then for every initial datum $\mu^0 \in \mathcal{P}_c(\mathbb{R}^d)$, the Cauchy problem

$$\begin{cases} \partial_t \mu(t) + \nabla \cdot (v[\mu(t)](t, \cdot) \mu(t)) = 0 \\ \mu(0) = \mu^0, \end{cases} \quad (1.32)$$

admits a unique solution $\mu(\cdot)$ in $\text{Lip}_{\text{loc}}(\mathbb{R}, \mathcal{P}_c(\mathbb{R}^d))$ locally Lipschitz with respect to the W_1 -metric. Besides, if μ^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 $R_T > 0$ depending on $\text{supp}(\mu^0)$ and $C_T > 0$ such that

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

for all times $t \in [0, T]$ and any solutions $\mu(\cdot), \nu(\cdot)$ of (1.32).

Let $(\Phi_{(0,t)}^v[\mu^0](\cdot))_{t \geq 0}$ be the family of flows of diffeomorphisms generated by the non-local vector field $v[\mu(t)](t, \cdot)$, defined as the unique solution of

$$\begin{cases} \partial_t \Phi_{(0,t)}^v[\mu^0](x) = v[\mu(t)]\left(t, \Phi_{(0,t)}^v[\mu^0](x)\right), \\ \Phi_{(0,0)}^v[\mu^0](x) = x \quad \text{for all } x \text{ in } \mathbb{R}^d. \end{cases} \quad (1.33)$$

Then, the unique solution of the Cauchy problem (1.32) is given by

$$\mu(t) = \Phi_{(0,t)}^v[\mu^0](\cdot) \# \mu^0 \quad (1.34)$$

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

In [11, 78], it is mentioned that the only regularity framework for (1.27) allowing to encompass both discrete and absolutely continuous measures is that of Theorem 1.7. Indeed, the powerful results of Theorem 1.6 are intrinsically macroscopic, and allow for solutions supported on crossing characteristics to exist. In the context of optimal control problems, providing a general sufficient conditions for such a system to be well-posed in the sense of Theorem 1.7 is then of major interest. This is precisely the object of our investigations in Chapter 4.

1.4 Directional derivatives of non-local flows

In this section, we prove the existence of directional derivatives along measure curves generated by suitable Lipschitz families of continuous and bounded maps for non-local flows. Such derivatives are characterized as the only solution of a linearized Cauchy

problem. This result can be seen as a generalization to the Wasserstein setting of Proposition 1.10. We recall in the following Proposition a standard result which links the differential of the flow of diffeomorphisms of an ODE at time $t \in [0, T]$ to the solution of a corresponding linearized Cauchy problem (see e.g. [37]).

Proposition 1.10 (Differential of a flow). *Let $(t, x) \mapsto v(t, x)$ be measurable in t as well as sublinear and C^1 in x . Define the family of C^1 -flows $(\Phi_t^v(\cdot))_{t \geq 0}$ associated to $v(\cdot, \cdot)$ by (1.33) in the case where $v(\cdot, \cdot)$ is independent from $\mu(\cdot)$.*

Then, it holds that the differential $D_x \Phi_{(s,t)}^v(x) \cdot h$ of the flow between times s and t , evaluated at x and applied to some vector $h \in \mathbb{R}^d$ is the unique solution $w(\cdot, x)$ of the linearized Cauchy problem

$$\partial_t w(t, x) = D_x v(t, \Phi_{(s,t)}^v(x)) \cdot w(t, x), \quad w(s, x) = h. \quad (1.35)$$

We shall prove in Theorem 1.9 a generalization of this result in the non-local case where the initial measure is perturbed by a continuous family of continuous and bounded maps. Such a result is crucial to study the first order perturbation induced by a needle-like variation on a measure curve in the non-local setting, and is used extensively in the proofs of Theorem 2.1 and Theorem 3.1.

Before stating our result, we recall the classical Banach Fixed Point Theorem with parameter (see e.g. [37, Theorem A.2.1]).

Theorem 1.8 (Banach fixed point theorem with parameter). *Let X be a Banach space, S be a metric space and $\Lambda : X \times S \rightarrow X$ be a continuous mapping such that, for some $\kappa < 1$,*

$$\|\Lambda(x, s) - \Lambda(y, s)\|_X \leq \kappa \|x - y\|_X \quad \text{for all } x, y \in X \text{ and } s \in S.$$

Then for each $s \in S$, there exists a unique fixed point $x(s) \in X$ of $\Lambda(\cdot, s)$. Moreover, the map $s \mapsto x(s)$ is continuous and for any $(s, y) \in S \times X$, it holds

$$\|y - x(s)\|_X \leq \frac{1}{1 - \kappa} \|y - \Lambda(s, y)\|_X. \quad (1.36)$$

We are now ready to state and prove the main result of this Section.

Theorem 1.9 (Directional derivative of a non-local flow with respect to the initial data). *Let $K \subset \mathbb{R}^d$ be a compact set and $\mu \in \mathcal{P}(K)$. Let $v[\cdot](\cdot, \cdot)$ be a non-local vector field satisfying hypotheses (H') of Theorem 1.7, $\Phi_{(0,\cdot)}^v\cdot$ be the corresponding family of non-local flows as defined in Theorem 1.7 and $\mu(\cdot)$ be the unique solution of the corresponding Cauchy problem (1.32) starting from μ .*

Let $\bar{\epsilon} > 0$ be a small parameter and $\mathcal{G}(\cdot, \cdot) \in \text{Lip}((-\bar{\epsilon}, \bar{\epsilon}), C^0(\mathbb{R}^d, \mathbb{R}^d))$ be a family of bounded maps satisfying the assumptions of Proposition 1.6. Furthermore, assume that the following holds.

- (a) *For any $i \in \{1, \dots, d\}$ and $(t, x) \in [0, T] \times \mathbb{R}^d$, the component maps $\mu \in \mathcal{P}_c(\mathbb{R}^d) \mapsto v^i[\mu](t, x)$ are regular in the sense of Definition 1.8, and the barycenters $\bar{\gamma}_{(t,x)}^{i,\circ}(\cdot)$ of their minimal subdifferentials at μ are continuous.*
- (b) *The curves $\epsilon \in (-\bar{\epsilon}, \bar{\epsilon}) \mapsto v^i[\mathcal{G}(\epsilon, \cdot) \# \mu](t, x)$ are differentiable at $\epsilon = 0$, uniformly with respect to $(t, x) \in [0, T] \times \mathbb{R}^d$.*

Then, the map $\epsilon \in (-\bar{\epsilon}, \bar{\epsilon}) \mapsto \Phi_{(0,t)}^v[\mathcal{G}(\epsilon, \cdot) \# \mu](x)$ admits a derivative at $\epsilon = 0$ for all $(t, x) \in [0, T] \times \overline{B(0, R_T)}$ that we denote by $w_\Phi(t, x)$. It can be characterised as the unique solution of the Cauchy problem

$$\begin{cases} \partial_t w(t, x) = D_x v[\mu(t)] \left(t, \Phi_{(0,t)}^v[\mu](x) \right) w(t, x) \\ + \int_{\mathbb{R}^d} \mathbb{I}_{\left(t, \Phi_{(0,t)}^v[\mu](x) \right)}^\circ \left(\Phi_{(0,t)}^v[\mu](y) \right) \cdot \left[D_x \Phi_{(0,t)}[\mu](y) \mathcal{F}(y) + w(t, y) \right] d\mu(y), \\ w(0, x) = 0 \end{cases} \quad \text{for all } x \in \mathbb{R}^d, \quad (1.37)$$

where for all (t, z) , $\mathbb{I}_{(t,z)}^\circ(\cdot)$ is the matrix-valued map made of the barycenters of the minimal selections $\partial_\mu^\circ v^i[\mu(t)](t, z)$ in the extended subdifferential of the components of $\mu \mapsto v^i[\mu](t, z)$ at $\mu(t)$.

Proof. We follow a classical scheme of proof used in the finite dimensional setting to show that flows of diffeomorphisms admit directional derivatives characterized as the unique solution of a linearized Cauchy problem (see e.g. [37, Theorem 2.3.1]).

We first define $\Omega = \overline{B(0, R_T)}$ where $R_T > 0$ is defined as in Theorem 1.7, and we introduce the operator $\Lambda_\Phi : w \in C^0([0, T] \times \Omega, \mathbb{R}^d) \mapsto \Lambda_\Phi(w) \in C^0([0, T] \times \Omega, \mathbb{R}^d)$ defined for all $(t, x) \in [0, T] \times \Omega$ by

$$\begin{aligned} \Lambda_\Phi(w)(t, x) &= \int_0^t D_x v[\mu(s)] \left(s, \Phi_{(0,s)}^v[\mu](x) \right) w(s, x) ds \\ &+ \int_0^t \int_{\mathbb{R}^d} \mathbb{I}_{\left(s, \Phi_{(0,s)}^v[\mu](x) \right)}^\circ \left(\Phi_{(0,s)}^v[\mu](y) \right) \cdot \left[D_x \Phi_{(0,s)}[\mu](y) \mathcal{F}(y) + w(s, y) \right] d\mu(y) ds. \end{aligned}$$

The right hand side of the previous equation is continuous in (t, x) as a consequence of our standing assumptions (\mathbf{H}°) of Theorem 1.7 and (a). We first show that this operator admits a unique fixed point and afterwards that it coincides with the map which to every (t, x) associates the derivative at $\epsilon = 0$ of the family of non-local flows $\epsilon \mapsto \Phi_{(0,t)}^v[\mathcal{G}(\epsilon, \cdot) \# \mu](x)$. With this goal, we introduce a parameter $\alpha > 0$ that will be chosen so that the operator $\Lambda_\Phi(\cdot)$ is contracting with respect to the equivalent norm

$$\|w\|_{C^0([0,T] \times \Omega)}^\alpha = \sup_{(t,x) \in [0,T] \times \Omega} e^{-2\alpha t} |w(t, x)|. \quad (1.38)$$

Remark that for any $w_1, w_2 \in C^0([0, T] \times \Omega, \mathbb{R}^d)$ and any $(t, x) \in [0, T] \times \Omega$, it holds

$$\begin{aligned} &|\Lambda_\Phi(w_2)(t, x) - \Lambda_\Phi(w_1)(t, x)| \\ &\leq \int_0^t \left| D_x v[\mu(s)] \left(s, \Phi_{(0,s)}^v[\mu](x) \right) \cdot (w_2(s, x) - w_1(s, x)) \right| ds \\ &+ \int_0^t \int_{\mathbb{R}^d} \left| \mathbb{I}_{\left(s, \Phi_{(0,s)}^v[\mu](x) \right)}^\circ \left(\Phi_{(0,s)}^v[\mu](y) \right) \cdot (w_2(s, y) - w_1(s, y)) \right| d\mu(y) ds \\ &\leq \int_0^t \left(L_1 |w_2(s, x) - w_1(s, x)| + L_2 \|w_2(s, \cdot) - w_1(s, \cdot)\|_{L^1(\mu)} \right) ds \\ &\leq \int_0^t (L_1 + L_2) \|w_2(s, \cdot) - w_1(s, \cdot)\|_{C^0(\Omega)} ds, \end{aligned}$$

since $\mu(\Omega) = 1$, and where we introduced

$$L_1 = \left\| D_x v[\mu(\cdot)](\cdot, \Phi_{(0,\cdot)}^v[\mu](\cdot)) \right\|_{L^\infty([0,T] \times \Omega; \mathcal{L}^1 \times \mu(\cdot))},$$

and

$$L_2 = \left\| \mathbb{I}^\circ \left(\cdot, \Phi_{(0,\cdot)}^v[\mu](\cdot) \right) \left(\Phi_{(0,\cdot)}^v[\mu](\cdot) \right) \right\|_{L^\infty([0,T] \times \Omega^2; \mathcal{L}^1 \times \mu(\cdot) \times \mu(\cdot))}$$

which exist again as a consequence of the differentiability assumptions on $v[\cdot](\cdot, \cdot)$ described in **(H')** and (a). It further holds by definition of $\|\cdot\|_{C^0([0,T] \times \Omega)}^\alpha$ that

$$\begin{aligned} |\Lambda_\Phi(w_2)(t, x) - \Lambda_\Phi(w_1)(t, x)| &\leq \int_0^t e^{2\alpha s} (L_1 + L_2) \|w_2(\cdot, \cdot) - w_1(\cdot, \cdot)\|_{C^0([0,T] \times \Omega)}^\alpha ds \\ &\leq \frac{e^{2\alpha t} - 1}{2\alpha} (L_1 + L_2) \|w_2(\cdot, \cdot) - w_1(\cdot, \cdot)\|_{C^0([0,T] \times \Omega)}^\alpha. \end{aligned}$$

Multiplying both sides of the inequality by $e^{-2\alpha t}$ and taking the supremum over $(t, x) \in [0, T] \times \Omega$ in the left-hand side yields the desired contractivity with a constant equal to $1/2$ provided that $\alpha \geq (L_1 + L_2)$. It is then possible to apply Theorem 1.8 to obtain the existence of a unique fixed point $w_\Phi(\cdot, \cdot) \in C^0([0, T] \times \Omega, \mathbb{R}^d)$ of $\Lambda_\Phi(\cdot)$.

Define for $\epsilon \in (-\bar{\epsilon}, \bar{\epsilon})$ the parametrized family of operators $\Psi^\epsilon : f \in C^0([0, T] \times \Omega, \mathbb{R}^d) \mapsto \Psi^\epsilon(f) \in C^0([0, T] \times \Omega, \mathbb{R}^d)$ defined by

$$\Psi^\epsilon(f)(t, x) = x + \int_0^t v[f(s, \cdot) \# (\mathcal{G}(\epsilon, \cdot) \# \mu)](s, f(s, x)) ds \quad (1.39)$$

for all $(t, x) \in [0, T] \times \Omega$. Up to defining again the equivalent norm $\|\cdot\|_{C^0([0,T] \times \Omega)}^\alpha$ as in (1.38) with a suitable $\alpha > 0$, it can be shown that this operator is contracting independently from ϵ as a direct consequence of the Lipschitzianity hypotheses given in hypotheses **(H')**. We can thus invoke again Theorem 1.8 to obtain the existence of a unique fixed point of $\Psi^\epsilon(\cdot)$ for each $\epsilon \in (-\bar{\epsilon}, \bar{\epsilon})$. Notice that by definition, this family of fixed points is precisely the parametrized family of non-local flows $(t, x) \mapsto \Phi_{(0,t)}^v[\mathcal{G}(\epsilon, \cdot) \# \mu](x)$. As a consequence of Theorem 1.7, we know that these maps are C^1 with respect to x for all ϵ .

We now define the map $\hat{\Phi}_{(0,\cdot)}^{v,\epsilon}[\mu](\cdot)$ by

$$\hat{\Phi}_{(0,\cdot)}^{v,\epsilon}[\mu](\cdot) : (t, x) \mapsto \Phi_{(0,t)}^v[\mu](x) + \epsilon w_\Phi(t, x).$$

To conclude, we then need to show that

$$\lim_{\epsilon \rightarrow 0} \left\| \frac{1}{\epsilon} \left(\hat{\Phi}_{(0,\cdot)}^{v,\epsilon}[\mu](\cdot) - \Phi_{(0,\cdot)}^v[\mathcal{G}(\epsilon, \cdot) \# \mu](\cdot) \right) \right\|_{C^0([0,T] \times \Omega)} = 0,$$

which will directly yield the existence and the characterization of the directional derivative of the flow along $(-\bar{\epsilon}, \bar{\epsilon}) \mapsto \mathcal{G}(\epsilon, \cdot) \# \mu^0$. By (1.36) in Theorem 1.8 and the equivalence of the C^0 -norms we introduced, there exists a constant $C > 0$ independent from ϵ such that

it holds

$$\frac{1}{|\epsilon|} \left\| \hat{\Phi}_{(0,\cdot)}^{v,\epsilon}[\mu](\cdot) - \Phi_{(0,\cdot)}^v[\mathcal{G}(\epsilon, \cdot) \# \mu](\cdot) \right\|_{C^0} \leq \frac{2C}{|\epsilon|} \left\| \hat{\Phi}_{(0,\cdot)}^{v,\epsilon}[\mu](\cdot) - \Psi^\epsilon(\hat{\Phi}_{(0,\cdot)}^{v,\epsilon}[\mu](\cdot))(\cdot, \cdot) \right\|_{C^0}.$$

We now want to perform a first order expansion on $\Psi^\epsilon(\hat{\Phi}_{(0,s)}^{v,\epsilon}[\mu](\cdot))(\cdot, \cdot)$ with respect to ϵ . Take $(s, x) \in [0, T] \times \Omega$. One has by definition of $\hat{\Phi}_{(0,\cdot)}^{v,\epsilon}[\mu](\cdot)$ that

$$\begin{aligned} \hat{\Phi}_{(0,s)}^{v,\epsilon}[\mu](\mathcal{G}(\epsilon, x)) &= \Phi_{(0,s)}^v[\mu](\mathcal{G}(\epsilon, x)) + \epsilon w_\Phi(s, \mathcal{G}(\epsilon, x)) \\ &= \Phi_{(0,s)}^v[\mu](x) + \epsilon \left(D_x \Phi_{(0,s)}^v[\mu](x) \mathcal{F}(x) + w_\Phi(s, x) \right) + o(\epsilon), \end{aligned}$$

by continuity of $w_\Phi(s, \cdot)$ for all $s \in [0, T]$.

As a consequence of the differentiability hypotheses on $v[\cdot](\cdot, \cdot)$ along with assumptions (a)-(b), we can apply the chain rule of Proposition 1.6 component-wise on the v^i to obtain that

$$\begin{aligned} v \left[\hat{\Phi}_{(0,s)}^{v,\epsilon}[\mu](\cdot) \circ \mathcal{G}(\epsilon, \cdot) \# \mu \right] (s, z) &= v[\Phi_{(0,s)}^v[\mu](\cdot) \# \mu](s, z) \\ + \epsilon \int_{\mathbb{R}^d} \mathbb{I}_{(s,z)}^\circ \left(\Phi_{(0,s)}^v[\mu](y) \right) &\left[D_x \Phi_{(0,s)}^v[\mu](y) \mathcal{F}(y) + w_\Phi(s, y) \right] d\mu(y) + o(\epsilon) \end{aligned} \quad (1.40)$$

where for all (s, z) the map $y \mapsto \mathbb{I}_{(s,z)}^\circ(y) = (\tilde{\gamma}_{(s,z)}^{i,\circ}(y))_{1 \leq i \leq d} \in \mathbb{R}^{d \times d}$ is made of the barycenters of the minimal selections in the extended subdifferentials of the components v^i 's.

Performing a Taylor expansion in the space variable for the non-local velocity field, it also holds that

$$\begin{aligned} v \left[\Phi_{(0,s)}^v[\mu](\cdot) \# \mu \right] \left(s, \Phi_{(0,s)}^v[\mu](x) + \epsilon w_\Phi(s, x) \right) &= v[\Phi_{(0,s)}^v[\mu](\cdot) \# \mu] \left(s, \Phi_{(0,s)}^v[\mu](x) \right) \\ + \epsilon D_x v[\Phi_{(0,s)}^v[\mu](\cdot) \# \mu] \left(s, \Phi_{(0,s)}^v[\mu](x) \right) &\cdot w_\Phi(s, x) + o(\epsilon), \end{aligned} \quad (1.41)$$

as well as

$$\mathbb{I}_{\left(s, \Phi_{(0,s)}^v[\mu](x) + \epsilon w_\Phi(s, x) \right)}^\circ \left(\Phi_{(0,s)}^v[\mu](y) \right) = \mathbb{I}_{\left(s, \Phi_{(0,s)}^v[\mu](x) \right)}^\circ \left(\Phi_{(0,s)}^v[\mu](y) \right) + o(1) \quad (1.42)$$

as a consequence of assumption (a) which implies in particular that $z \mapsto \mathbb{I}_{(t,z)}^\circ(x)$ is continuous for all $(s, x) \in [0, T] \times \mathbb{R}^d$.

Merging (1.39), (1.40) and (1.41), (1.42) and recalling the definition of $w_\Phi(\cdot, \cdot)$, it holds

$$\begin{aligned} \Psi^\epsilon \left(\hat{\Phi}_{(0,\cdot)}^{v,\epsilon}[\mu](\cdot) \right) (t, x) &= x + \int_0^t v[\Phi_{(0,s)}^v[\mu](\cdot) \# \mu](s, \Phi_{(0,s)}^v[\mu](x)) ds \\ + \epsilon \int_0^t \int_{\mathbb{R}^d} \mathbb{I}_{\left(s, \Phi_{(0,s)}^v[\mu](x) \right)}^\circ &\left(\Phi_{(0,s)}^v[\mu](y) \right) \left[D_x \Phi_{(0,s)}^v[\mu](y) \mathcal{F}(y) + w_\Phi(s, y) \right] d\mu(y) ds \\ + \epsilon \int_0^t D_x v[\Phi_{(0,s)}^v[\mu](\cdot) \# \mu] &\left(s, \Phi_{(0,s)}^v[\mu](x) \right) \cdot w_\Phi(s, x) ds + o(\epsilon) \\ = \Phi_{(0,t)}^v[\mu](x) + \epsilon w_\Phi(t, x) &+ o(\epsilon). \end{aligned}$$

Therefore, we finally recover that

$$\frac{1}{|\epsilon|} \left| \Psi^\epsilon(\hat{\Phi}_{(0,\cdot)}^\epsilon[\mu](\cdot))(t,x) - \hat{\Phi}_{(0,t)}^\epsilon[\mu](x) \right| \leq o(1)$$

as $\epsilon \rightarrow 0$ for all $(t,x) \in [0,T] \times \Omega$, and conclude that

$$\lim_{\epsilon \rightarrow 0} \left\| \frac{1}{\epsilon} \left(\Phi_{(0,\cdot)}[\mathcal{G}(\epsilon, \cdot) \# \mu](\cdot) - \hat{\Phi}_{(0,\cdot)}^\epsilon[\mu](\cdot) \right) \right\|_{C^0([0,T] \times \Omega)} = 0.$$

We have thus proven that the derivative of $\epsilon \in (-\bar{\epsilon}, \bar{\epsilon}) \mapsto \Phi_{(0,t)}^v[\mathcal{G}(\epsilon, \cdot) \# \mu](x)$ at $\epsilon = 0$ exists for any (t,x) and that it is the only solution of equation (1.37). \square

In the following proposition, we prove an extension of Theorem 1.9 to the case of multi-dimensional perturbations of the initial measure of a non-local continuity equation. This result is instrumental in the proof of Theorem 3.1.

Proposition 1.11 (Wasserstein differential of a non-local flow of diffeomorphisms). *Let $K \subset \mathbb{R}^d$ be a compact set, $\mu^0 \in \mathcal{P}(K)$, $v[\cdot](\cdot, \cdot)$ be a non-local velocity field satisfying hypotheses (H') of Theorem 1.7 and $(\Phi_{(0,t)}^v[\mu^0](\cdot))_{t \in [0,T]}$ be its associated flow of diffeomorphisms. Suppose moreover that the maps $\mu \mapsto v^i[\mu](t,x)$ are regular and continuously differentiable over $\mathcal{P}(K)$ uniformly with respect to $(t,x) \in [0,T] \times K$.*

Given $N \geq 1$ and a small parameter $\bar{\epsilon} > 0$, let $\mathcal{G} \in C^0([-\bar{\epsilon}, \bar{\epsilon}]^N \times K, \mathbb{R}^d)$ be a function satisfying the hypotheses of Proposition 1.8. Then, the map $e \in [-\bar{\epsilon}, \bar{\epsilon}]^N \mapsto \Phi_{(s,t)}^v[\mathcal{G}(e, \cdot) \# \mu](x)$ is Fréchet-differentiable at $e = 0$ and its differential $w_\sigma(\cdot, x)$ in an arbitrary direction $\sigma \in [-\bar{\epsilon}, \bar{\epsilon}]^N$ 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

$$\begin{cases} \partial_t w_k(t,x) = D_x v[\mu(t)] \left(t, \Phi_{(s,t)}^v[\mu](x) \right) w_k(t,x) \\ \quad + \int_{\mathbb{R}^d} \mathbb{I}_{\left(t, \Phi_{(s,t)}^v[\mu](x) \right)}^v \left(\Phi_{(s,t)}^v[\mu](y) \right) \left(D_x \Phi_{(s,t)}^v[\mu](y) \mathcal{F}_k(y) + w_k(t,y) \right) d\mu(y) \\ w_k(s,x) = 0 \end{cases} \quad \text{for all } x \in \mathbb{R}^d. \quad (1.43)$$

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

$$\left(\mathbb{I}_{(t,x)}^v(y) \right)_{i,j} = \left(\nabla_\mu \left(v^i[\cdot](t,x) \right) (\mu(t))(y) \right)_j \quad (1.44)$$

for any $i, j \in \{1, \dots, d\}$.

Proof. By Proposition 1.6 and as a consequence of our hypotheses on $v[\cdot](\cdot, \cdot)$, we know that the map $e \in [-\bar{\epsilon}, \bar{\epsilon}]^N \mapsto \Phi_{(s,t)}^v[\mathcal{G}(e, \cdot) \# \mu](x)$ 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.

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

Now, it has been proven in Theorem 1.9 that such one-dimensional variations could be characterized as the unique solution of the linearized Cauchy problems (1.43). \square

2 The Pontryagin Maximum Principle in the Wasserstein Space

In this chapter, we present a Pontryagin Maximum Principle for optimal mean-field optimal control problems written in the general form

$$(\mathcal{P}) \quad \begin{cases} \min_{u \in \mathcal{U}} \left[\int_0^T L(\mu(t), u(t)) dt + \varphi(\mu(T)) \right], \\ \text{s.t.} \quad \begin{cases} \partial_t \mu(t) + \nabla \cdot ((v[\mu(t)](t, \cdot) + u(t, \cdot))\mu(t)) = 0, \\ \mu(0) = \mu^0 \in \mathcal{P}_c(\mathbb{R}^d). \end{cases} \end{cases} \quad (2.1)$$

Here, the functions $(\mu, \omega) \mapsto L(\mu, \omega)$ and $\mu \mapsto \varphi(\mu)$ describe running and final costs. The velocity field $(t, x, \mu) \mapsto v[\mu](t, x)$ is a general non-local drift. The control $(t, x) \mapsto u(t, x)$ is vector-field which depends on both time and space, as customary in distributed control of partial differential equations (see e.g. [131]).

As discussed in the Introduction, the scheme of proof that we follow is close to that of the finite-dimensional setting (see e.g. [8, 37]). Yet, each step and idea requires a careful adaptation to the specific setting of Wasserstein spaces described in Chapter 1. As a result, this new PMP is formulated in the language of the subdifferential calculus in Wasserstein spaces described in Section 1.2. In this context, the state-costate variables are replaced by a measure on the product of the tangent and cotangent bundle. The dynamics is given by an Hamiltonian system in the space of measures, similar to what studied in [15], where the corresponding Hamiltonian is given by a maximization in an adapted space of controls functions satisfying Lipschitz constraints. This general result is stated in Theorem 2.1 below.

Theorem 2.1 (Pontryagin Maximum Principle for (\mathcal{P})). *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 for (\mathcal{P}) and assume that the following hypotheses **(H)** hold.*

(H)

(H1) *The set of admissible controls is $\mathcal{U} = L^1([0, T], U)$ where $U \subset C^1(\mathbb{R}^d, \mathbb{R}^d)$ is a non-empty and closed subset of $\{v \in C^1(\mathbb{R}^d, \mathbb{R}^d) \text{ s.t. } \|v(\cdot)\|_{C^1(\mathbb{R}^d)} \leq L_U\}$ for a given constant $L_U > 0$.*

(H2) *The running cost $L : (\mu, \omega) \in \mathcal{P}_c(\mathbb{R}^d) \times U \mapsto L(\mu, \omega) \in \mathbb{R}$ is Lipschitz in (μ, ω) with respect to the product metric $W_2 \times C^0$ over $\mathcal{P}(K) \times U$ for any*

compact set $K \subset \mathbb{R}^d$. The functional $\mu \in \mathcal{P}(K) \mapsto L(\mu, \omega)$ is proper, regular in the sense of Definition 1.8, bounded for any $\omega \in U$ and $K \subset \mathbb{R}^d$ compact.

(H3) The terminal cost $\varphi : \mu \in \mathcal{P}_c(\mathbb{R}^d) \mapsto \varphi(\mu) \in \mathbb{R}$ is proper, regular in the sense of Definition 1.8, Lipschitz with respect to the W_2 -metric, bounded from below over $\mathcal{P}(K)$ for any compact set $K \subset \mathbb{R}^d$.

(H4) The non-local velocity field $v : \mu \in \mathcal{P}_c(\mathbb{R}^d) \mapsto v[\mu](\cdot, \cdot) \in L^1([0, T], C^1(\mathbb{R}^d, \mathbb{R}^d)) \cap L^\infty(\mathbb{R}^d, \mathbb{R}^d)$ satisfies the hypotheses of Theorem 1.7, i.e.

$$|v[\mu](t, x)| \leq M(1 + |x|), \quad |v[\mu](t, x) - v[\mu](t, y)| \leq L_1|x - y|,$$

and $\|v[\mu](t, \cdot) - v[\nu](t, \cdot)\|_{C^0(\mathbb{R}^d)} \leq L_2W_1(\mu, \nu)$

for \mathcal{L}^1 -almost every $t \in [0, T]$ and all $(x, y) \in \mathbb{R}^{2d}$ where M, L_1 and L_2 are positive constants. For any compact set $K \subset \mathbb{R}^d$ and any $i \in \{1, \dots, d\}$, the components $\mu \in \mathcal{P}(K) \mapsto v^i[\mu](t, x)$ are regular in the sense of Definition 1.8. The differential in space $\mu \in \mathcal{P}_c(\mathbb{R}^d) \mapsto D_x v[\mu](t, x)$ is narrowly continuous for \mathcal{L}^1 -almost every $t \in [0, T]$ and all $x \in \mathbb{R}^d$.

(H5) The barycenter $x \mapsto \bar{\gamma}_\varphi^\circ(x)$ of the minimal selection $\partial^\circ \varphi(\mu)$ in the extended subdifferential of the terminal cost $\varphi(\cdot)$ at some measure $\mu \in \mathcal{P}_c(\mathbb{R}^d)$ is continuous.

The barycenter $x \mapsto \bar{\gamma}_L^\circ(x)$ of the minimal selection $\partial_\mu^\circ L(\mu, \omega)$ in the extended subdifferential of the running cost $L(\cdot, \omega)$ at some $\mu \in \mathcal{P}_c(\mathbb{R}^d)$ is continuous. The barycenters $(x, y) \mapsto \bar{\gamma}_{(t,x)}^{i,\circ}(y)$ of the minimal selections $\partial_\mu^\circ v^i[\mu](t, x)$ in the extended subdifferentials of the components v^i define continuous mappings for \mathcal{L}^1 -almost every $t \in [0, T]$.

(H6) The maps $\mu \in \mathcal{P}_c(\mathbb{R}^d) \mapsto \varphi(\mu)$, $\mu \in \mathcal{P}_c(\mathbb{R}^d) \mapsto L(\mu, \omega)$ and $\mu \in \mathcal{P}_c(\mathbb{R}^d) \mapsto v[\mu](t, x)$ are differentiable along measure curves generated by Lipschitz-in-time, continuous and bounded perturbations of the identity for $\mathcal{L}^1 \times \mu$ -almost every $(t, x) \in [0, T] \times \mathbb{R}^d$ and any $\omega \in U$, i.e.

$$\frac{d^+}{d\epsilon} [\phi(\mathcal{G}(\epsilon, \cdot) \# \mu)] = \frac{d^-}{d\epsilon} [\phi(\mathcal{G}(\epsilon, \cdot) \# \mu)]$$

$$\frac{d^+}{d\epsilon} [L(\mathcal{G}(\epsilon, \cdot) \# \mu, \omega)] = \frac{d^-}{d\epsilon} [L(\mathcal{G}(\epsilon, \cdot) \# \mu, \omega)]$$

and

$$\frac{d^+}{d\epsilon} [v[\mathcal{G}(\epsilon, \cdot) \# \mu](t, x)] = \frac{d^-}{d\epsilon} [v[\mathcal{G}(\epsilon, \cdot) \# \mu](t, x)],$$

whenever $(\mathcal{G}(\epsilon, \cdot))_{(-\bar{\epsilon}, \bar{\epsilon})}$ is a Lipschitz family of continuous and bounded maps, differentiable at $\epsilon = 0$ and such that $\mathcal{G}(0, \cdot) = \text{Id}$.

Then, there exist a constant $R > 0$ depending on μ^0 , T , \mathcal{U} , $v[\cdot](\cdot, \cdot)$, $\varphi(\cdot)$, $L(\cdot, \cdot)$ and a curve $\nu^*(\cdot) \in \text{Lip}([0, T], \mathcal{P}(\overline{B_{2d}(0, R)}))$ Lipschitz with respect to the W_1 -metric satisfying the following conditions.

(i) It solves the forward-backward system of continuity equations

$$\begin{cases} \partial_t \nu^*(t) + \nabla \cdot \left(\mathbb{J}_{2d} \tilde{\nabla}_\nu \mathbb{H}_c(t, \nu^*(t), u^*(t)) \nu^*(t) \right) = 0 \text{ in } [0, T] \times \mathbb{R}^{2d}, \\ \pi_{\#}^1 \nu^*(0) = \mu^0, \\ \pi_{\#}^2 \nu^*(T) = (-\bar{\gamma}_\varphi^\circ)_{\#} \mu^*(T), \end{cases} \quad (2.2)$$

where $\bar{\gamma}_\varphi^\circ(\cdot)$ is the barycenter of the minimal selection $\partial^\circ \varphi(\mu^*(T))$ and \mathbb{J}_{2d} is the symplectic matrix in \mathbb{R}^{2d} .

The compactified Hamiltonian of the system $\mathbb{H}_c(\cdot, \cdot, \cdot)$ is defined by

$$\mathbb{H}_c(t, \nu, \omega) = \begin{cases} \mathbb{H}(t, \nu, \omega) & \text{if } \nu \in \mathcal{P}(\overline{B_{2d}(0, R)}), \\ +\infty & \text{otherwise,} \end{cases} \quad (2.3)$$

for any $(t, \nu, \omega) \in [0, T] \times \mathcal{P}_c(\mathbb{R}^{2d}) \times U$, where

$$\mathbb{H}(t, \nu, \omega) = \int_{\mathbb{R}^{2d}} \langle r, v[\pi_{\#}^1 \nu](t, x) + \omega(x) \rangle d\nu(x, r) - L(\pi_{\#}^1 \nu, \omega) \quad (2.4)$$

is the infinite-dimensional Hamiltonian of the system.

The vector field $\tilde{\nabla}_\nu \mathbb{H}_c(t, \nu^*(t), u^*(t))(\cdot, \cdot)$ is defined by

$$\begin{aligned} \tilde{\nabla}_\nu \mathbb{H}_c(t, \nu^*(t), u^*(t)) : (x, r) \in \text{supp}(\nu^*(t)) &\mapsto \\ \left(\begin{array}{c} D_x u^*(t, x)^\top r + D_x v[\pi_{\#}^1 \nu^*(t)](t, x)^\top r + \Gamma_v^\circ[\nu^*(t)](t, x) - \bar{\gamma}_L^\circ(t, x) \\ v[\pi_{\#}^1 \nu^*(t)](t, x) + u^*(t, x) \end{array} \right) \end{aligned}$$

where $t \in [0, T] \mapsto \bar{\gamma}_L^\circ(t, \cdot)$ is a measurable selection of the barycenters of the minimal selections $\partial_\mu^\circ L(\mu^*(t), u^*(t))$.

The map $\Gamma_v^\circ[\nu](\cdot, \cdot)$ is defined for any $\nu \in \mathcal{P}(\overline{B_{2d}(0, R)})$ by

$$\Gamma_v^\circ[\nu] : (t, x) \in [0, T] \times \pi^1(\overline{B(0, R)}) \mapsto \int_{\mathbb{R}^{2d}} \left(\Gamma_{(t, y)}^\circ(x) \right)^\top p d\nu(y, p)$$

where for $\mathcal{L}^1 \times \pi_{\#}^1 \nu$ -almost every $(t, y) \in [0, T] \times \pi^1(\overline{B_{2d}(0, R)})$, we denote by $\Gamma_{(t, y)}^\circ : x \in \text{supp}(\pi_{\#}^1 \nu(t)) \mapsto (\bar{\gamma}_{(t, y)}^{i, \circ}(x))_{1 \leq i \leq d}$ the matrix-valued map made of the barycenters of the minimal selections $\partial_\mu^\circ v^i[\pi_{\#}^1 \nu](t, y)$ in the extended subdifferentials of the components (v^i) of the non-local velocity field.

(ii) It satisfies the Pontryagin maximization condition

$$\mathbb{H}_c(t, \nu^*(t), u^*(t)) = \max_{\omega \in U} [\mathbb{H}_c(t, \nu^*(t), \omega)], \quad (2.5)$$

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

The general hypotheses **(H)** are rather cumbersome and can sometimes be hard to

verify. Nevertheless, they are satisfied by a good score of functionals of great interest in various application fields. We present some relevant examples in Section 2.3.

Remark 2.1 (On the smoothness assumption **(H1)**). *The reason why we chose to impose the strong $C^{1,1}$ -smoothness assumption on the set of admissible controls is twofold.*

First, the main scope of this chapter is to provide first-order optimality conditions for infinite-dimensional problems arising as mean-field limits of finite-dimensional systems. Even though very general existence results à la DiPerna-Lions-Ambrosio [11, 68] are available for Cauchy problems of the form (1.32), they only deal with macroscopic quantities which are related to the underlying microscopic ones only for almost every curve in a suitable space of curves. The desired exact micro-macro correspondence which we aim at preserving can only hold in the presence of Cauchy-Lipschitz smoothness assumptions on the driving vector fields, see [11, 78].

Second, the geometric proof of the maximum principle consisting in performing local-in-time perturbations of optimal trajectories can only be carried out under C^1 -regularity assumptions, due to the non-linearity of the problem studied here. Even though the derivation of a maximum principle under a merely Lipschitz-regularity assumption on the optimal control in the spirit of the non-smooth maximum principle (see e.g. [59]) might be available in this context, it would require a completely different approach and much more technical arguments.

Let it be noted that these assumptions are verified in the classical setting of systems that are linear or affine with respect to the controls, i.e. where the controlled term is of the form $u : (t, x) \mapsto \sum_{k=1}^m u_k(t) F_k(x)$ where the $(F_k(\cdot))_{1 \leq k \leq m}$ are C^1 vector fields. They can also be leveraged to incorporate non-linear controlled dynamics of the form $(t, x) \mapsto v(t, x, u(t))$ where $u(t) \in \mathbb{R}^m$ for \mathcal{L}^1 -almost every $t \in [0, T]$ ([39, 118, 117]).

Remark 2.2 (Almost-Hamiltonian flow). *Observe that in our formulation of the PMP, the vector field $\tilde{\nabla}_{\nu} \mathbb{H}_c(t, \nu^*(t), u^*(t))$ is not the Wasserstein gradient of the compactified Hamiltonian $\mathbb{H}_c(\nu^*(t), u^*(t))$, since in general the barycenter of a minimal selection is not in the classical subdifferential. However, in any context where the minimal selections of the cost and dynamics functionals are induced by maps – which will automatically be their barycenters –, it can be shown by standard methods that $\tilde{\nabla}_{\nu} \mathbb{H}_c(t, \nu^*(t), u^*(t))$ is in fact the Wasserstein gradient of the compactified Hamiltonian at $(t, \nu^*(t), u^*(t))$ for \mathcal{L}^1 -almost every $t \in [0, T]$.*

Let it be noted that this “almost-Hamiltonian” flow still has a geometric meaning since $\tilde{\nabla}_{\nu} \mathbb{H}_c(t, \nu^(t), u^*(t)) \in \text{Tan}_{\nu^*(t)} \mathcal{P}_2(\mathbb{R}^{2d})$. Indeed, as it was already mentioned in Remark 1.3, the barycenter of a minimal selection always belongs to the analytical tangent space by construction.*

The structure of this chapter in the following. We first describe in Section 2.1 our scheme of proof – and in particular introduce the concept of *needle like variation* – on a simplified problem (\mathcal{P}_1) where there are neither interaction field $v[\cdot](\cdot, \cdot)$ nor running cost $L(\cdot, \cdot)$. We then proceed to prove the PMP for the more general problem (\mathcal{P}) in Section 2.2. In Section 2.3 we discuss more in details the set of hypotheses **(H)** of Theorem 2.1 and list some relevant examples of classical functionals satisfying them.

2.1 A Simpler Pontryagin Maximum Principle with no interaction field and no running cost

We start by proving the Pontryagin Maximum Principle for a simplified version of problem (\mathcal{P}) presented in the Introduction. We consider the following optimal control problem in the space of probability measures

$$(\mathcal{P}_1) \quad \begin{cases} \min_{u \in \mathcal{U}} [\varphi(\mu(T))], \\ \text{s.t.} \quad \begin{cases} \partial_t \mu(t) + \nabla \cdot (u(t, \cdot) \mu(t)) = 0, \\ \mu(0) = \mu^0 \in \mathcal{P}_c(\mathbb{R}^d), \end{cases} \end{cases} \quad (2.6)$$

and show that the Pontryagin-type optimality conditions provided in the following theorem hold.

Theorem 2.2 (Pontryagin Maximum Principle for (\mathcal{P}_1)).

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 for (\mathcal{P}_1) and assume that the relevant hypotheses of **(H)** hold. Then, there exists a constant $R > 0$ and a curve $\nu^*(\cdot) \in \text{Lip}([0, T], \mathcal{P}(B_{2d}(0, R)))$ satisfying the following statements :

(i) It solves the forward-backward system of continuity equations

$$\begin{cases} \partial_t \nu^*(t) + \nabla \cdot (\mathbb{J}_{2d} \nabla_\nu \mathbb{H}_c(\nu^*(t), u^*(t)) \nu^*(t)) = 0, & \text{in } [0, T] \times \mathbb{R}^{2d} \\ \pi_{\#}^1 \nu^*(t) = \mu^*(t), \\ \pi_{\#}^2 \nu^*(T) = (-\bar{\gamma}_\varphi^\circ)_{\#} \mu^*(T), \end{cases} \quad (2.7)$$

where \mathbb{J}_{2d} is the symplectic matrix of \mathbb{R}^{2d} .

The compactified Hamiltonian $\mathbb{H}_c(\cdot, \cdot)$ of the system is defined by

$$\mathbb{H}_c : (\nu, \omega) \in \mathcal{P}_c(\mathbb{R}^{2d}) \times U \mapsto \begin{cases} \mathbb{H}(\nu, \omega) & \text{if } \nu \in \mathcal{P}(\overline{B_{2d}(0, R)}) \\ +\infty & \text{otherwise} \end{cases}, \quad (2.8)$$

where

$$\mathbb{H} : (\nu, \omega) \in \mathcal{P}_c(\mathbb{R}^{2d}) \times U \mapsto \int_{\mathbb{R}^{2d}} \langle r, \omega(x) \rangle d\nu(x, r) \quad (2.9)$$

is the infinite-dimensional Hamiltonian of the system.

The vector field $(x, r) \mapsto \nabla_\nu \mathbb{H}_c(\nu^*(t), u^*(t))(x, r) = (\mathbb{D}_x u^*(t, x)^\top r, u^*(t, x))$ is the Wasserstein gradient of the compactified Hamiltonian for \mathcal{L}^1 -almost every $t \in [0, T]$, i.e. $\partial_\nu^\circ \mathbb{H}_c(\nu^*(t), u^*(t)) = (\text{Id} \times \nabla_\nu \mathbb{H}_c(\nu^*(t), u^*(t)))_{\#} \nu^*(t)$.

(ii) It satisfies the Pontryagin maximization condition

$$\mathbb{H}_c(\nu^*(t), u^*(t)) = \max_{\omega \in U} [\mathbb{H}_c(\nu^*(t), \omega)] \quad (2.10)$$

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

We split the proof of this result into several steps. In Step 1, we introduce the concept of *needle-like variation* of an optimal control and compute explicitly the corresponding family of perturbed measures. In Step 2 we study the first order perturbation of the

final cost induced by the needle-like variation. We introduce in Step 3 a suitable costate propagating this information backward to the base point of the needle-variation. In Step 4, we show that the curve introduced in Step 3 satisfies the conditions (i) and (ii) of the PMP.

Step 1 : Needle-like variations

We start by considering an optimal pair control-trajectory $(u^*(\cdot), \mu^*(\cdot)) \in \mathcal{U} \times \text{Lip}([0, T], \mathcal{P}_c(\mathbb{R}^d))$ along with the constant $R_T > 0$ such that $\text{supp}(\mu(t)) \subset B(0, R_T)$ for all times $t \in [0, T]$. Fix a control $\omega \in U$, a Lebesgue point $\tau \in [0, T]$ of $t \mapsto u^*(t) \in C^0(\mathbb{R}^d, \mathbb{R}^d)$ in the Bochner sense (see, e.g. [69]) and a parameter $\epsilon \in [0, \bar{\epsilon}]$ with $\bar{\epsilon} > 0$ small. We define the *needle-like variation* of parameters (ω, τ, ϵ) of u^* as follows

$$\tilde{u}_\epsilon^{\omega, \tau} \equiv \tilde{u}_\epsilon : t \mapsto \begin{cases} \omega & \text{if } t \in [\tau - \epsilon, \tau], \\ u^*(t) & \text{otherwise.} \end{cases} \quad (2.11)$$

We denote by $t \mapsto \tilde{\mu}_t(\epsilon)$ the corresponding solution of the continuity equation starting from μ^0 at time $t = 0$. Notice that $\tilde{u}_\epsilon(\cdot) \in L^1([0, T], C^1(\mathbb{R}^d, \mathbb{R}^d) \cap L^\infty(\mathbb{R}^d, \mathbb{R}^d))$, thus the corresponding continuity equation is still well-posed.

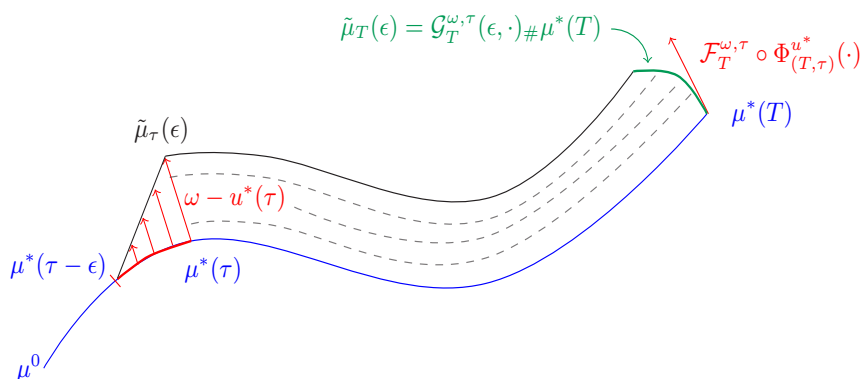


Figure 2.1: Illustration of the effect of a needle-like variation on a measure curve.

The explicit link between the perturbed measure $\tilde{\mu}_T(\epsilon)$ and the optimal measure $\mu^*(T)$ at time T is given in the following lemma.

Lemma 2.1. *There exists a family of functions $\mathcal{G}_T^{\omega, \tau}(\cdot, \cdot) \in \text{Lip}((-\bar{\epsilon}, \bar{\epsilon}), C^0(\mathbb{R}^d, \mathbb{R}^d))$ such that $\mathcal{G}_T^{\omega, \tau}(\epsilon, \cdot)$ is a C^1 -diffeomorphisms over $\overline{B(0, R_T)}$ for all $\epsilon \geq 0$ and it holds*

$$\tilde{\mu}_T(\epsilon) = \mathcal{G}_T^{\omega, \tau}(\epsilon, \cdot) \# \mu^*(T). \quad (2.12)$$

Moreover, there exists a constant $R_T^\Phi > 0$ depending on R_T, L_U and $v[\cdot](\cdot, \cdot)$ such that for all $\epsilon \in (-\bar{\epsilon}, \bar{\epsilon})$ one has $\text{supp}(\mathcal{G}_T^{\omega, \tau}(\epsilon, \cdot) \# \mu^*(T)) \subset B(0, R_T^\Phi)$.

This family of maps satisfies the following Taylor expansion with respect to the $L^2(\mathbb{R}^d, \mathbb{R}^d; \mu^*(T))$ -norm

$$\mathcal{G}_T^{\omega, \tau}(\epsilon, \cdot) = \text{Id} + \epsilon \mathcal{F}_T^{\omega, \tau} \circ \Phi_{(T, \tau)}^{u^*}(\cdot) + o(\epsilon),$$

where

$$\mathcal{F}_T^{\omega, \tau} : x \in \text{supp}(\mu^*(\tau)) \mapsto D_x \Phi_{(\tau, T)}^{u^*}(x) \cdot [\omega(x) - u^*(\tau, x)] \quad (2.13)$$

is a C^0 mapping.

Proof. By definition of $\tilde{u}_\epsilon(\cdot, \cdot)$ in (2.11), it holds that

$$\tilde{\mu}_T(\epsilon) = \Phi_{(\tau, T)}^{u^*} \circ \Phi_{(\tau-\epsilon, \tau)}^\omega \circ \Phi_{(\tau, \tau-\epsilon)}^{u^*} \circ \Phi_{(T, \tau)}^{u^*}(\cdot) \# \mu^*(T)$$

for all $\epsilon \in [0, \bar{\epsilon}]$. Thus, by choosing $\mathcal{G}_T^{\omega, \tau}(\epsilon, \cdot) = \Phi_{(\tau, T)}^{u^*} \circ \Phi_{(\tau-\epsilon, \tau)}^\omega \circ \Phi_{\tau-\epsilon}^{u^*} \circ \Phi_{(T, \tau)}^{u^*}(\cdot)$, formula (2.12) holds true for $\epsilon \in [0, \bar{\epsilon}]$. Moreover, since the definition of $\epsilon \mapsto \mathcal{G}_T^{\omega, \tau}(\epsilon, \cdot)$ only involves functions that are continuous and uniformly bounded over $\overline{B(0, R_T)}$, the perturbed measures $\tilde{\mu}_T(\cdot)$ are compactly supported in some bigger ball $\overline{B(0, R'_T)}$ for all $\epsilon \in [0, \bar{\epsilon}]$ as well.

Recalling the definition of the flow $x \mapsto \Phi_{(0, t)}^\nu(x)$, one has by Lebesgue's Differentiation Theorem (see e.g. [14, Corollary 2.23]) that

$$\begin{cases} \Phi_{(\tau, \tau-\epsilon)}^{u^*}(x) = x - \int_{\tau-\epsilon}^\tau u^*(t, \Phi_{(\tau, t)}^{u^*}(x)) dt = x - \epsilon u^*(\tau, x) + o(\epsilon) \\ \Phi_{(\tau-\epsilon, \tau)}^\omega(x) = x + \int_{\tau-\epsilon}^\tau \omega(\Phi_{(\tau-\epsilon, t)}^\omega(x)) dt = x + \epsilon \omega(x) + o(\epsilon) \end{cases}$$

since $t \mapsto \Phi_{(\tau-\epsilon, t)}^{u^*}(x)$ and $t \mapsto \Phi_{(t, \tau)}^{u^*}(x)$ are C^0 for any $x \in \overline{B(0, R_T)}$ and $\tau \in [0, T]$ Lebesgue point of $t \mapsto u^*(t) \in C^1(\mathbb{R}^d, \mathbb{R}^d)$. Chaining these two expansions and recalling that $\omega(\cdot)$ and $\Phi_{(\tau, T)}^{u^*}(\cdot)$ are C^1 -smooth yields

$$\Phi_{(\tau, T)}^{u^*} \circ \Phi_{(\tau-\epsilon, \tau)}^\omega \circ \Phi_{(\tau, \tau-\epsilon)}^{u^*}(x) = \Phi_{(\tau, T)}^{u^*}(x) + \epsilon D_x \Phi_{(\tau, T)}^{u^*}(x) \cdot [\omega(x) - u^*(\tau, x)] + o(\epsilon).$$

Thus,

$$\mathcal{G}_T^{\omega, \tau}(\epsilon, x) = x + \epsilon \mathcal{F}_T^{\omega, \tau} \circ \Phi_{(T, \tau)}^{u^*}(x) + o(\epsilon),$$

for any $x \in \text{supp}(\mu^*(T))$ where we choose

$$\mathcal{F}_T^{\omega, \tau} : x \mapsto D_x \Phi_{(\tau, T)}^{u^*}(x) \cdot [\omega(x) - u^*(\tau, x)].$$

We can now extend $\mathcal{G}_T^{\omega, \tau}(\cdot, \cdot)$ from $[0, \bar{\epsilon}]$ to $(-\bar{\epsilon}, \bar{\epsilon})$ in such a way that the left and right derivatives at $\epsilon = 0$ coincide, by defining e.g.

$$\mathcal{G}_T^{\omega, \tau}(\epsilon, \cdot) \equiv \text{Id} + \epsilon \mathcal{F}_T^{\omega, \tau} \circ \Phi_{(T, \tau)}^{u^*}(\cdot)$$

whenever $\epsilon \in (-\bar{\epsilon}, 0]$. Notice that since $\mathcal{F}_T^{\omega, \tau} \circ \Phi_{(T, \tau)}^{u^*}(\cdot)$ is C^0 , both $x \in \overline{B(0, R_T)} \mapsto \mathcal{G}_T^{\omega, \tau}(\epsilon, x)$ and $x \in \overline{B(0, R_T)} \mapsto \frac{d}{d\epsilon} [\mathcal{G}_T^{\omega, \tau}(\epsilon, x)]_{\epsilon=0}$ define C^0 mappings for all $\epsilon \in (-\bar{\epsilon}, \bar{\epsilon})$. Moreover, the continuity and uniform boundedness of $D_x \Phi_{(\tau, T)}^{u^*}(\cdot)$ over $\overline{B(0, R_T)}$ along with hypothesis **(H1)** imply that $\mathcal{F}_T^{\omega, \tau}(\cdot)$ is bounded. Hence, there exists a constant $R_T^\Phi > 0$ such that $\text{supp}(\mathcal{G}_T^{\omega, \tau}(\epsilon, \cdot) \# \mu^*(T)) \subset \overline{B(0, R_T^\Phi)}$ for all $\epsilon \in (-\bar{\epsilon}, \bar{\epsilon})$. Moreover, the fact that $\mathcal{G}(\epsilon, \cdot)$ and $\mathcal{F}_T^{\omega, \tau}(\cdot)$ are continuous and bounded yields that they are uniformly integrable with respect to the compactly supported measure $\mu^*(T)$. An application of Proposition 1.2-(iii) allows to conclude that this expansion holds in $L^2(\mathbb{R}^d, \mathbb{R}^d; \mu^*(T))$,

which achieves the proof. \square

We end this first step by a lemma which is a direct consequence of Proposition 1.10.

Lemma 2.2. *For any $x \in \text{supp}(\mu^*(\tau))$, the trajectory $t \in [\tau, T] \mapsto \mathcal{F}_t^{\omega, \tau}(x)$ is the unique solution of the Cauchy problem*

$$\partial_t \mathcal{F}_t^{\omega, \tau}(x) = D_x u^*(t, \Phi_{(\tau, t)}^{u^*}(x)) \mathcal{F}_t^{\omega, \tau}(x), \quad \mathcal{F}_\tau^{\omega, \tau}(x) = \omega(x) - u^*(\tau, x). \quad (2.14)$$

Proof. It is sufficient to apply Proposition 1.10 with $v[\cdot](\cdot, \cdot) \equiv u^*(\cdot, \cdot)$. \square

Step 2 : First-order optimality condition

Thanks to the optimality of $u^*(\cdot)$, for each $\epsilon \in (0, \bar{\epsilon})$ it holds

$$\frac{\varphi(\tilde{\mu}_T(\epsilon)) - \varphi(\mu^*(T))}{\epsilon} \geq 0, \quad (2.15)$$

where $\epsilon \in (0, \bar{\epsilon}) \mapsto \tilde{\mu}_T(\epsilon) = \mathcal{G}_T^{\omega, \tau}(\epsilon, \cdot) \# \mu^*(T)$.

Recalling that the measures $\epsilon \mapsto \tilde{\mu}_T(\epsilon)$ are uniformly compactly supported, that $\varphi(\cdot)$ satisfies hypotheses **(H3)** and that the map $\epsilon \in (-\bar{\epsilon}, \bar{\epsilon}) \mapsto \varphi(\tilde{\mu}_T(\epsilon))$ is differentiable at $\epsilon = 0$ by hypothesis **(H6)**, we can apply the chain rule given in Proposition 1.6 to the endpoint cost :

$$0 \leq \lim_{\epsilon \downarrow 0} \left[\frac{\varphi(\tilde{\mu}_T(\epsilon)) - \varphi(\mu^*(T))}{\epsilon} \right] = \int_{\mathbb{R}^d} \langle \bar{\gamma}_\varphi^\circ(x), \mathcal{F}_T^{\omega, \tau} \circ \Phi_{(T, \tau)}^{u^*}(x) \rangle d\mu^*(T)(x), \quad (2.16)$$

where $\bar{\gamma}_\varphi^\circ \in L^2(\mathbb{R}^d, \mathbb{R}^d; \mu^*(T))$ is the barycenter of the minimal selection $\partial^\circ \varphi(\mu^*(T))$ in the extended subdifferential of $\varphi(\cdot)$ at $\mu^*(T)$.

We recover a formula similar to the classical finite-dimensional case. The next step is to introduce a suitable *costate* along with its backward dynamics that will *propagate* this first-order information to the base-point τ of the needle-like variation while generating a Hamiltonian-like dynamical structure.

Step 3 : Backward dynamics and Pontryagin maximization condition

Equation (2.16) provides us with a first-order optimality condition which involves all the needle parameters $(\omega, \tau) \in U \times [0, T]$. We will show that it implies, along with the choice of a suitable costate, the maximization condition (2.10). To this aim, we build a curve $\nu^* \in \text{Lip}([0, T], \mathcal{P}_c(\mathbb{R}^{2d}))$ solution of the forward-backward system of continuity equations

$$\begin{cases} \partial_t \nu^*(t) + \nabla \cdot (\mathcal{V}^*(t, \cdot, \cdot)) \nu^*(t) = 0 & \text{in } [0, T] \times \mathbb{R}^{2d}, \\ \pi_{\#}^1 \nu^*(t) = \mu^*(t) & \text{for all } t \in [0, T], \\ \nu^*(T) = (\text{Id} \times (-\bar{\gamma}_\varphi^\circ)) \# \mu^*(T), \end{cases} \quad (2.17)$$

associated with the vector field

$$\mathcal{V}^* : (t, x, r) \in [0, T] \times \mathbb{R}^{2d} \mapsto (u^*(t, x), -D_x u^*(t, x)^\top r). \quad (2.18)$$

Notice that, contrarily to system (2.7), we impose the more restrictive product structure on the terminal datum.

This system is peculiar in the sense that the driving vector field $\mathcal{V}^*(\cdot, \cdot, \cdot)$ does not satisfy verbatim the hypotheses **(H')** of Theorem 1.7. However, it exhibits a *cascade* structure, in the sense that one can first determine uniquely $\mu^*(\cdot)$ and then build $\nu^*(\cdot)$ by disintegration. This fact is underlined by the condition $\pi_{\#}^1 \nu^*(t) = \mu^*(t)$ for all times $t \in [0, T]$. We make this statement precise in the next lemma.

Lemma 2.3 (Definition and well-posedness of solutions of (2.17)).

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 $(\Psi_{(T,t)}^x(\cdot))_{t \leq T}$ associated to the Cauchy problems

$$\partial_t w_x(t, r) = -D_x u^*(t, \Phi_{(T,t)}^{u^*}(x)) w_x(t, r), \quad w_x(T, r) = r, \quad (2.19)$$

and define the associated curves of measures $\sigma_x^* : t \mapsto \Psi_{(T,t)}^x(\cdot) \# \delta_{(-\tilde{\gamma}_\varphi^\circ(x))}$.

Then, $\nu^* : t \mapsto (\Phi_{(T,t)}^{u^*}, \text{Id})_{\#} \nu_T^*(t)$ is the unique solution of (2.17) with ¹

$$\nu_T^*(t) = \int_{\mathbb{R}^d} \sigma_x^*(t) d\mu^*(T)(x) \in \mathcal{P}_c(\mathbb{R}^{2d})$$

for all times $t \in [0, T]$. Moreover, there exists two constants $R'_T, L'_T > 0$ such that

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

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

Proof. By hypothesis **(H1)**, the elements of \mathcal{U} are uniformly sublinear and Lipschitz in space for \mathcal{L}^1 -almost every times $t \in [0, T]$. We recall that by Theorem 1.7, this implies the existence of a constant $R_T > 0$ depending on $\text{supp}(\mu^0), T$ and L_U such that $\text{supp}(\mu^*(\cdot)) \subset \overline{B(0, R_T)}$.

For $\mu^*(T)$ -almost every $x \in \mathbb{R}^d$, the Cauchy problem (2.19) has a unique solution and the corresponding curves $t \mapsto \sigma_x^*(t)$ are uniquely determined. Moreover, the uniform Lipschitzianity of the elements of \mathcal{U} implies that these curves are uniformly compactly supported and Lipschitz in the W_1 -metric uniformly with respect to $x \in \text{supp}(\mu^*(T))$ with constants \tilde{R}_T, \tilde{L}_T depending on L_U, T and $\varphi(\cdot)$.

We now define the curve $\nu^*(\cdot)$ as in the statement of Lemma 2.3 above and show that it is a uniformly compactly supported and Lipschitz solution of the forward-backward system (2.17). The fact that there exists $R_T > 0$ depending on R_T and \tilde{R}_T such that $\nu^*(\cdot)$ is uniformly compactly supported in $\overline{B_{2d}(0, R_T)}$ is a direct consequence of its definition. The Lipschitzianity in the W_1 -metric comes from the following computations. For any

1. Namely, $\nu_T^*(t)$ is defined as the $\mu^*(T)$ -almost uniquely determined measure which has $\mu^*(T)$ as its first marginal and which disintegration is given by $\{\sigma_x^*(t)\}_x$ (see Definition 1.5).

$\xi \in \text{Lip}(\mathbb{R}^{2d}, \mathbb{R})$ with $\text{Lip}(\xi, \mathbb{R}^{2d}) \leq 1$, it holds

$$\begin{aligned}
\int_{\mathbb{R}^{2d}} \xi(x, r) d(\nu^*(t) - \nu^*(s))(x, r) &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \xi(\Phi_{(T,t)}^{u^*}(x), r) d\sigma_x^*(t)(r) d\mu^*(T)(x) \\
&\quad - \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \xi(\Phi_{(T,s)}^{u^*}(x), r) d\sigma_x^*(s)(r) d\mu^*(T)(x) \\
&\leq \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |\Phi_{(T,t)}^{u^*}(x) - \Phi_{(T,s)}^{u^*}(x)| d\sigma_x^*(t)(r) d\mu^*(T)(x) \\
&\quad + \int_{\mathbb{R}^d} \text{Lip}(\Phi_{(T,s)}^{u^*}, \overline{B(0, R_T)}) W_1(\sigma_x^*(s), \sigma_x^*(t)) d\mu^*(T)(x) \\
&\leq L'_T |t - s|
\end{aligned}$$

where $L'_T > 0$ is a uniform constant depending on the time and space Lipschitz constants of the flows of diffeomorphisms $(\Phi_{(T,t)}^{u^*}(\cdot))_{t \in [0, T]}$ and on L_U . Taking the supremum over all the 1-Lipschitz functions $\xi(\cdot, \cdot)$ and using the Kantorovich-Rubinstein duality (1.4) yields the Lipschitzianity of $\nu^*(\cdot)$ in the W_1 -metric.

Finally, remark that for any $\xi \in C_c^\infty(\mathbb{R}^{2d}, \mathbb{R})$ it holds

$$\begin{aligned}
&\frac{d}{dt} \left[\int_{\mathbb{R}^{2d}} \xi(x, r) d\nu^*(t)(x, r) \right] \\
&= \frac{d}{dt} \left[\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \xi(\Phi_{(T,t)}^{u^*}(x), r) d\sigma_x^*(t)(r) d\mu^*(T)(x) \right] \\
&= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \langle \nabla_x \xi(\Phi_{(T,t)}^{u^*}(x), r), u^*(t, \Phi_{(T,t)}^{u^*}(x)) \rangle d\sigma_x^*(t)(r) d\mu^*(T)(x) \\
&\quad + \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \langle \nabla_r \xi(\Phi_{(T,t)}^{u^*}(x), r), -D_x u^*(t, \Phi_{(T,t)}^{u^*}(x))^\top r \rangle d\sigma_x^*(t)(r) d\mu^*(T)(x) \\
&= \int_{\mathbb{R}^{2d}} \left\langle \nabla_{(x,r)} \xi(x, r), \begin{pmatrix} u^*(t, x) \\ -D_x u^*(t, x)^\top r \end{pmatrix} \right\rangle d\nu^*(t)(x, r)
\end{aligned}$$

which along with the fact that $\nu^*(T) = \nu_T^*(T) = (\text{Id} \times (-\bar{\gamma}_\varphi^\circ))_{\#} \mu^*(T)$ achieves the proof. \square

Remark 2.3 (Wasserstein and classical costates). *In the finite-dimensional proof of the PMP, the optimal costates $p^*(\cdot)$ are defined as the solutions of the backward equations*

$$\dot{p}^*(t) = -\nabla_x \mathcal{H}(t, x^*(t), p^*(t), u^*(t)), \quad p^*(T) = -\nabla \varphi(x^*(T)). \quad (2.20)$$

In our statement of the PMP, one should think of $\pi_{\#}^2 \nu^(\cdot)$ as being concentrated on the characteristic curves of the backward costate dynamics. Indeed in Lemma 2.3, the curves $\sigma_x^*(\cdot)$ are concentrated on the unique characteristic of the linearized backward non-local dynamics (3.21) starting from $(-\bar{\gamma}_\varphi^\circ(x))$ for $\mu^*(T)$ -almost every $x \in \mathbb{R}^d$. The second marginal of the curve $\nu_T^*(\cdot) = \int_{\mathbb{R}^d} \sigma_x^*(t)(r) \mu^*(T)(x)$ can then be seen as a Lagrangian superposition of integral curves of (2.20) depending on the starting point of the curve in $\text{supp}(\mu^*(T))$.*

We now show that the curve of measures $\nu^*(\cdot)$ defined in Lemma 2.3 is such that the

map $\mathcal{K}_{\omega,\tau}(\cdot)$ defined by

$$\mathcal{K}_{\omega,\tau} : t \in [\tau, T] \mapsto \int_{\mathbb{R}^{2d}} \langle r, \mathcal{F}_t^{\omega,\tau} \circ \Phi_{(t,\tau)}^{u^*}(x) \rangle d\nu^*(t)(x, r), \quad (2.21)$$

is constant over $[\tau, T]$. We shall see in Step 4 that this is equivalent to the Pontryagin maximization condition (2.10).

Lemma 2.4. *The map $t \mapsto \mathcal{K}_{\omega,\tau}(t)$ defined in (2.21) is constant over $[\tau, T]$ for any couple of needle parameters (ω, τ) .*

Proof. Notice that by definition of $\nu^*(\cdot)$, the map $\mathcal{K}_{\omega,\tau}(\cdot)$ rewrites

$$\mathcal{K}_{\omega,\tau}(t) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \langle r, \mathcal{F}_t^{\omega,\tau} \circ \Phi_{(T,\tau)}^{u^*}(x) \rangle d\sigma_x^*(t)(r) d\mu^*(T)(x) \quad (2.22)$$

for all $t \in [\tau, T]$. The maps $t \in [\tau, T] \mapsto \mathcal{F}_t^{\omega,\tau} \circ \Phi_{(T,\tau)}^{u^*}(x)$ and $t \in [\tau, T] \mapsto \sigma_x^*(t)$ are Lipschitz, uniformly with respect to $x \in \text{supp}(\mu^*(T))$. The integrand $(x, r) \mapsto \langle r, \mathcal{F}_t^{\omega,\tau} \circ \Phi_{(T,\tau)}^{u^*}(x) \rangle$ is bounded with respect to x and Lipschitz with respect to r , uniformly with respect to $t \in [\tau, T]$. Hence, $t \mapsto \mathcal{K}_{\omega,\tau}(t)$ is Lipschitz as well. It will therefore be constant provided that its derivative - which exists \mathcal{L}^1 -almost everywhere - is equal to zero.

Observe that, using formula (1.30) and the definition of $\mathcal{V}^*(\cdot, \cdot, \cdot)$ in (1.39), it holds

$$\begin{aligned} \frac{d}{dt} \mathcal{K}_{\omega,\tau}(t) &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \langle r, \partial_t \mathcal{F}_t^{\omega,\tau} \circ \Phi_{(T,\tau)}^{u^*}(x) \rangle d\sigma_x^*(t)(r) d\mu^*(T)(x) \\ &+ \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \langle -D_x u^*(t, \Phi_{(T,t)}^{u^*}(x))^\top r, \mathcal{F}_t^{\omega,\tau} \circ \Phi_{(T,\tau)}^{u^*}(x) \rangle d\sigma_x^*(t)(r) d\mu^*(T)(x). \end{aligned} \quad (2.23)$$

Recalling the characterization of $\partial_t \mathcal{F}_t^{\omega,\tau}(\cdot)$ given in (2.14) and plugging it into (2.23), we obtain that $\frac{d}{dt} \mathcal{K}_{\omega,\tau}(t) = 0$ for \mathcal{L}^1 -almost every times $t \in [\tau, T]$, and thus that $\mathcal{K}_{\omega,\tau}(\cdot)$ is constant over $[\tau, T]$. \square

Step 4 : Proof of the Pontryagin Maximum Principle for (\mathcal{P}_1)

We proved in Lemma 2.3 the existence of a constant $R \equiv R'_T > 0$ such that the solution $\nu^*(\cdot)$ to (2.17) satisfies $\text{supp}(\nu^*(\cdot)) \subset \overline{B_{2d}(0, R)}$. We accordingly define the infinite-dimensional Hamiltonian $\mathbb{H} : (\nu, \omega) \in \mathcal{P}_c(\mathbb{R}^{2d}) \times U \mapsto \int_{\mathbb{R}^{2d}} \langle r, \omega(x) \rangle d\nu(x, r)$ of the system and the compactified Hamiltonian $\mathbb{H}_c(\cdot, \cdot)$ by (2.8).

In Lemma 2.4 we showed that, with this choice of forward-backward system (2.17), the map $\mathcal{K}_{\omega,\tau}(\cdot)$ defined in (2.21) is constant over $[\tau, T]$ for any choice of $\omega \in U$ and $\tau \in [0, T]$ Lebesgue point of $u^*(\cdot)$. This implies in particular that $\mathcal{K}_{\omega,\tau}(\tau) = \mathcal{K}_{\omega,\tau}(T)$. Since we proved in (2.16) that it holds

$$0 \leq \int_{\mathbb{R}^d} \langle \bar{\gamma}_\varphi^\circ(x), \mathcal{F}_T^{\omega,\tau} \circ \Phi_{(T,\tau)}^{u^*}(x) \rangle d\mu^*(T)(x) = -\mathcal{K}_{\omega,\tau}(T),$$

it directly follows that

$$\mathcal{K}_{\omega,\tau}(\tau) = \int_{\mathbb{R}^{2d}} \langle r, \omega(x) - u^*(\tau, x) \rangle d\nu^*(\tau)(x, r) \leq 0,$$

for all $\omega \in U$ and $\tau \in [0, T]$ Lebesgue point of $u^*(\cdot)$.

Recalling that \mathcal{L}^1 -almost $\tau \in [0, T]$ is a Lebesgue point in the Bochner sense for an L^1 -function defined over the real line (see e.g. [69, Chapter II - Theorem 9]), we recover the infinite-dimensional maximization condition

$$\mathbb{H}_c(\nu^*(t), u^*(t)) = \max_{\omega \in U} [\mathbb{H}_c(\nu^*(t), \omega)]$$

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

Invoking the C^1 -regularity of the elements of U , it can be seen using Proposition 2.1 that the minimal selection $\partial_\nu^\circ \mathbb{H}_c(\nu^*(t), u^*(t))$ in the extended subdifferential of $\mathbb{H}_c(\cdot, u^*(t))$ exists at $\nu^*(t) \in \mathcal{P}(\overline{B_{2d}(0, R)})$ for \mathcal{L}^1 -almost every $t \in [0, T]$ and that it is induced by the map

$$\nabla_\nu \mathbb{H}_c(\nu^*(t), u^*(t)) : (x, r) \in \text{supp}(\nu^*(t)) \mapsto \begin{pmatrix} D_x u^*(t, x)^\top r \\ u^*(t, x) \end{pmatrix}.$$

Therefore, we recognize the Wasserstein Hamiltonian structure generated by

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

for \mathcal{L}^1 -almost every $t \in [0, T]$, where \mathbb{J}_{2d} is the symplectic matrix in \mathbb{R}^{2d} . This ends our proof of Theorem 2.2.

2.2 Proof of Theorem 2.1

After having exhibited the main mechanisms of our proof for the Pontryagin Maximum Principle for the simplified problem (\mathcal{P}_1) , we are ready to tackle the general case proposed in (\mathcal{P}) . The latter is a generalization of (\mathcal{P}_1) in the sense that we add a general running cost $L(\cdot, \cdot)$ and a general non-local interaction vector field $v[\cdot](\cdot, \cdot)$.

Step 1 : Needle-like variations in the non-local case

As in Section 2.1, let us consider an optimal pair control-trajectory $(u^*(\cdot), \mu^*(\cdot))$, a Lebesgue point $\tau \in [0, T]$ of $t \in [0, T] \mapsto v[\mu^*(t)](t, \cdot) + u^*(t, \cdot)$ and an element $\omega \in U$. We introduce again the needle-like variation $\tilde{u}_\epsilon^{\omega, \tau}(\cdot)$ of $u^*(\cdot)$ with parameters (ω, τ, ϵ) for $\epsilon \in [0, \bar{\epsilon})$, as defined in (2.11).

In keeping with the notations introduced in (1.33) for flows associated to transport PDEs with non-local velocities, the family of perturbed measures $\epsilon \in [0, \bar{\epsilon}) \mapsto \tilde{\mu}_t(\epsilon)$ are defined for all times $t \in [\tau, T]$ by

$$\tilde{\mu}_t(\epsilon) = \Phi_{(\tau, t)}^{v, u^*}[\tilde{\mu}_\tau(\epsilon)] \circ \Phi_{(\tau - \epsilon, \tau)}^{v, \omega}[\mu^*(\tau - \epsilon)] \circ \Phi_{(t, \tau - \epsilon)}^{v, u^*}[\mu^*(t)](\cdot) \# \mu^*(t).$$

One can readily check that under the sublinearity and regularity hypotheses imposed in **(H1)** and **(H4)**, there exists again a constant $\tilde{R}_T > 0$ such that $\text{supp}(\tilde{\mu}_t(\epsilon)) \subset \overline{B(0, \tilde{R}_T)}$ for all $(t, \epsilon) \in [0, T] \times [0, \bar{\epsilon})$.

We now derive in Lemma 2.5 the perturbation stemming from the needle-like variation. We prove therein a result akin to Lemma 2.2 giving a precise ODE-type characterization of this perturbation. To do so, we use the results of Theorem 1.9 concerning the directional

derivatives of the non-local flow combined to the classical result stated in Lemma 2.2 and the definition of needle-like variation.

Lemma 2.5 (Perturbation induced by a needle-like variation in the non-local case). *Let $(u^*(\cdot), \mu^*(\cdot))$ be an optimal pair control-trajectory for problem (\mathcal{P}) and $\tilde{u}_\epsilon(\cdot)$ be the needle-like perturbation of $u^*(\cdot)$ as introduced in (2.11).*

Then, there exists a family of maps $(\mathcal{G}_t^{\omega, \tau}(\cdot, \cdot))_{t \in [\tau, T]} \subset \text{Lip}((-\bar{\epsilon}, \bar{\epsilon}), C^0(\mathbb{R}^d, \mathbb{R}^d))$ such that $\mathcal{G}_t^{\omega, \tau}(\epsilon, \cdot)$ is a C^1 -diffeomorphisms over $\overline{B(0, R_T)}$ for all $(t, \epsilon) \in [\tau, T] \times [0, \bar{\epsilon}]$ and it holds

$$\tilde{\mu}_t(\epsilon) = \mathcal{G}_t^{\omega, \tau}(\epsilon, \cdot) \# \mu^*(t).$$

Besides, there exists a constant $R_T^\Phi > 0$ depending on R_T, L_U and $v[\cdot](\cdot, \cdot)$ such that for all $(t, \epsilon) \in [\tau, T] \times (-\bar{\epsilon}, \bar{\epsilon})$ it holds $\text{supp}(\mathcal{G}_t^{\omega, \tau}(\epsilon, \cdot) \# \mu^(t)) \subset B(0, R_T^\Phi)$.*

This family of maps satisfies the Taylor expansion for all $t \in [\tau, T]$ with respect to the $L^2(\mathbb{R}^d, \mathbb{R}^d; \mu^(t))$ norm :*

$$\mathcal{G}_t^{\omega, \tau}(\epsilon, \cdot) = \text{Id} + \epsilon \mathcal{F}_t^{\omega, \tau} \circ \Phi_{(t, \tau)}^{v, u^*}[\mu^*(t)](\cdot) + o(\epsilon),$$

with

$$\mathcal{F}_t^{\omega, \tau} : x \in \text{supp}(\mu^*(\tau)) \mapsto D_x \Phi_{(\tau, t)}^{v, u^*}[\mu^*(\tau)](x) \cdot [\omega(x) - u^*(\tau, x)] + w_\Phi^{\omega, \tau}(t, x),$$

and where $w_\Phi^{\omega, \tau}(t, x)$ is the derivative at $\epsilon = 0$ of the map $\epsilon \in (-\bar{\epsilon}, \bar{\epsilon}) \mapsto \Phi_{(\tau, t)}^{v, u^*}[\tilde{\mu}_\tau(\epsilon)](x)$ as described in Theorem 1.9.

Moreover, the map $(t, x) \in [\tau, T] \times \text{supp}(\mu^*(\tau)) \mapsto \mathcal{F}_t^{\omega, \tau}(x)$ is the unique solution of the Cauchy problem

$$\begin{cases} \partial_t \mathcal{F}_t^{\omega, \tau}(x) = \int_{\mathbb{R}^d} \mathbb{I}^\circ \left(t, \Phi_{(\tau, t)}^{v, u^*}[\mu^*(\tau)](x) \right) \left(\Phi_{(\tau, t)}^{v, u^*}[\mu^*(\tau)](y) \right) \cdot \mathcal{F}_t^{\omega, \tau}(y) d\mu^*(\tau)(y) \\ + \left[D_x u^* \left(t, \Phi_{(\tau, t)}^{v, u^*}[\mu^*(\tau)](x) \right) + D_x v[\mu^*(t)] \left(t, \Phi_{(\tau, t)}^{v, u^*}[\mu^*(\tau)](x) \right) \right] \cdot \mathcal{F}_t^{\omega, \tau}(x) \\ \mathcal{F}_\tau^{\omega, \tau}(x) = \omega(x) - u^*(\tau, x). \end{cases} \quad (2.24)$$

Proof. We start by computing the measures $\tilde{\mu}_\tau(\epsilon)$ as a function of $\mu^*(\tau)$ for all $\epsilon \in [0, \bar{\epsilon}]$. By definition of the needle-like variation, it holds

$$\tilde{\mu}_\tau(\epsilon) = \Phi_{(\tau-\epsilon, \tau)}^{v, \omega}[\mu^*(\tau - \epsilon)] \circ \Phi_{(\tau, \tau-\epsilon)}^{v, u^*}[\mu^*(\tau)](\cdot) \# \mu^*(\tau)$$

Using Lebesgue's Differentiation Theorem, we obtain the following expansions at the first order with respect to ϵ

$$\begin{aligned} & \Phi_{(\tau-\epsilon, \tau)}^{v, \omega}[\mu^*(\tau - \epsilon)](y) \\ &= y + \int_{\tau-\epsilon}^\tau \left[v[\tilde{\mu}_t(\epsilon)] \left(t, \Phi_{(\tau-\epsilon, t)}^{v, \omega}[\mu^*(\tau - \epsilon)](y) \right) + \omega \left(\Phi_{(\tau-\epsilon, t)}^{v, \omega}[\mu^*(\tau - \epsilon)](y) \right) \right] dt, \\ &= y + \epsilon (v[\mu^*(\tau)](\tau, y) + \omega(y)) + o(\epsilon), \end{aligned}$$

as well as

$$\begin{aligned}
& \Phi_{(\tau, \tau-\epsilon)}^{v, u^*}[\mu^*(\tau)](y) \\
&= y - \int_{\tau-\epsilon}^{\tau} \left[v[\mu^*(t)] \left(t, \Phi_{(t, \tau-\epsilon)}^{v, u^*}[\mu^*(\tau)](y) \right) + u^* \left(t, \Phi_{(t, \tau-\epsilon)}^{v, u^*}[\mu^*(\tau)](y) \right) \right] dt, \\
&= y - \epsilon (v[\mu^*(\tau)](\tau, y) + u^*(\tau, y)) + o(\epsilon).
\end{aligned}$$

Chaining these two expressions together and recalling that $\omega(\cdot)$ and $v[\mu^*(\tau)](\tau, \cdot)$ are C^1 -smooth, it holds

$$\Phi_{(\tau-\epsilon, \tau)}^{v, \omega}[\mu^*(\tau-\epsilon)] \circ \Phi_{(\tau, \tau-\epsilon)}^{v, u^*}[\mu^*(\tau)](y) = y + \epsilon [\omega(y) - u^*(\tau, y)] + o(\epsilon)$$

and we deduce the expression that will prove useful in the sequel

$$\tilde{\mu}_\tau(\epsilon) = (\text{Id} + \epsilon [\omega(\cdot) - u^*(\tau, \cdot)] + o(\epsilon))_{\#} \mu^*(\tau). \quad (2.25)$$

We now want to obtain a similar expression but at some time $t \in [\tau, T]$. First, recall that $\tilde{\mu}_t(\epsilon) = \mathcal{G}_t^{\omega, \tau}(\epsilon, \cdot)_{\#} \mu^*(t)$ where

$$\begin{aligned}
\mathcal{G}_t^{\omega, \tau}(\epsilon, \cdot) : x \in \text{supp}(\mu^*(t)) &\mapsto \Phi_{(\tau, t)}^{v, u^*}[\tilde{\mu}_\tau(\epsilon)] \circ \Phi_{(\tau-\epsilon, \tau)}^{v, \omega}[\mu^*(\tau-\epsilon)] \\
&\circ \Phi_{(\tau, \tau-\epsilon)}^{v, u^*}[\mu^*(\tau)] \circ \Phi_{(t, \tau)}^{v, u^*}[\mu^*(t)](x).
\end{aligned} \quad (2.26)$$

By (2.25), one has the following expansion

$$\begin{aligned}
& \Phi_{(\tau, t)}^{v, u^*}[\tilde{\mu}_\tau(\epsilon)] \circ \Phi_{(\tau-\epsilon, \tau)}^{v, \omega}[\mu^*(\tau-\epsilon)] \circ \Phi_{(\tau, \tau-\epsilon)}^{v, u^*}[\mu^*(\tau)] \circ \Phi_{(t, \tau)}^{v, u^*}[\mu^*(t)](\cdot) \\
&= \Phi_{(\tau, t)}^{v, u^*}[\tilde{\mu}_\tau(\epsilon)] \circ \left(\Phi_{(t, \tau)}^{v, u^*}[\mu^*(\tau)](\cdot) \right) \\
&+ \epsilon \left[\omega \left(\Phi_{(t, \tau)}^{v, u^*}[\mu^*(\tau)](\cdot) \right) - u^* \left(\tau, \Phi_{(t, \tau)}^{v, u^*}[\mu^*(\tau)](\cdot) \right) \right] + o(\epsilon) \\
&= \Phi_{(\tau, t)}^{v, u^*}[\tilde{\mu}_\tau(\epsilon)] \left(\Phi_{(t, \tau)}^{v, u^*}[\mu^*(\tau)](\cdot) \right) + \epsilon D_x \Phi_{(\tau, t)}^{v, u^*}[\mu^*(\tau)] \left(\Phi_{(t, \tau)}^{v, u^*}[\mu^*(\tau)](\cdot) \right) \cdot \\
&\quad \left[\omega \left(\Phi_{(t, \tau)}^{v, u^*}[\mu^*(\tau)](\cdot) \right) - u^* \left(\tau, \Phi_{(t, \tau)}^{v, u^*}[\mu^*(\tau)](\cdot) \right) \right] + o(\epsilon).
\end{aligned} \quad (2.27)$$

since $\tilde{\mu}_\tau(\epsilon) \xrightarrow{W_1} \mu^*(\tau)$ as $\epsilon \downarrow 0$, $\mu \in \mathcal{P}_c(\mathbb{R}^d) \mapsto D_x \Phi_{(0, t)}^{v, u^*}[\mu](x)$ is continuous by **(H4)**, and we are only interested in a Taylor expansion at the first order in ϵ .

It then remains to compute the $O(\epsilon)$ term arising from $\Phi_{(\tau, t)}^{v, u^*}[\tilde{\mu}_\tau(\epsilon)](\cdot)$ in (2.27). Due to Theorem 1.9, the derivative of the non-local flow along directions induced by Lipschitz families of continuous and bounded maps exists. Recalling (2.25), this translates into

$$\Phi_{(\tau, t)}^{v, u^*}[\tilde{\mu}_\tau(\epsilon)](y) = \Phi_{(\tau, t)}^{v, u^*}[\mu^*(\tau)](y) + \epsilon w_{\Phi}^{\omega, \tau}(t, y) + o(\epsilon),$$

where $w_{\Phi}^{\omega, \tau}(t, y)$ is defined through (1.37) in the case where the non-local velocity field is given by $(t, x) \mapsto v[\mu^*(t)](t, x) + u^*(t, x)$.

Thus, we proved the following pointwise Taylor expansion at the first order with respect to ϵ

$$\mathcal{G}_t^{\omega, \tau}(\epsilon, x) = x + \epsilon \mathcal{F}_t^{\omega, \tau} \circ \Phi_{(t, \tau)}^{u^*, v}[\mu^*(t)](x) + o(\epsilon)$$

for $\mu^*(t)$ -almost every $x \in \mathbb{R}^d$, where

$$\mathcal{F}_t^{\omega, \tau} : x \in \text{supp}(\mu^*(\tau)) \mapsto D_x \Phi_{(\tau, t)}^{v, u^*}[\mu^*(\tau)](x) \cdot [\omega(x) - u^*(\tau, x)] + w_{\Phi}^{\omega, \tau}(t, x) \quad (2.28)$$

is a continuous mapping for all $t \in [\tau, T]$.

A standard application of Proposition 1.2-(iii) shows that this expansion holds in $L^2(\mathbb{R}^d, \mathbb{R}^d; \mu^*(t))$. One can then extend $\mathcal{G}_t^{\omega, \tau}(\cdot, \cdot)$ to $(-\bar{\epsilon}, \bar{\epsilon})$ while preserving this expansion around $\epsilon = 0$ by defining e.g.

$$\mathcal{G}_t^{\omega, \tau}(\epsilon, \cdot) \equiv \text{Id} + \epsilon \mathcal{F}_t^{\omega, \tau} \circ \Phi_{(\tau, t)}^{v, u^*}[\mu^*(\tau)](\cdot)$$

for $\epsilon \in (-\bar{\epsilon}, 0]$. The existence of a constant R_T^Φ depending on R_T, L_U and $v[\cdot](\cdot, \cdot)$ such that

$$\text{supp}(\mathcal{G}_t^{\omega, \tau}(\epsilon, \cdot) \# \mu^*(t)) \subset \overline{B(0, R_T^\Phi)}$$

follows from hypotheses **(H4)** and **(H5)**, which ensure the continuity and boundedness of the perturbation functions over $\overline{B(0, \tilde{R}_T)}$.

We finally prove the counterpart of Lemma 2.2 providing an ODE-type characterization for the perturbation induced by the needle-like variation in the non-local case. Recalling the definition of $(t, x) \mapsto \mathcal{F}_t^{\omega, \tau}(x)$ given in (2.28) and summing the ODE-type characterization of $t \mapsto w_{\Phi}^{\omega, \tau}(t, \cdot)$ and $D_x \Phi_{(\tau, \cdot)}^{v, u^*}[\mu^*(\tau)](\cdot) \cdot [\omega(\cdot) - u^*(\tau, \cdot)]$, we recover

$$\begin{cases} \partial_t \mathcal{F}_t^{\omega, \tau}(x) = \int_{\mathbb{R}^d} \mathbb{I}^\circ \left(t, \Phi_{(\tau, t)}^{v, u^*}[\mu^*(\tau)](x) \right) \left(\Phi_{(\tau, t)}^{v, u^*}[\mu^*(\tau)](y) \right) \cdot \mathcal{F}_t^{\omega, \tau}(y) d\mu^*(\tau)(y) \\ + \left[D_x u^* \left(t, \Phi_{(\tau, t)}^{v, u^*}[\mu^*(\tau)](x) \right) + D_x v[\mu^*(t)] \left(t, \Phi_{(\tau, t)}^{v, u^*}[\mu^*(\tau)](x) \right) \right] \cdot \mathcal{F}_t^{\omega, \tau}(x) \\ \mathcal{F}_\tau^{\omega, \tau}(x) = \omega(x) - u^*(\tau, x), \end{cases}$$

which concludes the proof of our result. \square

In the development of Steps 2, 3 and 4, we do not need to take into account the explicit dependence of the flows with respect to their starting measures. We shall henceforth write $\Phi_{(\cdot, \cdot)}^{v, u^*}(\cdot) \equiv \Phi_{(\cdot, \cdot)}^{v, u^*}[\mu^*(\cdot)](\cdot)$ for clarity and conciseness.

Step 2 : First-order optimality condition

In the framework of Problem (\mathcal{P}) , the optimality of $u^*(\cdot)$ writes

$$\begin{aligned} \frac{\varphi(\tilde{\mu}_T(\epsilon)) - \varphi(\mu^*(T))}{\epsilon} + \frac{1}{\epsilon} \int_{\tau-\epsilon}^{\tau} [L(\tilde{\mu}_t(\epsilon), \omega) - L(\mu^*(t), u^*(t))] dt \\ + \frac{1}{\epsilon} \int_{\tau}^T [L(\tilde{\mu}_t(\epsilon), u^*(t)) - L(\mu^*(t), u^*(t))] dt \geq 0 \end{aligned} \quad (2.29)$$

for all $\epsilon \in [0, \bar{\epsilon})$. The first order perturbation corresponding to the final cost $\varphi(\cdot)$ has already been treated in (2.15)-(2.16), Section 2.1. We study the integral terms arising from the running cost. Remark first that it holds

$$\lim_{\epsilon \downarrow 0} \left[\frac{1}{\epsilon} \int_{\tau-\epsilon}^{\tau} [L(\tilde{\mu}_t(\epsilon), \omega) - L(\mu^*(t), u^*(t))] dt \right] = L(\mu^*(\tau), \omega) - L(\mu^*(\tau), u^*(\tau)),$$

by the Lebesgue Differentiation Theorem, since $\tilde{\mu}_t(\epsilon) \xrightarrow{W_1} \mu^*(t)$ as $\epsilon \downarrow 0$ for all $t \in [0, T]$ and since τ is a Lebesgue point of $u^*(\cdot)$.

Equivalently to the proof of the PMP for Problem (\mathcal{P}_1) , the perturbed measures are uniformly supported in a compact set. Thus, under hypotheses **(H2)** and recalling that the function $\epsilon \in (-\bar{\epsilon}, \bar{\epsilon}) \mapsto L(\tilde{\mu}_t(\epsilon), u^*(t))$ is differentiable at $\epsilon = 0$ for \mathcal{L}^1 -almost every $t \in [\tau, T]$ by hypothesis **(H6)**, the chain rule of Proposition 1.6 can be applied to the running cost to obtain

$$\lim_{\epsilon \downarrow 0} \left[\frac{1}{\epsilon} [L(\tilde{\mu}_t(\epsilon), u^*(t)) - L(\mu^*(t), u^*(t))] \right] = \int_{\mathbb{R}^d} \langle \bar{\gamma}_L^\circ(t, x), \mathcal{F}_t^{\omega, \tau} \circ \Phi_{(t, \tau)}^{v, u^*}(x) \rangle d\mu^*(t)(x), \quad (2.30)$$

where $\bar{\gamma}_L^\circ(t, \cdot) \in L^2(\mathbb{R}^d, \mathbb{R}^d; \mu^*(t))$ is the barycenter of $\partial_\mu^\circ L(\mu^*(t), u^*(t))$ for \mathcal{L}^1 -almost every $t \in [0, T]$.

Moreover, the uniform compactness of the supports of the perturbed measures and hypothesis **(H2)** imply that the left hand side in (2.30) is uniformly bounded by a function in $L^1([0, T], \mathbb{R}_+)$ for any $\epsilon \in (0, \bar{\epsilon})$. Therefore, it holds by an application of Lebesgue Dominated Convergence Theorem that

$$\begin{aligned} & \lim_{\epsilon \downarrow 0} \left[\frac{1}{\epsilon} \int_\tau^T [L(\tilde{\mu}_t(\epsilon), u^*(t)) - L(\mu^*(t), u^*(t))] dt \right] \\ &= \int_\tau^T \int_{\mathbb{R}^d} \langle \bar{\gamma}_L^\circ(t, x), \mathcal{F}_t^{\omega, \tau} \circ \Phi_{(t, \tau)}^{v, u^*}(x) \rangle d\mu^*(t)(x) dt. \end{aligned}$$

Thus, the optimality of $(u^*(\cdot), \mu^*(\cdot))$ translates into the first-order condition

$$\begin{aligned} & \int_{\mathbb{R}^d} \langle \bar{\gamma}_\varphi^\circ(x), \mathcal{F}_T^{\omega, \tau} \circ \Phi_{(T, \tau)}^{v, u^*}(x) \rangle d\mu^*(T)(x) + [L(\mu^*(\tau), \omega) - L(\mu^*(\tau), u^*(\tau))] \\ &+ \int_\tau^T \int_{\mathbb{R}^d} \langle \bar{\gamma}_L^\circ(t, x), \mathcal{F}_t^{\omega, \tau} \circ \Phi_{(t, \tau)}^{v, u^*}(x) \rangle d\mu^*(t)(x) dt \geq 0, \end{aligned} \quad (2.31)$$

for any couple of needle parameters (ω, τ) .

Step 3 : Backward dynamics and Pontryagin maximization condition

We now build a solution $\nu^*(\cdot) \in \text{Lip}([0, T], \mathcal{P}_c(\mathbb{R}^{2d}))$ to the system of continuity equations

$$\begin{cases} \partial_t \nu^*(t) + \nabla \cdot (\mathcal{V}^*[\nu^*(t)](t, \cdot, \cdot) \nu^*(t)) = 0 & \text{in } [0, T] \times \mathbb{R}^{2d}, \\ \pi_{\#}^1 \nu^*(t) = \mu^*(t) & \text{for all } t \in [0, T], \\ \nu^*(T) = (\text{Id} \times (-\bar{\gamma}_\varphi^\circ))_{\#} \mu^*(T), \end{cases} \quad (2.32)$$

associated with the non-local vector field

$$\begin{aligned} \mathcal{V}^*[\nu^*(t)] : (t, x, r) \in [0, T] \times \text{supp}(\nu) \mapsto \\ \left(\begin{array}{l} v[\pi_{\#}^1 \nu^*(t)](t, x) + u^*(t, x) \\ \bar{\gamma}_L^\circ(t, x) - \Pi_v^\circ[\nu^*(t)](t, x) - D_x u^*(t, x)^\top r - D_x v[\mu^*(t)](t, x)^\top r \end{array} \right) \end{aligned}$$

where $\mathbb{I}_v^\circ[\cdot](\cdot, \cdot)$ is given by

$$\mathbb{I}_v^\circ[\nu](t, x) = \int_{\mathbb{R}^{2d}} \left(\mathbb{I}_{(t,y)}^\circ(x) \right)^\top p \, d\nu(y, p) \quad (2.33)$$

for any $(\nu, t, x) \in \mathcal{P}_c(\mathbb{R}^{2d}) \times [0, T] \times \text{supp}(\pi_{\#}^1 \nu)$. Here, $\mathbb{I}_{(\cdot, \cdot)}^\circ(\cdot)$ is defined as in Theorem 2.1.

As in Lemma 2.3 of Section 2.1, we build a solution $\nu^*(\cdot)$ of (2.32) by making use of the cascaded structure of the system. We then show that this solution is such that the map $\mathcal{K}_{\omega, \tau}(\cdot)$ defined in this context by

$$\begin{aligned} \mathcal{K}_{\omega, \tau} : t \in [\tau, T] \mapsto & \int_{\mathbb{R}^{2d}} \langle r, \mathcal{F}_t^{\omega, \tau} \circ \Phi_{(t, \tau)}^{v, u^*}(x) \rangle d\nu^*(t)(x, r) + [L(\mu^*(\tau), u^*(\tau)) - L(\mu^*(\tau), \omega)] \\ & - \int_{\tau}^t \int_{\mathbb{R}^{2d}} \langle \bar{\gamma}_L^\circ(t, x), \mathcal{F}_s^{\omega, \tau} \circ \Phi_{(s, \tau)}^{v, u^*}(x) \rangle d\mu^*(s)(x) ds \end{aligned} \quad (2.34)$$

is constant over $[\tau, T]$.

Lemma 2.6 (Well-posedness of solutions of (2.32)). *Let $(u^*(\cdot), \mu^*(\cdot))$ be an optimal pair control-trajectory for (\mathcal{P}) . We consider the family of maps $(t, x, r) \in [0, T] \times \mathbb{R}^{2d} \mapsto \Psi_t(x, r) \in \mathbb{R}^d$, defined to be the solution of*

$$\begin{cases} \partial_t w(t, x, r) = \bar{\gamma}_L^\circ(t, \Phi_{(T,t)}^{v, u^*}(x)) - D_x u^*(t, \Phi_{(T,t)}^{v, u^*}(x))^\top w(t, x, r) \\ \quad - D_x v[\mu^*(t)](t, \Phi_{(T,t)}^{v, u^*}(x))^\top w(t, x, r) \\ - \int_{\mathbb{R}^{2d}} \mathbb{I}_{(t, \Phi_{(T,t)}^{v, u^*}(y))}^\circ \left(\Phi_{(T,t)}^{v, u^*}(x) \right)^\top w(t, y, p) \, d(\text{Id} \times (-\bar{\gamma}_\varphi^\circ))_{\#} \mu^*(T)(y, p), \\ w(T, x, r) = r. \end{cases} \quad (2.35)$$

For $\mu^*(T)$ -almost every $x \in \mathbb{R}^d$, we define the curves of measures $\sigma_x^* : t \mapsto \Psi_t(x, \cdot)_{\#} \delta_{(-\bar{\gamma}_\varphi^\circ(x))}$ and denote by $\mathcal{V}_x^*(\cdot, \cdot)$ the corresponding non-local vector fields describing their evolution.

Then $\nu^* : t \mapsto (\Phi_{(T,t)}^{v, u^*}, \text{Id})_{\#} \nu_T^*(t)$ solves (2.32) with $\nu_T^*(t) = \int \sigma_x^*(t) d\mu^*(T)(x)$. Moreover, there exists two constants $R'_T, L'_T > 0$ such that

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

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

Proof. We denote by Ω the compact² subset $\text{supp}((\text{Id} \times (-\bar{\gamma}_\varphi^\circ))_{\#} \mu^*(T))$ of \mathbb{R}^{2d} . We first show that (2.35) admits a unique continuous solution $(t, x, r) \in [0, T] \times \Omega \mapsto \Psi_t(x, r) \in \mathbb{R}^d$. This can be done by reproducing the strategy of the proof of Theorem 1.9 which consists in defining a weighted $C^0([0, T] \times \Omega)$ -norm for which the right-hand side of (2.35) is contracting and applying Banach's Fixed Point Theorem. Notice that here, the coupling between the non-local flows arising from the integral term in (2.35) requires us to use explicitly the continuity of the right-hand side with respect to x . In Lemma 2.3, all the backward Cauchy problems were independent and we did not need any regularity assumption on x for the proof to work.

2. Recall that $\bar{\gamma}_\varphi^\circ(\cdot)$ is a continuous map by hypothesis **(H5)**.

Since $[0, T] \times \Omega$ is compact, $(t, x, r) \mapsto \Psi_t(x, r)$ is bounded. This implies by (2.35) along with hypotheses **(H1)**, **(H4)** and **(H5)** that $t \mapsto \Psi_t(x, r)$ is Lipschitz for all $(x, r) \in \Omega$. Moreover, a direct application of Grönwall Lemma along with (2.35) allows to show that for all $(t, x) \in [0, T] \times \pi^1(\Omega)$, ones has

$$|\Psi_t(x, r_2) - \Psi_t(x, r_1)| \leq C|r_2 - r_1|$$

for all $(r_1, r_2) \in \pi^2(\Omega)$. Hence, we showed that $(t, r) \in [0, T] \times \pi^2(\Omega) \mapsto \Psi_t(x, r)$ is Lipschitz for all $x \in \pi^1(\Omega)$.

Therefore, carrying out same computations as in Lemma 2.3, we show that the curves of measures $t \mapsto \sigma_x^*(t) = \Psi_t(x, \cdot) \# \delta_{-\bar{\gamma}_{\varphi(x)}^\circ}$ are well-defined, uniformly compactly supported and Lipschitz in the W_1 -metric for $\mu^*(T)$ -almost every $x \in \mathbb{R}^d$. This implies the existence of two constants $R'_T, L'_T > 0$ such that it holds

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

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

Moreover, for any $\xi \in C_c^\infty(\mathbb{R}^{2d}, \mathbb{R})$, it holds that

$$\begin{aligned} & \frac{d}{dt} \left[\int_{\mathbb{R}^{2d}} \xi(x, r) d\nu^*(t)(x, r) \right] = \frac{d}{dt} \left[\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \xi(\Phi_{(T,t)}^{v, u^*}(x), r) d\sigma_x^*(t)(r) d\mu^*(T)(x) \right] \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \langle \nabla_x \xi(\Phi_{(T,t)}^{v, u^*}(x), r), v[\mu^*(t)](t, \Phi_{(T,t)}^{v, u^*}(x)) \rangle d\sigma_x^*(t)(r) d\mu^*(T)(x) \\ &+ \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \langle \nabla_x \xi(\Phi_{(T,t)}^{v, u^*}(x), r), u^*(t, \Phi_{(T,t)}^{v, u^*}(x)) \rangle d\sigma_x^*(t)(r) d\mu^*(T)(x) \\ &+ \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \langle \nabla_r \xi(\Phi_{(T,t)}^{v, u^*}(x), r), \bar{\gamma}_L^\circ(t, \Phi_{(T,t)}^{v, u^*}(x)) \rangle d\sigma_x^*(t)(r) d\mu^*(T)(x) \\ &- \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \langle \nabla_r \xi(\Phi_{(T,t)}^{v, u^*}(x), r), D_x u^*(t, \Phi_{(T,t)}^{v, u^*}(x))^\top r \rangle d\sigma_x^*(t)(r) d\mu^*(T)(x) \\ &- \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \langle \nabla_r \xi(\Phi_{(T,t)}^{v, u^*}(x), r), D_x v[\mu^*(t)](t, \Phi_{(T,t)}^{v, u^*}(x))^\top r \rangle d\sigma_x^*(t)(r) d\mu^*(T)(x) \\ &- \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \langle \nabla_r \xi(\Phi_{(T,t)}^{v, u^*}(x), r), \mathbb{I}_v^\circ[\nu^*(t)](t, \Phi_{(T,t)}^{v, u^*}(x)) \rangle d\sigma_x^*(t)(r) d\mu^*(T)(x) \\ &= \int_{\mathbb{R}^{2d}} \langle \nabla_x \xi(x, r), v[\mu^*(t)](t, x) + u^*(t, x) \rangle d\nu^*(t)(x, r) \\ &+ \int_{\mathbb{R}^{2d}} \langle \nabla_r \xi(x, r), \bar{\gamma}_L^\circ(t, x) - D_x u^*(t, x)^\top r - D_x v[\mu^*(t)](t, x)^\top r \rangle d\nu^*(t)(x, r) \\ &- \int_{\mathbb{R}^{2d}} \langle \nabla_r \xi(x, r), \mathbb{I}_v^\circ[\nu^*(t)](t, x) \rangle d\nu^*(t)(x, r) \end{aligned}$$

where we used the fact that

$$\begin{aligned}
\mathbb{I}^\circ_v[\nu^*(t)](t, \Phi_{(T,t)}^{v,u^*}(x)) &= \int_{\mathbb{R}^{2d}} \mathbb{I}^\circ \left(t, \Phi_{(T,t)}^{v,u^*}(y) \right) \left(\Phi_{(T,t)}^{v,u^*}(x) \right)^\top p \, d\nu_T^*(t)(y, p) \\
&= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{I}^\circ \left(t, \Phi_{(T,t)}^{v,u^*}(y) \right) \left(\Phi_{(T,t)}^{v,u^*}(x) \right)^\top p \, d\sigma_y^*(t)(p) d\mu^*(T)(y) \\
&= \int_{\mathbb{R}^{2d}} \mathbb{I}^\circ \left(t, \Phi_{(T,t)}^{v,u^*}(y) \right) \left(\Phi_{(T,t)}^{v,u^*}(x) \right)^\top \Psi_t(y, p) \, d((\text{Id} \times (-\bar{\gamma}_\varphi^\circ)_{\#} \mu^*(T))(y, p))
\end{aligned}$$

We therefore recover that $t \mapsto \nu^*(t)$ solves (2.32), which ends the proof. \square

Lemma 2.7. *The map $t \mapsto \mathcal{K}_{\omega, \tau}(t)$ defined in (2.34) is constant over $[\tau, T]$ for any couple of needle parameters (ω, τ) .*

Proof. This proof follows the same steps as in the proof of Lemma 2.4, the difference lying in the fact that the flows $(\Phi_{(0,t)}^{v,u^*}(\cdot))_{t \in [0, T]}$ are associated to the non-local PDE. It can be verified again as in the proof of Lemma 2.4 that $t \mapsto \mathcal{K}_{\omega, \tau}(t)$ is Lipschitz. We compute $\frac{d}{dt} \mathcal{K}_{\omega, \tau}$ using (1.30) as in (2.23) while plugging in the expressions for $\mathcal{F}_t^{\omega, \tau}(\cdot)$ and its time-derivative provided by Lemma 2.5.

$$\begin{aligned}
&\frac{d}{dt} \mathcal{K}_{\omega, \tau}(t) \\
&= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left\langle r, \partial_t \mathcal{F}_t^{\omega, \tau} \circ \Phi_{(T, \tau)}^{v, u^*}(x) \right\rangle d\sigma_x^*(t)(r) d\mu^*(T)(x) \\
&+ \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left\langle \mathcal{V}_x^*(t, r), \mathcal{F}_t^{\omega, \tau} \circ \Phi_{(T, \tau)}^{v, u^*}(x) \right\rangle d\sigma_x^*(t)(r) d\mu^*(T)(x) \\
&- \int_{\mathbb{R}^{2d}} \left\langle \bar{\gamma}_L^\circ(t, \Phi_{(T,t)}^{v, u^*}(x)), \mathcal{F}_t^{\omega, \tau} \circ \Phi_{(T, \tau)}^{v, u^*}(x) \right\rangle d\mu^*(T)(x) \\
&= \int_{\mathbb{R}^{2d}} \left\langle r, \int_{\mathbb{R}^d} \mathbb{I}^\circ \left(t, \Phi_{(T,t)}^{v, u^*}(x) \right) \left(\Phi_{(T,t)}^{v, u^*}(y) \right) \mathcal{F}_t^{\omega, \tau} \circ \Phi_{(T, \tau)}^{v, u^*}(x) d\mu^*(T)(y) \right\rangle d\nu_T^*(t)(x, r) \\
&- \int_{\mathbb{R}^{2d}} \left\langle \int_{\mathbb{R}^{2d}} \mathbb{I}^\circ \left(t, \Phi_{(T,t)}^{v, u^*}(y) \right) \left(\Phi_{(T,t)}^{v, u^*}(x) \right)^\top p \, d\nu_T^*(t)(y, p), \mathcal{F}_t^{\omega, \tau} \circ \Phi_{(T, \tau)}^{v, u^*}(x) \right\rangle d\nu_T^*(t)(x, r) \\
&= 0
\end{aligned}$$

by plugging in the expressions of $\partial_t \mathcal{F}_t^{\omega, \tau}(\cdot)$ and $\mathcal{V}_x^*(t, \cdot)$. The two quantities are shown to be equal due to the uniform boundedness of the integrands given by **(H5)** and Fubini-Tonelli theorem. This altogether leads to $\frac{d}{dt} \mathcal{K}_{\omega, \tau}(t) = 0$ for \mathcal{L}^1 -almost every $t \in [\tau, T]$ and thus to $\mathcal{K}_{\omega, \tau}(\cdot)$ being constant over $[\tau, T]$. \square

Step 4 : Proof of the Pontryagin Maximum Principle for (\mathcal{P})

We proved in Lemma 2.6 that there exists a curve $\nu^* \in \text{Lip}([0, T], \mathcal{P}_c(\mathbb{R}^{2d}))$ solution of (2.32) along with a constant $R \equiv R'_T > 0$ such that $\text{supp}(\nu^*(\cdot)) \subset \overline{B_{2d}(0, R)}$. The non-local velocity field $\mathcal{V}^*[\nu^*(\cdot)](\cdot, \cdot, \cdot)$ is given in this context for $\mathcal{L}^1 \times \nu^*(\cdot)$ -almost every

$(t, x, r) \in [0, T] \times \overline{B_{2d}(0, R)}$ by

$$\mathcal{V}^*[\nu^*(t)](t, x, r) = \begin{pmatrix} \bar{\gamma}_L^\circ(t, x) - D_x u^*(t, x)^\top r - D_x v[\pi_{\#}^1 \nu^*(t)](t, x)^\top r - \mathbb{F}_v^\circ[\nu^*(t)](t, x) \\ v[\pi_{\#}^1 \nu^*(t)](t, x) + u^*(t, x) \end{pmatrix}.$$

We also define the infinite-dimensional Hamiltonian $\mathbb{H}(\cdot, \cdot, \cdot)$ of the system by

$$\mathbb{H} : (t, \nu, \omega) \in [0, T] \times \mathcal{P}_c(\mathbb{R}^{2d}) \times U \mapsto \int_{\mathbb{R}^{2d}} \langle r, \omega(x) \rangle d\nu(x, r) - L(\pi_{\#}^1 \nu, \omega).$$

along with its compactification $\mathbb{H}_c(\cdot, \cdot, \cdot)$ given by (2.3).

We proved in Lemma 2.7 that the solution $\nu^*(\cdot)$ that we built is such that the map $\mathcal{K}_{\omega, \tau}(\cdot)$ defined in (2.34) is constant over $[\tau, T]$ for any couple of needle parameters (ω, τ) . Hence, it holds in particular that $\mathcal{K}_{\omega, \tau}(\tau) = \mathcal{K}_{\omega, \tau}(T)$ which is a non-positive quantity by the first-order optimality condition (2.31). This fact implies that

$$\int_{\mathbb{R}^{2d}} \langle r, \omega(x) - u^*(\tau, x) \rangle d\nu^*(\tau)(x, r) - [L(\mu^*(t), \omega) - L(\mu^*(t), u^*(t))] \leq 0,$$

for all $\omega \in U$ and $\tau \in [0, T]$ Lebesgue point of $t \in [0, T] \mapsto v[\mu^*(t)](t, \cdot) + u^*(t, \cdot)$. This inequality straightforwardly rewrites as the Pontryagin Maximization condition (2.5), i.e.

$$\mathbb{H}_c(t, \nu^*(t), u^*(t)) = \max_{\omega \in U} [\mathbb{H}_c(t, \nu^*(t), \omega)]$$

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

Finally, one recognizes the pseudo-Hamiltonian structure generated by

$$\mathcal{V}^*[\nu^*(t)](t, x, r) = \mathbb{J}_{2d} \tilde{\nabla}_\nu \mathbb{H}_c(t, \nu^*(t), u^*(t))$$

for $\mathcal{L}^1 \times \nu^*(\cdot)$ -almost every $(t, x, r) \in [0, T] \times \overline{B_{2d}(0, R)}$ where one can check that the map $\tilde{\nabla}_\nu \mathbb{H}_c(t, \nu^*(t), u^*(t))(\cdot, \cdot)$ is precisely the non-local velocity field $\mathbb{J}_{2d} \mathcal{V}^*[\nu^*](t, \cdot, \cdot)$ for \mathcal{L}^1 -almost every $t \in [0, T]$. This concludes our proof of Theorem 2.1.

2.3 Examples of functionals satisfying hypotheses (H) of

Theorem 2.1

The aim of the general result stated in Theorem 2.1 is to provide first-order necessary optimality conditions that are adapted to a wide range of functionals. We give in the following Propositions some examples of classical functionals that are encompassed in hypotheses (H) and compute the minimal selection in their Wasserstein subdifferential.

Proposition 2.1 (Subdifferential of a smooth integral functional).

Let $V \in C^1(\mathbb{R}^d, \mathbb{R})$ and $K \subset \mathbb{R}^d$ be a compact set. Define $\mathcal{V} : \mu \in \mathcal{P}(K) \mapsto \int_{\mathbb{R}^d} V(x) d\mu(x)$. Then the functional $\mathcal{V}(\cdot)$ is regular at any $\mu \in \mathcal{P}(K)$ in the sense of Definition 1.8, Lipschitz in the W_1 -metric. Moreover, the minimal selection $\partial^\circ \mathcal{V}(\mu)$ in its extended subdifferential at μ is a classical strong subdifferential induced by a map and given explicitly by $\partial^\circ \mathcal{V}(\mu) = (\text{Id} \times \nabla V)_{\#} \mu$.

Proof. See e.g. [17, Proposition 10.4.2]. \square

Remark 2.4. Taking any power $\alpha > 0$ of $\mathcal{V}(\cdot)$ yields the same results provided that the functional $x \mapsto x^\alpha$ is differentiable at $\mathcal{V}(\mu)$. In which case, the minimal selection in the extended subdifferential is induced by the map

$$\nabla_\mu(\mathcal{V}^\alpha)(\mu) : x \in \text{supp}(\mu) \mapsto \alpha \mathcal{V}(\mu)^{\alpha-1} \nabla V(x). \quad (2.36)$$

Proposition 2.2 (Subdifferential of the variance functional). *Let $K \subset \mathbb{R}^d$ be a compact set and define the variance functional by*

$$\text{Var} : \mu \in \mathcal{P}(K) \mapsto \frac{1}{2} \int_{\mathbb{R}^d} |x - \bar{\mu}|^2 d\mu(x) = \frac{1}{2} \int_{\mathbb{R}^d} |x|^2 d\mu(x) - \frac{1}{2} |\bar{\mu}|^2$$

where $\bar{\mu} = \int_{\mathbb{R}^d} x d\mu(x)$ denotes the average of the measure μ .

Then, the functional $\text{Var}(\cdot)$ is regular at any $\mu \in \mathcal{P}(K)$, Lipschitz in the W_1 -metric and the minimal selection $\partial^\circ \text{Var}(\mu)$ in its extended subdifferential is a classical strong subdifferential induced by the map $\nabla_\mu \text{Var}(\mu) : x \in \text{supp}(\mu) \mapsto x - \bar{\mu}$.

Proof. It is clear by definition of the variance functional that it is bounded from below over $\mathcal{P}(K)$. Moreover, an application of the Kantorovich-Rubinstein duality formula (1.4) yields the Lipschitzianity in the W_1 -metric. The regularity in the sense of Definition 1.8 is a consequence of the convexity along Wasserstein geodesics of the variance functional (see [17, Lemma 10.3.8]).

We now show that $x \mapsto x - \bar{\mu}$ is in the classical strong subdifferential of the variance functional at $\mu \in \mathcal{P}(K)$. For any $\nu \in \mathcal{P}(K)$ and $\mu \in \Gamma(\mu, \nu)$, it holds that

$$\begin{aligned} & \int_{\mathbb{R}^{2d}} \langle x_1 - \bar{\mu}, x_2 - x_1 \rangle d\mu(x_1, x_2) \\ &= \int_{\mathbb{R}^{2d}} \langle x_1, x_2 \rangle d\mu(x_1, x_2) - \int_{\mathbb{R}^d} |x_1|^2 d\mu(x_1) + |\bar{\mu}|^2 - \langle \bar{\mu}, \bar{\nu} \rangle \\ &\leq \frac{1}{2} \int_{\mathbb{R}^d} |x_2|^2 d\nu(x_2) - \frac{1}{2} \int_{\mathbb{R}^d} |x_1|^2 d\mu(x_1) + |\bar{\mu}|^2 - \langle \bar{\mu}, \bar{\nu} \rangle \\ &\leq \text{Var}(\nu) - \text{Var}(\mu) + \frac{1}{2} |\bar{\mu} - \bar{\nu}|^2. \end{aligned}$$

Moreover, one can estimate the quantity $|\bar{\mu} - \bar{\nu}|^2$ as follows

$$\begin{aligned} |\bar{\mu} - \bar{\nu}|^2 &\leq \left(\int_{\mathbb{R}^{2d}} |x_1 - x_2| d\mu(x_1, x_2) \right)^2 \\ &\leq \int_{\mathbb{R}^{2d}} |x_1 - x_2|^2 d\mu(x_1, x_2) = W_{2, \mu}^2(\mu, \nu) = o(W_{2, \mu}(\mu, \nu)), \end{aligned}$$

since $\mu \in \Gamma(\mu, \nu)$, invoking Jensen's inequality and the definition of $W_{2, \mu}(\cdot, \cdot)$ given in Definition 1.7. Therefore, we conclude that for any $\nu \in \mathcal{P}(K)$ and any $\mu \in \Gamma(\mu, \nu)$ it holds

$$\text{Var}(\nu) - \text{Var}(\mu) \geq \int_{\mathbb{R}^{2d}} \langle x_1 - \bar{\mu}, x_2 - x_1 \rangle d\mu(x_1, x_2) + o(W_{2, \mu}(\mu, \nu)).$$

which is equivalent to $x \in \text{supp}(\mu) \mapsto x - \bar{\mu}$ being a classical strong subdifferential at μ .

Now, take in particular $\nu \equiv \nu_s = (\text{Id} + s\xi) \# \mu$ for some small $s > 0$ and $\xi \in C_c^\infty(\mathbb{R}^d)$ such that $\text{supp}(\nu_s) \subset \mathcal{P}(K)$. It then holds

$$+\infty > \lim_{s \downarrow 0} \left[\frac{\text{Var}((\text{Id} + s\xi) \# \mu) - \text{Var}(\mu)}{s} \right] \geq \int_{\mathbb{R}^d} \langle x_1 - \bar{\mu}, \xi(x_1) \rangle d\mu(x_1).$$

Furthermore, one can check that it holds

$$\begin{aligned} & \lim_{s \downarrow 0} \left[\frac{\text{Var}((\text{Id} + s\xi) \# \mu) - \text{Var}(\mu)}{s} \right] \\ & \leq \limsup_{s \downarrow 0} \left[\frac{(\text{Var}((\text{Id} + s\xi) \# \mu) - \text{Var}(\mu))^+}{W_2(\mu, (\text{Id} + s\xi) \# \mu)} \right] \limsup_{s \downarrow 0} \left[\frac{W_2(\mu, (\text{Id} + s\xi) \# \mu)}{s} \right] \\ & \leq |\partial \text{Var}|(\mu) \|\xi\|_{L^2(\mu)}, \end{aligned}$$

so that, for any $\xi \in C_c^\infty(\mathbb{R}^d)$ with $\|\xi\|_{L^2(\mu)} \leq 1$, one has

$$\int_{\mathbb{R}^d} \langle x_1 - \bar{\mu}, \xi(x_1) \rangle d\mu(x_1) \leq |\partial \text{Var}|(\mu).$$

By applying a density argument for test functions in the space $L^2(\mathbb{R}^d, \mathbb{R}^d; \mu)$ and using the dual characterization of the L^2 -norm of a functional, it finally holds that $\|\text{Id} - \bar{\mu}\|_{L^2(\mu)} \leq |\partial \text{Var}|(\mu)$, which amounts to state by Theorem 1.5 that the strong subdifferential $x \in \text{supp}(\mu) \mapsto x - \bar{\mu}$ is the minimal selection in the classical subdifferential $\partial \text{Var}(\mu)$ of the variance functional at μ . \square

Remark 2.5 (Possible extensions). *The analysis carried out in the previous Proposition for the variance functional can be applied in a similar fashion to integral functionals of the form*

$$\begin{cases} \mathcal{W}^k : \mu \in \mathcal{P}(K) \mapsto \int_{\mathbb{R}^d} W(x_1, \dots, x_k) d(\mu \times \dots \times \mu)(x_1, \dots, x_k) \\ \mathcal{V}_m : \mu \in \mathcal{P}(K) \mapsto \int_{\mathbb{R}^d} V(x, \int_{\mathbb{R}^d} m(y) d\mu(y)) d\mu(x) \end{cases}$$

for any $k \geq 1$, $W \in C^1(\mathbb{R}^{d \times k}, \mathbb{R})$, $V \in C^2(\mathbb{R}^d \times \mathbb{R}^n, \mathbb{R})$ and $m \in C^2(\mathbb{R}^d, \mathbb{R}^n)$ for some $n \geq 1$.

Proposition 2.3 (Subdifferential of a smooth convolution interaction). *Let $K \subset \mathbb{R}^d$ be a compact set, $H(\cdot, \cdot) \in C^1(\mathbb{R}^{2d}, \mathbb{R}^d)$ be a function with sublinear growth and consider the non-local velocity field $v\cdot : (\mu, x) \in \mathcal{P}_c(K) \times \mathbb{R}^d \mapsto \int_{\mathbb{R}^d} H(x, y) d\mu(y)$.*

*Then, $v[\cdot](\cdot, \cdot)$ satisfies **(H4)**, **(H5)** and **(H6)** and the first order variations $x \in \text{supp}(\mu) \mapsto D_x v[\mu](x)$ and $x \in \text{supp}(\mu) \mapsto \int_{\mathbb{R}^d} \Gamma_x^\circ(y) d\mu(y)$ can be computed explicitly as*

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

where $\Gamma_x^\circ(\cdot)$ is defined as in Theorem 2.1.

Proof. The Lipschitz estimates and the regularity in the sense of Definition 1.8 can be derived using Kantorovich duality and the results of Proposition 2.1. For the first order

variations, one can simply apply a classical differentiation under the integral sign result for the first one and Proposition 2.1 to the components $\mu \mapsto \int_{\mathbb{R}^d} H^i(x, y) d\mu(y)$ for the second one. \square

We summarize these results in the form of an overview of possible functions satisfying **(H)** in the following corollaries.

Corollary 2.1 (Example of terminal costs satisfying the hypotheses of Theorem 2.1). *If $\varphi : \mathcal{P}(K) \mapsto \mathbb{R}$ is either a (suitable) power of a smooth integral functional or the variance functional, then it satisfies hypotheses **(H3)**, **(H5)** and **(H6)**.*

Corollary 2.2 (Example of running costs satisfying the hypotheses of Theorem 2.1). *Let $l : (x, v) \in \mathbb{R}^{2d} \mapsto l(x, v) \in \mathbb{R}$ be a C^1 , function $K \subset \mathbb{R}^d$ be compact, and define the running cost $L : (\mu, \omega) \in \mathcal{P}(K) \times U \mapsto \mathbb{R}$ by*

$$L(\mu, \omega) = \int_{\mathbb{R}^d} l(x, \omega(x)) d\mu(x).$$

*Then, $L(\cdot, \cdot)$ satisfies hypotheses **(H2)**, **(H5)** and **(H6)**. Moreover, the barycenter of the minimal selection in its extended subdifferential $\partial_\mu^\circ L(\mu, \omega)$ is determined at any $\mu \in \mathcal{P}(K)$ by*

$$\bar{\gamma}_L^\circ \equiv \nabla_\mu L(\mu, \omega) : x \in \text{supp}(\mu) \mapsto \nabla_x l(x, \omega(x)) + D_x \omega(x)^\top \nabla_v l(x, \omega(x)).$$

Proof. The proof only involves elementary Lipschitz-type estimates and the use of Proposition 2.1. \square

Notice again that it is possible to take any power $\alpha \geq 1$ of the previous cost and any power $\alpha > 0$ provided that the functional does not vanish along the optimal pair control-trajectory $(u^*(\cdot), \mu^*(\cdot)) \in \text{Lip}([0, T], \mathcal{P}_c(\mathbb{R}^d)) \times \mathcal{U}$.

The following result shows that functionals based on kernels are regular. They appear in several mean-field models for interaction, see e.g. [25, 26, 61, 82, 111, 129, 134].

Corollary 2.3 (Non-local vector field satisfying hypotheses **(H)**). *If $v[\cdot](\cdot, \cdot) : \mathcal{P}(K) \times [0, T] \times K \rightarrow \mathbb{R}^d$ is defined for any (μ, t, x) by*

$$v[\mu](t, x) = (H(t, \cdot) \star \mu(t))(x) + v_l(t, x),$$

*for some sublinear interaction kernel $H \in L^\infty([0, T], C^1(\mathbb{R}^d, \mathbb{R}^d))$ and vector field $v_l(\cdot, \cdot)$ measurable in t as well as sublinear and Lipschitz in x , then it satisfies hypotheses **(H4)**, **(H5)** and **(H3)**.*

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

In this chapter, we further the line of research presented in Chapter 2 by extending the Wasserstein PMP of Theorem 2.1 to constrained problems. We will therefore study constrained optimal control problems in Wasserstein spaces, given in the general form

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

Compared to problem (\mathcal{P}) , we added here the maps $(t, \mu) \mapsto \Lambda(t, \mu)$ and $\mu \mapsto \Psi^I(\mu), \Psi^E(\mu)$ which are running and end-point constraints respectively, written in the form of functional equalities and inequalities.

Theorem 3.1 (Pontryagin Maximum Principle for $(\mathcal{P}_{\text{const}})$). *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 for $(\mathcal{P}_{\text{const}})$ 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 $\{\omega \in C^1(\mathbb{R}^d, \mathbb{R}^d) \text{ s.t. } \|\omega(\cdot)\|_{C^0(\mathbb{R}^d, \mathbb{R}^d)} + \text{Lip}(\omega(\cdot), \mathbb{R}^d) \leq L_U\}$ for a given constant $L_U > 0$.*

(H2) *The non-local velocity field $\mu \in \mathcal{P}_c(\mathbb{R}^d) \mapsto v[\mu] \in L^1([0, T], C^1(\mathbb{R}^d, \mathbb{R}^d))$ satisfies the classical Cauchy-Lipschitz assumptions in Wasserstein spaces, i.e. there exists positive constants M, L_1 and L_2 such that*

$$|v[\mu](t, x)| \leq M(1 + |x|) \text{ , } |v[\mu](t, x) - v[\mu](t, y)| \leq L_1|x - y|, \\ \|v[\mu](t, \cdot) - v[\nu](t, \cdot)\|_{C^0(\mathbb{R}^d)} \leq L_2W_1(\mu, \nu)$$

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

- (H3)** The maps $\mu \in \mathcal{P}_c(\mathbb{R}^d) \mapsto v^i[\mu](t, x)$ are regular in the sense of Definition 1.8 and Wasserstein-differentiable at $\mu^*(t)$ for all $i \in \{1, \dots, d\}$, uniformly with respect to $(t, x) \in [0, T] \times \mathbb{R}^d$. The maps $\mu \mapsto D_x v[\mu](t, x)$ and $(y, z) \mapsto \nabla_\mu (v^i[\cdot](t, y))(\mu)(z)$ are continuous for all $i \in \{1, \dots, d\}$, uniformly with respect to $(t, x) \in [0, T] \times \mathbb{R}^d$.
- (H4)** The final cost $\mu \mapsto \varphi(\mu)$ and the boundary constraints maps $\mu \mapsto (\Psi_i^I(\mu), \Psi_j^E(\mu))_{i,j} \in \mathbb{R}^{n+m}$ 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 regular in the sense of Definition 1.8 and Wasserstein-differentiable at $\mu^*(T)$. The 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 for any $(i, j) \in \{1, \dots, n\} \times \{1, \dots, m\}$.
- (H5)** The running cost $(t, \mu, \omega) \mapsto L(t, \mu, \omega) \in \mathbb{R}$ is \mathcal{L}^1 -measurable with respect to $t \in [0, T]$, bounded, 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 regular in the sense of Definition 1.8 and Wasserstein-differentiable at $\mu^*(t)$ for any $(t, \omega) \in [0, T] \times U$, and its Wasserstein gradient $\nabla_\mu L(t, \mu^*(t), \omega)(\cdot)$ is continuous.
- (H6)** The state constraints maps $(t, \mu) \mapsto (\Lambda_l(t, \mu))_l \in \mathbb{R}^r$ are bounded and Lipschitz-continuous over $[0, T] \times \mathcal{P}(K)$ for any compact set $K \subset \mathbb{R}^d$. Moreover, the maps $(t, \mu) \mapsto \partial_t \Lambda_l(t, \mu)$ and $(t, \mu) \mapsto \nabla_\mu \Lambda_l(t, \mu)(\cdot)$ are well-defined and continuously differentiable at $(t, \mu^*(t))$ with

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

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

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

$$\begin{cases} \partial_t \nu^*(t) + \nabla \cdot (\mathbb{J}_{2d} \nabla_\nu \mathcal{H}(t, \nu^*(t), \zeta^*(t), u^*(t)) \nu^*(t)) = 0 & \text{in } [0, T] \times \mathbb{R}^{2d}, \\ \pi_{\#}^1 \nu^*(t) = \mu^*(t), & \text{for all } t \in [0, T], \\ \pi_{\#}^2 \nu^*(T) = (-\nabla_\mu \mathcal{S}(\mu^*(T)))_{\#} \mu^*(T), \end{cases} \quad (3.1)$$

where \mathbb{J}_{2d} is the symplectic matrix of \mathbb{R}^{2d} . The augmented infinite-dimensional

Hamiltonian $\mathcal{H}(\cdot, \cdot, \cdot, \cdot)$ of the system is defined by

$$\mathcal{H}(t, \nu, \zeta, \omega) = \int_{\mathbb{R}^{2d}} \langle r, v[\pi_{\#}^1 \nu](t, x) + \omega(x) \rangle d\nu(x, r) - L(t, \pi_{\#}^1 \nu, \omega) - \mathcal{C}(t, \pi_{\#}^1 \nu, \zeta), \quad (3.2)$$

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

$$\mathcal{C}(t, \mu, \zeta, \omega) = \sum_{l=1}^r \zeta_l \left(\partial_t \Lambda_l(t, \mu) + \int_{\mathbb{R}^d} \langle \nabla_{\mu} \Lambda_l(t, \mu)(x), v[\mu](t, x) + \omega(x) \rangle d\mu(x) \right), \quad (3.3)$$

and

$$\nabla_{\mu} \mathcal{S}(\mu) = \lambda_0 \nabla_{\mu} \varphi(\mu) + \sum_{i=1}^n \lambda_i \nabla_{\mu} \Psi_i^I(\mu) + \sum_{j=1}^m \eta_j \nabla_{\mu} \Psi_j^E(\mu). \quad (3.4)$$

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

$$(\lambda_0, \dots, \lambda_n, \eta_1, \dots, \eta_m, \varpi_1, \dots, \varpi_r) \neq 0. \quad (3.5)$$

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

$$\mathcal{H}(t, \nu^*(t), \zeta^*(t), u^*(t)) = \max_{\omega \in U} [\mathcal{H}(t, \nu^*(t), \zeta^*(t), \omega)] \quad (3.6)$$

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

Remark 3.1 (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. [133, Chapter 9]).*

As already mentioned in the Introduction, this fact is crucial for our purpose since absolutely continuous curves in Wasserstein spaces are exactly those curves which solve a continuity equation. Hence, one cannot derive a Hamiltonian system such as (3.1) by sticking to the classical formulation of the PMP.

Remark 3.2 (On the regularity hypothesis **(H1)**). *The considerations and comments of Remark 2.1 concerning the regularity of the admissible controls still apply in the context*

of Theorem 3.1.

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 recover the maximum principle from a family of finite dimensional first-order optimality conditions by means of the introduction of a suitable costate. Even though classical in the unconstrained case, this direct approach does require 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 goes to infinity. Originally introduced in [22] for smooth optimal control problems with end-point constraints, this approach was extended in [126] to the case of non-smooth and state-constrained problems. When trying to further adapt this method to the setting of Wasserstein spaces, one is faced with an extra structural difficulty. In the classical statement of the maximum principle, the presence of state constraints implies a mere BV regularity in time for the covectors. However, a deep structural result in optimal transport theory which we recalled in Theorem 1.4 states that solutions of continuity equations in Wasserstein spaces coincide exactly with absolutely continuous curves. Whence, in order to write a well-defined Wasserstein Hamiltonian flow in the spirit of [15, 35], we choose to formulate a maximum principle in the so-called Gamkrelidze form (see e.g. [21]), which allows to recover a stronger absolute continuity of the costates at the price of an extra regularity assumption on the state constraints.

The structure of the chapter is the following. In Section 3.1, we introduce several tools of non-smooth analysis including a general Lagrange multiplier rule formulated in terms of *Michel-Penot* subdifferentials. In Section 3.2, we state and prove our main result, that is Theorem 3.1. The argument is again split into four steps which loosely follow the methodology already introduced in Chapter 2. In Section 3.3, we provide the analytical expression of the Wasserstein gradient of a functional involved in the statement of Theorem 3.1. We then exhibit in Section 3.4 a series of examples of functionals satisfying the structural assumptions **(H)** of Theorem 3.1.

3.1 Non-smooth multiplier rule and differentiable extension of functions

In this section, we recall some facts of non-smooth analysis, for which we refer the reader to the excellent monographs [59, 133], 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. [113, 94]. In the sequel, we denote by $(X, \|\cdot\|_X)$ a separable Banach space and by X^* its topological dual associated with the duality bracket $\langle \cdot, \cdot \rangle_X$. Given a map $f : X \rightarrow \mathbb{R}$, we denote by $D(f) = \{x \in X \text{ s.t. } f(x) < +\infty\}$ its effective domain.

Definition 3.1 (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

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

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 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 3.1 (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*

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

- (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*

$$d_{\text{MP}}(f \circ \mathcal{G})(0; \sigma) = \langle \xi, D\mathcal{G}(0)\sigma \rangle_X, \tag{3.7}$$

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

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. [126]. 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 [94] and refined in [126] where the author extended the result to the class of so-called *calm* functions.

Definition 3.2 (Calm functions). *A map $f : X \rightarrow \mathbb{R}$ is said to be calm at $x \in X$ provided that*

- (i) *There exists a constant $L > 0$ such that for any $\delta \in X$ with $\|\delta\|_X$ sufficiently small, it holds 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 3.2 (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

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

where $\mathcal{N}(\Omega, x_*)$ denotes the normal cone of convex analysis to Ω 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 to be useful in the sequel.

Lemma 3.1 (Pointwise convergence and restriction). *Let $f : [0, T] \rightarrow X$ be an L^1 -function in the sense of Bochner (see e.g. [69]) 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*

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

Proof. This result is a consequence of Lusin's Theorem applied to vector valued maps (see e.g. [69]). \square

Lemma 3.2 (A continuous extension preserving the derivative). *Let $\epsilon > 0$ and $f : [0, \epsilon]^N \rightarrow X$ be a continuous map differentiable at $e = 0$ relatively to \mathbb{R}_+^N . Then, there exists a continuous extension $\tilde{f} : [-\bar{\epsilon}_N, \bar{\epsilon}_N]^N \rightarrow X$ of $f(\cdot)$ which is Fréchet differentiable at $e = 0$ and 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 [126, Lemma 2.11]. Define the map

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

By definition, $g(\cdot)$ is continuous over $\mathbb{R}_+^N \setminus \{0\}$ and can be extended to \mathbb{R}_+^N by imposing that $g(0) = 0$ since $f(\cdot)$ is differentiable at $e = 0$ relatively to \mathbb{R}_+^N . Invoking Dugundji's extension theorem (see [72]), 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 map $\tilde{f}(\cdot)$ is Fréchet-differentiable at $e = 0$ with $D_e \tilde{f}(0) = D_e f(0)$. \square

3.2 Proof of Theorem 3.1

In this section we prove the main result of this article, that is Theorem 3.1. 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 the non-smooth Lagrange multiplier rule of Theorem 3.2 to the sequence of finite-dimensional optimization problem formulated on the length of the needle variations, to obtain a family of finite-dimensional optimality conditions at time T . We introduce in Step 3 a suitable notion of costate, allowing to propagate this family of 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 several approximation arguments.

Step 1 : Packages of needle-like variations

We start by considering an optimal pair control-trajectory $(u^*(\cdot), \mu^*(\cdot)) \in \mathcal{U} \times \text{Lip}([0, T], \mathcal{P}(\overline{B(0, R_T)}))$ where $R_T > 0$ is given by Theorem 1.7. Let $\mathcal{T} \subset [0, T]$ be the set of Lebesgue points of $t \in [0, T] \mapsto (v[\mu^*(t)](t, \cdot), u^*(t, \cdot), L(t, \mu^*(t), \cdot)) \in C^0(\overline{B(0, R_T)}, \mathbb{R}^d) \times U \times C^0(U, \mathbb{R})$ in the sense of Bochner's integral (see e.g. [69, Chapter II - Theorem 9]). This set has full Lebesgue measure in $[0, T]$, and Lemma 3.1 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 τ and such that

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

We further denote by U_D a countable and dense subset of the set of admissible control values U which is compact and separable in the C^0 -topology as a consequence of **(H1)**.

Definition 3.3 (Package of needle 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. Given $N \geq 1$, a family of elements $\{(\omega_k, \tau_k)\}_{k=1}^N \subset U_D \times \mathcal{A}$ and $e = (e_1, \dots, e_N) \in [0, \bar{e}_N]^N$ 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*

$$\tilde{u}_e^N \equiv \tilde{u}_e : t \mapsto \begin{cases} \omega_k & \text{if } t \in [\tau_k - e_k, \tau_k], \\ u^*(t) & \text{otherwise.} \end{cases} \quad (3.8)$$

We also denote by $t \mapsto \tilde{\mu}_e(\cdot)$ the corresponding perturbed trajectory, i.e. the solution of (1.32) associated with the controlled non-local velocity field $v[\cdot](\cdot, \cdot) + \tilde{u}_e(\cdot, \cdot)$.

The reader can check that Definition 3.3 of a package of needle-like variation is a direct generalization of that given in (2.11) in Chapter 2. In the following lemma, we make use of the geometric structure of solutions to non-local transport equations presented in Theorem 1.7 together with some notations borrowed from Proposition 1.11 to express $\tilde{\mu}_e(t)$ as a function of $\mu^*(t)$ for all times $t \in [0, T]$. In the sequel, we denote by $\Phi_{(s,t)}^{v,u}[\mu(s)](\cdot)$ the flow map generated by the non-local velocity field $v[\cdot](\cdot, \cdot) + u(\cdot, \cdot)$ between times s and t , as defined by (1.33) in Theorem 1.7.

Lemma 3.3 (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*

$$\tilde{\mu}_e(t) = \mathcal{G}_t^N(e, \cdot) \# \mu^*(t). \quad (3.9)$$

(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 $(t, e) \in [0, T] \times [-\bar{\epsilon}_N, \bar{\epsilon}_N]^N$ 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 $C^0(\overline{B(0, R_T)}, \mathbb{R}^d)$ -norm uniformly with respect to $t \in [0, T]$. The corresponding Taylor expansion can be written explicitly as*

$$\mathcal{G}_t^N(e, \cdot) = \text{Id} + \sum_{k=1}^{\iota(t)} e_k \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}[\mu^*(t)](\cdot) + o(|e|), \quad (3.10)$$

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

$$\begin{cases} \partial_t \mathcal{F}_t^{\omega_k, \tau_k}(x) = \left[D_x u^* \left(t, \Phi_{(\tau_k, t)}^{v, u^*}(x) \right) + D_x v[\mu^*(t)] \left(t, \Phi_{(\tau_k, t)}^{v, u^*}(x) \right) \right] \mathcal{F}_t^{\omega_k, \tau_k}(x) \\ \quad + \int_{\mathbb{R}^d} \mathbb{I}^v \left(t, \Phi_{(\tau_k, t)}^{v, u^*}(x) \right) \left(\Phi_{(\tau_k, t)}^{v, u^*}(y) \right) \mathcal{F}_t^{\omega_k, \tau_k}(y) d\mu^*(\tau_k)(y), \\ \mathcal{F}_{\tau_k}^{\omega_k, \tau_k}(x) = \omega_k(x) - u^*(\tau_k, x). \end{cases} \quad (3.11)$$

Proof. The proof of this result is similar to that of Lemma 2.5 in Chapter 2, 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 $(\Phi_{(0,t)}^{v, \tilde{u}_e}[\mu^0](\cdot))_{t \in [0, T]}$ which are 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 $\mathcal{G}_t^N(e, \cdot) = \Phi_{(0,t)}^{\tilde{u}_e, v}[\mu^0] \circ \Phi_{(t,0)}^{u^*, v}[\mu^*(t)](\cdot)$.

We focus our attention on the proof by induction of (iv). Let $t \in [0, T]$ be such that $\iota(t) = 1$. By (3.8), one has that

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

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

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

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 1.11, 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 **(H1)**-**(H2)**. 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 invoking the statements of Proposition 1.10 and Theorem 1.9, we have that both (3.10) and (3.11) hold for any $e_1 \in [0, \bar{\epsilon}_N]$ and all times $t \in [0, T]$ such that $\iota(t) = 1$.

Let us now assume that (3.10) and (3.11) hold for all times $t \in [0, T]$ such that $\iota(t) = k - 1$, i.e.

$$\tilde{\mu}_e(t) = \mathcal{G}_t^N(e, \cdot) \# \mu^*(t) = \left(\text{Id} + \sum_{l=1}^{k-1} e_l \mathcal{F}_t^{\omega_l, \tau_k} \circ \Phi_{(t, \tau_l)}^{v, u^*} [\mu^*(t)](\cdot) + o(|e|) \right) \# \mu^*(t), \quad (3.12)$$

for $e \in [0, \bar{\epsilon}_N]^N$. By definition (3.8) of an N -package of needle-like variations, 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

$$\Phi_{(\tau_k - e_k, \tau_k)}^{v, \omega_k} [\tilde{\mu}_e(\tau_k - e_k)] \circ \Phi_{(\tau_k, \tau_k - e_k)}^{v, u^*} [\tilde{\mu}_e(\tau_k)](x) = x + e_k [\omega_k(x) - u^*(\tau_k, x)] + o(e_k). \quad (3.13)$$

Furthermore, invoking the induction hypothesis (3.12) and the results of Proposition 1.8, we obtain that

$$\begin{aligned} & \Phi_{(\tau_{k-1}, \tau_k)}^{v, u^*} [\tilde{\mu}_e(\tau_{k-1})] \left(x + \sum_{l=1}^{k-1} \left(e_l \mathcal{F}_{\tau_{k-1}}^{\omega_l, \tau_l} \circ \Phi_{(\tau_{k-1}, \tau_l)}^{v, u^*} (x) + o(e_l) \right) \right) \\ &= \Phi_{(\tau_{k-1}, \tau_k)}^{v, u^*} [\tilde{\mu}_e(\tau_{k-1})](x) + \sum_{l=1}^{k-1} \left(e_l D_x \Phi_{(\tau_{k-1}, \tau_k)}^{v, u^*} [\tilde{\mu}_e(\tau_{k-1})](x) \mathcal{F}_{\tau_{k-1}}^{\omega_l, \tau_l} \circ \Phi_{(\tau_{k-1}, \tau_l)}^{v, u^*} (x) + o(e_l) \right) \\ &= \Phi_{(\tau_{k-1}, \tau_k)}^{v, u^*} [\mu^*(\tau_{k-1})](x) + \sum_{l=1}^{k-1} e_l \left(D_x \Phi_{(\tau_{k-1}, \tau_k)}^{v, u^*} [\mu^*(\tau_{k-1})](x) \mathcal{F}_{\tau_{k-1}}^{\omega_l, \tau_l} \circ \Phi_{(\tau_{k-1}, \tau_l)}^{v, u^*} (x) \right. \\ & \qquad \qquad \qquad \left. + w_l(\tau_k, x) + o(e_l) \right) \end{aligned} \quad (3.14)$$

where the maps $(w_l(\cdot, \cdot))_{1 \leq l \leq k-1}$ are defined as in Proposition 1.11 with $\mathcal{F}_l(\cdot) \equiv \mathcal{F}_{\tau_{k-1}}^{\omega_l, \tau_l} \circ \Phi_{(\tau_{k-1}, \tau_l)}^{u^*} [\mu^*(\tau_{k-1})](\cdot)$. Plugging together equation (1.35) of Proposition 1.10 and equation (1.37) of Theorem 1.9, one can see that the maps

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

are solutions of (3.11) on $[\tau_{k-1}, \tau_k]$ with initial condition $\mathcal{F}_{\tau_{k-1}}^{\omega_l, \tau_l}(\cdot)$ at time τ_{k-1} for any $l \in \{1, \dots, k-1\}$. By Cauchy-Lipschitz uniqueness, we can therefore extend the definition of the maps $t \mapsto \mathcal{F}_t^{\omega_l, \tau_l}(x)$ to the whole of $[\tau_l, \tau_k]$ for any $l \in \{1, \dots, k-1\}$.

Chaining the expansions (3.13) and (3.14) along with our previous extension argument, we obtain that both (3.10) and (3.11) hold up to time τ_k , i.e.

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

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 extension 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 3.2, we can now extend the map $e \in [0, \bar{\epsilon}_N]^N \mapsto \mathcal{G}_t^N(e, \cdot) \in C^0(\overline{B(0, R_T)}, \mathbb{R}^d)$ 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 3.3. \square

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

Step 2 : First-order optimality condition

In Lemma 3.3, 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

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

admits $e = 0$ as an optimal solution in $[0, \bar{\epsilon}_N]^N$ for $\bar{\epsilon}_N$ small enough.

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 3.2. We also compute their first-order variation induced by the package of needle-like variations at $e = 0$.

Lemma 3.4 (Differentiability and calmness of the functionals involved in (\mathcal{P}_N)). *The maps $e \in [-\bar{\epsilon}_N, \bar{\epsilon}_N]^N \mapsto \varphi(\tilde{\mu}_e(T))$, $\Psi^E(\tilde{\mu}_e(T))$, $\Psi^I(\tilde{\mu}_e(T))$ 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 Fréchet differentiable at $e = 0$. Their Fréchet derivative in a direction $\sigma \in [0, \bar{\epsilon}_N]^N$ are respectively given by*

$$\begin{cases} \nabla_e(\varphi(\tilde{\mu}_e(T)))|_{e=0}(\sigma) = \sum_{k=1}^N \sigma_k \int_{\mathbb{R}^d} \langle \nabla_\mu \varphi(\mu^*(T))(x), \mathcal{F}_T^{\omega_k, \tau_k} \circ \Phi_{(T, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(T)(x), \\ \nabla_e(\Psi_i^I(\tilde{\mu}_e(T)))|_{e=0}(\sigma) = \sum_{k=1}^N \sigma_k \int_{\mathbb{R}^d} \langle \nabla_\mu \Psi_i^I(\mu^*(T))(x), \mathcal{F}_T^{\omega_k, \tau_k} \circ \Phi_{(T, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(T)(x), \\ \nabla_e(\Psi_j^E(\tilde{\mu}_e(T)))|_{e=0}(\sigma) = \sum_{k=1}^N \sigma_k \int_{\mathbb{R}^d} \langle \nabla_\mu \Psi_j^E(\mu^*(T))(x), \mathcal{F}_T^{\omega_k, \tau_k} \circ \Phi_{(T, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(T)(x), \end{cases} \quad (3.15)$$

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

$$\begin{aligned} \nabla_e \left(\int_0^T L(t, \tilde{\mu}_e(t), \tilde{u}_e(t)) dt \right) \Big|_{e=0} (\sigma) &= \sum_{k=1}^N \sigma_k \left(L(\tau_k, \mu^*(\tau_k), u^*(\tau_k)) - L(\tau_k, \mu^*(\tau_k), \omega_k) \right) \\ &+ \sum_{k=1}^N \sigma_k \int_{\tau_k}^T \int_{\mathbb{R}^d} \left\langle \nabla_{\mu} L(t, \mu^*(t), u^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \right\rangle d\mu^*(t)(x) dt \end{aligned} \quad (3.16)$$

The maps $e \in [-\bar{\epsilon}_N, \bar{\epsilon}_N]^N \mapsto \max_{t \in [0, T]} \Lambda_1(t, \tilde{\mu}_e(t)), \dots, \max_{t \in [0, T]} \Lambda_r(t, \tilde{\mu}_e(t))$ 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

$$\begin{aligned} &d_{\text{MP}} \left(\max_{t \in [0, T]} \Lambda_l(t, \tilde{\mu}_e(t)) \right) (0; \sigma) \\ &= \sum_{k=1}^N \sigma_k \int_0^T \int_{\mathbb{R}^d} \left\langle \nabla_{\mu} \Lambda_l(t, \mu^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \right\rangle d\mu^*(t) d\varpi_l^N(t), \end{aligned} \quad (3.17)$$

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

Proof. The calmness property of the maps $e \mapsto \varphi(\tilde{\mu}_e(T))$, $\Psi^E(\tilde{\mu}_e(T))$, $\Psi^I(\tilde{\mu}_e(T))$ 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 3.3. The differentials of the final cost and boundary constraints can be computed with a direct application of Proposition 1.8.

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 (see e.g. [14, Corollary 2.23]), since τ_k is a Lebesgue points of $t \in [0, T] \mapsto L(t, \mu^*(t), \cdot)$ and 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 chainrule of Proposition 1.8 along with the results of Lemma 3.3, we have

that

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

for \mathcal{L}^1 -almost every $t \in [0, T]$. Combining these expansions with an application of Lebesgue's Dominated Convergence Theorem yields (3.16).

We now turn our attention to the state constraints functionals. By hypothesis **(H6)** and Proposition 1.8, 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

$$\gamma \in C^0([0, T], \mathbb{R}) \mapsto \max_{t \in [0, T]} \gamma(t)$$

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 3.1-(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. [95, Section 4.5.1]) as the set of Borel regular measures such that

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

Invoking again Proposition 1.8 and Lemma 3.3, we can write the differential of $e \mapsto \Lambda_l(t, \tilde{\mu}_e(t))$ at $e = 0$ evaluated in a direction $\sigma \in [0, \bar{\epsilon}_N]^N$ as

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

for all times $t \in [0, T]$. Using the measure-theoretic version of Lebesgue's Dominated Convergence Theorem (see e.g. [14, Theorem 1.21]), we conclude that (3.17) holds as well, which ends the proof of Lemma 3.4. \square

Using the results of Lemma 3.4, we can apply the Lagrange multiplier rule of Theorem 3.2 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

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

Since all the functions involved in the subdifferential inclusion (S) are calm, we can use the summation rule of Proposition 3.1-(b) along with the characterization of MP-subdifferentials for Fréchet differentiable stated in Proposition 3.1-(a) to obtain that

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

By combining the expressions of the gradients (3.15), (3.16) and the MP-derivative (3.17) derived in Lemma 3.4, along with the composition rule of Proposition 3.1-(c) for MP-subdifferentials, we obtain that

$$\begin{aligned} & \sum_{k=1}^N \sigma_k \int_{\mathbb{R}^d} \langle -\nabla_{\mu} \mathcal{S}_N(\mu^*(T))(x), \mathcal{F}_T^{\omega_k, \tau_k} \circ \Phi_{(T, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(T)(x) \\ & - \sum_{k=1}^N \sigma_k \lambda_0^N \left(L(\tau_k, \mu^*(\tau_k), u^*(\tau_k)) - L(\tau_k, \mu^*(\tau_k), \omega_k) \right) \\ & - \sum_{k=1}^N \sigma_k \int_{\tau_k}^T \int_{\mathbb{R}^d} \langle \lambda_0^N \nabla_{\mu} L(t, \mu^*(t), u^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(t)(x) dt \\ & - \sum_{k=1}^N \sigma_k \sum_{l=1}^r \theta_l^N \int_{\tau_k}^T \int_{\mathbb{R}^d} \langle \nabla_{\mu} \Lambda_l(t, \mu^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(t)(x) d\varpi_l^N(t) \leq 0 \end{aligned}$$

for any direction $\sigma \in [0, \bar{\epsilon}_N]^N$ where $\nabla_{\mu} \mathcal{S}_N(\mu^*(T))(\cdot)$ is defined as in (3.4). By choosing in particular vectors $\sigma \in [0, \bar{\epsilon}_N]^N$ which have all their components except one equal to 0,

this family of inequalities can be rewritten as

$$\begin{aligned}
& \int_{\mathbb{R}^d} \langle -\nabla_{\mu} \mathcal{S}_N(\mu^*(T))(x), \mathcal{F}_T^{\omega_k, \tau_k} \circ \Phi_{(T, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(T)(x) \\
& - \lambda_0^N \left(L(\tau_k, \mu^*(\tau_k), u^*(\tau_k)) - L(\tau_k, \mu^*(\tau_k), \omega_k) \right) \\
& - \int_{\tau_k}^T \int_{\mathbb{R}^d} \langle \lambda_0^N \nabla_{\mu} L(t, \mu^*(t), u^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(t)(x) dt \\
& - \sum_{l=1}^r \int_{\tau_k}^T \int_{\mathbb{R}^d} \langle \nabla_{\mu} \Lambda_l(t, \mu^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(t)(x) d\varpi_l^N(t) \leq 0
\end{aligned} \tag{3.18}$$

for all $k \in \{1, \dots, N\}$, where we redefined the notation $\varpi_l^N \equiv \theta_l^N \varpi_l^N$.

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 (3.18) derived at time T to the base points (τ_1, \dots, τ_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

$$\begin{cases} \partial_t \nu_N^*(t) + \nabla \cdot (\mathcal{V}_N^*[\nu_N^*(t)](t, \cdot, \cdot) \nu_N^*(t)) = 0 & \text{in } [0, T] \times \mathbb{R}^{2d}, \\ \pi_{\#}^1 \nu_N^*(t) = \mu^*(t) & \text{for all } t \in [0, T], \\ \nu_N^*(T) = (\text{Id} \times (-\nabla_{\mu} \mathcal{S}_N(\mu^*(T))))_{\#} \mu^*(T). \end{cases} \tag{3.19}$$

Here, 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 $(x, r, \nu) \in \mathbb{R}^{2d} \times \mathcal{P}_c(\mathbb{R}^{2d})$ by

$$\mathcal{V}_N^*[\nu](t, x, r) = \left(\begin{array}{c} v[\pi_{\#}^1 \nu](t, x) + u^*(t, x) \\ \lambda_0^N \nabla_{\mu} L(t, \pi_{\#}^1 \nu, u^*(t))(x) + \nabla_{\mu} \mathcal{C}(t, \pi_{\#}^1 \nu, \zeta_N^*(t), u^*(t))(x) \\ -\mathbb{I}^v[\nu](t, x) - \text{D}_x v[\pi_{\#}^1 \nu](t, x)^{\top} r - \text{D}_x u^*(t, x)^{\top} r \end{array} \right),$$

where we introduce the notation

$$\mathbb{I}^v[\nu](t, x) = \int_{\mathbb{R}^{2d}} \left(\mathbb{I}_{(t, y)}^v(x) \right)^{\top} p d\nu(y, p). \tag{3.20}$$

Notice that the transport equation (3.19) does not satisfy the classical hypotheses of Theorem 1.7. Following the methodology introduced in Lemma 2.3 and 2.6 of Chapter 2, it is possible to circumvent this difficulty by building explicitly a solution of (3.19) relying on the *cascade structure* of the equations.

Lemma 3.5 (Definition and well-posedness of solutions of (3.19)). *Let $(u^*(\cdot), \mu^*(\cdot))$ be an optimal pair control-trajectory for $(\mathcal{P}_{\text{const}})$. 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

$$\left\{ \begin{array}{l} \partial_t w_x(t, r) = \lambda_0^N \nabla_\mu L(t, \mu^*(t), u^*(t)) \left(\Phi_{(T,t)}^{v, u^*}(x) \right) + \nabla_\mu \mathcal{C}(t, \mu^*(t), \zeta_N^*(t), u^*(t)) \left(\Phi_{(T,t)}^{v, u^*}(x) \right) \\ \quad - D_x u^* \left(t, \Phi_{(T,t)}^{v, u^*}(x) \right)^\top w_x(t, r) - D_x v[\mu^*(t)] \left(t, \Phi_{(T,t)}^{v, u^*}(x) \right)^\top w_x(t, r) \\ \quad - \int_{\mathbb{R}^{2d}} \mathbb{I}^v \left(t, \Phi_{(T,t)}^{v, u^*}(y) \right) \left(\Phi_{(T,t)}^{v, u^*}(x) \right)^\top w_y(t, p) d((\text{Id} \times (-\nabla_\mu \mathcal{S}_N(\mu^*(T))))_{\#} \mu^*(T))(y, p) \\ w_x(T, r) = r, \end{array} \right. \quad (3.21)$$

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 $\nu_N^* : t \in [0, T] \mapsto (\Phi_{(T,t)}^{u^*}(\cdot), \text{Id})_{\#} \nu_{T,N}^*(t)$ is the unique solution of (3.19). 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

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

Such a set exists since the maps $(\nabla_\mu \mathcal{S}_N(\mu^*(T)))(\cdot)$ are continuous by **(H4)**, as well as uniformly bounded as a consequence of the non-triviality condition (NT) on the Lagrange multipliers $(\lambda_1^N, \dots, \lambda_n^N, \eta_1^N, \dots, \eta_m^N)$.

The existence and uniqueness of the maps $(t, x, r) \mapsto w_x(t, r)$ solving the family of non-local Cauchy problems (3.21) can be obtained under hypotheses **(H)**, as a consequence of Banach fixed point Theorem in the spirit of Theorem 1.9. In this context, the Banach space under consideration is that of all maps $w : [0, T] \times \Omega \rightarrow \mathbb{R}^d$ endowed with the norm

$$\|w(\cdot, \cdot)\|_* = \inf \left\{ M > 0 \text{ s.t. } \|w_x(\cdot, \cdot)\|_{C^0([0,T] \times \mathbb{R}^d)} \leq M \text{ for } \mu^*(T)\text{-almost every } x \in \mathbb{R}^d \right\}.$$

By an application of Grönwall's Lemma to (3.21), it holds that $(t, r) \in [0, T] \times \pi^2(\Omega) \mapsto \Psi_{(T,t)}^{x,N}(r)$ is bounded by a positive constant, uniformly with respect to $x \in \text{supp}(\mu^*(T))$ and $N \geq 1$. This follows in particular from the uniform boundedness of the sequences of multipliers (λ_N^0) and $(\zeta_N^*(\cdot))$. Therefore, there exists a uniform constant $R'_T > 0$ such that

$$\text{supp}(\nu_N^*(t)) \subset \overline{B_{2d}(0, R'_T)},$$

for all $t \in [0, T]$. This in turn implies that the right-hand side of (3.21) is uniformly bounded, so that the maps $t \in [0, T] \mapsto \Psi_{(T,t)}^{x,N}(r)$ are Lipschitz, uniformly with respect to $(x, r) \in \Omega$ and $N \geq 1$. By applying again Grönwall's Lemma to the difference $|\Psi_{(T,t)}^{x,N}(r_2) - \Psi_{(T,t)}^{x,N}(r_1)|$ with $r_1, r_2 \in \pi^2(\Omega)$, we further obtain that $(t, r) \in [0, T] \times \pi^2(\Omega) \mapsto \Psi_{(T,t)}^{x,N}(r)$ is also Lipschitz regular, uniformly with respect to $x \in \text{supp}(\mu^*(T))$

and $N \geq 1$. It can be checked by leveraging Kantorovich duality in the spirit of Lemma 2.6 that this in turn yields the Lipschitz regularity of $t \in [0, T] \mapsto \sigma_{x,N}^*(t)$ uniformly with respect to $x \in \text{supp}(\mu^*(T))$ and $N \geq 1$. An application of Proposition 1.4 to $\nu_N^*(\cdot)$ combined with the uniform Lipschitz regularity of $(t, x) \in [0, T] \times \pi^1(\Omega) \mapsto \Phi_{(T,t)}^{v,u^*}(x)$ provides the existence of a uniform constant $L'_T > 0$ such that

$$W_1(\nu_N^*(t), \nu_N^*(s)) \leq L'_T |t - s|,$$

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

In order to prove that $\nu_N^*(\cdot)$ is indeed a solution of (3.19), take $\xi \in C_c^\infty(\mathbb{R}^{2d})$ and compute the time derivative

$$\begin{aligned} & \frac{d}{dt} \int_{\mathbb{R}^{2d}} \xi(x, r) d\nu_N^*(t)(x, r) \\ &= \frac{d}{dt} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \xi \left(\Phi_{(T,t)}^{v,u^*}(x), r \right) d\sigma_{x,N}^*(t)(r) d\mu^*(T)(x) \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left\langle \nabla_x \xi \left(\Phi_{(T,t)}^{v,u^*}(x), r \right), v[\mu^*(t)] \left(t, \Phi_{(T,t)}^{v,u^*}(x) \right) \right\rangle d\sigma_{x,N}^*(t)(r) d\mu^*(T)(x) \\ &+ \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left\langle \nabla_x \xi \left(\Phi_{(T,t)}^{v,u^*}(x), r \right), u^* \left(t, \Phi_{(T,t)}^{v,u^*}(x) \right) \right\rangle d\sigma_{x,N}^*(t)(r) d\mu^*(T)(x) \\ &+ \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left\langle \nabla_r \xi \left(\Phi_{(T,t)}^{v,u^*}(x), r \right), \lambda_0^N \nabla_\mu L(t, \mu^*(t), u^*(t)) \left(\Phi_{(T,t)}^{v,u^*}(x) \right) \right\rangle d\sigma_{x,N}^*(t)(r) d\mu^*(T)(x) \\ &+ \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left\langle \nabla_r \xi \left(\Phi_{(T,t)}^{v,u^*}(x), r \right), \nabla_\mu \mathcal{C}(t, \mu^*(t), \zeta_N^*(t), u^*(t)) \left(\Phi_{(T,t)}^{v,u^*}(x) \right) \right\rangle d\sigma_{x,N}^*(t)(r) d\mu^*(T)(x) \\ &- \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left\langle \nabla_r \xi \left(\Phi_{(T,t)}^{v,u^*}(x), r \right), D_x v[\mu^*(t)] \left(t, \Phi_{(T,t)}^{v,u^*}(x) \right)^\top r \right\rangle d\sigma_{x,N}^*(t)(r) d\mu^*(T)(x) \\ &- \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left\langle \nabla_r \xi \left(\Phi_{(T,t)}^{v,u^*}(x), r \right), D_x u^* \left(t, \Phi_{(T,t)}^{v,u^*}(x) \right)^\top r \right\rangle d\sigma_{x,N}^*(t)(r) d\mu^*(T)(x) \\ &- \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left\langle \nabla_r \xi \left(\Phi_{(T,t)}^{v,u^*}(x), r \right), \mathbb{I}^v[\nu_N^*(t)] \left(t, \Phi_{(T,t)}^{v,u^*}(x) \right) \right\rangle d\sigma_{x,N}^*(t)(r) d\mu^*(T)(x) \end{aligned}$$

where we used the fact that by Fubini's Theorem

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

This can in turn be reformulated into the more concise expression

$$\begin{aligned} & \frac{d}{dt} \int_{\mathbb{R}^{2d}} \xi(x, r) d\nu_N^*(t)(x, r) \\ &= \int_{\mathbb{R}^{2d}} \langle \nabla_x \xi(x, r), v[\mu^*(t)](t, x) + u^*(t, x) \rangle d\nu_N^*(t)(x, r) \end{aligned}$$

$$\begin{aligned}
& + \int_{\mathbb{R}^{2d}} \langle \nabla_r \xi(x, r), \lambda_0^N \nabla_\mu L(t, \mu^*(t), u^*(t))(x) + \nabla_\mu \mathcal{L}(t, \mu^*(t), \zeta_N^*(t), u^*(t))(x) \rangle d\nu_N^*(t)(x, r) \\
& - \int_{\mathbb{R}^{2d}} \langle \nabla_r \xi(x, r), D_x v[\mu^*(t)](t, x)^\top r + D_x u^*(t, x)^\top r + \mathbb{I}^v[\nu_N^*(t)](t, x) \rangle d\nu_N^*(t)(x, r),
\end{aligned}$$

which by (1.30) precisely corresponds to the fact that $\nu_N^*(\cdot)$ is a solution of (3.19). \square

Now that we have built a suitable notion of solution for (3.19), 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 3.6 (A partial Pontryagin Maximum Principle). *For any $N \geq 1$, the curve of measures $\nu_N^*(\cdot)$ introduced in Lemma 3.5 is a solution of the Hamiltonian flow (3.1) 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*

$$\begin{aligned}
& \mathcal{H}(\tau_k, \nu_N^*(\tau_k), \zeta_N^*(\tau_k), \omega_k) - \sum_{l=1}^r \varpi_l^N(\{\tau_k\}) \int_{\mathbb{R}^d} \langle \nabla_\mu \Lambda_l(\tau_k, \mu^*(\tau_k))(x), \omega_k \rangle d\mu^*(\tau_k)(x) \\
& \leq \mathcal{H}(\tau_k, \nu_N^*(\tau_k), \zeta_N^*(\tau_k), u^*(\tau_k)) \\
& - \sum_{l=1}^r \varpi_l^N(\{\tau_k\}) \int_{\mathbb{R}^d} \langle \nabla_\mu \Lambda_l(\tau_k, \mu^*(\tau_k))(x), u^*(\tau_k, x) \rangle d\mu^*(\tau_k)(x)
\end{aligned} \tag{3.22}$$

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

Proof. Using the expression (3.2) of the augmented infinite-dimensional Hamiltonian $\mathcal{H}(\cdot, \cdot, \cdot, \cdot)$ along with the definition of the Wasserstein gradient (1.16) and its characterization given e.g. in Proposition 3.4 for general integral functionals, 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

$$\begin{aligned}
\mathcal{K}_{\omega_k, \tau_k}^N(t) & = [L(\tau_k, \mu^*(\tau_k), \omega_k) - L(\tau_k, \mu^*(\tau_k), u^*(\tau_k))] \\
& + \int_{\mathbb{R}^{2d}} \langle r, \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \rangle d\nu^*(t)(x, r) \\
& - \int_{\tau_k}^t \int_{\mathbb{R}^d} \langle \lambda_0^N \nabla_\mu L(s, \mu^*(s), u^*(s))(x), \mathcal{F}_s^{\omega_k, \tau_k} \circ \Phi_{(s, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(s)(x) ds \\
& - \sum_{l=1}^r \int_{\tau_k}^t \int_{\mathbb{R}^d} \langle \nabla_\mu \Lambda_l(s, \mu^*(s))(x), \mathcal{F}_s^{\omega_k, \tau_k} \circ \Phi_{(s, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(s)(x) d\varpi_l^N(s) \\
& - \sum_{l=1}^r \zeta_{l, N}^*(t) \int_{\mathbb{R}^d} \langle \nabla_\mu \Lambda_l(t, \mu^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(t)(x).
\end{aligned} \tag{3.23}$$

By construction, the maps $\mathcal{K}_{\omega_k, \tau_k}^N(\cdot)$ satisfy

$$\mathcal{K}_{\omega_k, \tau_k}^N(T) \leq 0, \tag{3.24}$$

since it can be verified that the evaluation of $\mathcal{K}_{\omega_k, \tau_k}^N(\cdot)$ at T coincides with the left-hand side of (3.18), which has been shown to be non-positive for all $k \in \{1, \dots, N\}$. Moreover, the evaluation of the maps $\mathcal{K}_{\omega_k, \tau_k}^N(\cdot)$ at τ_k can be written explicitly as

$$\begin{aligned} \mathcal{K}_{\omega_k, \tau_k}^N(\tau_k) &= \mathcal{H}(\tau_k, \nu_N^*(\tau_k), \zeta_N^*(\tau_k), \omega_k) - \mathcal{H}(\tau_k, \nu_N^*(\tau_k), \zeta_N^*(\tau_k), u^*(\tau_k)) \\ &\quad - \sum_{l=1}^r \varpi_l^N(\{\tau_k\}) \int_{\mathbb{R}^d} \langle \nabla_\mu \Lambda_l(\tau_k, \mu^*(\tau_k)), \omega_k(x) - u^*(\tau_k, x) \rangle d\mu^*(\tau_k)(x). \end{aligned} \quad (3.25)$$

We now aim to show that the maps $\mathcal{K}_{\omega_k, \tau_k}^N(\cdot)$ are constant over $[\tau_k, T]$. By definition, these functions are in $BV([0, T], \mathbb{R})$ and therefore admit a distributional derivative in the form of a finite Borel regular measure (see e.g. [14, Chapter 3]). A simple computation of the time derivatives of the last two terms in (3.23) shows that the non-absolutely continuous parts of the derivatives of the maps $\mathcal{K}_{\omega_k, \tau_k}^N(\cdot)$ 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 absolutely continuous and therefore differentiable \mathcal{L}^1 -almost everywhere. One can then compute their derivative at \mathcal{L}^1 -almost every $t \in [\tau_k, T]$ as follows.

$$\begin{aligned} &\frac{d}{dt} \mathcal{K}_{\omega_k, \tau_k}^N(t) \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \langle r, \partial_t \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(T, \tau_k)}^{v, u^*}(x) \rangle d\sigma_{x, N}^*(t)(r) d\mu^*(T)(x) \\ &\quad + \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \langle \mathcal{V}_N^*[u_N^*(t)](t, x, r), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(T, \tau_k)}^{v, u^*}(x) \rangle d\sigma_{x, N}^*(t)(r) d\mu^*(T)(x) \\ &\quad - \int_{\mathbb{R}^d} \left\langle \lambda_0^N \nabla_\mu L(t, \mu^*(t), u^*(t)) \left(\Phi_{(T, t)}^{v, u^*}(x) \right), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(T, \tau_k)}^{v, u^*}(x) \right\rangle d\mu^*(T)(x) \\ &\quad - \sum_{l=1}^r \zeta_{l, N}^*(t) \frac{d}{dt} \left[\int_{\mathbb{R}^d} \left\langle \nabla_\mu \Lambda_l(t, \mu^*(t)) \left(\Phi_{(T, t)}^{v, u^*}(x) \right), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(T, \tau_k)}^{v, u^*}(x) \right\rangle d\mu^*(T)(x) \right] \end{aligned} \quad (3.26)$$

The time-derivatives of the summands of the last term can be computed as follows using Proposition 1.8 and the geometric structure (1.33) of solutions of (1.32) associated with the non-local velocity field $v[\mu^*(t)](t, \cdot) + u^*(t, \cdot)$.

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

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

by applying Fubini's Theorem and identifying the analytical expressions of the Wasserstein gradients of the summands $\mathcal{E}_l(t, \mu^*(t), u^*(t))$ of $\mathcal{C}(t, \mu^*(t), \zeta^*(t), u^*(t))$, which computation is postponed to Proposition (3.2) below. Plugging this expression into (3.26) along with the characterization (3.11) of $\partial_t \mathcal{F}_t^{\omega_k, \tau_k}(\cdot)$ derived in Lemma 3.3, we obtain that

$$\begin{aligned}
\frac{d}{dt} \mathcal{K}_{\omega_k, \tau_k}^N(t) & = \int_{\mathbb{R}^{2d}} \left\langle D_x u^*(t, x)^\top r + D_x v[\mu^*(t)](t, x)^\top r, \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v,u^*}(x) \right\rangle d\nu_N^*(t)(x, r) \\
& + \int_{\mathbb{R}^{2d}} \left\langle \mathbb{I}^v[\nu_N^*(t)](t, x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v,u^*}(x) \right\rangle d\nu_N^*(t)(x, r) \\
& + \int_{\mathbb{R}^{2d}} \left\langle \mathcal{V}_N^*[\nu_N^*(t)](t, x, r), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v,u^*}(x) \right\rangle d\nu_N^*(t)(x) \\
& - \int_{\mathbb{R}^d} \left\langle \lambda_0^N \nabla_{\mu} L(t, \mu^*(t), u^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v,u^*}(x) \right\rangle d\mu^*(t)(x) \\
& - \sum_{l=1}^r \zeta_{l,N}^*(t) \int_{\mathbb{R}^d} \left\langle \nabla_{\mu} \mathcal{E}_l(t, \mu^*(t), u^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(T,t)}^{v,u^*}(x) \right\rangle d\mu^*(t)(x)
\end{aligned}$$

where in the first line we used the fact that

$$\begin{aligned}
& \int_{\mathbb{R}^{2d}} \left\langle r, \int_{\mathbb{R}^d} \mathbb{I}^v_{(t,x)}(y) \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v,u^*}(y) d\mu^*(t)(y) \right\rangle d\nu_N^*(t)(x, r) \\
& = \int_{\mathbb{R}^d} \int_{\mathbb{R}^{2d}} \left\langle \mathbb{I}^v_{(t,y)}(x)^\top p, \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v,u^*}(x) \right\rangle d\nu_N^*(t)(x, r) d\mu^*(t)(y) \\
& = \int_{\mathbb{R}^d} \left\langle \mathbb{I}^v[\nu_N^*(t)](t, x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v,u^*}(x) \right\rangle d\mu^*(t)(x),
\end{aligned}$$

as a consequence of Fubini's Theorem, where $\mathbb{I}^v[\nu_N^*(\cdot)](\cdot, \cdot)$ is defined as in (3.20). Recalling the definition of the vector field $\mathcal{V}_N^*[\cdot](\cdot, \cdot, \cdot)$ given in (3.19), we therefore 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. Merging this fact with (3.24) and (3.25) yields (3.22) and concludes the proof of our claim. \square

Step 4 : Limiting procedure

In Step 3, we have built for any $N \geq 1$ a suitable state-costate curve $\nu_N^*(\cdot)$ solution of the Hamiltonian system (3.1), and such that the relaxed Pontryagin maximization

condition (3.22) holds on an N -dimensional subset of needle parameters. The last step in the proof of Theorem 3.1 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.

The PMP for absolutely continuous state constraints multipliers

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

$$(\lambda_0^N, \dots, \lambda_n^N, \eta_1^N, \dots, \eta_m^N) \xrightarrow{N \rightarrow +\infty} (\lambda_0, \dots, \lambda_n, \eta_1, \dots, \eta_m).$$

A straightforward passage to the limit shows that these limit multipliers satisfy the Complementary Slackness condition

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

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 as a consequence of Lemma 3.4. By Banach-Alaoglu's Theorem (see e.g. [38, Theorem 3.16]), it therefore admits a weakly-* converging subsequence to some $(\varpi_1, \dots, \varpi_r) \in \mathcal{M}_+([0, T])^r$. Moreover for any $l \in \{1, \dots, r\}$, the measures (ϖ_l^N) are equi-supported in the sets $\{t \in [0, T] \text{ s.t. } \Lambda_l(t, \mu^*(t)) = 0\}$. It can then be shown by standard convergence properties on the supports of sequences of measures (see e.g. [17, Proposition 5.1.8]) that the limit measures $(\varpi_1, \dots, \varpi_r)$ are supported on these sets as well. Therefore, it holds that

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

for any $l \in \{1, \dots, r\}$. Furthermore, if all 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 $\varpi_l \neq 0$ in the limit, at least for some $l \in \{1, \dots, r\}$. Hence, we recover the Non-Degeneracy condition (3.5), i.e.

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

In Lemma 3.5, we have shown that the curves of measures $\nu_N^*(\cdot)$ are uniformly equi-compactly supported and equi-Lipschitz. Hence, $(\nu_N^*(\cdot))$ admits converging subsequences in the C^0 -topology by Ascoli-Arzelà Theorem (see e.g. [123, Theorem 11.28]).

We now prove that there exists an accumulation point $\nu^*(\cdot)$ of $(\nu_N^*(\cdot))$ which solves the system of equations (3.1) associated with 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 $\mathcal{L}_{[0, T]}^1$.

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. Proposition 1.2-(ii), 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 1.1 that

$$\zeta_{i,N}^*(t) = \mathbb{1}_{[0,T)}(t) \int_t^T d\varpi_i^N(s) \xrightarrow[N \rightarrow +\infty]{} \mathbb{1}_{[0,T)}(t) \int_t^T d\varpi_i(s) = \zeta_i^*(t)$$

for all times $t \in [0, T]$. By definition (1.29) of distributional solutions to transport equations, the fact that $\nu_N^*(\cdot)$ is a solution of (3.19) can be written as

$$\begin{aligned} & \int_0^T \int_{\mathbb{R}^{2d}} \partial_t \xi(t, x, r) d\nu_N^*(t)(x, r) dt \\ & + \int_0^T \int_{\mathbb{R}^{2d}} \langle \nabla \xi(t, x, r), \mathbb{J}_{2d} \nabla_{\nu} \mathcal{H}(t, \nu_N^*(t), \zeta_N^*(t), u^*(t))(x, r) \rangle d\nu_N^*(t)(x, r) dt = 0 \end{aligned} \quad (3.27)$$

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 $t \in [0, T]$. By using this fact along with the uniform equi-compactness of the supports of $(\nu_N^*(\cdot))$, we can take the limit as $N \rightarrow +\infty$ in (3.27) and apply Lebesgue's Dominated Convergence Theorem to recover that

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

for any $\xi \in C_c^\infty([0, T] \times \mathbb{R}^{2d})$. Hence, the accumulation point $\nu^*(\cdot)$ of $(\nu_N^*(\cdot))$ in the C^0 -topology is a solution of the Hamiltonian flow (3.19) associated with the limit multipliers $(\lambda_0, \dots, \lambda_n, \eta_1, \dots, \eta_m, \varpi_1, \dots, \varpi_r)$.

In order to complete our proof of Theorem 3.1, there remains to show that the limit curve $\nu^*(\cdot)$ is such that the maximization condition (3.6) 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 (3.22) holds. By the preliminary assumption **(H7)** that the limit measures $(\varpi_1, \dots, \varpi_r)$ are absolutely continuous with respect to \mathcal{L}^1 , we can apply Proposition 1.1 to recover

that

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

for any $l \in \{1, \dots, r\}$. Since the infinite-dimensional Hamiltonian is continuous with respect to its second argument in the W_1 -metric and linear with respect to 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 $k \leq N$. We can therefore pass to the limit as $N \rightarrow +\infty$ in the partial maximization condition (3.22) to obtain that

$$\mathcal{H}(\tau_k, \nu^*(\tau_k), \zeta^*(\tau_k), \omega_k) \leq \mathcal{H}(\tau_k, \nu^*(\tau_k), \zeta^*(\tau_k), u^*(\tau_k)) \quad (3.28)$$

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

$$\begin{cases} \|u^*(\tau, \cdot) - u^*(\tau_k, \cdot)\|_{C^0(\overline{B(0, R_T)}, \mathbb{R}^d)} \xrightarrow{k \rightarrow +\infty} 0, \\ \|L(\tau, \mu^*(\tau), \cdot) - L(\tau_k, \mu^*(\tau_k), \cdot)\|_{C^0(U, \mathbb{R})} \xrightarrow{k \rightarrow +\infty} 0. \end{cases} \quad (3.29)$$

Remark first that under assumption **(H7)**, the maps $t \rightarrow \zeta_l^*(t)$ are continuous on $[0, T]$. By (3.29) 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 (3.28). 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 statement of Theorem 3.1, we now 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 $\rho \in C_c^\infty([0, T])$ is such that $\int_0^T \rho(t) dt = 1$. Given $N \geq 1$, we define the mollified measure $(\varpi_1^{N, \epsilon}, \dots, \varpi_r^{N, \epsilon})$ by

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

where for any $l \in \{1, \dots, r\}$, the convolution maps are defined by $\rho_\epsilon \star \varpi_l^N : t \in [0, T] \mapsto \int_0^T \rho_\epsilon(t-s) d\varpi_l^N(s)$. 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 as a by-product of **(H6)** and Lemma 3.3, 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 (3.18) derived at time T as

$$\begin{aligned} & \int_{\mathbb{R}^d} \langle -\nabla_\mu \mathcal{S}_N(\mu^*(T))(x), \mathcal{F}_T^{\omega_k, \tau_k} \circ \Phi_{(T, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(T)(x) \\ & - \lambda_N^0 \left(L(\tau_k, \mu^*(\tau_k), u^*(\tau_k)) - L(\tau_k, \mu^*(\tau_k), \omega_k) \right) \\ & - \int_{\tau_k}^T \int_{\mathbb{R}^d} \langle \lambda_N^0 \nabla_\mu L(t, \mu^*(t), u^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(t)(x) dt \\ & - \sum_{l=1}^r \int_{\tau_k}^T \int_{\mathbb{R}^d} \langle \nabla_\mu \Lambda_l(t, \mu^*(t))(x), \mathcal{F}_t^{\omega_k, \tau_k} \circ \Phi_{(t, \tau_k)}^{v, u^*}(x) \rangle d\mu^*(t)(x) d\varpi_l^{N, \epsilon}(t) \leq C\epsilon \end{aligned} \quad (3.30)$$

By defining the family of measure curves $(\nu_{N, \epsilon}^*(\cdot))$ as in Lemma 3.5, we can prove that the corresponding maps $\mathcal{K}_{\omega_k, \tau_k}^{N, \epsilon}(\cdot)$ defined as in (3.26) 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 again the limiting arguments of Step 4 as $N \rightarrow +\infty$ and remarking that

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

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

$$\mathcal{H}(\tau, \nu_\epsilon^*(\tau), \zeta_\epsilon^*(\tau), \omega) \leq \mathcal{H}(\tau, \nu_\epsilon^*(\tau), \zeta_\epsilon^*(\tau), u^*(\tau)) + C\epsilon \quad (3.31)$$

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. [14, 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 1.1 it is

such that

$$\zeta_{l,\epsilon}^*(\tau) = \int_{\tau}^T d\varpi_l^\epsilon(s) \xrightarrow{\epsilon \downarrow 0} \int_{\tau}^T d\varpi_l(s) = \zeta_l^*(\tau).$$

for any $\tau \in \mathcal{M}^\circ$. By Ascoli-Arzelà Theorem, it holds that $\nu_\epsilon^*(\cdot) \rightarrow \nu^*(\cdot)$ in the C^0 -topology and it can be checked that this limit curve solves the forward-backward Hamiltonian continuity equations (3.1) 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 (3.31) implies that the Pontryagin maximization condition (3.6) holds on the restricted subset $U \times \mathcal{M}^\circ$.

3.3 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 $(t, \mu, \zeta, \omega) \mapsto \mathcal{C}(t, \mu, \zeta, \omega)$ defined in (3.3). We chose to postpone this result since its proof is quite heavy proof and gives no further insight on the methodology that we develop in this chapter.

Proposition 3.2 (Wasserstein differential of the state constraints penalization map). *Let $K \subset \mathbb{R}^d$ be a compact set and $\omega \in U$. Under hypotheses (H6), the map $\mu \in \mathcal{P}(K) \mapsto \mathcal{C}(t, \mu, \zeta^*(t), \omega)$ defined by*

$$\begin{aligned} \mathcal{C}(t, \mu, \zeta^*(t), \omega) &= \sum_{l=1}^r \zeta_l^*(t) \mathcal{C}_l(t, \mu, \omega) \\ &= \sum_{l=1}^r \zeta_l^*(t) \left(\partial_t \Lambda_l(t, \mu) + \int_{\mathbb{R}^d} \langle \nabla_\mu \Lambda_l(t, \mu)(x), v[\mu](t, x) + \omega(x) \rangle d\mu(x) \right) \end{aligned}$$

is regular and Wasserstein differentiable at any $\mu \in \mathcal{P}(K)$. The Wasserstein gradients of its summands $\mathcal{C}_l(t, \cdot, \omega)$ can be computed explicitly as

$$\begin{aligned} \nabla_\mu \mathcal{C}_l(t, \mu, \omega)(x) &= \partial_t \nabla_\mu \Lambda_l(t, \mu)(x) + D_x \nabla_\mu \Lambda_l(t, \mu)(x)^\top (v[\mu](t, x) + \omega(x)) \\ &\quad + \left(D_x v[\mu](t, x)^\top + D_x u^*(t, x)^\top \right) \nabla_\mu \Lambda_l(t, \mu)(x) \\ &\quad + \int_{\mathbb{R}^d} \mathbb{I}_{(t,y)}^v(x)^\top \nabla_\mu \Lambda_l(t, \mu)(y) d\mu(y) \\ &\quad + \int_{\mathbb{R}^d} \mathbb{I}_{(t,y)}^{\nabla_\mu \Lambda_l}(x)^\top (v[\mu](t, y) + u^*(t, y)) d\mu(y) \end{aligned} \tag{3.32}$$

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 rows 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

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

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

$$\nabla_{\mu} \mathcal{E}_t^1(t, \mu) = \nabla_{\mu} (\partial_t \Lambda_l(t, \cdot)) (\mu) = \partial_t \nabla_{\mu} \Lambda_l(t, \mu) \quad (3.33)$$

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

$$\begin{aligned} & \mathcal{E}_t^2(t, \nu) - \mathcal{E}_t^2(t, \mu) \\ &= \int_{\mathbb{R}^{2d}} \left(\langle \nabla_{\mu} \Lambda_l(t, \nu)(y), v[\nu](t, y) + u^*(t, y) \rangle - \langle \nabla_{\mu} \Lambda_l(t, \mu)(x), v[\mu](t, x) + \omega(x) \rangle \right) d\gamma(x, y) \\ &= \int_{\mathbb{R}^d} \left(\langle \nabla_{\mu} \Lambda_l(t, \nu)(x), v[\nu](t, x) + \omega(x) \rangle - \langle \nabla_{\mu} \Lambda_l(t, \mu)(x), v[\mu](t, x) + \omega(x) \rangle \right) d\mu(x) \\ &+ \int_{\mathbb{R}^{2d}} \left\langle D_x \nabla_{\mu} \Lambda_l(t, \mu)(x)^{\top} (v[\mu](x) + \omega(x)), y - x \right\rangle d\gamma(x, y) \\ &+ \int_{\mathbb{R}^{2d}} \left\langle \left(D_x v[\mu](t, x)^{\top} + D_x \omega(x)^{\top} \right) \nabla_{\mu} \Lambda_l(t, \mu)(x), y - x \right\rangle d\gamma(x, y) \\ &+ \int_{\mathbb{R}^{2d}} o(|x - y|) d\gamma(x, y) \end{aligned} \quad (3.34)$$

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

$$\begin{aligned} & \int_{\mathbb{R}^d} \left(\langle \nabla_{\mu} \Lambda_l(t, \nu)(x), v[\nu](t, x) + \omega(x) \rangle - \langle \nabla_{\mu} \Lambda_l(t, \mu)(x), v[\mu](t, x) + \omega(x) \rangle \right) d\mu(x) \\ &= \int_{\mathbb{R}^d} \left\langle \int_{\mathbb{R}^{2d}} \mathbb{I}_{(t,x)}^{\nabla_{\mu} \Lambda_l}(x')(y' - x') d\gamma'(x', y'), v[\mu](t, x) + \omega(x) \right\rangle d\mu(x) \\ &+ \int_{\mathbb{R}^d} \left\langle \int_{\mathbb{R}^{2d}} \mathbb{I}_{(t,x)}^v(x')(y' - x') d\gamma'(x', y'), \nabla_{\mu} \Lambda_l(t, \mu)(x) \right\rangle d\mu(x) + o(W_2(\mu, \nu)) \end{aligned} \quad (3.35)$$

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

by a standard application of Fubini's Theorem. Merging equations (3.33), (3.34) and (3.35), we recover the characterization (1.16) of the Wasserstein gradient $\nabla_\mu \mathcal{C}_l(t, \mu, \omega)(\cdot)$ of $\mathcal{C}_l(t, \cdot, \omega)$ at μ given by (3.32). \square

3.4 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 3.1. This list of examples has a non-empty intersection with that of Section 2.3.

Proposition 3.3 (Example of non-local velocity field). *Let $(t, x, y) \mapsto H(t, x, y) \in \mathbb{R}^d$ 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 (H2)-(H3) of Theorem 3.1. 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 3.4 (Example of cost and constraint functions). *Let $n \geq 1$ and $W \in C^1(\mathbb{R}^{nd}, \mathbb{R})$. Then, the functional*

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

with $\mu^{\otimes n} = \mu \times \dots \times \mu$ satisfies (H4) of Theorem 3.1 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 3.1 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 3.5 (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 3.1 and their derivatives can be computed using Propositions 3.4.

Remark 3.3. *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 $S \subset \mathbb{R}^d$ $\mu \mapsto \frac{1}{2} \int_{\mathbb{R}^d} d_S(x)^2 d\mu(x)$.*

4 Intrinsic Lipschitz Regularity of Mean-Field Optimal Controls

In this chapter, we investigate the intrinsic Lipschitz-in-space regularity of the solutions of general mean-field optimal control problems of the form

$$(\mathcal{P}) \quad \begin{cases} \min_{u \in \mathcal{U}} \left[\int_0^T \left(L(t, \mu(t)) + \int_{\mathbb{R}^d} \psi(u(t, x)) d\mu(t)(x) \right) dt + \varphi(\mu(T)) \right] \\ \text{s.t.} \quad \begin{cases} \partial_t \mu(t) + \nabla \cdot ((v[\mu(t)](t, \cdot) + u(t, \cdot))\mu(t)) = 0, \\ \mu(0) = \mu^0. \end{cases} \end{cases}$$

The set of admissible controls for (\mathcal{P}) is defined by $\mathcal{U} = L^1([0, T], L^1(\mathbb{R}^d, U; \mu(t)))$ where $U \subset \mathbb{R}^d$ is a convex and compact set. Remark that since we do not impose any a priori regularity assumptions on the control vector fields $u(\cdot, \cdot)$, there may not exist solutions to the non-local transport equation driving problem (\mathcal{P}) . Moreover, even if they do exist, these solution will not be classically well-posed and only defined in a weak sense (see Theorem 1.6).

As already discussed in the Introduction, the main motivation to study multi-agent systems in the mean-field setting is to be able to design scalable control laws that can in turn be applied to discrete systems. In general, such a procedure is only applicable in the case in which the controls are sufficiently regular with respect to the space variable.

The main contribution of this chapter is an existence result of intrinsically Lipschitz optimal controls for (\mathcal{P}) stated in Theorem 4.1 below. The proof of the latter is built around two main ingredients. The first one is an existence result for mean-field optimal controls which was derived in [80] and stated in Theorem 4.2 below. In this article, the authors prove under very general assumptions that there exist optimal solutions of problem $(\mathcal{P}_{\text{meas}})$ which can be recovered as Γ -limits in a suitable topology of sequences of solutions of the discrete problems

$$(\mathcal{P}_N) \quad \begin{cases} \min_{u(\cdot) \in \mathcal{U}_N} \left[\int_0^T \left(\mathbf{L}_N(t, \mathbf{x}(t)) + \frac{1}{N} \sum_{i=1}^N \psi(u_i(t)) \right) dt + \varphi_N(\mathbf{x}(T)) \right] \\ \text{s.t.} \quad \begin{cases} \dot{x}_i(t) = \mathbf{v}_N[\mathbf{x}(t)](t, x_i(t)) + u_i(t), \\ x_i(0) = x_i^0. \end{cases} \end{cases}$$

Here, $\mathcal{U}_N = L^\infty([0, T], U^N)$, and the functionals $(t, x, \mathbf{x}) \in [0, T] \times \mathbb{R}^d \times (\mathbb{R}^d)^N \mapsto \mathbf{v}_N[\mathbf{x}](t, x)$, $(t, \mathbf{x}) \in [0, T] \times (\mathbb{R}^d)^N \mapsto \mathbf{L}_N(t, \mathbf{x})$ and $\mathbf{x} \in (\mathbb{R}^d)^N \mapsto \varphi(\mathbf{x})$ are discrete approximating sequences (see Definition 4.1 below) for $v[\cdot](\cdot, \cdot)$, $L(\cdot, \cdot)$ and $\varphi(\cdot)$ respectively. To obtain this convergence result, it is necessary to introduce an intermediate relaxed

problem which encompasses both (\mathcal{P}) and the sequence (\mathcal{P}_N) . This problem is defined by

$$(\mathcal{P}_{\text{meas}}) \begin{cases} \min_{\boldsymbol{\nu} \in \mathcal{U}} \left[\int_0^T \left(L(t, \mu(t)) + \Psi(\boldsymbol{\nu}(t) | \mu(t)) \right) dt + \varphi(\mu(T)) \right] \\ \text{s.t.} \begin{cases} \partial_t \mu(t) + \nabla \cdot ((v[\mu(t)](t, \cdot) \mu(t) + \boldsymbol{\nu}(t)) = 0, \\ \mu(0) = \mu^0. \end{cases} \end{cases}$$

where $\mathcal{U} = \mathcal{M}([0, T] \times \mathbb{R}^d, U)$ is the set of *generalized measure controls*, $t \in [0, T] \mapsto \boldsymbol{\nu}(t) \in \mathcal{M}(\mathbb{R}^d, U)$ is a curve of control measure and $\Psi(\cdot | \mu)$ is an internal energy functional of the form

$$\Psi(\cdot | \mu) : \boldsymbol{\sigma} \in \mathcal{M}(\mathbb{R}^d, U) \mapsto \begin{cases} \int_{\mathbb{R}^d} \psi \left(\frac{d\boldsymbol{\sigma}}{d\mu}(x) \right) d\mu(x) & \text{if } \boldsymbol{\sigma} \ll \mu, \\ +\infty & \text{otherwise.} \end{cases}$$

As discussed more precisely in Section 4.1.2, the discrete problems (\mathcal{P}_N) are linked to $(\mathcal{P}_{\text{meas}})$ via the empirical state and control measures defined by

$$\mu_N(t) = \frac{1}{N} \sum_{i=1}^N \delta_{x_i(t)} \quad \text{and} \quad \boldsymbol{\nu}_N(t) = \frac{1}{N} \sum_{i=1}^N u_i(t) \delta_{x_i(t)},$$

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

The second key component of our approach is to adapt to the family of problems (\mathcal{P}_N) a methodology developed in [58, 71] which provides general metric regularity results (see Definition 4.4 below) for a large class of dynamical differential inclusion. This part relies crucially on the following *uniform mean-field coercivity estimate* (\mathbf{CO}_N) for the sequence of problems (\mathcal{P}_N)

$$\begin{aligned} \mathbf{Hess}_x \varphi_N[\mathbf{x}_N^*(T)](\mathbf{y}(T), \mathbf{y}(T)) &- \int_0^T \mathbf{Hess}_x \mathbb{H}_N[t, \mathbf{x}_N^*(t), \mathbf{r}_N^*(t), \mathbf{u}_N^*(t)](\mathbf{y}(t), \mathbf{y}(t)) dt \\ &- \int_0^T \mathbf{Hess}_u \mathbb{H}_N[t, \mathbf{x}_N^*(t), \mathbf{r}_N^*(t), \mathbf{u}_N^*(t)](\mathbf{w}(t), \mathbf{w}(t)) dt \\ &\geq \rho_T \int_0^T |\mathbf{w}(t)|_N^2 dt, \end{aligned}$$

along optimal mean-field Pontryagin triples $(\mathbf{u}_N^*(\cdot), \mathbf{x}_N^*(\cdot), \mathbf{r}_N^*(\cdot))$ (see Proposition 4.3 below), and where $(\mathbf{w}(\cdot), \mathbf{y}(\cdot))$ are control-trajectory pairs solution of a suitable linearized problem. In this context, $\mathbf{Hess}(\bullet)[\cdot](\cdot, \cdot)$ denotes the discretization of the Wasserstein Hessian bilinear form (see e.g. [57, 88]), which construction is detailed in Section 4.1. In essence, this uniform coercivity assumption allows one to inverse the maximization condition stemming from an application of the PMP to (\mathcal{P}_N) , with a uniform control on the Lipschitz constant of this inverse. The main subtlety lies in the fact that we need these estimates to be uniform with respect to N . Whence, we apply an adapted mean-field Pontryagin Maximum Principle to (\mathcal{P}_N) in the spirit of [31], which is in essence the discrete counterpart of the Wasserstein PMP studied in Chapter 2 and 3. We then express the coercivity condition in terms of Wasserstein calculus.

By following this strategy of proof, we are able to obtain the following result.

Theorem 4.1 (Existence of Lipschitz-in-space optimal controls for (\mathcal{P})). *Let $\mu^0 \in \mathcal{P}_c(\mathbb{R}^d)$, $(\mu_N^0) \subset \mathcal{P}_c(\mathbb{R}^d)$ be a sequence of empirical measures narrowly converging towards μ^0 , and suppose that the set of hypotheses **(H)** below hold.*

Hypotheses **(H)**

- (H1)** *The set of admissible control values $U \subset \mathbb{R}^d$ is a convex and compact set containing a neighbourhood of the origin.*
- (H2)** *The control cost $v \mapsto \psi(v) \in [0, +\infty]$ is coercive, $C_{\text{loc}}^{2,1}$ -regular, and strictly convex.*
- (H3)** *The non-local velocity field $(t, x, \mu) \mapsto v[\mu](t, x) \in \mathbb{R}^d$ is Lipschitz with respect to $t \in [0, T]$ and continuous in the product $|\cdot| \times W_2$ -topology with respect to $(x, \mu) \in \mathbb{R}^d \times \mathcal{P}_c(\mathbb{R}^d)$. For all times $t \in [0, T]$, it is such that*

$$|v[\mu](t, x)| \leq M \left(1 + |x| + \left(\int_{\mathbb{R}^d} |y|^2 d\mu(y) \right)^{1/2} \right),$$

for a given constant $M > 0$ and any $(x, \mu) \in \mathbb{R}^d \times \mathcal{P}_c(\mathbb{R}^d)$. It further satisfies the Cauchy-Lipschitz properties

$$|v[\mu](t, x) - v[\mu](t, y)| \leq L_1^K |x - y|, \quad \|v[\mu](t, \cdot) - v[\nu](t, \cdot)\|_{C^0(K)} \leq L_2^K W_2(\mu, \nu)$$

on any compact set $K \subset \mathbb{R}^d$ and for any pairs $x, y \in K$ and $\mu, \nu \in \mathcal{P}_c(\mathbb{R}^d)$.

- (H4)** *The map $v[\cdot](t, x)$ is $C_{\text{loc}}^{2,1}$ -Wasserstein regular in the sense of Definition 4.2, uniformly with respect to $(t, x) \in [0, T] \times K$ where $K \subset \mathbb{R}^d$ is compact.*
- (H5)** *The running cost $(t, \mu) \mapsto L(t, \mu)$ is Lipschitz with respect to $t \in [0, T]$ and $C_{\text{loc}}^{2,1}$ -Wasserstein regular with respect to $\mu \in \mathcal{P}_c(\mathbb{R}^d)$ in the sense of Definition 4.2.*
- (H6)** *The final cost $\mu \mapsto \varphi(\mu)$ is $C_{\text{loc}}^{2,1}$ -Wasserstein regular in the sense of Definition 4.2, uniformly with respect to $t \in [0, T]$.*

*Furthermore, assume that the mean-field coercivity assumption **(CO_N)** described in Section 4.2 holds. Then, there exists an optimal pair control-trajectory $(u^*(\cdot, \cdot), \mu^*(\cdot)) \in \mathcal{U} \times \text{Lip}([0, T], \mathcal{P}_c(\mathbb{R}^d))$ for problem (\mathcal{P}) . Moreover, the map $x \in \mathbb{R}^d \mapsto u^*(t, \cdot) \in U$ is \mathcal{L}_U -Lipschitz for \mathcal{L}^1 -almost every $t \in [0, T]$, where the uniform constant \mathcal{L}_U only depends on the datum of the problem (\mathcal{P}) .*

This chapter is structured as follows. In Section 4.1, we recall several general prerequisites dealing with the mean-field approximation of problem (\mathcal{P}) that will prove useful in the sequel. We also provide a precise statement of the existence result of locally optimal feedbacks for finite-dimensional optimal control problems. In Section 4.2, we introduce the mean-field coercivity estimate **(CO_N)** and prove Theorem 4.1. We provide a discussion of the coercivity assumptions in Section 4.3, by presenting in particular a sufficient condition for its validity as well as an example on which it is sharp.

4.1 Preliminary material

In this section, we introduce some tools which are specific to the combination that we make of mean-field limit approaches and optimal control in order to prove Theorem 4.1. In Section 4.1.1 we introduce an adapted differential structure for discrete mean-field optimal control problems along with the existence result of [80] in Section 4.1.2. We state a general result on locally optimal Lipschitz feedbacks in the context of finite-dimensional optimal control problems in Section 4.1.3.

4.1.1 Mean-field adapted structures and discrete measures

In this section, we introduce the notion of *mean-field approximating sequence* for a functional defined on measures, along with a discretized version of the differential structure described in Section 1.2.

We define the set $\mathcal{P}_N(\mathbb{R}^d) = \{\frac{1}{N} \sum_{i=1}^N \delta_{x_i} \text{ s.t. } (x_1, \dots, x_N) \in (\mathbb{R}^d)^N\}$ of *N-empirical probability measures*. It is a standard result in optimal transport theory (see e.g. [17, Chapter 7]) that $\cup_N \mathcal{P}_N(\mathbb{R}^d)$ is a dense and countable subset of $\mathcal{P}(\mathbb{R}^d)$ with respect to the narrow topology. For any $N \geq 1$, we denote by $\mathbf{x} = (x_1, \dots, x_N)$ a given element of $(\mathbb{R}^d)^N$ and by $\mu[\mathbf{x}] = \frac{1}{N} \sum_{i=1}^N \delta_{x_i} \in \mathcal{P}_N(\mathbb{R}^d)$ its associated empirical measure. A map $\phi : (\mathbb{R}^d)^N \rightarrow \mathbb{R}^m$ is said to be symmetric if $\phi \circ \sigma(\cdot) = \phi(\cdot)$ for any d -blockwise permutation $\sigma : (\mathbb{R}^d)^N \rightarrow (\mathbb{R}^d)^N$. This symmetry under permutation encodes the indistinguishability of the discrete particles (x_1, \dots, x_N) and is therefore needed to perform mean-field approximations. In the following definition, we recall the notion of *mean-field approximating sequence* for a continuous functional $\phi(\cdot)$ defined over $\mathcal{P}_c(\mathbb{R}^d)$.

Definition 4.1 (Mean-field approximating sequence). *Let $\phi \in C^0(\mathcal{P}_c(\mathbb{R}^d), \mathbb{R}^m)$. The mean-field approximating sequence of $\phi(\cdot)$ is the sequence $(\phi_N(\cdot)) \subset C^0((\mathbb{R}^d)^N, \mathbb{R}^m)$ of symmetric maps such that*

$$\phi(\mu[\mathbf{x}]) = \phi_N(\mathbf{x}), \quad (4.1)$$

for any $N \geq 1$ and $\mathbf{x} \in (\mathbb{R}^d)^N$. Given an integer $n \geq 1$ and a set $\Omega \subset \mathbb{R}^n$, we similarly define the mean-field approximating sequence of a functional $F \in C^0(\Omega \times \mathcal{P}_c(\mathbb{R}^d), \mathbb{R}^m)$ as the family of symmetric maps $(\mathbf{F}_N(\cdot, \cdot)) \subset C^0(\Omega \times (\mathbb{R}^d)^N, \mathbb{R}^m)$ such that

$$F(x, \mu[\mathbf{x}]) = \mathbf{F}_N(x, \mathbf{x})$$

for any $N \geq 1$ and $(x, \mathbf{x}) \in \Omega \times (\mathbb{R}^d)^N$.

Remark 4.1 (A more general definition of mean-field adaptedness). *In [80] and several other references in the literature, the authors consider an alternative definition of mean-field adaptedness which is valid for functionals $\phi(\cdot)$ that are only semicontinuous. This more general definition requires the existence of a sequence of maps $(\phi_N(\cdot)) \in C^0(\mathcal{P}_N(\mathbb{R}^d), \mathbb{R}^m)$ such that*

$$\phi_N(\mu[\mathbf{x}]) = \phi_N(\mathbf{x}), \quad |\phi(\mu) - \phi_N(\mu_N)| \xrightarrow{N \rightarrow +\infty} 0,$$

whenever $(\mu_N) \subset \mathcal{P}_N(\mathbb{R}^d)$ is narrowly converging towards $\mu \in \mathcal{P}_1(\mathbb{R}^d)$. This notion is

useful e.g. to encompass sequences of internal energy functionals of the form

$$\phi : \mu \in \mathcal{P}_c(\mathbb{R}^d) \mapsto \begin{cases} \int_{\mathbb{R}^d} F\left(\frac{d\mu}{d\nu_N}(x)\right) d\nu_N(x) & \text{if } \mu \ll \nu_N, \\ +\infty & \text{otherwise.} \end{cases}$$

where $(\nu_N) \subset \mathcal{P}_N(\mathbb{R}^d)$ is a sequence of empirical measures. To lighten the computations throughout the article, we restricted the notion presented in Definition 4.1 to the particular cases in which $\phi(\cdot)$ is continuous, so that one automatically has that $\phi_N(\cdot) \equiv \phi(\cdot)$. Let it be noted that the conclusions of Theorem 4.1 still hold provided that the relevant notions presented in the remainder of the paper – in particular the estimates of Definition 4.2 below – apply to the approximating sequence $(\phi_N(\cdot))$ uniformly with respect to N .

In what follows, we leverage the formalism of Wasserstein differential calculus described in Section 1.2 to define an adapted notion of differentiability for mean-field approximating sequences. We start by introducing the notion of $C_{\text{loc}}^{2,1}$ -Wasserstein regularity.

Definition 4.2 ($C_{\text{loc}}^{2,1}$ -Wasserstein regularity). *A functional $\phi : \mathcal{P}_c(\mathbb{R}^d) \rightarrow \mathbb{R}^m$ is said to be $C_{\text{loc}}^{2,1}$ -Wasserstein regular if for any compact set $K \subset \mathbb{R}^d$ the map $\phi(\cdot)$ is twice differentiable over $\mathcal{P}(K)$ in the sense of Definition 1.11 and such that*

$$\begin{aligned} & \phi(\mu) + \|\nabla_{\mu}\phi(\mu)(\cdot)\|_{C^0(K)} + \|D_x \nabla_{\mu}\phi(\mu)(\cdot)\|_{C^0(K)} + \left\| D_{\mu}^2 \phi(\mu)(\cdot, \cdot) \right\|_{C^0(K \times K)} \\ & + \text{Lip}(D_x \nabla_{\mu}\phi(\cdot)(\cdot); \mathcal{P}(K) \times K) + \text{Lip}(D_{\mu}^2 \phi(\cdot)(\cdot, \cdot); \mathcal{P}(K) \times K \times K) \leq C_K \end{aligned} \quad (4.2)$$

for all $\mu \in \mathcal{P}(K)$, where $C_K > 0$ is a constant depending on K .

We provide in what follows a series of examples of classical $C_{\text{loc}}^{2,1}$ -MF functionals that can be frequently encountered in applications.

Example 4.1 (Potential functionals). *Let $V \in C_{\text{loc}}^{2,1}(\mathbb{R}^d, \mathbb{R})$. Then, the functional on measures $\mathcal{V} : \mu \in \mathcal{P}_c(\mathbb{R}^d) \mapsto \int_{\mathbb{R}^d} V(x) d\mu(x)$ has a mean-field approximating sequence given by $\mathbf{V}_N : \mathbf{x} \in (\mathbb{R}^d)^N \mapsto \frac{1}{N} \sum_{i=1}^N V(x_i)$. It is twice differentiable in the sense of Definition 1.11, and its first and second order Wasserstein derivatives can be computed explicitly as*

$$\nabla_{\mu} \mathcal{V}(\mu)(x) = \nabla V(x), \quad D_x \nabla_{\mu} \mathcal{V}(\mu)(x) = \nabla^2 V(x), \quad D_{\mu}^2 \mathcal{V}(\mu)(x, y) = 0,$$

for any $(\mu, x, y) \in \mathcal{P}_c(\mathbb{R}^d) \times \mathbb{R}^{2d}$. Whence, we deduce that $\mathcal{V}(\cdot)$ is $C_{\text{loc}}^{2,1}$ -Wasserstein whenever $V \in C_{\text{loc}}^{2,1}(\mathbb{R}^d, \mathbb{R})$. The same conclusion still holds for more general functionals $\mathcal{W}, \mathcal{F} : \mathcal{P}_c(\mathbb{R}^d) \rightarrow (-\infty, +\infty]$ of the form

$$\mathcal{W}(\mu) = \int_{\mathbb{R}^d} W(x_1, \dots, x_n) d(\mu \times \dots \times \mu)(x), \quad \mathcal{F}(\mu) = \int_{\mathbb{R}^d} L(x, \int_{\mathbb{R}^d} m(y) d\mu(y)) d\mu(x),$$

provided that $W \in C_{\text{loc}}^{2,1}((\mathbb{R}^d)^n, \mathbb{R})$, $m \in C_{\text{loc}}^{2,1}(\mathbb{R}^d, \mathbb{R}^m)$ and $L \in C_{\text{loc}}^{2,1}(\mathbb{R}^d \times \mathbb{R}^m, \mathbb{R})$.

Example 4.2 (Variance functional). *We define the variance functional over $\mathcal{P}_2(\mathbb{R}^d)$ by $\text{Var}(\mu) = \frac{1}{2} \int_{\mathbb{R}^d} |x - \bar{\mu}|^2 d\mu(x)$ where $\bar{\mu} = \int_{\mathbb{R}^d} y d\mu(y)$. Its approximating sequence writes $\mathbf{Var}_N : \mathbf{x} \in (\mathbb{R}^d)^N \mapsto \frac{1}{2N} \sum_{i=1}^N |x_i - \bar{\mathbf{x}}|^2$. The variance functional is again twice*

differentiable in the sense of Definition 1.11 at any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, and its first and second order derivatives can be computed explicitly as

$$\nabla_\mu \text{Var}(\mu)(x) = x - \bar{\mu}, \quad D_x \nabla_\mu \text{Var}(\mu)(x) = \text{Id}, \quad D_\mu^2 \text{Var}(\mu)(x, y) = -\text{Id}.$$

It moreover satisfies the $C_{\text{loc}}^{2,1}$ -Wasserstein regularity assumption (4.2).

In the sequel, we endow the Euclidean space $(\mathbb{R}^d)^N$ with the rescaled inner product $\langle \cdot, \cdot \rangle_N$, defined by

$$\langle \mathbf{x}, \mathbf{y} \rangle_N = \frac{1}{N} \sum_{i=1}^N \langle x_i, y_i \rangle \quad (4.3)$$

for any $\mathbf{x}, \mathbf{y} \in (\mathbb{R}^d)^N$, where $\langle \cdot, \cdot \rangle$ is the standard Euclidean inner product of \mathbb{R}^d . We denote by $|\cdot|_N = \sqrt{\langle \cdot, \cdot \rangle_N}$ the rescaled Euclidean norm induced by $\langle \cdot, \cdot \rangle_N$ over $(\mathbb{R}^d)^N$, and remark that $((\mathbb{R}^d)^N, \langle \cdot, \cdot \rangle_N)$ has the structure of a Hilbert space.

In the following proposition, we show that the Wasserstein differential structure described in Section 1.2 for functionals defined on measures induces a natural differential structure on $(\mathbb{R}^d)^N$ adapted to the rescaled inner product $\langle \cdot, \cdot \rangle_N$.

Proposition 4.1 (Mean-field derivatives of symmetric maps). *Let $\phi \in C^0(\mathcal{P}_c(\mathbb{R}^d), \mathbb{R})$ be a $C_{\text{loc}}^{2,1}$ -Wasserstein regular map with mean-field approximating sequence $(\phi_N(\cdot)) \subset C^0((\mathbb{R}^d)^N, \mathbb{R})$. Then one has that $\phi_N \in C_{\text{loc}}^{2,1}((\mathbb{R}^d)^N, \mathbb{R})$ for any $N \geq 1$, and the following Taylor expansion holds*

$$\phi_N(\mathbf{x} + \mathbf{h}) = \phi_N(\mathbf{x}) + \langle \mathbf{Grad} \phi_N(\mathbf{x}), \mathbf{h} \rangle_N + \frac{1}{2} \mathbf{Hess} \phi_N[\mathbf{x}](\mathbf{h}, \mathbf{h}) + o(|\mathbf{h}|_N^2), \quad (4.4)$$

for any $\mathbf{x}, \mathbf{h} \in (\mathbb{R}^d)^N$, where we introduced the mean-field gradient $\mathbf{Grad} \phi_N(\cdot)$ and mean-field Hessian $\mathbf{Hess} \phi_N[\cdot]$ of $\phi_N(\cdot)$, defined respectively by

$$\mathbf{Grad} \phi_N(\mathbf{x}) = (\nabla_\mu \phi(\mu[\mathbf{x}])(x_i))_{1 \leq i \leq N} \quad (4.5)$$

and

$$\mathbf{Hess} \phi_N[\mathbf{x}](\mathbf{h}, \mathbf{h}) = \frac{1}{N} \sum_{i=1}^N \langle D_x \nabla_\mu \phi(\mu[\mathbf{x}])(x_i) h_i, h_i \rangle_N + \frac{1}{N^2} \sum_{i,j=1}^N \langle D_\mu^2 \phi(\mu[\mathbf{x}])(x_i, x_j) h_i, h_j \rangle. \quad (4.6)$$

Moreover for any compact set $K \subset \mathbb{R}^d$, there exists a constant $C_K > 0$ such that for any $N \geq 1$, one has that

$$\|\phi_N(\cdot)\|_{C^2(K^N)} + \text{Lip}(\mathbf{Hess} \phi_N[\cdot], K^N) \leq C_K \quad (4.7)$$

where the C^2 -norm here is defined by

$$\|\phi_N(\cdot)\|_{C^2(K)} = \max_{\mathbf{x} \in K} \phi_N(\mathbf{x}) + \max_{\mathbf{x} \in K} |\mathbf{Grad} \phi_N(\mathbf{x})|_N + \max_{\mathbf{x} \in K} \sup_{|\mathbf{h}|_N \leq 1} \mathbf{Hess} \phi_N[\mathbf{x}](\mathbf{h}, \mathbf{h}), \quad (4.8)$$

for any set $K \subset (\mathbb{R}^d)^N$.

Proof. Let $\mathbf{x}, \mathbf{h} \in (\mathbb{R}^d)^N$, $\epsilon = \frac{1}{4} \min_{x_i \neq x_j} |x_i - x_j|$ and $\zeta_N(\cdot)$ be the map defined by

$$\zeta_N : x \in \mathbb{R}^d \mapsto \begin{cases} \langle x, h_i \rangle & \text{if } x \in B(x_i, 2\epsilon), \\ 0 & \text{otherwise.} \end{cases}$$

Let $\eta \in C_c^\infty(\mathbb{R}^d)$ be a symmetric mollifier centered at the origin and supported on the closure of $B(0, \epsilon)$. We define the tangent vector $\xi_N \in \nabla C_c^\infty(\mathbb{R}^d) \subset \text{Tan}_{\mu[\mathbf{x}]} \mathcal{P}_2(\mathbb{R}^d)$ at $\mu[\mathbf{x}]$ by

$$\xi_N : x \in \mathbb{R}^d \mapsto \nabla(\eta * \zeta_N)(x). \quad (4.9)$$

Remark that by construction it holds that

$$\xi_N(x_i) = h_i, \quad D_x \xi_N(x_i) = 0, \quad (4.10)$$

so that in particular $\mu[\mathbf{x} + s\mathbf{h}] = (\text{Id} + s\xi_N) \# \mu[\mathbf{x}]$ for any $s \in \mathbb{R}$.

By assumption, the maps $\phi(\cdot)$ are differentiable at $\mu[\mathbf{x}] \in \mathcal{P}_c(\mathbb{R}^d)$. We can therefore apply the first-order chainrule derived in Proposition 1.5 along tangent vectors to recover that

$$\lim_{s \rightarrow 0} \left[\frac{\phi(\mu[\mathbf{x} + s\mathbf{h}]) - \phi(\mu[\mathbf{x}])}{s} \right] = \mathcal{L}_{\xi_N} \phi(\mu[\mathbf{x}]) = \int_{\mathbb{R}^d} \langle \nabla_\mu \phi(\mu[\mathbf{x}](x)), \xi_N(x) \rangle d\mu[\mathbf{x}](x).$$

We can now conclude by recalling the definition of the symmetric maps $\phi_N(\cdot)$ given in (4.1) that

$$\lim_{s \rightarrow 0} \left[\frac{\phi_N(\mathbf{x} + s\mathbf{h}) - \phi_N(\mathbf{x})}{s} \right] = \phi'_N(\mathbf{x}; \mathbf{h}) = \frac{1}{N} \sum_{i=1}^N \langle \nabla_\mu \phi(\mu[\mathbf{x}](x_i)), h_i \rangle, \quad (4.11)$$

where we used (4.10) along with the fact that $\mu[\mathbf{x}] = \frac{1}{N} \sum_{i=1}^N \delta_{x_i}$. It is straightforward to check that the directional derivative $\mathbf{h} \mapsto \phi'_N(\mathbf{x}; \mathbf{h})$ of $\phi_N(\cdot)$ defined in (4.11) is a linear form and that it is continuous with respect to the rescaled Euclidean metric $|\cdot|_N$. We therefore obtain that $\phi_N(\cdot)$ is Fréchet differentiable at \mathbf{x} and that its differential can be represented in $((\mathbb{R}^d)^N, \langle \cdot, \cdot \rangle_N)$ by the mean-field gradient $\mathbf{Grad} \phi_N(\mathbf{x}) = (\nabla_\mu \phi(\mu[\mathbf{x}](x_i)))_{1 \leq i \leq N}$ as a consequence of Riesz's Theorem [38, Theorem 5.5].

Consider now two elements $\mathbf{h}^1, \mathbf{h}^2 \in (\mathbb{R}^d)^N$ and the corresponding tangent vectors $\xi_N^1, \xi_N^2 \in \nabla C_c^\infty(\mathbb{R}^d)$ built as in (4.9). Since the maps $\phi(\cdot)$ are twice differentiable in the sense of Definition 1.11, we can use the the second-order differentiation formula (1.26) to obtain that

$$\lim_{s \rightarrow 0} \left[\frac{\mathcal{L}_{\xi_N^1} \phi((\text{Id} + s\xi_N^2) \# \mu[\mathbf{x}]) - \mathcal{L}_{\xi_N^1} \phi(\mu[\mathbf{x}])}{s} \right] = \text{Hess} \phi[\mu[\mathbf{x}]](\xi_N^1, \xi_N^2) + \mathcal{L}_{D\xi_N^1 \xi_N^2} \phi(\mu[\mathbf{x}]). \quad (4.12)$$

Recall now that by construction (4.9) of $\xi_N^1(\cdot)$, it holds that $D\xi_N^1(x) = 0$ for $\mu[\mathbf{x}]$ -almost every $x \in \mathbb{R}^d$, so that consequentially $\mathcal{L}_{D\xi_N^1 \xi_N^2} \phi(\mu[\mathbf{x}]) = 0$. Furthermore by definition of the symmetric maps $\phi_N(\cdot)$ along with that of their mean-field gradients $\nabla \phi_N(\cdot)$,

equation (4.12) can be equivalently written as

$$\begin{aligned} \lim_{s \rightarrow 0} \left[\frac{\langle \nabla \phi_N(\mathbf{x} + s\mathbf{h}^2) - \nabla \phi_N(\mathbf{x}), \mathbf{h}^1 \rangle_N}{s} \right] &= \frac{1}{N} \sum_{i=1}^N \langle D_x \nabla_\mu \phi(\mu[\mathbf{x}])(x_i) h_i^1, h_i^2 \rangle \\ &+ \frac{1}{N^2} \sum_{i,j=1}^N \langle D_\mu^2 \phi(\mu[\mathbf{x}])(x_i, x_j) h_i^1, h_j^2 \rangle \end{aligned}$$

where we used the analytical expression (1.25) of the Wasserstein Hessian. We accordingly introduce the mean-field Hessian of $\phi_N(\cdot)$ at \mathbf{x} , defined by

$$\begin{aligned} \mathbf{Hess} \phi_N[\mathbf{x}](\mathbf{h}^1, \mathbf{h}^2) &= \mathbf{Hess} \phi[\mu[\mathbf{x}]](\xi_N^1, \xi_N^2) \\ &= \frac{1}{N} \sum_{i=1}^N \langle D_x \nabla_\mu \phi(\mu[\mathbf{x}])(x_i) h_i^1, h_i^2 \rangle \\ &+ \frac{1}{N^2} \sum_{i,j=1}^N \langle D_\mu^2 \phi(\mu[\mathbf{x}])(x_i, x_j) h_i^1, h_j^2 \rangle. \end{aligned} \tag{4.13}$$

It is again possible to verify that $\mathbf{Hess} \phi_N[\mathbf{x}](\cdot, \cdot)$ defines a continuous bilinear form with respect to the rescaled metric $|\cdot|_N$, so that the map $\phi_N(\cdot)$ is twice Fréchet differentiable over $(\mathbb{R}^d)^N$.

The Taylor expansion formula (4.4) can be derived directly by expanding $\phi_N(\mathbf{x} + \mathbf{h})$ using the classical Taylor theorem in $(\mathbb{R}^d)^N$ along with (4.11) and (4.13). Defining the C^2 -norm of a functional $\phi_N(\cdot)$ as in (4.8), it follows directly from the uniform bounds (4.2) stemming from the $C_{\text{loc}}^{2,1}$ -MF regularity of $\phi(\cdot)$ that for any compact set $K \subset \mathbb{R}^d$, there exists a constant $C_K > 0$ such that

$$\|\phi_N(\cdot)\|_{C^2(K^N)} + \text{Lip}(\mathbf{Hess} \phi_N[\cdot]; K^N) \leq C_K,$$

which ends the proof of our claim. \square

Remark 4.2 (Link between mean-field and classical derivatives). *The notion of first and second order mean-field derivatives for a functional $\phi_N(\cdot)$ that we introduced in Proposition 4.1 can be linked to the classical derivatives in $(\mathbb{R}^d)^N$ endowed with its standard Euclidean product as follows*

$$\begin{cases} \nabla_{x_i} \phi_N(\mathbf{x}) = \frac{1}{N} \mathbf{Grad}_{x_i} \phi_N(\mathbf{x}) = \frac{1}{N} \nabla_\mu \phi(\mu[\mathbf{x}])(x_i), \\ \nabla_{x_i, x_j}^2 \phi_N(\mathbf{x}) = \frac{1}{N^2} D_\mu^2 \phi(\mu[\mathbf{x}])(x_i, x_j), \\ \nabla_{x_i, x_i}^2 \phi_N(\mathbf{x}) = \frac{1}{N} D_x \nabla_\mu \phi(\mu[\mathbf{x}])(x_i) + \frac{1}{N^2} D_\mu^2 \phi(\mu[\mathbf{x}])(x_i, x_i) \end{cases} \tag{4.14}$$

for any pair of indices $i, j \in \{1, \dots, N\}$ with $i \neq j$. While the formula linking the classical and mean-field gradients can be recovered by simple rescaling arguments, the formulas involving second-order derivatives can only be derived rigorously as a discrete counterpart of the Wasserstein differential structure.

Remark 4.3 (Matrix representation of the mean-field Hessian in $(\mathbb{R}^d)^N$). *The rescaled inner product $\langle \cdot, \cdot \rangle_N$ defined over $(\mathbb{R}^d)^N$ in (4.3) induces a rescaled matrix-vector product*

given by

$$\mathbf{A}\mathbf{x} = \left(\frac{1}{N} \sum_{j=1}^N A_{ij}x_j \right)_{1 \leq i \leq N}$$

for any matrix $\mathbf{A} \in \mathbb{R}^{dN \times dN}$ and any vector $\mathbf{x} \in (\mathbb{R}^d)^N$. By Riesz Theorem applied in the Hilbert space $((\mathbb{R}^d)^N, \langle \cdot, \cdot \rangle_N)$ (see e.g. [38, Theorem 5.5]), it is possible to represent the action of the Hessian bilinear form $\mathbf{Hess} \phi_N[\mathbf{x}](\cdot, \cdot)$ by using introducing the Hessian matrix defined by

$$\mathbf{Hess} \phi_N[\mathbf{x}](\mathbf{h}^1, \mathbf{h}^2) = \langle \mathbf{Hess} \phi_N(\mathbf{x})\mathbf{h}^1, \mathbf{h}^2 \rangle_N, \quad (4.15)$$

for any $\mathbf{x}, \mathbf{h}^1, \mathbf{h}^2 \in (\mathbb{R}^d)^N$. Moreover, we know by equation (4.14) in Remark 4.2 that the components of $\mathbf{Hess} \phi_N(\mathbf{x})$ are given explicitly by

$$\begin{cases} (\mathbf{Hess} \phi_N(\mathbf{x}))_{i,j} = D_\mu^2 \phi(\mu[\mathbf{x}])(x_i, x_j), \\ (\mathbf{Hess} \phi_N(\mathbf{x}))_{i,i} = N D_x \nabla_\mu \phi(\mu[\mathbf{x}])(x_i) + D_\mu^2 \phi(\mu[\mathbf{x}])(x_i, x_j), \end{cases}$$

for any pair of indices $i, j \in \{1, \dots, N\}$ such that $i \neq j$.

4.1.2 Existence of mean-field optimal controls for problem (\mathcal{P})

In this section, we show how problem (\mathcal{P}) can be reformulated so as to encompass both the measure theoretic formulation and its sequence of approximating problems. We subsequently recall a powerful existence result derived in [80] for general multi-agent optimal control problems formulated in the Wasserstein space $(\mathcal{P}_1(\mathbb{R}^d), W_1)$. Its main feature is to show that under mild structural conditions, there exist optimal solutions for (\mathcal{P}) which can be recovered as weak limits in a suitable topology of sequences of optimal solutions for finite dimensional problems.

Let us start by fixing an integer $N \geq 1$, an initial datum $\mathbf{x}_N^0 \in (\mathbb{R}^d)^N$ and the associated discrete measure $\mu_N^0 = \mu[\mathbf{x}_N^0]$ as defined in Section 4.1.1. As already sketched in the introduction, we are naturally brought to consider the family of discrete problems

$$(\mathcal{P}_N) \quad \begin{cases} \min_{\mathbf{u}(\cdot) \in \mathcal{U}_N} \left[\int_0^T \left(\mathbf{L}_N(t, \mathbf{x}(t)) + \frac{1}{N} \sum_{i=1}^N \psi(u_i(t)) \right) dt + \varphi_N(\mathbf{x}(T)) \right] \\ \text{s.t.} \quad \begin{cases} \dot{x}_i(t) = \mathbf{v}_N[\mathbf{x}(t)](t, x_i(t)) + u_i(t), \\ x_i(0) = x_i^0, \end{cases} \end{cases}$$

where $\mathcal{U}_N = L^\infty([0, T], U^N)$ and where we introduced the mean-field approximating functionals

$$\mathbf{v}_N[\mathbf{x}](\cdot, \cdot) = v[\mu[\mathbf{x}]](\cdot, \cdot), \quad \mathbf{L}_N(\cdot, \mathbf{x}) = L(t, \mu[\mathbf{x}]), \quad \varphi_N(\mathbf{x}) = \varphi(\mu[\mathbf{x}]),$$

in the sense of Definition 4.1. It can be checked that as a consequence of hypotheses (\mathbf{H}) displayed in Theorem 4.1, the problems (\mathcal{P}_N) satisfy in particular the set of hypotheses (\mathbf{H}_{oc}) of Section 4.1.3. We can then deduce the following lemma.

Lemma 4.1 (Existence of solutions for problem (\mathcal{P}_N)). *Under hypotheses (\mathbf{H}) , there*

exist optimal pairs control-trajectory $(\mathbf{u}_N^*(\cdot), \mathbf{x}_N^*(\cdot)) \in L^\infty([0, T], U^N) \times \text{Lip}([0, T], (\mathbb{R}^d)^N)$ for (\mathcal{P}_N) for all $N \geq 1$.

Proof. This result is a direct consequence of Proposition 4.2. \square

We proceed by recasting problem (\mathcal{P}) into a framework which also encompasses the sequence of problems (\mathcal{P}_N) . Let us consider a narrowly continuous curve of measures $\mu(\cdot) \in C^0([0, T], \mathcal{P}_1(\mathbb{R}^d))$ and its canonical lift $\tilde{\mu} \in \mathcal{P}_1([0, T] \times \mathbb{R}^d)$. Recall that by Definition 1.1, a vector-valued measure $\nu \in \mathcal{M}([0, T] \times \mathbb{R}^d, U)$ is absolutely continuous with respect to $\tilde{\mu}$ if and only if there exists a map $u(\cdot, \cdot) \in L^1([0, T] \times \mathbb{R}^d, U; \tilde{\mu})$ such that $\nu = u(\cdot, \cdot)\tilde{\mu}$. Moreover the absolute continuity of ν with respect to $\tilde{\mu} = \int_{[0, T]} \mu(t) d\lambda(t)$ implies the existence of a λ -almost unique measurable family of measures $\{\nu(t)\}_{t \in [0, T]}$ such that

$$\nu = \int_{[0, T]} \nu(t) d\lambda(t).$$

in the sense of disintegration for vector-valued measures recalled in Theorem 1.2.

Bearing this in mind, problem (\mathcal{P}) can be relaxed as

$$(\mathcal{P}_{\text{meas}}) \begin{cases} \min_{\nu \in \mathcal{U}} \left[\int_0^T \left(L(t, \mu(t)) + \Psi(\nu(t)|\mu(t)) \right) dt + \varphi(\mu(T)) \right] \\ \text{s.t.} \begin{cases} \partial_t \mu(t) + \nabla \cdot (v[\mu(t)](t, \cdot)\mu(t) + \nu(t)) = 0, \\ \mu(0) = \mu^0. \end{cases} \end{cases}$$

where we denote the set of *generalized measure controls* by $\mathcal{U} = \mathcal{M}([0, T] \times \mathbb{R}^d, U)$ and where the map $\sigma \in \mathcal{M}(\mathbb{R}^d, U) \mapsto \Psi(\sigma|\mu) \in [0, +\infty]$ is defined by

$$\Psi(\sigma|\mu) = \begin{cases} \int_{\mathbb{R}^d} \psi \left(\frac{d\sigma}{d\mu}(x) \right) d\mu(x) & \text{if } \sigma \ll \mu, \\ +\infty & \text{otherwise.} \end{cases} \quad (4.16)$$

This functional can be furthermore lifted back to a functional $\tilde{\Psi}(\cdot|\tilde{\mu}) : \mathcal{M}([0, T] \times \mathbb{R}^d, U) \rightarrow [0, +\infty]$ as a consequence of the common disintegration of $\tilde{\mu}$ and ν onto λ . This type of relaxation appears frequently in variational problems involving integral functional on measures. Indeed, functionals of the form of $\Psi(\cdot|\tilde{\mu})$ as defined in (4.16) possess a wide range of useful features, such as weak-* lower-semicontinuity, while also imposing an absolute continuity property on the measure. We refer the reader to [17, Section 9.4] for a detailed account on their properties.

Consider now an optimal pair control-trajectory $(\mathbf{u}_N^*(\cdot), \mathbf{x}_N^*(\cdot)) \in \mathcal{U}_N \times \text{Lip}([0, T], (\mathbb{R}^d)^N)$ for (\mathcal{P}_N) . One can canonically associate to any such solution the discrete measure curve $\mu_N^*(\cdot) \in \text{Lip}([0, T], \mathcal{P}_N(\mathbb{R}^d))$ defined by

$$\mu_N^*(\cdot) = \mu[\mathbf{x}_N^*(\cdot)] = \frac{1}{N} \sum_{i=1}^N \delta_{x_i^*(\cdot)} \quad (4.17)$$

for all times $t \in [0, T]$, along with the discrete control-measure

$$\nu_N^* = \nu[\mathbf{x}_N^*(\cdot), \mathbf{u}_N^*(\cdot)] = \int_{[0, T]} \left(\frac{1}{N} \sum_{i=1}^N u_i^*(t) \delta_{x_i^*(t)} \right) d\lambda(t). \quad (4.18)$$

In the following theorem, we state a condensed version of the main result of [80] which shows that this relaxation allows to prove the convergence of the discrete problems (\mathcal{P}_N) towards (\mathcal{P}) . This convergence result has to be understood both in terms of mean-field limit of the functional describing the dynamics and of Γ -convergence of the corresponding minimizers.

Theorem 4.2 (Existence of mean-field optimal controls for (\mathcal{P})). *Let $\mu^0 \in \mathcal{P}_c(\mathbb{R}^d)$, $(\mu_N^0) \subset \mathcal{P}_c(\mathbb{R}^d)$ be a sequence of empirical measures associated with $(\mathbf{x}_N^0) \subset (\mathbb{R}^d)^N$ such that $W_1(\mu_N^0, \mu^0) \rightarrow 0$, and assume that hypotheses **(H)** hold. For any $N \geq 1$, denote by $(\mathbf{u}_N^*(\cdot), \mathbf{x}_N^*(\cdot)) \in \mathcal{U}_N \times \text{Lip}([0, T], (\mathbb{R}^d)^N)$ an optimal pair control-trajectory for (\mathcal{P}_N) and by $(\nu_N^*, \mu_N^*(\cdot)) \in \mathcal{U} \times \text{Lip}([0, T], \mathcal{P}_N(\mathbb{R}^d))$ the corresponding pair of measure control-trajectory defined as in (4.17)-(4.18).*

Then, there exists $(\nu^, \mu^*(\cdot)) \in \mathcal{U} \times \text{Lip}([0, T], \mathcal{P}_c(\mathbb{R}^d))$ such that*

$$\nu_N^* \xrightarrow{N \rightarrow +\infty} \nu^* \quad \text{and} \quad \sup_{t \in [0, T]} W_1(\mu_N^*(t), \mu^*(t)) \xrightarrow{N \rightarrow +\infty} 0,$$

along a suitable subsequence. Moreover, the classical pair control-trajectory $(\frac{d\nu^}{d\mu^*}(\cdot, \cdot), \mu^*(\cdot)) \in L^\infty([0, T] \times \mathbb{R}^d, U; \tilde{\mu}) \times \text{Lip}([0, T], \mathcal{P}_c(\mathbb{R}^d))$ is optimal for problem (\mathcal{P}) .*

4.1.3 Existence of locally optimal Lipschitz feedbacks in finite-dimensional optimal control problems

In this section, we recall some classical facts about finite-dimensional optimal control problems, and we describe in Theorem 4.3 a result described in [71] providing sufficient conditions for the existence of locally optimal Lipschitz feedbacks in a neighbourhood of an optimal trajectory. The latter is based on deep metric regularity properties for dynamical differential inclusions explored recently in [58]. Throughout this section, we consider the finite-dimensional optimal control problem

$$(\mathcal{P}_{\text{oc}}) \quad \begin{cases} \min_{u \in \mathcal{U}} \left[\int_0^T (l(t, x(t)) + \psi(u(t))) dt + g(x(T)) \right] \\ \text{s.t.} \quad \begin{cases} \dot{x}(t) = f(t, x(t)) + u(t), \\ x(0) = x^0 \in \mathbb{R}^d, \end{cases} \end{cases}$$

under the following structural assumptions.

(H_{oc})

1. The set of admissible controls is defined by $\mathcal{U} = L^\infty([0, T], U)$ where $U \subset \mathbb{R}^d$ is a compact and convex set.

2. The control cost $u \mapsto \psi(u)$ is $C_{\text{loc}}^{2,1}$ -regular and strictly convex.
3. The map $(t, x) \mapsto f(t, x)$ is Lipschitz with respect to $t \in [0, T]$, sublinear and $C_{\text{loc}}^{2,1}$ -regular with respect to $x \in \mathbb{R}^d$.
4. The running cost $(t, x) \mapsto l(t, x)$ is Lipschitz with respect to $t \in [0, T]$ and $C_{\text{loc}}^{2,1}$ -regular with respect to $x \in \mathbb{R}^d$. Similarly, the final cost $x \mapsto g(x)$ is $C_{\text{loc}}^{2,1}$ -regular.

As a direct consequence of our regularity hypotheses and of the compactness of the set of admissible control values U , we can derive a uniform compactness estimate on the admissible trajectories which we state in the following lemma.

Lemma 4.2 (Uniform compactness of admissible trajectories). *There exists a compact set $K \subset \mathbb{R}^d$ such that any admissible curve $x(\cdot)$ for $(\mathcal{P}_{\text{oc}})$ associated with a control map $u(\cdot)$ satisfies $x(\cdot) \in \text{Lip}([0, T], K)$.*

The proof of this result is a direct consequence of Grönwall's Lemma. From now on, we fix such a compact set $K \subset \mathbb{R}^d$.

Proposition 4.2 (Existence of solutions for problem $(\mathcal{P}_{\text{oc}})$). *Under hypotheses (\mathbf{H}_{oc}) , there exists an optimal pair control-trajectory $(u^*(\cdot), x^*(\cdot)) \in L^\infty([0, T], U) \times \text{Lip}([0, T], K)$ for problem $(\mathcal{P}_{\text{oc}})$.*

This result is standard in optimal control theory under our working hypotheses and can be found e.g. in [59, Theorem 23.11]). We can further define the *Hamiltonian* associated to $(\mathcal{P}_{\text{oc}})$ by

$$H : (t, x, p, u) \in [0, T] \times (\mathbb{R}^d)^3 \mapsto \langle p, f(t, x) + u \rangle - (l(t, x) + \psi(u)).$$

Given an optimal pair control-trajectory $(u^*(\cdot), x^*(\cdot))$ for $(\mathcal{P}_{\text{oc}})$, there exists by Pontryagin's Maximum Principle (see e.g. [59, Theorem 22.2]) a curve $p^*(\cdot) \in \text{Lip}([0, T], \mathbb{R}^d)$ such that the couple $(x^*(\cdot), p^*(\cdot))$ is a solution of the *forward-backward Hamiltonian system*

$$\begin{cases} \dot{x}^*(t) = \nabla_p H(t, x^*(t), p^*(t), u^*(t)), & x^*(0) = x^0, \\ \dot{p}^*(t) = -\nabla_x H(t, x^*(t), p^*(t), u^*(t)), & p^*(T) = -\nabla g(x^*(T)). \end{cases} \quad (4.19)$$

Moreover, the *Pontryagin maximization condition*

$$H(t, x^*(t), p^*(t), u^*(t)) = \max_{v \in U} [H(t, x^*(t), p^*(t), v)], \quad (4.20)$$

holds along this extremal pair for \mathcal{L}^1 -almost every $t \in [0, T]$. Such a collection of optimal control, state and costate $(x^*(\cdot), p^*(\cdot), u^*(\cdot))$ is called an *optimal Pontryagin triple* for $(\mathcal{P}_{\text{oc}})$. Let it be noted that since the problem $(\mathcal{P}_{\text{oc}})$ is unconstrained, there are no abnormal curves stemming from the maximum principle.

Remark now that, as a by-product of the local Lipschitz regularity of $f(\cdot, \cdot)$, $l(\cdot, \cdot)$ and $g(\cdot)$, there exists a compact set $K' \subset \mathbb{R}^d$ such that any covector $p(\cdot)$ associated with an admissible pair $(u(\cdot), x(\cdot))$ via (4.19) satisfies $p \in \text{Lip}([0, T], K')$. We henceforth denote

by $\mathcal{K} = [0, T] \times K \times K' \times U$ the uniform compact set containing the admissible times, states, costates and controls for $(\mathcal{P}_{\text{oc}})$. Moreover, we denote by $\mathcal{L}_{\mathcal{K}}$ the Lipschitz constant over \mathcal{K} of the maps $f(\cdot, \cdot)$, $l(\cdot, \cdot)$, $\psi(\cdot)$ and $H(\cdot, \cdot, \cdot, \cdot)$ and of their derivatives with respect to the variables (x, u) up to the second order.

We now present the central and somewhat less standard assumption which allows for the construction of locally optimal feedbacks in a neighbourhood of $\text{Graph}(x^*(\cdot))$.

Definition 4.3 (Uniform coercivity property). *We say that an optimal Pontryagin triple $(x^*(\cdot), p^*(\cdot), u^*(\cdot))$ for $(\mathcal{P}_{\text{oc}})$ satisfies the uniform coercivity property with constant $\rho > 0$ if the following estimate holds*

$$\begin{aligned} \langle \nabla_x^2 g(x^*(T))y(T), y(T) \rangle - \int_0^T \langle \nabla_x^2 H(t, x^*(t), p^*(t), u^*(t))y(t), y(t) \rangle dt \\ - \int_0^T \langle \nabla_u^2 H(t, x^*(t), p^*(t), u^*(t))w(t), w(t) \rangle dt \\ \geq \rho \int_0^T |w(t)|^2 dt \end{aligned} \quad (4.21)$$

for any pair of maps $(y(\cdot), w(\cdot)) \in W^{1,2}([0, T], \mathbb{R}^d) \times L^2([0, T], \mathbb{R}^d)$ which satisfy the linearized control-state equation

$$\dot{y}(t) = D_x f(t, x^*(t))y(t) + w(t), \quad y(0) = 0, \quad (4.22)$$

along with the compatibility condition $u^*(t) + w(t) \in U$ for \mathcal{L}^1 -almost every $t \in [0, T]$.

As we shall see later on, the uniform coercivity estimate (4.21) can be interpreted as a strong positive-definiteness condition for the linearization of $(\mathcal{P}_{\text{oc}})$ in a neighbourhood of $(x^*(\cdot), p^*(\cdot), u^*(\cdot))$. It was first introduced in this form in [90], and further used in [70] to derive regularity properties of optimal control. Let it be noted that in [58, 70, 71, 90], the authors define the Hamiltonian of the problem as $\tilde{H}(t, x, p, u) = \langle p, f(t, x) + u \rangle + l(t, x) + \psi(u)$. Hence, some signs differ between their condition and (4.21), but the results are identical.

We can now state main result of this section which can be found in [71, Theorem 5.2].

Theorem 4.3 (Existence of locally optimal feedbacks for $(\mathcal{P}_{\text{oc}})$). *Let $(x^*(\cdot), p^*(\cdot), u^*(\cdot)) \in \text{Lip}([0, T], K) \times \text{Lip}([0, T], K') \times \mathcal{U}$ be an optimal Pontryagin triple for problem $(\mathcal{P}_{\text{oc}})$. Suppose that hypotheses (\mathbf{H}_{oc}) hold and that $(x^*(\cdot), p^*(\cdot), u^*(\cdot))$ satisfies the uniform coercivity estimate (4.21)-(4.22) with constant $\rho > 0$.*

Then, there exists a representative in the L^∞ -equivalence class of $u^(\cdot)$ such that the maximization condition (4.20) holds for all times $t \in [0, T]$. There further exists constants $\epsilon, \eta > 0$, an open subset $\mathcal{N} \subset [0, T] \times \mathbb{R}^d$ and a map $\bar{u}(\cdot, \cdot) \in \text{Lip}(\mathcal{N}, \mathbb{R}^d)$ which Lipschitz constant depends only on ρ and \mathcal{L}_K , such that the following conditions hold.*

- (i) $\bar{u}(\cdot, x^*(\cdot)) = u^*(\cdot)$.
- (ii) $(\text{Graph}(x^*(\cdot)) + \{0\}) \times B(0, \epsilon) \subset \mathcal{N}$.
- (iii) For every $(\tau, \xi) \in \mathcal{N}$, the equation

$$\dot{x}(t) = f(t, x(t)) + \bar{u}(t, x(t)), \quad x(\tau) = \xi, \quad (4.23)$$

has a unique solution $\hat{x}_{(\tau, \xi)}(\cdot)$ such that $\text{Graph}(\hat{x}_{(\tau, \xi)}(\cdot)) \subset \mathcal{N}$.

(iv) The map $\hat{u}_{(\tau,\xi)} : t \in [\tau, T] \mapsto \bar{u}(t, \hat{x}_{(\tau,\xi)}(t))$ is such that

$$\|u^*(\cdot) - \hat{u}_{(\tau,\xi)}(\cdot)\|_{L^\infty([\tau, T], \mathbb{R}^d)} \leq \eta,$$

and

$$\int_{\tau}^T l(t, \hat{x}_{(\tau,\xi)}(t), \hat{u}_{(\tau,\xi)}(t)) dt + g(\hat{x}_{(\tau,\xi)}(T)) \leq \int_{\tau}^T l(t, x(t), u(t)) dt + g(x(T))$$

among all the admissible open loop pairs $(u(\cdot), x(\cdot)) \in \mathcal{U} \times \text{Lip}([\tau, T], \mathbb{R}^d)$ solving (4.23) and such that $\|u(\cdot) - \hat{u}_{(\tau,\xi)}(\cdot)\|_{L^\infty([\tau, T], \mathbb{R}^d)} \leq \eta$.

The statements of Theorem 4.3 can be heuristically summed up as follows. As a consequence of the uniform coercivity condition, there exists a non-empty neighbourhood \mathcal{N} of the graph of the optimal trajectory $x^*(\cdot)$ on which it is possible to define a locally optimal feedback $\bar{u}(\cdot, \cdot)$. Here, local optimality is understood in the sense that the closed-loop system (4.23) generated by $\bar{u}(\cdot, \cdot)$ starting from any point $\xi \in \pi_{\mathbb{R}^d}(\mathcal{N})$ produces a lower cost than any admissible open-loop control. This locally optimal map $\bar{u}(\cdot, \cdot)$ can moreover be defined in such a way that $\bar{u}(\cdot, x^*(\cdot)) = u^*(\cdot)$, i.e. $\bar{u}(\cdot, \cdot)$ coincides with the optimal open-loop control $u^*(\cdot)$ when evaluated along the corresponding optimal trajectory $x^*(\cdot)$.

To better illustrate our subsequent use of this result in the proof of Theorem 4.1, we provide here an overview of the strategy used to prove Theorem 4.3 in [71], based on the earlier work [58]. We start our heuristic exposition by recalling the concept of *strong metric regularity* for a multi-function.

Definition 4.4 (Strong metric regularity). *Let \mathcal{Y}, \mathcal{Z} be two Banach spaces. A multi-valued mapping $\mathcal{G} : Y \rightrightarrows Z$ is said to be strongly metrically regular at $y^* \in Y$ for $z^* \in Z$ if $z^* \in \mathcal{F}(y^*)$ and if there exists $a, b > 0$ and $\kappa \geq 0$ such that*

$$\mathcal{G}^{-1} : B(z^*, b) \rightarrow B(y^*, a)$$

is single-valued and κ -Lipschitz.

We start by fixing a time $\tau \in [0, T]$. In (4.19)-(4.20), we wrote the Pontryagin maximum principle for $(\mathcal{P}_{\text{oc}})$. Since $v \in U \mapsto H(t, x^*(t), p^*(t), v)$ is differentiable, we can reformulate the maximization condition (4.20) as

$$\nabla_u H(t, x^*(t), p^*(t), u^*(t)) \in N_U(u^*(t)),$$

for all times $t \in [0, T]$, where $N_U(v)$ denotes the normal cone of convex analysis to U at v . Then, any optimal Pontryagin triple $(x^*(\cdot), p^*(\cdot), u^*(\cdot))$ can be seen as a solution of the *differential generalized equation*

$$0 \in F_\tau(x(\cdot), p(\cdot), u(\cdot)) + G_\tau(x(\cdot), p(\cdot), u(\cdot)) \quad (4.24)$$

where the maps $F_\tau : \mathcal{Y}_\tau \rightarrow \mathcal{Z}_\tau$ and $G_\tau : \mathcal{Y}_\tau \rightrightarrows \mathcal{Z}_\tau$ are defined by

$$F_\tau(x(\cdot), p(\cdot), u(\cdot)) = \begin{pmatrix} \dot{x}(\cdot) - f(\cdot, x(\cdot)) - u(\cdot) \\ x(\tau) - x^*(\tau) \\ \dot{p}(t) + \nabla_x H(\cdot, x(\cdot), p(\cdot), u(\cdot)) \\ p(T) + \nabla g(x(T)) \\ -\nabla_u H(\cdot, x(\cdot), p(\cdot), u(\cdot)) \end{pmatrix},$$

and

$$G_\tau(x(\cdot), p(\cdot), u(\cdot)) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ N_{\mathcal{U}}^\infty(u(\cdot)) \end{pmatrix}.$$

Here, we introduced the two Banach spaces

$$\begin{cases} \mathcal{Y}_\tau = W^{1,\infty}([\tau, T], \mathbb{R}^d) \times W^{1,\infty}([\tau, T], \mathbb{R}^d) \times L^\infty([\tau, T], U), \\ \mathcal{Z}_\tau = L^\infty([\tau, T], \mathbb{R}^d) \times \mathbb{R}^d \times L^\infty([\tau, T], \mathbb{R}^d) \times \mathbb{R}^d \times L^\infty([\tau, T], \mathbb{R}^d). \end{cases}$$

as well as the set $N_{\mathcal{U}}^\infty(u(\cdot)) = \{v \in L^\infty([0, T], U) \text{ s.t. } v(t) \in N_U(u(t)) \text{ for } \mathcal{L}^1\text{-almost every } t \in [0, T]\}$.

In [71], it is proven that Theorem 4.3 can be derived as a consequence of the *strong metric regularity* of $F_\tau(\cdot) + G_\tau(\cdot)$ at the restriction to $[\tau, T]$ of the Pontryagin triple $(x^*(\cdot), p^*(\cdot), u^*(\cdot))$ for 0, uniformly with respect to τ . A standard strategy for proving metric regularity of mappings of the form of $F(\cdot) + G(\cdot)$ where $F(\cdot)$ is Fréchet-differentiable, is to apply the *Robinson's inverse function theorem*, which states the following fact.

Theorem 4.4 (Robinson's inverse function theorem). *Let $y^* \in \mathcal{Y}$ and $z^* \in \mathcal{G}(y^*)$. Suppose that $\mathcal{F} : \mathcal{Y} \rightarrow \mathcal{Z}$ is Fréchet differentiable at y^* . Then, the multi-valued mapping $y \in \mathcal{Y} \mapsto \mathcal{F}(y) + \mathcal{G}(y)$ is strongly metrically regular at y^* for $\mathcal{F}(y^*) + z^*$ if and only if the partially linearized mapping $y \mapsto \mathcal{F}(y^*) + D\mathcal{F}(y^*)(y - y^*) + \mathcal{G}(y^*)$ is strongly metrically regular at y^* for $\mathcal{F}(y^*) + z^*$.*

The strong metric regularity of (4.24) can therefore be equivalently derived from that of its partial linearization involving the Fréchet differential of $F_\tau(\cdot)$, which is given by

$$DF_\tau(x^*(\cdot), p^*(\cdot), u^*(\cdot))(y(\cdot), q(\cdot), w(\cdot)) = \begin{pmatrix} \dot{y}(t) - D_x f(\cdot, x^*(\cdot), u^*(\cdot))y(\cdot) - w(\cdot) \\ y(\tau) \\ \dot{q}(\cdot) + \nabla_x^2 H(\cdot, x^*(\cdot), p^*(\cdot), u^*(\cdot))y(\cdot) + \nabla_{px}^2 H(\cdot, x^*(\cdot), p^*(\cdot), u^*(\cdot))q(\cdot) \\ q(T) + \nabla_x^2 g(x^*(T))y(T) \\ -\nabla_u^2 H(\cdot, x^*(\cdot), p^*(\cdot), u^*(\cdot))w(\cdot) - \nabla_{pu}^2 H(\cdot, x^*(\cdot), p^*(\cdot), u^*(\cdot))q(\cdot) \end{pmatrix}. \quad (4.25)$$

Notice that since in our problem the control and state are decoupled, there are no crossed derivatives in (x, u) . Now, the key point is to remark that the partially linearized generalized differential inclusion

$$0 \in DF_\tau(x^*(\cdot), p^*(\cdot), u^*(\cdot))(y(\cdot), q(\cdot), w(\cdot)) + G_\tau(y(\cdot), q(\cdot), w(\cdot))$$

can be equivalently seen as the Pontryagin maximum principle for the linear-quadratic optimal control problem

$$\begin{cases} \min_{w(\cdot) \in \mathcal{U}_\tau} \left[\int_\tau^T \left(\frac{1}{2} \langle A(t)y(t), y(t) \rangle + \frac{1}{2} \langle B(t)w(t), w(t) \rangle \right) dt + \frac{1}{2} \langle C(T)y(T), y(T) \rangle \right] \\ \text{s.t.} \begin{cases} \dot{y}(t) = Df(t, x^*(t))y(t) + w(t), \\ y(\tau) = 0, \end{cases} \end{cases} \quad (4.26)$$

where $\mathcal{U}_\tau = \{v \in L^2([\tau, T], U) \text{ s.t. } u^*(t) + v(t) \in U \text{ for } \mathcal{L}^1\text{-almost every } t \in [\tau, T]\}$ and

$$\begin{cases} A(t) = -\nabla_x^2 H(t, x^*(t), p^*(t), u^*(t)), \\ B(t) = -\nabla_u^2 H(t, x^*(t), p^*(t), u^*(t)), \\ C(T) = \nabla_x^2 g(x^*(T)). \end{cases}$$

The coercivity estimate (4.21)-(4.22) is still valid on $[\tau, T]$ up to choosing $w(\cdot) \equiv 0$ on $[0, \tau]$, and one can see that it indeed is a second-order strict positive-definiteness condition for the linearized problem (4.26). It was proven in [90] that under the uniform coercivity estimate (4.21), the solutions of (4.26) are unique and $\frac{1}{\rho}$ -Lipschitz with respect to the data of the problem. In [58], it was proven that by applying Robinson's inverse function theorem, one can then recover the strong metric regularity of (4.24) uniformly with respect to τ , which in turn yields Theorem 4.3 in [71] after some extra efforts.

4.2 Proof of Theorem 4.1

In this section, we prove the main result of this article stated in Theorem 4.1. We suppose that hypotheses (H) of Theorem 4.1 hold, along with the following additional *mean-field coercivity* assumption.

Hypothesis (CO_N)

There exists a constant $\rho_T > 0$ such that for every $(\mathbf{w}(\cdot), \mathbf{y}(\cdot)) \in L^2([0, T], (\mathbb{R}^d)^N) \times W^{1,2}([0, T], (\mathbb{R}^d)^N)$ solution of the linearized control-to-state equations

$$\begin{cases} \dot{y}_i(t) = D_x \mathbf{v}_N[\mathbf{x}^*(t)](t, x_i^*(t))y_i(t) + \frac{1}{N} \sum_{j=1}^N D_{x_j} \mathbf{v}_N[\mathbf{x}^*(t)](t, x_i^*(t))y_j(t) + w_i(t), \\ y_i(0) = 0, \quad \mathbf{u}_N^*(t) + \mathbf{w}(t) \in U^N \text{ for } \mathcal{L}^1\text{-almost every } t \in [0, T], \end{cases}$$

the following uniform mean-field coercivity estimate

$$\begin{aligned} \mathbf{Hess} \varphi_N[\mathbf{x}_N^*(T)](\mathbf{y}(T), \mathbf{y}(T)) &- \int_0^T \mathbf{Hess}_x \mathbb{H}_N[t, \mathbf{x}_N^*(t), \mathbf{r}_N^*(t), \mathbf{u}_N^*(t)](\mathbf{y}(t), \mathbf{y}(t)) dt \\ &- \int_0^T \mathbf{Hess}_u \mathbb{H}_N[t, \mathbf{x}_N^*(t), \mathbf{r}_N^*(t), \mathbf{u}_N^*(t)](\mathbf{w}(t), \mathbf{w}(t)) dt \\ &\geq \rho_T \int_0^T |\mathbf{w}(t)|_N^2 dt \end{aligned}$$

holds along any mean-field optimal Pontryagin triple $(\mathbf{x}_N^*(\cdot), \mathbf{r}_N^*(\cdot), \mathbf{u}_N^*(\cdot))$ for (\mathcal{P}_N) .

We set this hypothesis apart from the more classical assumptions **(H)** to emphasize that it is both original and new in the setting of mean-field optimal control, as well as instrumental in the proof of Theorem 4.1.

The argument is split into three main steps. In Step 1, we write a Pontryagin Maximum Principle adapted to the mean-field structure of the problem (\mathcal{P}_N) . We proceed by building in Step 2 a sequence of Lipschitz-in-space optimal control maps for the discrete problems (\mathcal{P}_N) by leveraging Theorem 4.3. We then show in Step 3 that this sequence of Lipschitz control maps is compact in a suitable weak topology preserving its regularity in space, and that its limit points coincide with the mean-field optimal control introduced in Theorem 4.2.

Step 1 : Solutions of (\mathcal{P}_N) and mean-field Pontryagin Maximum Principle

In this first step, we derive uniform characterizations and estimates on the optimal pairs $(\mathbf{u}_N^*(\cdot), \mathbf{x}_N^*(\cdot))$ for (\mathcal{P}_N) . Our analysis is based on the finite-dimensional Pontryagin maximum principle applied to $(\mathbb{R}^d)^N$ and written as a Hamiltonian flow with respect to the rescaled mean-field inner product $\langle \cdot, \cdot \rangle_N$.

Proposition 4.3 (Characterization of the solutions of (\mathcal{P}_N)). *Let $N \geq 1$ and $(\mathbf{u}_N^*(\cdot), \mathbf{x}_N^*(\cdot)) \in L^\infty([0, T], U^N) \times \text{Lip}([0, T], (\mathbb{R}^d)^N)$ be an optimal pair control-trajectory for (\mathcal{P}_N) . Then, there exists a rescaled covector $\mathbf{r}_N^*(\cdot) \in \text{Lip}([0, T], (\mathbb{R}^d)^N)$ such that $(\mathbf{x}_N^*(\cdot), \mathbf{r}_N^*(\cdot), \mathbf{u}_N^*(\cdot))$ satisfies the mean-field Pontryagin Maximum Principle*

$$\begin{cases} \dot{\mathbf{x}}_N^*(t) = \mathbf{Grad}_{\mathbf{r}} \mathbb{H}_N(t, \mathbf{x}_N^*(t), \mathbf{r}_N^*(t), \mathbf{u}_N^*(t)), & \mathbf{x}_N^*(0) = \mathbf{x}_N^0, \\ \dot{\mathbf{r}}_N^*(t) = -\mathbf{Grad}_{\mathbf{x}} \mathbb{H}_N(t, \mathbf{x}_N^*(t), \mathbf{r}_N^*(t), \mathbf{u}_N^*(t)), & \mathbf{r}_N^*(T) = -\mathbf{Grad}_{\mathbf{x}} \varphi_N(\mathbf{x}_N^*(T)), \\ \mathbf{u}_N^*(t) \in \underset{\mathbf{v} \in U^N}{\text{argmax}} \mathbb{H}_N(t, \mathbf{x}_N^*(t), \mathbf{r}_N^*(t), \mathbf{v}) & \text{for } \mathcal{L}^1\text{-almost every } t \in [0, T], \end{cases} \quad (4.27)$$

where the mean-field Hamiltonian $\mathbb{H}_N(\cdot, \cdot, \cdot, \cdot)$ of the system is defined by

$$\mathbb{H}_N(t, \mathbf{x}, \mathbf{r}, \mathbf{u}) = \frac{1}{N} \sum_{i=1}^N \left(\langle r_i, \mathbf{v}_N[\mathbf{x}](t, x_i) + u_i \rangle - \psi(u_i) \right) - \mathbf{L}_N(t, \mathbf{x}) \quad (4.28)$$

for all $(t, \mathbf{x}, \mathbf{r}, \mathbf{u}) \in [0, T] \times (\mathbb{R}^d)^N \times (\mathbb{R}^d)^N \times U^N$. Furthermore, there exists uniform constants $R_T, L_T > 0$ which are independent from N , such that

$$\text{Graph}\left((\mathbf{x}^*(\cdot), \mathbf{r}^*(\cdot))\right) \subset [0, T] \times \overline{\mathbf{B}(0, R_T)^{2N}} \quad , \quad \text{Lip}\left((\mathbf{x}^*(\cdot), \mathbf{r}^*(\cdot)); [0, T]\right) \leq L_T, \quad (4.29)$$

Proof. By an application of the standard PMP to (\mathcal{P}_N) (see for instance [59, Theorem 22.2]), there exists a family of costate variables $\mathbf{p}^*(\cdot) = (p_i^*(\cdot))_{i=1}^N \in \text{Lip}([0, T], (\mathbb{R}^d)^N)$

such that

$$\begin{cases} \dot{x}_i^*(t) = \nabla_{p_i} \mathcal{H}_N(t, \mathbf{x}^*(t), \mathbf{p}^*(t), \mathbf{u}^*(t)), & x_i^*(0) = x_i^0, \\ \dot{p}_i^*(t) = -\nabla_{x_i} \mathcal{H}_N(t, \mathbf{x}^*(t), \mathbf{p}^*(t), \mathbf{u}^*(t)), & p_i^*(T) = -\nabla_{x_i} \varphi_N(\mathbf{x}^*(T)), \\ u_i^*(t) \in \operatorname{argmax}_{v \in U} [\langle p_i^*(t), v \rangle - \frac{1}{N} \psi(v)]. \end{cases} \quad (4.30)$$

Here, the classical Hamiltonian $\mathcal{H}_N(\cdot, \cdot, \cdot, \cdot)$ of the system is defined by

$$\mathcal{H}_N(t, \mathbf{x}, \mathbf{p}, \mathbf{u}) = \sum_{i=1}^N \langle p_i, \mathbf{v}_N[\mathbf{x}](t, x_i) + u_i \rangle - \frac{1}{N} \sum_{i=1}^N \psi(u_i) - \mathbf{L}_N(t, \mathbf{x}),$$

for every $(t, \mathbf{x}, \mathbf{p}, \mathbf{u}) \in [0, T] \times (\mathbb{R}^d)^N \times (\mathbb{R}^d)^N \times U$. By introducing the rescaled variables $r_i^*(\cdot) = N p_i^*(\cdot)$, one can check that

$$\dot{x}_i^*(t) = N \nabla_{r_i} \mathbb{H}_N(t, \mathbf{x}^*(t), \mathbf{r}^*(t), \mathbf{u}^*(t)) = \mathbf{Grad}_{r_i} \mathbb{H}_N(t, \mathbf{x}^*(t), \mathbf{r}^*(t), \mathbf{u}^*(t)), \quad (4.31)$$

as well as

$$\dot{r}_i^*(t) = -N \nabla_{x_i} \mathbb{H}_N(t, \mathbf{x}^*(t), \mathbf{r}^*(t), \mathbf{u}^*(t)) = -\mathbf{Grad}_{x_i} \mathbb{H}_N(t, \mathbf{x}^*(t), \mathbf{r}^*(t), \mathbf{u}^*(t)). \quad (4.32)$$

and

$$r_i^*(T) = -N \nabla_{x_i} \varphi(\mathbf{x}^*(T)) = -\mathbf{Grad}_{x_i} \varphi(\mathbf{x}^*(T)). \quad (4.33)$$

as a consequence of Proposition 4.1 and Remark 4.2. Moreover, it can be seen easily from the maximization condition in (4.30) that $u_i^*(t) \in \operatorname{argmax}[\langle r_i^*(t), v \rangle - \psi(v)]$. Merging this condition with (4.31), (4.32) and (4.33), we recover the desired claim that $(\mathbf{x}^*(\cdot), \mathbf{r}^*(\cdot), \mathbf{u}^*(\cdot))$ satisfies the mean-field Pontryagin Maximum Principle (4.27) associated to the mean-field Hamiltonian $\mathbb{H}_N(\cdot, \cdot, \cdot, \cdot)$ for all times $t \in [0, T]$.

In the spirit of [31, 81], we introduce the discrete L^∞ -type function

$$X_N : t \in [0, T] \mapsto \max_{i \in \{1, \dots, N\}} |x_i^*(t)|$$

By Danskin's Theorem (see e.g. [65]), the map $X_N(\cdot)$ is differentiable \mathcal{L}^1 -almost everywhere and it holds that

$$\begin{aligned} X_N(t) X_N'(t) &= \frac{d}{dt} \left[\frac{1}{2} X_N^2(t) \right] \leq \langle x_{I(t)}^*(t), \dot{x}_{I(t)}^*(t) \rangle \\ &= \langle x_{I(t)}^*(t), \mathbf{v}_N[\mathbf{x}_N^*(t)](t, x_{I(t)}^*(t)) + u_{I(t)}^*(t) \rangle \\ &\leq |x_{I(t)}^*(t)| \left(M \left(1 + |x_{I(t)}^*(t)| + |\mathbf{x}_N^*(t)|_N \right) + L_U \right) \end{aligned}$$

by **(H1)**, **(H3)** and Cauchy-Schwarz inequality. Here, $I(t) \in \operatorname{argmax}_{i \in \{1, \dots, N\}} |x_i^*(t)|$ is any of the indices realizing the value of $X_N(t)$ for \mathcal{L}^1 -almost every $t \in [0, T]$. Remarking now that $|\mathbf{x}_N^*(t)|_N \leq X_N(t)$, we recover that

$$X_N'(t) \leq L_U + M(1 + 2X_N(t))$$

so that by Grönwall Lemma, there exists a constant $R_T^1 > 0$ depending only on $\operatorname{supp}(\mu^0)$,

T , M and L_U such that

$$\sup_{t \in [0, T]} |x_i^*(t)| \leq R_T^1, \quad (4.34)$$

for all $i \in \{1, \dots, N\}$. Plugging this uniform bound into (4.31), we recover the existence of a uniform constant $L_T^1 > 0$ such that

$$\text{Lip}(x_i^*(\cdot); [0, T]) \leq L_T^1, \quad (4.35)$$

for all $i \in \{1, \dots, N\}$.

We now prove a similar estimate on the costate variable $(\mathbf{r}_N^*(\cdot))$. By invoking the $C_{\text{loc}}^{2,1}$ -MF regularity assumptions of **(H4)**-**(H6)** as well as the uniform bound (4.34)-(4.35), we can derive by a similar application of Grönwall Lemma that

$$\sup_{t \in [0, T]} |r_i^*(t)| \leq C' \left(T + |\mathbf{Grad}_{x_i} \varphi_N(\mathbf{x}_N^*(T))| \right) e^{C'T} \quad (4.36)$$

for all $i \in \{1, \dots, N\}$ where $C' > 0$ is a given uniform constant, independent from N . By hypothesis **(H6)**, we know that $\varphi_N(\cdot)$ is locally Lipschitz over $(\mathbb{R}^d)^N$ with a uniform constant on products of compact sets, so that

$$\sup_{t \in [0, T]} |r_i^*(t)| \leq R_T^2, \quad \text{Lip}(r_i^*(\cdot); [0, T]) \leq L_T^2, \quad (4.37)$$

for all $i \in \{1, \dots, N\}$ and for some positive constants $R_T^2, L_T^2 > 0$. Subsequently, there exists uniform constants $R_T, L_T > 0$ which are again independent from N , such that

$$\text{Graph}\left((\mathbf{x}^*(\cdot), \mathbf{r}^*(\cdot))\right) \subset [0, T] \times \overline{B(0, R_T)^{2N}}, \quad \text{Lip}\left((\mathbf{x}^*(\cdot), \mathbf{r}^*(\cdot)); [0, T]\right) \leq L_T,$$

which concludes the proof of our claim. \square

We end this first step of our proof by a simple corollary in which we provide a common Lipschitz constant for all the maps involved in (\mathcal{P}_N) that is uniform with respect to N .

Corollary 4.1. *Let $\mathcal{K} = [0, T] \times \overline{B(0, R_T)^{2N}} \times U^N$ where $R_T > 0$ is defined as in (4.29). Then, there exists a constant $\mathcal{L}_{\mathcal{K}} > 0$ such that the $C^{2,1}$ -norms of the maps $\mathbb{H}_N(t, \cdot, \mathbf{r}, \cdot)$, $L_N(t, \cdot)$, $\frac{1}{N} \sum_{i=1}^N \psi(\cdot)$ and $\varphi_N(\cdot)$ with respect to the variables (\mathbf{x}, \mathbf{u}) are bounded by $\mathcal{L}_{\mathcal{K}}$ over \mathcal{K} , uniformly with respect to $(t, \mathbf{r}) \in [0, T] \times \overline{B(0, R_T)^{2N}}$.*

Proof. This result follows directly from the $C_{\text{loc}}^{2,1}$ -Wasserstein regularity hypotheses **(H3)**-**(H6)** on the datum of (\mathcal{P}_N) along with the uniform compactness of the optimal Pontryagin triples derived in Proposition 4.3. \square

Step 2 : Construction of a Lipschitz-in-space optimal controls for (\mathcal{P}_N)

In this second step, we wish to associate to any solution $(\mathbf{u}_N^*(\cdot), \mathbf{x}_N^*(\cdot))$ of (\mathcal{P}_N) a mean-field optimal control map $u_N^* \in L^\infty([0, T], \text{Lip}(\mathbb{R}^d, U))$ which $W^{1, \infty}$ -norm in space is bounded uniformly with respect to N for \mathcal{L}^1 -almost every $t \in [0, T]$.

We have seen in Proposition 4.3 that as a consequence of **(H)**, any optimal pair $(\mathbf{u}_N^*(\cdot), \mathbf{x}_N^*(\cdot))$ satisfies a PMP adapted to the mean-field structure of (\mathcal{P}_N) . In Proposition

4.4 below, we show that this result along with the coercivity assumption **(CO_N)** and Theorem 4.3 allows us to build a sequence of optimal controls $(u_N(\cdot, \cdot)) \subset L^\infty([0, T], \text{Lip}(\Omega, U))$ which Lipschitz constant in space exist and are uniformly bounded with respect to $N \geq 1$.

Proposition 4.4 (Existence of mean-field locally optimal Lipschitz feedback). *Let $(\mathbf{u}_N^*(\cdot), \mathbf{x}_N^*(\cdot)) \in \mathcal{U}_N \times \text{Lip}([0, T], \overline{B(0, R_T)^N})$ be an optimal pair control-trajectory for (\mathcal{P}_N) and assume that hypotheses **(H)** hold. Then, there exists a Lipschitz map $u_N^*(\cdot, \cdot) \in \text{Lip}([0, T] \times \mathbb{R}^d, U)$ such that $u_N^*(t, x_i(t)) = u_i^*(t)$ for all times $t \in [0, T]$ and which Lipschitz constant \mathcal{L}_U with respect to the space variable is independent from N .*

Proof. The first step of this proof is to apply Theorem 4.3 to (\mathcal{P}_N) seen as an optimal control problem in the rescaled Euclidean space $((\mathbb{R}^d)^N, \langle \cdot, \cdot \rangle_N)$ introduced in (4.3). As it was already mentioned in the proof of Proposition 4.3, (\mathcal{P}_N) satisfies the structural assumptions **(H_{oc})** of Section 4.1.3.

Given a rescaled covector $\mathbf{r}_N^*(\cdot)$ associated to $(\mathbf{u}_N^*(\cdot), \mathbf{x}^*(\cdot))$ via (4.27), the mean-field Pontryagin triple $(\mathbf{x}_N^*(\cdot), \mathbf{r}_N^*(\cdot), \mathbf{u}_N^*(\cdot))$ is bounded in $L^\infty([0, T], (\mathbb{R}^{2d})^N \times U^N)$ uniformly with respect to N as a consequence of **(H1)** and Proposition 4.3. By Corollary 4.1, the $C^{2,1}$ -norms of the datum of (\mathcal{P}_N) , defined in the sense of (4.7)-(4.8), are uniformly bounded over $\mathcal{K} = [0, T] \times \overline{B(0, R_T)^{2N}} \times U^N$ by a constant $\mathcal{L}_{\mathcal{K}} > 0$.

Similarly to what was presented in Section 4.1.3, the mean-field Pontryagin optimality system (4.27) can be written as a solution of the differential generalized equation

$$0 \in \mathbf{F}_\tau^N(\mathbf{x}(\cdot), \mathbf{r}(\cdot), \mathbf{u}(\cdot)) + \mathbf{G}_\tau^N(\mathbf{x}(\cdot), \mathbf{r}(\cdot), \mathbf{u}(\cdot)) \quad (4.38)$$

for any $\tau \in [0, T]$. Here, the mappings $\mathbf{F}_\tau^N : \mathcal{Y}_\tau^N \rightarrow \mathcal{Z}_\tau^N$ and $\mathbf{G}_\tau^N : \mathcal{Y}_\tau^N \rightrightarrows \mathcal{Z}_\tau^N$ are respectively defined by

$$\mathbf{F}_\tau^N(\mathbf{x}(\cdot), \mathbf{r}(\cdot), \mathbf{u}(\cdot)) = \begin{pmatrix} \dot{\mathbf{x}}(\cdot) - \mathbf{V}_N[\mathbf{x}(\cdot)](\cdot, \mathbf{x}(\cdot)) - \mathbf{u}(\cdot) \\ x_i(\tau) - x_i^*(\tau) \\ \dot{\mathbf{r}}(\cdot) + \mathbf{Grad}_x \mathbb{H}_N(\cdot, \mathbf{x}(\cdot), \mathbf{r}(\cdot), \mathbf{u}(\cdot)) \\ \mathbf{r}(T) + \mathbf{Grad}_x \varphi(\mathbf{x}(T)) \\ -\mathbf{Grad}_u \mathbb{H}_N(\cdot, \mathbf{x}(\cdot), \mathbf{r}(\cdot), \mathbf{u}(\cdot)) \end{pmatrix}, \quad (4.39)$$

where $\mathbf{V}_N[\mathbf{x}(\cdot)](t, \mathbf{x}(\cdot)) \equiv (\mathbf{v}_N[\mathbf{x}(\cdot)](t, x_i(\cdot)))_{1 \leq i \leq N} \in (\mathbb{R}^d)^N$ and

$$\mathbf{G}_\tau^N(\mathbf{x}(\cdot), \mathbf{r}(\cdot), \mathbf{u}(\cdot)) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ N_{\mathcal{U}_N}^\infty(\mathbf{u}(\cdot)) \end{pmatrix}.$$

The two Banach spaces $\mathcal{Y}_\tau^N, \mathcal{Z}_\tau^N$ are defined in this context by

$$\begin{cases} \mathcal{Y}_\tau^N = W^{1,1,\infty}([\tau, T], (\mathbb{R}^d)^N) \times W^{1,1,\infty}([\tau, T], (\mathbb{R}^d)^N) \times L^\infty([\tau, T], U^N), \\ \mathcal{Z}_\tau^N = L^\infty([\tau, T], (\mathbb{R}^d)^N) \times (\mathbb{R}^d)^N \times L^\infty([\tau, T], (\mathbb{R}^d)^N) \times (\mathbb{R}^d)^N \times L^\infty([\tau, T], (\mathbb{R}^d)^N). \end{cases}$$

Following [58], we now compute the first-order variation of the map $\mathbf{F}_\tau^N(\cdot)$ with respect to the adapted differential structure introduced in Section 4.1.1. Let $(\mathbf{y}(\cdot), \mathbf{s}(\cdot), \mathbf{w}(\cdot)) \in \mathcal{Y}_0^N$,

$i \in \{1, \dots, N\}$ and $t \in [0, T]$. One has that

$$\begin{aligned} \mathbf{v}_N[\mathbf{x} + \mathbf{y}](t, x_i + sy_i) &= \mathbf{v}_N[\mathbf{x}](t, x_i) + \mathbf{D}_x \mathbf{v}_N[\mathbf{x}](t, x_i) y_i \\ &\quad + \frac{1}{N} \sum_{j=1}^N \mathbf{D}_{x_j} \mathbf{v}_N[\mathbf{x}](t, x_i) y_j + o(|y_i|) + o(|\mathbf{y}|_N), \end{aligned} \quad (4.40)$$

where $\mathbf{D}_{x_j} \mathbf{v}_N[\mathbf{x}](t, x_i)$ is the matrix which rows are the mean-field gradients with respect to x_j of the components $(\mathbf{v}_N^k[\mathbf{x}](t, x_i))_{1 \leq k \leq d}$. Analogously, it holds that

$$\begin{aligned} \mathbf{Grad}_x \mathbb{H}_N(t, \mathbf{x}(t) + \mathbf{y}, \mathbf{r}(t) + \mathbf{s}(t), \mathbf{u}(t) + \mathbf{w}(t)) &= \mathbf{Grad}_x \mathbb{H}_N(t, \mathbf{x}(t), \mathbf{r}(t), \mathbf{u}(t)) \\ &\quad + \mathbf{Hess}_x \mathbb{H}_N(t, \mathbf{x}(t), \mathbf{r}(t), \mathbf{u}(t)) \mathbf{y}(t) \\ &\quad + \mathbf{Hess}_{rx} \mathbb{H}_N(t, \mathbf{x}(t), \mathbf{r}(t), \mathbf{u}(t)) \mathbf{s}(t) \\ &\quad + o(|\mathbf{y}(t)|_N) + (|\mathbf{w}(t)|_N) \end{aligned}$$

as well as

$$\begin{aligned} \mathbf{Grad}_u \mathbb{H}_N(t, \mathbf{x}(t) + \mathbf{y}, \mathbf{r}(t) + \mathbf{s}(t), \mathbf{u}(t) + \mathbf{w}(t)) &= \mathbf{Grad}_u \mathbb{H}_N(t, \mathbf{x}(t), \mathbf{r}(t), \mathbf{u}(t)) \\ &\quad + \mathbf{Hess}_u \mathbb{H}_N(t, \mathbf{x}(t), \mathbf{r}(t), \mathbf{u}(t)) \mathbf{w}(t) \\ &\quad + \mathbf{Hess}_{ru} \mathbb{H}_N(t, \mathbf{x}(t), \mathbf{r}(t), \mathbf{u}(t)) \mathbf{s}(t) \\ &\quad + o(|\mathbf{s}(t)|_N) + (|\mathbf{w}(t)|_N) \end{aligned}$$

and

$$\mathbf{Grad}_x \varphi_N(\mathbf{x}(T) + \mathbf{y}(T)) = \mathbf{Grad}_x \varphi(\mathbf{x}(T)) + \mathbf{Hess}_x \varphi_N(\mathbf{x}(T)) \mathbf{y}(T) + o(|\mathbf{y}(T)|_N)$$

as a consequence of the chainrule of Proposition 4.1. Here for convenience, we used the matrix representation (4.15) for mean-field Hessians in $(\mathbb{R}^d)^N$ introduced in Remark 4.3.

It is again possible to interpret the partial linearization of the differential generalized inclusion (4.38) as the Pontryagin maximum principle for the linear-quadratic problem

$$\left\{ \begin{array}{l} \min_{\mathbf{w}(\cdot) \in \mathcal{U}_\tau^N} \left[\int_\tau^T \left(\frac{1}{2} \langle \mathbf{A}(t) \mathbf{y}(t), \mathbf{y}(t) \rangle_N + \frac{1}{2} \langle \mathbf{B}(t) \mathbf{w}(t), \mathbf{w}(t) \rangle_N \right) dt + \frac{1}{2} \langle \mathbf{C}(T) \mathbf{y}(T), \mathbf{y}(T) \rangle_N \right] \\ \text{s.t.} \begin{cases} \dot{y}_i(t) = \mathbf{D}_x \mathbf{v}_N[\mathbf{x}_N^*(t)](t, x_i^*(t)) y_i(t) + \frac{1}{N} \sum_{j=1}^N \mathbf{D}_{x_j} \mathbf{v}_N[\mathbf{x}_N^*(t)](t, x_i^*(t)) y_j(t) \\ y_i(\tau) = 0. \end{cases} \end{array} \right.$$

where $\mathcal{U}_\tau^N = \left\{ v \in L^\infty([\tau, T], U^N) \text{ s.t. } \mathbf{u}_N^*(t) + \mathbf{w}(t) \in U^N \text{ for } \mathcal{L}^1\text{-almost every } t \in [\tau, T] \right\}$ and

$$\left\{ \begin{array}{l} \mathbf{A}(t) = -\mathbf{Hess}_x \mathbb{H}_N(t, \mathbf{x}_N^*(t), \mathbf{r}_N^*(t), \mathbf{u}_N^*(t)), \\ \mathbf{B}(t) = -\mathbf{Hess}_u \mathbb{H}_N(t, \mathbf{x}_N^*(t), \mathbf{r}_N^*(t), \mathbf{u}_N^*(t)), \\ \mathbf{C}(T) = \mathbf{Hess}_x \varphi_N(\mathbf{x}_N^*(T)). \end{array} \right.$$

Moreover, we assumed in (\mathbf{CO}_N) that there exists a constant ρ_T , which is independent

from N , such that the mean-field coercivity estimate

$$\begin{aligned} \mathbf{Hess}_x \varphi_N[\mathbf{x}_N^*(T)](\mathbf{y}(T), \mathbf{y}(T)) &= \int_0^T \mathbf{Hess}_x \mathbb{H}_N[t, \mathbf{x}_N^*(t), \mathbf{r}_N^*(t), \mathbf{u}_N^*(t)](\mathbf{y}(t), \mathbf{y}(t)) dt \\ &\quad - \int_0^T \mathbf{Hess}_u \mathbb{H}_N[t, \mathbf{x}_N^*(t), \mathbf{r}_N^*(t), \mathbf{u}_N^*(t)](\mathbf{w}(t), \mathbf{w}(t)) dt \\ &\geq \rho_T \int_0^T |\mathbf{w}(t)|_N^2 dt \end{aligned}$$

holds for any admissible pair $(\mathbf{w}(\cdot), \mathbf{y}(\cdot)) \in L^2([0, T], (\mathbb{R}^d)^N) \times W^{1,2}([0, T], (\mathbb{R}^d)^N)$ solution of the linearized mean-field dynamics.

We can therefore apply Theorem 4.3 to (\mathcal{P}_N) and recover the existence of an open neighbourhood $\mathcal{N} \subset [0, T] \times (\mathbb{R}^d)^N$ of $\text{Graph}(\mathbf{x}^*(\cdot))$ along with that of a locally optimal Lipschitz feedback $\tilde{\mathbf{u}}(\cdot, \cdot)$ defined over $\mathcal{N} \cap ([0, T] \times B(0, R_T)^N)$ which Lipschitz constant \mathcal{L}_U depends only on the structural constant \mathcal{L}_K introduced in Corollary 4.1 and on the coercivity constant ρ_T introduced in (\mathbf{CO}_N) . In particular, \mathcal{L}_U is independent from N .

For any $i \in \{1, \dots, N\}$, we associate to $x_i^*(\cdot)$ the projected control maps $\tilde{u}_i : \mathcal{N}_i \equiv \pi^i(\mathcal{N}) \rightarrow \mathbb{R}^d$ defined by

$$\tilde{u}_i(t, x) = \tilde{\mathbf{u}}_i(t, \hat{\mathbf{x}}_i^x(t)),$$

where $\hat{\mathbf{x}}_i^x(t) = (x_1^*(t), \dots, x_{i-1}^*(t), x, x_{i+1}^*(t), \dots, x_N^*(t))$ denotes the element in $(\mathbb{R}^d)^N$ which has all its components matching that of $\mathbf{x}^*(t)$ except the i -th one which is free and equal to x . By construction, each $\tilde{u}_i(\cdot, \cdot)$ defines a locally optimal feedback in the neighbourhood \mathcal{N}_i of $\text{Graph}(x_i^*(\cdot))$. Furthermore, we can derive the following uniform estimate for the projected control maps

$$\begin{aligned} |\tilde{u}_i(t, y) - \tilde{u}_i(t, x)| &= |\tilde{\mathbf{u}}_i(t, \hat{\mathbf{x}}_i^y(t)) - \tilde{\mathbf{u}}_i(t, \hat{\mathbf{x}}_i^x(t))| \\ &\leq \left(\sum_{j=1}^N |\tilde{\mathbf{u}}_j(t, \hat{\mathbf{x}}_i^y(t)) - \tilde{\mathbf{u}}_j(t, \hat{\mathbf{x}}_i^x(t))|^2 \right)^{1/2} \\ &= \sqrt{N} |\tilde{\mathbf{u}}(t, \hat{\mathbf{x}}_i^y(t)) - \tilde{\mathbf{u}}(t, \hat{\mathbf{x}}_i^x(t))|_N \\ &\leq \sqrt{N} \mathcal{L}_U |\hat{\mathbf{x}}_i^y(t) - \hat{\mathbf{x}}_i^x(t)|_N = \mathcal{L}_U |y - x|, \end{aligned}$$

so that we recover the uniform Lipschitz estimate

$$|\tilde{u}_i(t, y) - \tilde{u}_i(t, x)| \leq \mathcal{L}_U |y - x|$$

for all $(t, x, y) \in [0, T] \times \pi_{\mathbb{R}^d}(\mathcal{N}_i)^2$. This shows that the projected optimal control $\tilde{u}_i(\cdot, \cdot)$ maps are Lipschitz-regular in space uniformly with respect to N .

Therefore, each $\tilde{u}_i(\cdot, \cdot)$ can be defined unequivocally on a closed neighbourhood of $\text{Graph}(x_i^*(\cdot))$ contained in \mathcal{N}_i . By using e.g. Kirszbraun's Extension Theorem (see e.g. [76, Theorem 2.10.43]) combined with a projection on the convex and compact set U , one can define a global optimal control map $u_N^* : [0, T] \times \mathbb{R}^d \rightarrow U$ such that $u_N^*(t, x_i^*(t)) = u_i^*(t)$ for all $t \in [0, T]$ and $\sup_{t \in [0, T]} \text{Lip}(u_N^*(t, \cdot); \mathbb{R}^d) \leq \mathcal{L}_U$. \square

Remark 4.4 (Absence of collisions and crossings of characteristics). *Notice that as a consequence of our results, particles cannot collide into one another. This comes from the fact that under Cauchy-Lipschitz well-posedness, the solutions defined in the superposition*

sense of Theorem 1.6 are concentrated on well-defined and non-crossing characteristic curves. Let it be remarked that such an absence of collision is not sufficient for the Lipschitz regularity of the optimal controls. This fact is highlighted in Section 4.3.2 where we provide an example in which no collisions can occur between particles, and yet the uniform coercivity estimate (\mathbf{CO}_N) is necessary and sufficient for the Lipschitz regularity in space of the optimal control.

Step 3 : Existence of Lipschitz optimal controls for problem (\mathcal{P})

In this third and last step, we show that the sequence of optimal maps $(u_N^*(\cdot, \cdot))$ that we constructed in Proposition 4.4 is compact in a suitable topology and that the limits along suitable subsequences are optimal solutions of problem (\mathcal{P}) which are Lipschitz-regular in space. We state in the following proposition a variation of the classical *Dunford-Pettis* compactness criterion (see e.g. [14, Theorem 1.38]).

Proposition 4.5 (Compactness of Lipschitz-in-space optimal maps). *Let $\mathcal{L}_U > 0$ be a positive constant and $\Omega \subset \mathbb{R}^d$ be a bounded set. Then, the set*

$$\mathcal{U}_{\mathcal{L}_U} = \left\{ u(\cdot, \cdot) \in L^2([0, T], \text{Lip}(\Omega, U)) \text{ s.t. } \sup_{t \in [0, T]} \|u(t, \cdot)\|_{W^{1, \infty}(\Omega, U)} \leq \mathcal{L}_U \right\}$$

is compact in the weak topology of $L^2([0, T], W^{1, p}(\Omega, \mathbb{R}^d))$ for any $p \in (1, +\infty)$.

Proof. See [82, Theorem 2.5] in which this result is also used in the context of mean-field optimal control. \square

This compactness result allows to derive the following convergence result on the sequence of mean-field controls $(u_N^*(\cdot, \cdot))$ built in Step 2.

Corollary 4.2 (Convergence of Lipschitz optimal control). *There exists a map $u^*(\cdot, \cdot) \in L^\infty([0, T], \text{Lip}(\mathbb{R}^d, U))$ such that the sequence of Lipschitz optimal controls maps $(u_N^*(\cdot, \cdot))$ defined in Proposition 4.4 converges weakly towards $u^*(\cdot, \cdot)$ along a subsequence.*

Proof. This result comes from a direct application of Proposition 4.5 to the sequence of optimal maps built in Proposition 4.4 up to redefining $\mathcal{L}_U \equiv \max\{L_U, \mathcal{L}_U\}$. \square

We now prove that the generalized optimal control $\nu^* \in \mathcal{U}$ for problem $(\mathcal{P}_{\text{meas}})$ is in fact induced by the Lipschitz-in-space optimal control $u^*(\cdot, \cdot) \in L^\infty([0, T], \text{Lip}(\mathbb{R}^d, U))$ which has been defined in Corollary 4.2. Remark first that by construction of the maps $(u_N^*(\cdot, \cdot))$, it holds that

$$\begin{aligned} \nu_N^* &= \int_{[0, T]} \left(\frac{1}{N} \sum_{i=1}^N u_i^*(t) \delta_{x_i^*(t)} \right) d\lambda(t) \\ &= \int_{[0, T]} \left(\frac{1}{N} \sum_{i=1}^N u_N^*(t, x_i^*(t)) \delta_{x_i^*(t)} \right) d\lambda(t) = u_N^*(\cdot, \cdot) \tilde{\mu}_N^*, \end{aligned}$$

for any $N \geq 1$, where $\nu_N^* \in \mathcal{U}$ denotes the generalized discrete control measure introduced in Theorem 4.2. In the following proposition, we prove that the sequence $(u_N^*(\cdot, \cdot) \tilde{\mu}_N^*)$ converges towards $u^*(\cdot, \cdot) \tilde{\mu}^*$ in the weak-* topology of $\mathcal{M}([0, T] \times \mathbb{R}^d, U)$.

Proposition 4.6 (Convergence of generalized Lipschitz optimal controls). *Let $(\mu_N^*(\cdot)) \subset \text{Lip}([0, T], \mathcal{P}_N(\mathbb{R}^d))$ be the sequence of optimal measure curves associated with (\mathcal{P}_N) and $(u_N^*(\cdot, \cdot)) \subset L^\infty([0, T], \text{Lip}(\mathbb{R}^d, U))$ be the sequence of Lipschitz controls built in Proposition 4.4. Then, the sequence $(\nu_N^*) = (u_N^*(\cdot, \cdot)\tilde{\mu}_N^*)$ converges towards $\nu^* = u^*(\cdot, \cdot)\tilde{\mu}^*$ in the weak-* topology of $\mathcal{M}([0, T] \times \Omega, \mathbb{R}^d)$.*

Proof. We know by Proposition 4.5 that for any $p \in (1, +\infty)$, there exists a subsequence of $(u_N^*(\cdot, \cdot))$ which converges in the weak-topology of $L^2([0, T], W^{1,p}(\Omega, U))$ towards $u^*(\cdot, \cdot) \in \mathcal{U}_{\mathcal{L}_U}$. Recalling that one can identify the topological dual $(L^2([0, T], W^{1,p}(\Omega, U)))'$ with $L^2([0, T], W^{-1,p'}(\Omega, U))$ where p' is the conjugate exponent of p , the fact that $u_N(\cdot, \cdot) \rightharpoonup u(\cdot, \cdot)$ can be written as

$$\int_0^T \langle \xi(t), u_N^*(t, \cdot) \rangle_{W^{1,p}} dt \xrightarrow{N \rightarrow +\infty} \int_0^T \langle \xi(t), u^*(t, \cdot) \rangle_{W^{1,p}} dt \quad (4.41)$$

for any $\xi \in L^2([0, T], W^{-1,p'}(\Omega, \mathbb{R}^d))$ and where $\langle \cdot, \cdot \rangle_{W^{1,p}}$ denotes the duality bracket of $W^{1,p}(\Omega, U)$.

Let us now fix in particular a real number $p > d$ so that by Morrey's Embedding (see e.g. [38, Theorem 9.12]) it holds that $W^{1,p}(\Omega, U) \subset C^0(\Omega, U)$. By taking the topological dual of each spaces, we recover the reverse inclusion $\mathcal{M}(\Omega, U) \subset W^{-1,p'}(\Omega, U)$. The latter relation combined with the definition (1.1) of the duality pairing for vector measures and (4.41) yields that

$$\int_0^T \int_{\mathbb{R}^d} \langle \xi(t, x), u_N^*(t, x) \rangle d\sigma(t)(x) dt \xrightarrow{N \rightarrow +\infty} \int_0^T \int_{\mathbb{R}^d} \langle \xi(t, x), u^*(t, x) \rangle d\sigma(t)(x) dt, \quad (4.42)$$

for any measure-valued curve $\sigma(\cdot) \in C^0([0, T], \mathcal{M}(\Omega, \mathbb{R}_+))$ and any $\xi \in C_c^\infty([0, T] \times \Omega, \mathbb{R}^d)$. Remark now that for any $N \geq 1$, one has that

$$\begin{aligned} & \left| \int_0^T \int_{\mathbb{R}^d} \langle \xi(t, x), u^*(t, x) \rangle d\mu^*(t)(x) dt - \int_0^T \int_{\mathbb{R}^d} \langle \xi(t, x), u_N^*(t, x) \rangle d\mu_N^*(t)(x) dt \right| \\ & \leq \left| \int_0^T \int_{\mathbb{R}^d} \langle \xi(t, x), u^*(t, x) - u_N^*(t, x) \rangle d\mu^*(t)(x) dt \right| \\ & + \left| \int_0^T \int_{\mathbb{R}^d} \langle \xi(t, x), u_N^*(t, x) \rangle d(\mu^*(t) - \mu_N^*(t))(x) dt \right|. \end{aligned} \quad (4.43)$$

The first term in the right-hand side of (4.43) vanishes as $N \rightarrow +\infty$ as a consequence of (4.42). By invoking Kantorovich's duality formula (1.4) along with the uniform Lipschitz-regularity of the maps $(u_N^*(\cdot, \cdot))$, we can obtain the following upper bound on the second term in the right-hand side of (4.43)

$$\begin{aligned} & \left| \int_0^T \int_{\mathbb{R}^d} \langle \xi(t, x), u_N^*(t, x) \rangle d(\mu^*(t) - \mu_N^*(t))(x) dt \right| \\ & \leq \left(\mathcal{L}_U \sup_{t \in [0, T]} \text{Lip}(\xi(t, \cdot); \Omega) \right) \sup_{t \in [0, T]} W_1(\mu(t), \mu_N(t)) \xrightarrow{N \rightarrow +\infty} 0. \end{aligned}$$

Therefore, we recover that

$$\int_0^T \int_{\mathbb{R}^d} \langle \xi(t, x), u_N^*(t, x) \rangle d\mu_N^*(t)(x) dt \xrightarrow{N \rightarrow +\infty} \int_0^T \int_{\mathbb{R}^d} \langle \xi(t, x), u^*(t, x) \rangle d\mu^*(t)(x) dt,$$

which precisely amounts to saying that $\nu_N^* \rightharpoonup^* u^*(\cdot, \cdot) \tilde{\mu}^*$ as $N \rightarrow +\infty$ along the same subsequence. \square

By uniqueness of the weak-* limit in $\mathcal{M}([0, T] \times \mathbb{R}^d, U)$, we obtain by combining Proposition 4.6 with Theorem 4.2 that the optimal solution $\nu^* \in \mathcal{U}$ of $(\mathcal{P}_{\text{meas}})$ is induced by $u^*(\cdot, \cdot)$. This allows to conclude that the pair $(u^*(\cdot, \cdot), \mu^*(\cdot)) \in L^\infty([0, T], \text{Lip}(\mathbb{R}^d, U)) \times \text{Lip}([0, T], \mathcal{P}(\overline{B(0, R_T)}))$ is a classical optimal pair for (\mathcal{P}) , which concludes the proof of Theorem 4.1.

4.3 Discussions on the coercivity assumption (\mathbf{CO}_N)

In this section, we discuss more in detail the mean-field coercivity assumptions \mathbf{CO}_N . In Section 4.3.1, we derive a generic sufficient condition on the control cost $\psi : U \rightarrow [0, +\infty]$ for the uniform mean-field coercivity to hold. We further develop in Section 4.3.2 an example in which hypotheses (\mathbf{CO}_N) is both necessary and sufficient for the Lipschitz regularity in space of the optimal control.

4.3.1 A generic sufficient condition for coercivity

In this section, we prove a simple and general sufficient condition for the coercivity estimate (\mathbf{CO}_N) to hold, stated in the following proposition.

Proposition 4.7 (A sufficient condition for mean-field coercivity). *Consider a mean-field optimal control problem of the form (\mathcal{P}) , and suppose that hypotheses $(\mathbf{H1})$ - $(\mathbf{H6})$ hold. Then, there exists a constant $\lambda_{\mathcal{P}} > 0$ such that, if the control cost $\psi(\cdot)$ is λ_ψ -strongly convex with $\lambda_\psi > \lambda_{\mathcal{P}}$, then the mean-field coercivity estimate (\mathbf{CO}_N) holds with $\rho_T = \lambda_\psi - \lambda_{\mathcal{P}}$.*

The main ingredient involved in the proof of this result is contained in the following technical lemma, which proof is provided for the sake of completeness.

Lemma 4.3 (C^2 -functions are λ -convex on convex and compact sets). *Let $K \subset \mathbb{R}^d$ be a convex and compact set and $\phi \in C^0(\mathcal{P}_c(\mathbb{R}^d), \mathbb{R})$ be $C_{\text{loc}}^{2,1}$ -MF regular. Then, it holds that*

$$\mathbf{Hess} \phi_N[\mathbf{x}](\mathbf{h}, \mathbf{h}) \geq -\text{Lip}(\mathbf{Grad} \phi_N(\cdot); K^N) |\mathbf{h}|_N^2$$

for any $\mathbf{x}, \mathbf{y} \in K^N$, where $(\phi_N(\cdot))$ denotes the discrete approximating sequence of $\phi(\cdot)$.

Proof. The proof of this result is split into two steps. We first prove a λ -convexity inequality for $\phi_N(\cdot)$, and then proceed by showing that the latter implies the claimed lower bound on the Hessian matrix.

Let $\mathbf{x}, \mathbf{y} \in K^N$ and $t \in [0, 1]$. As a consequence of Proposition 4.1, we can write the following integral Taylor formulas for $\phi_N(\cdot)$

$$\begin{cases} \phi_N((1-t)\mathbf{x} + t\mathbf{y}) = \phi_N(\mathbf{x}) + \int_0^1 \langle \mathbf{Grad} \phi_N(\mathbf{x} + st(\mathbf{y} - \mathbf{x})), t(\mathbf{y} - \mathbf{x}) \rangle_N ds, \\ \phi_N(\mathbf{y}) = \phi_N(\mathbf{x}) + \int_0^1 \langle \mathbf{Grad} \phi_N(\mathbf{x} + s(\mathbf{y} - \mathbf{x})), \mathbf{y} - \mathbf{x} \rangle_N ds. \end{cases} \quad (4.44)$$

Combining the two equations of (4.44), we obtain that

$$\begin{aligned} & \phi_N((1-t)\mathbf{x} + t\mathbf{y}) - (1-t)\phi_N(\mathbf{x}) - t\phi_N(\mathbf{y}) \\ &= \int_0^1 \langle \mathbf{Grad} \phi_N(\mathbf{x} + st(\mathbf{y} - \mathbf{x})), t(\mathbf{y} - \mathbf{x}) \rangle_N ds \\ & - \int_0^1 \langle \mathbf{Grad} \phi_N(\mathbf{x} + s(\mathbf{y} - \mathbf{x})), t(\mathbf{y} - \mathbf{x}) \rangle_N ds \leq \frac{t(1-t)}{2} \text{Lip}(\mathbf{Grad} \phi_N(\cdot); K^N) |\mathbf{y} - \mathbf{x}|_N^2, \end{aligned}$$

where we used the fact that both $\mathbf{x} + st(\mathbf{y} - \mathbf{x})$ and $\mathbf{x} + s(\mathbf{y} - \mathbf{x})$ belong to K^N since this set is convex. Therefore, we have shown that the map $\phi_N(\cdot)$ is λ -convex over K^N with $\lambda = -\text{Lip}(\mathbf{Grad}(\phi_N(\cdot)); K^N)$.

Choose now in particular $\mathbf{y} = \mathbf{x} + \mathbf{h}$ with $|\mathbf{h}|_N$ small. The λ -convexity inequality of $\phi_N(\cdot)$ can be rewritten as

$$\phi_N(\mathbf{x} + t\mathbf{h}) \leq t(\phi_N(\mathbf{x} + \mathbf{h}) - \phi_N(\mathbf{x})) + \frac{t(1-t)}{2} \text{Lip}(\mathbf{Grad} \phi_N(\cdot); K^N) |\mathbf{h}|_N^2. \quad (4.45)$$

By applying (4.4) of Proposition 4.1 to (4.45), we recover that

$$t^2 \mathbf{Hess} \phi_N[\mathbf{x}](\mathbf{h}, \mathbf{h}) \leq t \mathbf{Hess} \phi_N[\mathbf{x}](\mathbf{h}, \mathbf{h}) + t(1-t) \text{Lip}(\mathbf{Grad} \phi_N(\cdot); K^N) |\mathbf{h}|_N^2,$$

so that dividing by t and letting $t \rightarrow 0^+$, we obtain

$$\mathbf{Hess} \phi_N[\mathbf{x}](\mathbf{h}, \mathbf{h}) \geq -\text{Lip}(\mathbf{Grad} \phi_N(\cdot); K^N) |\mathbf{h}|_N^2,$$

which concludes the proof of our claim. \square

Proof. (of Proposition 4.7) As a consequence of hypotheses **(H3)**-**(H6)** and of the uniform bounds derived in Proposition 4.3 on the Pontryagin triples $(\mathbf{u}^*(\cdot), \mathbf{x}^*(\cdot), \mathbf{r}^*(\cdot))$, we know that the mean-field Hamiltonian $\mathbb{H}_N(\cdot, \cdot, \cdot, \cdot)$ defined in (4.28) and the final cost $\varphi_N(\cdot)$ are $C_{\text{loc}}^{2,1}$ -MF regular with respect to $(\mathbf{x}, \mathbf{u}) \in (\mathbb{R}^d)^N \times U^N$. By Grönwall Lemma, we furthermore know that solutions $\mathbf{y}(\cdot)$ of the mean-field linearized system described in **(CO_N)** are contained in a product of compact sets $K^N \subset (\mathbb{R}^d)^N$ and that

$$\max \left\{ |\mathbf{y}(T)|_N^2, \int_0^T |\mathbf{y}(t)|_N^2 dt \right\} \leq C_T \int_0^T |\mathbf{w}(t)|_N^2 dt,$$

for a given uniform constant $C_T > 0$. Merging these few facts together along with the

statement of Lemma 4.3, we obtain the existence of a uniform constant $\lambda_{\mathcal{P}} > 0$ such that

$$\begin{aligned} \mathbf{Hess}_x \varphi_N[\mathbf{x}_N^*(T)](\mathbf{y}(T), \mathbf{y}(T)) &- \int_0^T \mathbf{Hess}_x \mathbb{H}_N[t, \mathbf{x}_N^*(t), -\mathbf{r}_N^*(t), \mathbf{u}_N^*(t)](\mathbf{y}(t), \mathbf{y}(t)) dt \\ &\geq -\lambda_{\mathcal{P}} \int_0^T |\mathbf{w}(t)|_N^2 dt. \end{aligned} \quad (4.46)$$

Notice now that by the strong convexity assumption made on $\psi(\cdot)$ in , it holds that

$$\begin{aligned} &- \int_0^T \mathbf{Hess}_u \mathbb{H}_N[t, \mathbf{x}_N^*(t), -\mathbf{r}_N^*(t), \mathbf{u}_N^*(t)](\mathbf{w}(t), \mathbf{w}(t)) dt \\ &= \frac{1}{N} \sum_{i=1}^N \int_0^T \langle \nabla^2 \psi(u_i^*(t)) w_i(t), w_i(t) \rangle dt \leq \lambda_{\psi} \int_0^T |\mathbf{w}(t)|_N^2 dt. \end{aligned} \quad (4.47)$$

Combining (4.46) and (4.47), we obtain the uniform coercivity-type estimate

$$\begin{aligned} \mathbf{Hess}_x \varphi_N[\mathbf{x}_N^*(T)](\mathbf{y}(T), \mathbf{y}(T)) &- \int_0^T \mathbf{Hess}_x \mathbb{H}_N[t, \mathbf{x}_N^*(t), -\mathbf{r}_N^*(t), \mathbf{u}_N^*(t)](\mathbf{y}(t), \mathbf{y}(t)) dt \\ &- \int_0^T \mathbf{Hess}_u \mathbb{H}_N[t, \mathbf{x}_N^*(t), -\mathbf{r}_N^*(t), \mathbf{u}_N^*(t)](\mathbf{w}(t), \mathbf{w}(t)) dt \\ &\geq (\lambda_{\psi} - \lambda_{\mathcal{P}}) \int_0^T |\mathbf{w}(t)|_N^2 dt \end{aligned}$$

Therefore, up to choosing a control cost with strong convexity constant $\lambda_{\psi} > \lambda_{\mathcal{P}}$, the mean-field coercivity estimate of hypothesis (\mathbf{CO}_N) is ensured to hold with $\rho_T = \lambda_{\psi} - \lambda_{\mathcal{P}}$. \square

4.3.2 Sharpness of the mean-field coercivity condition (\mathbf{CO}_N) on a 1D example

In this section, we look into a simple example of on the real line for which the mean-field coercivity condition (\mathbf{CO}_N) is both necessary and sufficient for the existence of a mean-field Lipschitz optimal control. With this aim, we consider the following Wasserstein optimal control problem

$$(\mathcal{P}_V) \quad \begin{cases} \min_{u \in \mathcal{U}} \left[\frac{\lambda}{2} \int_0^T \int_{\mathbb{R}} |u(t, x)|^2 d\mu(t)(x) dt - \frac{1}{2} \int_{\mathbb{R}} |x - \bar{\mu}(T)|^2 d\mu(T)(x) \right] \\ \text{s.t.} \quad \begin{cases} \partial_t \mu(t) + \nabla \cdot (u(t, \cdot) \mu(t)) = 0, \\ \mu(0) = \mu^0 = \frac{1}{2} \mathbb{1}_{[-1, 1]} \mathcal{L}^1, \end{cases} \end{cases}$$

consisting in maximizing the variance at time $T > 0$ of a measure curve $\mu(\cdot)$ starting from the indicator function of $[-1, 1]$ at time $t = 0$, while penalizing the L^2 -norm of the control u . Here, the set of admissible control values is $U = [-C, C]$ for a positive constant $C > 0$, and the parameter $\lambda > 0$ is the relative weight between the final cost and the control penalization.

It can be verified straightforwardly that this problem fits the hypotheses $(\mathbf{H1})$ - $(\mathbf{H6})$ of Theorem 4.1. Given a sequence of empirical measures $(\mu_N^0) \equiv (\mu[\mathbf{x}_N]) \subset \mathcal{P}_N(\mathbb{R})$ converging narrowly towards μ^0 , we can define the family (\mathcal{P}_V^N) of discretized multi-agent

problems as

$$(\mathcal{P}_V^N) \quad \begin{cases} \min_{\mathbf{u}(\cdot) \in \mathcal{U}_N} \left[\frac{\lambda}{2N} \sum_{i=1}^N \int_0^T u_i^2(t) dt - \frac{1}{2N} \sum_{i=1}^N |x_i(T) - \bar{\mathbf{x}}(T)|^2 \right] \\ \text{s.t.} \quad \begin{cases} \dot{x}_i(t) = u_i(t), \\ x_i(0) = x_i^0. \end{cases} \end{cases}$$

where $\bar{\mathbf{x}}(\cdot) = \frac{1}{N} \sum_{i=1}^N x_i(\cdot)$ and $\mathcal{U}_N = L^\infty([0, T], U)$. As a consequence of Proposition 4.2, there exists for any $N \geq 1$ an optimal pair control-trajectory $(\mathbf{u}_N^*(\cdot), \mathbf{x}_N^*(\cdot)) \in L^\infty([0, T], U^N) \times \text{Lip}([0, T], (\mathbb{R}^d)^N)$ for (\mathcal{P}_V^N) .

The mean-field Hamiltonian associated to (\mathcal{P}_V^N) is given by

$$\mathbb{H}_N : (t, \mathbf{x}, \mathbf{r}, \mathbf{u}) \in [0, T] \times (\mathbb{R}^3)^N \mapsto \frac{1}{N} \sum_{i=1}^N \left(\langle r_i, u_i \rangle - \frac{1}{2} |u_i|^2 \right). \quad (4.48)$$

By applying the mean-field Pontryagin Maximum Principle displayed in Proposition 4.3, we obtain the existence of a covector $\mathbf{r}_N^*(\cdot) \in \text{Lip}([0, T], \mathbb{R}^N)$ such that

$$\begin{cases} \dot{r}_i^*(t) = -\mathbf{Grad}_{x_i} \mathbb{H}_N(t, \mathbf{x}_N^*(t), \mathbf{r}_N^*(t), \mathbf{u}_N^*(t)) = 0, \\ r_i^*(T) = \mathbf{Grad}_{x_i} \text{Var}_N(\mathbf{x}_N^*(T)) = x_i^*(T) - \bar{\mathbf{x}}^*(T), \\ u_i^*(t) \in \underset{v \in U}{\text{argmax}} [\langle r_i^*(t), v \rangle - \frac{1}{2} |v|^2]. \end{cases}$$

Therefore, the optimal covector $\mathbf{r}_N^*(\cdot)$ is constant and uniquely determined via

$$r_i^*(t) = x_i^*(T) - \bar{\mathbf{x}}^*(T).$$

Moreover, the optimal control $\mathbf{u}_N^*(\cdot)$ is also uniquely determined, and its components write explicitly as

$$u_i^*(t) = \pi_U(r_i^*(t)) \equiv \pi_{[-C, C]} \left(\frac{x_i^*(T) - \bar{\mathbf{x}}^*(T)}{\lambda} \right), \quad (4.49)$$

for all $i \in \{1, \dots, N\}$. It follows directly from this expression that

$$\dot{\bar{\mathbf{x}}}^*(t) = \frac{1}{N} \sum_{i=1}^N \dot{u}_i^* = \frac{1}{N} \sum_{i=1}^N \pi_{[-C, C]} \left(\frac{x_i^*(T) - \bar{\mathbf{x}}^*(T)}{\lambda} \right) = 0.$$

Without loss of generality, we can therefore choose $\mathbf{x}^0 \in \mathbb{R}^N$ such that $\bar{\mathbf{x}}^*(\cdot) \equiv \bar{\mathbf{x}}^0 = 0$.

In the following lemma, we derive a simple analytical necessary and sufficient condition for the mean-field coercivity assumption to hold for (\mathcal{P}_V) .

Lemma 4.4 (Charaterization of the coercivity condition for (\mathcal{P}_V)). *The mean-field coercivity condition (\mathbf{CO}_N) holds for (\mathcal{P}_V) if and only if $\lambda > T$. In which case, the optimal coercivity constant is given by $\rho_T = \lambda - T$.*

Proof. We start by computing the mean-field Hessians involved in the coercivity estimate.

For any $\mathbf{x}, \mathbf{y}, \mathbf{u}, \mathbf{w} \in \mathbb{R}^N$, we have as a consequence of Example 4.2 and (4.48) that

$$\begin{cases} \mathbf{Hess} \text{Var}_N[\mathbf{x}](\mathbf{y}, \mathbf{y}) = |\mathbf{y}|_N^2 - |\bar{\mathbf{y}}|^2 \leq |\mathbf{y}|_N^2, \\ \mathbf{Hess}_u \mathbb{H}_N[t, \mathbf{x}, \mathbf{r}, \mathbf{u}](\mathbf{w}, \mathbf{w}) = \lambda |\mathbf{w}|_N^2. \end{cases}$$

Let $(\mathbf{w}(\cdot), \mathbf{y}(\cdot)) \in L^2([0, T], U^N) \times W^{1,2}([0, T], \mathbb{R}^N)$ be the solution of the linearized control-state problem

$$\dot{\mathbf{y}}(t) = \mathbf{w}(t), \quad \mathbf{y}(0) = 0, \quad (4.50)$$

with $\mathbf{u}_N^*(t) + \mathbf{w}(t) \in U^N$. By Cauchy-Schwarz inequality, one can further estimate $|\mathbf{y}(T)|_N^2$ as

$$|\mathbf{y}(T)|_N^2 = \left| \int_0^T \mathbf{w}(t) dt \right|_N^2 \leq T \int_0^T |\mathbf{w}(t)|_N^2 dt,$$

so that we recover

$$\begin{aligned} -\mathbf{Hess} \text{Var}_N[\mathbf{x}_N^*(T)](\mathbf{y}(T), \mathbf{y}(T)) - \int_0^T \mathbf{Hess}_u \mathbb{H}_N[t, \mathbf{x}_N^*(t), \mathbf{r}_N^*(t), \mathbf{u}_N^*(t)](\mathbf{w}(t), \mathbf{w}(t)) dt \\ \geq (\lambda - T) \int_0^T |\mathbf{w}(t)|_N^2 dt, \end{aligned}$$

and we obtain that the mean-field coercivity condition (\mathbf{CO}_N) holds whenever $\lambda > T$.

Conversely, let us choose a constant admissible control perturbation $\mathbf{w}_c(\cdot) \equiv \mathbf{w}_c$ such that $\bar{\mathbf{w}}_c = 0$. It is always possible to make such a choice since by (4.49), there exists at least two indices i, j such that $\text{sign}(u_i) = -\text{sign}(u_j)$ for all times $t \in [0, T]$. It is then sufficient to choose \mathbf{w}_c such that

$$\begin{cases} (\mathbf{w}_c)_i = -\text{sign}(u_i)\epsilon, \quad (\mathbf{w}_c)_j = -(\mathbf{w}_c)_i, \\ (\mathbf{w}_c)_k = 0 \quad \text{if } k \in \{1, \dots, N\} \text{ and } k \neq i, j, \end{cases}$$

where $\epsilon > 0$ is a small parameter. As a consequence of (4.50), the corresponding state perturbation $\mathbf{y}_c(\cdot)$ is such that $\bar{\mathbf{y}}_c(\cdot) \equiv 0$. Moreover, it also holds that

$$|\mathbf{y}_c(T)|_N^2 = T^2 |\mathbf{w}_c|_N^2 = T \int_0^T |\mathbf{w}_c|_N^2 dt.$$

We therefore obtain that for this particular choice of linearized pair control-state, it holds that

$$\begin{aligned} -\mathbf{Hess} \text{Var}_N[\mathbf{x}_N^*(T)](\mathbf{y}_c(T), \mathbf{y}_c(T)) \\ - \int_0^T \mathbf{Hess}_u \mathbb{H}_N[t, \mathbf{x}_N^*(t), \mathbf{r}_N^*(t), \mathbf{u}_N^*(t)](\mathbf{w}_c(t), \mathbf{w}_c(t)) dt = (\lambda - T) \int_0^T |\mathbf{w}_c(t)|_N^2 dt, \end{aligned}$$

so that $\rho_T = \lambda - T$ is the sharp mean-field coercivity constant, and the mean-field coercivity condition holds only if $\lambda > T$. \square

Remark 4.5. *In the context of problem (\mathcal{P}_V) , the constant $\lambda_{\mathcal{P}}$ introduced in Proposition 4.7 is given explicitly by $\lambda_{\mathcal{P}} = T$, and we have proven that it is sharp.*

We can now use this characterization of the coercivity condition to show that it is

itself equivalent to the Lipschitz regularity in space of the optimal controls, uniformly with respect to time.

Proposition 4.8 (Coercivity and regularity). *The following statements are equivalent*

- (i) *The mean-field coercivity condition $\lambda > T$ holds.*
- (ii) *For any sequence of empirical measures (μ_N^0) converging narrowly towards $\mu^0 = \frac{1}{2}\mathbb{1}_{[-1,1]}\mathcal{L}^1$ along with the associated discrete optimal pairs $(\mathbf{u}_N^*(\cdot), \mathbf{x}_N^*(\cdot))$, it holds that*

$$|u_i^*(t) - u_j^*(t)| \leq \frac{1}{\rho_T} |x_i^*(t) - x_j^*(t)|,$$

for all $t \in [0, T]$, where $\rho_T = \lambda - T$ is the sharp mean-field coercivity constant of (\mathcal{P}_V) .

Proof. Suppose first that the uniform coercivity estimate does not hold, i.e. $\lambda \leq T$. Since the optimal controls are constant over $[0, T]$ as a consequence of (4.49), the total cost of (\mathcal{P}_V^N) can be rewritten as

$$\mathcal{C}(u_1, \dots, u_N) = \frac{\lambda T}{2N} \sum_{i=1}^N u_i^2 - \frac{1}{2N} \sum_{i=1}^N |x_i^0 + T u_i|^2 = \frac{1}{2N} \sum_{i=1}^N (T(\lambda - T)u_i^2 - 2T x_i^0 u_i - |x_i^0|^2).$$

for any N -tuple $\mathbf{u} = (u_1, \dots, u_N) \in [-C, C]^N$. Since $\lambda \leq T$, the minimum of \mathcal{C} is achieved by taking $u_i^* = \text{sign}(x_i^0)C$ for all $i \in \{1, \dots, N\}$. This further implies that

$$|u_i^* - u_j^*| = \begin{cases} 0 & \text{if } \text{sign}(x_i) = \text{sign}(x_j), \\ 2C & \text{otherwise,} \end{cases}$$

so that for any pair of indices such that $\text{sign}(x_i^0) = -\text{sign}(x_j^0)$, it holds that

$$|u_i^* - u_j^*| = \frac{2C}{|x_i^0 - x_j^0| + 2Ct} |x_i^*(t) - x_j^*(t)|. \quad (4.51)$$

The fact that $\mu_N \rightharpoonup^* \mu^0 = \frac{1}{2}\mathbb{1}_{[-1,1]}\mathcal{L}^1$ as $N \rightarrow +\infty$ implies that for all $\epsilon > 0$, there exists $N_\epsilon \geq 1$ such that for any $N \geq N_\epsilon$, there exists at least one pair of indices $i, j \in \{1, \dots, N\}$ such that $\text{sign}(x_i^0) = -\text{sign}(x_j^0)$ and $|x_i^0 - x_j^0| \leq \epsilon$. Thus, it follows from (4.51) that (ii) necessarily fails to hold some pairs of indices and at least for small times.

Suppose now that the mean-field coercivity estimate hold, i.e. $\lambda > T$, and denote by $\rho_T = \lambda - T$ the sharp coercivity constant. Let $I_N, J_N \subset \{1, \dots, N\}$ be the set of indices defined by

$$I_N = \left\{ i \in \{1, \dots, N\} \text{ s.t. } |x_i^0| \leq \rho_T C \right\}, \quad J_N = \{1, \dots, N\} \setminus I_N.$$

For N sufficiently big, I_N is necessarily non-empty since $\rho_T > 0$ and as a consequence of the narrow convergence of (μ_N^0) towards μ^0 . Then for any $i \in I_N$, one has that

$$|x_i^*(T)| \leq |x_i^0| + CT \leq (\rho_T + T)C = \lambda C,$$

whence for any such indices, the optimal controls are given by $u_i^* = \frac{1}{\lambda} x_i^*(T)$. In which

case, one has that

$$x_i^*(T) = x_i^0 + Tu_i^* \iff x_i^*(T) = \frac{x_i^0}{1 - T/\lambda} \quad \text{and} \quad u_i^* = \frac{x_i^*(t)}{\rho_T + t}$$

so that

$$|u_i^* - u_j^*| \leq \frac{|x_i^*(t) - x_j^*(t)|}{\rho_T + t}, \quad (4.52)$$

for any pair of indices $i, j \in I_N$. It can be checked reciprocally that $u_i^* = \text{sign}(x_i^0)C$ for any $i \in J_N$, which furthermore yields by (4.51) that

$$|u_i^* - u_j^*| \leq \begin{cases} 0 & \text{if } \text{sign}(x_i) = \text{sign}(x_j), \\ \frac{|x_i^*(t) - x_j^*(t)|}{\rho_T + t} & \text{otherwise,} \end{cases} \quad (4.53)$$

since in this case $|x_i^0 - x_j^0| \geq 2\rho_T C$ whenever $i, j \in J_N$ and $\text{sign}(x_i) = -\text{sign}(x_j)$. Suppose now that we are given a pair of indices $i, j \in \{1, \dots, N\}$ such that $i \in I_N$ and $j \in J_N$. If $\text{sign}(x_i^0) = \text{sign}(x_j^0)$, it holds that

$$\begin{aligned} |u_i^* - u_j^*| &= u_j^* - u_i^* = \text{sign}(x_j^0)C - \frac{x_i^*(t)}{\rho_T + t} \\ &= \frac{x_j^*(t)C}{|x_j^*(t)|} - \frac{x_i^*(t)}{\rho_T + t} \leq \frac{x_j^*(t) - x_i^*(t)}{\rho_T} = \frac{|x_i^*(t) - x_j^*(t)|}{\rho_T}, \end{aligned} \quad (4.54)$$

since $|x_j^*(t)| \geq \rho_T C$ by definition of J_N . Symmetrically if $\text{sign}(x_i^0) = -\text{sign}(x_j^0)$, one can easily show that

$$|u_i^* - u_j^*| \leq \frac{|x_i^*(t) - x_j^*(t)|}{\rho_T}. \quad (4.55)$$

By merging (4.52), (4.53), (4.54) and (4.55), we conclude that (ii) holds with the uniform constant $\frac{1}{\rho_T} > 0$ whenever the mean-field coercivity estimate holds, which ends the proof of our claim. \square

In Proposition 4.8, we have proven that the mean-field coercivity estimate is both necessary and sufficient for the existence of a uniform Lipschitz constant for the finite-dimensional optimal controls. It is clear when this condition fails that it is not possible to build a sequence of uniformly Lipschitz optimal maps $(u_N^*(\cdot, \cdot))$ for problem (\mathcal{P}_V^N) . Since the discrete optimal pairs control-trajectory $(\mathbf{u}_N^*(\cdot), \mathbf{x}_N^*(\cdot)) \in L^\infty([0, T], U^N) \times \text{Lip}([0, T], \mathbb{R}^N)$ are uniquely determined, we conclude that the mean-field coercivity condition (\mathbf{CO}_N) is necessary and sufficient in the limit for the existence of a Lipschitz-in-space mean-field optimal control for the Wasserstein optimal control problem (\mathcal{P}_V) .

Moreover, the explicit computations in the proof of Proposition 4.8 illustrates the statement of Remark 4.4. Namely, it is perfectly possible for a mean-field optimal control problem in which particle trajectories are not crossing each other to give rise to non-Lipschitz optimal controls.

5 Convergence Analysis and Sparse Control of Weakly Cooperative Alignment Models

The study of emerging patterns in dynamical systems describing collective behaviour has been the object of an increasing attention in the last decades. There is now a large literature devoted to the analysis of *consensus formation* in the class of so-called *cooperative systems*, see e.g. [127]. These systems are widely used, for example, to study crowd motion [61], robot swarms [26, 28, 75] and animal groups [9, 23, 29] such as bird flocks or fish schools.

Since the seminal paper [62] by Cucker and Smale, a great deal of interest has been manifested towards the analysis of the so-called *flocking behaviour* (see Definition 5.4 below), which describes the appearance of alignment patterns in second-order cooperative multi-agent systems. In [89], the authors proposed an alternative proof of the emergence of asymptotic flocking, based on Lyapunov methods. This method was then extended both to finite and infinite dimension [49], and to specific time and state-dependent interaction topologies [108]. It also allowed to control key models towards consensus and flocking [41, 42, 43, 115].

The structure of this chapter is the following. In Section 5.1, we prove a convergence towards consensus and flocking for first and second order weakly cooperative systems respectively, under a persistence of excitation condition. Then in Section 5.2, we build an explicit control strategy enabling to drive weakly-cooperative kinetic systems towards approximate alignment, which is a natural relaxation of flocking.

5.1 Convergence to consensus and flocking of randomly failed cooperative systems

In real world applications, the agents of a system may not systematically be able to interact with each others. Their communication channels may be subject e.g. to random failures, occurring during possibly extended periods of time. It is then crucial to verify whether convergence towards consensus or flocking is still guaranteed in this perturbed setting. For discrete-time first and second order systems, opinion formation models have been thoroughly investigated in a graph theoretic framework, see for instance the seminal paper [110]. Further results allowed to incorporate asymmetric communication rates and random communication failures e.g. in [64, 122]. However, to the best of our knowledge, there is no general proof of convergence for general time-continuous systems subject to random communication failures.

In this section, we investigate sufficient conditions for both asymptotic consensus and flocking formation based on Lyapunov methods. The main ingredient is the introduction of a condition of *persistence of excitation* (see Definition 5.3 below). This type of condition appeared quite recently in stability theory, and has proven its adaptability to

build strict Lyapunov functions, see [102, 103, 107]. Besides the interest of having a strict Lyapunov function, e.g. for studying input-to-state stability, persistence of excitation has a both deep and simple signification in cooperative dynamics. Indeed, it transcribes the fact that, on average on a given time window, the interaction graph describing a multi-agent system is connected with a prescribed lower bound on the intensity of this averaged interaction. This type of average connectedness assumption is standard when studying general time-varying interaction topologies (see e.g. [24, 110]), and it is even proven to be necessary for consensus in a large number of cases in [110]. In the way we formulate it, this condition further encodes the idea that one only requires the system to be persistently exciting with respect to the agents which have not yet reached consensus.

The structure of this section is the following. In Section 5.1.1, we prove the asymptotic consensus for persistently excited first-order dynamics. We then extend this result in Section 5.1.2 to asymptotic flocking for Cucker-Smale type systems with strongly interacting kernels in the sense of (5.13), which is the main result of this paper.

5.1.1 Consensus under persistent excitation for first-order dynamics

In this section, we introduce the tools used in the article, in the particular case of consensus formation. We study first-order cooperative systems of the form

$$(CS_1) \quad \begin{cases} \dot{x}_i(t) = \frac{1}{N} \sum_{i=1}^N \xi_{ij}(t) \phi(|x_i(t) - x_j(t)|) (x_j(t) - x_i(t)), \\ x_i(0) = x_i^0, \end{cases}$$

where $(x_1^0, \dots, x_N^0) \in (\mathbb{R}^d)^N$ is a given initial datum. We assume that the interaction kernel $\phi \in \text{Lip}(\mathbb{R}_+, \mathbb{R}_+^*)$ is strictly positive.

The functions $\xi_{ij} \in L^\infty(\mathbb{R}_+, [0, 1])$ represent communication rates, taking into account potential communication failures that can occur in the system (when $\xi_{ij}(t) < 1$). We require them to be symmetric, i.e. $\xi_{ij}(\cdot) = \xi_{ji}(\cdot)$. One of the main motivations for this choice of communication rates is to study consensus and flocking when *random interaction failures* occur. This article is the first step towards a more general theory for such systems, in which the $\xi_{ij}(\cdot)$ will be realizations of stochastic processes.

From now on, we use the notation $\mathbf{x} = (x_1, \dots, x_N)$ for the state in $(\mathbb{R}^d)^N$ and $\bar{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^N x_i$ for its mean value. For systems of the form (CS₁), we aim to study the formation of *asymptotic consensus*, defined as follows.

Definition 5.1 (Convergence to consensus). *A solution $\mathbf{x}(t)$ of (CS₁) asymptotically converges to consensus if $\lim_{t \rightarrow +\infty} |x_i(t) - \bar{\mathbf{x}}(t)| = 0$ for all $i \in \{1, \dots, N\}$.*

As a consequence of the symmetry of the rates $\xi_{ij}(\cdot)$, (CS₁) can be rewritten as

$$\dot{\mathbf{x}}(t) = -\mathbf{L}(t, \mathbf{x}(t))\mathbf{x}(t), \quad \mathbf{x}(0) = \mathbf{x}_0, \quad (CSM_1)$$

where $\mathbf{L} : \mathbb{R}_+ \times (\mathbb{R}^d)^N \rightarrow \mathcal{L}((\mathbb{R}^d)^N)$ is the so-called *graph Laplacian*, defined by

$$(\mathbf{L}(t, \mathbf{x})\mathbf{y})_i := \frac{1}{N} \sum_{j=1}^N \xi_{ij}(t) \phi(|x_i - x_j|) (y_i - y_j). \quad (5.1)$$

In the following, we will also use $\mathbf{L}_\xi(\cdot)$ defined by

$$(\mathbf{L}_\xi(t)\mathbf{y})_i := \frac{1}{N} \sum_{j=1}^N \xi_{ij}(t)(y_i - y_j). \quad (5.2)$$

Observe that both $\mathbf{L}(\cdot, \cdot)$ and $\mathbf{L}_\xi(\cdot)$ depend on the time-dependent communication rates $\xi_{ij}(\cdot)$, that are L^∞ functions, thus defined for almost every $t \in [0, +\infty)$.

The structure displayed in (5.1) is fairly general and allows for a comprehensive study of both consensus and flocking problems in a unified way via Lyapunov methods. With this goal in mind, we introduce the following bilinear form in the spirit of [41, 42].

Definition 5.2. *The variance bilinear form $B(\cdot, \cdot)$ is*

$$B(\mathbf{x}, \mathbf{y}) := \frac{1}{N} \sum_{i=1}^N \langle x_i, y_i \rangle - \langle \bar{\mathbf{x}}, \bar{\mathbf{y}} \rangle. \quad (5.3)$$

It is symmetric and positive semi-definite.

Such bilinear form is the squared distance of a given $\mathbf{x} \in (\mathbb{R}^d)^N$ from the so-called *consensus manifold* $\mathcal{C} = \{\mathbf{x} \in (\mathbb{R}^d)^N \text{ s.t. } x_1 = \dots = x_N\}$. As a consequence, $B(\mathbf{x}, \mathbf{x}) = 0$ if and only if $x_i = \bar{\mathbf{x}}$ for any index $i \in \{1, \dots, N\}$, i.e. if \mathbf{x} is a consensus.

We now list useful properties linking $B(\cdot, \cdot)$ and $\mathbf{L}(\cdot, \cdot)$.

Proposition 5.1 (Elementary properties of the graph Laplacian). *The graph Laplacian $\mathbf{L}(t, \mathbf{x})$ is positive-semi definite with respect to $B(\cdot, \cdot)$. Moreover, vectors of the form $\mathbf{L}(t, \mathbf{x})\mathbf{y}$ have zero mean.*

Proof. By summing over $i \in \{1, \dots, N\}$ the components in (5.1), the mean of $\mathbf{L}(t, \mathbf{x})\mathbf{y}$ is zero. As a consequence, and by symmetry of the communication rates $\xi_{ij}(\cdot)$, it holds

$$\begin{aligned} B(\mathbf{L}(t, \mathbf{x})\mathbf{y}, \mathbf{y}) &= \frac{1}{N^2} \sum_{i,j=1}^N \xi_{ij}(t) \phi(|x_i - x_j|) \langle y_i, y_i - y_j \rangle \\ &= \frac{1}{2N^2} \sum_{i,j=1}^N \xi_{ij}(t) \phi(|x_i - x_j|) |y_i - y_j|^2 \geq 0. \end{aligned}$$

which concludes the proof of our claim. \square

We now introduce a definition of *persistence of excitation* adapted to the analysis of consensus and flocking formation.

Definition 5.3 (Persistence of excitation condition). *Let $\tau, \mu > 0$ be given parameters. We say that the persistence of excitation condition (PE $_{\tau, \mu}$) holds for (CSM $_1$) if*

$$B\left(\left(\frac{1}{\tau} \int_t^{t+\tau} \mathbf{L}_\xi(s) ds\right) \mathbf{x}, \mathbf{x}\right) \geq \mu B(\mathbf{x}, \mathbf{x}). \quad (\text{PE}_{\tau, \mu})$$

for \mathcal{L}^1 -almost every $t \geq 0$.

Remark 5.1 (About the persistence of excitation). *The condition $(\text{PE}_{\tau,\mu})$ only involves the communication weights $\xi_{ij}(\cdot)$ through $\mathbf{L}_\xi(\cdot)$ and not the state of the system. Moreover, it is formulated using the bilinear form $B(\cdot, \cdot)$, representing the fact that one only needs the persistence to hold along directions orthogonal to the consensus manifold \mathcal{C} . Finally, $(\text{PE}_{\tau,\mu})$ can be interpreted as a connectivity condition of the time-averaged interaction graph of the system, which is fairly common in the literature, see [110] and [24, Chapter 2].*

In the following theorem, we state a general result of convergence for (CSM_1) , when $(\text{PE}_{\tau,\mu})$ holds.

Theorem 5.1 (Consensus in randomly failed cooperative systems). *Let $\phi(\cdot)$ be positive and non-increasing. Let $\xi_{ij} \in L^\infty(\mathbb{R}_+, [0, 1])$ satisfy $(\text{PE}_{\tau,\mu})$ for some $\tau, \mu > 0$. Then, any solution $\mathbf{x}(\cdot)$ of (CSM_1) asymptotically converges to consensus.*

Proof. Let c^2 be the operator norm of $\mathbf{L}(\cdot, \cdot)$ with respect to $B(\cdot, \cdot)$, i.e.

$$\begin{aligned} c^2 &:= \sup_{(t, \mathbf{x})} \|\mathbf{L}(\cdot, \cdot)\|_B \\ &= \sup_{(t, \mathbf{x}, \mathbf{y})} \left\{ \sqrt{B(\mathbf{L}(t, \mathbf{x})\mathbf{y}, \mathbf{L}(t, \mathbf{x})\mathbf{y})} \text{ s.t. } B(\mathbf{y}, \mathbf{y}) = 1 \right\}. \end{aligned}$$

Here, we made a slight abuse of notation by denoting by “sup” the essential supremum with respect to \mathcal{L}^1 . Denote by $X(\cdot)$ the *standard deviation* of $\mathbf{x}(\cdot)$, given by

$$X(t) := \sqrt{B(\mathbf{x}(t), \mathbf{x}(t))},$$

By definition of $B(\cdot, \cdot)$, $\mathbf{x}(\cdot)$ asymptotically converges to consensus if and only if one has that $X(t) \xrightarrow[t \rightarrow +\infty]{} 0$.

We define the time-dependent family of linear operators $\psi_\tau : \mathbb{R}_+ \rightarrow \mathcal{L}((\mathbb{R}^d)^N)$ along the trajectory $\mathbf{x}(\cdot)$ by

$$\psi_\tau(t) := (1 + c^2)\tau \text{Id} - \frac{1}{\tau} \int_t^{t+\tau} \int_t^s \mathbf{L}(\sigma, \mathbf{x}(\sigma)) d\sigma ds. \quad (5.4)$$

Then, $\psi_\tau(\cdot)$ is Lipschitz with pointwise derivative

$$\dot{\psi}_\tau(t) = \mathbf{L}(t, \mathbf{x}(t)) - \frac{1}{\tau} \int_t^{t+\tau} \mathbf{L}(s, \mathbf{x}(s)) ds. \quad (5.5)$$

By definition, it further holds

$$\sqrt{\tau}X(t) \leq \sqrt{B(\psi_\tau(t)\mathbf{x}, \mathbf{x})} \leq \sqrt{(1 + c^2)\tau}X(t). \quad (5.6)$$

Consider the candidate Lyapunov function

$$\mathcal{X}_\tau(t) := \lambda X(t) + \sqrt{B(\psi_\tau(t)\mathbf{x}(t), \mathbf{x}(t))} \quad (5.7)$$

where $\lambda > 0$ is a tuning parameter. This type of construction is rather recent in the design of strict Lyapunov function under persistent excitation assumption, see e.g. [102, 103, 107].

By (5.6), it holds that

$$(\lambda + \sqrt{\tau})X(t) \leq \mathcal{X}_\tau(t) \leq (\lambda + \sqrt{(1+c^2)\tau})X(t). \quad (5.8)$$

By Proposition 5.1, any solution $\mathbf{x}(\cdot)$ of (CSM₁) has constant mean, i.e. $\bar{\mathbf{x}}(\cdot) \equiv \bar{\mathbf{x}}_0$. By invariance with respect to translation of (CSM₁), we assume without loss of generality $\bar{\mathbf{x}}(\cdot) \equiv 0$ from now on. We now aim to prove a strict-dissipation inequality of the form

$$\dot{\mathcal{X}}_\tau(t) \leq -\alpha \mathcal{X}_\tau(t), \quad (5.9)$$

for some $\alpha > 0$. With this goal, we start by computing the time-derivative of $\mathcal{X}_\tau(\cdot)$

$$\begin{aligned} \dot{\mathcal{X}}_\tau(t) = & -\frac{\lambda}{X(t)} B(\mathbf{L}(t, \mathbf{x}(t)), \mathbf{x}(t)) + \frac{B(\dot{\psi}_\tau(t)\mathbf{x}(t), \mathbf{x}(t))}{2\sqrt{B(\psi_\tau(t)\mathbf{x}(t), \mathbf{x}(t))}} \\ & - \frac{B(\mathbf{L}(t, \mathbf{x}(t))\mathbf{x}(t), \psi_\tau(t)\mathbf{x}(t))}{\sqrt{B(\psi_\tau(t)\mathbf{x}(t), \mathbf{x}(t))}}. \end{aligned}$$

By (5.5)-(5.6), it holds

$$\begin{aligned} \dot{\mathcal{X}}_\tau(t) \leq & -\frac{1}{2\sqrt{(1+c^2)\tau}X(t)} B\left(\left(\frac{1}{\tau} \int_t^{t+\tau} \mathbf{L}(s, \mathbf{x}(s)) ds\right) \mathbf{x}(t), \mathbf{x}(t)\right) \\ & + \frac{1}{\sqrt{\tau}X(t)} B\left(\left(\frac{1}{\tau} \int_t^{t+\tau} \int_t^s \mathbf{L}(\sigma, \mathbf{x}(\sigma)) d\sigma ds\right) \mathbf{x}(t), \mathbf{L}(t, \mathbf{x}(t))\mathbf{x}(t)\right) \\ & + \frac{1}{\sqrt{\tau}X(t)} \left(\frac{1}{2} - \sqrt{(1+c^2)\tau} - \sqrt{\tau}\lambda\right) B(\mathbf{L}(t, \mathbf{x}(t))\mathbf{x}(t), \mathbf{x}(t)). \end{aligned} \quad (5.10)$$

To estimate the first line of (5.10), recall that first-order cooperative systems have uniformly compactly supported trajectories, see e.g. [115, Lemma 1]. Since $\phi(\cdot)$ is positive and continuous, there exists a positive constant $C(\mathbf{x}_0)$ such that

$$\min_{x_i \neq x_j} \phi(|x_i(t) - x_j(t)|) \geq C(\mathbf{x}_0)$$

for all times $t \geq 0$. By definition of $\mathbf{L}(\cdot, \cdot)$, this implies

$$B(\mathbf{L}(t, \mathbf{x}(t))\mathbf{y}, \mathbf{y}) \geq \frac{C(\mathbf{x}_0)}{2N^2} \sum_{i,j=1}^N \xi_{ij}(t) |y_i - y_j|^2 = C(\mathbf{x}_0) B(\mathbf{L}_\xi(t)\mathbf{y}, \mathbf{y}),$$

for any $\mathbf{y} \in (\mathbb{R}^d)^N$. By using (PE _{τ, μ}), it holds

$$B\left(\left(\frac{1}{\tau} \int_t^{t+\tau} \mathbf{L}(s, \mathbf{x}(s)) ds\right) \mathbf{x}(t), \mathbf{x}(t)\right) \geq C(\mathbf{x}_0) \mu X^2(t). \quad (5.11)$$

For the second line of (5.10), one has that

$$\begin{aligned}
& B\left(\frac{1}{\tau}\left(\int_t^{t+\tau}\int_t^s\mathbf{L}(\sigma,\mathbf{x}(\sigma))d\sigma ds\right)\mathbf{x}(s),\mathbf{L}(t,\mathbf{x}(t))\mathbf{x}(t)\right) \\
& \leq \tau c^2 X(t)\sqrt{B(\mathbf{L}(t,\mathbf{x}(t))\mathbf{x}(t),\mathbf{L}(t,\mathbf{x}(t))\mathbf{x}(t))} \\
& \leq \tau c^2 X(t)\|\mathbf{L}(t,\mathbf{x}(t))^{1/2}\|_B\sqrt{B(\mathbf{L}(t,\mathbf{x}(t))\mathbf{x}(t),\mathbf{x}(t))} \\
& \leq \tau c^3\left(\frac{\epsilon}{2}X(t)^2+\frac{1}{2\epsilon}B(\mathbf{L}(t,\mathbf{x}(t))\mathbf{x}(t),\mathbf{x}(t))\right),
\end{aligned} \tag{5.12}$$

for any $\epsilon > 0$, by definition of $\|\cdot\|_B$ and Young's inequality. Merge (5.10)-(5.11)-(5.12) and recall that $\mathbf{L}(\cdot,\cdot)$ is positive semi-definite to obtain

$$\begin{aligned}
\dot{\mathcal{X}}_\tau(t) & \leq -\left(\frac{C(\mathbf{x}_0)\mu}{2\sqrt{(1+c^2)}\tau}-\frac{c^3\sqrt{\tau}}{2}\epsilon\right)X(t) \\
& +\frac{1}{X(t)}\left(\frac{1}{2\sqrt{\tau}}+\frac{c^3\sqrt{\tau}}{2\epsilon}-\lambda\right)B(\mathbf{L}(t,\mathbf{x}(t))\mathbf{x}(t),\mathbf{x}(t)).
\end{aligned}$$

Choose

$$\epsilon=\frac{C(\mathbf{x}_0)\mu}{2c^3\tau\sqrt{(1+c^2)}},\quad\lambda=\frac{1}{2\sqrt{\tau}}+\frac{c^3\sqrt{\tau}}{2\epsilon}.$$

Using (5.8), we recover (5.9) for a positive constant $\alpha(\mu,c,\tau,C(\mathbf{x}_0))$. Then, it holds $\lim_{t\rightarrow+\infty}\mathcal{X}_\tau(t)=0$, and thus

$$\lim_{t\rightarrow+\infty}X(t)=0,$$

by (5.8). By definition of $X(\cdot)$, this implies that $\mathbf{x}(\cdot)$ converges to consensus. \square

5.1.2 Flocking for Cucker-Smale type systems with strong interactions

In this section, we derive sufficient conditions for the asymptotic convergence to flocking of general Cucker-Smale type dynamics subject to random communication failures. These systems are of the form

$$(\text{CS}_2)\quad\begin{cases}\dot{x}_i(t)=v_i(t), & x_i(0)=x_i^0, \\ \dot{v}_i(t)=\frac{1}{N}\sum_{j=1}^N\xi_{ij}(t)\phi(|x_i(t)-x_j(t)|)(v_j(t)-v_i(t)), & v_i(0)=v_i^0.\end{cases}$$

Similarly to Section 5.1.1, (CS₂) can be rewritten in matrix form using the graph Laplacian defined in (5.1):

$$\text{CSM}_2\quad\begin{cases}\dot{\mathbf{x}}(t)=\mathbf{v}(t), & \mathbf{x}(0)=\mathbf{x}_0, \\ \dot{\mathbf{v}}(t)=-\mathbf{L}(t,\mathbf{x}(t))\mathbf{v}(t), & \mathbf{v}(0)=\mathbf{v}_0.\end{cases}$$

We now recall the definition of flocking.

Definition 5.4 (Convergence to flocking). *A solution $(\mathbf{x}(\cdot),\mathbf{v}(\cdot))$ of (CSM₂) converges*

to flocking if for any $i \in \{1, \dots, N\}$ it holds

$$\sup_{t \geq 0} |x_i(t) - \bar{\mathbf{x}}(t)| < +\infty \quad \text{and} \quad \lim_{t \rightarrow +\infty} |v_i(t) - \bar{\mathbf{v}}(t)| = 0.$$

For this problem, we always assume to have $\phi(\cdot) \in \text{Lip}(\mathbb{R}_+, \mathbb{R}_+^*)$ non-increasing and such that

$$\int_0^{+\infty} \phi(r) dr = +\infty. \quad (5.13)$$

Equation (5.13) is known as a strong interaction condition, since it describes the fact that the interaction between agents does not decrease too fast when their distance goes to infinity.

Remark 5.2 (Flocking in the full communication setting). *When $\phi(\cdot)$ is uniformly bounded from below by a positive constant, then flocking in the full-communication setting occurs, see e.g. [62, 89, 115]. In our framework, this result is a simple consequence of Theorem 5.2. For positive kernels not satisfying (5.13), one can easily construct examples of initial conditions $(\mathbf{x}_0, \mathbf{v}_0)$ for which flocking does not occur, see [62].*

It can be checked easily that solutions of (CSM_2) satisfy

$$\dot{\bar{\mathbf{x}}}(t) = \bar{\mathbf{v}}(t), \quad \dot{\bar{\mathbf{v}}}(t) = 0.$$

Without loss of generality, we can therefore assume from now on that $\bar{\mathbf{x}}(\cdot) = \bar{\mathbf{v}}(\cdot) \equiv 0$. We define the candidate Lyapunov functionals by

$$X(t) := \sqrt{B(\mathbf{x}(t), \mathbf{x}(t))}, \quad V(t) := \sqrt{B(\mathbf{v}(t), \mathbf{v}(t))}$$

As a consequence of symmetry of $\xi_{ij}(\cdot)$, system (CSM_2) is *weakly dissipative* in the sense that

$$\dot{X}(t) \leq V(t), \quad \dot{V}(t) \leq 0, \quad (5.14)$$

along any solution $(\mathbf{x}(\cdot), \mathbf{v}(\cdot))$.

In their seminal paper [89], Ha and Liu produced a concise proof of the Cucker-Smale flocking based on the analysis of a system of *strictly dissipative inequalities*: if it holds that

$$\dot{X}(t) \leq V(t), \quad \dot{V}(t) \leq -\phi(2\sqrt{N}X(t))V(t), \quad (5.15)$$

where $\phi(\cdot)$ satisfies the strong interaction condition (5.13), then the system converges to flocking. Our aim is to adapt their strategy while taking into account possible communication failures. We prove the following, main result of this paper.

Theorem 5.2 (Flocking in randomly failed cooperative systems). *Let $(\text{PE}_{\tau, \mu})$ hold and $\phi(\cdot)$ be positive, non-increasing, and such that the following “very strong” interaction condition holds*

$$\int_0^{+\infty} \exp\left(-\sqrt{r} \phi \circ \Phi^{-1}(\sqrt{r})\right) dr < +\infty, \quad (5.16)$$

where $\Phi(\cdot)$ is the unique primitive of $r \mapsto \phi(2\sqrt{N}(r + \tau V(0)))$ which vanishes at $X(0)$. Then, any solution of (CSM_2) converges to flocking.

Remark 5.3 (On the “very strong” interaction condition). *The interaction condition (5.16) is satisfied e.g. whenever $\phi(\cdot)$ can be asymptotically bounded from below by a map of the form $r \mapsto \alpha \left(1 + \frac{1}{r^\beta}\right)$ with $\beta \in (0, \frac{1}{2})$. We stress that this condition – and possibly our proof strategy – is probably not sharp, since the expected exponent range for convergence to occur in this context should be $\beta \in (0, 1)$. It is not yet clear whether flocking can arise in the critical case $\beta = 1$ under the persistent excitation condition $(\text{PE}_{\tau, \mu})$.*

The proof of this result relies on the construction of a strict Lyapunov function for (CSM_2) , for which a system of inequalities akin to (5.15) holds only on a bounded time interval. This finite-time strict dissipation allows us to recover the asymptotic flocking of the system as a consequence of the weak dissipativity (5.14) of (CSM_2) . To the best of our knowledge, this combination of strict Lyapunov design and flocking analysis via systems of dissipative inequalities has not been covered by the existing literature.

We first prove a technical lemma.

Lemma 5.1. *Let $(\mathbf{x}(\cdot), \mathbf{v}(\cdot))$ be a solution of (CSM_2) . If $(\text{PE}_{\tau, \mu})$ holds, then one has*

$$B \left(\left(\frac{1}{\tau} \int_t^{t+\tau} \mathbf{L}(s, \mathbf{x}(s)) ds \right) \mathbf{w}, \mathbf{w} \right) \geq \mu \phi(2\sqrt{N}(X(t) + \tau V(0))) B(\mathbf{w}, \mathbf{w}) \quad (5.17)$$

for any $\mathbf{w} \in (\mathbb{R}^d)^N$.

Proof. By definition of $\mathbf{L}(\cdot, \cdot)$, it holds that

$$\begin{aligned} & B \left(\left(\frac{1}{\tau} \int_t^{t+\tau} \mathbf{L}(s, \mathbf{x}(s)) ds \right) \mathbf{w}, \mathbf{w} \right) \\ &= \frac{1}{2N^2} \sum_{i,j=1}^N \left(\frac{1}{\tau} \int_t^{t+\tau} \xi_{ij}(s) \phi(|x_i(s) - x_j(s)|) ds \right) |v_i - v_j|^2 \\ &\geq \frac{1}{2N^2} \sum_{i,j=1}^N \left(\frac{1}{\tau} \int_t^{t+\tau} \xi_{ij}(s) \phi(2\sqrt{N}X(s)) ds \right) |v_i - v_j|^2, \end{aligned}$$

where we used the fact that $\phi(\cdot)$ is non-increasing. As a consequence of the weak dissipation (5.14), one further has

$$X(s) = X(t) + \int_t^s \dot{X}(\sigma) d\sigma \leq X(t) + \tau V(0).$$

for all $s \in [t, t + \tau]$. By (5.18), and recalling again that $\phi(\cdot)$ is non-increasing, it holds

$$\begin{aligned} & B \left(\left(\frac{1}{\tau} \int_t^{t+\tau} \mathbf{L}(s, \mathbf{x}(s)) ds \right) \mathbf{w}, \mathbf{w} \right) \\ &\geq \frac{\phi(2\sqrt{N}(X(t) + \tau V(0)))}{2N^2} \sum_{i,j=1}^N \left(\frac{1}{\tau} \int_t^{t+\tau} \xi_{ij}(s) ds \right) |v_i - v_j|^2 \\ &= \phi(2\sqrt{N}(X(t) + \tau V(0))) B \left(\left(\frac{1}{\tau} \int_t^{t+\tau} \mathbf{L}_\xi(s) ds \right) \mathbf{w}, \mathbf{w} \right) \\ &= \mu \phi(2\sqrt{N}(X(t) + \tau V(0))) B(\mathbf{w}, \mathbf{w}), \end{aligned}$$

where we used $(\text{PE}_{\tau,\mu})$ in the last inequality. \square

We now define the candidate Lyapunov function

$$\mathcal{V}_\tau(t) := \lambda(t)V(t) + \sqrt{B(\psi_\tau(t)\mathbf{v}(t), \mathbf{v}(t))} \quad (5.18)$$

where $\psi_\tau(\cdot)$ is defined in (5.4) and $\lambda(\cdot)$ is a tuning curve, smooth with the respect to time. We have the following lemma.

Lemma 5.2. *For any $\epsilon_0 > 0$, there exists $T_{\epsilon_0}^* > 0$ such that for almost every $t \in [0, 2T_{\epsilon_0}^*]$, it holds*

$$\dot{\mathcal{V}}_\tau(t) \leq -\frac{\mu \phi(2\sqrt{N}(X(t) + \tau V(0)))}{2\sqrt{(1+c^2)\tau}} V(t). \quad (5.19)$$

Proof. Following the proof of Theorem 5.1, we have

$$\begin{aligned} \dot{\mathcal{V}}_\tau(t) &\leq \frac{\left(\frac{1}{2\sqrt{\tau}} + \frac{c^3\tau}{2\epsilon(t)} - \sqrt{\tau}\lambda(t)\right)}{\sqrt{\tau}V(t)} B(\mathbf{L}(t, \mathbf{x}(t))\mathbf{v}(t), \mathbf{v}(t)) \\ &\quad - \left(\frac{\mu \phi(2\sqrt{N}(X(t) + \tau V(0)))}{2\sqrt{(1+c^2)\tau}} - \frac{c^3\sqrt{\tau}}{2}\epsilon(t) - \dot{\lambda}(t)\right) V(t). \end{aligned} \quad (5.20)$$

The two differences with respect to the proof of Theorem 5.1 are the choice of *time-dependent families of parameters* $(\lambda(\cdot), \epsilon(\cdot))$ and the use of (5.17) instead of $(\text{PE}_{\tau,\mu})$.

Choose

$$\lambda(t) = \frac{1}{2\sqrt{\tau}} + \frac{c^3\sqrt{\tau}}{2\epsilon(t)}. \quad (5.21)$$

This implies in particular that $\dot{\lambda}(t) = -\frac{c^3\sqrt{\tau}}{2\epsilon^2(t)}\dot{\epsilon}(t)$. Choose now $\epsilon(\cdot)$ as the solution of

$$\dot{\epsilon}(t) = \epsilon^3(t), \quad \epsilon(0) = \epsilon_0,$$

for a given constant $\epsilon_0 > 0$, i.e. $\epsilon(\cdot)$ defined by

$$\epsilon(t) = \frac{\epsilon_0}{\sqrt{1 - 2\epsilon_0^2 t}}, \quad (5.22)$$

for $t \in [0, 1/2\epsilon_0^2]$. Then, (5.20) reads as

$$\dot{\mathcal{V}}_\tau(t) \leq -\frac{\mu \phi(2\sqrt{N}(X(t) + \tau V(0)))}{2\sqrt{(1+c^2)\tau}} V(t),$$

and (5.19) holds with $2T_{\epsilon_0}^* = 1/2\epsilon_0^2$. \square

Observe that (5.19) involves both $\mathcal{V}_\tau(\cdot)$ and $V(\cdot)$. We now aim to find an estimate involving $V(\cdot)$ only.

Proposition 5.2. *For any $\epsilon_0 > 0$, there exists a constant $X_M(\epsilon_0) > 0$ such that*

$\sup_{t \geq 0} X(t) \leq X_M(\epsilon_0)$. Moreover, it holds that

$$V(T_{\epsilon_0}^*) \leq \left(\frac{\alpha_1 + \beta_1 \epsilon_0}{\alpha_2 + \beta_2 \epsilon_0} \right) V(0) \exp \left(- \frac{\mu \phi(2\sqrt{N}(X_M(\epsilon_0) + \tau V(0)))}{4(\alpha_3 + \beta_3 \epsilon_0) \epsilon_0} \right), \quad (5.23)$$

where $\{\alpha_k, \beta_k\}_{k=1}^3$ are positive constants depending on (c, τ) only.

Proof. Choose $\epsilon_0 \in (0, 1]$ and denote by $(\lambda(\cdot), \epsilon(\cdot))$ the corresponding functions given by (5.21)-(5.22) respectively.

Similarly to (5.6), it holds that

$$\sqrt{\tau} V(t) \leq \sqrt{B(\psi_\tau(t) \mathbf{v}(t), \mathbf{v}(t))} \leq \sqrt{(1 + c^2) \tau} V(t).$$

By definition of $\mathcal{V}_\tau(\cdot)$ in (5.18), it then holds that

$$\begin{cases} \mathcal{V}_\tau(0) \leq \left(\sqrt{(1 + c^2) \tau} + \frac{1}{2\sqrt{\tau}} + \frac{c^3 \sqrt{\tau}}{2\epsilon_0} \right) V(0), \\ \mathcal{V}_\tau(t) \geq \left(\sqrt{\tau} + \frac{1}{2\sqrt{\tau}} + \frac{c^3 \sqrt{2\tau}}{4\epsilon_0} \right) V(t), \end{cases}$$

for any $t \in [0, T_{\epsilon_0}^*]$, where we used the fact that $\epsilon(t) \leq \sqrt{2}\epsilon_0$ on this time interval. By simple identification of the coefficients, these estimates can be rewritten as

$$\mathcal{V}_\tau(0) \leq \left(\frac{\alpha_1}{\epsilon_0} + \beta_1 \right) V(0) \quad \text{and} \quad \mathcal{V}_\tau(t) \geq \left(\frac{\alpha_2}{\epsilon_0} + \beta_2 \right) V(t)$$

for positive constants $\{\alpha_k, \beta_k\}_{k=1}^2$ depending on (c, τ) .

We then integrate (5.19) on $[0, t]$ to obtain

$$\mathcal{V}_\tau(t) \leq \mathcal{V}_\tau(0) - \frac{\mu}{2\sqrt{(1 + c^2) \tau}} \int_0^t \phi(2\sqrt{N}(X(s) + \tau V(0))) V(s) ds.$$

This in turn implies that

$$V(t) \leq \left(\frac{\alpha_1 + \beta_1 \epsilon_0}{\alpha_2 + \beta_2 \epsilon_0} \right) V(0) - \frac{\mu \epsilon_0}{\alpha_3 + \beta_3 \epsilon_0} \int_0^t \phi(2\sqrt{N}(X(s) + \tau V(0))) V(s) ds, \quad (5.24)$$

where $(\alpha_3, \beta_3) = 2\sqrt{(1 + c^2) \tau}(\alpha_2, \beta_2)$. Recall now that $\dot{X}(s) \leq V(s)$ by (5.14). Hence by applying the change of variable $r = X(s)$ in (5.24), we further obtain that

$$V(t) \leq \left(\frac{\alpha_1 + \beta_1 \epsilon_0}{\alpha_2 + \beta_2 \epsilon_0} \right) V(0) - \frac{\mu \epsilon_0}{\alpha_3 + \beta_3 \epsilon_0} \int_{X(0)}^{X(t)} \phi(2\sqrt{N}(r + \tau V(0))) dr. \quad (5.25)$$

Remark that by the strong interaction condition (5.13), there exists a radius $X_M(\epsilon_0) > 0$ such that

$$\frac{\mu \epsilon_0}{\alpha_3 + \beta_3 \epsilon_0} \int_{X(0)}^{X_M(\epsilon_0)} \phi(2\sqrt{N}(r + \tau V(0))) dr = \left(\frac{\alpha_1 + \beta_1 \epsilon_0}{\alpha_2 + \beta_2 \epsilon_0} \right) V(0). \quad (5.26)$$

Since $V(\cdot)$ is a non-negative quantity by definition, (5.26) automatically implies along with (5.25) that

$$\sup_{t \in [0, T_{\epsilon_0}^*]} X(t) \leq X_M(\epsilon_0), \quad (5.27)$$

where $X_M(\epsilon_0)$ can be expressed synthetically as

$$X_M(\epsilon_0) = \Phi^{-1} \left(\frac{2\sqrt{(1+c^2)\tau}(\alpha_1 + \beta_1\epsilon_0)}{\mu\epsilon_0} \right), \quad (5.28)$$

with $\Phi(\cdot)$ defined as in the statement of Theorem 5.2.

Recalling now that $\phi(\cdot)$ is non-increasing, it further holds that

$$\phi(2\sqrt{N}(X(t) + \tau V(0))) \geq \phi(2\sqrt{N}(X_M(\epsilon_0) + \tau V(0))).$$

for all times $t \in [0, T_{\epsilon_0}^*]$. Thus, (5.24) yields that

$$V(t) \leq \left(\frac{\alpha_1 + \beta_1\epsilon_0}{\alpha_2 + \beta_2\epsilon_0} \right) V(0) - \frac{\mu\epsilon_0}{\alpha_3 + \beta_3\epsilon_0} \phi(2\sqrt{N}(X_M(\epsilon_0) + \tau V(0))) \int_0^t V(s) dr.$$

By Grönwall's Lemma, it further holds

$$V(T_{\epsilon_0}^*) \leq \left(\frac{\alpha_1 + \beta_1\epsilon_0}{\alpha_2 + \beta_2\epsilon_0} \right) V(0) \exp \left(- \frac{\mu\phi(2\sqrt{N}(X_M(\epsilon_0) + \tau V(0)))}{4(\alpha_3 + \beta_3\epsilon_0)\epsilon_0} \right)$$

where we used the fact that $T_{\epsilon_0}^* = 1/4\epsilon_0^2$. □

We now prove Theorem 5.2, that is our main result.

Proof. We start by proving the formation of consensus in the velocity variables. We first fix a parameter $\delta > 0$. Observe that as a consequence of (5.16) and of the monotonicity of $\phi \circ \Phi^{-1}(\cdot)$, it holds in particular that

$$\begin{aligned} \lim_{\epsilon \rightarrow 0^+} V(T_{\epsilon_0}^*) &= \left(\frac{\alpha_1}{\alpha_2} \right) \lim_{\epsilon_0 \rightarrow 0^+} \exp \left(- \frac{\mu\phi(2\sqrt{N}(X_M(\epsilon_0) + \tau V(0)))}{4(\alpha_3 + \beta_3\epsilon_0)\epsilon_0} \right) \\ &= \left(\frac{\alpha_1}{\alpha_2} \right) \lim_{\epsilon_0 \rightarrow 0^+} \exp \left(- \frac{\mu\alpha_1\phi \left(2\sqrt{N} \left(\Phi^{-1} \left(\frac{2\sqrt{(1+c^2)\tau}}{\mu\epsilon_0} \right) + \tau V(0) \right) \right)}{4\alpha_3\epsilon_0} \right) = 0. \end{aligned}$$

Whence there exists a constant ϵ_0 such that $V(T_{\epsilon_0}^*) \leq \delta$ where $T_{\epsilon_0}^* = 1/4\epsilon_0^2$. Since (CSM₂) is dissipative with respect to the velocity variable, one has that $V(t) \leq \delta$ for all times $t \geq T_{\epsilon_0}^*$. Since this estimate holds for any $\delta > 0$, we can conclude that $\lim_{t \rightarrow +\infty} V(t) = 0$. By definition of $V(\cdot)$, this in turn yields

$$\lim_{t \rightarrow +\infty} |v_i(t) - \bar{v}(t)| = 0,$$

for all $i \in \{1, \dots, N\}$.

We now turn our attention to the uniform boundedness of the position mean-deviation. Recall that by the weak dissipativity of (CSM₂), it holds that

$$X(t) \leq X(0) + \int_0^t V(s)ds \leq X(0) + \int_0^{+\infty} V(t)dt.$$

As a consequence of Lemma 5.2 we know that for any $\epsilon > 0$, one has that

$$V(T_{\epsilon_0}^*) \leq C \exp \left(- \frac{\mu T_{\epsilon_0}^* \phi \left(2\sqrt{N} \left(\Phi^{-1} \left(\frac{4}{\mu} \sqrt{(1+c^2)\tau T_{\epsilon_0}^*} \right) + \tau V(0) \right) \right)}{8(\beta_3 + 2\alpha_3 \sqrt{T_{\epsilon_0}^*})} \right)$$

where $C_1 > 0$ is a positive constant. Since $T_{\epsilon_0}^*$ spans \mathbb{R}_+^* as ϵ_0 spans \mathbb{R}_+^* , it is possible to reparametrize the time variable using $T \equiv T_{\epsilon_0}^*$ and to recover that

$$X(t) \leq V(0) + C_1 \int_0^{+\infty} \exp \left(-C_2 \sqrt{T} \phi \circ \Phi^{-1} \left(C_3 + C_4 \sqrt{T} \right) \right) dT < +\infty,$$

for all times $t \geq 0$. As a consequence of the “very strong” interaction condition (5.16), we finally conclude that

$$\sup_{t \geq 0} |x_i(t) - \bar{x}(t)| < +\infty.$$

for each $i \in \{1, \dots, N\}$. □

5.2 Sparse control of kinetic cooperative systems to approximate alignment

In Section 5.1, we have seen that flocking does not automatically arise in weakly cooperative systems. Even for Cucker-Smale type kernels in the full-communication setting, asymptotic flocking arises conditionally to a confinement condition on the initial support of the crowd, see e.g. [62]. It is therefore relevant to search for control strategies allowing to stir weakly cooperative systems towards alignment. A strong emphasis in this endeavour has been put on the *sparsity* of such control strategies, see e.g. [41, 42, 115, 116]. In this section, we focus on second-order multi-agent systems of the form

$$\begin{cases} \dot{x}_i(t) = v_i(t), & x_i(0) = x_i^0, \\ \dot{v}_i(t) = \frac{1}{N} \sum_{j=1}^N \psi(x_j(t) - x_i(t), v_j(t) - v_i(t)), & v_i(0) = v_i^0, \end{cases} \quad (5.29)$$

where the map $\psi \in \text{Lip}(\mathbb{R}^{2d}, \mathbb{R}^d)$ satisfies

$$\langle \psi(x, v), v \rangle \leq 0, \quad (5.30)$$

for any $v \in \mathbb{R}^d$. Notice that (5.29)-(5.30) define a more general notion of weakly-cooperative system than the Cucker-Smale variant (CS₂) presented before where we explicitly attributed the loss of strict cooperation to communication failures.

In this section, we are interested in designing a sparse control strategy for the kinetic approximation of (5.29). As already sketched in the Introduction and throughout this

manuscript, such systems are the *mean field limits* of finite-dimensional cooperative systems when the number N of agents goes to infinity. We have seen that the natural way to represent the crowd in this formalism is by means of a curve of measures $t \in \mathbb{R}_+ \mapsto \mu(t) \in \mathcal{P}_c(\mathbb{R}^{2d})$. The time-evolution of this curve is then described by a continuity equation with non-local terms of the general form

$$\partial_t \mu(t) + \nabla \cdot \left((v, \Psi[\mu(t)](x, v))^\top \mu(t) \right) = 0, \quad (5.31)$$

where the non-local velocity field $\Psi[\cdot](\cdot, \cdot)$ is defined by

$$\Psi[\mu] : (x, v) \in \mathbb{R}^{2d} \mapsto \int_{\mathbb{R}^{2d}} \psi(y - x, w - v) d\mu(y, w).$$

for any $\mu \in \mathcal{P}_c(\mathbb{R}^{2d})$. Throughout this section, we will restrict our attention to initial distributions μ^0 which are absolutely continuous with respect to \mathcal{L}^{2d} . In keeping with the notations of Theorem 1.7, we denote by $(\Phi_{(0,t)}[\mu^0](\cdot, \cdot))_{t \geq 0}$ the family of flows of diffeomorphisms generated by the non-local velocity field $(x, v) \mapsto (v, \Psi[\mu](x, v))$.

By now, it should be clear that the most natural way of controlling the velocity part of (5.31) is by means of a vector field $(t, x, v) \in \mathbb{R}_+ \times \mathbb{R}^{2d} \mapsto u(t, x, v) \in \mathbb{R}^d$. We explicitly denote time-varying space support of $u(\cdot, \cdot, \cdot)$ by the curve of Borel sets $t \in \mathbb{R}_+ \mapsto \omega(t)$. The controlled version of (5.31) can then be written as

$$\partial_t \mu(t) + \nabla \cdot \left((v, \Psi[\mu(t)](x, v) + \mathbf{1}_{\omega(t)} u(t, x, v)) \mu(t) \right) = 0. \quad (5.32)$$

To ensure the well-posedness of the controlled dynamics in the sense of Theorem 1.7, we shall design our control strategy in such a way that $\chi_{\omega(t)} u(t, \cdot, \cdot)$ defines a Lipschitz vector field for all times $t \geq 0$.

One of the key interests of formulating a control problem of the form (5.32) is that the resulting control strategies are inherently bound to be N -free. Moreover, under Cauchy-Lipschitz regularity requirements, the mean-field control strategy designed for (5.32) can be applied to approximately control finite-dimensional systems of the form (5.29). The corresponding control error can then be easily estimated as a function of N and of the Lipschitz constants of the vector fields driving the dynamics. As advertised previously, we are also going to impose *sparsity* and *boundedness* constraint on our control strategies, so that it can only act on a small portion of the crowd at all times, and with limited amplitude. In our developments, these constraints can be expressed as shown below.

◇ The following **population sparsity constraint**

$$\int_{\omega(t)} d\mu(t)(x, v) \leq c, \quad (5.33)$$

holds for all times $t \geq 0$.

◇ The following **boundedness constraint**

$$\|u(t, \cdot, \cdot)\|_{L^\infty(\mathbb{R}^d \times \mathbb{R}^d)} \leq 1, \quad (5.34)$$

holds for all times $t \geq 0$

Remark 5.4 (Space sparsity constraint). *Another sparsity constraint for kinetic systems is the space sparsity constraint, which writes*

$$\int_{\omega(t)} dx dv \leq c, \quad (5.35)$$

for all $t \geq 0$. This type of constraint has already been treated in [116]. The constraint (5.33) is harder to handle, since it requires to monitor accurately the evolution of the density $\mu(\cdot)$ in order to ensure that the control acts only on a small portion of the crowd. The constraint (5.35) is less restrictive since one only needs to act on a small area of the phase space, with no restriction on the crowd density within it.

In Section 5.1, we studied *flocking* behaviour, i.e. the asymptotic convergence towards alignment in the velocity variable for a Cucker-Smale type system. In the present case, we are going to focus on a relaxation of this notion, called the *approximate alignment*, which definition we introduce below.

Definition 5.5 (Approximate alignment). *A solution $\mu(\cdot) \in C^0(\mathbb{R}_+, \mathcal{P}_c(\mathbb{R}^{2d}))$ of (5.32) is said to be ϵ -approximately aligned around a given velocity $v^* \in \mathbb{R}^d$ starting from time T provided that*

$$\text{supp}(\mu(t)) \subset \mathbb{R}^d \times B(v^*, \epsilon)$$

for all $t \geq T$.

Before moving on to the statement of the main result of this section, let us comment on the differences between approximate alignment and flocking as they are presented in this manuscript. First, the notion of approximate alignment only involves a bound on the velocity part of the support of the measure whereas in the definition of flocking, one also requires that the positions do not diverge. Second, the notion of flocking introduced in Definition 5.4 measures the variance of the measure in the velocity component – and is therefore an L^2 -type quantity – whereas the approximate alignment of Definition 5.5 measures the size of the support – hence an L^∞ -type quantity. While both concepts are equivalent for finite-dimensional cooperative systems, only the latter allows for the definition of flocking and other related alignment patterns for kinetic equations of the form (5.31), see e.g. [89].

In the following Theorem, we state the main result of this section.

Theorem 5.3 (Sparse ϵ -alignment). *Let $\mu^0 \in \mathcal{P}_c^{ac}(\mathbb{R}^{2d})$ be a given initial datum for (5.32), and let $c, \epsilon > 0$ be positive constants and $v^* \in \mathbb{R}^d$. Then, there exists a time $T > 0$ depending on μ^0, c, v^* and ϵ and a Lipschitz-in-space control strategy*

$$(t, x, v) \in [0, T] \times \mathbb{R}^{2d} \mapsto \mathbf{1}_{\omega(t)}(x, v)u(t, x, v),$$

satisfying (5.33)-(5.34), and such that the corresponding solution $\mu(\cdot) \in C^0(\mathbb{R}_+, \mathcal{P}_c^{ac}(\mathbb{R}^{2d}))$ of (5.32) is approximately aligned around v^* with precision ϵ starting from time T .

This section is structured as follows. In Section 5.2.1, we review some of the basic properties of kinetic cooperative systems described by (5.32). We then prove Theorem 5.3 in Section 5.2.2.

5.2.1 Invariance properties of kinetic cooperative systems

In the following paragraphs, we establish some structural lemmas on kinetic cooperative systems. Some of the most fundamental properties of (5.29) are its invariance with respect to translations and the fact that the size of its velocity support cannot increase. These properties are canonically inherited by its mean-field counterpart (5.31) and can be stated as follows. The proof of both results is the matter of a few lines of computations.

Lemma 5.3. *Let $\mu(\cdot)$ be a solution of (5.31) with initial datum $\mu^0 \in \mathcal{P}_c^{ac}(\mathbb{R}^{2d})$ and for all times $t \geq 0$ let $f(t, \cdot, \cdot) \in L^\infty(\mathbb{R}^{2d}, \mathbb{R}^{2d})$ be its density with respect to \mathcal{L}^{2d} . Then for any $(y, w) \in \mathbb{R}^d \times \mathbb{R}^d$, the curve of measures $\tilde{\mu} : t \in \mathbb{R}_+ \mapsto f(t, x + y + tw, v + w) \cdot \mathcal{L}^{2d}$ is the unique solution of (5.31) with initial datum $(\text{Id} + y, \text{Id} + w)_\# \mu^0$.*

Besides, the attractivity of the interaction forces of (5.29) allows us to establish an easy estimate of the evolution through time of the support of a solution of (5.31).

Lemma 5.4. *Let $\mu(\cdot)$ be a solution of (5.31) with initial datum $\mu^0 \in \mathcal{P}_c^{ac}(\mathbb{R}^d)$ and assume that there exists a collection of real numbers $(\underline{X}^i, \bar{X}^i, \underline{V}^i, \bar{V}^i)_{1 \leq i \leq d} \in \mathbb{R}^{4d}$ such that*

$$\text{supp}(\mu^0) \subseteq \prod_{i=1}^d \left([\underline{X}^i, \bar{X}^i] \times [\underline{V}^i, \bar{V}^i] \right).$$

Then, it holds that

$$\text{supp}(\mu(t)) \subseteq \prod_{i=1}^d \left([\underline{X}^i + t\underline{V}^i, \bar{X}^i + t\bar{V}^i] \times [\underline{V}^i, \bar{V}^i] \right)$$

for all times $t \geq 0$.

Proof. This invariance is a direct consequence of the fact that $(x, v) \mapsto (v, \Psi[\mu](x, v))^T$ always points inward $\mathbb{R} \times [\underline{V}^i, \bar{V}^i]$ along the v direction, and this for each $i \in \{1, \dots, N\}$. This combined with the fact that

$$\text{supp}(\mu(t)) = \Phi_{(0,t)}[\mu^0] \left(\text{supp}(\mu^0) \right)$$

yields the proof of our claim. \square

We end these preparatory paragraphs by the statement of an L^∞ -estimate on the density of a solution of (5.32).

Proposition 5.3. *Let $\mu(\cdot) \in C^0(\mathbb{R}_+, \mathcal{P}_c^{ac})$ be a solution of (5.32) with initial datum $\mu^0 \in \mathcal{P}_c^{ac}$ and $f(t, \cdot) \in L^\infty(\mathbb{R}^{2d}, \mathbb{R}^{2d})$ be its density with respect to \mathcal{L}^{2d} . Then, it holds that*

$$\begin{aligned} \frac{d}{dt} \|f(t, \cdot, \cdot)\|_{L^\infty(\mathbb{R}^{2d})} \leq & \|f(t, \cdot, \cdot)\|_{L^\infty(\mathbb{R}^{2d})} \left(\|\nabla_v \cdot (\mathbb{1}_{\omega(t)}(\cdot, \cdot)u(t, \cdot, \cdot))\|_{L^\infty(\mathbb{R}^{2d})} \right. \\ & \left. + \|\nabla_v \cdot \Psi[\mu(t)](\cdot, \cdot)\|_{L^\infty(\text{supp}(f(t, \cdot)))} \right) \end{aligned} \quad (5.36)$$

for all times $t \geq 0$.

Proof. The proof of this result is elementary and can be found in [115, Section 4.2]. \square

5.2.2 Proof of Theorem 5.3

In this section we prove our main result Theorem 5.3, using a constructive algorithmic approach, in the spirit of [115, 116]. We start by showing in Step 1 how the invariance properties of cooperative systems listed hereinabove allow us to reduce the problem to the case where $v^* = 0$ and $\text{supp}(\mu^0) \subset [0, +\infty)^{2d}$. We proceed by defining in Step 2 the fundamental step of our control strategy in the particular case of space dimension $d = 1$, and we then show in Step 3 how its iteration steers the dynamics to approximate alignment in the sense of Definition 5.5. In Step 4, we finally show how this procedure can be generalized to the higher dimensions in which $d > 1$.

Step 1 : Reduction to the case $v^* = 0$ and $\text{supp}(\mu^0) \subset [0, +\infty)^{2d}$

Let $\mu^0 \in \mathcal{P}_c^{ac}(\mathbb{R}^{2d})$ be given as in Theorem 5.3 and $(\underline{X}^i, \bar{X}^i, \underline{V}^i, \bar{V}^i)_{1 \leq i \leq d} \in \mathbb{R}^{4d}$ be such that

$$\text{supp}(\mu^0) \subseteq \prod_{i=1}^d \left([\underline{X}^i, \bar{X}^i] \times [\underline{V}^i, \bar{V}^i] \right)$$

As a consequence of the invariance properties stated in Lemma 5.3, it is possible to restrict the proof of Theorem 5.3 without loss of generality to the case where

$$v^* = 0, \quad \underline{X}^i = 0, \quad \bar{X}^i > 0, \quad \underline{V}^i = 0 \text{ and } \bar{V}^i > 0,$$

for any $i \in \{1, \dots, N\}$. Indeed, one can always achieve approximate alignment in the sense of Definition 5.5 in dimension j with $\underline{V}^j < 0 < \bar{V}^j$ by applying the following strategy.

- (1) Define ¹ $\tilde{\mu}_j(t, x, v) = \mu(t, x - t\underline{V}^j e_j, v - \underline{V}^j e_j)$ – where e_j stands for the j -th unitary vector of \mathbb{R}^d – and notice that it follows the dynamics (5.31) by Lemma 5.3. It is then possible to perform an approximate alignment around $v_j^* = 0$ with precision $\frac{\epsilon}{\sqrt{d}} - \underline{V}^j$. The velocity support of the system in dimension j is now $[\underline{V}^j, \frac{\epsilon}{\sqrt{d}}]$.
- (2) Define $\hat{\mu}_j(t, x, v) = \mu(t, -x - t\frac{\epsilon}{\sqrt{d}}e_j, -v - \frac{\epsilon}{\sqrt{d}}e_j)$ and notice that it satisfies the dynamics (5.31) with $\psi'(x, v) = -\psi(-x, -v)$. Proceed again by performing an approximate alignment around 0 with precision $\frac{2\epsilon}{\sqrt{d}}$. The velocity support for the initial measure is included in $(-\frac{\epsilon}{\sqrt{d}}, \frac{\epsilon}{\sqrt{d}})$ and approximate alignment is achieved for the j -th component.

Step 2 : Fundamental step in the case $d = 1$

Throughout Step 2 and Step 3, we make the assumption that $d = 1$. As a consequence of Step 1, it is furthermore possible to restrict our attention to the case where $\text{supp}(\mu^0) \subset [0, X_0] \times [0, V_0]$.

Our aim is now to build a Lipschitz-in-space control $\chi_{\omega(\cdot)} u^0(\cdot, \cdot, \cdot)$ and a time T_0 such that the control satisfies the constraints (5.33) and (5.34) at all times in $[0, T_0]$, and such that $\text{supp}(\mu(T_0))$ is strictly smaller than $\text{supp}(\mu^0)$ along the velocity component. To this

1. Here we make a slight abuse of notations by identifying the curves of measures $\tilde{\mu}_j(\cdot)$ and $\hat{\mu}_j(\cdot)$ with their densities with respect to \mathcal{L}^{2d} .

end, we define a partition of our initial domain $[0, X_0] \times [0, V_0]$ into $n = \lceil \frac{2}{c} \rceil$ rectangles $(\Omega_{[i]}^0)_{1 \leq i \leq n}$ defined by

$$\Omega_{[i]}^0 = [x_{[i-1]}, x_{[i]}] \times [0, V_0],$$

for any $i \in \{1, \dots, n\}$. The points $(x_{[i]})_{1 \leq i \leq n-1}$ are defined recursively via

$$x_{[i]} = \operatorname{argmin}_{x \in [x_{[i-1]}, X_0]} \left\{ x \text{ s.t. } \mu^0([x_{[i-1]}, x] \times [0, V_0]) = \frac{c}{2} \right\} \quad (5.37)$$

for all $i \in \{1, \dots, n\}$, starting from $x_{[0]} = 0$. We also define $x_{[n]} = X_0$. Remark that the points $x_{[i]}$ are well defined since $\mu^0 \in \mathcal{P}_c^{ac}(\mathbb{R}^2)$, which ensures that $x \mapsto \mu^0([x_{[i-1]}, x] \times [0, V_0])$ is continuous. We further define the parameter δ_0 as the biggest real number such that

$$\mu^0([x_{[i-1]} - 3\delta_0, x_{[i]} + 3\delta_0] \times \mathbb{R}) \leq c \quad (5.38)$$

for any $i \in \{1, \dots, n\}$. We finally fix a parameter $\eta_0 \in (0, V_0)$ and a time $T_0 \in \left(0, \frac{\delta_0}{V_0}\right]$, which precise choices shall be detailed in Step 3.

We define the control sets $(\omega_{[i]}^0)_{1 \leq i \leq n}$ by

$$\omega_{[i]}^0 = [x_{[i-1]} - 2\delta_0, x_{[i]} + 2\delta_0] \times [0, V_0 + \eta_0], \quad (5.39)$$

and we introduce the corresponding control functions $u_{[i]} : (t, x, v) \mapsto -\zeta_{[i]}(t, x, v)$ where the maps $\zeta_{[i]}(\cdot, \cdot, \cdot)$ are given by

$$\zeta_{[i]}(t, x, v) = \begin{cases} 1 & \text{on } [x_{[i-1]} - \delta_0, x_{[i]} + \delta_0] \times [\eta_0, V_0], \\ \text{linearly decreasing to 0} & \text{on } \omega_{[i]}^0 \setminus ([x_{[i-1]} - \delta_0, x_{[i]} + \delta_0] \times [\eta_0, V_0]), \\ 0 & \text{on } \mathbb{R}^{2d} \setminus \omega_{[i]}^0. \end{cases} \quad (5.40)$$

for all $i \in \{1, \dots, n\}$. A picture of this domain decomposition along with the definition of $\zeta_{[i]}(\cdot, \cdot, \cdot)$ for a given i is given in Figure 1.

We consider the time partition $[0, T_0] = \bigcup_{i=0}^{n-1} \left[\frac{iT_0}{n}, \frac{(i+1)T_0}{n} \right]$ and apply each control $u_{[i]}(\cdot, \cdot, \cdot)$ on the set $\omega_{[i]}^0$ for $t \in \left[\frac{iT_0}{n}, \frac{(i+1)T_0}{n} \right)$. This control design ensures that the following properties are satisfied.

- (1) The control is Lipschitz-in-space and satisfies (5.34) at all times by definition of the functions $(\zeta_{[i]}(\cdot, \cdot, \cdot))_{1 \leq i \leq n}$.
- (2) By Lemma 5.4, one can easily check that $\operatorname{supp}(\mu(t)) \subset [0, X_0 + tV_0] \times [0, V_0]$, yielding in turn

$$\operatorname{supp}(\mu(t)) \subset [0, X_0 + T_0V_0] \times [0, V_0] \quad (5.41)$$

for all times $t \in [0, T_0]$. Moreover, choosing $T_0 \leq \delta_0/V_0$, we have by construction that all the points in $\operatorname{supp}(\mu^0)$ will locally undergo a displacement of amplitude at most equal to δ_0 in the variable x .

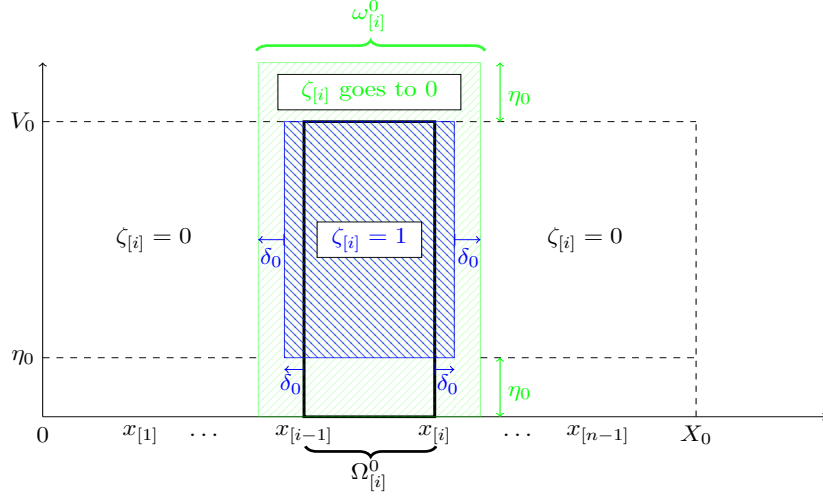


Figure 5.1: Construction of a control set $\omega_{[i]}^0$ and of the corresponding $\Omega_{[i]}^0$ and $\zeta_{[i]}(\cdot, \cdot, \cdot)$ for a given $i \in \{1, \dots, n\}$

(3) The population constraint (5.33) is respected at all times. Indeed, one has that

$$\int_{\omega_{[i]}^0} d\mu(t)(x, v) \leq \int_{x_{[i-1]}-3\delta_0}^{x_{[i]}+3\delta_0} \int_0^{V_0} d\mu^0(x, v) = c, \quad (5.42)$$

for any $i \in \{1, \dots, n\}$.

After having defined a proper control satisfying our constraints, we are interested in building estimates for the size of $\text{supp}(\mu(T_0))$. To do so, we monitor the evolution through time of the points (x, v) such that v realizes the maximum of velocity in $\Omega_{[i]}(t)$ (similar estimates were given in [116]). Here for $i \in \{1, \dots, n\}$ and $t \geq 0$, we define the sets $\Omega_{[i]}(t)$ by

$$\Omega_{[i]}(t) = \Phi_{(0,t)}[\mu^0](\Omega_{[i]}^0)$$

where we recall that $(\Phi_{(0,t)}[\mu^0](\cdot, \cdot))_{t \geq 0}$ is the family of non-local flows generated by $(x, v) \mapsto (v, \Psi[\mu] + \chi_\omega u)^T$.

We now introduce monitoring functions $b_i(\cdot)$ defined by

$$b_i(t) = \sup \left\{ v \in \mathbb{R} \text{ s.t. } (x, v) \in \Omega_{[i]}(t) \right\}$$

for all $t \in [0, T_0]$. Our goal is to derive upper estimates on these maps that will in turn be useful to prove that the velocity support of the measure shrinks. We proceed by a case disjunction.

◇ If $b_i(t) \geq \eta_0$ for all $t \in [0, T_0]$, then using (5.32) and the Lipschitzianity of $\psi(\cdot, \cdot)$ one has that

$$\begin{aligned}
\dot{b}_i(t) &= \Psi[\mu(t)](x_i(t), b_i(t)) + \mathbb{1}_{\omega_{[i]}^0}(x_i(t), b_i(t))u_{[i]}(t, x_i(t), b_i(t)) \\
&\leq L \int_{\mathbb{R}^2} (w - b_i(t))d\mu(t)(y, w) + \mathbb{1}_{\omega_{[i]}^0}(x_i(t), b_i(t))u_{[i]}(t, x_i(t), b_i(t)) \\
&\leq (V_0 - b_i(t)) + \chi_{\omega_{[i]}^0} u_{[i]}(t, x_i(t), b_i(t)),
\end{aligned}$$

where $L = \text{Lip}(\psi(\cdot, \cdot), \mathbb{R}^2)$. Applying Grönwall's Lemma to $b_i(\cdot) - V_0$ and noticing that $b_i(0) - V_0 \leq 0$, we recover that

$$b_i(T_0) \leq V_0 + e^{-LT_0} \int_0^{T_0} \mathbb{1}_{\omega_{[i]}^0}(x_i(t), b_i(t))u_{[i]}(t, x_i(t), b_i(t))dt. \quad (5.43)$$

By construction of the set $\omega_{[i]}^0$ and of the control $u_{[i]}(\cdot, \cdot, \cdot)$, the fact that $b_i(\cdot) \geq \eta_0$ on $[0, T_0]$ implies that $u_{[i]}(\cdot, x_i(\cdot), b_i(\cdot))$ is equal to (-1) on $[\frac{iT_0}{n}, \frac{(i+1)T_0}{n})$ and to 0 on $[0, \frac{(i-1)T_0}{n}) \cup [\frac{(i+1)T_0}{n}, T_0]$, leading to

$$b_i(T_0) \leq V_0 - \frac{e^{-LT_0}T_0}{n}. \quad (5.44)$$

◇ On the other hand if $b_i(t) < \eta_0$ for some $t \in [0, T_0]$, define \bar{t} to be the biggest time for which $b_i(\cdot) \leq \eta_0$. Notice then that $v - b_i(s) \leq V_0 - \eta_0$ for all $v \in [0, V_0]$ and $s \geq \bar{t}$. By a similar argument as in the previous point, one gets

$$b_i(T_0) \leq \eta_0(1 - LT_0) + LV_0T_0. \quad (5.45)$$

This holds in particular if $\bar{t} = T_0$, i.e. if $b_i(\cdot) \leq \eta_0$.

Since $(\Omega_{[i]}(t))_{1 \leq i \leq n}$ defines a covering of $\text{supp}(\mu(t))$ for all $t \geq 0$, these estimates together with (5.41) allow us to write that $\text{supp}(\mu(T_0)) \subset [0, X_1] \times [0, V_1]$ with

$$\begin{cases} V_1 = \max \left\{ V_0 - \frac{e^{-LT_0}T_0}{n}, \eta_0(1 - LT_0) + LT_0V_0 \right\}, \\ X_1 = X_0 + T_0V_0. \end{cases} \quad (5.46)$$

Step 3 : Iteration of the fundamental step and convergence of the algorithm

In this third Step, we show how a sequence of the fundamental steps defined in Step 2 steers the system to approximate alignment. To this end, we are going to iterate the fundamental step and show that the velocity support must decrease strictly at each step up to choosing the free parameters (η, T) in a proper way.

We start by fixing $\epsilon > 0$ and by denoting for any $k \in \mathbb{N}$ the measure $\mu^{k+1} = \mu \left(\sum_{l=0}^k T_l \right)$. The estimate (5.46) shows us that $\text{supp}(\mu^{k+1}) \subseteq [0, X_{k+1}] \times [0, V_{k+1}]$ where

$$\begin{cases} V_{k+1} = \max \left\{ V_k - \frac{e^{-LT_k}T_k}{n}, \eta_k(1 - LT_k) + LT_kV_k \right\}, \\ X_{k+1} = X_k + V_kT_k. \end{cases} \quad (5.47)$$

We build the corresponding partition $(\Omega_{[i]}^k)_{1 \leq i \leq n}$ of $[0, X_k] \times [0, V_k]$ according to (5.37) and define the corresponding parameters δ_k as in (5.38). We also define the sets $(\omega_{[i]}^k)_{1 \leq i \leq n}$, along with the corresponding controls $(u_{[i]}^k)_{1 \leq i \leq n}$ as in Step 2. We set define the auxiliary constant $\alpha = (1 + \frac{3}{nL\epsilon})$ where we recall that $L = \text{Lip}(\psi(\cdot, \cdot))$, and choose

$$T_k = \min \left\{ \frac{\delta_k}{V_k}, \frac{1}{\alpha L} \right\}, \quad \eta_k = \frac{1}{2} \left(V_k - \frac{e^{-LT_k} T_k}{n(1 - LT_k)} \right). \quad (5.48)$$

We now want to show that the sequence $(V_k)_{k \in \mathbb{N}}$ defined in (5.47) becomes smaller than ϵ **within a finite number of iterations** of our fundamental step when (η_k, T_k) are chosen as in (5.48). To do so, we prove the slightly stronger result that $(V_k)_{k \in \mathbb{N}}$ converges to a limit $V_* < \epsilon/2$. This implies that there exists $K \in \mathbb{N}$ such that $V_k < \epsilon$ for all $k \geq K$, which yields our claim.

We start by proving the following useful estimate.

Lemma 5.5. *For a given $k \in \mathbb{N}$, let $f^k \in L^\infty(\mathbb{R}^{2d}, \mathbb{R}^{2d})$ be the density of μ^k with respect to \mathcal{L}^{2d} . Then, it holds that*

$$12 \|f^k(\cdot, \cdot)\|_{L^\infty(\mathbb{R}^{2d})} \delta_k V_k \geq c, \quad (5.49)$$

where V_k is the size of the support of μ^k along the velocity component and δ_k is defined as in (5.38).

Proof. The proof follows from a simple geometric argument. For any $k \in \mathbb{N}$, one has that

$$\mu^k \left([x_{[i-1]}^k, x_{[i]}^k] \times [0, V_k] \right) \leq \frac{c}{2}$$

by construction of the partition $(\Omega_{[i]}^k)_{1 \leq i \leq n}$. It follows that for any real number $r > 0$, one has

$$\mu^k \left([x_{[i-1]}^k - r, x_{[i]}^k + r] \times [0, V_k] \right) \leq \frac{c}{2} + 2rV_k \|f(\cdot, \cdot)\|_{L^\infty(\mathbb{R}^{2d})}$$

Besides by definition of δ_k in (5.38), there exists an index $i_k \in \{1, \dots, n\}$ such that

$$\mu^k \left([x_{[i-1]}^k - 3\delta_k, x_{[i]}^k + 3\delta_k] \times [0, V_k] \right) = c.$$

Taking $r = 3\delta_k$ yields the desired estimate. \square

One can check that the following facts hold true along our iteration.

- (1) Choosing $T_k \leq \frac{1}{\alpha L}$ ensures that η_k is strictly positive as long as $V_k \geq \epsilon/2$. This comes from the following easy estimates

$$V_k - \frac{T_k e^{-LT_k}}{n(1 - LT_k)} \geq V_k - \frac{\alpha T_k e^{-LT_k}}{n(\alpha - 1)} \geq V_k - \frac{\epsilon}{3} e^{-\frac{1}{\alpha}},$$

in which we used the fact that $\alpha = (1 + \frac{3}{nL\epsilon})$.

- (2) Choosing T_k and η_k as in (5.48) ensures that

$$\eta_k = \frac{1}{2} \left(V_k - \frac{e^{-LT_k} T_k}{n(1 - LT_k)} \right) \leq V_k - \frac{e^{-LT_k} T_k}{n(1 - LT_k)}$$

which in turn yields

$$(1 - LT_k)\eta_k + LT_k V_k \leq V_k - \frac{e^{-LT_k T_k}}{n}.$$

By the definition of η_k given in (5.48), we obtain that $V_{k+1} = V_k - \frac{e^{-LT_k T_k}}{n}$ as long as $V_k \geq \epsilon/2$. This implies in particular that the sequence (V_k) strictly decreases as long as $V_k \geq \epsilon/2$.

- (3) Since $T_k \leq \frac{1}{\alpha L}$, then $V_k - V_{k+1} = \frac{e^{-LT_k T_k}}{n} \geq \frac{e^{-\frac{1}{\alpha}} T_k}{n}$ for any $k \in \mathbb{N}$ such that $V_k \geq \epsilon/2$. This implies that for any K such that $V_K \geq \epsilon/2$ one has

$$V_0 \geq \frac{e^{-\frac{1}{\alpha}}}{n} \sum_{k=0}^K T_k. \quad (5.50)$$

This last fact allows us to show by contradiction that our iteration procedure terminates within a finite number of steps. Assume that $V_k \geq \epsilon/2$ for all $k \in \mathbb{N}$. This necessarily implies that the sequence (V_k) is strictly decreasing by item (2), so that it converges to a limit $V_* \geq \frac{\epsilon}{2}$. Remark now that as a consequence of (5.50), one necessarily has that

$$T_k \xrightarrow[k \rightarrow +\infty]{} 0.$$

Taking the limit as $k \rightarrow +\infty$ in the definition (5.48) of (η_k) yields

$$\eta_k \xrightarrow[k \rightarrow +\infty]{} \frac{V_*}{2} > 0$$

Thus, we infer that there exists a constant $\bar{\eta} > 0$ such that $\eta_k \geq \bar{\eta}$ for any $k \in \mathbb{N}$.

By the very definition of the controls $u^k(\cdot, \cdot, \cdot)$ – which are chosen to be piecewise linear with slope at most equal to $\frac{1}{\eta_k}$ –, it further holds that

$$\|\nabla_v \cdot u_{[i]}^k(t, \cdot, \cdot)\|_{L^\infty(\omega_{[i]}^0)} \leq \frac{1}{\eta_k} \leq \frac{1}{\bar{\eta}}, \quad (5.51)$$

for any $i \in \{1, \dots, n\}$. Moreover, it also holds for any $d \in \mathbb{N}^*$ that

$$|\nabla_v \cdot \Psi[\mu(t)](x, v)| = \left| \int_{\mathbb{R}^2} \partial_v \psi(y - x, w - v) d\mu(t)(y, w) \right| \leq L, \quad (5.52)$$

for any $(x, v) \in \mathbb{R}^2$ by Lipschitzianity of $(x, v) \mapsto \psi(x, v)$. Merging (5.51)-(5.52), we recover the following estimate

$$\|\nabla_v \cdot (\mathbf{1}_{\omega_{[i]}^0}(\cdot, \cdot) u_{[i]}^k(t, \cdot, \cdot))\|_{L^\infty(\mathbb{R}^2)} + \|\nabla_v \cdot \Psi[\mu(t)](\cdot, \cdot)\|_{L^\infty(\mathbb{R}^2)} \leq \bar{F}_1, \quad (5.53)$$

for all times $t \geq 0$, where we introduced $\bar{F}_1 = (L + 1/\bar{\eta}) > 0$.

Recall that $\bar{T} = \sum_{k=0}^{\infty} T_k$ must be finite as a consequence of (5.50). Then, combining (5.53) with the estimate (5.36) of Proposition 5.3, one has

$$\|f^k(\cdot, \cdot)\|_{L^\infty(\mathbb{R}^2)} \leq \|f^0(\cdot, \cdot)\|_{L^\infty(\mathbb{R}^2)} e^{\bar{F}_1 \bar{T}},$$

for any $k \in \mathbb{N}$. This, together with (5.49) implies that

$$\delta_k \geq \frac{ce^{-\bar{F}_1 \bar{T}}}{12 \|f^0(\cdot, \cdot)\|_{L^\infty(\mathbb{R}^2)} V_0} = \bar{\delta}_1 > 0$$

again for every $k \in \mathbb{N}$. Plugging this into the definition (5.48) allows us to finally obtain

$$T_k = \min \left\{ \frac{\delta_k}{V_k}, \frac{1}{\alpha L} \right\} \geq \min \left\{ \frac{\bar{\delta}}{V_0}, \frac{1}{\alpha L} \right\} > 0 \quad (5.54)$$

for all $k \in \mathbb{N}$.

This strictly positive lower bound (5.54) on the sequence (T_k) implies that $\bar{T} = \sum_{k=0}^{+\infty} T_k$ diverges to infinity, which contradicts (5.50). Hence, one concludes that $V_* < \epsilon/2$. Since the sequence (V_k) is decreasing, there exists a step index $K \in \mathbb{N}$ such that $V_k < \epsilon$ for all $k \geq K$, which is equivalent to saying that our control strategy steers the system to ϵ -approximate alignment around 0 starting from time $\bar{T} = \sum_{k=0}^K T_k \leq e^{\frac{1}{\alpha}} \lceil \frac{2}{c} \rceil V_0$.

Step 4 : Proof of Theorem 5.3 in dimensions $d > 1$

In this last Step, we show how the control design of Step 2 and Step 3 in the one dimensional setting can be extended to the case in which $d > 1$. We therefore fix an integer $d \geq 1$, a precision $\epsilon > 0$, and suppose that

$$\text{supp}(\mu^0) \subset \prod_{i=1}^d \left([0, X_0^i] \times [0, V_0^i] \right)$$

The proof of Theorem 5.3 in dimensions $d \geq 1$ consists in applying iteratively a d -dimensional fundamental step which is analogous to the 1-dimensional fundamental step, but which acts on one component of the system at a time. Once approximate alignment is obtained in a given direction, the sequence of fundamental steps is applied to another one until the whole system is steered to approximate alignment. Remark here that the global invariance of the velocity support of μ^0 under the dynamics (5.31) given by Lemma 5.4 is essential to ensure that this component-wise sequential approach converges.

We define the fundamental step acting only on the j -th component as in Step 2 and Step 3, with the following adaptations. Let us fix a coordinate $j \in \{1, \dots, d\}$.

(i) The partition $(\Omega_{[i]}^{0,j})_{1 \leq i \leq n}$ associated to j -th component is defined via

$$\Omega_i^{0,j} = \prod_{l=1}^{j-1} [0, X^l] \times [x_{j,[i-1]}, x_{j,[i]}] \times \prod_{l=j+1}^d [0, X^l] \times \prod_{l=1}^d [0, V^l].$$

(ii) The parameters δ_0^j are defined as the biggest real numbers such that

$$\mu^0 \left(\mathbb{R}^{j-1} \times [x_{j,[i-1]} - 3\delta_0, x_{j,[i]} + 3\delta_0] \times \mathbb{R}^{d-j} \times \mathbb{R}^d \right) \leq c$$

$i \in \{1, \dots, n\}$.

(iii) The control sets $\omega_{[i]}^{0,j}$ are defined on $[x_{j,[i-1]} - 2\delta_0, x_{j,[i]} + 2\delta_0]$ on the j -th coordinate only.

- (iv) The maps $\zeta_{[i]}^j(t, \cdot, \cdot)$ depend only on (x_j, v_j) for all times $t \geq 0$ and the corresponding controls are defined by $u_{[i]}^j(\cdot, \cdot, v) = -\zeta_{[i]}^j(\cdot, \cdot, v) \frac{v_j}{|v_j|}$.
- (v) The density estimate (5.49) is rewritten

$$12\delta_k^j \|f^k(\cdot, \cdot)\|_{L^\infty(\mathbb{R}^{2d})} \prod_{\substack{i=1 \\ i \neq j}}^d X_k^i \prod_{\substack{i=1 \\ i \neq j}}^d V_k^i \geq c,$$

while (5.53) becomes

$$\|\nabla_v \cdot (\mathbb{1}_{\omega_{[i]}^0}(\cdot, \cdot) u_{[i]}^k(t, \cdot, \cdot))\|_{L^\infty(\mathbb{R}^2)} + \|\nabla_v \cdot \Psi[\mu(t)](\cdot, \cdot)\|_{L^\infty(\mathbb{R}^2)} \leq \bar{F}_d,$$

where $\bar{F}_d = (Ld + 1/\bar{\eta})$.

We iterate the d -dimensional fundamental step on each component of our system, choosing the values for the parameters T_k^j and η_k^j

$$T_k^j = \min \left\{ \frac{\delta_k^j}{V_k^j}, \frac{1}{\alpha L} \right\}, \quad \eta_k^j = \max \left\{ 0, \frac{1}{2} \left(V_k^j - \frac{e^{-LT_k^j} T_k^j}{n(1 - LT_k^j)} \right) \right\} \quad (5.55)$$

with $\alpha = 1 + \frac{3}{nL\epsilon}$ and again with the convention that $u^k(\cdot, \cdot, \cdot) = 0$ if $\eta_k = 0$. This ensures that for any $j \in \{1, \dots, d\}$, there exists a constant $K_j \in \mathbb{N}$ and a final time $\bar{T}^j = \sum_{k=0}^{K_j} T_k^j$ such that the system is steered to ϵ -approximate alignment around 0 starting from time \bar{T}^j on the j -th component. Moreover, (5.50) implies $\bar{T}^j \leq e^{\frac{1}{\alpha}} \lceil \frac{2}{\epsilon} \rceil V_0^j$.

By invariance of the velocity support of the crowd under the dynamics (5.32), we have by applying the d -dimensional fundamental step sequentially on each component that there exists a time $\bar{T} = \sum_{j=1}^d \bar{T}^j$ from which our system is steered to ϵ -approximate alignment around 0. Moreover we know by (5.50) that $\bar{T} \leq \lceil \frac{2}{\epsilon} \rceil e^{\frac{1}{\alpha}} \sum_{j=1}^d V_0^j$. As already described in Step 1, this result can then be generalized to produce an ϵ -alignment around any given $v^* \in \mathbb{R}^d$ from time \bar{T} using translations and the invariance properties of the system (5.31), which concludes the proof Theorem 5.3.

6 Generic Singularities of the 3D-Contact Sub-Riemannian Conjugate Locus

In this chapter, we present a results at the crossroad between *singularity theory* in the spirit of Arnold's school [20] and *sub-Riemannian geometry*. This chapter is somewhat apart from the rest of the manuscript both due to its mathematical goals and the tools it involves. Hence before moving on to the statement of Theorem 6.1, we provide below a short introduction to this problem and to the corresponding literature.

One of the major discrepancies between Riemannian and sub-Riemannian geometry is the fact that sub-Riemannian geodesics can stop being globally optimal in an arbitrary small vicinity of their starting point. As a consequence, the sub-Riemannian distance function is never smooth in a neighbourhood of its base point. This uncanny property translates into the presence of a wide amount of singularities along the so-called *caustic* surfaces generated by the critical points of the corresponding exponential map. The analysis of the least degenerate generic behaviour of the 3D-contact sub-Riemannian caustic was carried out for the first time in the seminal paper [3]. The further degenerate generic situations were then studied independently in [4, 52], and the full classification of the generic singularities of the *semi-caustics* – i.e. the intersections of the caustic with adequate half-spaces – was completed in [6]. In the present note, we accomplish this research by providing a complete classification of the generic singularities of the full caustic. Let it be noted that the question of *stability* for these singularities is very delicate, and was only studied subsequently in [7]. Furthermore, the analysis of the less degenerate configurations of the contact sub-Riemannian caustics has been recently extended to the case of dimension greater or equal to 5 in [124].

It was shown in [6, 52] that for the caustic, the generic degenerate situations appear when one of the fundamental geometric invariants of the sub-Riemannian distribution vanishes. Starting from the generic non-degenerate situation in which the semi-caustics exhibit 4 plea lines (see Figure 6.1-left), one can observe the formation of a *swallow tail* (see Figure 6.1-center) as the geometric invariant gets closer to zero. When this invariant vanishes, the swallow tail folds itself and becomes degenerate, and the semi-caustics become distinct closed surfaces exhibiting 6 plea lines (see Figure 6.1-right). The intersections of these surfaces with horizontal planes are closed curves which present 6 cusp points as well as *self-intersections* (see Figure 1-right and Figure 6.2), the arrangement of which fully characterizes the corresponding singularity of the semi-caustic. The classification of these singularities, i.e. of the distribution of the corresponding cuspidal points and self-intersections, can be synthetically represented by means of *symbols*. Here, a symbol is a six-tuple of rational numbers (s_1, \dots, s_6) where each s_i is half the number of self-intersections appearing along the piece of curve joining two consecutive cusp points. Using this notation, we can state the main result of this article.

Theorem 6.1 (Main result). *Let M be a 3-dimensional smooth and connected manifold and $SubR(M)$ be the space of contact sub-Riemannian distributions over M endowed with*

the Whitney topology. There exists an open and dense subset $\mathcal{E} \subset \text{SubR}(M)$ such that for any $(\Delta, \mathbf{g}) \in \mathcal{E}$ the following holds.

- (i) There exists a smooth curve $\mathcal{C} \subset M$ such that, outside \mathcal{C} , the intersections of the caustic with horizontal planes $\{h = \pm\epsilon\}$ are closed curves exhibiting 4 cusp points (see Figure 6.1-left).
- (ii) There exists an open and dense subset $\mathcal{O} \subset \mathcal{C}$ on which the intersections of the caustic with horizontal planes $\{h = \pm\epsilon\}$ are described by pairs of symbols $(\mathcal{S}_i, \mathcal{S}_j)$ with $i, j \in \{1, 2, 3\}$ (see Figure 6.2-A and Figure 6.2-B), where

$$\mathcal{S}_1 = (0, 1, 1, 1, 1, 1), \quad \mathcal{S}_2 = (2, 1, 1, 1, 1, 1), \quad \mathcal{S}_3 = (2, 1, 1, 2, 1, 0).$$

- (iii) There exists a discrete subset $\mathcal{D} \subset \mathcal{C}$ complement of \mathcal{O} in \mathcal{C} on which the intersections of the caustic with horizontal planes $\{h = \pm\epsilon\}$ are described by pairs of symbols $(\mathcal{S}_i, \mathcal{S}_j)$ with $i \in \{1, 2, 3\}, j \in \{4, 5, 6, 7\}$ (see Figure 6.2-C and Figure 6.2-D), where

$$\begin{aligned} \mathcal{S}_4 &= \left(\frac{1}{2}, \frac{1}{2}, 1, 0, 0, 1\right), & \mathcal{S}_5 &= \left(1, \frac{1}{2}, \frac{1}{2}, 1, 1, 1\right), \\ \mathcal{S}_6 &= \left(\frac{3}{2}, \frac{1}{2}, 1, 1, 0, 1\right), & \mathcal{S}_7 &= \left(2, \frac{1}{2}, \frac{1}{2}, 2, 0, 0\right). \end{aligned}$$

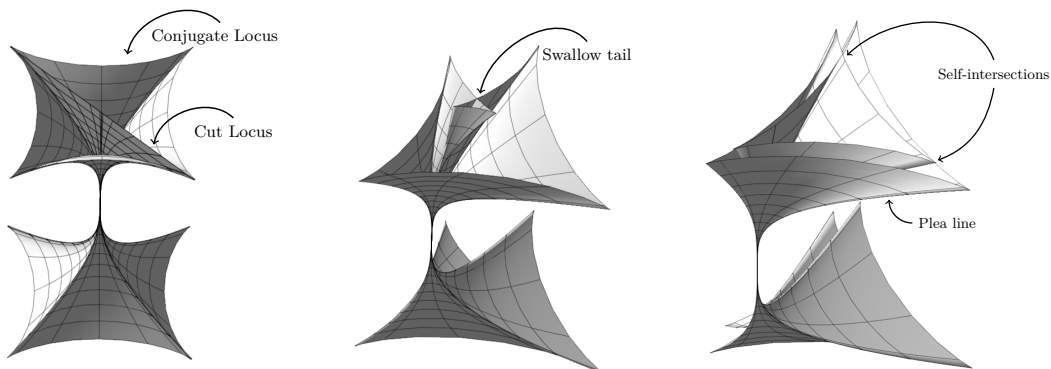


Figure 6.1: Generic 4-cusp conjugate and cut loci outside \mathcal{C} (left), formation of the swallow tail near \mathcal{C} (center), folding of the swallow tail and formation of the 6-cusp singularity along \mathcal{C} (right)

This result is in a sense the most natural one to be expected after the classification of the half conjugate loci displayed in [6]. Indeed, it transcribes the fact that the upper and lower semi-caustics are independent and that there are no extra couplings appearing between the two structures. Indeed, the only possible obstruction to the combination of two given symbols is that the corresponding codimension in the space of Taylor coefficients – which is preserved by standard arguments of transversality theory – is strictly larger than 3. In particular, this generically prevents pairs of the form $(\mathcal{S}_i, \mathcal{S}_j)$ with $i, j \in \{4, 5, 6, 7\}$ from appearing.

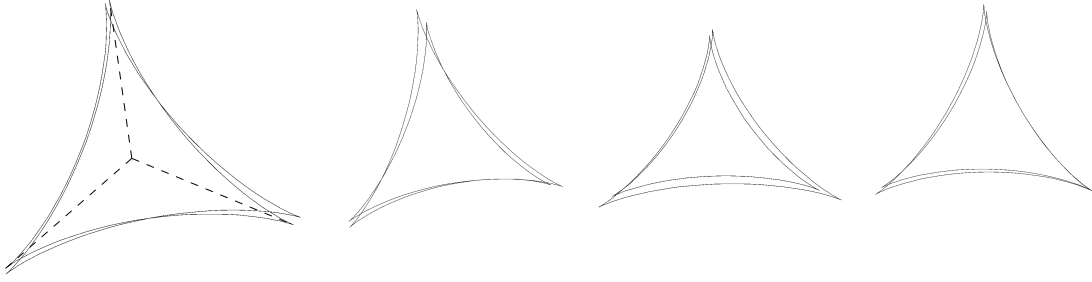


Figure 6.2: Intersections of the upper semi-caustics corresponding to the symbols \mathcal{S}_1 (A), \mathcal{S}_2 (B), \mathcal{S}_4 (C) and \mathcal{S}_6 (D) with the planes $h = \epsilon$, and intersection of the corresponding cut locus (dashed lines on the left)

6.1 3D-Contact sub-Riemannian manifolds and their conjugate locus

In this section, we recall some elementary facts about sub-Riemannian geometry defined over 3-dimensional manifolds. For a complete introduction, see e.g. [5].

Definition 6.1 (Sub-Riemannian manifold). *A 3D-contact sub-Riemannian manifold is defined by a triple (M, Δ, \mathbf{g}) where*

- M is an 3-dimensional smooth and connected differentiable manifold,
- Δ is a smooth 2-dimensional distribution over M with step 1, i.e.

$$\text{Span} \{X_1(q), X_2(q), [X_1(q), X_2(q)]\} = T_q M,$$

for all $q \in M$ and $(X_1(q), X_2(q))$ spanning $\Delta(q)$.

- \mathbf{g} is a Riemannian metric over Δ .

Definition 6.2 (Horizontal curves and sub-Riemannian metric). *An absolutely continuous curve $\gamma(\cdot)$ is said to be horizontal if $\dot{\gamma}(t) \in \Delta(\gamma(t))$ for \mathcal{L}^1 -almost every $t \in [0, T]$. We define the length $l(\gamma(\cdot))$ of a horizontal curve $\gamma(\cdot)$ as*

$$l(\gamma(\cdot)) = \int_0^T \sqrt{\mathbf{g}_{\gamma(t)}(\dot{\gamma}(t), \dot{\gamma}(t))} dt.$$

For $(q_0, q_1) \in M$, it is then possible to define the sub-Riemannian distance $d_{\text{SR}}(q_0, q_1)$ as the infimum of the length of the horizontal curves connecting q_0 and q_1 .

Given a local orthonormal frames (X_1, X_2) for the metric \mathbf{g} which spans Δ , the Carnot-Carathéodory distance $d_{\text{SR}}(q_0, q_1)$ can be alternatively computed by solving the optimal control problem

$$\begin{cases} \min_{(u_1(\cdot), u_2(\cdot))} \int_0^T (u_1^2(t) + u_2^2(t)) dt \\ \text{s.t. } \dot{\gamma}(t) = u_1(t)X_1(\gamma(t)) + u_2(t)X_2(\gamma(t)), \quad u_1^2(t) + u_2^2(t) \leq 1, \\ \text{and } (\gamma(0), \gamma(T)) = (q_0, q_1). \end{cases} \quad (6.1)$$

We detail in the following proposition the explicit form of 3D-contact sub-Riemannian geodesics obtained by applying the maximum principle to (6.1).

Proposition 6.1 (The Pontryagin Maximum Principle in the 3D contact case).

Let $\gamma(\cdot) \in \text{Lip}([0, T], M)$ be a horizontal curve and $\mathcal{H} : T^*M \rightarrow \mathbb{R}$ be the Hamiltonian associated to the contact geodesic problem, defined by

$$\mathcal{H}(q, \lambda) = \frac{1}{2} \langle \lambda, X_1(q) + X_2(q) \rangle,$$

for any $(q, \lambda) \in T^*M$. Then, the curve $\gamma(\cdot)$ is a contact geodesic parametrized by sub-Riemannian arclength if and only if there exists a Lipschitzian curve $t \in [0, T] \mapsto \lambda(t) \in T_{\gamma(t)}^*M$ such that $t \mapsto (\gamma(t), \lambda(t))$ is a solution of the Hamiltonian system

$$\dot{\gamma}(t) = \partial_\lambda \mathcal{H}(\gamma(t), \lambda(t)), \quad \dot{\lambda}(t) = -\partial_q \mathcal{H}(\gamma(t), \lambda(t)), \quad \mathcal{H}(\gamma(t), \lambda(t)) = \frac{1}{2}. \quad (6.2)$$

We denote by $\vec{\mathcal{H}} \in \text{Vec}(T^*M)$ the corresponding Hamiltonian vector field defined over the cotangent bundle and by $(\gamma(t), \lambda(t)) = e^{t\vec{\mathcal{H}}}(q_0, \lambda_0)$ the corresponding solution of (6.2).

Both the absence of abnormal lifts and the sufficiency of the maximum principle are consequences of the contact hypothesis made on the sub-Riemannian structure (see e.g. [5, Chapter 4]).

Definition 6.3 (Exponential map). Let $q_0 \in M$ and $\Lambda_{q_0} = \{\lambda \in T_{q_0}^*M \text{ s.t. } \mathcal{H}(q_0, \lambda_0) = \frac{1}{2}\}$. We define the exponential map from q_0 as

$$E_{q_0} : (t, \lambda_0) \in \mathbb{R}_+ \times \Lambda_{q_0} \mapsto \pi_M \left(e^{t\vec{\mathcal{H}}}(q_0, \lambda_0) \right),$$

where $\pi_M : T^*M \rightarrow M$ is the canonical projection.

In this paper, we study the germ at the origin – i.e. equivalence classes of maps defined by equality of derivatives up to a certain order – of the *cut* and *conjugate loci* associated to contact sub-Riemannian structures.

Definition 6.4 (Cut and conjugate locus). Let (M, Δ, \mathbf{g}) be a 3D contact sub-Riemannian manifold, $(q_0, \lambda_0) \in T^*M$ and $\gamma(\cdot) = E_{q_0}(\cdot, \lambda_0)$ be a geodesic parametrized by sub-Riemannian arclength. The cut time associated to $\gamma(\cdot)$ is defined by

$$\tau_{\text{cut}} = \sup\{t \in \mathbb{R}_+ \text{ s.t. } \gamma_{[0,t]}(\cdot) \text{ is optimal}\},$$

and the corresponding cut locus is

$$\text{Cut}(q_0) = \{\gamma(\tau_{\text{cut}}) \text{ s.t. } \gamma(\cdot) \text{ is a sub-Riemannian geodesic from } q_0\}.$$

The first conjugate time τ_{conj} associated to the curve $\gamma(\cdot)$ is defined by

$$\tau_{\text{conj}} = \inf\{t \in \mathbb{R}_+ \text{ s.t. } (t, \lambda_0) \text{ is a critical point of } E_{q_0}(\cdot, \cdot)\},$$

and the corresponding conjugate locus is then defined by

$$\text{Conj}(q_0) = \{\gamma(\tau_{\text{conj}}) \text{ s.t. } \gamma(\cdot) \text{ is a sub-Riemannian geodesic from } q_0\}.$$

We recall in Theorem 6.2 below the *normal form* introduced formally in [52] and then derived geometrically in [6] for 3D-contact sub-Riemannian structures. Up to a simple change of coordinates, we can assume that $0 \in M$ and study the germ of the conjugate locus in a neighbourhood of the origin.

Theorem 6.2 (Normal form for 3D-contact sub-Riemannian distributions). *Let (M, Δ, \mathbf{g}) be a 3D-contact sub-Riemannian structure and (X_1, X_2) be a local orthonormal frame for Δ in a neighbourhood of the origin. Then, there exists a smooth system of so-called normal coordinates (x, y, w) on M along with two maps $\beta, \gamma \in C^\infty(M, \mathbb{R})$ such that (X_1, X_2) can be written in normal form as*

$$\begin{cases} X_1(x, y, w) = (1 + y^2\beta(x, y, w))\partial_x - xy\beta(x, y, w)\partial_y + \frac{y}{2}(1 + \gamma(x, y, w))\partial_w, \\ X_2(x, y, w) = (1 + x^2\beta(x, y, w))\partial_y - xy\beta(x, y, w)\partial_x - \frac{x}{2}(1 + \gamma(x, y, w))\partial_w, \\ \beta(0, 0, w) = \gamma(0, 0, w) = \partial_x\gamma(0, 0, w) = \partial_y\gamma(0, 0, w) = 0. \end{cases} \quad (6.3)$$

This system of coordinates is unique up to an action of $SO(2)$ and adapted to the contact structure, i.e. it induces a gradation with respective weights $(1, 1, 2)$ on the space of formal power series in (x, y, w) .

The truncation at the order $(-1, -1)$ of this normal form is precisely the orthonormal frame associated with the usual left-invariant metric on the Heisenberg group, i.e. $(X_1, X_2) = (\partial_x + \frac{y}{2}\partial_w, \partial_y - \frac{x}{2}\partial_w)$.

Given a Heisenberg geodesic with initial covector $\lambda_0 = (p(0), q(0), r(0)) \in T_0^*M$, the corresponding conjugate time is exactly $\tau_{\text{conj}} = 2\pi/r(0)$. In the general contact case (see e.g. [5, Chapter 16]), the first conjugate time is of the form

$$\tau_{\text{conj}} = 2\pi/r(0) + O(1/r(0)^3)$$

where the higher order terms can be expressed via the coefficients of the Taylor expansions

$$\begin{cases} \beta(x, y, w) = \sum_{l=1}^k \beta^l(x, y, w) + O^{k+1}(x, y, w), \\ \gamma(x, y, w) = \sum_{l=2}^k \gamma^l(x, y, w) + O^{k+1}(x, y, w) \end{cases}$$

of the maps β and γ with respect to the gradation $(1, 1, 2)$. We introduce in the following equations the coefficients (c_i, c_{jk}, c_{lmn}) of the polynomial functions γ^2, γ^3 and γ^4 appearing in these expansions.

$$\begin{cases} \gamma^2(x, y) = (c_0 + c_2)(x^2 + y^2) + (c_0 - c_2)(x^2 - y^2) - 2c_1xy, \\ \gamma^3(x, y) = (c_{11}x + c_{12}y)(x^2 + y^2) + 3(c_{31}x - c_{32}y)(x^2 - y^2) - 2(c_{31}x^3 + c_{32}y^3) \\ \gamma^4(x, y, w) = \frac{w}{2} \left((c_{421} + c_{422})(x^2 + y^2) + (c_{421} - c_{422})(x^2 - y^2) - 2c_{423}xy \right) \\ \quad + c_{441}(x^2 + y^2)^2 + c_{442}(x^4 + y^4 - 6x^2y^2) + 4c_{443}xy(x^2 - y^2) \\ \quad + c_{444}(x^4 - y^4) - 2c_{445}xy(x^2 + y^2) \end{cases} \quad (6.4)$$

These coefficients derive from the irreducible decompositions of γ^2 , γ^3 and γ^4 under the action of $SO(2)$, and are precisely the fundamental invariants discriminating the singularities listed in Theorem 6.1. Indeed, it was proven in [52] that the caustic becomes degenerate along a smooth curve $\mathcal{C} \subset M$ on which the coefficients (c_0, c_1, c_2) in the decomposition of γ^2 in (6.4) satisfy $c_0 = c_2$ and $c_1 = 0$. Moreover, the decompositions of β, γ^2 and γ^3 have an interpretation in terms of canonical sub-Riemannian *connection* and *curvature*, as detailed in [52].

Following the methodology developed in [6], we introduce the coordinates (h, φ) defined by

$$(p, q) = (\cos(\varphi), \sin(\varphi)), \quad h = \text{sign}(w)\sqrt{|w|/\pi}.$$

We can then express the Taylor expansion of order $k \geq 3$ of the semi conjugate loci as the suspension

$$\text{Conj}_{\pm}(\varphi, h) = (x(\varphi, h), y(\varphi, h), h) = \left(\sum_{l=3}^k h^l f_l^{\pm}(\varphi), h \right), \quad (6.5)$$

where the \pm symbol highlights the dependency of the expression on the sign of w . By carrying out explicitly the computations necessary to put the conjugate locus in the suspended form (6.5), it can be verified that $f_3^+ = f_3^-$ and $f_4^+ = f_4^-$. Hence, the generic behaviour of the full conjugate locus is already known outside the smooth curve $\mathcal{C} \subset M$ on which the caustic becomes degenerate. We therefore restrict our attention to the more degenerate singularities arising in the form of *self-intersections* along the curve \mathcal{C} .

To prove that the symbols listed in Theorem 6.1 are generic, one needs to compute the Taylor expansions of order $k = 7$ of the metric in a neighbourhood of the origin. All these expressions were obtained using Maple software using a piece of code that can be found at the following address:

http://www.lsis.org/bonnetb/depots/Sub-Riemannian_Conjugate_Locus

6.2 Generic singularities of the full-conjugate locus

In this section, we prove Theorem 6.1. We show that the coefficients defined in (6.4), characterizing the generic conjugate locus on the curve \mathcal{C} , generate independent structures for Conj_+ and Conj_- .

Definition 6.5 (Self-intersection set of the semi-conjugate loci). *We define the self-intersection set of Conj_+ as*

$$\text{Self}(\text{Conj}_+) = \{(h, \varphi_1, \varphi_2) \in \mathbb{R}_+^* \times \mathbb{S}^1 \times \mathbb{S}^1 \text{ s.t. } \varphi_1 \neq \varphi_2, \text{Conj}_+(h, \varphi_1) = \text{Conj}_+(h, \varphi_2)\}.$$

An angle $\varphi \in \mathbb{S}^1$ is said to be adherent to $\text{Self}(\text{Conj}_+)$ provided that $(0, \varphi, \varphi + \pi) \in \text{Self}(\text{Conj}_+)$. The set of such angles is denoted by $\text{A-Self}(\text{Conj}_+)$. We define in the same way the self-intersection and adherent angles sets of Conj_- .

We recall in Theorem 6.3 below the complete classification of the generic self-intersections of Conj_+ (or equivalently of Conj_-) which was derived in [6].

Theorem 6.3 (Generic self-intersections of the positive semi conjugate locus). *Let $\mathcal{C} \subset M$ be the curve defined in Theorem 6.1. Then, \mathcal{C} is a generically smooth curve, and outside \mathcal{C} the semi conjugate loci are the standard 4-cusp semi-caustics. On \mathcal{C} , the following situations can occur.*

- (i) *There exists an open and dense subset $\mathcal{O}^+ \subset \mathcal{C}$ on which the self-intersections of Conj_+ are described by the symbols $(\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3)$.*
- (ii) *There exists a discrete subset \mathcal{D}^+ complement of \mathcal{O}^+ in \mathcal{C} on which the self-intersections of Conj_+ are described by the symbols $(\mathcal{S}_4, \mathcal{S}_5, \mathcal{S}_6, \mathcal{S}_7)$.*

We recall in the following lemma the structural result allowing to describe the sets $\text{A-Self}(\text{Conj}_\pm)$.

Lemma 6.1 (Structure of the self-intersection). *The set of adherent angle to $\text{Self}(\text{Conj}_\pm)$ satisfies the inclusion $\text{A-Self}(\text{Conj}_\pm) \subset \{\varphi \in \mathbb{S}^1 \text{ s.t. } f'_4(\varphi) \wedge f_5^\pm(\varphi) = 0\}$, where $a \wedge b \equiv \det(a, b)$ for vectors $a, b \in \mathbb{R}^2$.*

An explicit computation based on the expressions of the maps f_4, f_5^- and f_5^+ defined in (6.5) shows that

$$f'_4(\varphi) \wedge f_5^\pm(\varphi) = -20\pi^2 \tilde{b} \sin(3\varphi + \omega_b) P_\pm(\varphi)$$

where we introduced the notations $b = (c_{31}, c_{32}) = \tilde{b}(\sin(\omega_b), -\cos(\omega_b)) \in \mathbb{C}$ and

$$P_\pm(\varphi) = A_\pm \cos(2\varphi) + B_\pm \sin(2\varphi) + C \cos(4\varphi) + D \sin(4\varphi).$$

Here, the coefficients (A_\pm, B_\pm, C, D) are **independent linear combinations** of the fourth-order coefficients $(c_{421}, c_{422}, c_{423}, c_{442}, c_{443}, c_{444}, c_{445})$ introduced in (6.4). Their analytical expressions were derived again by using the Maple software and write as follows:

$$\begin{cases} A_\pm = \frac{35}{8}(c_{422} - c_{421}) \pm 3\pi c_{423} + 45c_{445}, & C = 36c_{443}, \\ B_\pm = \frac{35}{8}c_{423} \pm 3\pi(c_{421} - c_{422}) + 45c_{444}, & D = -36c_{442}. \end{cases} \quad (6.6)$$

In order to describe the roots of $f'_4 \wedge f_5^\pm$ in \mathbb{S}^1 , it is convenient to introduce the complex polynomials

$$\tilde{P}_\pm(z) = \mu z^4 + \nu_\pm z^3 + \bar{\nu}_\pm z + \bar{\mu}, \quad \tilde{T}(z) = bz^3 + \bar{b},$$

where the complex coefficients ν_-, ν_+ and μ are defined by

$$\nu_\pm = \frac{1}{2}(A_\pm - iB_\pm), \quad \mu = \frac{1}{2}(C - iD).$$

It can be checked that $f'_4(\varphi) \wedge f_5^\pm(\varphi) = 0$ if and only if $\tilde{P}_\pm(e^{i\varphi})\tilde{T}(e^{i\varphi}) = 0$. Since (A_\pm, B_\pm, C, D) independently span \mathbb{R}^6 as a consequence of (6.6), the classification of the generic singularities of the degenerate semi-caustics reduces to understanding the distribution of the unit roots of \tilde{P}_\pm and \tilde{T} as (μ, ν_\pm, b) span \mathbb{C}^4 . In the following lemma, we compile some of the facts highlighting the relationship between these roots and the symbols describing $\text{Self}(\text{Conj}_\pm)$ that can be found in [6, 52].

Lemma 6.2 (Algebraic equations describing the singularities). *Let γ^2, γ^3 and γ^4 be given as in (6.4) along with their decomposition under the action of $SO(2)$. Then, the following hold.*

- (i) The polynomials \tilde{P}_\pm have either 2 or 4 distinct simple roots on the unit circle.
- (ii) The polynomials \tilde{P}_\pm share a unit root with \tilde{T} if and only if $\text{Res}(\mu, \nu_\pm, b) = 0$ where $\text{Res}(\mu, \nu_\pm, b)$ is the resultant polynomial of \tilde{P}_\pm and \tilde{T} .

The symbols \mathcal{S}_i introduced in Theorem 6.1 can be understood as follows: $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3$ describe all the possible situations in which \tilde{P}_\pm have 2 or 4 simple roots, none of which is shared with \tilde{T} , while $\mathcal{S}_4, \mathcal{S}_5, \mathcal{S}_6, \mathcal{S}_7$ refer to all the combinations in which \tilde{P}_\pm and \tilde{T} share a single simple root. All the other possible situations are non-generic even for the semi conjugate loci.

For the full conjugate locus, the situation is the following. It can be shown that the sets $S_\pm = \{(c_{jk}, c_{lmn}) \in \mathbb{R}^{16} \text{ s.t. } \mu, b \neq 0, \text{Res}(\mu, \nu_\pm, b) = 0\}$ are smooth manifolds of codimension 1. Moreover, it holds that $S_- \cap S_+$ has codimension 2 as a consequence of the fact that $\text{Res}(\mu, \nu_\pm, b)$ is a homogeneous polynomial of degree 7 in (μ, ν_\pm, b) which only vanishes at isolated points where its differential is surjective when both $b, \mu \neq 0$.

Since M is a 3-dimensional manifold and $\mathcal{C} \subset M$ already has codimension 2, it can be proven that the set of sub-Riemannian distributions with canonical orthonormal frame (X_1, X_2) defined in (6.4) such that $(c_{jk}, c_{lmn}) \notin S_- \cap S_+$ is open and dense. This stems from standard transversality and preservation of codimension arguments, following [6, Section 3.3]. Hence, it cannot happen generically that the singularities of the conjugate locus are described by a pair of symbols $(\mathcal{S}_i, \mathcal{S}_j)$ with both $i, j \in \{4, 5, 6, 7\}$.

This independence result implies in particular the following structural corollary for the cut locus.

Corollary 6.1 (Generic behaviour of the 3D-contact sub-Riemannian cut locus). *There exists an open and dense subset $\mathcal{E} \subset \text{SubR}(M)$ for the Whitney topology such that for any $(\Delta, \mathbf{g}) \in \mathcal{E}$, the following holds.*

- (i) Outside the smooth curve $\mathcal{C} \subset M$ defined in Theorem 6.1, the semi cut loci are independent : each of the two is a portion of plane joining two opposite plea lines along the corresponding semi caustic (see Figure 6.1-left).
- (ii) On the curve \mathcal{C} , the semi cut loci are the union of three portions of planes connecting the h -axis with alternate plea lines along the corresponding semi caustic (see Figure 6.2-A dotted lines). There is no interdependence whatsoever between these planes for the positive and negative semi cut loci.

The generic behaviour of the 3D-contact semi cut loci outside the degenerate curve \mathcal{C} is already known (see e.g. [5, Chapter 16]). As described in the introduction, swallow tails appear along the semi caustics as one approaches the curve \mathcal{C} (see Figure 6.1-center), the four plea lines then degenerating into the 6-cusp structure. The corresponding cusps are regrouped by pairs of the form $(k, k + 1)$, appearing respectively in a small vicinity of the cuspidal angles $\frac{k}{3}(\pi - \omega_b)$ with $k \in \{0, 1, 2\}$.

Generically, the semi cut loci are the unions of three portions of planes joining three of the six plea lines to the h -axis (see Figure 2-left dotted lines). Moreover, in the absence of an interior cusp, these supporting plea lines can be chosen freely for each semi cut loci to be either one of the two numberings $(1, 3, 5)$ or $(2, 4, 6)$ of the six plea lines. The result of this paper stating that the cusps of the conjugate locus are independently distributed for its upper and lower parts therefore yields Corollary 6.1.

Bibliography

- [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. Methods of Control Theory in Nonholonomic Geometry. In *Proceedings of the ICM*. Zürich, Birkhäuser, 1995.
- [4] A. Agrachev. Exponential Mappings for Contact Sub-Riemannian Structures. *Journal of Dynamical Control Systems*, 2(3):321–358, 1996.
- [5] A. Agrachev, D. Barilari, and U. Boscain. *A Comprehensive Introduction to Sub-Riemannian Geometry*. Cambridge Studies in Advanced Mathematics, 2019.
- [6] A. Agrachev, H. Chakir, and J.P. Gauthier. Subriemannian metrics on \mathbb{R}^3 . *Geometric control and nonholonomic Mechanics, Proceedings of Canadian Mathematical Society*, 25:29–76, 1998.
- [7] A. Agrachev, G. Charlot, J.P. Gauthier, and Zakalyukin V. On Subriemannian Caustics and Wave Fronts for Contact Distributions in the Three Space. *Journal of Dynamical and control Systems*, 6(3):365–395, 2000.
- [8] A. Agrachev and Y. Sachkov. *Control Theory from the Geometric Viewpoint*, volume 87 of *Encyclopaedia of Mathematical Sciences*. Springer, 2004.
- [9] G. Albi, Balagué.D, J.A. Carrillo, and J. von Brecht. Stability Analysis of Flock and Mill Rings for Second Order Models in Swarming. *SIAM J. App. Math.*, 74(3):794–818, 2014.
- [10] 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, 2016.
- [11] L. Ambrosio. Transport Equation and Cauchy Problem for BV Vector Fields. *Inventiones Mathematicae*, 158(2):227–260, 2004.
- [12] L. Ambrosio and G. Crippa. Continuity Equations and ODE Flows with Non-Smooth Velocities. *Proceedings of the Royal Society of Edinburgh*, 144(6):1191–1244, 2014.
- [13] L. Ambrosio, G. Crippa, C. De Lellis, F. Otto, and M. Westdickenberg. *Transport Equations and Multi-D Hyperbolic Conservation Laws*, volume 5 of *Lecture Notes of the Unione Matematica Italiana*. Springer, 2008.
- [14] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variations and Free Discontinuity Problems*. Oxford Mathematical Monographs, 2000.
- [15] L. Ambrosio and W. Gangbo. Hamiltonian ODEs in the Wasserstein space of probability measures. *Comm. Pure Appl. Math.*, 61(1):18–53, 2008.

- [16] L. Ambrosio and N. Gigli. Construction of the Parallel Transport in the Wasserstein Space. *Methods and Applications of Analysis*, 15(1):1–30, 2008.
- [17] 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.
- [18] L. Ambrosio, N. Gigli, and G. Savaré. Calculus and Heat Flow in Metric Measure Spaces and Applications to Spaces with Ricci Bounds from Below. *Inventiones Mathematicae*, 195(2):289–391, 2014.
- [19] M. Arjovsky, S. Chintala, and L. Bottou. Wasserstein Generative Adversarial Networks. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70, pages 214–223, 2017.
- [20] V.I. Arnold. *Singularity Theory*, volume 53 of *Lectures Notes*. London Mathematical Society, 1981.
- [21] 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.
- [22] 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.
- [23] M. Ballerini, N. Cabibbo, R. Candelier, et al. Interaction Ruling Animal Collective Behavior Depends on Topological Rather than Metric Distance: Evidence from a Field Study. *Proceedings of the national academy of sciences*, 105(4):1232–1237, 2008.
- [24] R.W. Beard and W. Ren. *Distributed Consensus in Multi-Vehicle Cooperative Control*. Springer-Verlag, 2008.
- [25] N. Bellomo, P. Degond, E. Tadmor, et al. *Active Particles, Volume 1: Advances in Theory, Models, and Applications*. Springer, 2017.
- [26] 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.
- [27] J.D. Benamou and Y. Brenier. A Computational Fluid Mechanics Solution to the Monge-Kantorovich Mass Transfer Problem. *Numerische Mathematik*, 84(3):375–393, 2000.
- [28] S. Berman, A. Halasz, M.A. Hsieh, and V. Kumar. Optimized Stochastic Policies for Task Allocation in Swarms of Robots. *IEEE Trans. Rob.*, 25(4):927–937, 2009.
- [29] A.L. Bertozzi and C.M. Topaz. Swarming Patterns in a Two-Dimensional Kinematic Model for Biological Groups. *SIAM J. App. Math.*, 65(1):152–174, 2004.
- [30] R. Bonalli, B. Hérisse, and E. Trélat. Optimal Control of Endo-Atmospheric Launch Vehicle Systems: Geometric and Computational Issues. 2019. arXiv:1710.11501.
- [31] M. Bongini, M. Fornasier, F. Rossi, and F. Solombrino. Mean Field Pontryagin Maximum Principle. *Journal of Optimization Theory and Applications*, 175:1–38, 2017.
- [32] B Bonnard, L. Faubourg, and E. Trélat. Optimal Control of the Atmospheric Arc of a Space Shuttle and Numerical Simulation with Multiple Shooting Methods. *Mathematical Models and Methods in Applied Sciences*, 15(1):109–140, 2005.

- [33] B. Bonnard, L. Faubourg, and E. Trélat. *Mécanique céleste et contrôle de systèmes spatiaux*, volume 51 of *Mathématiques et Applications*. Springer, 2006.
- [34] B Bonnet. A Pontryagin Maximum Principle in Wasserstein Spaces for Constrained Optimal Control Problems. *submitted*. <https://arxiv.org/abs/1810.13117>.
- [35] B. Bonnet and F. Rossi. The Pontryagin Maximum Principle in the Wasserstein Space. *Calculus of Variations and Partial Differential Equations*, 58:11, 2019.
- [36] U. Boscaïn and B. Piccoli. *Optimal Synthesis for Control Systems on 2D Manifolds*. Mathématiques et Applications. Springer, 2004.
- [37] 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.
- [38] H. Brézis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Universitext. Springer, 2010.
- [39] M. Burger, R. Pinnau, O. Totzeck, and O. Tse. Mean-Field Optimal Control and Optimality Conditions in the Space of Probability Measures. arXiv:1902.05339.
- [40] L. Caffarelli. Some Regularity Properties of Solutions of Monge Ampère Equation. *Communications in Pure and Applied Mathematics*, 44(8-9):965–969, 1991.
- [41] M. Caponigro, M. Fornasier, B. Piccoli, and E. Trélat. Sparse Stabilization and Optimal Control of the Cucker-Smale Model. *Mathematical Control and Related Fields*, 3(4):447–466, 2013.
- [42] M. Caponigro, M. Fornasier, B. Piccoli, and E. Trélat. Sparse Stabilization and Control of Alignment Models. *Math. Mod. Meth. Appl. Sci.*, 25 (3):521–564, 2015.
- [43] M. Caponigro, B. Piccoli, F. Rossi, and E. Trélat. Mean-Field Sparse Jurdjevic-Quinn Control. *Math. Mod. Meth. Appl. Sci.*, 27(7):1223–1253, 2017.
- [44] M. Caponigro, B. Piccoli, F. Rossi, and E. Trélat. Sparse Jurdjevic-Quinn Stabilization of Dissipative Systems. *Automatica*, 86:110–120, 2017.
- [45] P. Cardaliaguet. Notes on Mean-Field Games. 2012. <https://www.ceremade.dauphine.fr/~cardaliaguet/MFG20130420.pdf>.
- [46] P Cardaliaguet, F. Delarue, J-M. Lasry, and P-L. Lions. The Master Equation and the Convergence Problem in Mean Field Games. <https://arxiv.org/abs/1509.02505>.
- [47] P Cardaliaguet, A. Poretta, and D. Tonon. Sobolev Regularity for the First Order Hamilton–Jacobi Equation. *Calculus of Variations and Partial Differential Equations*, 54:3037–3065, 2015.
- [48] P. Cardaliaguet and L. Silvester. Hölder Continuity to Hamilton-Jacobi Equations with Super-Quadratic Growth in the Gradient and Unbounded Right-Hand Side. *Communications in Partial Differential Equations*, 37(9):1668–1688, 2012.
- [49] J.A. Carrillo, M. Fornasier, J. Rosado, and G. Toscani. Asymptotic Flocking for the Kinetic Cucker-Smale Model. *SIAM Journal on Mathematical Analysis*, 42(1):218–236, 2010.
- [50] 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.

- [51] G. Cavagnari, A. Marigonda, and B. Piccoli. Averaged Time-Optimal Control Problem in the Space of Positive Borel Measures. *ESAIM COCV*, 24(2):721–740, 2018.
- [52] H. Chakir, J.P. Gauthier, and I. Kupka. Small Sub-Riemannian Balls on \mathbb{R}^3 . *Journal of Dynamical and Control Systems*, 2(3), 1996.
- [53] Y. Chitour, F. Colonius, and M. Sigalotti. Growth Rates for Persistently Excited Linear Systems. *Mathematics of Control, Signals and Systems*, 26(4):589–616, 2012.
- [54] Y. Chitour and M. Sigalotti.
- [55] L. Chizat and F. Bach. On the Global Convergence of Gradient Descent for Over-parameterized Models using Optimal Transport.
- [56] L. Chizat, G. Peyré, B. Schmitzer, and F.X. Vialard. Scaling Algorithms for Unbalanced Optimal Transport Problems. *Mathematics of Computations*, 87:2563–2609, 2018.
- [57] Y.T. Chow and W. Gangbo. A Partial Laplacian as an Infinitesimal Generator on the Wasserstein Space. arXiv:1710.10536.
- [58] R. Cibulka, A.L. Dontchev, M.I. Krastanov, and V.M. Veliov. Metrically Regular Differential Generalized Equations. *SIAM J. Cont. Opt.*, 56(1):316–342, 2018.
- [59] F. Clarke. *Functional Analysis, Calculus of Variations and Optimal Control*. Springer, 2013.
- [60] J. Cortés and Egerstedt. Coordinated Control of Multi-Robot Systems: A Survey. *SICE Journal of Control, Measurement, and System Integration*, 10(6):495–503, 2017.
- [61] E. Cristiani, B. Piccoli, and A. Tosin. *Multiscale Modeling of Pedestrian Dynamics*, volume 12. Springer, 2014.
- [62] F. Cucker and S. Smale. Emergent Behavior in Flocks. *IEEE Trans. Automat. Control*, 52(5):852–862, 2007.
- [63] F. Cucker and S. Smale. On the Mathematics of Emergence. *Japanese Journal of Mathematics*, 2(1):197–227, 2007.
- [64] F. Dalmao and E. Mordecki. Cucker-Smale Flocking under Hierarchical Leadership and Random Interactions. *SIAM J. Appl. Math.*, 71(4):1307–1316, 2011.
- [65] J.M. Danskin. *The Theory of Max-Min and its Application to Weapons Allocation Problems*, volume 5 of *Ökonometrie und Unternehmensforschung Econometrics and Operations Research*. Springer-Verlag Berlin Heidelberg, 1967.
- [66] G. De Phillipis. Regularity of Optimal Transport Maps and Applications, 2012. <http://cvgmt.sns.it/media/doc/paper/1978/PHDthesis.pdf>.
- [67] G. De Phillipis and A. Figalli. Regularity for Solutions of the Monge-Ampère Equation. *Inventiones Mathematicae*, 192(1):55–69, 2013.
- [68] R.L. Di Perna and Lions P.L. Ordinary Differential Equations, Transport Theory and Sobolev Spaces. *Inventiones Mathematicae*, 98(3):511–548, 1989.
- [69] J. Diestel and J.J.Jr Uhl. *Vector Measures*, volume 15. American Mathematical Society, 1977.

- [70] A.L. Dontchev and W.W. Hager. Lipschitzian Stability in Nonlinear Control and Optimization. *SIAM Journal on Control and Optimization*, 31(3):569–603, 1993.
- [71] A.L. Dontchev, M.I. Krastanov, and V.M. Veliov. On the Existence of Lipschitz Continuous Optimal Feedback Control. *Research Report*. ISSN 2521-313X.
- [72] J. Dugundji. An Extension of Tietze’s Theorem. *Pacific Journal of Mathematics*, 1(3):353–367, 1951.
- [73] M. Duprez, M. Morancey, and F. Rossi. Approximate and Exact Controllability of the Continuity Equation with a Localized Vector Field. *Siam Journal on Control and Optimization*.
- [74] M. Duprez, M. Morancey, and F. Rossi. Minimal Time Problem for Crowd Models with a Localized Vector Field. 2019. hal-01676679.
- [75] K. Elamvazhuthi and S. Berman. Optimal Control of Stochastic Coverage Strategies for Robotic Swarms. In *IEEE Int. Conf. Rob. Aut.*, 2015.
- [76] H. Federer. *Geometric Measure Theory*. Springer, 1996.
- [77] A. Ferscha and K. Zia. Lifebelt: Crowd evacuation based on vibro-tactile guidance. *IEEE Pervasive Computing*, 9(4):33–42, 2010.
- [78] A. Figalli. Optimal Transport, Euler Equations, Mather and DiPerna-Lions Theory, 2009.
- [79] A. Figalli, Y.H. Kim, and R.J. McCann. Hölder Continuity and Injectivity of Optimal Maps. *Archives of Rational Mechanics and Analysis*, 209(3):747–795, 2013.
- [80] M. Fornasier, S. Lisini, C. Orrieri, and G. Savaré. Mean-Field Optimal Control as Gamma-Limit of Finite Agent Controls. *European Journal of Applied Mathematics*, pages 1–34, 2019.
- [81] M. Fornasier, B. Piccoli, and F. Rossi. Mean-Field Sparse Optimal Control. *Phil. Trans. R. Soc. A*, 372(2028):20130400, 2014.
- [82] M. Fornasier and F. Solombrino. Mean Field Optimal Control. *Esaim COCV*, 20(4):1123–1152, 2014.
- [83] G.N. Galbraith and R.B. Vinter. Lipschitz Continuity of Optimal Controls for State-Constrained Problems. *SIAM Journal on Control and Optimization*, 42(5):1727–1744, 2003.
- [84] R.V. Gamkrelidze. Optimal Control Processes under Bounded Phase Coordinates. *Izv. Akad. Nauk SSSR Math.*, 24(3):315–356, 1959.
- [85] W. Gangbo, T. Nguyen, , and A. Tudorascu. Hamilton-Jacobi Equations in the Wasserstein Space. *Methods and Applications of Analysis*, 15(2):155–184, 2008.
- [86] W. Gangbo and A. Tudorascu. On Differentiability in Wasserstein Spaces and Well-Posedness for Hamilton-Jacobi Equations. *Technical Report*, 2017.
- [87] Ugo Gianazza, Giuseppe Savaré, and Giuseppe Toscani. The Wasserstein Gradient Flow of the Fisher Information and the Quantum Drift-Diffusion Equation. *Archives of Rational Mechanics and Analysis*, 1(194):133–220, 2009.
- [88] N. Gigli. *Second Order Analysis on $(\mathcal{P}_2(M), W_2)$* , volume 216 of *Memoirs of the American Mathematical Society*. AMS, 2012.

- [89] S.Y. Ha and J.G. Liu. A Simple Proof of the Cucker-Smale Flocking Dynamics and Mean-Field Limit. *Comm. Math. Sci.*, 7(2):297–325, 2009.
- [90] W.W. Hager. Multiplier Method for Nonlinear Optimal Control. *SIAM Journal on Control and Optimization*, 27(4):1061–1080, 1990.
- [91] R. Hegselmann and U. Krause. Opinion Dynamics and Bounded Confidence Models, Analysis, and Simulation. *Journal of artificial societies and social simulation*, 5(3), 2002.
- [92] M.Y. Huang, R. Malhamé, and P.E. Caines. Large Population Stochastic Dynamic Games : Closed-Loop McKean-Vlasov Systems and the Nash Certainty Equivalence Principle. *Communications in Information and Systems*, 6(3):221–252, 2006.
- [93] Shvartsman I.A. and Vinter R.B. Regularity Properties of Optimal Controls for Problems with Time-Varying State and Control Constraints. *Nonlinear Analysis*, 65:448–474, 2006.
- [94] 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.
- [95] A.D. Ioffe and V.M. Tihomirov. *Theory of Extremal Problems*. North Holland Publishing Company, Elsevier, 1979.
- [96] R. Jordan, D. Kinderlehrer, and F. Otto. The Variational Formulation of the Fokker-Planck Equation. *SIAM Journal of Mathematical Analysis*, 29(1):1–17, 1998.
- [97] L.V. Kantorovich. On the Translocation of Mass. *Dokl. Akad. Nauk. USSR*, 37:199–201, 1942.
- [98] J-M. Lasry and P-L. Lions. Mean Field Games. *Japanese Journal of Mathematics*, 2(1):229–260, 2007.
- [99] H. Lavenant and F. Santambrogio. Optimal Density Evolution with Congestion: L^∞ Bounds via Flow Interchange Techniques and Applications to Variational Mean Field Games. *Communications in Partial Differential Equations*, 43(12):1761–1802, 2018.
- [100] M. Lavrentiev. Sur Quelques Problèmes du Calcul des Variations. *Annali di Matematica Pura e Applicata*, 4(1):7–28, 1927.
- [101] J. Lott and C. Villani. Ricci Curvature for Metric-Measure Spaces via Optimal Transport. *Annals of Mathematics*, 169(3):903–991, 2009.
- [102] M.-A. Maghenem and A. Loría. Strict Lyapunov Functions for Time-Varying Systems with Persistency of Excitation. *Automatica*, 78:274–279, 2017.
- [103] Mohamed Maghenem, Antonio Loría, and Elena Panteley. Formation-tracking control of autonomous vehicles under relaxed persistency of excitation conditions. *IEEE Transactions on Control Systems Technology*, 26(5):1860–1865, 2018.
- [104] B. Maury and J. Venel. A mathematical framework for a crowd motion model. *Comptes Rendus Mathématique*, 346(23):1245–1250, 2008.
- [105] B. Maury and J. Venel. A discrete contact model for crowd motion. *ESAIM: Mathematical Modelling and Numerical Analysis*, 45(1):145–168, 2011.

- [106] G. Mazanti and F. Santambrogio. Minimal-Time Mean-Field Games. *Submitted*, 2018.
- [107] F. Mazenc and M. Malisoff. *Construction of Strict Lyapunov Functions*. Springer-Verlag, 2009.
- [108] S. McQuade, B. Piccoli, and N. Pouradier Duteil. Social Dynamics Models with Time-varying Influence. *Accepted in Math. Models Methods Appl. Sci.*, 2019.
- [109] G. Monge. Mémoire sur la Théorie des Déblais et des Remblais. *Histoire de l'Académie Royale des Sciences de Paris*, 1781.
- [110] L. Moreau. Stability of Multiagent Systems with Time-Dependent Communication Links. *IEEE Trans. Automat. Control*, pages 169–182, 2005.
- [111] A. Muntean, J. Rademacher, and A. Zagaris. *Macroscopic and large scale phenomena: coarse graining, mean field limits and ergodicity*. Springer, 2016.
- [112] F. Otto. The Geometry of Dissipative Equations : The Porous Medium Equation. *Communications in Partial Differential Equations*, 26:101–174.
- [113] 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.
- [114] 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.
- [115] B. Piccoli, F. Rossi, and E. Trélat. Control of the kinetic Cucker-Smale model. *SIAM J. Math. Anal.*, 47(6):4685–4719, 2015.
- [116] B. Piccoli, F. Rossi, and E. Trélat. Sparse control of second-order cooperative systems and partial differential equations to approximate alignment. In *22nd International Symposium on Mathematical Theory of Networks and Systems, 2016*, 2016.
- [117] N. Pogodaev. Numerical Algorithm for Optimal Control of Continuity Equations. arXiv:1708.05516.
- [118] N. Pogodaev. Optimal Control of Continuity Equations. *Nonlinear Differential Equations and Applications*, 23:21, 2016.
- [119] L.S. Pontryagin. *The Mathematical Theory of Optimal Processes*, volume 4. CRC Press, 1987.
- [120] A. Prosinski and F. Santambrogio. Global-in-Time Regularity via Duality for Congestion-Penalized Mean Field Games. *Stochastics*, 89(6-7):923–942, 2017.
- [121] J. Rabin, G. Peyré, J. Delon, and M. Bernot. Wasserstein Barycenter and Its Application to Texture Mixing. In *Scale Space and Variational Methods in Computer Vision*, volume 6667 of *Lecture Notes in Computer Science*, pages 435–446.
- [122] L.N. Ru, Li Z.C., and Xue X.P. Cucker-Smale flocking with randomly failed interactions. *Journal of the Franklin Institute*, 352 (3):1099–1118, 2015.
- [123] W. Rudin. *Real and Complex Analysis*. Mathematical Series. McGraw-Hill International Editions, 1987.
- [124] L. Sacchelli. Short Geodesics Losing Optimality in Contact Sub-Riemannian Manifolds and Stability of the 5-Dimensional Caustic. *Accepted for publication in SIAM Journal of Control*, 2019.

- [125] F. Santambrogio. *Optimal Transport for Applied Mathematicians*, volume 87. Birkhauser Basel, 2015.
- [126] 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.
- [127] H.L. Smith. *Monotone dynamical systems*. AMS, 1995.
- [128] J. Solomon, F. De Goes, G. Peyré, M. Cuturi, A. Butscher, A. Nguyen, T. Du, and L. Guibas. Convolutional Wasserstein Distances: Efficient Optimal Transportation on Geometric Domains. *ACM Transactions on Graphics*, 34(4), 2015.
- [129] H. Spohn. *Large scale dynamics of interacting particles*. Springer Science & Business Media, 2012.
- [130] K.T. Sturm. On the Geometry of Metric Measure Spaces. *Acta Mathematica*, 196(1):65–131, 2006.
- [131] F Tröltzsch. *Optimal Control of Partial Differential Equations*. American Mathematical Society, 2010.
- [132] C. Villani. *Optimal Transport : Old and New*. Springer-Verlag, Berlin, 2009.
- [133] R.B. Vinter. *Optimal Control*. Modern Birkhauser Classics. Birkhauser Basel, 2000.
- [134] A.A. Vlasov. *Many-Particle Theory and its Application to Plasma*. New York, Gordon and Breach, 1961.