Tesi di Dottorato

Luciano Marzufero

Some optimal visiting problems: from a single player to a mean-field type model

Dottorato in Matematica, Trento (2022).

<http://www.bdim.eu/item?id=tesi_2022_MarzuferoLuciano_1>

University of Trento

DOCTORAL PROGRAMME IN MATHEMATICS

Degree of Doctor of Philosophy in Mathematics
Cycle XXXIV

Some optimal visiting problems: from a single player to a mean-field type model

Ph.D. Thesis in Mathematical Analysis (S.S.D. MAT/05)

Candidate:
Luciano Marzufero

Supervisor:
Prof. Fabio Bagagiolo

YEARS 2018-2022

Preface

This thesis is the result of the author's Ph.D. research, supervised by professor Fabio Bagagiolo. This research was financially supported by means of a fellowship awarded by the University of Trento, Dipartimenti di Eccellenza.

First of all, I want to express my gratitude to Fabio Bagagiolo for his support during this Ph.D., for giving me the freedom to choose own topics in the field of Optimal Control and Mean-Field Games Theory and for encouraging me to collaborate with prominent researchers in Italy and abroad. The results in this thesis would not have been possible without his helpful suggestions, insightful conversations and useful comments. Moreover, I am very grateful to the Department of Mathematics for giving me the opportunity to do a Ph.D. in my country and to all the professors and researchers who in these years gave a contribution to this work with valuable talks and advices. In particular, I thank my other co-author Adriano Festa I have collaborated with during this Ph.D. Furthermore, I am grateful to all my office colleagues at the Department of Mathematics, for the fruitful discussions which sometimes have led to comical jokes. A special mention to Christiaan and Francesco.

As a mathematical analyst, it is an honor and privilege to do this work, and I can say that I feel very lucky and satisfied to have been in the position to do this Ph.D. in Trento. In particular, the freedom of choice of research topics that came with this Ph.D. fellowship allowed me to combine several of my interests: mathematics and real-world applications. I had a great time the last three years in Trento, during which I learned a great deal about mathematical analysis, but I also developed myself as a person.

Finally, I would like to thank all my longtime friends and my family for their care in my development and supporting my decisions. A special thought is due also to my Bachelor's Degree and Master's Degree supervisors, Andrea Spiro and Cristina Giannotti, for stimulating me to follow my dream to find a Ph.D. position, and for always being there to help and give positive advices.

Contents

Introduction

In this thesis, we deal with the problem of optimizing a trajectory to "visit", i.e., to touch or at least to pass as close as possible, to a collection of targets. In the following, we refer to this problem as *optimal visiting*. The issue presents various inherent difficulties: some related to its high computational complexity (shared with other well-known optimization problems as the "Traveling salesman problem" [27]) and other related to its possible continuous/discontinuous nature.

Let us state the problem more precisely: consider the controlled dynamics

$$\begin{cases} y'(s) = f(y(s), \alpha(s), s), & s \in]t, T] \\ y(t) = x, & x \in \mathbb{R}^d \end{cases}, \tag{0.0.1}$$

where $t \in [0, T]$, $\alpha : [t, T] \longrightarrow A$ is a measurable control function, and the dynamics $f : \mathbb{R}^d \times A \times [0, +\infty[\longrightarrow \mathbb{R}^d$ is suitably regular. Consider a collection of N compact disjoint target sets in $\mathbb{R}^d$, $\{\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_N\}$, and $y_{(x,t)}(\cdot; \alpha)$ a solution of (0.0.1) related to a starting point x, a starting time t and a control α. We can write the optimal visiting problem just considering, for instance, the minimization of a cost functional of the form, in a finite horizon feature,

$$J(x, t, \alpha) = \int_t^T \ell(y(s), \alpha(s), s) ds,$$

with the running cost ℓ suitably designed in order to keep trace of the distances from the targets of the trajectory. The visiting cost is then defined as

$$v(x, t) = \inf_{\alpha \in \mathcal{A}} J(x, t, \alpha).$$

Actually, the problem requires a particular framework as a standard continuous optimal control setting fails to describe the problem correctly. Let us illustrate this difficulty using the following toy example.

Let us consider the one-dimensional problem with $A = \mathbb{R}$, $f(x, a) = a$, and $\mathcal{T}_1 =]-\infty, -1]$, $\mathcal{T}_2 = [1, +\infty[$. We focus on giving an optimal visit formulation in the interval $[-1, 1]$. At first, we consider the easiest running cost

$$\ell(y, a, t) := \frac{1}{2}\left(\sum_{j=1}^{2} d(y, \mathcal{T}_j)^2 + \|a\|^2\right),$$

which penalizes quadratically the distance from the targets and the norm of the control. It is easy to verify that we have a feedback formula for the optimal control as

$$\alpha(x, t) = -\frac{x}{(1-t)^2 + 1/2},$$

which means that the trajectory is led to zero, which is the middle point between the two targets. Since we want to model a slightly different problem, i.e., a visit more than a compromise between the distances, we are unsatisfied with this result.

Therefore, we should include the information about the visit of the targets in a different way in the model, allowing us to focus on a single target, as well as on a subfamily of targets, at once. If, for example, we consider the problem of visiting first target $\mathcal{T}_1$ and then $\mathcal{T}_2$, we can easily observe that we would obtain a different problem just swapping the order of the visit. This is a consequence because no Dynamical Programming Principle would be available for the value function $v(x, t)$, since the only information brought by the state-position x does not give information about the already visited targets (see also [3, 2, 5, 6]). Hence, at this level, we can not in general characterize the visiting cost as a suitable solution of a Hamilton-Jacobi equation. Consequently, it is quite challenging to perform a global study of the problem or obtain a feedback optimal control map.

The argument above suggests that we need to include in the model a "memory" of the targets already visited. This can be done using various tools. Here, we opt for a hybrid control-based construction. In particular, we introduce additional discrete state-variables as switching N-strings p of 0 and 1, where 1 in the i-position means that the target $\mathcal{T}_i$ has been already visited and viceversa for 0, and they have a discontinuous evolution in time. Hence, starting from $p_o = (0, \ldots, 0)$, the goal can be seen as obtaining the string $\bar{p} = (1, \ldots, 1)$, i.e., visiting all the N targets, paying as less as possible. We then split the optimal visiting problem into several problems, labeled by the N-strings p, and we suitably interpret it as a collection of several optimal stopping/switching problems coupled to each other by the stopping/switching cost (a switch between N-strings corresponds to the visit of a target or the choice to forgo visiting one or more targets). This leads to the study of the following Hamilton-Jacobi equation

$$\begin{cases} \max\{v(x,t)-\psi(x,t), -v_t(x,t)+\lambda v(x,t)+H(x,t,D_x v(x,t))\}=0, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (x,t)\in\mathbb{R}^d\times[0,T[\\ v(x,T)=\psi(x,T), \qquad\qquad\qquad\qquad\qquad\qquad x\in\mathbb{R}^d \end{cases}, \tag{0.0.2}$$

where ψ is the stopping/switching cost and H is the Hamiltonian, which characterizes the value function v as the unique viscosity solution.

0.1 A network representation

We study also an optimal visiting problem in a quite different framework than $\mathbb{R}^d$. Indeed, we interpret the additional discrete state-variables p, which give a memory feature to the model, as nodes of a direct network, where p_o is the origin and $\bar{p}$ is the final destination. The problem can be seen then as the search for an optimal origin-destination path. Due to the dynamical feature of the multi-dimensional optimal visiting problem described before, in the network representation we keep the possibility to choose the sequence of instant to perform the switches. Our idea is then to study the problem without a real dynamics as (0.0.1), i.e., without a controlled continuous trajectory for visiting the N targets. In particular, the state of the system is represented by the discrete variable p, which basically corresponds to the node of the network on which the agent is. Such a variable acts also as a switching discrete control at the agent's disposal, that is, once the agent is on the node p, it has to choose optimally the next admissible subsequent node p' after p. In this way, the agent switches to p', i.e., visit a new target, and the state of the system becomes p'. In performing such a switch, the agent incurs a switching cost. A time-variable is accounted for the problem too. In particular, besides the switching discrete control variable p, the agent has to choose the optimal time it is convenient to switch to the next node of the network. Moreover, all the admissible switches have to be performed within a fixed time $T>0$: if the agent reaches the final node before T, for example it pays an earliness penalization cost, while if it does not reach the final node and the time is over, it pays a time-loseness penalization cost.

0.2 The mean-field case

In this thesis, we are also interested in a possible study of a mean-field game model, that is when a huge population of agents plays the optimal visiting problem with/without a controlled dynamics and with costs also depending on the distribution of the population. Actually, the study of the interacting motion of many agents with more than one target seems to be rather new in the literature, especially for what concerns the corresponding continuity equation for the mass distribution. As regards the multi-dimensional optimal visiting problem (i.e., in $\mathbb{R}^d$ with controlled dynamics (0.0.1)), we start such a kind of study investigating a single continuity equation with a mass-sink in $\mathbb{R}^d$ (corresponding to the case where some agents, possibly labeled by the same N-strings p, visit or forgo visiting some targets and then pass to a subsequent level, labeled by another N-string p'), and then we extend the analysis to a more general transport equation with a mass-source too (corresponding to the previous case with the addition of agents coming from an admissible preceding level). The single continuity equation with just a mass-sink is formally described by

$$\begin{cases} \mu_t(x,t) + \operatorname{div}(\mu(x,t)b(x,t)) + \mathbb{1}_{\{(\mathcal{S},t):t\in[0,T]\}}\mu(x,t) = 0, \\ \hfill (x,t) \in \mathbb{R}^d \times [0,T] \\ \mu(x,0) = \mu_0(x), \hfill x \in \mathbb{R}^d \end{cases}, \tag{0.2.3}$$

which models the transport of a mass μ of agents with initial distribution μ_0, subject to a given suitably regular field b (possibly depending on the mass μ too) with the presence of a region $\mathcal{S} \subset \mathbb{R}^d$ (possibly depending on time) acting as a sink: the portion of mass possibly entering the sink instantaneously disappears. Roughly speaking, the first two terms give the evolution of the mass μ and the third $\mathbb{1}_{\{(\mathcal{S},t):t\in[0,T]\}}\mu(x,t)$ represents the leaving rate of agents which stop in $\mathcal{S}$ and then vanish. We prove that the unique candidate for suitably solving (0.2.3) is the following measure in $\mathbb{R}^d$

$$\tilde{\mu}(s) = \begin{cases} \Phi(\cdot,0,s)\sharp\tilde{m}_0 & \text{on } \mathbb{R}^d \setminus \mathcal{K}(s) \\ 0, & \text{otherwise} \end{cases}, \quad s \in [0,T],$$

where Φ is the flow generated by the field b, and $\mathcal{K}(s)$ is suitably defined in such a way that it takes account of the agents who passed through the sink $\mathcal{S}$ at least once in the time interval $[0,T]$ and then disappeared.

We generalize then the analysis of the continuity equation with a sink to the case of the evolution in $\mathbb{R}^d$ of a mass μ with the presence of a region acting as a source: the portion of mass possibly coming out from the source starts flowing immediately, according to the field b. In particular, starting from the ideas for the study of (0.2.3), we find a suitable candidate for the evolution and we prove its uniqueness.

The results that are shown for what concerns the study of a single continuity equation with sinks and sources (including the possible dependence of the field b on the measure μ), also constitute the first step toward a general theory of mean-field games in the presence of switches in the formulation of the problem. To complete the theory, a major point to be investigated would be the introduction of a real coupling between the continuity equations with sinks and sources and the Hamilton-Jacobi equations (0.0.2) via the optimal feedback (regularity assumptions on the vector field should be probably adjusted). In any case, as shown in the tests in [6], we may suppose, due to some promising numerical evidence, the existence of an equilibrium for such a coupling. We postpone this study to future research.

We investigated also a mean-field game model for the optimal visiting problem on a network in §0.1. In particular, after studying the single-player optimization problem and the properties of the value function, we face the problem of the existence of a mean-field equilibrium. This is done by performing a suitable fixed-point procedure for an approximated problem, and then we address the passage to the limit in the approximation. We need first an approximated problem because the switching mass-evolution, solution of the mean-field equilibrium problem, turns out to be piecewise continuous (even piecewise constant in some particular case) and this fact makes the standard compactness and convexity requirements for fixed-point results lacking in our case. Moreover, possibly due to non-uniqueness of the optimal control, we have to work with set-valued functions and, similarly as in [2, 12], we must consider agents splitting into fractions, each one of them following one of the optimal behaviors. That rather new approximation allows us to prove the existence of an approximated mean-field equilibrium via a fixed-point procedure for a suitable set-valued map. The passage to the limit in the approximation is then investigated by assuming a suitable hypothesis on the optimal switching instants. Anyway, such a hypothesis can be satisfied by requiring some proper conditions on the costs. A more general investigation for avoiding this assumption is left to future works. As regards the uniqueness of the equilibrium, usually, in the mean-field games theory, it is guaranteed by imposing a kind of monotonicity condition satisfied by the costs with respect to the mass of the agents (see [31]). In several cases, the adaptation of that property to uniqueness results does not require too much work because the studied problem almost naturally fits that condition. Our problem, due to many of its aspects, does not provide instead an immediate evident way to adapt it. However, inspired by [31], using a monotonicity-type property, we give some easy examples and calculations which seem to be promising for a future and deeper study of the uniqueness.

0.3 Discussion

In general, as aforementioned, the study of single-player optimal visiting problems requires an hybrid control framework in order to recover a dynamic programming property, and hence to derive a Hamilton-Jacobi equation. More precisely, it requires a special framework able to include a memory of the targets already visited. For the formulation, we adapt to our setting some classic results of viscosity solutions theory that can be found, e.g., in [9, 25] (see also Appendix B, §B.1). In particular, the hybrid framework that we propose is related to hybrid control [15] and somehow to the mathematical switching hysteresis models [38]. The need of a memory feature, associated with the optimal visiting, dynamic programming and Hamilton-Jacobi equations, has been presented in [3], where a continuous hysteresis memory was introduced. The use of a switching/discontinuous/hybrid memory, as in the present thesis, was instead used for a one-dimensional optimal visiting problem on a network in [2], which basically inspired the model introduced in §0.1. For switching hybrid control problems related to the models here presented, and in connection with Hamilton-Jacobi equations, we refer to [11] (similar formulations for the deterministic case have also been proposed in [15, 22]).

The literature concerning the Traveling Salesman Problem, to which our optimal visiting problems are related, is very large. We only quote an early paper by R. Bellman [10] devoted to the problem and dynamic programming.

The model for a crowd of indistinguishable players is taken from the framework of mean-field games [31, 30, 29, 19], while the adaptation of the same hybrid structure to networks has been only very recently attempted, as in [2, 4] and, more generally, in [16, 17]. Some works which share the same ideas to treat the mean-field case in the presence of switches in the dynamics of the problem are [13, 12], where the author discusses a mean-field optimal stopping and impulse control problem, and [24], where a hybrid mean-field game is presented to model a multi-lane traffic flux of vehicles. Moreover, another similar mean-field model can be found in [33], where a continuous and a discrete set of switching labels are introduced to study the case of a leader-follower dynamics.

There are several applicative motivations for our models. For instance, the multi-dimensional optimal visiting problem complies with tourists' flow which has to visit some points of interest both in an heritage city and in a museum environment (see for example [2, 20]). In [14] it is instead given an example of a situation where, in a crowded environment, people have to perform a sequence of different operations in different places, such as in big airports or train stations. Still concerning the multi-dimensional setting, but related to a single-player one, in [7] it is given an example of a framework which is used to solve a series of applied problems arising from the sport of orienteering races. Finally, as we already said in §0.1, we rewrite the optimal visiting problem as an origin-to-destination one on a network, and of course the possible applications and literature on this kind of problems are very huge. Other possible interesting interpretations of such a model could be found in the optimal job scheduling and in the similar open-shop scheduling problem, or in the optimal co-flow scheduling in operations research (see for example [37, 28] and [21]).

0.4 Reading this thesis

The body of the thesis is organized as follows. In Chapter 1, we introduce the optimal visiting problem in $\mathbb{R}^d$ for a single agent, reporting all the theoretical elements that justify the use of a Hamilton-Jacobi formulation. In support of this, in Appendix A, §A.1 we state and prove several preliminary results. The main contributions have been published in [5, 6] (together with some numerical tests in order to verify in practice the model), while in [7] it is given an original approach for a sport game known as orienteering problem, based on the hybrid control techniques in this chapter.

In Chapter 2, we consider a crowd of indistinguishable players focusing on the good position of the continuity equation (at first with a sink and then with a source too), which models the motion of the density of players. This is a first essential step for a possible study of a more general mean-field game model. The main contributions in this chapter have been published in [6], where, however, just the continuity equation with a sink is studied. Furthermore, in the same paper, the model in action through a collection of some promising numerical tests is also examined.

In Chapter 3, we introduce a network representation of the multi-dimensional optimal visiting problem in Ch. 1, for a single agent and for a crowd, giving all the theoretical elements and hypotheses that motivate the use of a switching feature on a different framework. In particular, we focus on the well-position of such a problem and we address also the mean-field case. In Appendix A, §A.2 and §A.3, some examples and calculations are given to support the main results. The contributions in this chapter have been submitted for publication in [8].

In Appendix B, we state some important tools and useful results which are needed in the thesis.

Notations

$\mathbb{R}^d$	the Euclidean d-dimensional space
$\|x\|$	the absolute value or modulus of a real number x
$x \cdot y$ or $\langle x, y\rangle$	the scalar product $\sum_{i=1}^{d} x_i y_i$ of vectors $x = (x_1, \dots, x_d)$ and $y = (y_1, \dots, y_d)$
$\\|x\\|$	the Euclidean norm of $x \in \mathbb{R}^d$, $\\|x\\| = \langle x, x\rangle^{\frac{1}{2}}$
$\\|x\\|_\infty$	the supremum norm $\sup_{x\in E} \|u(x)\|$ of a function $u : E \longrightarrow \mathbb{R}$
$d(x, E)$	the distance from x to E (i.e., $d(x, E) = \inf_{y\in E} \\|x - y\\|$)
$\mathrm{d_H}(E, S)$	the Hausdorff distance between the sets E and S (§2.1.1)
$B(x_0, r)$	the open ball $\{x \in \mathbb{R}^d : \\|x - x_0\\| < r\}$
$\overline{B}(x_0, r)$	the closed ball $\{x \in \mathbb{R}^d : \\|x - x_0\\| \le r\}$
$u_t(x, t)$ or $\partial_t u(x, t)$	the time derivative of the function u, i.e. the derivative w.r.t. t
$D_x u(x, t)$	the spatial gradient of the function u, i.e. the gradient w.r.t. x
$D^+u(x, t)$, $D^-u(x, t)$	the super- and the subdifferential (or semidifferentials) of u at (x, t) (§1.1.3)
Jf	the Jacobian matrix of a function $f : \mathbb{R}^d \longrightarrow \mathbb{R}^m$
$\mathcal{P}(E)$	the power set of a set E
$C^0(E)$	the space of continuous functions $u : E \longrightarrow \mathbb{R}$
$C^0(E, D)$	the space of continuous functions $u : E \longrightarrow D$, $D \neq \mathbb{R}$
$C^1(E)$	the space of continuously differentiable functions $u : E \longrightarrow \mathbb{R}$

$C^\infty(E)$	the space of infinitely differentiable functions (i.e., possess derivatives of all orders in E) $u : E \longrightarrow \mathbb{R}$
$C_c^k(E)$	the space of compactly supported and k-times continuously differentiable functions $u : E \longrightarrow \mathbb{R}$, $0 \leq k \leq \infty$
$L^p(E)$	the Lebesgue space of integrable functions $u : E \longrightarrow \mathbb{R}$, $1 \leq p < \infty$
$L^p(E, D)$	the Lebesgue space of integrable functions $u : E \longrightarrow D$, $D \neq \mathbb{R}$, $1 \leq p < \infty$
$L^\infty(E)$	the space of essentially bounded functions $u : E \longrightarrow \mathbb{R}$
$L^\infty(E, D)$	the space of essentially bounded functions $u : E \longrightarrow D$, $D \neq \mathbb{R}$
$\mathrm{BUC}(E)$	the space of bounded and uniformly continuous functions $u : E \longrightarrow \mathbb{R}$
$\mathcal{PC}(E, [0,1])$	the space of piecewise constant functions $u : E \longrightarrow [0, 1]$
$\mathcal{L}^d(E)$	the d-dimensional Lebesgue measure of E
$\mathcal{H}^d(E)$	the d-dimensional Hausdorff measure of E
$\mathbb{1}_E$	the indicator function of a set E, that is $\mathbb{1}_E(x) = 1$ if $x \in E$ and $\mathbb{1}_E(x) = 0$ if $x \notin E$

Chapter 1

Hybrid control for a single-player optimal visiting problem

The optimal visiting problem is the optimization of a trajectory that has to touch or pass as close as possible to a collection of target points. The problem does not verify the dynamic programming principle, and it needs a specific formulation to keep track of the visited target points. In this chapter, we introduce a hybrid approach by adding a discontinuous part of the trajectory switching between a group of discrete states related to the targets. Then, we show the well-position of the related Hamilton-Jacobi problem, by reformulating the optimal visiting as a collection of time-dependent optimal stopping problems.

1.1 The optimal visiting problem

Given N disjoint compact target sets $\{\mathcal{T}_j\}_{j=1,\dots,N} \subset \mathbb{R}^d$, we represent the state of the system by the pair $(x,p) \in \mathbb{R}^d \times \mathcal{I}$, where $p = (p^1, p^2, \dots, p^N) \in \mathcal{I} = \{0,1\}^N$. Therefore, x is the continuous state variable (i.e., the position in $\mathbb{R}^d$) and p is the switching discrete state variable. The evolution of the continuous variable is described by the controlled dynamics

$$\begin{cases} y'(s) = f(y(s), \alpha(s), q(s)), & \text{a.e. } s \in]t, T] \\ y(t) = x, \ q(t) = p \end{cases}, \tag{1.1.1}$$

where $(x, p) \in \mathbb{R}^d \times \mathcal{I}$ is the initial state, $t \in [0, T]$ the initial instant, $T > 0$ the fixed finite horizon. The measurable control is (for $A \subset \mathbb{R}^m$ compact)

$$\alpha \in \mathcal{A} := \{\alpha : [0, +\infty[\longrightarrow A \text{ measurable}\}$$

and the dynamics $q(\cdot)$ of the switching variable (which represents here the memory) is subject to

$$\exists \tau \in [t, s],\ y(\tau) \in \mathcal{T}_j \Rightarrow\ q^j(s) = 1;\quad q^j(s) = p^j \text{ otherwise.} \tag{1.1.2}$$

Formally, $q^j(s) = 0$ means that the target $\mathcal{T}_j$ has not been visited yet in $[t, s]$ and viceversa for $q^j(s) = 1$. The dynamics $f : \mathbb{R}^d \times A \times \mathcal{I} \longrightarrow \mathbb{R}^d$ is continuous, bounded and Lipschitz continuous w.r.t. $x \in \mathbb{R}^d$ uniformly w.r.t. $(a, p) \in A \times \mathcal{I}$, i.e., there exists $L > 0$ such that

$$\|f(x, a, p) - f(y, a, p)\| \leq L\|x - y\| \text{ for all } (x, y) \in \mathbb{R}^d \times \mathbb{R}^d \text{ and } (a, p) \in A \times \mathcal{I}.$$

The state of the system at time s is the pair $(y(s), q(s))$ and, for every initial state (x, t, p) and control α, by our hypotheses the existence of a unique solution $(y(s), q(s))$ of (1.1.1)-(1.1.2) is guaranteed. In particular, note that the number of switches of the variable q is necessarily finite, hence q is piecewise constant and the solution $y^\alpha_{(x,t,p)}(s)$ (or simply $y(s)$) of (1.1.1)-(1.1.2) is in the sense of absolutely continuous function.

The optimal visiting problem consists then in reaching, if possible, the discrete state $\bar{p} = (1, 1, \ldots, 1)$ (i.e., to visit all the targets) at a time $t \leq \bar{t} \leq T$, minimizing the cost

$$\int_t^{\bar{t}} e^{-\lambda(s-t)} \ell(y(s), \alpha(s), q(s), s) ds,$$

for a given running cost ℓ and a discount factor $\lambda > 0$.

1.1.1 A hybrid-control relaxation: optimal switching

The optimal control problem described above requires to "exactly touch" all the targets. This makes the evolution of the discrete variable q rather complicated, in particular in view of the corresponding Hamilton-Jacobi equation. We then relax the problem asking instead for "to pass as close as possible" to each target. Then we assume that we can definitely get rid of some targets and take into account only the remaining ones. In doing that, we also pay an additional cost depending, for instance, on the actual distance from the discarded targets. In this way, the evolution $q(\cdot)$ of the discrete

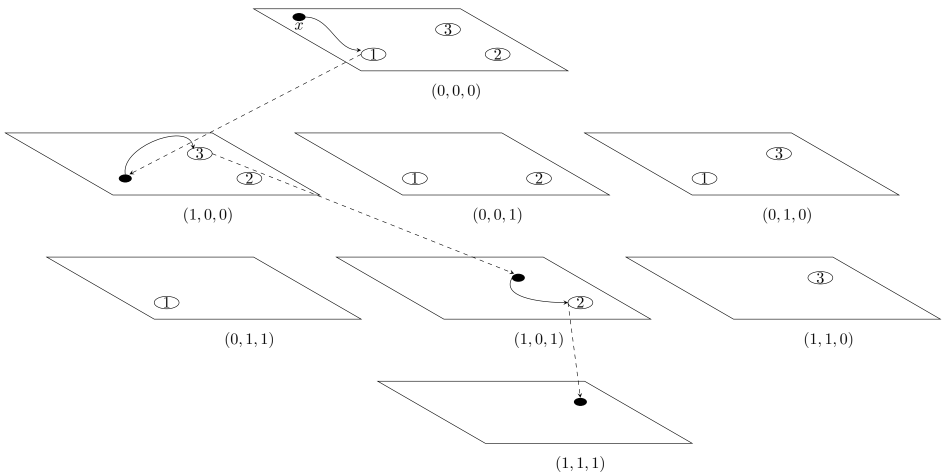


Figure 1.1: An optimal visiting problem with three targets: 1, 2, 3. The initial state is $(x, (0, 0, 0))$: no target visited/discarded yet. The agent first visits/discards target 1 and then the label switches to $(1, 0, 0)$. The second visited/discarded target is 3, and hence the second switch is to $(1, 0, 1)$. After visiting/discarding target 2, the final switch is to $(1, 1, 1)$. The rectangular indicates $\mathbb{R}^d$ and, for every label, the corresponding already visited/discarded targets are not displayed.

variables is no more a solution of (1.1.1)-(1.1.2), but instead, it becomes a control at our disposal. Clearly, there are some constraints: for example, for $N = 3$, if $p = (1, 0, 0)$, $p' = (1, 0, 1)$, $p'' = (0, 1, 1)$ and $p''' = \bar{p} = (1, 1, 1)$, then from p we can not switch to p'' otherwise we lose the information about the already visited/discarded target $\mathcal{T}_1$. However, we can switch to p''' directly. The process above is sketched in the Figure 1.1. In particular, by an optimization criterium, such a process is feasible because at every switching instant we get rid of a maximal quantity of targets, and hence no infinitesimal accumulation of subsequent switches is possible (no Zeno phenomenon). See also Figure 2.2.

Therefore, for any p, we denote by $\mathcal{I}_p$ the set of all possible new variables in $\mathcal{I}$ after a switch from p:

$$\mathcal{I}_p = \{\tilde{p} \in \mathcal{I} : p^i = 1 \Rightarrow \tilde{p}^i = 1 \text{ and } \exists l = 1, \ldots, N : p^l = 0, \ \tilde{p}^l = 1\}.$$

Note that in particular $\mathcal{I}_{\bar{p}} = \emptyset$, where $\bar{p} = (1, 1, \ldots 1)$.

For a given p, the number of the admissible subsequent switches is at most $N - \sum_i p^i \leq N$. Given the state (x, p) at the time t with $p \neq \bar{p}$, the controller chooses: the measurable control $\alpha \in \mathcal{A}$, and the discrete one $q : [0, +\infty[\longrightarrow \mathcal{I}$ which contains: the number $1 \leq m \leq N - \sum_i p^i$ of switches to be performed in order to reach $\bar{p}$, the switching instants $t \leq t_1 < t_2 < \ldots < t_m \leq T$ and the switching destinations $p_1, \ldots, p_{m-1}$, $p_m = \bar{p}$. Such destinations must satisfy $p_1 \in \mathcal{I}_p$, $p_{i+1} \in \mathcal{I}_{p_i}$, $i = 1, \ldots, m-1$. To resume, the control at disposal is then

$$(\alpha, m, t_1, \ldots, t_m, p_1, \ldots, p_{m-1}) =: u,$$

and note that for any (x, p, t) as above such a string belongs to a set depending on p and t denoted by $\mathcal{U}_{(p,t)}$. The cost to be minimized is

$$J(x, t, p, u) = \sum_{j=1}^{m} \Bigg(\int_{t_{j-1}}^{t_j} e^{-\lambda(s-t)} \ell(y(s), \alpha(s), p_{j-1}, s) ds \\ + e^{-\lambda(t_j - t)} C(y(t_j), p_{j-1}, p_j) \Bigg),$$

with $\lambda \geq 0$, $p_0 = p$, $t_0 = t$ and $y(s)$ is the solution of (1.1.1) where $q(s) = p_{j-1}$ if $s \in [t_{j-1}, t_j]$.

We assume $\ell : \mathbb{R}^d \times A \times \mathcal{I} \times [0, T] \longrightarrow [0, +\infty[$ bounded, continuous and uniformly continuous w.r.t. x uniformly w.r.t. $a \in A$, $p \in \mathcal{I}$ and $t \in [0, T]$. Moreover $C : \mathbb{R}^d \times \mathcal{I} \times \mathcal{I} \longrightarrow [0, +\infty[$ is uniformly continuous w.r.t. $x \in \mathbb{R}^d$, uniformly w.r.t. $p, p' \in \mathcal{I} \times \mathcal{I}_p$. Note that $C(x, p, p')$ represents the switching cost from p to p' when the state position is $x \in \mathbb{R}^d$. For example, it may depend on the distance from the discarded targets, that is $C(x, p, p') = \sum_j \chi_j(p, p') d(x, \mathcal{T}_j)$, where

$$\chi_j(p, p') = \begin{cases} 0, & p^j = p'^j \\ 1, & \text{otherwise} \end{cases}.$$

The value function of the problem is

$$V(x, t, p) = \inf_{u \in \mathcal{U}_{(p,t)}} J(x, t, p, u). \tag{1.1.3}$$

1.1.2 Another possible interpretation: a family of optimal stopping problems

Our aim is to make the optimal switching problem of the previous subsection more prone to be solved by an algorithmic procedure using Hamilton-Jacobi type problems. We then introduce here a possible formulation as a family of time-dependent optimal stopping subproblems, one per every switching variable p, suitably coupled by the stopping costs. For example, suppose $N = 3$ and take p such that $\sum_i p^i = N - 1 = 2$ (i.e., from p we can switch only to $\bar{p}$). Then, for a (x, t, p), the controller has only to choose $u = (\alpha \in \mathcal{A}, \tau \in [t, T])$ and minimize the cost

$$J_p(x, t, \alpha, \tau) = \int_t^\tau e^{-\lambda(s-t)} \ell(y(s), \alpha(s), p, s) ds + e^{-\lambda(\tau - t)} C(y(\tau), p, \bar{p}). \tag{1.1.4}$$

Note that in this representation p is fixed, that is does not change in the time interval $[t, \tau]$. Hence (1.1.4) gives a time-dependent optimal stopping problem in the state space $\mathbb{R}^d$, whose value function is

$$V_p(x, t) = \inf_{(\alpha,\tau)} J_p(x, t, \alpha, \tau).$$

Now, take p such that $\sum_i p^i = N-2 = 1$. Then consider the time-dependent optimal stopping problem in the state space $\mathbb{R}^d$ where, for a given (x, t), the control is $u = (\alpha \in \mathcal{A}, \tau \in [t, T], p' \in \mathcal{I}_p)$ and the cost to be minimized is

$$\begin{aligned} J_p(x, t, \alpha, \tau, p') = &\int_t^\tau e^{-\lambda(s-t)} \ell(y(s), \alpha(s), p, s) ds \\ &+ e^{-\lambda(\tau-t)} \Big(C(y(\tau), p, p') + V_{p'}(y(\tau), \tau) \Big). \end{aligned} \tag{1.1.5}$$

Note that from p' we can only switch to the final state $\bar{p}$, and hence $V_{p'}$ can be a priori evaluated as in the previous step. Since when $p = \bar{p}$, the game stops, we set $V_{\bar{p}} \equiv 0$. Hence (1.1.4) can be seen formulated as (1.1.5). The value function is then

$$V_p(x, t) = \inf_{(\alpha,\tau,p')} J_p(x, t, \alpha, \tau, p'). \tag{1.1.6}$$

Proceeding backwardly in this way, we consider a suitable time-dependent optimal stopping problem in $\mathbb{R}^d$ for any $p \in \mathcal{I}$, and we can at least formally compute the corresponding value functions V_p.

We will see in §1.1.4 the equivalence between the optimal control problem formulated in §1.1.1 and the family of optimal stopping problems here formulated.

1.1.3 Time-dependent optimal stopping problem: position and theoretical results

In this section, we collect some theoretical results for a time-dependent optimal stopping problem with a fixed finite horizon $T > 0$. We suitably generalize to our finite horizon time-dependent model the results in [9] for an optimal stopping problem with no time-dependence and infinite horizon feature. In order to prove the uniqueness of the solution of the related Hamilton-Jacobi equation, we need a key result, that is Lemma 1.1.1, which is proved in Appendix A, §A.1. In the same section, we also state and prove several preliminary results. Moreover here, and in the following, we use the notion of *viscosity solution*. For the definition and the first properties, we refer to Appendix B, §B.1.

Let us consider the dynamical system

$$\begin{cases} y'(s) = f(y(s), \alpha(s)), & s \in]t, T] \\ y(t) = x \end{cases}, \tag{1.1.7}$$

where $x \in \mathbb{R}^d$, $t \in [0, T]$ and

- $\alpha \in \mathcal{A} := \{\alpha : [0, +\infty[\longrightarrow A : \alpha \text{ is measurable}\}$, $A \subset \mathbb{R}^m$ compact,
- $f : \mathbb{R}^d \times A \longrightarrow \mathbb{R}^d$ is continuous, bounded and Lipschitz w.r.t. x uniformly w.r.t. a, that is there exists $L > 0$ such that

$$\|f(x, a) - f(y, a)\| \leq L\|x - y\| \quad \text{for all } x, y \in \mathbb{R}^d,\ a \in A. \tag{1.1.8}$$

We recall the following basic estimates on the trajectory $y_{(x,t)}(\cdot; \alpha)$:

- for all $x \in \mathbb{R}^d$, $\alpha \in \mathcal{A}$ and $s \in [t, T]$,

$$\|y_{(x,t)}(s; \alpha) - x\| \leq M(s - t), \tag{1.1.9}$$

  where $M := \sup\{\|f(z, a)\| : (z, a) \in \mathbb{R}^d \times A\}$;
- for all $x, z \in \mathbb{R}^d$, $\alpha \in \mathcal{A}$, $t, \tau \in [0, T]$ and $s \in [\max(t, \tau), T]$,

$$\|y_{(x,t)}(s; \alpha) - y_{(z,\tau)}(s; \alpha)\| \leq e^{L(T - \max(t,\tau))}(\|x - z\| + M|t - \tau|). \tag{1.1.10}$$

The cost to be minimized is

$$J(x, t, \alpha, \tau) = \int_t^\tau e^{-\lambda(s-t)} \ell(y_{(x,t)}(s; \alpha), \alpha(s), s) ds + e^{-\lambda(\tau - t)} \psi(y_{(x,t)}(\tau; \alpha), \tau),$$

where $\tau \leq T$ is the stopping time and $\lambda \geq 0$ the discount factor. We assume that

- $\psi : \mathbb{R}^d \times [0,T] \longrightarrow [0,+\infty[$ is bounded and uniformly continuous;
- $\ell : \mathbb{R}^d \times A \times [0,T] \longrightarrow [0,+\infty[$ is bounded, continuous and such that there exists a modulus of continuity ω_ℓ for which $|\ell(x,a,t)-\ell(y,a,t)| \leq \omega_\ell(\|x-y\|)$ for every $x, y \in \mathbb{R}^d$, $a \in A$ and $t \in [0,T]$.

The value function is

$$V(x,t) = \inf_{(\alpha\in\mathcal{A},\tau\geq t)} J(x,t,\alpha,\tau). \tag{1.1.11}$$

In the following, K, G and ω_ψ are respectively the bounds for ℓ, ψ and the modulus of continuity of ψ.

Proposition 1.1.1. *Under the previous hypotheses, V as in* (1.1.11) *is in* $\mathrm{BUC}(\mathbb{R}^d \times [0,T])$.

Proof. Fix $x, z \in \mathbb{R}^d$, $t, \tilde{t} \in [0,T]$, $\varepsilon > 0$ and $\tilde{\alpha} \in \mathcal{A}$, $\tilde{\tau} \geq \tilde{t}$ such that

$$V(z,\tilde{t}) \geq \int_{\tilde{t}}^{\tilde{\tau}} e^{-\lambda(s-\tilde{t})}\ell(y_{(z,\tilde{t})}(s;\tilde{\alpha}),\tilde{\alpha}(s),s)ds + e^{-\lambda(\tilde{\tau}-\tilde{t})}\psi(y_{(z,\tilde{t})}(\tilde{\tau};\tilde{\alpha}),\tilde{\tau}) - \varepsilon.$$

Then, recalling (1.1.10) too, we have

$$\begin{aligned}
|V(x,t) - V(z,\tilde{t})| &\leq |J(x,t,\tilde{\alpha},\tilde{\tau}) - J(z,\tilde{t},\tilde{\alpha},\tilde{\tau}) + \varepsilon| \\
&= \Bigg| \int_t^{\tilde{\tau}} e^{-\lambda(s-t)}\ell(y_{(x,t)}(s;\tilde{\alpha}),\tilde{\alpha}(s),s)ds + e^{-\lambda(\tilde{\tau}-t)}\psi(y_{(x,t)}(\tilde{\tau};\tilde{\alpha}),\tilde{\tau}) \\
&\quad - \int_{\tilde{t}}^{\tilde{\tau}} e^{-\lambda(s-\tilde{t})}\ell(y_{(z,\tilde{t})}(s;\tilde{\alpha}),\tilde{\alpha}(s),s)ds - e^{-\lambda(\tilde{\tau}-\tilde{t})}\psi(y_{(z,\tilde{t})}(\tilde{\tau};\tilde{\alpha}),\tilde{\tau}) + \varepsilon \Bigg| \\
&\leq K(1+T)|t-\tilde{t}| + \int_{\max(t,\tilde{t})}^{\tilde{\tau}} \omega_\ell(|y_{(x,t)}(s;\tilde{\alpha}) - y_{(z,\tilde{t})}(s;\tilde{\alpha})|)ds + G|t-\tilde{t}| \\
&\quad + \omega_\psi(|y_{(x,t)}(\tilde{\tau};\tilde{\alpha}) - y_{(z,\tilde{t})}(\tilde{\tau};\tilde{\alpha})|) + \varepsilon \leq K(1+T)|\tilde{t}-t| \\
&\quad + T\omega_\ell\left(e^{L(T-\max(t,\tilde{t}))}(\|x-z\| + M|t-\tilde{t}|)\right) + G|t-\tilde{t}| \\
&\quad + \omega_\psi\left(e^{L(T-\max(t,\tilde{t}))}(\|x-z\| + M|t-\tilde{t}|)\right) + \varepsilon.
\end{aligned}$$

By the arbitrariness of ε, we get the uniform continuity of V. The boundedness follows from the ones of ℓ and ψ. □

We have the following Dynamic Programming Principle.

Proposition 1.1.2. *Assume the hypotheses of Proposition 1.1.1. For every $x \in \mathbb{R}^d$ and $t \in [0,T]$, we have* *(i)* $V(x,t) \le \psi(x,t)$;

(ii) *for every* $\tilde t \ge t$, $\alpha \in \mathcal{A}$,

$$V(x,t) \le \int_t^{\tilde t} e^{-\lambda(s-t)} \ell(y_{(x,t)}(s;\alpha),\alpha(s),s)ds + e^{-\lambda(\tilde t - t)} V(y_{(x,t)}(\tilde t;\alpha),\tilde t);$$

(iii) *for any* (x,t) *for which the strict inequality in* (i) *holds, there exists* $t_0 = t_0(x,t) > 0$ *such that, for every* $\zeta \in [t, t+t_0]$,

$$V(x,t) = \inf_{\alpha\in\mathcal{A}} \left(\int_t^{\zeta} e^{-\lambda(s-t)} \ell(y_{(x,t)}(s;\alpha),\alpha(s),s)ds \right.$$
$$\left. + e^{-\lambda(\zeta-t)} V(y_{(x,t)}(\zeta;\alpha),\zeta) \right).$$

Proof. Inequality (i) is clear since, in particular,

$$V(x,t) \le J(x,t,\alpha,t) = \psi(x,t) \quad \text{for all } (x,t) \in \mathbb{R}^d \times [0,T].$$

For (ii), fix $\alpha \in \mathcal{A}$, $\tilde t \ge t$, $\varepsilon > 0$ and let $(\tilde\alpha, \tilde\tau \ge \tilde t)$ be ε-optimum for $V(y_{(x,t)}(\tilde t;\alpha),\tilde t)$, that is

$$J(y_{(x,t)}(\tilde t;\alpha),\tilde t,\tilde\alpha,\tilde\tau) \le V(y_{(x,t)}(\tilde t;\alpha),\tilde t) + \varepsilon.$$

Now define

$$\hat\alpha(\tau) = \begin{cases} \alpha(\tau), & \tau \le \tilde t \\ \tilde\alpha(\tau - \tilde t), & \tau > \tilde t \end{cases}.$$

Observe that, calling $z := y_{(x,t)}(\tilde t;\alpha)$, we have

$$\begin{aligned} V(x,t) \le J(x,t,\hat\alpha,\tilde\tau) &= \int_t^{\tilde\tau} e^{-\lambda(s-t)} \ell(y_{(x,t)}(s;\hat\alpha),\hat\alpha(s),s)ds \\ + e^{-\lambda(\tilde\tau-t)} \psi(y_{(x,t)}(\tilde\tau;\hat\alpha),\tilde\tau) &= \int_t^{\tilde t} e^{-\lambda(s-t)} \ell(y_{(x,t)}(s;\alpha),\alpha(s),s)ds \\ + \int_{\tilde t}^{\tilde\tau} e^{-\lambda(s-t)} \ell(y_{(z,\tilde t)}(s;\tilde\alpha),\tilde\alpha(s),s)ds &+ e^{-\lambda(\tilde\tau-t)} \psi(y_{(z,\tilde t)}(\tilde\tau;\tilde\alpha),\tilde\tau) \\ &= \int_t^{\tilde t} e^{-\lambda(s-t)} \ell(y_{(x,t)}(s;\alpha),\alpha(s),s)ds \\ &+ e^{-\lambda(\tilde t-t)} \int_{\tilde t}^{\tilde\tau} e^{-\lambda(s-\tilde t)} \ell(y_{(z,\tilde t)}(s;\tilde\alpha),\tilde\alpha(s),s)ds \end{aligned}$$

$$
\begin{aligned}
&+ e^{-\lambda(\tilde\tau-\tilde t)}e^{-\lambda(\tilde t-t)}\psi(y_{(z,\tilde t)}(\tilde\tau;\tilde\alpha),\tilde\tau) = \int_t^{\tilde t} e^{-\lambda(s-t)}\ell(y_{(x,t)}(s;\alpha),\alpha(s),s)ds \\
&+ e^{-\lambda(\tilde t-t)}\left(\int_{\tilde t}^{\tilde\tau} e^{-\lambda(s-\tilde t)}\ell(y_{(z,\tilde t)}(s;\tilde\alpha),\tilde\alpha(s),s)ds + e^{-\lambda(\tilde\tau-\tilde t)}\psi(y_{(z,\tilde t)}(\tilde\tau;\tilde\alpha),\tilde\tau)\right) \\
&= \int_t^{\tilde t} e^{-\lambda(s-t)}\ell(y_{(x,t)}(s;\alpha),\alpha(s),s)ds + e^{-\lambda(\tilde t-t)}J(z,\tilde t,\tilde\alpha,\tilde\tau) \\
&\le \int_t^{\tilde t} e^{-\lambda(s-t)}\ell(y_{(x,t)}(s;\alpha),\alpha(s),s)ds + e^{-\lambda(\tilde t-t)}\Big(V(y_{(x,t)}(\tilde t;\alpha),\tilde t)+\varepsilon\Big).
\end{aligned}
$$

Then, from the arbitrariness of ε, the inequality follows.

Now let us prove (iii). Suppose that $(x,t)\in\mathbb{R}^d\times[0,T]$ is such that

$$V(x,t)<\psi(x,t) \tag{1.1.12}$$

and let $\{(\alpha_n,t_n)\}\subset\mathcal{A}\times[0,T]$ be a minimizing sequence, that is

$$\lim_{n\to\infty} J(x,t,\alpha_n,t_n)=V(x,t). \tag{1.1.13}$$

We claim that there exists $t_0>t$ such that

$$t_n\ge t+t_0>t \tag{1.1.14}$$

for n sufficiently large. To see this, set $\delta_n := J(x,t,\alpha_n,t_n)-V(x,t)$. Then

$$
\begin{aligned}
V(x,t)+\delta_n \ge -C\int_t^{t_n} e^{-\lambda(s-t)}ds \\
+ e^{-\lambda(t_n-t)}\Big(\psi(x,t)-\omega_\psi\Big(\|y_{(x,t)}(t_n;\alpha_n)-x\|+|t_n-t|\Big)\Big).
\end{aligned}
$$

If for some subsequence $t_n\to t$, the previous inequality would imply that

$$V(x,t)\ge\psi(x,t),$$

a contradiction with (1.1.12). Then (1.1.14) holds. Now observe that, for $\tilde t\in[t,t+t_0]$,

$$
\begin{aligned}
J(x,t,\alpha_n,t_n) = \int_t^{\tilde t} e^{-\lambda(s-t)}\ell(y_{(x,t)}(s;\alpha_n),\alpha_n(s),s)ds \\
+ e^{-\lambda(\tilde t-t)}J(y_{(x,t)}(\tilde t;\alpha_n),\tilde t,\alpha_n,t_n).
\end{aligned}
$$

Since $t_n\ge\tilde t$ from (1.1.14), by definition of V it follows that

$$
\begin{aligned}
J(x,t,\alpha_n,t_n) &\geq \int_t^{\tilde t} e^{-\lambda(s-t)}\ell(y_{(x,t)}(s;\alpha_n),\alpha_n(s),s)ds \\
&\qquad\qquad + e^{-\lambda(\tilde t-t)}V(y_{(x,t)}(\tilde t;\alpha_n),\tilde t) \\
&\geq \inf_{\alpha\in\mathcal{A}}\left(\int_t^{\tilde t} e^{-\lambda(s-t)}\ell(y_{(x,t)}(s;\alpha),\alpha(s),s)ds + e^{-\lambda(\tilde t-t)}V(y_{(x,t)}(\tilde t;\alpha),\tilde t)\right)
\end{aligned}
$$

for any $\tilde t \in [t, t+t_0]$. Letting $n\to\infty$, from (1.1.13) we get

$$
\begin{aligned}
V(x,t) \geq \inf_{\alpha\in\mathcal{A}}\Bigg(&\int_t^{\tilde t} e^{-\lambda(s-t)}\ell(y_{(x,t)}(s;\alpha),\alpha(s),s)ds \\
&+ e^{-\lambda(\tilde t-t)}V(y_{(x,t)}(\tilde t;\alpha),\tilde t)\Bigg) \quad \text{for any } \tilde t\in[t,t+t_0].
\end{aligned}
$$

From this inequality and (ii), statement (iii) follows. □

For $x,\xi\in\mathbb{R}^d$ and $t\in[0,T]$, we define the Hamiltonian function by

$$
H(x,t,\xi) = \sup_{a\in A}\{-f(x,a)\cdot\xi - \ell(x,a,t)\}.
$$

Theorem 1.1.1. *Under the hypotheses of Proposition 1.1.2, the value function V is a viscosity solution of*

$$
\begin{cases}
\max\{u(x,t)-\psi(x,t), -u_t(x,t)+\lambda u(x,t)+H(x,t,D_xu(x,t))\} = 0, \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (x,t)\in\mathbb{R}^d\times[0,T[\\
u(x,T) = \psi(x,T), \qquad\qquad\qquad\qquad\qquad x\in\mathbb{R}^d
\end{cases}
. \tag{1.1.15}
$$

Proof. Let $(x_1,t_1)\in\mathbb{R}^d\times[0,T[$ be a local maximum point of $V-\varphi$, $\varphi\in C^1(\mathbb{R}^d\times[0,T])$. Then, for some $r>0$,

$$
V(x_1,t_1) - V(z,t) \geq \varphi(x_1,t_1) - \varphi(z,t)
$$

for every $(z,t)\in B((x_1,t_1),r)$. Fix an arbitrary $a\in A$ and let $y_{(x_1,t_1)}(\cdot)$ be the solution corresponding to the constant control $\alpha(\zeta)=a$ for all ζ. For ζ sufficiently close to t_1, $(y_{(x_1,t_1)}(\zeta),\zeta)\in B((x_1,t_1),r)$ by (1.1.9), and then

$$
\varphi(x_1,t_1) - \varphi(y_{(x_1,t_1)}(\zeta),\zeta) \leq V(x_1,t_1) - V(y_{(x_1,t_1)}(\zeta),\zeta).
$$

By (ii) of Proposition 1.1.2, we obtain

$$
\begin{aligned}
\varphi(x_1,t_1) - \varphi(y_{(x_1,t_1)}(\zeta),\zeta) &\le \int_{t_1}^{\zeta} e^{-\lambda(s-t_1)}\ell(y_{(x_1,t_1)}(s),\alpha(s),s)ds \\
&\quad + e^{-\lambda(\zeta-t_1)}V(y_{(x_1,t_1)}(\zeta),\zeta) - V(y_{(x_1,t_1)}(\zeta),\zeta) \\
&= \int_{t_1}^{\zeta} e^{-\lambda(s-t_1)}\ell(y_{(x_1,t_1)}(s),\alpha(s),s)ds + (e^{-\lambda(\zeta-t_1)}-1)V(y_{(x_1,t_1)}(\zeta),\zeta).
\end{aligned}
$$

Dividing now by $\zeta - t_1$ and letting $\zeta \to t_1$, by the differentiability of φ w.r.t. x and t we get

$$
-\varphi_t(x_1,t_1) - D_x\varphi(x_1,t_1)\cdot f(x_1,a) \le \ell(x_1,a,t_1) - \lambda V(x_1,t_1).
$$

Since $V(x,t) \le \psi(x,t)$ for every $(x,t) \in \mathbb{R}^d \times [0,T]$ by (i) of Proposition 1.1.2 and $a \in A$ is arbitrary, the subsolution condition follows.

Next suppose that $(x_2,t_2) \in \mathbb{R}^d \times [0,T[$ is a local minimum point of $V-\varphi$, $\varphi \in C^1(\mathbb{R}^d \times [0,T])$, that is, for some $r>0$,

$$
V(x_2,t_2) - V(z,t) \le \varphi(x_2,t_2) - \varphi(z,t) \tag{1.1.16}
$$

for every $(z,t) \in B((x_2,t_2),r)$. If $V(x_2,t_2) = \psi(x_2,t_2)$, then, obviously,

$$
\begin{aligned}
\max\{&V(x_2,t_2) - \psi(x_2,t_2), \\
&- V_t(x_2,t_2) + \lambda V(x_2,t_2) + H(x_2,t_2,D_xV(x_2,t_2))\} \\
&\qquad\qquad\qquad \ge V(x_2,t_2) - \psi(x_2,t_2) = 0
\end{aligned}
$$

and V is a supersolution of (1.1.15). Assume then $V(x_2,t_2) < \psi(x_2,t_2)$ (the only other possibility by (i) of Proposition 1.1.2). For each $\varepsilon > 0$ and $\zeta \ge t_2$, by (iii) of Proposition 1.1.2 there exists $\bar{\alpha} \in \mathcal{A}$ such that

$$
\begin{aligned}
V(x_2,t_2) \ge &\int_{t_2}^{\zeta} e^{-\lambda(s-t_2)}\ell(\bar{y}_{(x_2,t_2)}(s),\bar{\alpha}(s),s)ds \\
&+ e^{-\lambda(\zeta-t_2)}V(\bar{y}_{(x_2,t_2)}(\zeta),\zeta) - (\zeta - t_2)\varepsilon,
\end{aligned} \tag{1.1.17}
$$

where $\bar{y}_{(x_2,t_2)}(s) = y_{(x_2,t_2)}(s;\bar{\alpha})$ is the trajectory of (1.1.7) corresponding to $\bar{\alpha}$. Now, by the hypotheses on ℓ and by (1.1.9), we have

$$
|\ell(\bar{y}_{(x_2,t_2)}(s),\bar{\alpha}(s),s) - \ell(x_2,\bar{\alpha}(s),s)| \le \omega_\ell(M(s-t_2)), \tag{1.1.18}
$$

and, by (1.1.8) and (1.1.9) again,

$$
\|f(\bar{y}_{(x_2,t_2)}(s),\bar{\alpha}(s)) - f(x_2,\bar{\alpha}(s))\| \le LM(s-t_2). \tag{1.1.19}
$$

By (1.1.18), the integral in (1.1.17) can be written as

$$\int_{t_2}^{\zeta} e^{-\lambda(s-t_2)}\ell(x_2,\bar\alpha(s),s)ds + o(\zeta-t_2) \quad \text{as } \zeta\to t_2,$$

where $o(\zeta-t_2)$ indicates a function $g(\zeta-t_2)$ such that $\lim_{\zeta\to t_2} g(\zeta-t_2)/(\zeta-t_2)=0$ and, in this case, $|g(\zeta-t_2)|\le(\zeta-t_2)\omega_\ell(M(\zeta-t_2))$. Then, by (1.1.16) with $(z,t)=(\bar y_{(x_2,t_2)}(\zeta),\zeta)$ and by (1.1.17), we obtain

$$\begin{aligned}\varphi(x_2,t_2)-\varphi(\bar y_{(x_2,t_2)}(\zeta),\zeta)-\int_{t_2}^{\zeta} e^{-\lambda(s-t_2)}\ell(x_2,\bar\alpha(s),s)ds&\\ +(1-e^{-\lambda(\zeta-t_2)})V(\bar y_{(x_2,t_2)}(\zeta),\zeta)\ge-(\zeta-t_2)\varepsilon+o(\zeta-t_2).&\end{aligned} \tag{1.1.20}$$

Moreover, by (1.1.9), (1.1.19) and the fact that $\varphi\in C^1$, we have

$$\begin{aligned}\varphi(x_2,t_2)-\varphi(\bar y_{(x_2,t_2)}(\zeta),\zeta)&=-\int_{t_2}^{\zeta}\frac{d}{ds}\varphi(\bar y_{(x_2,t_2)}(s),s)ds\\ &=-\int_{t_2}^{\zeta}(D_x\varphi(\bar y_{(x_2,t_2)}(s),s)\cdot f(\bar y_{(x_2,t_2)}(s),\bar\alpha(s))+\varphi_t(\bar y_{(x_2,t_2)}(s),s))ds\\ &=-\int_{t_2}^{\zeta}(D_x\varphi(x_2,s)\cdot f(x_2,\bar\alpha(s))+\varphi_t(x_2,s))ds+o(\zeta-t_2).\end{aligned} \tag{1.1.21}$$

Plugging (1.1.21) into (1.1.20) and adding $\pm\int_{t_2}^{\zeta}\ell(x_2,\bar\alpha(s),s)ds$, we get

$$\begin{aligned}&\int_{t_2}^{\zeta}\{-D_x\varphi(x_2,s)\cdot f(x_2,\bar\alpha(s))-\varphi_t(x_2,s)-\ell(x_2,\bar\alpha(s),s)\}ds\\ &\quad+\int_{t_2}^{\zeta}(1-e^{-\lambda(s-t_2)})\ell(x_2,\bar\alpha(s),s)ds+(1-e^{-\lambda(\zeta-t_2)})V(\bar y_{(x_2,t_2)}(\zeta),\zeta)\\ &\qquad\ge-(\zeta-t_2)\varepsilon+o(\zeta-t_2).\end{aligned} \tag{1.1.22}$$

The first integral is estimated from above by

$$\int_{t_2}^{\zeta}\sup_{a\in A}\{-D_x\varphi(x_2,s)\cdot f(x_2,a)-\varphi_t(x_2,s)-\ell(x_2,a,s)\}ds$$

and the second one is $o(\zeta-t_2)$ by the hypotheses on ℓ. Dividing (1.1.22) by $\zeta-t_2$ and letting $\zeta\to t_2$, we obtain

$$-\varphi_t(x_2,t_2)+\sup_{a\in A}\{-D_x\varphi(x_2,t_2)\cdot f(x_2,a)-\ell(x_2,a,t_2)\}+\lambda V(x_2,t_2)\ge-\varepsilon,$$

where we also used the continuity of V and $\bar y_{(x_2,t_2)}$ at (x_2,t_2) and t_2 respectively. Since ε is arbitrary, the supersolution condition follows. $\square$

For the uniqueness, we show that if u is a viscosity solution of (1.1.15), then

$$u(x,t) = \inf_{\alpha\in\mathcal{A}} J(x,t,\alpha,\tau^*_{(x,t)}(\alpha))$$

for some $\tau^*_{(x,t)}(\alpha)$ such that

$$\inf_{\alpha\in\mathcal{A}} J(x,t,\alpha,\tau^*_{(x,t)}(\alpha)) = \inf_{(\alpha\in\mathcal{A},\tau\geq t)} J(x,t,\alpha,\tau) = V(x,t),$$

and hence V is the unique viscosity solution. We need at first the following Lemma, which is proved in Appendix A, §A.1.

Lemma 1.1.1. *Let $\Omega \subseteq \mathbb{R}^d$ be an open subset. For $(x,t) \in \Omega\times[0,T]$ and $\alpha\in\mathcal{A}$, we set*

$$\tau_{(x,t)}(\alpha) := \min\{\inf\{\tau \geq t : y_{(x,t)}(\tau;\alpha) \notin \Omega\}, T\}.$$

Then, under the hypotheses of Theorem 1.1.1, for $u \in \mathrm{BUC}(\Omega\times[0,T])$ the following statements are equivalent:

(i) for all $x\in\Omega$, $\alpha\in\mathcal{A}$ and $t \leq s \leq \tau < \tau_{(x,t)}(\alpha)$,

$$e^{-\lambda(s-t)}u(y_{(x,t)}(s;\alpha),s) - e^{-\lambda(\tau-t)}u(y_{(x,t)}(\tau;\alpha),\tau) \leq \int_s^\tau e^{-\lambda(\zeta-t)}\ell(y_{(x,t)}(\zeta;\alpha),\alpha(\zeta),\zeta)d\zeta;$$

(ii) $u_t(x,t) - \lambda u(x,t) - H(x,t,D_x u(x,t)) \geq 0$, $(x,t)\in\Omega\times[0,T[$, in the viscosity sense;

(iii) $-u_t(x,t) + \lambda u(x,t) + H(x,t,D_x u(x,t)) \leq 0$, $(x,t)\in\Omega\times[0,T[$, in the viscosity sense.

Proof. See the proof of Lemma A.1.7 in Appendix A, §A.1. □

Theorem 1.1.2. *Let $u \in \mathrm{BUC}(\mathbb{R}^d\times[0,T])$ be a viscosity solution of* (1.1.15). *Then, under the hypotheses of Theorem 1.1.1,*

$$u(x,t) = \inf_{\alpha\in\mathcal{A}} J(x,t,\alpha,\tau^*_{(x,t)}(\alpha)) = V(x,t)$$

for every $(x,t)\in\mathbb{R}^d\times[0,T]$, where

$$\tau^* := \tau^*_{(x,t)}(\alpha) = \inf\{\tau\in[t,T] : u(y_{(x,t)}(\tau;\alpha),\tau) = \psi(y_{(x,t)}(\tau;\alpha),\tau)\}. \tag{1.1.23}$$

Proof. At first we observe that, since $u(x,T)=\psi(x,T)$ for every $x\in\mathbb{R}^d$, the set in (1.1.23) is always non-empty, and hence $\tau^*\leq T$ always exists. Now let $u\in \mathrm{BUC}(\mathbb{R}^d\times[0,T])$ be a viscosity solution of (1.1.15) and consider the open set $\mathcal{C}=\{(x,t)\in\mathbb{R}^d\times[0,T[: u(x,t)<\psi(x,t)\}$. At first we prove that

$$u(x,t)\leq\psi(x,t),\quad (x,t)\in\mathbb{R}^d\times[0,T] \tag{1.1.24}$$

and that

$$-u_t(x,t)+\lambda u(x,t)+H(x,t,D_xu(x,t))\leq 0,\ (x,t)\in\mathbb{R}^d\times[0,T[, \tag{1.1.25}$$

$$-u_t(x,t)+\lambda u(x,t)+H(x,t,D_xu(x,t))=0,\quad (x,t)\in\mathcal{C}, \tag{1.1.26}$$

in the viscosity sense. By contradiction, suppose that $u(x_0,t_0)>\psi(x_0,t_0)$ at some $(x_0,t_0)\in\mathbb{R}^d\times[0,T]$. Then, by continuity,

$$u(x,t)>\psi(x,t)\quad\text{for every } (x,t)\in B((x_0,t_0),\delta),\ \delta>0. \tag{1.1.27}$$

It can be easily proved that $u-\varphi$ has a local maximum at some point $(\bar{x},\bar{t})\in B((x_0,t_0),\delta)$ for some $\varphi\in C^1(\mathbb{R}^d\times[0,T])$, so that, since u is a viscosity solution of (1.1.15),

$$\max\{u(\bar{x},\bar{t})-\psi(\bar{x},\bar{t}),-u_t(\bar{x},\bar{t})+\lambda u(\bar{x},\bar{t})+H(\bar{x},\bar{t},D_x\varphi(\bar{x},\bar{t}))\}\leq 0.$$

This contradicts (1.1.27) and hence (1.1.24) holds.

By (1.1.24), the inequality (1.1.25) immediately follows since u is a viscosity solution of (1.1.15).

To prove (1.1.26), it is sufficient to show that

$$-u_t(x,t)+\lambda u(x,t)+H(x,t,D_xu(x,t))\geq 0,\quad (x,t)\in\mathcal{C}, \tag{1.1.28}$$

in the viscosity sense. At any local minimum $(x_1,t_1)\in\mathcal{C}$ of $u-\varphi$, $\varphi\in C^1(\mathbb{R}^d\times[0,T])$, by (1.1.15) we have

$$\begin{aligned}\max\{&u(x_1,t_1)-\psi(x_1,t_1),\\ &-u_t(x_1,t_1)+\lambda u(x_1,t_1)+H(x_1,t_1,D_x\varphi(x_1,t_1))\}\geq 0,\end{aligned}$$

and (1.1.28) is proved since $(x_1,t_1)\in\mathcal{C}$.

Now we apply Lemma 1.1.1 with $\Omega=\mathbb{R}^d$, $s=t$ and, by (1.1.24), (1.1.25), we get

$$\begin{aligned}u(x,t) &\le e^{-\lambda(\tau-t)}u(y_{(x,t)}(\tau;\alpha),\tau) + \int_t^\tau e^{-\lambda(\zeta-t)}\ell(y_{(x,t)}(\zeta;\alpha),\alpha(\zeta),\zeta)d\zeta \\ &\le e^{-\lambda(\tau-t)}\psi(y_{(x,t)}(\tau;\alpha),\tau) + \int_t^\tau e^{-\lambda(\zeta-t)}\ell(y_{(x,t)}(\zeta;\alpha),\alpha(\zeta),\zeta)d\zeta\end{aligned}$$

for all $t \le \tau \le T$ and $\alpha \in \mathcal{A}$. Then

$$u(x,t) \le \inf_{(\alpha\in\mathcal{A},\tau\ge t)} J(x,t,\alpha,\tau) = V(x,t).$$

For the reverse inequality, assume at first $(x,t) \notin \mathcal{C}$. In this case, $u(x,t) = \psi(x,t)$ and $\tau^* = t$. Then

$$u(x,t) = \psi(x,t) = J(x,t,\alpha,\tau^*) \ge \inf_{(\alpha\in\mathcal{A},\tau\ge t)} J(x,t,\alpha,\tau) = V(x,t).$$

Now suppose $(x,t) \in \mathcal{C}$, so that (1.1.26) holds. Applying Lemma 1.1.1 with $\Omega = \mathcal{C}$ and $s = t$, we obtain

$$\begin{aligned}u(x,t) = \inf_{\alpha\in\mathcal{A}} \Bigg(& e^{-\lambda(\tau-t)}u(y_{(x,t)}(\tau;\alpha),\tau) \\ & + \int_t^\tau e^{-\lambda(\zeta-t)}\ell(y_{(x,t)}(\zeta;\alpha),\alpha(\zeta),\zeta)d\zeta\Bigg)\end{aligned}$$

for every $t \le \tau < \tau^* \le T$. Letting $\tau \to \tau^*$, we get

$$\begin{aligned}u(x,t) = \inf_{\alpha\in\mathcal{A}} J(x,t,\alpha,\tau^*) &\ge \inf_{(\alpha\in\mathcal{A},\tau\in[t,\tau^*])} J(x,t,\alpha,\tau) \\ &\ge \inf_{\alpha\in\mathcal{A},\tau\in[t,T])} J(x,t,\alpha,\tau) = V(x,t)\end{aligned}$$

since $u(y_{(x,t)}(\tau^*;\alpha),\tau^*) = \psi(y_{(x,t)}(\tau^*;\alpha),\tau^*)$. □

1.1.4 Equivalence of the two formulations

In this section, we show the equivalence between the optimal switching problem and the family of the optimal stopping ones, i.e., $V(x,t,p) = V_p(x,t)$ for every $(x,t,p) \in \mathbb{R}^d \times [0,T] \times \mathcal{I}$. Here, and in the sequel, V is the value function defined in (1.1.3) and V_p is the value function defined backwardly as in (1.1.6).

Proposition 1.1.3. *Under the hypotheses in §1.1.1, we have* *(i)* $V \in \mathrm{BUC}(\mathbb{R}^d \times [0,T])$ *for every* $p \in \mathcal{I}$*;*

(ii) *for every* p*, the value functions* V_p *are bounded and uniformly continuous too.*

Proof. The proof of (i) goes as the one in Proposition 1.1.1 in §1.1.3. For (ii), by the backward definition of V_p, as in §1.1.2, note that at the levels p with $\sum_i p^i = N-1$, the stopping cost is just C and hence does not depend on the value function $V_{p'}$ at lower levels $p' \in \mathcal{I}_p$. For higher levels p such that $\sum_i p^i < N-1$, let us define

$$\psi_p(x,t) := \inf_{p' \in \mathcal{I}_p} (C(x,p,p') + V_{p'}(x,t)), \quad x \in \mathbb{R}^d, \tag{1.1.29}$$

and recalling that ℓ does not depend on $p' \in \mathcal{I}_p$, we have

$$V_p(x,t) = \inf_{(\alpha,\tau)} \left(\int_t^\tau e^{-\lambda(s-t)} \ell(y(s),\alpha(s),p,s)ds + e^{-\lambda(\tau-t)} \psi_p(y(\tau),\tau) \right).$$

By the backward definition of V_p at every level $p \in \mathcal{I}$, the stopping cost $\psi_p(x,t)$ can be assumed as known and hence, in particular, bounded and uniformly continuous. Again, the thesis comes from the results in §1.1.3. $\square$

Proposition 1.1.4. *Under the hypotheses of Proposition 1.1.3, for all* $x \in \mathbb{R}^d$*,* $t \in [0,T]$ *and* $p \in \mathcal{I}$*, we have*

$$\begin{aligned} V(x,t,p) = \inf_{(\alpha,\tau,p' \in \mathcal{I}_p)} \Bigg(\int_t^\tau & e^{-\lambda(s-t)} \ell(y(s),\alpha(s),p,s)ds \\ & + e^{-\lambda(\tau-t)} \Big(C(y(\tau),p,p') + V_{p'}(y(\tau),\tau) \Big) \Bigg). \end{aligned}$$

As a consequence, $V(x,t,p) = V_p(x,t)$ *for all* (x,t,p)*.*

Proof. We follow a procedure as the one used in §1.1.2. Consider p with $\sum_i p^i = N-1$. By definition we have

$$\begin{aligned} V(x,t,p) &= \inf_{(\alpha,\tau)} \left(\int_t^\tau e^{-\lambda(s-t)} \ell(y(s),\alpha(s),p,s)ds + e^{-\lambda(\tau-t)} C(y(\tau),p,\bar{p}) \right) \\ &= V_p(x,t) \end{aligned}$$

for every $(x,t) \in \mathbb{R}^d \times [0,T]$ since $V(\cdot,\cdot,\bar{p}) = V_{\bar{p}}(\cdot,\cdot) \equiv 0$. Consider now p with $\sum_i p^i = N-2$. We need to show that

$$V(x,t,p) = \inf_{(\alpha,\tau,p' \in \mathcal{I}_p)} \left(\int_t^\tau e^{-\lambda(s-t)} \ell(y(s),\alpha(s),p,s)ds \right.$$
$$\left. + e^{-\lambda(\tau-t)} \Big(C(y(\tau),p,p') + V_{p'}(y(\tau),\tau) \Big) \right). \quad (1.1.30)$$

We recall that, calling $x_1 := y^\alpha_{(x,t,p)}(t_1)$, we have

$$V(x,t,p) = \inf_{(\alpha, t \le t_1 \le t_2, p_1 \in \mathcal{I}_p)} \left(\int_t^{t_1} e^{-\lambda(s-t)} \ell(y^\alpha_{(x,t,p)}(s),\alpha(s),p,s)ds \right.$$
$$+ e^{-\lambda(t_1-t)} C(x_1,p,p_1) + \int_{t_1}^{t_2} e^{-\lambda(s-t)} \ell(y^\alpha_{(x,t,p_1)}(s),\alpha(s),p_1,s)ds$$
$$\left. + e^{-\lambda(t_2-t)} C(y^\alpha_{(x,t,p_1)}(t_2),p_1,\bar{p}) \right). \quad (1.1.31)$$

So we have to prove that the inf in (1.1.30) coincides with the inf in (1.1.31). At first we show the inequality $(\le)$. For every $x \in \mathbb{R}^d$, $t \in [0,T]$, $\alpha \in \mathcal{A}$ and $p_1 \in \mathcal{I}_p$, we have

$$\int_t^{t_1} e^{-\lambda(s-t)} \ell(y^\alpha_{(x,t,p)}(s),\alpha(s),p,s)ds + e^{-\lambda(t_1-t)} C(x_1,p,p_1)$$
$$+ \int_{t_1}^{t_2} e^{-\lambda(s-t)} \ell(y^\alpha_{(x,t,p_1)}(s),\alpha(s),p_1,s)ds + e^{-\lambda(t_2-t)} C(y^\alpha_{(x,t,p_1)}(t_2),p_1,\bar{p})$$
$$= \int_t^{t_1} e^{-\lambda(s-t)} \ell(y^\alpha_{(x,t,q)}(s),\alpha(s),p,s)ds + e^{-\lambda(t_1-t)} C(x_1,p,p_1)$$
$$+ e^{-\lambda(t_1-t)} \int_{t_1}^{t_2} e^{-\lambda(s-t_1)} \ell(y^\alpha_{(x_1,t_1,p_1)}(s),\alpha(s),p_1,s)ds$$
$$+ e^{-\lambda(t_1-t)} e^{-\lambda(t_2-t_1)} C(y^\alpha_{(x_1,t_1,p_1)}(t_2),p_1,\bar{p})$$
$$= \int_t^{t_1} e^{-\lambda(s-t)} \ell(y^\alpha_{(x,t,p)}(s),\alpha(s),p,s)ds + e^{-\lambda(t_1-t)} C(x_1,p,p_1)$$
$$+ e^{-\lambda(t_1-t)} \left(\int_{t_1}^{t_2} e^{-\lambda(s-t_1)} \ell(y^\alpha_{(x_1,t_1,p_1)}(s),\alpha(s),p_1,s)ds \right.$$
$$\left. + e^{-\lambda(t_2-t_1)} C(y^\alpha_{(x_1,t_1,p_1)}(t_2),p_1,\bar{p}) \right)$$

$$
\begin{aligned}
&= \int_t^{t_1} e^{-\lambda(s-t)}\ell(y^{\alpha}_{(x,t,p)}(s),\alpha(s),p,s)ds + e^{-\lambda(t_1-t)}C(x_1,p,p_1)\\
&+ e^{-\lambda(t_1-t)}J(x_1,t_1,p_1,\alpha,t_2,\bar p) \geq \int_t^{t_1} e^{-\lambda(s-t)}\ell(y^{\alpha}_{(x,t,p)}(s),\alpha(s),p,s)ds\\
&\qquad + e^{-\lambda(t_1-t)}\Big(C(x_1,p,p_1)+V(x_1,t_1,p_1)\Big)\\
&\qquad = \int_t^{t_1} e^{-\lambda(s-t)}\ell(y^{\alpha}_{(x,t,p)}(s),\alpha(s),p,s)ds\\
&\qquad\qquad + e^{-\lambda(t_1-t)}\Big(C(x_1,p,p_1)+V_{p_1}(x_1,t_1)\Big).
\end{aligned}
$$

Passing to the inf over $(\alpha, t_1, p_1 \in \mathcal{I}_p)$, we get $V(x,t,p) \geq V_p(x,t)$.

Let us prove $(\geq)$. Fix $\alpha \in \mathcal{A}$, $\tau \geq t$, $p' \in \mathcal{I}_p$ and let $(\tilde\alpha_1, \tilde t \geq \tau \geq t)$ be ε-optimum for $V_{p'}(y^{\alpha}_{(x,t,p)}(\tau),\tau)$. Then, calling $x' := y^{\alpha}_{(x,t,p)}(\tau)$, we have

$$
\begin{aligned}
&\inf_{(\alpha,\tau,p')}\Bigg(\int_t^{\tau} e^{-\lambda(s-t)}\ell(y^{\alpha}_{(x,t,p)}(s),\alpha(s),p,s)ds\\
&\qquad + e^{-\lambda(\tau-t)}\Big(C(x',p,p')+V_{p'}(x',\tau)\Big)\Bigg)\\
&\geq \int_t^{\tau} e^{-\lambda(s-t)}\ell(y^{\alpha}_{(x,t,p)}(s),\alpha(s),p,s)ds + e^{-\lambda(\tau-t)}\Bigg(C(x',p,p')\\
&+\int_\tau^{\tilde t} e^{-\lambda(s-\tau)}\ell(y^{\tilde\alpha_1}_{(x',\tau,p')}(s),\tilde\alpha_1(s),p',s)ds + e^{-\lambda(\tilde t-\tau)}C(y^{\tilde\alpha_1}_{(x',\tau,p')}(\tilde t),p',\bar p)-\varepsilon\Bigg)\\
&= \int_t^{\tau} e^{-\lambda(s-t)}\ell(y^{\alpha}_{(x,t,p)}(s),\alpha(s),p,s)ds + e^{-\lambda(\tau-t)}C(x',p,p')\\
&\qquad + e^{-\lambda(\tau-t)}\int_\tau^{\tilde t} e^{-\lambda(s-\tau)}\ell(y^{\tilde\alpha_1}_{(x',\tau,p')}(s),\tilde\alpha_1(s),p',s)ds\\
&\qquad + e^{-\lambda(\tau-t)}e^{-\lambda(\tilde t-\tau)}C(y^{\tilde\alpha_1}_{(x',\tau,p')}(\tilde t),p',\bar p) - e^{\lambda(\tau-t)}\varepsilon\\
&= \int_t^{\tau} e^{-\lambda(s-t)}\ell(y^{\alpha}_{(x,t,p)}(s),\alpha(s),p,s)ds + e^{-\lambda(\tau-t)}C(x',p,p')\\
&\qquad + \int_\tau^{\tilde t} e^{-\lambda(s-t)}\ell(y^{\tilde\alpha_1}_{(x',\tau,p')}(s),\tilde\alpha_1(s),p',s)ds\\
&\qquad\qquad + e^{-\lambda(\tilde t-t)}C(y^{\tilde\alpha_1}_{(x',\tau,p')}(\tilde t),p',\bar p) - e^{\lambda(\tau-t)}\varepsilon. \quad (1.1.32)
\end{aligned}
$$

By defining

$$
\hat\alpha(s) = \begin{cases} \alpha(s), & s \leq \tau \\ \tilde\alpha_1(s-\tau), & s > \tau \end{cases},
$$

we can write the right-hand side of (1.1.32) as

$$\int_t^{\tau} e^{-\lambda(s-t)}\ell(y^{\hat\alpha}_{(x,t,p)}(s),\hat\alpha(s),p,s)ds + e^{-\lambda(\tau-t)}C(y^{\hat\alpha}_{(x,t,p)}(\tau),p,p')$$
$$+\int_{\tau}^{\tilde t} e^{-\lambda(s-t)}\ell(y^{\hat\alpha}_{(x,t,p')}(s),\hat\alpha(s),p',s)ds + e^{-\lambda(\tilde t-t)}C(y^{\hat\alpha}_{(x,t,p')}(\tilde t),p',\bar p) - e^{\lambda(\tau-t)}\varepsilon$$
$$\geq \int_t^{\tau} e^{-\lambda(s-t)}\ell(y^{\hat\alpha}_{(x,t,p)}(s),\hat\alpha(s),p,s)ds + e^{-\lambda(\tau-t)}C(y^{\hat\alpha}_{(x,t,p)}(\tau),p,p')$$
$$+\int_{\tau}^{\tilde t} e^{-\lambda(s-t)}\ell(y^{\hat\alpha}_{(x,t,p')}(s),\hat\alpha(s),p',s)ds + e^{-\lambda(\tilde t-t)}C(y^{\hat\alpha}_{(x,t,p')}(\tilde t),p',\bar p) - \varepsilon$$
$$\geq \inf_{(t\leq t_1\leq t_2,\alpha,p_1\in\mathcal{I}_p)} \Bigg(\int_t^{t_1} e^{-\lambda(s-t)}\ell(y^{\alpha}_{(x,t,p)}(s),\alpha(s),p,s)ds$$
$$+ e^{-\lambda(t_1-t)}C(y^{\alpha}_{(x,t,p)}(t_1),p,p_1) + \int_{t_1}^{t_2} e^{-\lambda(s-t)}\ell(y^{\alpha}_{(x,t,p_1)}(s),\alpha(s),p_1,s)ds$$
$$+ e^{-\lambda(t_2-t)}C(y^{\alpha}_{(x,t,p_1)}(t_2),p_1,\bar p)\Bigg) - \varepsilon = V(x,t,p)-\varepsilon.$$

By the arbitrariness of ε we get the desired inequality.

The same arguments can be repeated for the others p such that $\sum_i p^i < N-2$. □

1.1.5 Optimality condition for V_p in PDE form

In view of the results in §1.1.4, we are able to obtain a differential characterization of the value function V_p as viscosity solution of an Hamilton-Jacobi equation. For $x,\xi\in\mathbb{R}^d$, $t\in[0,T]$ and $p\in\mathcal{I}$, we define then the Hamiltonian function by

$$H^p(x,t,\xi) = \sup_{a\in A}\{-f(x,a,p)\cdot\xi - \ell(x,a,p,t)\}.$$

Theorem 1.1.3. *Under the hypotheses of Proposition 1.1.4, for any $p\in\mathcal{I}$ the value function V_p is the unique bounded and uniformly continuous viscosity solution u of (ψ_p as in (1.1.29))*

$$\begin{cases} \max\{u(x,t)-\psi_p(x,t), -u_t(x,t)+\lambda u(x,t)+H^p(x,t,D_xu(x,t))\} = 0, \\ \hfill (x,t)\in\mathbb{R}^d\times[0,T[\\ u(x,T)=\psi_p(x,T), \hfill x\in\mathbb{R}^d \end{cases} . \tag{1.1.33}$$

Proof. See Theorem 1.1.1 and 1.1.2 in §1.1.3. $\square$

Theorem 1.1.4. *The family of functions $V := \{V_p : p \in \mathcal{I}\}$ is the unique family of bounded and uniformly continuous functions $U := \{u_p : p \in \mathcal{I}\}$ that solves the problem*

$$\begin{cases} \textit{for any } p \in \mathcal{I},\ u_p \textit{ is the unique viscosity solution of } (1.1.33) \textit{ with} \\ \psi_p \textit{ replaced by } \psi_p^U(x,t) := \inf\limits_{p' \in \mathcal{I}_p} (C(x,p,p') + u_{p'}(x,t)), \\ u_{\bar{p}} = 0 \end{cases} . \tag{1.1.34}$$

Proof. Note that $\psi_p^V = \psi_p$ as in (1.1.29). For $p \in \mathcal{I}$ such that $\sum_i p^i = N-1$, we have that the problem (1.1.34) is the same as (1.1.33) because $\psi_p = \psi_p^U$. Therefore, by Theorem 1.1.2 in §1.1.3, $V_p = u_p$. Hence, if $p \in \mathcal{I}$ is such that $\sum_i p^i = N-2$, we also have $\psi_p = \psi_p^U$ and again $u_p = V_p$. We then conclude backwardly. $\square$

Chapter 2

Towards a mean-field type optimal visiting problem

In a possible study of a mean-field game for a population of agents of density μ, each one of them playing a p-labeled optimal stopping problem like the one in Ch. 1, §1.1, we would be led to consider the coupling of the system (1.1.33) of Hamilton-Jacobi equations (coupled by the stopping costs) with a system of continuity equations (one per each level p and coupled by a transfer through some sinks and sources). In particular, the sink at level p is the region where the agents stop running at level p and pass to a new subsequent level $p' \in \mathcal{I}_p$, and similarly for the sources. Such a coupling should provide the optimal vector field $b_p(x,t)$, giving the optimal flow, and the optimal switching time-dependent sets $\mathcal{S}_p^t$ for the evolution of the masses of the agent μ_p, labeled by p. The vector field b_p and the switching sets $\mathcal{S}_p^t$ will depend on the value function $V_p(x,t)$, in particular b_p is typically $-D_x V_p$ (see also Remark 2.1.1 and Figure 2.1 for an illustrative scheme). A sketch of the motion rules of μ_p is represented in Figure 2.2.

2.1 Optimal visiting for a crowd of agents: the continuity equation

Here, from an analytical point of view, we focus only on a single continuity equation for a given suitably regular field (possibly depending on the measure), with possible sinks and sources, and we left further analysis to future studies. In particular, in §2.1.1, we investigate at first the model with just a sink, and then, in §2.1.4, the one with a source too. In [6], some numerical tests are also shown for the cases addressed here and for more general

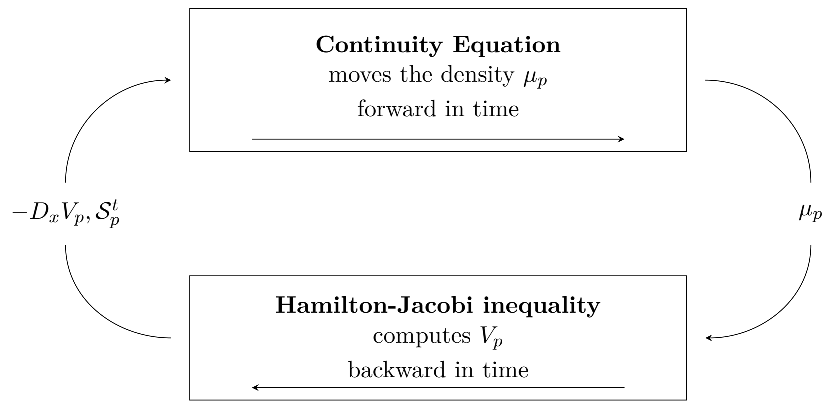


Figure 2.1: The coupling between the Hamilton-Jacobi and the continuity equation.

situations too.

2.1.1 The continuity equation with a sink

We want to model the evolution on $\mathbb{R}^d$ of a mass μ subject to a given flow with a presence of a given region of $\mathbb{R}^d$ acting as sink: the portion of mass that possibly enters the sink instantaneously disappears. The region representing the sink can be in general moving in time but here, for simplicity, we consider it as constant. Under suitable hypotheses, the generalization to the moving case works with same ideas and calculations as explained in Remark 2.1.8.

For the notation and the construction of the following setting, we mostly rely on [18]. We consider a flow $\Phi : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}^d$ given by the solutions of the ordinary differential system for $t \in \mathbb{R}$ and $x \in \mathbb{R}^d$,

$$\begin{cases} y'(s) = b(y(s), s), \quad s > t \\ y(t) = x \end{cases}, \tag{2.1.1}$$

that is $\Phi(x, t, s) = y(s)$ solving (2.1.1). We will be mostly concerned with $\Phi(\cdot, 0, \cdot)$. In (2.1.1), the field $b : \mathbb{R}^d \times \mathbb{R} \longrightarrow \mathbb{R}^d$ is assumed to be bounded, continuous and Lipschitz continuous w.r.t. $x \in \mathbb{R}^d$ uniformly w.r.t. $t \in \mathbb{R}$. Then, the flow $\Phi(\cdot, 0, \cdot)$ is Lipschitz continuous.

The sink is represented by a subset $\mathcal{S} \subset \mathbb{R}^d$, which is assumed to be closed with compact and C^1 boundary. Then, for a point $x \in \mathbb{R}^d$, we define

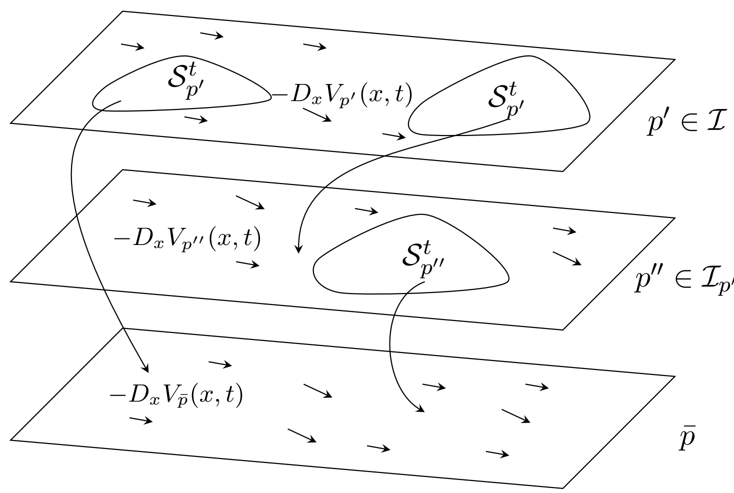


Figure 2.2: The evolution of the densities μ_p: when outside the switching sets, for a label p, the density moves accordingly to the direction $-D_x V_p$; when inside a switching set, the new label is also detected by any optimization criterium which may also depend on space and time.

the possible first arrival time (to the sink) as

$$t_x := \inf\{t \geq 0 : \Phi(x, 0, t) \in \mathcal{S}\} \quad (\inf \emptyset = +\infty), \tag{2.1.2}$$

and the set of possible arrival points to the sink, for a given t, as

$$\mathcal{S}^t := \{z \in \partial\mathcal{S} : \exists x \in \mathbb{R}^d \text{ such that } t_x = t \text{ and } \Phi(x, 0, t) = z\}.$$

We will see in §2.1.3 that it is possible to characterize the possible first arrival time t_x as the unique (viscosity) solution of an Hamilton-Jacobi equation with suitable boundary conditions.

We work in the set $\mathcal{G}$ of positive Radon measures μ on $\mathbb{R}^d$ with finite first order moment, bounded by a constant G (i.e., $\int_{\mathbb{R}^d} d\mu \leq G$ for every $\mu \in \mathcal{G}$). Such a space can be endowed with the generalized Wasserstein distance (see [35, 36]), which, for simplicity, we write in the following equivalent form:

$$\mathcal{W}(\mu, \mu') = \sup \left\{ \int_{\mathbb{R}^d} \varphi d(\mu - \mu') : \varphi \in C^0_c(\mathbb{R}^d),\ \|\varphi\|_\infty \leq 1,\ \mathrm{Lip}(\varphi) \leq 1 \right\}. \tag{2.1.3}$$

Let

$$\tilde{m}_0(x) := \begin{cases} m_0(x), & x \in \mathbb{R}^d \setminus \mathcal{S} \\ 0, & \text{otherwise} \end{cases}, \qquad \tilde{m}_0 \in \mathcal{G}, \tag{2.1.4}$$

where $m_0 \in \mathcal{G}$ is given, be the initial distribution on $\mathbb{R}^d$. We suppose that it is absolutely continuous with a density, still denoted $\tilde{m}_0$, which is bounded and has a compact support.

Then, the continuity equation with a sink (to be interpreted in a suitable weak formulation), with a finite horizon $T > 0$, is

$$\begin{cases} \mu_t(x,t) + \operatorname{div}(\mu(x,t)b(x,t)) + \mathbb{1}_{\{(\mathcal{S}^t,t):t\in[0,T]\}}\mu(x,t) = 0, \\ \hfill (x,t) \in \mathbb{R}^d \times [0,T] \\ \mu(x,0) = \tilde{m}_0(x), \hfill x \in \mathbb{R}^d \end{cases} . \tag{2.1.5}$$

Remark 2.1.1. *In the possible mean-field game problem, using the notation of previous sections in Ch. 1, at every level p the sink would be given by the evolutive stopping set $\mathcal{S}_p(t) = \{x \in \mathbb{R}^d : V_p(x,t) = \psi_p(x,t)\}$ and the field b by the gradient of the value function V_p. Hence, the regularity assumptions above should be probably adjusted. In particular, the presence of more than one target leads to possible multiplicity of the optimal control, which makes the population split into several fractions, each one of them following one of the optimal behaviors. A similar situation is studied in [2], [12] and in Ch. 3 (see also [8]). Anyway, we may expect that the value function V_p will be suitably regular along the optimal trajectories.*

In view of the mean-field case, in §2.1.2 we will study a possible dependence of the field b on the measure. In [6], some numerical experiments are performed including also this possibility.

In the sequel, we denote by Ψ the inverse of the flow Φ starting from $\partial\mathcal{S}$, i.e., all the states backwardly reached by the trajectories starting from the points of $\partial\mathcal{S}$ in the time interval $[0,T]$. That is

$$(z,\tau) \longmapsto \Psi(z,\tau,\tau), \quad 0 \le \tau \le T, \quad z \in \partial\mathcal{S}$$

with $\Psi(z,\tau,\tau) = \zeta(\tau)$ satisfying

$$\begin{cases} \zeta'(s) = -b(\zeta(s),\tau - s) = \beta_\tau(\zeta(s),s), \quad 0 < s < \tau \\ \zeta(0) = z \end{cases} . \tag{2.1.6}$$

By hypotheses, Ψ is Lipschitz continuous as Φ and it is such that

$$\begin{aligned} \Phi(\Psi(z,\tau,\tau),0,\tau) &= z, \\ \Psi(\Phi(x,0,\tau),\tau,\tau) &= x. \end{aligned} \tag{2.1.7}$$

Now, fixed $s \in [0, T]$, we define the sink-reaching-points set, at time s, as

$$\mathcal{B}(s) := \{x \in \mathbb{R}^d : t_x \le s\}, \tag{2.1.8}$$

that is the set of all initial points $x \in \mathbb{R}^d$ from which the agents enter the sink before s. Observe that

$$\mathcal{B}(s) = \bigcup_{\tau \in [0,s]} \underbrace{\{x \in \mathbb{R}^d : t_x = \tau\}}_{=:\mathcal{B}^\tau} = \bigcup_{\tau \in [0,s]} \{x = \Psi(z, \tau, \tau) : z \in \mathcal{S}^\tau\}$$

and that

$$s_1 \le s_2 \Rightarrow \mathcal{B}(s_1) \subseteq \mathcal{B}(s_2). \tag{2.1.9}$$

Definition 2.1.1. *We say that μ is a weak solution of (2.1.5) if $\mu \in L^1([0,T], \mathcal{G})$ is such that, for any test function $\varphi \in C_c^\infty(\mathbb{R}^d \times [0,T[)$, we have*

$$\begin{aligned} &\int_{\mathbb{R}^d} \varphi(x,0) d\tilde{m}_0(x) \\ &\quad + \int_0^T \int_{\mathbb{R}^d \setminus \Phi(\mathcal{B}(t),0,t)} (\varphi_t(x,t) + \langle D_x\varphi(x,t), b(x,t)\rangle) d\mu(t)(x) dt \\ &\qquad - \int_0^T \int_{\mathbb{R}^d} \mathbb{1}_{\{(\mathcal{S}^t,t):t\in[0,T]\}} \varphi(x,t) d\mu^t(t)(x) dt = 0, \end{aligned}$$

where μ^t entering the last integral is $\mu^t(t) = g(t)\mu^t(0)$. The measure $\mu^t(0)$ is the disintegration of $\mu(0)$ on the fibers $\mathcal{B}^\tau$ that compose $\mathcal{B}(t)$, and $g(\cdot)$ is the density of the measure ν on the indices τ of the fibers $\mathcal{B}^\tau$ such that

$$E \subset \mathcal{B}(t) \Rightarrow \mu(0)(E) = \int_0^t \mu^\tau(0)(\mathcal{B}^\tau \cap E) d\nu(\tau) = \int_0^t g(\tau)\mu^\tau(0)(\mathcal{B}^\tau \cap E) d\tau.$$

For $s \in [0,T]$, we define the following measure on $\mathbb{R}^d$

$$\tilde{\mu}(s) = \begin{cases} \Phi(\cdot,0,s)\sharp\tilde{m}_0 & \text{on } \mathbb{R}^d \setminus \Phi(\mathcal{B}(s),0,s) \\ 0, & \text{otherwise} \end{cases}. \tag{2.1.10}$$

Observe that the set $\Phi(\mathcal{B}(s),0,s)$ takes into account all the positions of the agents who passed through the sink $\mathcal{S}$ at least once in the time interval $[0,T]$ and then disappeared. For simplicity, we set $\mu(s) := \Phi(\cdot,0,s)\sharp\tilde{m}_0$.

By the hypotheses on b, $\tilde{m}_0$ and by (2.1.10), we have that $\tilde{\mu}(s)$ is a positive Radon measure on $\mathbb{R}^d$ with finite first order moment. Moreover, it certainly satisfies the constraint $\int_{\mathbb{R}^d} d\tilde{\mu} \le G$ because, with respect to $\mu(s)$, it may only lose mass through the sink. Hence $\tilde{\mu}(s) \in \mathcal{G}$. Furthermore, it is absolutely continuous with a density which is bounded and has a compact support.

Lemma 2.1.1. *For $s \in [0, T]$, consider the function*

$$\pi : \mathcal{B}(s) \longrightarrow [0, s], \qquad \pi(x) := t_x \tag{2.1.11}$$

and suppose that it is Lipschitz continuous (see Remark 2.1.2). Then, the map $s \longmapsto \tilde{\mu}(s)$ is Lipschitz continuous in $\mathcal{G}$ (with respect to the metrics (2.1.3)).

Proof. Let $s, s' \in [0, T]$, $s \geq s'$. Then, recalling (2.1.9), we have

$$\begin{aligned}
\mathcal{W}(\tilde{\mu}(s'), \tilde{\mu}(s)) &= \sup_{\substack{\|\varphi\|_\infty \leq 1 \\ \mathrm{Lip}(\varphi) \leq 1}} \left\{ \int_{\mathbb{R}^d} \varphi(x) d\tilde{\mu}(s')(x) - \int_{\mathbb{R}^d} \varphi(x) d\tilde{\mu}(s)(x) \right\} \\
&= \sup_{\substack{\|\varphi\|_\infty \leq 1 \\ \mathrm{Lip}(\varphi) \leq 1}} \left\{ \int_{\mathbb{R}^d \setminus \Phi(\mathcal{B}(s'),0,s')} \varphi(x) d\mu(s')(x) - \int_{\mathbb{R}^d \setminus \Phi(\mathcal{B}(s),0,s)} \varphi(x) d\mu(s)(x) \right\} \\
&= \sup_{\substack{\|\varphi\|_\infty \leq 1 \\ \mathrm{Lip}(\varphi) \leq 1}} \left\{ \int_{\mathbb{R}^d} \varphi(x) d(\mu(s') - \mu(s))(x) + \int_{\Phi(\mathcal{B}(s),0,s) \setminus \Phi(\mathcal{B}(s'),0,s)} \varphi(x) d\mu(s)(x) \right. \\
&\qquad \left. + \int_{\Phi(\mathcal{B}(s'),0,s)} \varphi(x) d\mu(s)(x) - \int_{\Phi(\mathcal{B}(s'),0,s')} \varphi(x) d\mu(s')(x) \right\} \\
&= \sup_{\substack{\|\varphi\|_\infty \leq 1 \\ \mathrm{Lip}(\varphi) \leq 1}} \left\{ \int_{\mathbb{R}^d} (\varphi(\Phi(x, 0, s')) - \varphi(\Phi(x, 0, s))) d\tilde{m}_0(x) \right. \\
&\qquad + \int_{\mathcal{B}(s) \setminus \mathcal{B}(s')} \varphi(\Phi(x, 0, s)) d\tilde{m}_0(x) \\
&\qquad \left. + \int_{\mathcal{B}(s')} (\varphi(\Phi(x, 0, s)) - \varphi(\Phi(x, 0, s'))) d\tilde{m}_0(x) \right\} \\
&\leq \sup_{\substack{\|\varphi\|_\infty \leq 1 \\ \mathrm{Lip}(\varphi) \leq 1}} \left\{ \int_{\mathbb{R}^d} \|\Phi(x, 0, s') - \Phi(x, 0, s)\| d\tilde{m}_0(x) \right. \\
&\qquad + \int_{\mathcal{B}(s) \setminus \mathcal{B}(s')} \varphi(\Phi(x, 0, s)) d\tilde{m}_0(x) \\
&\qquad \left. + \int_{\mathcal{B}(s')} \|\Phi(x, 0, s) - \Phi(x, 0, s')\| d\tilde{m}_0(x) \right\} \\
&\leq 2GM|s' - s| + \|\tilde{m}_0\|_\infty \mathcal{L}^d(\mathcal{B}(s) \setminus \mathcal{B}(s')),
\end{aligned}$$

where M is the time-Lipschitz constant for Φ (i.e., the bound for b). Hence we conclude if we estimate $\mathcal{L}^d(\mathcal{B}(s) \setminus \mathcal{B}(s'))$. In particular, if we prove that the map $s \longmapsto \mathcal{L}^d(\mathcal{B}(s))$ is Lipschitz continuous, we are done. Observe that the function

$$f : \sigma \longmapsto \pi(\Phi(y, 0, \sigma)) = \pi(y) - \sigma$$

is such that

$$1 = |f'(0)| = |\nabla\pi(y) \cdot b(y, 0)| \leq \|\nabla\pi\| \|b\|_\infty \leq M\|\nabla\pi\| \quad \text{a.e.}$$

and then $\|\nabla\pi\| \geq \frac{1}{M} > 0$ almost everywhere. Therefore

$$\begin{aligned}\frac{\mathcal{L}^d(\mathcal{B}(s) \setminus \mathcal{B}(s'))}{M} &\leq \int_{\mathcal{B}(s)\setminus\mathcal{B}(s')} \|\nabla\pi\| dx = \int_{s'}^{s} \mathcal{H}^{d-1}(\pi^{-1}(\tau))d\tau \\ &= \int_{s'}^{s} \mathcal{H}^{d-1}(\Psi(\mathcal{S}^\tau, \tau, \tau))d\tau \leq K|s - s'|, \quad (2.1.12)\end{aligned}$$

where we used the Coarea Formula (see Theorem B.2.2) and the fact that the $(d-1)$-dimensional Hausdorff measure $\mathcal{H}^{d-1}(\Psi(\mathcal{S}^\tau, \tau, \tau))$ is bounded by a constant $K > 0$ since, by hypotheses, Ψ is Lipschitz continuous and $\mathcal{S}^\tau$ is compact. Then, the map $s \longmapsto \mathcal{L}^d(\mathcal{B}(s))$ is Lipschitz continuous and the thesis follows. □

Remark 2.1.2. *In general, the map π is not Lipschitz continuous in $\mathcal{B}(s)$. But, in view of the possible mean-field game model, the field b will be the optimal feedback for an optimal control problem with controlled dynamics from a system $y' = \alpha$, and hence with total controllability. Then we expect that such a Lipschitz continuity may hold and, at the moment, it is not too restrictive to assume it. Anyway, future investigations will be made on this direction.*

Theorem 2.1.1. *The map $s \longmapsto \tilde{\mu}(s)$ is a weak solution of* (2.1.5)*.*

Proof. Let $\varphi \in C_c^\infty(\mathbb{R}^d \times [0, T[)$. By Lemma 2.1.1, the map

$$s \longmapsto \int_{\mathbb{R}^d} \varphi(x, s) d\tilde{\mu}(s)(x)$$

is absolutely continuous and then we have

$$\begin{aligned}\frac{d}{ds}\int_{\mathbb{R}^d} \varphi(x, s)d\tilde{\mu}(s)(x) &= \frac{d}{ds}\int_{\mathbb{R}^d\setminus\Phi(\mathcal{B}(s),0,s)} \varphi(x, s)d\mu(s)(x) \\ = \frac{d}{ds}\int_{\mathbb{R}^d} \varphi(\Phi(x, 0, s), s)d\tilde{m}_0(x) &- \frac{d}{ds}\int_{\mathcal{B}(s)} \varphi(\Phi(x, 0, s), s)d\tilde{m}_0(x)\end{aligned}$$

$$
\begin{aligned}
&= \int_{\mathbb{R}^d} (\varphi_s(\Phi(x,0,s),s) + \langle D_x\varphi(\Phi(x,0,s),s), b(\Phi(x,0,s),s)\rangle) d\tilde{m}_0(x) \\
&\qquad - \frac{d}{ds}\int_{\mathcal{B}(s)} \varphi(\Phi(x,0,s),s) d\tilde{m}_0(x) \\
&= \int_{\mathbb{R}^d} (\varphi_s(y,s) + \langle D_x\varphi(y,s), b(y,s)\rangle) d\mu(s)(y) \\
&\qquad - \frac{d}{ds}\int_{\mathcal{B}(s)} \varphi(\Phi(x,0,s),s) d\tilde{m}_0(x).
\end{aligned}
$$

We have to compute

$$
\frac{d}{ds}\int_{\mathcal{B}(s)} \varphi(\Phi(x,0,s),s) d\tilde{m}_0(x).
$$

By the Disintegration Theorem (see Remark 2.1.3), we get

$$
\begin{aligned}
&\frac{d}{ds}\int_{\mathcal{B}(s)} \varphi(\Phi(x,0,s),s) d\tilde{m}_0(x) \\
&\quad = \frac{d}{ds}\int_0^s \int_{\{x\in\mathbb{R}^d : t_x=\tau\}} \varphi(\Phi(x,0,s),s) d\tilde{m}_0^\tau(x) d\nu(\tau) \\
&\quad = \frac{d}{ds}\int_0^s \int_{\{x\in\mathbb{R}^d : t_x=\tau\}} \varphi(\Phi(x,0,s),s) g(\tau) d\tilde{m}_0^\tau(x) d\tau \\
&\quad = \int_{\{x\in\mathbb{R}^d : t_x=s\}} \varphi(\Phi(x,0,s),s) g(s) d\tilde{m}_0^s(x) \\
&\quad + \int_{\mathcal{B}(s)} (\varphi_s(\Phi(x,0,s),s) + \langle D_x\varphi(\Phi(x,0,s),s), b(\Phi(x,0,s),s)\rangle) d\tilde{m}_0(x).
\end{aligned}
$$

Now, recalling that $\{x \in \mathbb{R}^d : t_x = s\} = \Psi(\mathcal{S}^s, s, s)$ by definition, we have

$$
\begin{aligned}
&\int_{\{x\in\mathbb{R}^d : t_x=s\}} \varphi(\Phi(x,0,s),s) g(s) d\tilde{m}_0^s(x) \\
&\quad = \int_{\Psi(\mathcal{S}^s,s,s)} \varphi(\Phi(x,0,s),s) g(s) d\tilde{m}_0^s(x) = \int_{\mathcal{S}^s} \varphi(y,s) d\mu^s(s)(y),
\end{aligned}
$$

where $\mu^s(s) := g(s)(\Phi(\cdot,0,s)\sharp\tilde{m}_0^s)$. Finally, we obtain

$$
\begin{aligned}
&\frac{d}{ds}\int_{\mathbb{R}^d} \varphi(y,s) d\tilde{\mu}(s)(y) \\
&\quad = \int_{\mathbb{R}^d\setminus\Phi(\mathcal{B}(s),0,s)} (\varphi_s(y,s) + \langle D_x\varphi(y,s), b(y,s)\rangle) d\mu(s)(y) \\
&\qquad - \int_{\mathcal{S}^s} \varphi(y,s) d\mu^s(s)(y).
\end{aligned}
$$

Since $\tilde{\mu}(0) = \tilde{m}_0$, integrating this between 0 and T we get

$$\int_{\mathbb{R}^d} \varphi(y,0)d\tilde{m}_0(y) + \int_0^T \int_{\mathbb{R}^d \setminus \Phi(\mathcal{B}(s),0,s)} (\varphi_s(y,s) + \langle D_x\varphi(y,s), b(y,s)\rangle)d\mu(s)(y)ds - \int_0^T \int_{\mathcal{S}^s} \varphi(y,s)d\mu^s(s)(y)ds = 0,$$

that is

$$\int_{\mathbb{R}^d} \varphi(y,0)d\tilde{m}_0(y) + \int_0^T \int_{\mathbb{R}^d \setminus \Phi(\mathcal{B}(s),0,s)} (\varphi_s(y,s) + \langle D_x\varphi(y,s), b(y,s)\rangle)d\mu(s)(y)ds - \int_0^T \int_{\mathbb{R}^d} \mathbb{1}_{\{(\mathcal{S}^s,s):s\in[0,T]\}}\varphi(y,s)d\mu^s(s)(y)ds = 0.$$

□

Remark 2.1.3. *The Disintegration Theorem (see also Theorem B.2.3) in the previous proof is applied as follows: we set $Y = \mathcal{B}(s)$, $X = [0,s]$ and we consider the map* (2.1.11)

$$\pi : Y \longrightarrow X, \qquad \pi(x) = t_x$$

and $\nu = \pi\sharp\tilde{m}_0 \in \mathcal{G}(X)$. In this way $\pi^{-1}(\tau) = \{x \in \mathbb{R}^d : t_x = \tau\}$ for every $\tau \in [0,s]$. Then, there exists a ν-almost everywhere uniquely determined family $\{\tilde{m}_0^\tau\}_{\tau\in[0,s]} \subset \mathcal{G}(Y)$ such that for every $f \in C_c^0(Y)$,

$$\int_Y f(y)d\tilde{m}_0(y) = \int_0^s \int_{\{x\in\mathbb{R}^d:t_x=\tau\}} f(y)d\tilde{m}_0^\tau(y)d\nu(\tau).$$

Moreover, in view of (2.1.12)*, that is the Lipschitz continuity of the map $s \longmapsto \mathcal{L}^d(\mathcal{B}(s))$, the measure ν is absolutely continuous on X with a L^∞ density denoted by g.*

Remark 2.1.4. *The absolute continuity of the measure ν on $[0,s]$ can be proved also without assuming the Lipschitz continuity of π. In this case, it turns out to have just a L^1 density. For the proof, we have to show at first the continuity of the map $s \longmapsto \mathcal{L}^d(\mathcal{B}(s))$. By* (2.1.9)*, it follows that*

$s \longmapsto \mathcal{L}^d(\mathcal{B}(s))$ is increasing. It is right continuous too. Indeed, recalling that $\mathrm{d_H}(\cdot,\cdot)$ is the Hausdorff distance in $\mathbb{R}^d$, we have

$$\lim_{s'\to s^+} \mathrm{d_H}(\mathcal{B}(s'),\mathcal{B}(s)) = 0.$$

In fact, since $\mathcal{B}(s) \subseteq \mathcal{B}(s')$, we have

$$\mathrm{d_H}(\mathcal{B}(s'),\mathcal{B}(s)) = \sup_{x\in\mathcal{B}(s')} d(x,\mathcal{B}(s)).$$

Now, if $x \in \mathcal{B}(s')$, then $x \in \mathcal{B}^\tau$ for some $0 \le \tau \le s'$. If $\tau \le s$, then $x \in \mathcal{B}(s)$ and $d(x,\mathcal{B}(s)) = 0$. Otherwise if $s < \tau \le s'$, then $x = \Psi(z,\tau,\tau) = \zeta(\tau)$ for $z \in \partial\mathcal{S}$ and $\tilde{x} := \Psi(z,s,s) = \tilde{\zeta}(s) \in \mathcal{B}(s)$. Moreover, by the well-position of the dynamical system (2.1.6) *(that is if $\tau \to t$, then $\beta_\tau \longrightarrow \beta_t$ uniformly at least on compact sets by the regularity of the field b), we have*

$$\|x-\tilde{x}\| = \|\Psi(z,\tau,\tau) - \Psi(z,s,s)\| \le O(|\tau - s|) \le O(|s'-s|).$$

Then

$$\lim_{s'\to s^+} \mathrm{d_H}(\mathcal{B}(s'),\mathcal{B}(s)) = 0.$$

By (2.1.9) *and the upper-semicontinuity of $\mathcal{L}^d$ w.r.t.* $\mathrm{d_H}$, *we obtain*

$$\mathcal{L}^d(\mathcal{B}(s)) \le \liminf_{s'\to s^+} \mathcal{L}^d(\mathcal{B}(s')) \le \limsup_{s'\to s^+} \mathcal{L}^d(\mathcal{B}(s')) \le \mathcal{L}^d(\mathcal{B}(s))$$

and hence the right continuity holds.

Let us prove the left continuity. Let $s' \to s^-$. Observe that

$$\mathcal{B}(s) \setminus \mathcal{B}(s') \to \mathcal{B}^s = \{x = \Psi(z,s,s) : z \in \partial\mathcal{S}\}$$

in the Hausdorff metrics as $s' \to s^-$. Indeed if $x \in \mathcal{B}(s) \setminus \mathcal{B}(s')$, then $x \in \mathcal{B}^\tau = \{x = \Psi(z,\tau,\tau) : z \in \partial\mathcal{S}\}$ for some $s' < \tau \le s$ and, again by (2.1.9),

$$\mathrm{d_H}(\mathcal{B}(s)\setminus\mathcal{B}(s'),\mathcal{B}^s) = \sup_{x\in\mathcal{B}(s)\setminus\mathcal{B}(s')} d(x,\mathcal{B}^s).$$

Setting again $\tilde{x} := \Psi(z,s,s) \in \mathcal{B}^s$, we have

$$\|x-\tilde{x}\| = \|\Psi(z,\tau,\tau) - \Psi(z,s,s)\| \le O(|\tau - s|) \le O(|s-s'|).$$

Then, by (2.1.9) *and the upper-semicontinuity of the Lebesgue measure w.r.t. the Hausdorff metrics,*

$$\mathcal{L}^d(\mathcal{B}^s) \le \liminf_{s'\to s^-} \mathcal{L}^d(\mathcal{B}(s)\setminus\mathcal{B}(s')) \le \limsup_{s'\to s^-} \mathcal{L}^d(\mathcal{B}(s)\setminus\mathcal{B}(s')) \le \mathcal{L}^d(\mathcal{B}(s)).$$

It follows that $\lim_{s'\to s^-}\mathcal{L}^d(\mathcal{B}(s)\setminus\mathcal{B}(s'))=\mathcal{L}^d(\mathcal{B}^s)$.

Now, $\mathcal{B}^s$ *is the image of* $\partial\mathcal{S}$ *by* $z\longmapsto\Psi(z,s,s)$. *Such a function is Lipschitz continuous and hence it preserves zero Lebesgue measure sets, that is if* $\mathcal{L}^d(\partial\mathcal{S})=0$, *then* $\mathcal{L}^d(\mathcal{B}^s)=0$. *Therefore*

$$\lim_{s'\to s^-}\mathcal{L}^d(\mathcal{B}(s)\setminus\mathcal{B}(s'))=0.$$

Then

$$\begin{aligned}\lim_{s'\to s^-}\mathcal{L}^d(\mathcal{B}(s'))&=\lim_{s'\to s^-}\Big(\mathcal{L}^d(\mathcal{B}(s))-\mathcal{L}^d(\mathcal{B}(s)\setminus\mathcal{B}(s'))\Big)\\&=\mathcal{L}^d(\mathcal{B}(s))-\lim_{s'\to s^-}\mathcal{L}^d(\mathcal{B}(s)\setminus\mathcal{B}(s'))\\&=\mathcal{L}^d(\mathcal{B}(s))\end{aligned}\quad.$$

Finally

$$\lim_{s'\to s}\mathcal{L}^d(\mathcal{B}(s'))=\mathcal{L}^d(\mathcal{B}(s)),$$

that is the function $s\longmapsto\mathcal{L}^d(\mathcal{B}(s))$ *is continuous.*

Now, we prove the absolute continuity of ν *on the intervals of* $[0,s]$. *So let* $[s_1,s_2]\subset[0,s]$. *By the definition of* t_x (2.1.2), *we have*

$$\pi^{-1}([s_1,s_2])\subseteq\mathcal{B}(s_2)\setminus\mathcal{B}(s_1).$$

Hence

$$\begin{aligned}\nu([s_1,s_2])&=m_0(\pi^{-1}([s_1,s_2]))\le\|m_0\|_{L^\infty}\mathcal{L}^d(\pi^{-1}([s_1,s_2]))\\&\le\|m_0\|_{L^\infty}\mathcal{L}^d(\mathcal{B}(s_2)\setminus\mathcal{B}(s_1))\le\|m_0\|_{L^\infty}\omega(|s_2-s_1|),\end{aligned}\qquad(2.1.13)$$

where ω *is the modulus of continuity of the map* $s\longmapsto\mathcal{L}^d(\mathcal{B}(s))$ *(which exists globally since such a function is uniformly continuous on* $[s_1,s_2]$*). It follows that* ν *is absolutely continuous on the intervals of* $[0,s]$.

Now, let I *be a Borel subset of* $[0,s]$ *such that* $\mathcal{L}(I)=0$. *Then, for every* $\varepsilon>0$ *there exists an open set* $\mathcal{O}$ *such that* $\mathcal{O}\supset I$ *and* $\mathcal{L}(\mathcal{O})<\varepsilon$. *But every open set in the real line can be expressed as a disjoint countable union of open intervals and then, thanks to* (2.1.13), *we have that* $\nu(\mathcal{O})<\varepsilon$. *By the arbitrariness of* ε, *we conclude that* ν *is absolutely continuous on* $[0,s]$.

Notice that, in general, these calculations do not allow to bypass the hypothesis of Lipschitz continuity of π. *Indeed, the continuity of the map* $s\longmapsto\mathcal{L}^d(\mathcal{B}(s))$ *is not sufficient to prove Lemma 2.1.1, which anyway needs a Lipschitz property of* π. *This would guarantee a Lipschitz continuity of* $s\longmapsto\mathcal{L}^d(\mathcal{B}(s))$, *which is essential to prove such a lemma (see also Remark 2.1.2). Nevertheless, the same arguments allow us to weaken the hypotheses in Theorem 2.1.3 (see also Remark 2.1.5).*

Theorem 2.1.2. *The continuity equation* (2.1.5) *has a unique solution given by* $s \longmapsto \tilde{\mu}(s)$.

Proof. Let $\varphi \in C^\infty(\mathbb{R}^d)$ with $\text{supp}(\varphi) \subset \mathbb{R}^d \setminus \Phi(\mathcal{B}(t), 0, t)$ for every $t \leq T$. Fix $t \leq T$ and let us consider the map

$$w : \mathbb{R}^d \times [0, t] \longrightarrow \mathbb{R}, \qquad w(x, s) := \varphi(\Phi(x, 0, t - s)). \tag{2.1.14}$$

Then, w is Lipschitz continuous in both variables $(x, s) \in \mathbb{R}^d \times [0, t]$ with $\text{supp}(w) \subset (\mathbb{R}^d \setminus \Phi(\mathcal{B}(s), 0, s)) \times [0, t]$. Moreover, by (2.1.7) we have

$$\varphi(x) = w(\Psi(x, t - s, t), s) = \varphi(\Phi(\Psi(x, t - s, t), 0, t - s))$$

and, recalling that $\Psi(x, t, t)$ is the solution of (2.1.6) with $\tau = t$, the function w satisfies

$$\begin{aligned} 0 = \frac{d}{ds}\varphi(x) = w_s(&\Psi(x, t - s, t), s) \\ &+ \langle D_x w(\Psi(x, t - s, t), s), b(\Psi(x, t - s, t), s)\rangle \quad \text{a.e.} \end{aligned}$$

and hence, in general,

$$w_s(y, s) + \langle D_x w(y, s), b(y, s)\rangle = 0 \quad \text{a.e. in } \mathbb{R}^d \times]0, t[.$$

Using w as a test function for a generic μ satisfying Definition 2.1.1, for almost all $s \leq t$ we have

$$\begin{aligned} \frac{d}{ds}\int_{\mathbb{R}^d} w(y, s) d\mu(s)(y) &= \int_{\mathbb{R}^d} w_s(y, s) d\mu(s)(y) + \int_{\mathbb{R}^d} w(y, s) d\mu_s(s)(y) \\ &= \int_{\mathbb{R}^d} (-\langle D_x w(y, s), b(y, s)\rangle + \langle D_x w(y, s), b(y, s)\rangle) d\mu(s)(y) = 0 \end{aligned}$$

since $\text{supp}(w) \subset (\mathbb{R}^d \setminus \Phi(\mathcal{B}(s), 0, s)) \times [0, t]$, which implies

$$\begin{aligned} &\int_{\mathcal{S}^s} w(y, s) d\mu^s(s)(y) \\ &\qquad = \int_{\Phi(\mathcal{B}(s), 0, s)} (w_s(y, s) + \langle D_x w(y, s), b(y, s)\rangle) d\mu(s)(y) = 0, \quad s \leq t. \end{aligned}$$

Therefore, integrating between 0 and t, we get

$$\int_{\mathbb{R}^d} w(y, t) d\mu(t)(y) = \int_{\mathbb{R}^d} w(y, 0) d\mu(0)(y)$$

and then

$$\int_{\mathbb{R}^d} \varphi(y)d\mu(t)(y) = \int_{\mathbb{R}^d} \varphi(\Phi(y,0,t))d\tilde{m}_0(y),$$

which shows that $\mu(t) = \Phi(\cdot, 0, t)\sharp\tilde{m}_0$ on $\mathbb{R}^d \setminus \Phi(\mathcal{B}(t), 0, t)$.

Now we have to prove that $\mu(t) = 0$ on $\Phi(\mathcal{B}(t), 0, t)$, that is

$$\int_{\mathbb{R}^d} \varphi(x)d\mu(t)(x) = 0$$

for any $\varphi \in C^\infty(\mathbb{R}^d)$ with $\mathrm{supp}(\varphi) \subset \Phi(\mathcal{B}(t), 0, t)$ for every $t \leq T$. Again fix $t \leq T$ and let us consider the map (2.1.14). Then, proceeding as before, we get

$$w_s(y,s) + \langle D_x w(y,s), b(y,s)\rangle = 0 \quad \text{a.e. in } \mathbb{R}^d \times]0,t[.$$

Using w as a test function for a generic μ satisfying Definition 2.1.1, for almost all $s \leq t$ we have

$$\begin{aligned}
\frac{d}{ds}\int_{\mathbb{R}^d} w(y,s)d\mu(s)(y) &= \int_{\mathbb{R}^d} w_s(y,s)d\mu(s)(y) + \int_{\mathbb{R}^d} w(y,s)d\mu_s(s)(y)\\
&= \int_{\mathbb{R}^d} (-\langle D_x w(y,s), b(y,s)\rangle + \langle D_x w(y,s), b(y,s)\rangle)d\mu(s)(y)\\
&\quad - \int_{\Phi(\mathcal{B}(s),0,s)} (w_s(y,s) + \langle D_x w(y,s), b(y,s)\rangle)d\mu(s)(y) - \int_{\mathcal{S}^s} w(y,s)d\mu^s(s)(y)\\
&= -\int_{\Phi(\mathcal{B}(s),0,s)} (w_s(y,s) + \langle D_x w(y,s), b(y,s)\rangle)d\mu(s)(y)\\
&\qquad - \int_{\mathcal{S}^s} w(y,s)d\mu^s(s)(y).
\end{aligned}$$

Now, observe that

$$\begin{aligned}
&\int_{\mathcal{S}^s} w(y,s)d\mu^s(s)(y) + \int_{\Phi(\mathcal{B}(s),0,s)} (w_s(y,s) + \langle D_x w(y,s), b(y,s)\rangle)d\mu(s)(y)\\
&= \int_{\{x\in\mathbb{R}^d : t_x = s\}} w(\Phi(y,0,s),s)g(s)d\mu^s(0)(y)\\
&\quad + \int_{\mathcal{B}(s)} (w_s(\Phi(y,0,s),s) + \langle D_x w(\Phi(y,0,s),s), b(\Phi(y,0,s),s)\rangle)d\mu(0)(y)\\
&= \frac{d}{ds}\int_{\mathcal{B}(s)} w(\Phi(y,0,s),s)d\mu(0)(y).
\end{aligned}$$

Then, by (2.1.14) and the semigroup property of the flow Φ, since $\mu(0) = \tilde{m}_0$ we obtain

$$\frac{d}{ds}\int_{\mathbb{R}^d} w(y,s)\mu(s)(y) = -\frac{d}{ds}\int_{\mathcal{B}(s)} \varphi(\Phi(y,0,t))d\tilde{m}_0(y)$$

and hence, integrating between 0 and t,

$$\int_{\mathbb{R}^d} w(y,t)d\mu(t)(y) = \int_{\mathbb{R}^d} w(y,0)d\mu(0)(y) - \int_0^t \frac{d}{ds}\int_{\mathcal{B}(s)} \varphi(\Phi(y,0,t))d\tilde{m}_0(y).$$

Therefore

$$\begin{aligned}\int_{\mathbb{R}^d} \varphi(y)d\mu(t)(y) = \int_{\mathbb{R}^d} \varphi(\Phi(y,0,t))d\tilde{m}_0(y) - \int_{\mathcal{B}(t)} \varphi(\Phi(y,0,t))d\tilde{m}_0(y) \\ + \int_{\mathcal{B}(0)} \varphi(\Phi(y,0,t))d\tilde{m}_0(y).\end{aligned}$$

Since $\operatorname{supp}(\varphi) \subset \Phi(\mathcal{B}(t),0,t)$ (and $\tilde{m}_0 = 0$ in $\mathcal{S} = \mathcal{B}(0)$), the thesis follows. □

2.1.2 On the measure dependence of the field b

We consider now the field b depending on the measure, that is

$$\begin{aligned} b : C^0([0,T],\mathcal{G}) \times \mathbb{R}^d \times [0,T] &\longrightarrow \mathbb{R}^d, \\ (\mu,x,t) &\longmapsto b(\mu,x,t). \end{aligned}$$

We suppose that it is bounded and continuous in the whole entry (μ,x,t) and Lipschitz continuous w.r.t. $x \in \mathbb{R}^d$ uniformly w.r.t. $(\mu,t) \in C^0([0,T],\mathcal{G}) \times [0,T]$, that is, there exists $L > 0$ such that

$$\|b(\mu,x,t) - b(\mu,y,t)\| \le L\|x-y\|, \quad \forall x,y \in \mathbb{R}^d,\ (\mu,t) \in C^0([0,T],\mathcal{G}) \times [0,T].$$

In the evolution of the flow given by b, the sink is always represented by $\mathcal{S}$. In the sequel, for every $\mu \in C^0([0,T],\mathcal{G})$ fixed, we will also use the notation

$$\begin{aligned} b[\mu] : \mathbb{R}^d \times [0,T] &\longrightarrow \mathbb{R}^d \\ (x,t) &\longmapsto b[\mu](x,t) := b(\mu,x,t) \end{aligned}$$

and we consider the corresponding flow with sink evolution given by the field $b[\mu]$. As, for every fixed μ, we denote by $\tilde{\mu}$ the unique solution of the corresponding problem (2.1.5), which is, for every $t \in [0,T]$,

$$\tilde{\mu}(t) = \begin{cases} \Phi[\mu](\cdot,0,t)\sharp\tilde{m}_0 & \text{on } \mathbb{R}^d \setminus \Phi[\mu](\mathcal{B}[\mu](t),0,t) \\ 0, & \text{otherwise} \end{cases}, \tag{2.1.15}$$

where $\Phi[\mu]$ is the flow generated by the field $b[\mu]$ and $\mathcal{B}[\mu](\cdot)$ is the corresponding sink-reaching-points set. We also denote by $\pi[\mu]$ the map as in (2.1.11).

Theorem 2.1.3. *Let us suppose that $\pi[\mu]$ is Lipschitz continuous uniformly in $\mu \in C^0([0,T],\mathcal{G})$ (see Remark 2.1.5). For every $\mu \in C^0([0,T],\mathcal{G})$, we have $\tilde{\mu} \in C^0([0,T],\mathcal{G})$. Moreover, the function*

$$\psi : C^0([0,T],\mathcal{G}) \longrightarrow C^0([0,T],\mathcal{G}), \qquad \psi(\mu) := \tilde{\mu}$$

has a fixed point in $C^0([0,T],\mathcal{G})$. This means that the problem of flow with sink and with field depending on the measure has a solution.

Proof. At first observe that, under the previous hypotheses, by analogous considerations as in the case with no measure dependence we have $\tilde{\mu}(t) \in \mathcal{G}$ for all $t \in [0,T]$.

Let us prove that $\tilde{\mu} \in C^0([0,T],\mathcal{G})$. In particular we have to prove that, whenever $t_n \to t$ in $[0,T]$, then $\tilde{\mu}(t_n) \to \tilde{\mu}(t)$ weakly-star. This comes from standard regularity results for push-forward measures (see (2.1.15)) and from the fact that $\mathcal{B}[\mu](t_n) \to \mathcal{B}[\mu](t)$ in the Hausdorff metrics and as d-dimensional Lebesgue measure, and hence similarly for $\Phi[\mu](\mathcal{B}[\mu](t_n),0,t_n)$ and $\Phi[\mu](\mathcal{B}[\mu](t),0,t)$.

Now, to prove the second statement of the theorem, we have to show that the function ψ is continuous and compact. In this way, we can conclude by the Schauder-Tychonoff fixed-point Theorem (see Theorem B.2.5).

At first we prove the continuity of ψ. Let $\mu_n \to \mu$ in $C^0([0,T],\mathcal{G})$. We have to prove that $\tilde{\mu}_n \to \tilde{\mu}$. Let us consider the two trajectories

$$x_n(t) = x_0 + \int_0^t b[\mu_n](x_n(s),s)ds, \qquad x(s) = x_0 + \int_0^t b[\mu](x(s),s)ds$$

and we prove that x_n uniformly converges to x on compact sets. Indeed, by the continuity and boundedness hypotheses, the field b is bounded and uniformly continuous on $(\{\mu_n\}_n \cup \{\mu\}) \times K \times [0,T]$, where $\{\mu_n\}_n \cup \{\mu\}$ is the compact set in $C^0([0,T],\mathcal{G})$ given by the whole sequence with its limit, and $K \subset \mathbb{R}^d$ is compact. Then, the sequence $b[\mu_n]$ is bounded and equicontinuous on $K \times [0,T]$ and moreover it pointwise converges to $b[\mu]$. Hence, by the Ascoli-Arzelà Theorem, the convergence is uniform on $K \times [0,T]$. From this we deduce the desired uniform convergence of the trajectories. Therefore, also the flows $\Phi[\mu_n]$ uniformly converge to $\Phi[\mu]$ on compact sets. From this we have that $\mathcal{B}[\mu_n](\cdot)$ converges to $\mathcal{B}[\mu](\cdot)$ in the Hausdorff distance and uniformly in time, and we conclude the convergence of $\tilde{\mu}_n$ to $\tilde{\mu}$ in $C^0([0,T],\mathcal{G})$, that is, for every $\varphi \in C_c^0(\mathbb{R}^d)$,

$$\sup_{t\in[0,T]} \left| \int_{\mathbb{R}^d} \varphi(x)d\mu_n(t)(x) - \int_{\mathbb{R}^d} \varphi(x)d\mu(t)(x) \right| \to 0 \quad \text{as } n \to +\infty.$$

It remains to prove that ψ has compact image. We can restrict to measures on a compact set $K \subset \mathbb{R}^d$ independent of μ, which contains all the possible compact supports of the measures $\tilde{\mu}$ (because of bounded dynamics and finite horizon). Then, since $s \longmapsto \tilde{\mu}(s)$ is Lipschitz continuous uniformly in μ (similarly as in Lemma 2.1.1 and using the hypothesis on $\pi[\mu]$) with values in the compact set $\mathcal{G}$ of Radon measures on K, we get the desired conclusion by Ascoli-Arzelà Theorem. $\square$

Remark 2.1.5. *In the previous proof, we used the fact that $s \longmapsto \tilde{\mu}(s)$ is Lipschitz continuous uniformly in μ, and this comes from the Lipschitz continuity of $\pi[\mu]$ uniformly w.r.t. μ. However, just assuming the Lipschitz continuity of $\pi[\mu]$, and not necessarily uniformly in μ, after some calculations similar to the ones in Remark 2.1.4, we can prove the equicontinuity w.r.t. μ of $s \longmapsto \tilde{\mu}(s)$ and then still apply Ascoli-Arzelà Theorem.*

2.1.3 A differential characterization of the first arrival time

In this section, we see a characterization of the possible first arrival time t_x (2.1.2) to the sink $\mathcal{S}$ as the unique viscosity solution of an Hamilton-Jacobi equation with suitable boundary conditions. For the statements and the proofs, we mostly refer to [9]. We define

$$\begin{aligned} &\mathrm{T} : \mathbb{R}^d \times [0,T] \longrightarrow [0,+\infty[, \\ &\mathrm{T}(x,t) := t_{(x,t)} = \inf\{\zeta \geq t : \Phi(x,t,\zeta) \in \mathcal{S}\}. \end{aligned}$$

Fixed $s \in [t,T]$, the sink-reaching-points set, that is the set from which the agents are able to reach the target $\mathcal{S}$, is

$$\mathcal{B}(s) = \{(x,t) \in \mathbb{R}^d \times [0,T] : \mathrm{T}(x,t) \leq s\}.$$

Let us set

$$\mathcal{B} := \bigcup_{s\in[t,T]} \mathcal{B}(s) = \{(x,t) \in \mathbb{R}^d \times [0,T] : \mathrm{T}(x,t) \leq T\}.$$

We suppose that

(i) T is continuous in $\mathcal{B}$;

(ii) $\mathrm{T}(x,t) = T$ for any $(x,t) \in \partial\mathcal{B}$.

Remark 2.1.6. *Hypotheses* (i) *and* (ii) *are not too restrictive to be assumed. Indeed, a priori,* T *may be discontinuous in* $\mathcal{B}$*. But, as we noticed in Remark 2.1.2 too, in view of the possible mean-field game model, the field* b *comes as the optimal feedback from a system like* $y' = \alpha$*, with total controllability. Then it is not too strong to assume that* T *is continuous in* $\mathcal{B}$*. Moreover, again in view of the possible mean-field model, condition* (ii) *holds only on the boundary which is not viable.*

Remark 2.1.7. *Observe that* $\mathrm{T}(x,t) > 0$ *if and only if* $(x,t) \notin (\mathcal{S}, t)$*. Indeed, by the properties of the trajectory (see* (1.1.9)*),*

$$d(x, \mathcal{S}) \le \|\Phi(x,t,t_{(x,t)}) - x\| \le M(t_{(x,t)} - t) = M(\mathrm{T}(x,t) - t),$$

where M *is the bound for* b*, and* $d(x,\mathcal{S}) > 0$ *for* $(x,t) \notin (\mathcal{S}, t)$ *because* $\mathcal{S}$ *is closed.*

We have the following Dynamic Programming Principle for T.

Proposition 2.1.1. *For every* $(x,t) \in \mathcal{B}$*,*

$$\mathrm{T}(x,t) = (\tau - t) \wedge t_{(x,t)} + \mathbb{1}_{\{\tau - t \le t_{(x,t)}\}} \mathrm{T}(\Phi(x,t,\tau), \tau) \quad \textit{for all } \tau \ge t \tag{2.1.16}$$

and

$$\mathrm{T}(x,t) = (\tau - t) + \mathrm{T}(\Phi(x,t,\tau),\tau) \quad \textit{for all } \tau \in [t, \mathrm{T}(x,t)]. \tag{2.1.17}$$

Proof. Let B and C be respectively the right-hand sides of (2.1.16) and (2.1.17). Note that (2.1.16) reduces to the definition of $\mathrm{T}(x,t)$ for $\tau - t > \mathrm{T}(x,t)$. For $\tau - t \le \mathrm{T}(x,t)$, we have $B = C$. To show that $\mathrm{T}(x,t) = C$, observe that for all $t \le \tau \le \mathrm{T}(x,t) = t_{(x,t)}$ we have

$$\mathrm{T}(x,t) = t_{(x,t)} = (\tau - t) + t_{(\Phi(x,t,\tau),\tau)} = (\tau - t) + \mathrm{T}(\Phi(x,t,\tau),\tau).$$

Hence $\mathrm{T}(x,t) = C = B$ and the proof is complete. $\square$

Now, taking into account that T is time dependent (and then the associated Hamilton-Jacobi equation is evolutive) and that there is no control, we have the following

Theorem 2.1.4. *The function* T *is the unique viscosity solution of*

$$\begin{cases} -\mathrm{T}_t(x,t) - b(x,t) \cdot D_x \mathrm{T}(x,t) - 1 = 0, & (x,t) \in \mathcal{B} \\ \mathrm{T}(x,t) = t, & (x,t) \in (\partial\mathcal{S}, t) \ . \\ \mathrm{T}(x,t) = T, & (x,t) \in \partial\mathcal{B} \end{cases} \tag{2.1.18}$$

Proof. Let us prove at first that T is a viscosity solution of (2.1.18). Let $\varphi \in C^1(\mathcal{B})$ and (x_1, t_1) be a local maximum point of $\mathrm{T} - \varphi$, that is for some $r > 0$,

$$\mathrm{T}(x_1, t_1) - \mathrm{T}(x, t) \geq \varphi(x_1, t_1) - \varphi(x, t) \quad \text{for every } (x, t) \in B((x_1, t_1), r).$$

For τ sufficiently close to t_1, we have that $(\Phi(x_1, t_1, \tau), \tau) \in B((x_1, t_1), r)$ by the properties of the trajectory ((1.1.9)) and then

$$\varphi(x_1, t_1) - \varphi(\Phi(x_1, t_1, \tau), \tau) \leq \mathrm{T}(x_1, t_1) - \mathrm{T}(\Phi(x_1, t_1, \tau), \tau).$$

Now, by Proposition 2.1.1 (note that $\mathrm{T}(x, t) > 0$ by Remark 2.1.7) we have

$$\begin{aligned}\varphi(x_1, t_1) - \varphi(\Phi(x_1, t_1, \tau), \tau) &\leq \mathrm{T}(x_1, t_1) - \mathrm{T}(\Phi(x_1, t_1, \tau), \tau)\\ &= (\tau - t_1) + \mathrm{T}(\Phi(x_1, t_1, \tau), \tau) - \mathrm{T}(\Phi(x_1, t_1, \tau), \tau) = \tau - t_1,\end{aligned}$$

that is

$$\frac{\varphi(x_1, t_1) - \varphi(\Phi(x_1, t_1, \tau), \tau)}{\tau - t_1} \leq 1.$$

Letting $\tau \to t_1$, we get

$$-\varphi_t(x_1, t_1) - b(x_1, t_1) \cdot D_x\varphi(x_1, t_1) \leq 1.$$

We conclude that T is a subsolution of (2.1.18).

Next assume that (x_2, t_2) is a local minimum point of $\mathrm{T} - \varphi$, $\varphi \in C^1(\mathcal{B})$. As above, for τ sufficiently close to t_2, by (1.1.9) we have

$$\varphi(x_2, t_2) - \varphi(\Phi(x_2, t_2, \tau), \tau) \geq \mathrm{T}(x_2, t_2) - \mathrm{T}(\Phi(x_2, t_2, \tau), \tau).$$

By Proposition 2.1.1, we have

$$\varphi(x_2, t_2) - \varphi(\Phi(x_2, t_2, \tau), \tau) \geq \tau - t_2,$$

that is

$$\frac{\varphi(x_2, t_2) - \varphi(\Phi(x_2, t_2, \tau), \tau)}{\tau - t_2} \geq 1.$$

Letting $\tau \to t_2$, we obtain

$$-\varphi_t(x_2, t_2) - b(x_2, t_2) \cdot D_x\varphi(x_2, t_2) \geq 1.$$

We conclude that T is a supersolution of (2.1.18).

Now we prove that T is the unique viscosity solution of (2.1.18). Let u_1 and u_2 be, respectively, a subsolution and a supersolution of (2.1.18). We define

$$v_i(x,t) = 1 - e^{-u_i(x,t)}, \qquad i = 1,2.$$

By Proposition B.1.2 (taking into account that the Hamilton-Jacobi equation in (2.1.18) is evolutive), v_1 and v_2 are, respectively, a sub- and a supersolution of

$$-v_t(x,t) + v(x,t) - b(x,t)\cdot D_x v(x,t) - 1 = 0, \qquad (x,t)\in\mathcal{B}.$$

Moreover, v_i is bounded if u_i is bounded below and, by the boundary conditions on $\partial\mathcal{B}$ in (2.1.18), v_i can be uniquely extended to $v_i \in \mathrm{BUC}(\overline{\mathcal{B}})$ satisfying the boundary conditions

$$\begin{aligned} v_i(x,t) &= 1 - e^{-t}, \quad (x,t)\in(\partial\mathcal{S},t), \quad i=1,2\\ v_i(x,t) &= 1 - e^{-T}, \qquad (x,t)\in\partial\mathcal{B}, \quad i=1,2 \end{aligned}.$$

Then, by comparison results (see Theorem B.1.1 and Remark B.1.2) we obtain that $v_1 \le v_2$ and therefore $u_1 \le u_2$. The thesis follows by exchanging the roles of u_1 and u_2. □

Hence, to conclude, the possible first arrival time t_x (2.1.2) to the sink $\mathcal{S}$ is given by the solution $\mathrm{T}(x,t)$ of (2.1.18) at the time $t=0$.

Remark 2.1.8. *Until now, the region representing the sink, i.e., $\mathcal{S}$, was considered for simplicity as constant. Indeed, as mentioned at the beginning of §2.1.1, it can be in general moving in time but the generalization works with the same ideas and calculations. Anyway, for the evolutive case, suitable hypotheses have to be assumed. The sink is represented by the image of a multifunction*

$$\mathcal{S}:\mathbb{R}\longrightarrow\mathcal{P}(\mathbb{R}^d), \qquad t\longmapsto\mathcal{S}(t)\subset\mathbb{R}^d, \tag{2.1.19}$$

which describes its evolution, and it has to be supposed continuous with respect to the Hausdorff distance. Moreover, similarly to the non-moving case, we have to assume $\mathcal{S}(t)$ closed with compact and C^1 boundary $\partial\mathcal{S}(t)$ for every $t\in[0,T]$. The continuity w.r.t. the Hausdorff distance is essential in the proof of Theorem 2.1.1, in particular when we integrate between 0 and T the map $s\longmapsto\int_{\mathcal{S}^s(s)}\varphi(y,s)d\mu^s(s)(y)ds$, which has to be at least continuous (in order to be integrable). This is guaranteed by Lemma 2.1.1 and by the continuity of the map (2.1.19) w.r.t. the Hausdorff distance. The other assumptions on $\mathcal{S}(t)$ and its boundary are as necessary as the ones in the constant case.

2.1.4 The continuity equation with a sink and a source

Here, we extend the case of the continuity equation with just a sink, studied in §2.1.1, to a more general one. In particular, we want to model the evolution on $\mathbb{R}^d$ of a mass μ subject to a given flow with a presence of two regions in $\mathbb{R}^d$ acting as a sink and a source: the portion of mass that possibly enters the sink instantaneously disappears and the portion of mass that possibly exits the source starts flowing immediately. We consider the regions representing the source and the sink as constant. In particular, the source is represented by a subset $\Gamma \subset \mathbb{R}^d$, which has the same properties of the sink $\mathcal{S}$ in §2.1.1.

In order to preserve the hierarchical feature in Ch. 1, §1.1.2 (see also the comments at the beginning of Ch. 2) and to make the problem of flow with a sink and a source more close to the one with a sink only, we study it at a fixed level $p_1 \in \mathcal{I}$, assuming that the agents flowing out from the source come from a previous level p_0 (i.e., $p_1 \in \mathcal{I}_{p_0}$), in which Γ represents a sink. We need to introduce the following quantities:

$$\begin{aligned}
b_1 &:= \text{field at level } p_1,\\
b_0 &:= \text{field at level } p_0,\\
\tilde{m}_0^{p_1} &:= \text{initial distribution at } p_1 \text{ (as in (2.1.4) with } m_0 = m_0^{p_1}),\\
\tilde{m}_0^{p_0} &:= \text{initial distribution at } p_0 \text{ (as in (2.1.4) with } m_0 = m_0^{p_0} \text{ and } \mathcal{S} = \Gamma).
\end{aligned}$$

The hypotheses on $\mathcal{S}$, b_1, b_0, $\tilde{m}_0^{p_1}$ and $\tilde{m}_0^{p_0}$ are the same as $\mathcal{S}$, b and $\tilde{m}_0$ in §2.1.1. Roughly speaking, at level p_0 the agents flow with Φ_0 (solving (2.1.1) with $b = b_0$), reach the sink Γ and pass to the subsequent level p_1 (in which Γ is now a source) starting to flow with Φ_1 (solving (2.1.1) with $b = b_1$) and possibly entering the sink $\mathcal{S}$. Such agents, at level p_1 are then detected flowing with Φ_1 "concatenated" to Φ_0.

In order to formally formulate the problem, we have to basically redefine the same quantities as in §2.1.1 in view of the presence of a source and of a flow of agents coming from a previous level. In the following, we denote the starting positions of agents on p_0 and on p_1 by $y \in \mathbb{R}^d$ and $x \in \mathbb{R}^d$ respectively. In particular, we define

$$\bar{t} := t_y^{p_0} = \inf\{t \geq 0 : \Phi_0(y, 0, t) \in \Gamma\}, \quad y \in \mathbb{R}^d \quad (\inf \emptyset = +\infty), \tag{2.1.20}$$

that is the possible first arrival time to the sink Γ at level p_0, and

$$t_{(x,t)}^{p_1} := \inf\{\tau \geq t : \Phi_1(x, t, \tau) \in \mathcal{S}\}, \quad (x,t) \in \mathbb{R}^d \times [0,T] \quad (\inf \emptyset = +\infty),$$

that is the possible first arrival time to the sink $\mathcal{S}$ starting from time t at level p_1. Moreover, we introduce the flow $\Phi_1 \otimes \Phi_0$ as the concatenation of the flows Φ_1 and Φ_0, that is $(\Phi_1 \otimes \Phi_0)(\cdot, 0, \cdot) = y(s)$ as the solution of

$$\begin{cases} y'(s) = b_0(y(s), s), & 0 < s \le \bar{t} \\ y(0) = y \in \mathbb{R}^d \end{cases}, \qquad \begin{cases} y'(s) = b_1(y(s), s), & s > \bar{t} \\ y(\bar{t}) = x_{\bar{t}} \end{cases},$$

where $x_{\bar{t}} \in \Gamma$ is such that $x_{\bar{t}} = \bar{y} = \Phi_0(y, 0, \bar{t})$, that is the starting flowing point from Γ at level p_1 corresponding to the arrival point to Γ at level p_0. By hypotheses on the fields b_0 and b_1, the flow $(\Phi_1 \otimes \Phi_0)(\cdot, 0, \cdot)$ is Lipschitz continuous. Here, and in the sequel, we use the notation "$\otimes$" even though it is typically used for the tensor product or the measure product.

We then define

$$t_y^{p_1 \otimes p_0} := \inf\{\tau \ge 0 : (\Phi_1 \otimes \Phi_0)(y, 0, \tau) \in \mathcal{S}\}, \quad y \in \mathbb{R}^d \quad (\inf \emptyset = +\infty),$$

that is the possible first arrival time to the sink $\mathcal{S}$ at level p_1 for agents coming from level p_0 and starting from y at time 0. Note that, if $t_y^{p_0} < +\infty$, we have

$$t_y^{p_1 \otimes p_0} = t^{p_1}_{(\bar{y}, t_y^{p_0})} = t^{p_1}_{(\bar{y}, \bar{t})}.$$

The condition $t_y^{p_0} < +\infty$ means that the set in (2.1.20) is nonempty, and hence there exists at least an instant $t \ge 0$ at which the agents reach the sink Γ at level p_0 and pass to the new subsequent level p_1 (see also Remark 2.1.9).

The set of possible arrival points to the sink $\mathcal{S}$, for a given $t \in [0, T]$, for agents starting from level p_0, is given by

$$\mathcal{S}^t_{p_1 \otimes p_0} := \{z \in \partial\mathcal{S} : \exists y \in \mathbb{R}^d \text{ such that } t_y^{p_1 \otimes p_0} = t \text{ and } (\Phi_1 \otimes \Phi_0)(y, 0, t) = z\}.$$

We call instead $\Gamma^t_{p_0}$ and $\mathcal{S}^t_{p_1}$ the sets of possible arrival points to the sinks Γ and $\mathcal{S}$, for a given $t \in [0, T]$, for agents starting from levels p_0 and p_1 respectively.

Similarly to §2.1.1, we denote by Ψ_1 the inverse of the flow Φ_1 starting from $\partial\mathcal{S}$ at level p_1, and by Ψ_0 the inverse of the flow Φ_0 starting from $\partial\Gamma$ at level p_0. Moreover, we denote by $\Psi_0 \otimes \Psi_1$ the inverse of the flow $\Phi_1 \otimes \Phi_0$ starting from $\partial\mathcal{S}$, i.e., all the states at level p_0 backwardly reached by the trajectories starting from the points of $\partial\mathcal{S}$ at level p_1, and passing through Γ, in the time interval $[0, T]$. That is

$$(z, \tau) \longmapsto (\Psi_0 \otimes \Psi_1)(z, \tau, \tau), \quad z \in \partial\mathcal{S}, \ \ 0 \le \tau \le T$$

with $(\Psi_0 \otimes \Psi_1)(z,\tau,\tau) = \zeta(\tau)$ solving

$$\begin{cases} \zeta'(s) = -b_1(\zeta(s), \tau - s), & 0 < s < \tau \\ \zeta(0) = z \end{cases}, \qquad \begin{cases} \zeta'(s) = -b_0(\zeta(s), \bar{s} - s), & 0 < s < \tau - \bar{s} \\ \zeta(\bar{s}) = \bar{z} \in \Gamma \end{cases}, \tag{2.1.21}$$

where $\bar{s} = \inf\{s \geq 0 : \zeta(s) \in \Gamma\}$. It is Lipschitz continuous as $\Phi_1 \otimes \Phi_0$ and such that

$$\begin{aligned} &(\Phi_1 \otimes \Phi_0)((\Psi_0 \otimes \Psi_1)(x,\tau,\tau), 0, \tau) = x, \\ &(\Psi_0 \otimes \Psi_1)((\Phi_1 \otimes \Phi_0)(y, 0, \tau), \tau, \tau) = y. \end{aligned} \tag{2.1.22}$$

Now, fixed $s \in [0,T]$, we redefine the following sink-reaching-points sets for levels p_1 (starting from $t \neq 0$) and p_0:

$$\begin{aligned} \mathcal{B}^{p_0}(s) &:= \{y \in \mathbb{R}^d : t_y^{p_0} \leq s\} = \{y \in \mathbb{R}^d : \bar{t} \leq s\}, \\ \mathcal{B}_t^{p_1}(s) &:= \{(x,t) \in \mathbb{R}^d \times [0,s] : t_{(x,t)}^{p_1} \leq s\}, \\ \mathcal{B}^{p_1 \otimes p_0}(s) &:= \{y \in \mathbb{R}^d : t_y^{p_1 \otimes p_0} \leq s\}. \end{aligned}$$

Observe that

$$\begin{aligned} \mathcal{B}^{p_1 \otimes p_0}(s) &= \bigcup_{\tau \in [0,s]} \left\{ y \in \mathbb{R}^d : t_y^{p_1 \otimes p_0} = \tau \right\} \\ &= \bigcup_{\tau \in [0,s]} \left\{ y = (\Psi_0 \otimes \Psi_1)(z,\tau,\tau) : z \in \mathcal{S}^{\tau}_{p_1 \otimes p_0} \right\} \end{aligned}$$

and that property (2.1.9) clearly holds for $\mathcal{B}^{p_0}$, $\mathcal{B}_t^{p_1}$ and $\mathcal{B}^{p_1 \otimes p_0}$ too. Furthermore, note that the sink-reaching points set at level p_1 which takes into account all the initial positions $x \in \mathbb{R}^d$ at time $t = 0$ from which the agents enter the sink $\mathcal{S}$ before s corresponds to $\mathcal{B}_0^{p_1}(s)$ (and in this case t_x in (2.1.2), corresponds to $t_{(x,0)}^{p_1}$). Clearly $\mathcal{B}_0^{p_1}(s)$ satisfies all the same properties as (2.1.8).

As in §2.1.1, we work in the set $\mathcal{G}$ of positive Radon measures equipped with the generalized Wasserstein distance $\mathcal{W}$ (see also Remark 2.1.10 below).

The continuity equation with a sink and a source (to be interpreted in a suitable formulation we will see in the next Definition 2.1.2), with a finite horizon $T > 0$, is

$$\begin{cases} \mu_t(x,t) + \operatorname{div}(\mu(x,t)b(x,t)) + \mathbb{1}_{\left\{(\mathcal{S}^t_{p_1},t):t\in[0,T]\right\}}\mu(x,t) \\ \ +\mathbb{1}_{\left\{(\mathcal{S}^t_{p_1\otimes p_0},t):t\in[0,T]\right\}}\mu(x,t) = \mathbb{1}_{\left\{(\Gamma^t_{p_0},t):t\in[0,T]\right\}}\mu(x,t), \quad (x,t)\in\mathbb{R}^d\times[0,T] \\ \mu(x,0) = \tilde{m}^{p_1}_0(x) + m^{p_0}_0(x)\big|_{\mathcal{B}^{p_0}(0)}, \qquad\qquad\qquad x\in\mathbb{R}^d. \end{cases} \tag{2.1.23}$$

Definition 2.1.2. *We say that μ is a weak solution of (2.1.23) if $\mu \in L^1([0,T],\mathcal{G})$ is such that, for any test function $\varphi \in C^\infty_c(\mathbb{R}^d\times[0,T[)$, we have*

$$\begin{aligned} &\int_{\mathbb{R}^d}\varphi(x,0)d\tilde{m}^{p_1}_0(x) + \int_{\mathcal{B}^{p_0}(0)}\varphi(y,0)dm^{p_0}_0(y) \\ &\quad + \int_0^T\int_{\mathbb{R}^d\setminus\Phi_1(\mathcal{B}^{p_1}_0(t),0,t)}(\varphi_t(x,t) + \langle D_x\varphi(x,t), b_1(x,t)\rangle)d\mu(t)(x)dt \\ &\qquad + \int_0^T\int_{\mathbb{R}^d}\mathbb{1}_{\left\{(\Gamma^t_{p_0},t):t\in[0,T]\right\}}\varphi(y,t)d\mu^{p_0,t}(t)(y) \\ &\quad + \int_0^T\int_{(\Phi_1\otimes\Phi_0)(\mathcal{B}^{p_0}(t),0,t)\setminus(\Phi_1\otimes\Phi_0)(\mathcal{B}^{p_1\otimes p_0}(t),0,t)}(\varphi_s(x,t) \\ &\qquad\qquad + \langle D_x\varphi(x,t), b_1(x,t)\rangle)d\mu(t)(x) \\ &\qquad - \int_0^T\int_{\mathbb{R}^d}\mathbb{1}_{\left\{(\mathcal{S}^t_{p_1},t):t\in[0,T]\right\}}\varphi(y,t)d\mu^{p_1,t}(t)(y) \\ &\qquad\qquad - \int_0^T\int_{\mathbb{R}^d}\mathbb{1}_{\left\{(\mathcal{S}^t_{p_1\otimes p_0},t):t\in[0,T]\right\}}\varphi(y,t)d\mu^{p_1\otimes p_0,t}(t)(y), \end{aligned}$$

where $\mu^{p_0,t}(t) = g^{p_0}(t)\mu^{p_0,t}(0)$, $\mu^{p_1,t}(t) = g^{p_1}(t)\mu^{p_1,t}(0)$ and $\mu^{p_1\otimes p_0,t}(t) = g^{p_1\otimes p_0}(t)\mu^{p_1\otimes p_0,t}(0)$. The measure $\mu^{p_0,t}(0)$ is the disintegration of $\mu(0)$ on the fibers $\mathcal{B}^{p_0,\tau} := \left\{y\in\mathbb{R}^d : t^{p_0}_y = \tau\right\}$ that compose $\mathcal{B}^{p_0}(t)$, and $g^{p_0}(\cdot)$ is the density of the measure ν^{p_0} on the indices τ of the fibers $\mathcal{B}^{p_0,\tau}$ such that

$$\begin{aligned} E\subset\mathcal{B}^{p_0}(t) \ \Rightarrow\ \mu(0)(E) &= \int_0^t \mu^{p_0,\tau}(0)(\mathcal{B}^{p_0,\tau}\cap E)d\nu^{p_0}(\tau) \\ &= \int_0^t g^{p_0}(\tau)\mu^{p_0,\tau}(0)(\mathcal{B}^{p_0,\tau}\cap E)d\tau. \end{aligned}$$

Similarly for $\mu^{p_1,t}(0)$ and $\mu^{p_1\otimes p_0,t}(0)$.

For $s\in[0,T]$, we define a new measure on $\mathbb{R}^d$

$$\tilde{\tilde{\mu}}(s) = \begin{cases} (\Phi_1\otimes\Phi_0)(\cdot,0,s)\sharp\tilde{\tilde{m}}^{p_0}_0 & \text{on } \mathbb{R}^d\setminus(\Phi_1\otimes\Phi_0)(\mathcal{B}^{p_1\otimes p_0}(s),0,s) \\ 0 & \text{otherwise} \end{cases}, \tag{2.1.24}$$

where

$$\tilde{\tilde{m}}_0^{p_0}(\cdot) := \begin{cases} m_0^{p_0}(\cdot) & \text{in } \mathcal{B}^{p_0}(s) \\ 0 & \text{otherwise} \end{cases}, \qquad \tilde{\tilde{m}}_0^{p_0} \in \mathcal{G}.$$

By the hypotheses on $\tilde{m}_0^{p_0}$, the distribution $\tilde{\tilde{m}}_0^{p_0}$ is absolutely continuous with a density which is bounded and has a compact support.

Note that the set $(\Phi_1 \otimes \Phi_0)(\mathcal{B}^{p_1 \otimes p_0}(s), 0, s)$ takes into account all the positions of the agents at level p_1, coming from level p_0, who passed through the sink $\mathcal{S}$ at least once in the time interval $[\bar{t}, T]$ and then disappeared. The measure $\tilde{\tilde{\mu}}(s)$ is indeed detected at level p_1 only for times $s \geq \bar{t}$. By hypotheses on b_0, b_1, by the properties of $\tilde{\tilde{m}}_0^{p_0}$ and by (2.1.24), we have that $\tilde{\tilde{\mu}}(s) \in \mathcal{G}$. In fact, as (2.1.10), it certainly satisfies the constraint $\int_{\mathbb{R}^d} d\tilde{\tilde{\mu}} \leq G$ because, with respect to $\tilde{\tilde{\mu}}_{p_1}(s) := (\Phi_1 \otimes \Phi_0)(\cdot, 0, s)\sharp\tilde{\tilde{m}}_0^{p_0}$, it may only lose mass through the sink $\mathcal{S}$ at level p_1. Furthermore, it is absolutely continuous with a density which is bounded and has a compact support.

Lemma 2.1.2. *For $s \in [0, T]$, consider the maps*

$$\pi^{p_0} : \mathcal{B}^{p_0}(s) \longrightarrow [0, s], \qquad \pi(y) = t_y^{p_0} \quad \textit{at level } p_0, \tag{2.1.25}$$

$$\pi_t^{p_1} : \mathcal{B}_t^{p_1}(s) \longrightarrow [0, s], \qquad \pi_t^{p_1}(x) = t_{(x,t)}^{p_1} \quad \textit{at level } p_1, \tag{2.1.26}$$

and assume that they are Lipschitz continuous with Lipschitz constants L^{p_0} and L^{p_1} respectively (see Remark 2.1.9 too). Then, the map $s \longmapsto \tilde{\tilde{\mu}}(s)$ is Lipschitz continuous in $\mathcal{G}$ (with respect to the metrics $\mathcal{W}$ (2.1.3)).

Proof. Let $s, s' \in [0, T]$, $s \geq s'$. We have

$$\begin{aligned}
\mathcal{W}(\tilde{\tilde{\mu}}(s'), \tilde{\tilde{\mu}}(s)) = \sup_{\substack{\|\varphi\|_\infty \leq 1 \\ \mathrm{Lip}(\varphi) \leq 1}} \left\{ \int_{\mathbb{R}^d} \varphi(y) d\tilde{\tilde{\mu}}(s')(y) - \int_{\mathbb{R}^d} \varphi(y) d\tilde{\tilde{\mu}}(s)(y) \right\} \\
= \sup_{\substack{\|\varphi\|_\infty \leq 1 \\ \mathrm{Lip}(\varphi) \leq 1}} \left\{ \int_{\mathbb{R}^d \setminus (\Phi_1 \otimes \Phi_0)(\mathcal{B}^{p_1 \otimes p_0}(s'), 0, s')} \varphi(y) d\tilde{\tilde{\mu}}_{p_1}(s')(y) \right. \\
\left. - \int_{\mathbb{R}^d \setminus (\Phi_1 \otimes \Phi_0)(\mathcal{B}^{p_1 \otimes p_0}(s), 0, s)} \varphi(y) d\tilde{\tilde{\mu}}_{p_1}(s)(y) \right\} \\
= \sup_{\substack{\|\varphi\|_\infty \leq 1 \\ \mathrm{Lip}(\varphi) \leq 1}} \left\{ \int_{\mathbb{R}^d} \varphi(y) d(\tilde{\tilde{\mu}}_{p_1}(s') - \tilde{\tilde{\mu}}_{p_1}(s))(y) \right. \\
+ \int_{(\Phi_1 \otimes \Phi_0)(\mathcal{B}^{p_1 \otimes p_0}(s), 0, s) \setminus (\Phi_1 \otimes \Phi_0)(\mathcal{B}^{p_1 \otimes p_0}(s'), 0, s)} \varphi(y) d\tilde{\tilde{\mu}}_{p_1}(s)(y)
\end{aligned}$$

$$
\begin{aligned}
&\qquad + \int_{(\Phi_1\otimes\Phi_0)(\mathcal{B}^{p_1\otimes p_0}(s'),0,s)} \varphi(y)d\tilde{\tilde{\mu}}_{p_1}(s)(y) \\
&\qquad - \int_{(\Phi_1\otimes\Phi_0)(\mathcal{B}^{p_1\otimes p_0}(s'),0,s')} \varphi(y)d\tilde{\tilde{\mu}}_{p_1}(s')(y)\Bigg\} \\
&= \sup_{\substack{\|\varphi\|_\infty\le 1\\ \mathrm{Lip}(\varphi)\le 1}} \Bigg\{ \int_{\mathbb{R}^d}(\varphi((\Phi_1\otimes\Phi_0)(y,0,s')) - \varphi((\Phi_1\otimes\Phi_0)(y,0,s)))d\tilde{\tilde{m}}_0^{p_0}(y) \\
&\qquad + \int_{\mathcal{B}^{p_1\otimes p_0}(s)\setminus\mathcal{B}^{p_1\otimes p_0}(s')} \varphi((\Phi_1\otimes\Phi_0)(y,0,s))d\tilde{\tilde{m}}_0^{p_0}(y) \\
&\qquad + \int_{\mathcal{B}^{p_1\otimes p_0}(s')}(\varphi((\Phi_1\otimes\Phi_0)(y,0,s)) - \varphi((\Phi_1\otimes\Phi_0)(y,0,s')))d\tilde{\tilde{m}}_0^{p_0}(y)\Bigg\} \\
&= \sup_{\substack{\|\varphi\|_\infty\le 1\\ \mathrm{Lip}(\varphi)\le 1}} \Bigg\{ \int_{\mathcal{B}^{p_0}(s)}(\varphi((\Phi_1\otimes\Phi_0)(y,0,s')) - \varphi((\Phi_1\otimes\Phi_0)(y,0,s)))dm_0^{p_0}(y) \\
&\qquad + \int_{\mathcal{B}^{p_1\otimes p_0}(s)\setminus\mathcal{B}^{p_1\otimes p_0}(s')} \varphi((\Phi_1\otimes\Phi_0)(y,0,s))d\tilde{\tilde{m}}_0^{p_0}(y) \\
&\qquad + \int_{\mathcal{B}^{p_1\otimes p_0}(s')}(\varphi((\Phi_1\otimes\Phi_0)(y,0,s)) - \varphi((\Phi_1\otimes\Phi_0)(y,0,s')))d\tilde{\tilde{m}}_0^{p_0}(y)\Bigg\} \\
&\le \sup_{\substack{\|\varphi\|_\infty\le 1\\ \mathrm{Lip}(\varphi)\le 1}} \Bigg\{ \int_{\mathcal{B}^{p_0}(s)} \|(\Phi_1\otimes\Phi_0)(y,0,s') - (\Phi_1\otimes\Phi_0)(y,0,s)\|dm_0^{p_0}(y) \\
&\qquad + \int_{\mathcal{B}^{p_1\otimes p_0}(s)\setminus\mathcal{B}^{p_1\otimes p_0}(s')} \varphi((\Phi_1\otimes\Phi_0)(y,0,s))d\tilde{\tilde{m}}_0^{p_0}(y) \\
&\qquad + \int_{\mathcal{B}^{p_1\otimes p_0}(s')} \|(\Phi_1\otimes\Phi_0)(y,0,s) - (\Phi_1\otimes\Phi_0)(y,0,s')\|d\tilde{\tilde{m}}_0^{p_0}(y)\Bigg\} \\
&\qquad \le 2GM_{b_1}|s'-s| + \|\tilde{\tilde{m}}_0^{p_0}\|_\infty \mathcal{L}^d(\mathcal{B}^{p_1\otimes p_0}(s)\setminus\mathcal{B}^{p_1\otimes p_0}(s')),
\end{aligned}
$$

where M_{b_1} is the time-Lipschitz constant for $\Phi_1\otimes\Phi_0$ in $[\bar{t},T]$ (i.e., the bound for b_1). Note that the integrals

$$\int_{\mathcal{B}^{p_0}(s)} \|(\Phi_1\otimes\Phi_0)(y,0,s') - (\Phi_1\otimes\Phi_0)(y,0,s)\|dm_0^{p_0}(y)$$

and

$$\int_{\mathcal{B}^{p_1\otimes p_0}(s')} \|(\Phi_1\otimes\Phi_0)(y,0,s) - (\Phi_1\otimes\Phi_0)(y,0,s')\|d\tilde{\tilde{m}}_0^{p_0}(y)$$

are both bounded by $GM_{b_1}|s'-s|$ since $s, s' \ge \bar{t}$ by definition of (2.1.24) (recall that $m_0^{p_0}$ is restricted to $\mathcal{B}^{p_0}(s)$), and hence the time-Lipschitz constant of $\Phi_1 \times \Phi_0$ is given by the bound for b_1. Then, to conclude, we need the Lipschitz continuity of the map $s \longmapsto \mathcal{L}^d(\mathcal{B}^{p_1 \otimes p_0}(s))$.

For this we have to prove at first the Lipschitz continuity of the map

$$\pi^{p_1 \otimes p_0} : \mathcal{B}^{p_1 \otimes p_0}(s) \longrightarrow [0, s], \qquad \pi^{p_1 \otimes p_0}(y) = t_y^{p_1 \otimes p_0}. \tag{2.1.27}$$

Let $y_1, y_2 \in \mathbb{R}^d$ such that $\pi^{p_0}(y_2) \leq \pi^{p_0}(y_1)$, and let $\bar{y}_1 = \Phi_0(y_1, 0, \pi^{p_0}(y_1))$, $\bar{y}_2 = \Phi_0(y_2, 0, \pi^{p_0}(y_2))$ and $\tilde{y}_1 = \Phi_0(y_1, 0, \pi^{p_0}(y_2))$. Observe that by definition of π^{p_0} (2.1.25), we clearly have $\pi^{p_0}(y_1), \pi^{p_0}(y_2) < +\infty$. By the Lipschitz continuity of π^{p_0} and $\pi_t^{p_1}$, we have (M_{b_0} and L_{b_0} are the bound and the Lipschitz constant of b_0 respectively)

$$\begin{aligned}
|\pi^{p_1 \otimes p_0}(y_1) - \pi^{p_1 \otimes p_0}(y_2)| &= |\pi^{p_1}_{\pi^0(y_1)}(\bar{y}_1) - \pi^{p_1}_{\pi^{p_0}(y_2)}(\bar{y}_2)| \\
&\leq L^{p_1}(\|\bar{y}_1 - \bar{y}_2\| + |\pi^{p_0}(y_1) - \pi^{p_0}(y_2)|) \\
&\leq L^{p_1}\left(\|\bar{y}_1 - \tilde{y}_1\| + \|\tilde{y}_1 - \bar{y}_2\| + L^{p_0}\|y_1 - y_2\|\right) \\
&\leq L^{p_1}\left(M_{b_0}(\pi^{p_0}(y_1) - \pi^{p_0}(y_2)) + e^{L_{b_0}\pi^{p_0}(y_2)}\|y_1 - y_2\| + L^{p_0}\|y_1 - y_2\|\right) \\
&\leq L^{p_1}\left(M_{b_0}L^{p_0}\|y_1 - y_2\| + e^{L_{b_0}T}\|y_1 - y_2\| + L^{p_0}\|y_1 - y_2\|\right) \\
&\leq L^{p_1 \otimes p_0}\|y_1 - y_2\| \quad \text{for some } L^{p_1 \otimes p_0} > 0.
\end{aligned} \tag{2.1.28}$$

Now, note that the function

$$f : \sigma \longmapsto \pi^{p_1 \otimes p_0}((\Phi_1 \otimes \Phi_0)(y, 0, \sigma)) = \pi^{p_1 \otimes p_0}(y) - \sigma, \quad \sigma \in [\bar{t}, T]$$

is such that

$$1 = |f'(0)| = |\nabla \pi^{p_1 \otimes p_0}(y) \cdot b_1(y, 0)| \leq \|\nabla \pi^{p_1 \otimes p_0}\| \|b_1\|_\infty \leq M_{b_1}\|\nabla \pi\| \quad \text{a.e.} \tag{2.1.29}$$

and then $\|\nabla \pi^{p_1 \otimes p_0}\| \geq \frac{1}{M_{b_1}} > 0$ almost everywhere. Observe that inequality (2.1.29) holds also for $\sigma \leq \bar{t}$, but we are not considering such values of σ because of definition of (2.1.24), which is detected at level p_1 only after $\bar{t}$ (in other words, the agents reach the sink Γ at time $\bar{t} < +\infty$ at level p_0, pass to the new level p_1 and start flowing with field b_1). Therefore

$$\begin{aligned}
\frac{\mathcal{L}^d(\mathcal{B}^{p_1 \otimes p_0}(s) \setminus \mathcal{B}^{p_1 \otimes p_0}(s'))}{M} &\leq \int_{\mathcal{B}^{p_1 \otimes p_0}(s) \setminus \mathcal{B}^{p_1 \otimes p_0}(s')} \|\nabla \pi^{p_1 \otimes p_0}\| dx \\
&= \int_{s'}^{s} \mathcal{H}^{d-1}\left(\left(\pi^{p_1 \otimes p_0}\right)^{-1}(\tau)\right) d\tau \\
&= \int_{s'}^{s} \mathcal{H}^{d-1}\left((\Psi_0 \otimes \Psi_1)\left(\mathcal{S}^{\tau}_{p_1 \otimes p_0}, \tau, \tau\right)\right) d\tau \leq \tilde{K}|s - s'|,
\end{aligned} \tag{2.1.30}$$

where, similarly as in the proof of Lemma 2.1.1, we used the Coarea Formula (see Theorem B.2.2) and the fact that the Hausdorff measure in the last integral is bounded by a constant $\tilde{K} > 0$ since, by hypotheses, $\Psi_0 \otimes \Psi_1$ is Lipschitz continuous and $\mathcal{S}^{\tau}_{p_1 \otimes p_0}$ is compact. Hence, the map $s \longmapsto \mathcal{L}^d(\mathcal{B}^{p_1 \otimes p_0}(s))$ is Lipschitz continuous and the thesis follows. □

Remark 2.1.9. *Due to the same reasons explained in Remark 2.1.2, at the moment it is not restrictive to assume the Lipschitz continuity of the maps* π^{p_0} *and* $\pi^{p_1}_t$ *on* $\mathcal{B}^{p_0}(s)$ *and* $\mathcal{B}^{p_1}_t(s)$. *Moreover, observe that, in the previous proof, the Lipschitz continuity of* $\pi^{p_1 \otimes p_0}$*, which we proved on* $\mathcal{B}^{p_1 \otimes p_0}(s)$*, makes sense only if the map* π^{p_0} *is finite, that is the agents reach the sink* Γ *at level* p_0 *and pass to the new level* p_1*: the points* y_1 *and* y_2 *are indeed such that* $\pi^{p_0}(y_1), \pi^{p_0}(y_2) < +\infty$ *by definition of the map* π^{p_0} (2.1.25)*, and this allows us to write* $\pi^{p_1 \otimes p_0}(y_i) = \pi^{p_1}_{\pi^{p_0}(y_i)}(\bar{y}_i)$*, where* $\bar{y}_i = \Phi_0(y_i, 0, \pi^{p_0}(y_i))$*,* $i = 1, 2$.

The candidate for solving (2.1.23) is then the following measure on $\mathbb{R}^d$

$$\bar{\mu}(s) = \tilde{\mu}(s) + \tilde{\tilde{\mu}}(s), \qquad s \in [0, T],$$

where

$$\tilde{\mu}(s) = \begin{cases} \Phi_1(\cdot, 0, s)\sharp \tilde{m}_0^{p_1} & \text{on } \mathbb{R}^d \setminus \Phi_1(\mathcal{B}_0^{p_1}(s), 0, s) \\ 0, & \text{otherwise} \end{cases}$$

is defined exactly as in (2.1.10) with $\Phi = \Phi_1$, $\mathcal{B}(s) = \mathcal{B}_0^{p_1}(s)$ and $\tilde{m}_0 = \tilde{m}_0^{p_1}$. We set $\tilde{\mu}_{p_1}(s) := \Phi_1(\cdot, 0, s)\sharp \tilde{m}_0^{p_1}$.

Remark 2.1.10. *Note that, a priori,* $\bar{\mu}$ *may not belong to* $\mathcal{G}$*. Indeed, it may happen that* $\int_{\mathbb{R}^d} d\bar{\mu} > G$ *because, at the level* p_1*, the flowing mass from the source (coming from the previous level* p_0*), detected by* $\tilde{\tilde{\mu}}(s)$*, can be greater than the mass entering the sink. To avoid this issue, it is sufficient to take the constant* G *such that the total mass at* p_1 *(i.e., the initial distribution and the flowing mass from the source, which is a datum) is less or equal than* G*. In the general case with more than one optimal stopping problem with a sink and a source, that is with more than one level (and more than one sink and one source), by solving the problem forwardly in time, the sources are data and* G *corresponds to the initial distribution at all levels.*

Moreover, observe that $\bar{\mu}$ *is absolutely continuous with a density which is bounded and has compact support and it is also Lipschitz continuous in* $\mathcal{G}$ *(since* $\tilde{\mu}(s)$ *and* $\tilde{\tilde{\mu}}(s)$ *are Lipschitz continuous as proved in Lemma 2.1.1 and Lemma 2.1.2).*

Theorem 2.1.5. *The map* $s \longmapsto \bar{\mu}(s)$ *is a weak solution of* (2.1.23).

Proof. Let $\varphi \in C_c^\infty(\mathbb{R}^d \times [0,T[)$. By Lemma (2.1.1) and Lemma (2.1.2), the map

$$s \longmapsto \int_{\mathbb{R}^d} \varphi(x,s)d\bar{\mu}(s)(x)$$

is absolutely continuous and then we have

$$\begin{aligned}
\frac{d}{ds}\int_{\mathbb{R}^d} \varphi(x,s)d\bar{\mu}(s)(x) &= \frac{d}{ds}\int_{\mathbb{R}^d} \varphi(x,s)d\tilde{\mu}(s)(x) + \frac{d}{ds}\int_{\mathbb{R}^d} \varphi(x,s)d\tilde{\tilde{\mu}}(s)(x)\\
&= \frac{d}{ds}\int_{\mathbb{R}^d \setminus \Phi_1(\mathcal{B}_0^{p_1}(s),0,s)} \varphi(x,s)d\tilde{\mu}_{p_1}(s)(x)\\
&\quad + \frac{d}{ds}\int_{\mathbb{R}^d \setminus (\Phi_1\otimes\Phi_0)(\mathcal{B}^{p_1\otimes p_0}(s),0,s)} \varphi(y,s)d\tilde{\tilde{\mu}}_{p_1}(s)(y)\\
&= \frac{d}{ds}\int_{\mathbb{R}^d} \varphi(\Phi_1(x,0,s),s)d\tilde{m}_0^{p_1}(x) - \frac{d}{ds}\int_{\mathcal{B}_0^{p_1}(s)} \varphi(\Phi_1(x,0,s),s)d\tilde{m}_0^{p_1}(x)\\
&\quad + \frac{d}{ds}\int_{\mathbb{R}^d} \varphi((\Phi_1\otimes\Phi_0)(y,0,s),s)d\tilde{\tilde{m}}_0^{p_0}(y)\\
&\quad - \frac{d}{ds}\int_{\mathcal{B}^{p_1\otimes p_0}(s)} \varphi((\Phi_1\otimes\Phi_0)(y,0,s),s)d\tilde{\tilde{m}}_0^{p_0}(y)\\
&= \int_{\mathbb{R}^d} (\varphi_s(\Phi_1(x,0,s),s) + \langle D_x\varphi(\Phi_1(x,0,s),s), b_1(\Phi_1(x,0,s),s)\rangle)d\tilde{m}_0^{p_1}(x)\\
&\quad - \frac{d}{ds}\int_{\mathcal{B}_0^{p_1}(s)} \varphi(\Phi_1(x,0,s),s)d\tilde{m}_0^{p_1}(x)\\
&\quad + \frac{d}{ds}\int_{\mathcal{B}^{p_0}(s)} \varphi((\Phi_1\otimes\Phi_0)(y,0,s),s)dm_0^{p_0}(y)\\
&\quad - \frac{d}{ds}\int_{\mathcal{B}^{p_1\otimes p_0}(s)} \varphi((\Phi_1\otimes\Phi_0)(y,0,s),s)d\tilde{\tilde{m}}_0^{p_0}(y)\\
&= \int_{\mathbb{R}^d} (\varphi_s(z,s) + \langle D_x\varphi(z,s), b_1(z,s)\rangle)d\tilde{\mu}_{p_1}(s)(z)\\
&\quad - \frac{d}{ds}\int_{\mathcal{B}_0^{p_1}(s)} \varphi(\Phi_1(x,0,s),s)d\tilde{m}_0^{p_1}(x)\\
&\quad + \frac{d}{ds}\int_{\mathcal{B}^{p_0}(s)} \varphi((\Phi_1\otimes\Phi_0)(y,0,s),s)dm_0^{p_0}(y)\\
&\quad - \frac{d}{ds}\int_{\mathcal{B}^{p_1\otimes p_0}(s)} \varphi((\Phi_1\otimes\Phi_0)(y,0,s),s)d\tilde{\tilde{m}}_0^{p_0}(y).
\end{aligned}$$

We have to compute

$$\frac{d}{ds}\int_{\mathcal{B}_0^{p_1}(s)} \varphi(\Phi_1(x,0,s),s)d\tilde{m}_0^{p_1}(x), \tag{2.1.31}$$

$$\frac{d}{ds}\int_{\mathcal{B}^{p_0}(s)}\varphi((\Phi_1\otimes\Phi_0)(y,0,s),s)dm_0^{p_0}(y), \tag{2.1.32}$$

$$\frac{d}{ds}\int_{\mathcal{B}^{p_1\otimes p_0}(s)}\varphi((\Phi_1\otimes\Phi_0)(y,0,s),s)d\tilde{\tilde{m}}_0^{p_0}(y). \tag{2.1.33}$$

The first integral (2.1.31) is computed exactly as in the proof of Theorem 2.1.1 by the Disintegration Theorem (see Remark 2.1.11). We have indeed

$$\begin{aligned}
&\frac{d}{ds}\int_{\mathcal{B}_0^{p_1}(s)}\varphi(\Phi_1(x,0,s),s)d\tilde{m}_0^{p_1}(x)\\
&\quad=\frac{d}{ds}\int_0^s\int_{\left\{x\in\mathbb{R}^d:t_{(x,0)}^{p_1}=\tau\right\}}\varphi(\Phi_1(x,0,s),s)d\tilde{m}_0^{p_1,\tau}(x)d\nu^{p_1}(\tau)\\
&\quad=\frac{d}{ds}\int_0^s\int_{\left\{x\in\mathbb{R}^d:t_{(x,0)}^{p_1}=\tau\right\}}\varphi(\Phi_1(x,0,s),s)g^{p_1}(\tau)d\tilde{m}_0^{p_1,\tau}(x)d\tau\\
&\quad=\int_{\left\{x\in\mathbb{R}^d:t_{(x,0)}^{p_1}=s\right\}}\varphi(\Phi_1(x,0,s),s)g^{p_1}(s)d\tilde{m}_0^{p_1,s}(x)\\
&+\int_{\mathcal{B}_0^{p_1}(s)}(\varphi_s(\Phi_1(x,0,s),s)+\langle D_x\varphi(\Phi_1(x,0,s),s),b_1(\Phi_1(x,0,s),s)\rangle)d\tilde{m}_0^{p_1}(x).
\end{aligned}$$

Recalling that $\left\{x\in\mathbb{R}^d:t_{(x,0)}^{p_1}=s\right\}=\Psi_1(\mathcal{S}_{p_1}^s,s,s)$ by definition, we have

$$\begin{aligned}
&\int_{\left\{x\in\mathbb{R}^d:t_{(x,0)}^{p_1}=s\right\}}\varphi(\Phi_1(x,0,s),s)g^{p_1}(s)d\tilde{m}_0^{p_1,s}(x)\\
&=\int_{\Psi_1(\mathcal{S}_{p_1}^s,s,s)}\varphi(\Phi_1(x,0,s),s)g^{p_1}(s)d\tilde{m}_0^{p_1,s}(x)=\int_{\mathcal{S}_{p_1}^s}\varphi(z,s)d\mu^{p_1,s}(s)(z),
\end{aligned}$$

where $\mu^{p_1,s}(s):=g^{p_1}(s)(\Phi_1(\cdot,0,s)\sharp\tilde{m}_0^{p_1,s})$.

The second integral (2.1.32) is computed again by the Disintegration Theorem (see Remark 2.1.11). We have

$$\begin{aligned}
&\frac{d}{ds}\int_{\mathcal{B}^{p_0}(s)}\varphi((\Phi_1\otimes\Phi_0)(y,0,s),s)dm_0^{p_0}(y)\\
&\quad=\frac{d}{ds}\int_0^s\int_{\left\{y\in\mathbb{R}^d:t_y^{p_0}=\tau\right\}}\varphi((\Phi_1\otimes\Phi_0)(y,0,s),s)dm_0^{p_0,\tau}(y)d\nu^{p_0}(\tau)\\
&\quad=\frac{d}{ds}\int_0^s\int_{\left\{y\in\mathbb{R}^d:t_y^{p_0}=\tau\right\}}\varphi((\Phi_1\otimes\Phi_0)(y,0,s),s)g^{p_0}(\tau)dm_0^{p_0,\tau}(y)d\tau
\end{aligned}$$

$$
\begin{aligned}
&= \int_{\left\{y\in\mathbb{R}^d : t_y^{p_0}=s\right\}} \varphi((\Phi_1\otimes\Phi_0)(y,0,s),s) g^{p_0}(s) dm_0^{p_0,s}(y) \\
&\qquad + \int_{\mathcal{B}^{p_0}(s)} (\varphi_s((\Phi_1\otimes\Phi_0)(y,0,s),s) \\
&\quad + \langle D_x\varphi((\Phi_1\otimes\Phi_0)(y,0,s),s), b_1((\Phi_1\otimes\Phi_0)(y,0,s),s)\rangle) dm_0^{p_0}(y).
\end{aligned}
$$

Recalling that $\left\{y\in\mathbb{R}^d : t_y^{p_0}=s\right\} = \Psi_0(\Gamma_{p_0}^s,s,s)$ by definition, we have

$$
\begin{aligned}
&\int_{\left\{y\in\mathbb{R}^d : t_y^{p_0}=s\right\}} \varphi((\Phi_1\otimes\Phi_0)(y,0,s),s) g^{p_0}(s) dm_0^{p_0,s}(y) \\
&\qquad = \int_{\Psi_0(\Gamma_{p_0}^s,s,s)} \varphi((\Phi_1\otimes\Phi_0)(y,0,s),s) g^{p_0}(s) dm_0^{p_0,s}(y) \\
&\qquad\qquad = \int_{\Gamma_{p_0}^s} \varphi(w,s) d\mu^{p_0,s}(s)(w),
\end{aligned}
$$

where $\mu^{p_0,s}(s) := g^{p_0}(s)((\Phi_1\otimes\Phi_0)(\cdot,0,s)\sharp m_0^{p_0,s})$. Notice that in the last step we used the fact that $(\Phi_1\otimes\Phi_0)(\Psi_0(\Gamma_{p_0}^s,s,s),s,s) = \Gamma_{p_0}^s$. This is true because $t_y^{p_0}=s$, and then $(\Phi_1\otimes\Phi_0)^{-1}(y,0,s) = \Psi_0(y,s,s)$ (the mass, starting from Γ, backwardly flows with field b_0).

For the third integral (2.1.33), by the Disintegration Theorem (see Remark 2.1.11) we have

$$
\begin{aligned}
&\frac{d}{ds}\int_{\mathcal{B}^{p_1\otimes p_0}(s)} \varphi((\Phi_1\otimes\Phi_0)(y,0,s),s) d\tilde{\tilde{m}}_0^{p_0}(y) \\
&= \frac{d}{ds}\int_0^s \int_{\left\{y\in\mathbb{R}^d : t_y^{p_1\otimes p_0}=\tau\right\}} \varphi((\Phi_1\otimes\Phi_0)(y,0,s),s) d\tilde{\tilde{m}}_0^{p_0,\tau}(y) d\nu^{p_1\otimes p_0}(\tau) \\
&= \frac{d}{ds}\int_0^s \int_{\left\{y\in\mathbb{R}^d : t_y^{p_1\otimes p_0}=\tau\right\}} \varphi((\Phi_1\otimes\Phi_0)(y,0,s),s) g^{p_1\otimes p_0}(\tau) d\tilde{\tilde{m}}_0^{p_0,\tau}(y) d\tau \\
&\quad = \int_{\left\{y\in\mathbb{R}^d : t_y^{p_1\otimes p_0}=s\right\}} \varphi((\Phi_1\otimes\Phi_0)(y,0,s),s) g^{p_1\otimes p_0}(s) d\tilde{\tilde{m}}_0^{p_0,s}(y) \\
&\qquad\qquad + \int_{\mathcal{B}^{p_1\otimes p_0}(s)} (\varphi_s((\Phi_1\otimes\Phi_0)(y,0,s),s) \\
&\quad + \langle D_x\varphi((\Phi_1\otimes\Phi_0)(y,0,s),s), b_1((\Phi_1\otimes\Phi_0)(y,0,s),s)\rangle) d\tilde{\tilde{m}}_0^{p_0}(y).
\end{aligned}
$$

Recalling that $\left\{y\in\mathbb{R}^d : t_y^{p_1\otimes p_0}=s\right\} = (\Psi_0\otimes\Psi_1)(\mathcal{S}_{p_1\otimes p_0}^s,s,s)$ by definition, we have

$$
\int_{\left\{y\in\mathbb{R}^d : t_y^{p_1\otimes p_0}=s\right\}} \varphi((\Phi_1\otimes\Phi_0)(y,0,s),s) g^{p_1\otimes p_0}(s) d\tilde{\tilde{m}}_0^{p_0,s}(y)
$$

$$
\begin{aligned}
&= \int_{(\Psi_0\otimes\Psi_1)(\mathcal{S}^s_{p_1\otimes p_0},s,s)} \varphi((\Phi_1\otimes\Phi_0)(y,0,s),s) g^{p_1\otimes p_0}(s) d\tilde{\tilde{m}}_0^{p_0,s}(y) \\
&\qquad = \int_{\mathcal{S}^s_{p_1\otimes p_0}} \varphi(w,s) d\mu^{p_1\otimes p_0,s}(s)(w),
\end{aligned}
$$

where $\mu^{p_1\otimes p_0,s}(s) := g^{p_1\otimes p_0}(s)((\Phi_1\otimes\Phi_0)(\cdot,0,s)\sharp\tilde{\tilde{m}}_0^{p_0,s})$.

Finally, we obtain

$$
\begin{aligned}
&\frac{d}{ds}\int_{\mathbb{R}^d}\varphi(x,s)d\bar{\mu}(s)(x) \\
&\quad = \int_{\mathbb{R}^d\setminus\Phi_1(\mathcal{B}_0^{p_1}(s),0,s)} (\varphi_s(z,s)+\langle D_x\varphi(z,s),b_1(z,s)\rangle) d\tilde{\mu}_{p_1}(s)(z) \\
&\qquad - \int_{\mathcal{S}^s_{p_1}} \varphi(z,s) d\mu^{p_1,s}(s)(z) \\
&\qquad + \int_{(\Phi_1\otimes\Phi_0)(\mathcal{B}^{p_0}(s),0,s)\setminus(\Phi_1\otimes\Phi_0)(\mathcal{B}^{p_1\otimes p_0}(s),0,s)} (\varphi_s(w,s) \\
&\qquad\qquad + \langle D_x\varphi(w,s),b_1(w,s)\rangle) d\tilde{\tilde{\mu}}_{p_1}(s)(w) \\
&\qquad + \int_{\Gamma^s_{p_0}} \varphi(w,s) d\mu^{p_0,s}(s)(w) \\
&\qquad - \int_{\mathcal{S}^s_{p_1\otimes p_0}} \varphi(w,s) d\mu^{p_1\otimes p_0,s}(s)(w) \\
&\quad = \int_{\mathbb{R}^d\setminus\Phi_1(\mathcal{B}_0^{p_1}(s),0,s)} (\varphi_s(z,s)+\langle D_x\varphi(z,s),b_1(z,s)\rangle) d\tilde{\mu}_{p_1}(s)(z) \\
&\qquad - \int_{\mathcal{S}^s_{p_1}} \varphi(z,s) d\mu^{p_1,s}(s)(z) \\
&\qquad + \int_{(\Phi_1\otimes\Phi_0)(\mathcal{B}^{p_0}(s),0,s)\setminus(\Phi_1\otimes\Phi_0)(\mathcal{B}^{p_1\otimes p_0}(s),0,s)} (\varphi_s(w,s) \\
&\qquad\qquad + \langle D_x\varphi(w,s),b_1(w,s)\rangle) d\tilde{\tilde{\mu}}_{p_1}(s)(w) \\
&\quad + \int_{\Gamma^s_{p_0}} \varphi(w,s) d\mu^{p_0,s}(s)(w) - \int_{\mathcal{S}^s_{p_1\otimes p_0}} \varphi(w,s) d\mu^{p_1\otimes p_0,s}(s)(w).
\end{aligned} \tag{2.1.34}
$$

Since $\bar{\mu}(0)=\tilde{\mu}(0)+\tilde{\tilde{\mu}}(0)=\tilde{m}_0^{p_1}+\tilde{\tilde{m}}_0^{p_0}$, integrating (2.1.34) between 0 and T we get

$$
\begin{aligned}
&\int_{\mathbb{R}^d}\varphi(x,0)d\tilde{m}_0^{p_1}(x) + \int_{\mathbb{R}^d}\varphi(y,0)d\tilde{\tilde{m}}_0^{p_0}(y) \\
&\quad + \int_0^T \int_{\mathbb{R}^d\setminus\Phi_1(\mathcal{B}_0^{p_1}(s),0,s)} (\varphi_s(z,s)+\langle D_x\varphi(z,s),b_1(z,s)\rangle) d\tilde{\mu}_{p_1}(s)(z)
\end{aligned}
$$

$$+\int_0^T\int_{\Gamma^s_{p_0}}\varphi(w,s)d\mu^{p_0,s}(s)(w)$$
$$+\int_0^T\int_{(\Phi_1\otimes\Phi_0)(\mathcal{B}^{p_0}(s),0,s)\setminus(\Phi_1\otimes\Phi_0)(\mathcal{B}^{p_1\otimes p_0}(s),0,s)}(\varphi_s(w,s)$$
$$+\langle D_x\varphi(w,s),b_1(w,s)\rangle)d\tilde{\tilde{\mu}}_{p_1}(s)(w)$$
$$-\int_0^T\int_{\mathcal{S}^s_{p_1}}\varphi(z,s)d\mu^{p_1,s}(s)(z)-\int_0^T\int_{\mathcal{S}^s_{p_1\otimes p_0}}\varphi(w,s)d\mu^{p_1\otimes p_0,s}(s)(w),$$

that is

$$\int_{\mathbb{R}^d}\varphi(x,0)d\tilde{m}^{p_1}_0(x)+\int_{\mathcal{B}^{p_0}(0)}\varphi(y,0)dm^{p_0}_0(y)$$
$$+\int_0^T\int_{\mathbb{R}^d\setminus\Phi_1(\mathcal{B}^{p_1}_0(s),0,s)}(\varphi_s(z,s)+\langle D_x\varphi(z,s),b_1(z,s)\rangle)d\tilde{\mu}_{p_1}(s)(z)$$
$$+\int_0^T\int_{\mathbb{R}^d}\mathbb{1}_{\left\{(\Gamma^s_{p_0},s):s\in[0,T]\right\}}\varphi(w,s)d\mu^{p_0,s}(s)(w)$$
$$+\int_0^T\int_{(\Phi_1\otimes\Phi_0)(\mathcal{B}^{p_0}(s),0,s)\setminus(\Phi_1\otimes\Phi_0)(\mathcal{B}^{p_1\otimes p_0}(s),0,s)}(\varphi_s(w,s)$$
$$+\langle D_x\varphi(w,s),b_1(w,s)\rangle)d\tilde{\tilde{\mu}}_{p_1}(s)(w)$$
$$-\int_0^T\int_{\mathbb{R}^d}\mathbb{1}_{\left\{(\mathcal{S}^s_{p_1},s):s\in[0,T]\right\}}\varphi(z,s)d\mu^{p_1,s}(s)(z)$$
$$-\int_0^T\int_{\mathbb{R}^d}\mathbb{1}_{\left\{(\mathcal{S}^s_{p_1\otimes p_0},s):s\in[0,T]\right\}}\varphi(w,s)d\mu^{p_1\otimes p_0,s}(s)(w).$$

□

Remark 2.1.11. *The Disintegration Theorem (see also Theorem B.2.3) for integrals* (2.1.31)*,* (2.1.32) *and* (2.1.33) *in the previous proof is applied as follows: for integral* (2.1.31)*, we set $Y=\mathcal{B}^{p_1}_0(s)$, $X=[0,s]$ and we consider the map* (2.1.26) *with $t=0$*

$$\pi^{p_1}_0:Y\longrightarrow X,\qquad \pi^{p_1}_0(x)=t^{p_1}_{(x,0)}$$

and $\nu^{p_1}=\pi^{p_1}_0\sharp\tilde{m}^{p_1}_0\in\mathcal{G}(X)$. In this way $(\pi^{p_1}_0)^{-1}(\tau)=\left\{x\in\mathbb{R}^d:t^{p_1}_{(x,0)}=\tau\right\}$ for every $\tau\in[0,s]$. Then, there exists a ν^{p_1}-almost everywhere uniquely determined family $\{\tilde{m}^{p_1,\tau}_0\}_{\tau\in[0,s]}\subset\mathcal{G}(Y)$ such that for every $f\in C^0_c(Y)$,

$$\int_Y f(y)d\tilde{m}_0(y)=\int_0^s\int_{\left\{x\in\mathbb{R}^d:t^{p_1}_{(x,0)}=\tau\right\}}f(y)d\tilde{m}^{p_1,\tau}_0(y)d\nu^{p_1}(\tau).$$

Moreover, in view of the Lipschitz continuity of the map $s \longmapsto \mathcal{L}^d(\mathcal{B}_0^{p_1}(s))$*, which is proved by the Lipschitz continuity of* $\pi_0^{p_1}$ *(see Lemma 2.1.2 and* (2.1.12) *for the calculations, recalling that we are considering just the level* p_1 *with sink* $\mathcal{S}$*), the measure* ν^{p_1} *is absolutely continuous on* X *with a* L^∞ *density denoted by* g^{p_1}*.*

For integral (2.1.32)*, we set* $Y = \mathcal{B}^{p_0}(s)$*,* $X = [0, s]$ *and we consider the map* (2.1.25)

$$\pi^{p_0} : Y \longrightarrow X, \qquad \pi^{p_0}(y) = t_y^{p_0}$$

and $\nu^{p_0} = \pi^{p_0} \sharp m_0^{p_0} \in \mathcal{G}(X)$*. In this way* $(\pi^{p_0})^{-1}(\tau) = \left\{ y \in \mathbb{R}^d : t_y^{p_0} = \tau \right\}$ *for every* $\tau \in [0, s]$*. Then, there exists a* ν^{p_0}*-almost everywhere uniquely determined family* $\{m_0^{p_0,\tau}\}_{\tau \in [0,s]} \subset \mathcal{G}(Y)$ *such that for every* $f \in C_c^0(Y)$*,*

$$\int_Y f(y) dm_0^{p_0}(y) = \int_0^s \int_{\left\{y \in \mathbb{R}^d : t_y^{p_0} = \tau\right\}} f(y) dm_0^{p_0,\tau}(y) d\nu^{p_0}(\tau).$$

Moreover, in view of the Lipschitz continuity of the map $s \longmapsto \mathcal{L}^d(\mathcal{B}^{p_0}(s))$*, which is proved by the Lipschitz continuity of* π^{p_0} *(see Lemma 2.1.2 and* (2.1.12) *for the calculations, recalling that we are considering just the level* p_0 *with sink* Γ*), the measure* ν^{p_0} *is absolutely continuous on* X *with a* L^∞ *density denoted by* g^{p_0}*.*

For integral (2.1.33)*, we set* $Y = \mathcal{B}^{p_1 \otimes p_0}(s)$*,* $X = [0, s]$ *and we consider the map* (2.1.27)

$$\pi^{p_1 \otimes p_0} : Y \longrightarrow X, \qquad \pi^{p_1 \otimes p_0}(y) = t_y^{p_1 \otimes p_0}$$

and $\nu^{p_1 \otimes p_0} = \pi^{p_1 \otimes p_0} \sharp \tilde{\tilde{m}}_0^{p_0} \in \mathcal{G}(X)$*. In this way*

$$(\pi^{p_1 \otimes p_0})^{-1}(\tau) = \left\{ y \in \mathbb{R}^d : t_y^{p_1 \otimes p_0} = \tau \right\} \quad \textit{for every } \tau \in [0, s].$$

Then, there exists a $\nu^{p_1 \otimes p_0}$*-almost everywhere uniquely determined family* $\{m_0^{p_0,\tau}\}_{\tau \in [0,s]} \subset \mathcal{G}(Y)$ *such that for every* $f \in C_c^0(Y)$*,*

$$\int_Y f(y) d\tilde{\tilde{m}}_0^{p_0}(y) = \int_0^s \int_{\left\{y \in \mathbb{R}^d : t_y^{p_1 \otimes p_0} = \tau\right\}} f(y) d\tilde{\tilde{m}}_0^{p_0,\tau}(y) d\nu^{p_1 \otimes p_0}(\tau).$$

Moreover, in view of (2.1.30)*, that is the Lipschitz continuity of the map* $s \longmapsto \mathcal{L}^d(\mathcal{B}^{p_1 \otimes p_0}(s))$*, the measure* $\nu^{p_1 \otimes p_0}$ *is absolutely continuous on* X *with a* L^∞ *density denoted by* $g^{p_1 \otimes p_0}$*.*

Theorem 2.1.6. *The continuity equation* (2.1.23) *has a unique solution given by* $s \longmapsto \bar{\mu}(s)$*.*

Proof. Let $\varphi \in C^\infty(\mathbb{R}^d)$. The proof of the cases with supp $(\varphi) \subset \mathbb{R}^d \setminus \Phi_1(\mathcal{B}_0^{p_1}(t), 0, t)$ for every $t \le T$ and supp $(\varphi) \subset \Phi_1(\mathcal{B}_0^{p_1}(t), 0, t)$ for every $t \le T$, goes basically as the one of Theorem 2.1.2.

Then suppose that supp $(\varphi) \subset \mathbb{R}^d \setminus (\Phi_1 \otimes \Phi_0)(\mathcal{B}^{p_1 \otimes p_0}(t), 0, t)$ for every $t \le T$. Fix $t \le T$ and let us consider the map

$$w : \mathbb{R}^d \times [0, t] \longrightarrow \mathbb{R}, \qquad w(y, s) := \varphi((\Phi_1 \otimes \Phi_0)(y, 0, t - s)). \tag{2.1.35}$$

Therefore, w is Lipschitz continuous in both variables $(y, s) \in \mathbb{R}^d \times [0, t]$ with $\mathrm{supp}(w) \subset (\mathbb{R}^d \setminus (\Phi_1 \otimes \Phi_0)(\mathcal{B}^{p_1 \otimes p_0}(s), 0, s)) \times [0, t]$. Moreover, by (2.1.22) we have

$$\varphi(x) = w((\Psi_0 \otimes \Psi_1)(x, t-s, t), s) = \varphi((\Phi_1 \otimes \Phi_0)((\Psi_0 \otimes \Psi_1)(x, t-s, t), 0, t-s))$$

and, recalling that $(\Psi_0 \otimes \Psi_1)(x, t, t)$ is the solution of (2.1.21) with $\tau = t$, the function w satisfies

$$\begin{aligned} 0 = \frac{d}{ds}\varphi(x) &= w_s((\Psi_0 \otimes \Psi_1)(x, t - s, t), s) \\ &+ \langle D_x w((\Psi_0 \otimes \Psi_1)(x, t - s, t), s), b_1((\Psi_0 \otimes \Psi_1)(x, t - s, t), s)\rangle \quad \text{a.e.} \end{aligned}$$

and hence, in general,

$$w_s(y, s) + \langle D_x w(y, s), b_1(y, s)\rangle = 0 \quad \text{a.e. in } \mathbb{R}^d \times]0, t[.$$

Using w as a test function for a generic μ satisfying Definition 2.1.2, for almost all $s \le t$ we have

$$\begin{aligned} \frac{d}{ds}\int_{\mathbb{R}^d} w(y, s)d\mu(s)(y) &= \int_{\mathbb{R}^d} w_s(y, s)d\mu(s)(y) + \int_{\mathbb{R}^d} w(y, s)d\mu_s(s)(y) \\ &= \int_{\mathbb{R}^d} (-\langle D_x w(y, s), b_1(y, s)\rangle + \langle D_x w(y, s), b_1(y, s)\rangle)d\mu(s)(y) = 0 \end{aligned}$$

since $\mathrm{supp}(w) \subset (\mathbb{R}^d \setminus (\Phi_1 \otimes \Phi_0)(\mathcal{B}^{p_1 \otimes p_0}(s), 0, s)) \times [0, t]$, which implies

$$\begin{aligned} \int_{\Gamma^s_{p_0}} w(y, s)d\mu^{p_0,s}(s)(y) &= \int_{\mathcal{S}^s_{p_1 \otimes p_0}} w(y, s)d\mu^{p_1 \otimes p_0,s}(s)(y) \\ &= \int_{(\Phi_1 \otimes \Phi_0)(\mathcal{B}^{p_1 \otimes p_0}(s),0,s)} (w_s(y, s) + \langle D_x w(y, s), b_1(y, s)\rangle)d\mu(s)(y) = 0. \end{aligned}$$

Therefore, integrating between 0 and t, we get

$$\int_{\mathbb{R}^d} w(y, t)d\mu(t)(y) = \int_{\mathbb{R}^d} w(y, 0)d\mu(0)(y)$$

and then

$$
\begin{aligned}
\int_{\mathbb{R}^d} \varphi(y) d\mu(t)(y) &= \int_{\mathbb{R}^d} \varphi((\Phi_1 \otimes \Phi_0)(y, 0, t-0)) dm_0^{p_0}(y)\big|_{\mathcal{B}^{p_0}(t-0)} \\
&= \int_{\mathbb{R}^d} \varphi((\Phi_1 \otimes \Phi_0)(y, 0, t)) dm_0^{p_0}(y)\big|_{\mathcal{B}^{p_0}(t)},
\end{aligned}
$$

which shows that

$$
\mu(t) = (\Phi_1 \otimes \Phi_0)(\cdot, 0, t) \sharp \tilde{\tilde{m}}_0^{p_0} \quad \text{on } \mathbb{R}^d \setminus (\Phi_1 \otimes \Phi_0)(\mathcal{B}^{p_1 \otimes p_0}(t), 0, t).
$$

Now we have to prove that $\mu(t) = 0$ on $(\Phi_1 \otimes \Phi_0)(\mathcal{B}^{p_1 \otimes p_0}(t), 0, t)$, that is

$$
\int_{\mathbb{R}^d} \varphi(x) d\mu(t)(x) = 0
$$

for any $\varphi \in C^\infty(\mathbb{R}^d)$ with $\operatorname{supp}(\varphi) \subset (\Phi_1 \otimes \Phi_0)(\mathcal{B}^{p_1 \otimes p_0}(t), 0, t)$ for every $t \leq T$. Fix again $t \leq T$ and let us consider the map (2.1.35). Then, proceeding as before, we obtain

$$
w_s(y, s) + \langle D_x w(y, s), b_1(y, s) \rangle = 0 \quad \text{a.e. in } \mathbb{R}^d \times]0, t[.
$$

Using w as a test function for a generic μ satisfying Definition 2.1.2, for almost all $s \leq t$ we have

$$
\begin{aligned}
\frac{d}{ds} \int_{\mathbb{R}^d} w(y, s) d\mu(s)(y) &= \int_{\mathbb{R}^d} w_s(y, s) d\mu(s)(y) + \int_{\mathbb{R}^d} w(y, s) d\mu_s(s)(y) \\
&= \int_{\mathbb{R}^d} (-\langle D_x w(y, s), b_1(y, s) \rangle + \langle D_x w(y, s), b_1(y, s) \rangle) d\mu(s)(y) \\
&\quad + \int_{(\Phi_1 \otimes \Phi_0)(\mathcal{B}^{p_0}(s), 0, s)} (w_s(y, s) + \langle D_x w(y, s), b_1(y, s) \rangle) d\mu(s)(y) \\
&\quad - \int_{(\Phi_1 \otimes \Phi_0)(\mathcal{B}^{p_1 \otimes p_0}(s), 0, s)} (w_s(y, s) + \langle D_x w(y, s), b_1(y, s) \rangle) d\mu(s)(y) \\
&\quad + \int_{\Gamma_{p_0}^s} w(y, s) d\mu^{p_0, s}(s)(y) - \int_{\mathcal{S}_{p_1 \otimes p_0}^s} w(y, s) d\mu^{p_1 \otimes p_0, s}(s)(y) \\
&= \int_{(\Phi_1 \otimes \Phi_0)(\mathcal{B}^{p_0}(s), 0, s)} (w_s(y, s) + \langle D_x w(y, s), b_1(y, s) \rangle) d\mu(s)(y) \\
&\quad - \int_{(\Phi_1 \otimes \Phi_0)(\mathcal{B}^{p_1 \otimes p_0}(s), 0, s)} (w_s(y, s) + \langle D_x w(y, s), b_1(y, s) \rangle) d\mu(s)(y) \\
&\quad + \int_{\Gamma_{p_0}^s} w(y, s) d\mu^{p_0, s}(s)(y) - \int_{\mathcal{S}_{p_1 \otimes p_0}^s} w(y, s) d\mu^{p_1 \otimes p_0, s}(s)(y).
\end{aligned}
$$

Now, observe that

$$
\begin{aligned}
&\int_{\Gamma^s_{p_0}} w(y,s)d\mu^{p_0,s}(s)(y) \\
&\qquad + \int_{(\Phi_1\otimes\Phi_0)(\mathcal{B}^{p_0}(s),0,s)} (w_s(y,s) + \langle D_x w(y,s), b_1(y,s)\rangle) d\mu(s)(y) \\
&\qquad + \int_{(\Phi_1\otimes\Phi_0)(\mathcal{B}^{p_1\otimes p_0}(s),0,s)} (w_s(y,s) + \langle D_x w(y,s), b_1(y,s)\rangle) d\mu(s)(y) \\
&\qquad\qquad + \int_{\mathcal{S}^s_{p_1\otimes p_0}} w(y,s)d\mu^{p_1\otimes p_0,s}(s)(y) \\
&\qquad = \int_{\left\{y\in\mathbb{R}^d : t^{p_0}_y = s\right\}} w((\Phi_1\otimes\Phi_0)(y,0,s),s) g^{p_0}(s) d\mu^{p_0,s}(0)(y) \\
&\qquad\qquad + \int_{\mathcal{B}^{p_0}(s)} (w_s((\Phi_1\otimes\Phi_0)(y,0,s),s) \\
&\qquad + \langle D_x w((\Phi_1\otimes\Phi_0)(y,0,s),s), b_1((\Phi_1\otimes\Phi_0)(y,0,s),s)\rangle) d\mu(0)(y) \\
&\qquad + \int_{\left\{y\in\mathbb{R}^d : t^{p_1\otimes p_0}_y = s\right\}} w((\Phi_1\otimes\Phi_0)(y,0,s),s) g^{p_1\otimes p_0}(s) d\mu^{p_1\otimes p_0,s}(0)(y) \\
&\qquad\qquad + \int_{\mathcal{B}^{p_1\otimes p_0}(s)} (w_s((\Phi_1\otimes\Phi_0)(y,0,s),s) \\
&\qquad + \langle D_x w((\Phi_1\otimes\Phi_0)(y,0,s),s), b_1((\Phi_1\otimes\Phi_0)(y,0,s),s)\rangle) d\mu(0)(y) \\
&\qquad = \frac{d}{ds}\int_{\mathcal{B}^{p_0}(s)} w((\Phi_1\otimes\Phi_0)(y,0,s),s) d\mu(0)(y) \\
&\qquad\qquad + \frac{d}{ds}\int_{\mathcal{B}^{p_1\otimes p_0}(s)} w((\Phi_1\otimes\Phi_0)(y,0,s),s) d\mu(0)(y).
\end{aligned}
$$

Then, by (2.1.35) and the semigroup property of the flow $\Phi_1\otimes\Phi_0$, since $\mu(0) = \tilde{\tilde{m}}^{p_0}_0$ we obtain

$$
\begin{aligned}
\frac{d}{ds}\int_{\mathbb{R}^d} w(y,s)d\mu(s)(y) &= \frac{d}{ds}\int_{\mathcal{B}^{p_0}(s)} \varphi((\Phi_1\otimes\Phi_0)(y,0,t)) dm^{p_0}_0(y) \\
&\qquad - \frac{d}{ds}\int_{\mathcal{B}^{p_1\otimes p_0}(s)} \varphi((\Phi_1\otimes\Phi_0)(y,0,t)) d\tilde{\tilde{m}}^{p_0}_0(y)
\end{aligned}
$$

and hence, integrating between 0 and t,

$$
\begin{aligned}
\int_{\mathbb{R}^d} w(y,t)d\mu(t)(y) &= \int_{\mathbb{R}^d} w(y,0)d\mu(0)(y) \\
&\qquad + \int_0^t \frac{d}{ds}\int_{\mathcal{B}^{p_0}(s)} \varphi((\Phi_1\otimes\Phi_0)(y,0,t)) dm^{p_0}_0(y)
\end{aligned}
$$

$$-\int_0^t \frac{d}{ds}\int_{\mathcal{B}^{p_1\otimes p_0}(s)} \varphi((\Phi_1\otimes\Phi_0)(y,0,t))d\tilde{\tilde{m}}_0^{p_0}(y).$$

Therefore

$$\int_{\mathbb{R}^d}\varphi(y)d\mu(t)(y) = \int_{\mathbb{R}^d}\varphi((\Phi_1\otimes\Phi_0)(y,0,t))d\tilde{\tilde{m}}_0^{p_0}(y)$$
$$+\int_{\mathcal{B}^{p_0}(t)}\varphi((\Phi_1\otimes\Phi_0)(y,0,t))dm_0^{p_0}(y) - \int_{\mathcal{B}^{p_0}(0)}\varphi((\Phi_1\otimes\Phi_0)(y,0,t))dm_0^{p_0}(y)$$
$$-\int_{\mathcal{B}^{p_1\otimes p_0}(t)}\varphi((\Phi_1\otimes\Phi_0)(y,0,t))d\tilde{\tilde{m}}_0^{p_0}(y)$$
$$+\int_{\mathcal{B}^{p_1\otimes p_0}(0)}\varphi((\Phi_1\otimes\Phi_0)(y,0,t))d\tilde{\tilde{m}}_0^{p_0}(y).$$

Since $\mathrm{supp}(\varphi)\subset(\Phi_1\otimes\Phi_0)(\mathcal{B}^{p_1\otimes p_0}(t),0,t)$ (and $\tilde{\tilde{m}}_0^{p_0}=0$ in $\Gamma=\mathcal{B}^{p_0}(0)$ and $\mathcal{S}=\mathcal{B}^{p_1\otimes p_0}(0)$), the thesis follows. $\square$

Remark 2.1.12. *The case of a possible dependence of the fields b_1, b_2 on the measure is straightforward and very similar to the one with just a sink in §2.1.2. Indeed, as in Theorem* (2.1.3)*, the essential hypothesis is the Lipschitz continuity of the maps $\pi^0[\mu]$ and $\pi_t^{p_1}[\mu]$ uniformly w.r.t. $\mu\in C^0([0,T],\mathcal{G})$. This, with similar calculations as* (2.1.28)*, allows to prove the Lipschitz continuity of the map $\pi^{p_1\otimes p_0}[\mu]$ uniformly w.r.t. μ, which gives in turn the Lipschitz continuity of the map $s\longmapsto\tilde{\tilde{\mu}}(s)$ uniformly w.r.t. μ for applying Ascoli-Arzelà Theorem.*

Chapter 3

Mean-field type optimal visiting problems on networks

In this chapter, motivated by the optimal visiting problem in Ch. 1, we investigate a pure switching mean-field game model on a network, where both a decisional and a switching time-variable are controls at disposal of the agents for what concerns, respectively, the instant to decide and to perform the switch. The presence of such time variables gives to the problem a dynamical feature, which, a priori, is not accounted for due to the absence of a controlled trajectory. Every switch between the nodes of the network corresponds to a flip from 0 to 1 of one component of the string $p = (p_1, \dots, p_n)$ which, as in §1.1, possibly represents the visited targets, being labeled by $i = 1, \dots, n$. The goal is to reach the final string $(1, \dots, 1)$ (i.e., to visit all the targets) within a fixed final time T, minimizing a switching cost also depending on the congestion on the nodes. In particular, after introducing the problem, we show the existence of a suitable approximated ε-mean-field equilibrium and then we address the limit as $\varepsilon \to 0$.

3.1 A time-dependent optimal switching problem on network

Let $\{\mathcal{T}_j\}_{j=1,\dots,N} \subset \mathbb{R}^d$ be a collection of N targets of the optimal visiting problem as in Ch. 1, §1.1. We consider the set of the N-strings $p = (p^1, p^2, \dots, p^N) \in \mathcal{I} = \{0,1\}^N$, which we detect as the nodes of our network. We recall that $p^i = 1$ means that the target $\mathcal{T}_i$ has already been visited and viceversa for $p^i = 0$. The node $(1, 1, \ldots, 1)$ is the final destination and, once reached, the game ends.

Quite differently from §1.1.1, here, at every switch, just one component of p may change and it can do that only from 0 to 1. Such a component corresponds to the visited target. For example, for $N = 3$, if $p = (1, 0, 0)$, $p' = (1, 0, 1)$, $p'' = (0, 1, 1)$ and $p''' = (1, 1, 1)$, then from p we can not switch to p'' otherwise we lose the information that the first target has been already visited. Moreover, we can not switch to p''' directly since, as we said, at every switch just one component flips.

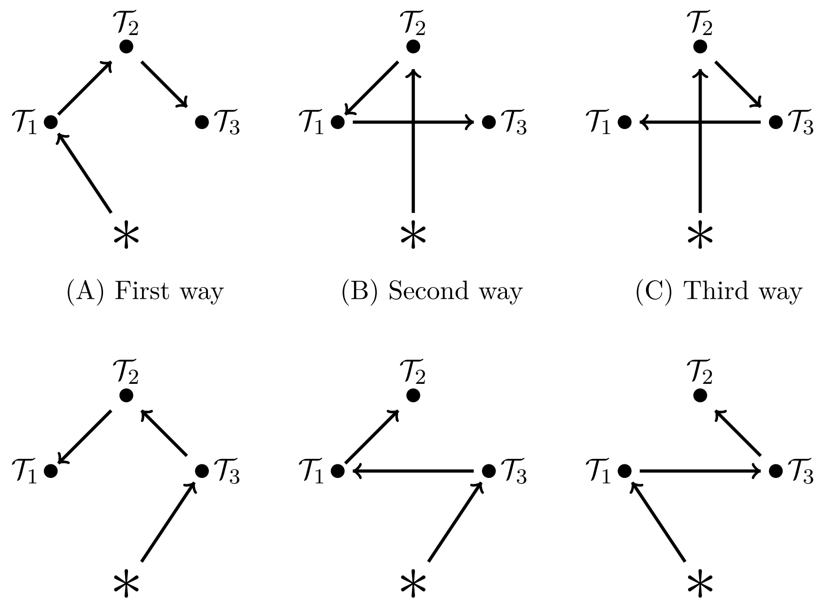
Hence, to any $p \in \mathcal{I}$ we associate the number k_p given by the sum of the components of p, that is $k_p = p^1 + \ldots + p^N$. In other words, k_p is the number of "1" in p, that is the number of the visited targets. Then, as in §1.1.1, for any $p \in \mathcal{I}$ we denote by $\mathcal{I}_p$ the set of all possible new variables (nodes) in $\mathcal{I}$ after a switch from p:

$$\begin{aligned}\mathcal{I}_p := \{\tilde{p} \in \mathcal{I} : &\text{ for every } i = 1, \ldots, N, \\ &\tilde{p}^i = p^i + 1 \;\text{ if }\; p^i \neq 1 \;\text{ and }\; k_{\tilde{p}} = k_p + 1\}.\end{aligned}$$

Clearly $\mathcal{I}_{\bar{p}} = \emptyset$, where $\bar{p} = (1, 1, \ldots, 1)$.

Example 3.1.1. *For $N = 3$ targets, all the possible ways to visit them are $N! = 3! = 6$ as we can see in Figure 3.1. The corresponding direct network is represented in Figure 3.2, where $p_o = (0, 0, 0)$ is the origin and $\bar{p} = (1, 1, 1)$ is the final destination. We then have for example $\mathcal{I}_{p_o} = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ and $\mathcal{I}_{\tilde{p}=(0,0,1)}\{(1, 0, 1), (0, 1, 1)\}$.*

The optimal visiting problem can be seen then as the search for an optimal path from an origin node p to $\bar{p}$, which must be performed within a fixed final time $T > 0$. However, we will assume that an agent at the time T may be still on an intermediate node and then, in that case, it will pay a final cost. Hence, similarly to §1.1.1, for an agent on the origin node $p \neq \bar{p}$ at time $t < T$, the number of the admissible subsequent switches is at most $N - \sum_i p^i \leq N$. The control at disposal of an agent on p at time t is then: the number of switches $0 \leq r \leq N - \sum_i p^i$; the decision/switching instants $\sigma = (t = t_0 < t_1 < t_2 < \ldots < t_r \leq T)$ and the switching path π given by the sequence of nodes $p = p_0, p_1, \ldots, p_r$, satisfying $p_1 \in \mathcal{I}_{p_0}$, $p_{i+1} \in \mathcal{I}_{p_i}$, $i = 0, 1, \ldots, r - 1$. We assume that the choice $1 \leq r < N - \sum_i p^i$ requires that $t_r = T$ and obviously $p_r \neq \bar{p}$ (because the number of switches r is not sufficient in order to reach $\bar{p}$ from $p = p_0$). Moreover, if the choice is $r = 0$, then, necessarily, either $p = p_0 \neq \bar{p}$ and $t = t_0 = T$ (that is the time is



(A) First way (B) Second way (C) Third way

(D) Fourth way (E) Fifth way (F) Sixth way

Figure 3.1: The six possible ways to visit all the three targets

already over) or $p = p_0 = \bar{p}$ and $t = t_0 \leq T$ (that is the agent may still have time at disposal but instead no more switches: it is already on $\bar{p}$). In particular, this implies that an agent can not decide to permanently stand still on a node p along a switching path unless $p = p_r = \bar{p}$ (or $t_r = T$). To resume, the control at disposal of the agent, which is on p at time t, is a triple as

$$(r, \sigma, \pi) = (r, t_0, t_1, \ldots, t_r, p_0, p_1, \ldots, p_r), \tag{3.1.1}$$

where $t_0 = t$ and $p_0 = p$. Actually, it is the switching evolution inside the network at disposal of the agent with constraints as specified here above. For example, referring to the network in Figure 3.2, the following switching evolutions/controls are admissible

$$\big(2, t_0, t_1, t_2 \leq T, (0,1,0), (1,1,0), (1,1,1)\big),$$
$$\big(2, t_0, t_1, t_2 = T, (0,0,0), (1,0,0), (1,0,1)\big),$$

whereas the following ones are not admissible

$$\big(0, t_0 < T, (1,1,0)\big), \quad \big(2, t_0, t_1, t_2 < T, (0,0,0), (1,0,0), (1,0,1)\big).$$

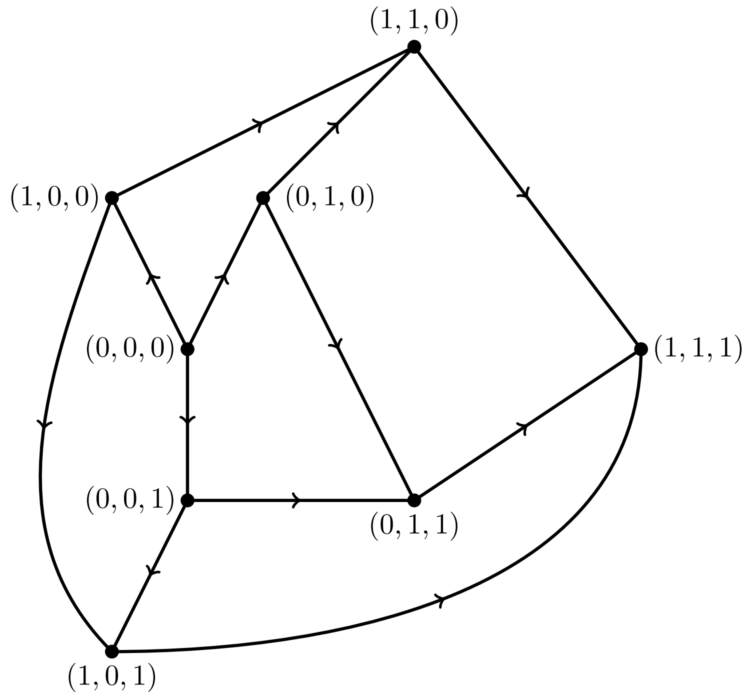


Figure 3.2: The direct network corresponding to $N = 3$ targets

In particular, $t_0 \dots, t_{r-1}$ are seen as decision instants and $t_1, \dots, t_r$ are seen as switching instants. That is the agent at time $t_i \in \{t_0, \dots t_{r-1}\}$ decides to switch from p to p_{i+1} and to perform such a switch at the time $t_{i+1} \in \{t_1, \dots, t_r\}$. Note that $t_1, \dots t_{r-1}$ are both decision and switching instants, and this means that the decision about the next switch occurs exactly at the actual switching time.

The cost to be minimized is (note that by the argumentation above if $p \neq \bar{p}$ and $t < T$, then necessarily $r \geq 1$)

$$J(p, t, (r, \sigma, \pi), \rho) = \begin{cases} \sum_{i=1}^{r} C(p_{i-1}, p_i, t_{i-1}, t_i, \rho) + \tilde{C}(p_r, t_r) & \text{if } p \neq \bar{p}, t < T \\ \tilde{C}(p, t) & \text{if } p = \bar{p} \text{ or} \\ & (p \neq \bar{p}, t = T) \end{cases} \tag{3.1.2}$$

where:

– $\rho = \left(\rho_0, \dots, \rho_{(2^N-1)}\right) \in L^2([0, T], [0, 1])^{2^N}$ is a (2^N)-uple of L^2 functions $\rho_j : [0, T] \longrightarrow [0, 1]$. Here we are using a possible enumeration of the nodes, and every $\rho_j(t)$ represents the mass of the agents at the j-node at time t. In particular, in the optimal visiting problem in §1.1.1, this would give the mass of agents with the same remaining targets to be visited as detected by the positions of the zeros in the string representing the node.

–

$$\begin{aligned} C : \mathcal{D} \subset \mathcal{I} \times \mathcal{I} \times [0,T] \times]0,T] \times L^2([0,T],[0,1])^{2^N} &\longrightarrow [0,+\infty[\\ (p,p',t,\tau,\rho) &\longmapsto C(p,p',t,\tau,\rho) \end{aligned}$$

is (for a suitable domain $\mathcal{D}$) the cost function, that is the cost that an agent incurs when, at the (decision) time t, being on the node p, decides that it will switch to a new node $p' \in \mathcal{I}_p$ at the (switching) time $\tau > t$. We assume that

(i) for every $(p,p') \in \mathcal{I} \times \mathcal{I}_p$ and $\tau \in]0,T]$, the map $(t,\rho) \longmapsto C(p,p',t,\tau,\rho)$ is bounded and Lipschitz continuous in $[0,\tau-h] \times L^2([0,T],[0,1])^{2^N}$, for all sufficiently small $h > 0$ and independently of τ, that is, there exists $L > 0$, depending only on h, such that

$$\begin{aligned} &|C(p,p',t',\tau,\rho') - C(p,p',t'',\tau,\rho'')| \\ &\qquad \le L\left(|t'-t''| + \|\rho'-\rho''\|_{L^2([0,T],[0,1])}\right); \end{aligned}$$

(ii) for every fixed ρ, $(p,p') \in \mathcal{I} \times \mathcal{I}_p$ and $t \in [0,T]$, C is decreasing in $\tau \in]t,T]$ and $\lim_{\tau \to t^+} C(p,p',t,\tau,\rho) = +\infty$;

(iii) $C(p,p,\cdot,\cdot) = 0$ for every $p \in \mathcal{I}$ and $C(\cdot,\cdot,T,T) = 0$. These assumptions correspond to the cases when the agent is on $p = p_r = \bar{p}$ and t_r is not necessarily T and when $t_r = T$ but the agent is on $p = p_r \neq \bar{p}$, and moreover give some kind of continuity of (3.1.2).

– The cost $\tilde{C}$ is bounded and Lipschitz continuous in time and it represents the final cost that an agent incurs at the end of the switching path (p_r, t_r). For example

  – if $t_r = T$, it depends on the number of the zeros in p_r (that is the number of the remaining targets to be visited);

  – if $p_r = \bar{p}$, it depends on the remaining time $T - t_r$ (that is the agent is penalized if $\bar{p}$ is obtained before T);

– if $p_r = \bar{p}$ and $t_r = T$, then it is null.

Definition 3.1.1. *Let $p \in \mathcal{I}$, $p' \in \mathcal{I}_p$, and $t < T$ be fixed. We say that the switch from p to p' with decision instant t optimally generates $\tau \in]t, T]$ as switching instant if there exists a control $(\bar{r}, \bar{\sigma}, \bar{\pi})$, with $\bar{r} \geq 1$, $\bar{\sigma} = (t_0 = t, t_1 = \tau, t_2, \dots, t_{\bar{r}})$ and $\bar{\pi} = (p_0 = p, p_1 = p', p_2, \dots, p_{\bar{r}})$, which minimizes the cost J among all controls (r, σ, π) such that $r \geq 1$, $\sigma = (t_0 = t, t_1, \dots, t_r)$, $\pi = (p_0 = p, p_1 = p', p_2, \dots, p_r)$. In other words: if whenever an agent on p at the time t decides to switch to p' (independently of the optimality of such a choice), then τ is an optimal choice as switching instant.*

Note that the optimally generated τ may be not unique. Hence the function $\varphi := \varphi_{p,p'} : t \longmapsto \varphi_{p,p'}(t) = \tau$ may be multivalued.

Moreover, other modeling assumptions are the following:

(iv) if at the decision time t, an agent on a node p chooses the switching time τ in order to switch to p', then, in the time interval $[t, \tau[$, it is assumed that such an agent continues to concur to the total mass present on the node p (coherently with the fact that the switch will occur at time τ and hence the agent will be on p in the time interval $[t, \tau[$). However, the agent can not change its decision (switching to p' at time τ) or take another decision in the time interval $]t, \tau[$. In other words, in the time interval $]t, \tau[$ it must stay on p;

(v) for the switching from p to p', if we have two different decision times t_1, t_2 with $t_1 < t_2$, which optimally generate the switching times $\tau_1, \tau_2 < T$ respectively (see Definition 3.1.1), then $\tau_1 < \tau_2$.

Assumption (iv) suggests the following useful definition

Definition 3.1.2. *An agent which is on p at time t and uses the control*

$$(r, t_0 = t, t_1, \dots, t_r, p_0 = p, p_1, \dots, p_r)$$

is called a decision-making agent at the decision instants $t_0, \dots, t_{r-1}$. Actually, since there is no incoming flow in our network (all the agents are already present at $t = 0$), all the agents are decision-making at $t = 0$. In particular, any single agent will take a new decision, mandatory, at time τ when it will switch to the new node; in other words: all agents are decision-making at $t = 0$ and they will return to be decision-making again exactly when, and only when, they switch to a new node.

Remark 3.1.1. *Assumption (ii) means that, if the switching time is too much close to the corresponding decision time, then the agent pays an high cost.*

The second part of assumption (iv) (the agent can not change the decision in $[t,\tau[$) is certainly due to the discrete feature of the time-dependent component σ of the global control (r,σ,π), but it may also be justified by a possible overlying optimal visiting problem (as the one in §1.1.1) where, when an agent is moving from one target to another, then, under some assumptions, it is not optimal to change destination or to come back to the previous node (see also [2]).

From assumption (v), it follows (v′): any optimal switching time less than T originates from a unique decision time. This can be also directly proved by assuming further hypotheses (see Remark 3.1.3). Moreover, suppose that the decision time t optimally generates the switching times τ_1, τ_2 with $\tau_1 < \tau_2$ for the switching from p to p'. Then, in view of assumption (v), in the time interval $[\tau_1, \tau_2[$ only the agents with decision time t can switch from p to p'. More generally, if we define $\tau^- := \inf_\tau\{\tau \text{ is optimal for } t\}$ and $\tau^+ := \sup_\tau\{\tau \text{ is optimal for } t\}$, in the time interval $[\tau^-, \tau^+[$, only the agents with decision time t can switch from p to p'. Hence, we can consider the function $\varphi : t \longmapsto \tau$, giving the optimal switching instant τ for the decisional instant t, as a maximal monotone graph filling the jumps by vertical segments, and so, in this case, φ is a multivalued function.

All the previous assumptions and arguments can be justified by a possible overlying optimal visiting problem, similar to the one in §1.1.1, with suitable energy and congestion costs (see [2] too). See also Remark 3.1.3.

The value function of the problem is

$$V(p,t,\rho) = \inf_{(r,\sigma,\pi)} J(p,t,(r,\sigma,\pi),\rho) \tag{3.1.3}$$

and a control (r,σ,π) is said to be optimal for (p,t) if

$$V(p,t,\rho) = J(p,t,(r,\sigma,\pi),\rho).$$

Definition 3.1.3. *Let $p \neq \bar{p}$, $t \in [0,T[$ and $\tau \in]t,T]$ be fixed. We say that τ is optimal for $V(p,t,\rho)$ if there exists a control $(\bar{r},\bar{\sigma},\bar{\pi})$ with $\bar{r} \geq 1$, $\bar{\sigma} = (t_0 = t, t_1 = \tau, t_2, \dots, t_{\bar{r}})$ and $\bar{\pi} = (p_0 = p, p_1, p_2, \dots, p_{\bar{r}})$ which is optimal, that is minimizes the cost J among all controls. In other words, there exists an optimal control whose first switching instant is τ.*

Given next Proposition 3.1.2 (and in particular looking at its proof), the previous definition is equivalent to require that there exists $p' \in \mathcal{I}_p$ such that the pair (p', τ) realizes the minimum in

$$V(p,t,\rho) = \inf_{\substack{p' \in \mathcal{I}_p \\ \tau \in]t,T]}} \left\{ V(p',\tau) + C(p,p',t,\tau,\rho) \right\}.$$

3.1.1 The optimal switching problem with fixed mass ρ

In this section, we mostly assume that the mass $\rho \in L^2([0,T],[0,1])^{2^N}$ is a priori fixed and then, when not needed, we do not display it as entry of the cost J and of the value function V.

Proposition 3.1.1. *The value function V in* (3.1.3) *is bounded and Lipschitz continuous in time, uniformly in ρ. Moreover, if ρ^n converges to ρ in $L^2([0,T],[0,1])$, then $V(p,\cdot,\rho^n)$ uniformly converges to $V(p,\cdot,\rho)$ on $[0,T]$, for all p. Also, if t'^n is optimal for $V(p,t^n,\rho^n)$ and t'^n, t^n converge to t', t respectively, then t' is optimal for $V(p,t,\rho)$ (see Definition 3.1.3 for t optimal).*

Proof. First of all note that, by (3.1.3) and by the definition of the control triple (r,σ,π), V is increasing with respect to time. Fix $p \in \mathcal{I}$, $t', t'' \in [0,T]$, with $t' > t''$, and $\varepsilon > 0$. Let (r,σ,π) be ε-optimal for (p,t''), that is $V(p,t'') \geq J(p,t'',(r,\sigma,\pi)) - \varepsilon$. Hence the control triple (r,σ,π) is also admissible for t' (all the instants in σ are larger than t') and, by increasingness, we have

$$\begin{aligned}
|V(p,t') - V(p,t'')| &= V(p,t') - V(p,t'') \\
&\leq J(p,t',(r,\sigma,\pi)) - J(p,t'',(r,\sigma,\pi)) + \varepsilon \\
&= C(p,p_1,t',t_1) + \sum_{i=2}^{r} C(p_{i-1},p_i,t_{i-1},t_i) + \tilde{C}(p_r,t_r) - C(p,p_1,t'',t_1) \\
&\quad - \sum_{i=2}^{r} C(p_{i-1},p_i,t_{i-1},t_i) - \tilde{C}(p_r,t_r) + \varepsilon \\
&= C(p,p_1,t',t_1) - C(p,p_1,t'',t_1) + \varepsilon \leq L|t' - t''| + \varepsilon,
\end{aligned}$$

where L is Lipschitz constant of the cost C (see assumption (i)), which is independent of ρ. By the arbitrariness of ε and changing the role of t' and t'', we get the Lipschitz continuity of V in time. The boundedness follows from the fact that, taking $r = 0$, it is $V(p,t) \leq \tilde{C}(p,t)$, which is bounded. The uniformity on ρ is then obtained.

For the convergence of $V(p, \cdot, \rho^n)$, note that, by the previous points and by Ascoli-Arzelà Theorem, at least for a subsequence, we have the uniform convergence to a limit function $\tilde{V}$. Taking $h > 0$ as in the next Remark 3.1.2 (and hence, for all t, the optimal t' belongs to $[t+h, T]$), by the Lipschitz continuity hypotheses on C and $\tilde{C}$ (in particular the continuity of C with respect to $\rho \in L^2$), we get the pointwise convergence to $V(p, \cdot, \rho)$, which then turns out to be the uniform limit $\tilde{V}$, independently of the subsequence. The final point on t'^n, t^n and t', t also comes, for example using the characterization of V by Proposition 3.1.2 which is independent of Proposition 3.1.1. □

Proposition 3.1.2. *The value function V is the unique solution of the following*

$$\begin{cases} V(p,t) = \inf_{\substack{p' \in \mathcal{I}_p \\ t' \in]t,T]}} \{V(p',t') + C(p,p',t,t')\}, & (p,t) \in (\mathcal{I} \setminus \{\bar{p}\}) \times [0,T[\\ V(\bar{p},t) = \tilde{C}(\bar{p},t), & t \in [0,T] \\ V(p,T) = \tilde{C}(p,T), & p \in \mathcal{I} \end{cases} . \tag{3.1.4}$$

Proof. First of all, let us note that the second and third equalities come from the definition of J (3.1.2). We have to prove the first equality. Suppose that $p_r = \bar{p}$, that is $p_r = (1, 1, \ldots, 1)$.

Case 1) Let $p \in \mathcal{I}$ be such that $\sum_i p^i = N - 1$, for instance $p = (1, 1, \ldots, 1, 0)$, so $r = 1$, $\pi = (p, \bar{p})$ and $\sigma = (t, t')$ for some arbitrary $t' \in]t, T]$. Thus we have to prove that

$$V(p,t) = \inf_{t' \in]t,T]} \{V(\bar{p},t') + C(p,\bar{p},t,t')\} = \inf_{t' \in]t,T]} [C(p,\bar{p},t,t') + \tilde{C}(\bar{p},t')] \tag{3.1.5}$$

since $V(\bar{p}, \cdot) = \tilde{C}(\bar{p}, \cdot)$. The last term in the above equality is

$$\inf_{(r,\sigma,\pi)} J(p,t,(r,\sigma,\pi)) = V(p,t),$$

being the controls $(1, (t, t'), (p, \bar{p}))$ the only admissible ones for (p, t).

Case 2) Let $p \in \mathcal{I}$ be such that $\sum_i p^i = N - 2$, that is, for instance, $p = (0, 0, 1, \ldots, 1)$. In this case, the admissible controls must have either $r = 2$ or $r = 1$, and so V is the minimum of the infimum of the cost over the controls with $r = 2$ and the infimum of the cost over the controls with $r = 1$. In the first case, setting $t_{r-2} = t$ and $p_{r-2} = p$, we have

$$
\begin{aligned}
V(p_{r-2}, t_{r-2}) &= \inf_{\substack{(r,\sigma,\pi)\text{ s.t.}\\ r=2\\ \sigma=(t_{r-2},t_{r-1},t_r)\\ \pi=(p_{r-2},p_{r-1},\bar{p})}} J(p,t,(r,\sigma,\pi))\\
&= \inf_{\substack{t_{r-1}\in]t_{r-2},T[\\ t_r\in]t_{r-1},T]\\ p_{r-1}\in\mathcal{I}_{p_{r-2}}}} \Big[C(p_{r-2},p_{r-1},t_{r-2},t_{r-1}) + C(p_{r-1},\bar{p},t_{r-1},t_r) + \tilde{C}(\bar{p},t_r)\Big]\\
&= \inf_{\substack{t_{r-1}\in]t_{r-2},T[\\ p_{r-1}\in\mathcal{I}_{p_{r-2}}}} \Bigg[C(p_{r-2},p_{r-1},t_{r-2},t_{r-1})\\
&\quad + \inf_{t_r\in]t_{r-1},T]} \Big[C(p_{r-1},\bar{p},t_{r-1},t_r) + \tilde{C}(\bar{p},t_r)\Big]\Bigg]\\
&= \inf_{\substack{t_{r-1}\in]t_{r-2},T[\\ p_{r-1}\in\mathcal{I}_{p_{r-2}}}} \left[V(p_{r-1},t_{r-1}) + C(p_{r-2},p_{r-1},t_{r-2},t_{r-1})\right],
\end{aligned}
$$

where the last equality comes from Case 1). The desired result follows.

In the second case, $r = 1$, we must necessarily have $p_r \neq \bar{p}$ and $t_r = T$. Thus we have only to prove that

$$
V(p,t) = \inf_{p_r\in\mathcal{I}_p}\{V(p_r,T) + C(p,p_r,t,T)\} = \inf_{p_r\in\mathcal{I}_p}\{C(p,p_r,t,T) + \tilde{C}(p_r,T)\}
$$

since $V(\cdot,T) = \tilde{C}(\cdot,T)$. The last term in the above equality is

$$
\inf_{(r,\sigma,\pi)} J(p,t,(r,\sigma,\pi)) = V(p,t)
$$

being, in this case, the controls $(1,(t,T),(p,p_r))$ the only ones we are taking account of.

Up to now, we proved the equality for every (p,t) such that $\sum_i p^i = N-1$ and $\sum_i p^i = N-2$. Proceeding backwardly in this way, we then can prove all the other cases with $\sum_i p^i = N-s$ for $s = 3,\dots,N$.

Still arguing backwardly, the uniqueness comes from the fact that any other function satisfying (3.1.4), by (3.1.5) must coincide with V on the nodes (p,t) with $t \in [0,T[$ and p such that $\sum_i p^i = N-1$. □

Remark 3.1.2. *By system* (3.1.4)*, by the boundedness of V (Proposition 3.1.1) and by conditions* (*i*)*,* (*ii*) *in §3.1, there exists $h > 0$ such that the infimum in the first line of* (3.1.4) *is indeed a minimum and t' belongs to* $[t+h,T]$*. The presence of this sort of minimal waiting time h between two consecutive switches will lead to a piecewise continuous/constant feature of* *the evolution of the masses ρ with a uniform bounded number of pieces in $[0, T]$.*

Also justified by Remark 3.1.2, we define

$$P(p,t) = \arg\min_{\substack{p' \in \mathcal{I}_p \\ t' \in]t,T]}} \{V(p',t') + C(p,p',t,t')\}. \tag{3.1.6}$$

In other words, $P(p,t)$ is the couple (p',t') of the node p' where it is optimal to switch at the switching instant $t' > t$. As above, we do not display the dependence on ρ.

Remark 3.1.3 (still on assumption (v) in §3.1)**.** *Assumption (v) may hold for example in the case where the cost C, besides (ii), is derivable w.r.t. the switching time-variable τ with derivative C_τ strictly increasing w.r.t. the quantity $\tau - t$. A possible cost satisfying the previous hypotheses may be for example of the form*

$$C(p,p',t,\tau,\rho) = \frac{\bar{C}(p,p',\rho)}{\tau - t}. \tag{3.1.7}$$

Moreover, we assume that V is convex in time. It follows that it is two times derivable in time almost everywhere (see Theorem B.2.4). For the following counterexample, we are going to assume that the first derivative exists everywhere. By contradiction, let us suppose that if, for the switching from p to p', the decision times t_1, t_2 with $t_1 < t_2$ optimally generate the switching times $\tau_1, \tau_2 < T$ respectively, then $\tau_2 < \tau_1$. Hence it follows that $\tau_1 > \tau_2 \geq t_2 > t_1$. This means that

$$\inf_{\tau \geq t_1} \{V(p',\tau,\rho) + C(p,p',t_1,\tau,\rho)\} = V(p',\tau_1,\rho) + C(p,p',t_1,\tau_1,\rho),$$

$$\inf_{\tau \geq t_2} \{V(p',\tau,\rho) + C(p,p',t_2,\tau,\rho)\} = V(p',\tau_2,\rho) + C(p,p',t_2,\tau_2,\rho).$$

First order conditions give

$$V'(p',\tau_1,\rho) + C_\tau(p,p',t_1,\tau_1,\rho) = 0,$$

$$V'(p',\tau_2,\rho) + C_\tau(p,p',t_2,\tau_2,\rho) = 0.$$

Therefore

$$V'(p',\tau_1,\rho) = -C_\tau(p,p',t_1,\tau_1,\rho) < -C_\tau(p,p',t_2,\tau_2,\rho) = V'(p',\tau_2,\rho),$$

which contradicts the convexity of V in time.

Recall that assumption (v) *implies* (v')*: any optimal switching time less than* T *originates from a unique decision time. With the same hypotheses on* C *as above,* (v') *can be also inferred, without assuming* (v)*, just assuming that* V *is derivable w.r.t. the time-variable without any convexity property. Indeed, suppose that at the decision times* t_1, t_2*,* $t_1 < t_2$*, the agents are optimally switching from* p *to* p' *with the same switching time* $\tau < T$*. Arguing as above, with* $\tau_1 = \tau_2 = \tau$*, we obtain*

$$V'(p', \tau, \rho) = -C_\tau(p, p', t_1, \tau, \rho) = -C_\tau(p, p', t_2, \tau, \rho),$$

contradicting $t_1 \neq t_2$*.*

Without the convexity assumption on V*, we can infer property* (v) *by* (v') *if, besides the derivability of* V*, we assume that the map* $\varphi : t \longmapsto \tau$ *is continuous. Note that, by definition of the optimal switching instant,* $\varphi(t) \to T$ *as* $t \to T$*. By contradiction, suppose that if, for the switching from* p *to* p'*, the decision times* t_1, t_2 *with* $t_1 < t_2$ *optimally generate the switching times* $\tau_1, \tau_2 < T$ *respectively, then* $\varphi(t_2) = \tau_2 < \tau_1 = \varphi(t_1) < T$*. Hence, the function* φ *is somehow decreasing in* $[t_1, t_2]$ *but, by continuity and the limit property above, we must have the existence of* $t' \neq t''$ *such that* $\tau = \varphi(t') = \varphi(t'')$*, contradicting* (v')*.*

Finally, for what concerns the convexity of V*, note that if* C *is strictly convex in* t *and* $\tilde{C}$ *is decreasing in time, due to the decreasingness of* C *with respect to* τ *((ii)) (and the example in* (3.1.7) *satisfies both hypotheses), then for all* p *with one* 0 *only (i.e., directly linked to the destination* $\bar{p}$*),* $V(p, \cdot, \rho)$ *is strictly convex and the functions* φ *are constantly equal to* T*. Proceeding backwardly, we can then prove that for all the other nodes the value functions are all strictly convex and, in particular, the function* φ *is single-valued and increasing (see the example in Appendix A.3).*

Remark 3.1.4. *Let us note that equation* (3.1.4) *is in some sense the Dynamic Programming Principle for the value function* V*. However, we can not differentiate it in the time-variable* t *and obtain an Hamilton-Jacobi equation because our model does not take account of a continuous dynamic evolution of the agents.*

3.2 On the continuity equations for the flow

For what concerns the ρ functions for the masses, using the same possible enumeration of nodes as in §3.1, for every $j = 0, \ldots, 2^N - 1$ we will have, at least formally, a system of 2^N continuity equations in the variables ρ_j^{dm}, the mass of decision-making agents (see assumption (iv) in §3.1), to be interpreted in a suitable formulation that we will see later:

$$\begin{cases}(\rho_j^{\mathrm{dm}})'(t) = \sum_{p_k|p_j\in\mathcal{I}_{p_k}} \lambda_{k,j}(s(t),t)\rho_k^{\mathrm{dm}}(s(t))\delta_t \\ \qquad\qquad\qquad\qquad - \sum_{p_h|p_h\in\mathcal{I}_{p_j}} \lambda_{j,h}(t,\varphi(t))\rho_j^{\mathrm{dm}}(t)\delta_t, \quad t\in]0,T]\ , \\ \rho_j^{\mathrm{dm}}(0) = \rho_j^0 \end{cases} \tag{3.2.8}$$

where ρ_j^0 is fixed for every j, φ is the (possibly multivalued) function introduced in Remark 3.1.1 and $t \longmapsto s(t) \in [0,T]$ takes into account the decision instant s at which an agent switches from p_i to p_j at the switching time t. By assumption (v), $s(t)$ is continuous and non-decreasing (being the inverse of the function φ in Remark 3.1.1) and satisfies $s(t) \le t$ for every t and $s(0) = 0$. Formally such a function s (as well as φ) should be indexed by i, j but, for simplicity, we omit that. The first term in the right-hand side of (3.2.8) represents the mass of decision-making agents arriving to p_j at the switching instant t and the second one, the mass of decision-making agents leaving p_j at the decisional instant t. The unknowns are the 2^N functions ρ_j^{dm} and the functions $\lambda_{k,j} : \mathcal{D} \subset [0,T]\times]0,T] \longrightarrow [0,1]$, $(s,t) \longmapsto \lambda_{k,j}(s,t)$, which indicate how many decision-making agents, on p_k at time s, have chosen $P(p_k,s) = (p_j,t)$, (3.1.6), that is the percentage of mass of decision-making agents which is on p_k and at time s optimally decides to switch to p_j at $t > s$. Of course, if $\lambda_{k,j}(s,t) > 0$, then, at time s, deciding to switch from p_k to p_j at time t is optimal, and we also have $\sum_{p_j|p_j\in\mathcal{I}_{p_k}} \lambda_{k,j}(s,\xi) = 1$, where $\xi \in \varphi(s)$ is any possible selection for the switch from p_k to p_j. Similarly for $\lambda_{j,h}$.

Note that the previous sum equal to 1 means that every instant s is a decisional instant for all the decision-making agents present on the node. The fact that those λ activate a real switch obviously depends on the real presence of decision-making agents on the node at the time s. Indeed, roughly speaking, the interpretation of (3.2.8) is the following one. The functions $\lambda_{i,j}$, for every i, j, give the right way to interpret it. Such functions are basically values between 0 and 1 along the curve $t \longmapsto (s(t),t)$, that is $\lambda_{i,j}$ is concentrated on the curve and it is elsewhere null. From a distributional point-of-view, $\lambda_{i,j}$ is a concentration of Dirac deltas on that curve. In other words, if at the switching instant t the switches from p_k to p_j and from p_j to p_h are both optimal, then $\lambda_{k,j}$ and $\lambda_{j,h}$ are possibly nonzero at $(s(t),t)$ and consequently activate the Dirac deltas, which give the corresponding accumulation of mass (of decision-making agents only) on the arrival node at time t. In the case when the function $t \longmapsto \tau = \varphi(t)$ (Remark 3.1.1) is always a singleton, i.e. not multivalued, then system (3.2.8) may be also interpreted as system of impulsive delayed equations (see for instance [32]). The solutions ρ_j^{dm} are somehow collections of possibly nonzero values on switching (incoming as well as outgoing) instants, and equal to zero elsewhere. The real mass evolution ρ_j, taking into account both decision-making and non-decision making agents, is just the right-continuous constant interpolation of those values. In other words, the 2^N solutions ρ_j are constructed node-by-node for every switching time according to the λ functions, and this process gives piecewise constant functions on $[0, T]$ (see also Remark 3.1.2).

In the next section we are going to make a suitable approximation of the problem, in order to be able to work with piecewise constant functions. Moreover, in that case, we will see a possible direct construction of such functions λ also explaining their presence and roles in (3.2.8), and then the construction of the functions ρ. Actually, we will not use the formal equations (3.2.8) but directly construct step-by-step (switch-by-switch) the solutions. In Figure 3.3, §3.3.1, we graphically represent the construction of a possible ρ^{dm} and its constant interpolation ρ.

3.3 The approximated mean-field problem

As argued at the end of the previous section, we are going to make a suitable approximation in order to allow us to look for solutions ρ of (3.2.8) in $\mathcal{PC}([0,T],[0,1])^{2^N}$, where we recall that $\mathcal{PC}([0,T],[0,1])$ is the set of piecewise constant functions from $[0,T]$ to $[0,1]$. In order to possibly simplify the notation, using the same enumeration of the nodes in §1.1, we consider all the functions ρ_j as forming a unique function in a juxtaposed sequence of 2^N intervals of length T. We then define $\mathcal{K} := \mathcal{PC}([0, 2^N T], [0,1])$ whose elements ρ are still thought as $(\rho_0, \ldots, \rho_{2^N-1})$. The mean-field game system we are going to study is formally described by

$$\begin{cases} V(p,t,\rho) \\ \quad = \inf_{\substack{p'\in\mathcal{I}_p \\ t'\in]t,T]}} \{V(p',t',\rho)+C(p,p',t,t',\rho)\}, \\ \hfill (p,t,\rho)\in(\mathcal{I}\setminus\{\bar{p}\})\times[0,T[\times\mathcal{K} \\ V(\bar{p},t,\rho)=\tilde{C}(\bar{p},t), \hfill (t,\rho)\in[0,T]\times\mathcal{K} \\ V(p,T,\rho)=\tilde{C}(p,T), \hfill (p,\rho)\in\mathcal{I}\times\mathcal{K} \\ \lambda_{i,j}(s,t)=0 \ \text{ if } (p_j,t)\notin P(p_i,s), \\ (\rho_j^{\mathrm{dm}})'(t)=\sum_{p_k|p_j\in\mathcal{I}_{p_k}}\lambda_{k,j}(s(t),t)\rho_k^{\mathrm{dm}}(s(t))\delta_t \\ \hfill -\sum_{p_h|p_h\in\mathcal{I}_{p_j}}\lambda_{j,h}(t,\varphi(t))\rho_j^{\mathrm{dm}}(t)\delta_t, \quad t\in[0,T] \\ \rho_j^{\mathrm{dm}}(0)=\rho_j^0 \\ \rho_j \text{ constant interpolation of } \rho_j^{\mathrm{dm}} \end{cases} . \tag{3.3.9}$$

Note that the fourth line of (3.3.9) stands for the fact that if a switch is not optimal, then the corresponding fraction λ is zero: no one is following that switch.

Next section is devoted to prove the existence of a solution $(\rho_j, \lambda_{j,k})$ of an approximated version of (3.3.9) and hence of an ε-approximated equilibrium of the mean-field game. Such an approximation is mainly consistent in a suitable approximation of the function P in (3.1.6).

3.3.1 Existence of an ε-approximated mean-field equilibrium

As usual in mean-field game problems, we are going to identify the solution ρ of (3.3.9) as a fixed point of a suitable function. At first sight, given also Remark 3.1.2, the space where to search for a fixed point would seem to be the following one:

$$X = \{\rho \in \mathcal{K} : \rho \text{ has at most } M \text{ pieces of constancy}\},$$

where M is a priori fixed, for example $M = \left(\frac{2^N T}{h}\right)^{2^N}$. Note that such a space can be made compact with respect to a suitable convergence but it is certainly not convex (every ρ has different pieces from the others) and, to perform a fixed-point procedure, we need that X satisfies a convexity property. Therefore, to overcome this difficulty, we fix $\varepsilon > 0$ and we consider the partition $\mathcal{P}_\varepsilon$ of $[0, 2^N T]$, given by the nodes $0 < \varepsilon < 2\varepsilon < \ldots \leq 2^N T$ with $\varepsilon = \frac{T}{m}$ for some $m \in \mathbb{N}$. We then consider the space

$$C_\varepsilon = \Big\{\rho \in L^2([0, 2^N T], [0, 1]) :$$
$$\rho \text{ is piecewise constant on the open intervals of } \mathcal{P}_\varepsilon \text{ and } \|\rho\|_\infty \leq \|\rho_0\|_\infty\Big\}.$$

Now, C_ε is convex and compact with respect to the L^2 topology. Indeed, since the partition $\mathcal{P}_\varepsilon$ is fixed and all the functions ρ are constant on it, from every interval of $\mathcal{P}_\varepsilon$ we can extract a convergent constant subsequence whose limit belongs to L^2.

We then look for a fixed point of a suitable multifunction $\psi_\varepsilon : C_\varepsilon \longrightarrow \mathcal{P}(C_\varepsilon)$, $\rho \longmapsto \psi_\varepsilon(\rho)$, that is we look for $\rho_\varepsilon \in C_\varepsilon$ such that $\rho_\varepsilon \in \psi_\varepsilon(\rho_\varepsilon)$. Roughly speaking, the idea is to construct ψ_ε as follows:

(i) ρ is put into (3.1.4) and the value function V is derived;

(ii) V is inserted in (3.1.6) and the variable P, which is not necessarily unique (that is, a priori, there may exist more than one optimal switching instant and more than one admissible subsequent node where it is optimal to switch), is derived;

(iii) we suitably approximate the optimal switching instants given by P at point (ii) with the nodes of the partition $\mathcal{P}_\varepsilon$;

(iv) with such approximated variables P_ε as in (iii), we construct all the possible optimal switching paths with their decision and switching times;

(v) for each optimal switching path π of point (iv), we construct the corresponding functions λ in (3.2.8), as all the agents were following π, that is
$$\lambda_{i,j}^{\pi,\varepsilon}(s,t) = \begin{cases} 1, & (p_j, t) \in P_\varepsilon(p_i, s) \cap \pi \\ 0, & \text{otherwise} \end{cases};$$

(vi) for any π, we insert the functions λ^π into (3.2.8), obtaining the evolution of the mass $\rho^\pi \in C_\varepsilon$;

(vii) by a suitable convexification (interval by interval of the partition $\mathcal{P}_\varepsilon$) of the functions ρ^π of (vi), we construct a set of functions $\psi_\varepsilon(\rho)$, which is contained in $\mathcal{P}(C_\varepsilon)$;

$(viii)$ by proving that $\psi_\varepsilon(\rho)$ is a non-empty and convex subset of C_ε and that the map $\rho \longmapsto \psi_\varepsilon(\rho)$ has closed graph, we can apply the fixed-point Kakutani-Ky Fan Theorem (see Theorem B.2.6) to find a desired ρ_ε.

Note that, by construction, ρ_ε, together with the coefficients λ of the convex combinations of the extremal ρ^π as in point (vii), gives what can be considered as an approximated solution of (3.3.9) and hence an ε-mean-field equilibrium.

We divide the construction of $\psi_\varepsilon(\rho)$ into some steps. For simplicity, we suppose $N = 3$ (compare with Figure 3.2) and consider only paths starting from $p_0 = (0,0,0)$ and that, at the initial time $t = 0$, all the agents are on p_0 (see Remark 3.3.2 below for the general situation). Moreover note that all the paths start at time $t = 0$. In the sequel, we use the following further notation: $p_1 = (1,0,0)$, $p_2 = (0,1,0)$, $p_3 = (0,0,1)$, $p_4 = (1,1,0)$, $p_5 = (0,1,1)$, $p_6 = (1,0,1)$, $p_7 = \bar{p} = (1,1,1)$.

Step 1 (points $(i) - (iv)$). Let $\rho = (\rho_{p_0}, \rho_{p_1}, \rho_{p_2}, \rho_{p_3}, \rho_{p_4}, \rho_{p_5}, \rho_{p_6}, \rho_{p_7}) \in C_\varepsilon$ be fixed. Consider the finite set

$$\tilde{P}_{p_0} = \{(p_1,\tau_1),(p_2,\tau_2),(p_3,\tau_3),(p_4,\tau_4),(p_5,\tau_5),(p_6,\tau_6),(p_7,\tau_7)\},$$

whose elements are the couples composed by all the possible optimal admissible nodes $p_1,\dots,p_7$ (starting from $p_0 = (0,0,0)$), and the possible optimal switching instants $\tau_1,\dots,\tau_7$, as derived in point (ii), that is, for example, τ_2 is the optimal switching instant in order to switch to $p_2 = (0,1,0)$ with decision at $t = 0$ on $p_0 = (0,0,0)$ (independently whether the choice of p_2 is optimal or not).

For point (iii), we argument as follows. At first observe that, at point (ii), the multiplicity of the variables P lies on the admissible subsequent node, but may also lie on the optimal switching instant (for a fixed node), if $\tau^- < \tau^+$, as in Remark 3.1.1. In order to make the solution ρ consistent with the partition $\mathcal{P}_\varepsilon$, and to overcome the possible difficulties of the multivalued feature in time (making it at most discrete), we approximate the possible optimal switching instants $\tau_1,\dots,\tau_7$ with the nodes of $\mathcal{P}_\varepsilon$. In particular, for a generic switching instant τ_i, we set

$$\begin{aligned} \underline{m}(\tau_i,\varepsilon) &:= \max\{n \in \mathbb{N} : n\varepsilon \le \tau_i\}, \\ \underline{m}(\tau_i,\varepsilon)\varepsilon &= \text{the largest node not larger than } \tau_i, \end{aligned}$$

$$\begin{aligned} \overline{m}(\tau_i,\varepsilon) &:= \min\{n \in \mathbb{N} : n\varepsilon \ge \tau_i\}, \\ \overline{m}(\tau_i,\varepsilon)\varepsilon &= \text{the smallest node not smaller than } \tau_i. \end{aligned}$$

Then, if in the switching from p to p', the optimal switching instant τ_i

belongs to the interval $[\underline{m}(\tau_i,\varepsilon)\varepsilon, \overline{m}(\tau_i,\varepsilon)\varepsilon]$, we select

$$\tilde{\tau}_{i,\varepsilon} \in F(\tau_i) = \begin{cases} \{\underline{m}(\tau_i,\varepsilon)\varepsilon\}, & \tau_i \in [\underline{m}(\tau_i,\varepsilon)\varepsilon, \underline{m}(\tau_i,\varepsilon)\varepsilon + \frac{\varepsilon}{2}[\\ \{\underline{m}(\tau_i,\varepsilon)\varepsilon, \overline{m}(\tau_i,\varepsilon)\varepsilon\}, & \tau_i = \underline{m}(\tau_i,\varepsilon)\varepsilon + \frac{\varepsilon}{2} \\ \{\overline{m}(\tau_i,\varepsilon)\varepsilon\}, & \tau_i \in]\underline{m}(\tau_i,\varepsilon)\varepsilon + \frac{\varepsilon}{2}, \overline{m}(\tau_i,\varepsilon)\varepsilon] \end{cases}. \tag{3.3.10}$$

In this way, the approximated variables P_ε in (iii) replace every optimal pair $(p_i,\tau_i) \in P \subseteq \tilde{P}_{p_o}$ by the pairs (which we call ε-optimal) $(p_i, \tilde{\tau}_{i,\varepsilon})$, $\tilde{\tau}_{i,\varepsilon} \in F(\tau_i)$. Therefore, we construct all the possible ε-optimal switching paths π with decision and switching times given by those approximated $\tilde{\tau}_{i,\varepsilon}$, just taking, switch by switch, one and only one of the pairs above. For example, if $p_0 \to p_1 \to p_4 \to p_7$ is an optimal path with τ_1, τ_4, τ_7 the corresponding optimal switching instants, that is

$$(p_1,\tau_1) \in P(p_0,0),\ (p_4,\tau_4) \in P(p_1,\tau_1),\ (p_7,\tau_7) \in P(p_4,\tau_4),$$

then we consider all the possible ε-optimal paths $p_0 \to p_1 \to p_4 \to p_7$ with ε-optimal switching instants $\tilde{\tau}_{j,\varepsilon} \in F(\tau_j)$, $j = 1,4,7$, that is

$$(p_1,\tilde{\tau}_{1,\varepsilon}) \in P_\varepsilon(p_0,0),\ (p_4,\tilde{\tau}_{4,\varepsilon}) \in P_\varepsilon(p_1,\tau_1),\ (p_7,\tilde{\tau}_{7,\varepsilon}) \in P_\varepsilon(p_4,\tau_4),$$

where

$$P_\varepsilon(p_i,s) = \{(p_j, F(\tau_j)) : (p_j,\tau_j) \in P(p_i,s)\}. \tag{3.3.11}$$

In particular, note that, if $\varphi(s) = [\tau_j^-, \tau_j^+]$ as in Remark 3.1.1, then $P_\varepsilon(p_i,s)$ contains all the pairs $(p_j,\tilde{\tau}_j)$ with $\tilde{\tau}_j =$ nodes of $\mathcal{P}_\varepsilon$ in $[\underline{m}(\tau_j^-,\varepsilon)\varepsilon, \overline{m}(\tau_j^+,\varepsilon)\varepsilon]$.

Step 2 (points $(v)-(vii)$). The aim is to build a multifunction $\rho \longmapsto \psi_\varepsilon(\rho) \subset C_\varepsilon$ with (compact and) convex images and closed graph, to which we will apply the fixed-point Kakutani-Ky Fan Theorem.

For each ε-optimal switching path π of point (iv), Step 1, we construct the corresponding evolution of the mass, assuming that all the agents (which here are assumed to be all at p_0 at time $t = 0$) are following π. For example, for the possible ε-optimal path $p_0 \to p_1 \to p_4 \to p_7$ as in Step 1, we would get

$$\rho_0(t) = \begin{cases} \rho^0, & 0 \le t < \tilde\tau_{1,\varepsilon} \\ 0, & \tilde\tau_{1,\varepsilon} \le t \le T \end{cases}, \qquad \rho_1(t) = \begin{cases} 0, & 0 \le t < \tilde\tau_{1,\varepsilon} \\ \rho^0, & \tilde\tau_{1,\varepsilon} \le t < \tilde\tau_{4,\varepsilon} \\ 0, & \tilde\tau_{4,\varepsilon} \le t \le T \end{cases},$$

$$\rho_4(t) = \begin{cases} 0, & 0 \le t < \tilde\tau_{4,\varepsilon} \\ \rho^0, & \tilde\tau_{4,\varepsilon} \le t < \tilde\tau_{7,\varepsilon} \\ 0, & \tilde\tau_{7,\varepsilon} \le t \le T \end{cases}, \qquad \rho_7(t) = \begin{cases} 0, & 0 \le t < \tilde\tau_{7,\varepsilon} \\ \rho^0, & \tilde\tau_{7,\varepsilon} \le t \le T \end{cases},$$

$$\rho_i \equiv 0, \;\; i = 2, 3, 5, 6,$$

and note that, by juxtaposition, $\rho^{\pi,\varepsilon} = (\rho_0, \rho_1, \dots, \rho_7) \in C_\varepsilon$. Formally, as explained in §3.2, such an evolution $\rho^{\pi,\varepsilon}$ can be seen as the constant interpolation of a decision-making solution ρ^{dm} of (3.2.8), with coefficients λ (to be understood associated to π, ε and hence to the corresponding selection in P_ε) satisfying

$$\lambda_{i,j}^{\pi,\varepsilon}(s,t) = \begin{cases} 1, & (p_j, t) \in P_\varepsilon(p_i, s) \cap \pi \\ 0, & \text{otherwise} \end{cases}.$$

The aim is to construct $\psi_\varepsilon(\rho)$ as a suitable convexification of all those "extremal" evolutions $\rho^{\pi,\varepsilon}$. Such a convexification is constructed by taking into account the decision-making nodes $(p_0, 0)$ and $(p_j, \tilde\tau_{j,\varepsilon})$. Still considering an example, suppose that the following paths (nodes p_i and switching time $\tilde\tau_i$) are ε-optimal

$$\begin{aligned} \pi^1 &: (p_0, 0) \to (p_1, \tilde\tau_1^1) \to (p_4, \tilde\tau_4^1) \to (p_7, \tilde\tau_7^1), \\ \pi^2 &: (p_0, 0) \to (p_1, \tilde\tau_1^2) \to (p_6, \tilde\tau_6^2) \to (p_7, \tilde\tau_7^2), \\ \pi^3 &: (p_0, 0) \to (p_3, \tilde\tau_3^3) \to (p_6, \tilde\tau_6^3) \to (p_7, \tilde\tau_7^3), \end{aligned}$$

where we suppose

$$0 < \tilde\tau_1^1 = \tilde\tau_1^2 < \tilde\tau_6^2 < \tilde\tau_3^3 < \tilde\tau_6^3 < \tilde\tau_4^1 < \tilde\tau_7^1 = \tilde\tau_7^2 = \tilde\tau_7^3 = T.$$

We have a first decisional split in p_0 at $t = 0$ between agents switching to p_1 and to p_3, respectively. We then have the convex coefficients $\lambda_{0,1}(0), \lambda_{0,3}(0) \in [0,1]$ with sum equal to 1. Then another decisional split occurs in p_1 at $\tilde\tau_1 = \tilde\tau_1^1 = \tilde\tau_1^2$, giving the convex coefficients $\lambda_{1,4}(\tilde\tau_1)$, $\lambda_{1,6}(\tilde\tau_1)$, and no other decisional split occurs. We then obtain the evolutions

$$
\rho_0(t) = \begin{cases} \rho^0, & 0 \le t < \tilde{\tau}_1^1 \\ \lambda_{0,3}(0)\rho^0, & \tilde{\tau}_1^1 \le t < \tilde{\tau}_3^3 \\ 0, & \tilde{\tau}_3^3 \le t \le T \end{cases},
$$

$$
\rho_1(t) = \begin{cases} 0, & 0 \le t < \tilde{\tau}_1^1 \\ \lambda_{0,1}(0)\rho^0, & \tilde{\tau}_1^1 \le t < \tilde{\tau}_6^2 \\ \lambda_{1,4}(\tilde{\tau}_1)\lambda_{0,1}\rho^0, & \tilde{\tau}_2^6 \le t < \tilde{\tau}_4^1 \\ 0, & \tilde{\tau}_4^1 \le t \le T \end{cases},
$$

$$
\rho_3(t) = \begin{cases} 0, & 0 \le t < \tilde{\tau}_3^3 \\ \lambda_{0,3}(0)\rho^0, & \tilde{\tau}_3^3 \le t < \tilde{\tau}_6^3 \\ 0, & \tilde{\tau}_6^3 \le t \le T \end{cases},
$$

$$
\rho_4(t) = \begin{cases} 0, & 0 \le t < \tilde{\tau}_4^1 \\ \lambda_{1,4}(\tilde{\tau}_1)\lambda_{0,1}(0)\rho^0, & \tilde{\tau}_4^1 \le t < T \\ 0, & t = T \end{cases}, \tag{3.3.12}
$$

$$
\rho_6(t) = \begin{cases} 0, & 0 \le t < \tilde{\tau}_6^2 \\ \lambda_{1,6}(\tilde{\tau}_1)\lambda_{0,1}(0)\rho^0, & \tilde{\tau}_6^2 \le t < \tilde{\tau}_6^3 \\ (\lambda_{1,6}(\tilde{\tau}_1)\lambda_{0,1}(0) + \lambda_{0,3}(0))\rho^0, & \tilde{\tau}_6^3 \le t < T \\ 0, & t = T \end{cases},
$$

$$
\rho_7(t) = \begin{cases} 0, & 0 \le t < T \\ \rho^0, & t = T \end{cases},
$$

$$
\rho_2 = \rho_5 \equiv 0.
$$

Again, by juxtaposition, we get an element of C_ε. The set $\psi_\varepsilon(\rho) \subseteq C_\varepsilon$ is then constructed by all the possible convexifications as above of all sets of extremal evolutions $\rho^{\pi,\varepsilon}$. See Figure 3.3 for a graphic representation of ρ_6^{dm} and its constant interpolation ρ_6.

Remark 3.3.1. *The functions $\lambda_{i,j}$ and their products as shown in the example above, together with the decisional and switching instants, give the coefficients $\lambda_{i,j}$ in the formal equations* (3.2.8)*, for the decision-making part ρ^{dm} of the evolution.*

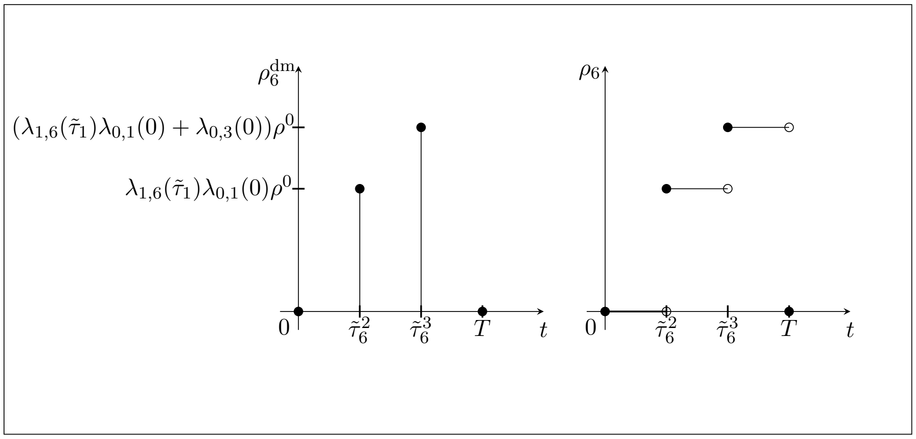


Figure 3.3: Representation of ρ_6^{dm} and of its constant interpolation ρ_6

Lemma 3.3.1 (point $(viii)$)**.** *For any $\rho \in C_\varepsilon$, the set $\psi_\varepsilon(\rho)$ is a non-empty convex (and compact) subset of C_ε. Moreover, the map $\rho \longmapsto \psi_\varepsilon(\rho)$ has closed graph.*

Proof. Clearly the set $\psi_\varepsilon(\rho)$ is non-empty and moreover it is convex. Indeed, if $\rho^1, \rho^2 \in \psi_\varepsilon(\rho)$ and $\lambda \in [0,1]$, then $\lambda\rho^1 + (1-\lambda)\rho^2 \in \psi_\varepsilon(\rho)$. First, note that the extremal evolutions are in a finite quantity $\{\rho^{\pi_{1,\varepsilon}}, \ldots, \rho^{\pi_{r,\varepsilon}}\}$ because the number of ε-optimal paths, $\pi^{k,\varepsilon}$, $k = 1, \ldots, r$, is finite. Hence we can consider both ρ^1 and ρ^2 as a convex combination, decisional node by decisional node (as described in Step 2), of all extremal evolutions, with convex coefficients sets Λ^1 and Λ^2 (note that the decisional nodes $(p_i, \tilde{\tau}_i)$ are determined by the fixed $\rho \in C_\varepsilon$ via (3.1.6), (3.3.11)). This gives that $\lambda\rho^1 + (1-\lambda)\rho^2$ is a same kind of convex combination of the extremal evolutions with set of convex coefficients $\lambda\Lambda^1 + (1-\lambda)\Lambda^2$ (the sum is performed ε-optimal path by ε-optimal path, $\pi^{k,\varepsilon}$, and decisional node by decisional node), and hence it belongs to $\psi_\varepsilon(\rho)$, which turns out to be convex.

Now, we prove that the multifunction $\rho \longmapsto \psi_\varepsilon(\rho)$ has closed graph. From this, we also get the closedness of $\psi_\varepsilon(\rho)$ and, since C_ε is compact, it follows that $\psi_\varepsilon(\rho)$ is compact too.

Consider a sequence $\{\rho^n\}_n \subset C_\varepsilon$ with $\rho^n \longrightarrow \rho$ in C_ε, that is $\rho \in C_\varepsilon$ and the convergence is in L^2. We want to show that for every $\rho'^n \in \psi_\varepsilon(\rho^n)$ with $\rho'^n \longrightarrow \rho'$ in C_ε, we have $\rho' \in \psi_\varepsilon(\rho)$.

Let us prove that, up to a subsequence, $\rho'^n \longrightarrow \tilde{\rho}'$ in L^2 with $\tilde{\rho}' \in C_\varepsilon$ and $\tilde{\rho}' \in \psi_\varepsilon(\rho)$. By the uniqueness of the limit in L^2, it must hold $\rho' = \tilde{\rho}'$, ending the proof. By Proposition 3.1.1, we have $V^n \longrightarrow V$ uniformly on $[0,1]$

(i.e., $V(p,\cdot,\rho^n) \longrightarrow V(p,\cdot,\rho)$ uniformly on $[0,T]$) and if t'^n is optimal for $V(p,t^n,\rho^n)$ and $t^n \to t$, $t'^n \to t'$, then t' is optimal for $V(p,t,\rho)$. Therefore, denoting by $P^n, P^n_\varepsilon, P, P_\varepsilon$ the functions (3.1.6) and (3.3.11) corresponding to ρ^n and ρ, respectively, we have

$$(p'^n, t'^n) \in P^n(p,t^n) \text{ and } (p'^n,t'^n) \to (p',t') \Rightarrow \ (p',t') \in P(p,t), \quad (3.3.13)$$

and hence, by definition of P_ε, (3.3.11) (see also the comment below it), in particular by the definition of F in (3.3.10), for every choice of $(p'^n, \tilde{t}'^n) \in P^n_\varepsilon(p,t^n)$ there exists $(p',\tilde{t}') \in P_\varepsilon(p,t)$ such that

$$(p'^n, \tilde{t}'^n) \to (p',\tilde{t}') \ \text{ up to a subsequence (with } p'^n, t'^n, p', t' \text{ as in (3.3.13))}. \quad (3.3.14)$$

Moreover, since the nodes are finite, there exists $\bar{n} \in \mathbb{N}$ such that for every p,

$$p^n \to p \ \Rightarrow \ p^n = p \quad \text{for every } n \geq \bar{n}. \quad (3.3.15)$$

Let $(\rho^{\pi_1,\varepsilon},\dots,\rho^{\pi_r,\varepsilon})$ be the extremal points of $\psi_\varepsilon(\rho)$, where $\pi_1,\dots,\pi_r$ are the ε-optimal paths. By (3.3.15), we can assume that for n sufficiently large, also in $\psi_\varepsilon(\rho^n)$ the extremal points are exactly in the quantity r and their sequences of nodes are the same as the ones of $\pi_1,\dots,\pi_r$ and only the decisional and switching instants may change with n. Let us denote by $\rho^{\pi_1,n,\varepsilon},\dots,\rho^{\pi_r,n,\varepsilon}$ those extremal points. Then, for n sufficiently large, $\rho'^n \in \psi_\varepsilon(\rho^n)$ is a convex combination, constructed as in Step 2, of the extremal points $\rho^{\pi_1,n,\varepsilon},\dots,\rho^{\pi_r,n,\varepsilon}$. Let $\lambda^n_{i,j}(\tilde{t}^n) \in [0,1]$ be the corresponding coefficients for the generic decisional instant $\tilde{t}^n$. Up to a subsequence, we can assume that $\tilde{t}^n \to \tilde{t}$ and $\lambda^n_{i,j}(\tilde{t}^n) \to \lambda_{i,j} =: \lambda_{i,j}(\tilde{t}) \in [0,1]$ and also $\tilde{t}'^n \to \tilde{t}'$ with $(p',\tilde{t}'^n) \in P^n_\varepsilon(p,\tilde{t}^n)$ and, by (3.3.14), $(p',\tilde{t}') \in P_\varepsilon(p,\tilde{t})$. Since $\tilde{t}'^n, \tilde{t}^n$ assume only discrete values on partition $\mathcal{P}_\varepsilon$, we can also assume $\tilde{t}'^n = \tilde{t}'$ and $\tilde{t}^n = \tilde{t}$ for n sufficiently large. Hence the extremal points $\rho^{\pi_1,n,\varepsilon},\dots,\rho^{\pi_r,n,\varepsilon}$ are exactly the same as the ones of the limit case $\psi_\varepsilon(\rho)$: the same ε-optimal paths $\pi_1,\dots,\pi_r$ with the same decisional and switching instants. The only convergence is in the convex coefficients.

Now, we construct $\tilde{\rho}'$ as the convex combination of the extremal points with limit coefficients $\lambda_{i,j}$. Obviously $\tilde{\rho}' \in C_\varepsilon$ and $\tilde{\rho}'^n \longrightarrow \tilde{\rho}'$ in L^2. To conclude, we have to prove that $\tilde{\rho}' \in \psi_\varepsilon(\rho)$. In particular, we have to show that if $(p_j,\tilde{t}') \notin P_\varepsilon(p_i,\tilde{t})$, then the corresponding $\lambda_{i,j}(\tilde{t}) = 0$. This is true because, if $\lambda_{i,j}(\tilde{t})$ was greater than 0, then $\lambda^n_{i,j}(\tilde{t}^n) > 0$ by convergence and hence $(p_j,\tilde{t}'^n) \in P^n_\varepsilon(p_i,\tilde{t}^n)$, and this is in contradiction with (3.3.14). Therefore $\tilde{\rho}' \in \psi_\varepsilon(\rho)$ and we conclude because, by construction, $\rho'^n \longrightarrow \tilde{\rho}'$ in L^2 since the convergence of the coefficients $\lambda^n_{i,j}$ gives the convergence of the constant values of ρ'^n on the partition $\mathcal{P}_\varepsilon$ to the constant values of $\tilde{\rho}'$. $\square$

Remark 3.3.2. *Observe that the general case $N > 3$ works with the same ideas and tools, being careful that we will have a more complex network (i.e, many more nodes and paths, that is a more complex topology of the network), which makes the fixed-point procedure above certainly harder from a computational point-of-view but even just from a notational one, already for what concerns the analytical description of ψ_ε (see for example the description of ρ_6 in the simple case in* (3.3.12)*). Moreover, here above, for simplicity, we considered only paths starting from $p_0 = (0,0,0)$ and that, at the initial time $t = 0$, all the agents are on p_0, that is $\rho_i(0) = 0$ for all $i \neq 0$. The case where at the initial time the mass is possible distributed to different nodes, up to suitably construct the evolutions as in Step 2, which will be more knotty, does not change the proof too much (we may have more involved intersections and overlaps of switches, still in a finite number, as ρ_6 in* (3.3.12) *but probably in a more complicated way).*

Still considering the network in Figure (3.2) *as in* (3.3.12)*, with the same enumeration of nodes $p_0, p_1, \ldots, p_7 = \bar{p}$, in order to give an idea of the descriptive and notational complexity of the construction of ψ_ε, already in the case of that simple network, but with a generic initial distribution $\rho^0 = (\rho_0^0, \rho_1^0, \ldots \rho_7^0)$, if we consider, for instance, the flow ρ_6 through the node p_6, we have*

$$\rho_6 = \rho^{6,6} + \rho^{1,6} + \rho^{3,6} + \rho^{0,1,6} + \rho^{0,3,6}.$$

The term $\rho^{6,6}$ corresponds to the flow of the agents that at time $t = 0$ are already on p_6: all of them, at the decisional instant $t = 0$, choose a switching instant $\tau_{6,7}$ optimally generated as in (3.3.10) *in order to switch from p_6 to p_7.*

The term $\rho^{1,6}$ corresponds to the flow, through p_6, of the agents that at $t = 0$ were on p_1: all of them, at the decisional instant $t = 0$, choose a switching instant $\tau_{1,6}$ optimally generated as in (3.3.10) *in order to switch from p_1 to p_6, together with the corresponding fraction $\lambda_{1,6}$ of agents performing such a switch. Hence, at the instant $\tau_{1,6}$, the mass of agents $\lambda_{1,6}\rho_1^0$ switches from p_1 to p_6. Such a mass of agents, at the (decisional) instant $\tau_{1,6}$, optimally chooses a switching instant $\tau_{1,6,7}$ in order to switch from p_6 to p_7.*

The term $\rho^{3,6}$ is constructed similarly to $\rho^{1,6}$ by replacing p_1 with p_3.

The term $\rho^{0,1,6}$ corresponds to the flow, through p_6, of the agents that at $t = 0$ were on p_0: all of them, at the decisional instant $t = 0$, choose a switching instant $\tau_{0,1}$ optimally generated as in (3.3.10) *in order to switch from p_0 to p_1, together with the corresponding fraction $\lambda_{0,1}$ of agents performing such a switch. Hence, at the instant $\tau_{0,1}$, the mass of agents $\lambda_{0,1}\rho_0^0$*

switches from p_0 to p_1. Such a mass of agents, at the (decisional) instant $\tau_{0,1}$, optimally chooses a switching instant $\tau_{0,1,6}$ in order to switch from p_1 to p_6, together with the fraction $\lambda_{0,1,6}$ of agents performing such a switch. Therefore, at the instant $\tau_{0,1,6}$, the mass of agents $\lambda_{0,1,6}\lambda_{0,1}\rho_0^0$, switches from p_1 to p_6. Such a mass of agents, at the (decisional) instant $\tau_{0,1,6}$, optimally chooses a switching instant $\tau_{0,1,6,7}$ in order to switch from p_6 to p_7.

The term $\rho^{0,3,6}$ is constructed similarly to $\rho^{0,1,6}$ by replacing p_1 with p_3.

Obviously, the coefficients λ above must be constrained to have sum equal to 1 with the other corresponding coefficients. For instance, $\lambda_{0,1,6} + \lambda_{0,1,4} = 1$. Finally note that in the simple case (3.3.12), ρ_6 corresponds to $\rho^{0,1,6} + \rho^{0,3,6}$ only, and, in particular, $\lambda_{3,6} = 1$, which means that $\lambda_{3,5} = 0$, for the optimality hypotheses assumed in that example.

Theorem 3.3.1. *Under all the hypotheses stated in §3.1, there exists an ε-mean-field equilibrium of system (3.3.9).*

Proof. The proof follows from Lemma 3.3.1, Remark 3.3.2 and the fixed-point Kakutani-Ky Fan Theorem. □

3.4 On the limit $\varepsilon \to 0$ and the existence and uniqueness of a mean-field equilibrium

In the sequel, we denote by ρ_ε a fixed point for $\psi_\varepsilon(\rho)$, i.e., a total mass satisfying $\rho_\varepsilon \in \psi_\varepsilon(\rho_\varepsilon)$. The existence of such fixed points is proved in the previous section and now we will perform the limit procedure as $\varepsilon \to 0$, obtaining as limit $\rho \in L^2([0,T],[0,1])^{2^N}$ such that $\rho \in \psi(\rho)$, where ψ is constructed as in the previous points (i)–$(viii)$ with the only difference that we do not perform the approximation P_ε in (iii), but we just consider the function P, (3.1.6). Hence ρ, together with its convexity coefficients, will be a solution of (3.3.9) and a mean-field equilibrium.

One of the main problems in performing such a limit is the fact that the functions $t \longmapsto \tau = \varphi(t)$ (see Remark 3.1.1) may be multivalued, and, in particular, with a continuum (an interval) as image of t. This problem was bypassed in the previous section using the time-discretizetion given by the partition $\mathcal{P}_\varepsilon$. We first assume that the functions φ are not multivalued and we prove, in such a case, the existence of a mean-field equilibrium, that is of a function $\rho \in L^2$ such that $\rho \in \psi(\rho)$.

Theorem 3.4.1. *Under all the hypotheses stated in §3.1 and assuming the single-valued feature of φ, there exists a mean-field equilibrium of system* (3.3.9)*, that is there exists $\rho \in L^2$ such that $\rho \in \psi(\rho)$.*

Proof. First of all note that, fixed ρ, under the hypothesis on φ, for every decisional instant t and node p_i, there exists a unique optimal switching instant τ for the switch to p_j, that is $(p_j, \tau) \in P(t, p_i)$. This fact gives that the mass evolution $\rho' \in \psi(\rho)$ is also piecewise constant and similarly constructed as in Step 2, §3.3.1, with the only difference that now the pieces of constancy are not fixed a priori (we do not have the partition $\mathcal{P}_\varepsilon$). Moreover, for all $\varepsilon > 0$, the function P_ε, (3.3.11), evaluated at (t, p_i), generates at most two ε-approximated switching instants for the switch to p_j: the possible approximation $\tilde{\tau}_\varepsilon$ of τ by the function F in (3.3.10) (and not the whole intersection of the nodes of the partition with the interval $\varphi(t)$ in the case of multivalued feature). Finally, $\tilde{\tau}_\varepsilon \to \tau$ as $\varepsilon \to 0$.

Now, recall that (see the beginning of §3.3.1) the fixed points ρ_ε are piecewise constant with at most a fixed number M of pieces of constancy. Hence, possibly extracting a subsequence, we can make such intervals of constancy converge as well as the corresponding values of the constants. We then obtain a function ρ such that, up to a subsequence, $\rho_\varepsilon \to \rho$ in L^2. The convergence of the constant values is obviously constructed by the convergence, up to a subsequence, of the convex coefficients $\lambda_{i,j}^\varepsilon \in \mathbb{R}$ evaluated on the decisional instants and implemented at the corresponding ε-approximated instants as in Step 2, §3.3.1. Note that the decisional and switching instants are the extremal points of the intervals of constancy, and also that, being the number of possible cases finite, we may assume, up to a subsequence, that those ones are decisional and switching instants for the same switch from p_i to p_j, i.e. for the same i and j for all ε. Finally note that ρ_ε, being a fixed point of ψ_ε, is exactly constructed by its coefficients $\lambda_{i,j}^\varepsilon$ implemented on the nodes that are generated by ρ_ε itself via P_ε.

Arguing as in the proof of Lemma 3.3.1, using Proposition 3.1.1 and similar convergence for $\varepsilon \to 0$ as in (3.3.13) and (3.3.14), we have that $\rho \in \psi(\rho)$ (i.e.: ρ is constructed by the coefficients $\lambda_{i,j}$ implemented on the nodes that are generated by ρ itself via P, and moreover if the switch is not optimal, then $\lambda_{i,j} = 0$). $\square$

3.4.1 The general case: φ multivalued

Without the single-valued hypothesis on φ, the passage to the limit as $\varepsilon \to 0$ is more involved. Indeed, if the image of the decisional time t is an interval

$[\tau^-, \tau^+]$, in the ε-approximation case we discretize it through the partition $\mathcal{P}_\varepsilon$ and, on every node, we get a value $\lambda^\varepsilon_{i,j}(t, \cdot)$ which composes with the others. Formally, we have a sum of weighted delta functions on the nodes of $\mathcal{P}_\varepsilon$ inside $[\tau^-, \tau^+]$. In the the limit as $\varepsilon \to 0$, we obtain instead a possible sum of functions $\lambda_{i,j}(t, \cdot)$, defined on the whole interval $[\tau^-, \tau^+]$ and other sums of delta functions. Hence the situation is more complex, including the interpretation of system (3.2.8). A deeper investigation of this situation is going to be the subject of future works. Again considering a particular case, where $\rho_\varepsilon \longrightarrow \rho$ in L^2 and ρ, via the functions P, generates functions φ not multivalued, then ρ may be a mean-field equilibrium because the proof of Theorem 3.4.1 can be probably adapted. Also for this case the details have not been checked. However, in Appendix A.3, we give an explicit example of possible costs that guarantee the single-valued feature of φ.

3.4.2 On the uniqueness of the equilibrium

The uniqueness of the equilibrium is often proved by assuming the Lasry-Lions monotonicity condition on the cost (see [31]). Our problem does not immediately fit into such a property because of its deterministic and network-type features, and the presence of two kinds of time variables. Anyway, in Appendix A.2, we try to show, by two simple examples, how a monotonicity-type condition can be promising in order to study the uniqueness of the equilibrium, but the real implementation of that condition in our model is completely left to future studies.

Conclusions

The aim of this thesis was to present some optimal visiting problems in different frameworks (multi-dimensional and network). For each of them, we studied the model for a single player, proving rather exhaustively the well-position of the problem, and then the model for a huge population of agents. For the latter, several difficulties arose for both frameworks. Some mainly due to the non-uniqueness of the optimal control, for which we had to propose different approaches to address the mean-field case and to study the existence of an equilibrium. Others essentially due to the deterministic features and to the presence of switches as well as more than one target in the dynamics of the problem. This led us to a step-by-step study, focusing at first on giving a suitable formulation to the continuity equation for the distribution of agents, and proving some important results which, although partial, are necessary for the continuation of the work and for a possible extension to a more general theory of mean-field games. To this purpose, a rigorous investigation, together with other questions left open in the thesis, is going to be the subject of future research.

Appendix A

Auxiliary results and proofs

A.1 Preliminary results on the time-dependent optimal stopping problem in Ch. 1, §1.1.3

The aim of this section is to prove Lemma 1.1.1 in Ch 1, §1.1.3, which is Lemma A.1.7 here. In particular, at first we prove it in the case with no control, which is Lemma A.1.5. Then, we observe that the general case can be proved by combining Lemma A.1.5 and an approximation argument, which is Lemma A.1.6, and hence we give the proof. As we explained also at the beginning of §1.1.3, for the results and the proofs in this section we suitably generalize the results in [9] for an optimal stopping problem with no time-dependence and infinite horizon feature.

Let $\Omega \subset \mathbb{R}^d$ be an open subset. Let us consider the system

$$\begin{cases} y'(s) = f(y(s)), & s \in]t, T] \\ y(t) = x \in \Omega \end{cases}, \tag{A.1.1}$$

where $T > 0$, $t \in [0, T]$ and $f : \Omega \longrightarrow \mathbb{R}^d$ satisfies the same hypotheses in §1.1.3.

Let us associate to a function $u \in C^0(\Omega \times [0, T])$ and $(x, t) \in \Omega \times [0, T[$, the sets

$$D^+ u(x,t) = \Bigg\{ (p, g) \in \mathbb{R}^{d+1} : \limsup_{(y,s)\to(x,t),(y,s)\in\Omega\times[0,T[} \frac{u(y,s) - u(x,t) - g(s-t) - p \cdot (y-x)}{\|y-x\| + |s-t|} \le 0 \Bigg\},$$

$$D^-u(x,t) = \Bigg\{(p,g) \in \mathbb{R}^{d+1} : \\ \liminf_{(y,s)\to(x,t),(y,s)\in\Omega\times[0,T[} \frac{u(y,s)-u(x,t)-g(s-t)-p\cdot(y-x)}{\|y-x\|+|s-t|} \le 0\Bigg\},$$

that is the super- and the subdifferential (or semidifferentials) of u at (x,t) respectively.

The following lemma provides a description of $D^+u(x,t)$ and $D^-u(x,t)$ in terms of test functions.

Lemma A.1.1. *Let $u \in C^0(\Omega \times [0,T])$. Then,*

(a) $(p,g) \in D^+u(x,t)$ if and only if there exists $\varphi \in C^1(\Omega \times [0,T])$ such that $D_x\varphi(x,t) = p$, $\varphi_t(x,t) = g$ and $u - \varphi$ has a local maximum at (x,t);

(b) $(p,g) \in D^-u(x,t)$ if and only if there exists $\varphi \in C^1(\Omega \times [0,T])$ such that $D_x\varphi(x,t) = p$, $\varphi_t(x,t) = g$ and $u - \varphi$ has a local minimum at (x,t).

Proof. At first we prove (a). Let $(p,g) \in D^+u(x,t)$. Then, for some $\delta > 0$,

$$u(y,s) \le u(x,t) + p\cdot(y-x) + g(s-t) \\ + \sigma(\|y-x\|+|s-t|)(\|y-x\|+|s-t|) \quad \text{for any } (y,s) \in B((x,t),\delta),$$

where σ is a continuous increasing function on $[0,+\infty[$ such that $\sigma(0) = 0$. Now define a C^1 function ρ by

$$\rho(r) = \int_0^r \sigma(\tau)d\tau.$$

It is not difficult to check that the properties

$$\rho(0) = \rho'(0) = 0, \quad \rho(2r) \ge \sigma(r)r$$

imply that the function φ defined by

$$\varphi(y,s) = u(x,t) + p\cdot(y-x) + g(s-t) + \rho\left(2(\|y-x\|+|s-t|)\right)$$

belongs to $C^1(\mathbb{R}^d \times [0,T])$ and $D_x\varphi(x,t) = p$, $\varphi_t(x,t) = g$. Moreover, for $(y,s) \in B((x,t),\delta)$,

$$
\begin{aligned}
(u-\varphi)(y,s) \le \sigma(\|x-y\|+|s-t|)(\|y-x\|+|s-t|)& \\
-\rho\left(2(\|y-x\|-|s-t|)\right) \le 0 = (u-\varphi)(x,t).&
\end{aligned}
$$

For the opposite implication, it is sufficient to observe that

$$
\begin{aligned}
u(y,s)-u(x,t)-D_x\varphi(x,t)\cdot(y-x)-\varphi_t(x,t)(s-t)& \\
\le \varphi(y,s)-\varphi(x,t)-D_x\varphi(x,t)\cdot(y-x)-\varphi_t(x,t)(s-t)&
\end{aligned}
$$

for $(y,s) \in B((x,t),\delta)$, and the proof of (a) is complete.

Since $D^-u(x,t) = -(D^+(-u)(x,t))$, the proof of (b) follows from the above argument applied to $-u$. □

A fundamental property of the super- and subdifferential and the semidifferential versions of a useful fact in elementary calculus are shown in the following lemma.

Lemma A.1.2. *Let $u \in C^0(\Omega \times [0,T])$.*

(i) The sets $A^+ = \{(x,t) \in \Omega \times [0,T[\colon D^+u(x,t) \neq \emptyset\}$, $A^- = \{(x,t) \in \Omega \times [0,T[\colon D^-u(x,t) \neq \emptyset\}$ are dense.

(ii) For $v(x,t,r) = \varphi(r)u(x,t)$ ($(x,t) \in \Omega \times [0,T]$, $r \in \mathbb{R}$), we have

$$
\begin{aligned}
D^+v(x,t,r) = \{(q,g,\sigma) \in \mathbb{R}^{d+2} :& \\
(q,g) \in \varphi(r)D^+u(x,t),\ \sigma = \varphi'(r)u(x,t)\},&
\end{aligned}
$$

provided $\varphi \in C^1(\mathbb{R})$, $\varphi(r) \ge 0$ for all $r \in \mathbb{R}$.

Proof. Let us prove (i). Let $(\bar{x},\bar{t}) \in \Omega \times [0,T[$ and consider the smooth function

$$
\varphi_\varepsilon(x,t) = \frac{\|x-\bar{x}\|^2 + |t-\bar{t}|^2}{\varepsilon}.
$$

For any $\varepsilon > 0$, $u - \varphi_\varepsilon$ attains its maximum over $\overline{B} = \overline{B}((\bar{x},\bar{t}),R)$ at some point $(x_\varepsilon,t_\varepsilon)$. From the inequality

$$
(u-\varphi_\varepsilon)(x_\varepsilon,t_\varepsilon) \ge (u-\varphi_\varepsilon)(\bar{x},\bar{t}) = u(\bar{x},\bar{t})
$$

we get, for all $\varepsilon > 0$,

$$
\|x_\varepsilon-\bar{x}\|^2 + |t_\varepsilon-\bar{t}|^2 \le 2\varepsilon \sup_{(x,t)\in\overline{B}} |u(x,t)|.
$$

Hence $(x_\varepsilon, t_\varepsilon)$ is not on the boundary of $\overline{B}$ for ε sufficiently small, and, by (a) of Lemma A.1.1,

$$(D_x\varphi_\varepsilon(x_\varepsilon,t_\varepsilon), \partial_t\varphi_\varepsilon(x_\varepsilon,t_\varepsilon)) = \left(2\frac{x_\varepsilon - \bar{x}}{\varepsilon}, 2\frac{t_\varepsilon - \bar{t}}{\varepsilon}\right)$$

belongs to $D^+u(x_\varepsilon, t_\varepsilon)$. This proves that A^+ is dense. Similar arguments show that A^- is dense too.

Now we prove (ii). Since $\varphi \in C^1(\mathbb{R})$ we have

$$\begin{aligned} v(y,\tau,s) - v(x,t,r) &= \varphi(s)u(y,\tau) - \varphi(r)u(x,t) \\ &= \varphi(s)u(y,\tau) - (\varphi(s) + \varphi'(s)(r-s) + o(|r-s|))u(x,t) \\ &= \varphi(s)u(y,\tau) - \varphi(s)u(x,t) - \varphi'(s)u(x,t)(r-s) + o(|s-r|)u(x,t) \\ &= \varphi'(s)u(x,t)(s-r) + \varphi(s)(u(y,\tau) - u(x,t)) + o(|s-r|). \end{aligned}$$

Hence $(q,g,\sigma) \in D^+v(x,t,r)$ if and only if

$$\begin{aligned} &v(y,\tau,s) - v(x,t,r) \\ &\qquad \le p\cdot(x-y) + g(\tau - t) + \sigma(s-r) + o(\|x-y\| + |\tau - t| + |s-r|), \end{aligned}$$

that is

$$\begin{aligned} &\varphi'(s)u(x,t)(s-r) + \varphi(s)(u(y,\tau) - u(x,t)) + o(|s-r|) \\ &\qquad \le p\cdot(x-y) + g(\tau - t) + \sigma(s-r) + o(\|x-y\| + |\tau - t| + |s-r|), \end{aligned}$$

that is

$$\begin{aligned} &(\varphi'(s)u(x,t) - \sigma)(s-r) + \varphi(s)(u(y,\tau) - u(x,t)) \\ &\qquad\qquad - p\cdot(x-y) - g(\tau - t) \le o(\|x-y\| + |\tau - t| + |s-r|) \end{aligned}$$

for any (y,τ,s) in a neighborhood of (x,t,r). This easily implies that $\sigma = \varphi'(s)u(x,t)$.

Assume now $\varphi(r) \neq 0$ and that $(q/\varphi(r), g/\varphi(r)) \notin D^+u(x,t)$. This implies

$$u(y,\tau) - u(x,t) > \frac{q}{\varphi(r)}\cdot(x-y) + \frac{g}{\varphi(r)}(\tau - t) + o(\|x-y\| + |\tau - t|),$$

i.e.,

$$\varphi(r)(u(y,\tau) - u(x,t)) - q\cdot(x-y) - g(\tau - t) > \varphi(r)o(\|x-y\| + |\tau - t|),$$

a contradiction with the above inequality with $r = s$. If $\varphi(s) = 0$, the choice $r = s$ gives

$$-q \cdot (x - y) - g(\tau - t) \leq o(\|x - y\| + |\tau - t|),$$

thus $q = 0$ and $g = 0$. Hence the thesis follows. $\square$

Remark A.1.1. *A similar result of Lemma A.1.2, (ii), holds for D^-. The sign condition on φ is essential. Indeed, in general, if $u \in C^0(\Omega \times [0, T])$, then*

(1) $D^+(\alpha u)(x, t) = \alpha D^+ u(x, t)$ *if* $\alpha > 0$*;*

(2) $D^+(\alpha u)(x, t) = \alpha D^- u(x, t)$ *if* $\alpha < 0$*.*

In the following lemma, the term T does not stand for the fixed finite horizon above.

Lemma A.1.3. *Let $u \in C^0(]0, T[)$, $T > 0$. Then the following statements are equivalent*

(i) *u is nondecreasing in $]0, T[$;*

(ii) *$u' \geq 0$ in $]0, T[$ in the viscosity sense;*

(iii) *$-u' \leq 0$ in $]0, T[$ in the viscosity sense.*

Proof. See for example [9], Ch. II, §5.5, Lemma 5.15. $\square$

Remark A.1.2. *From Lemma A.1.3 it follows that if $\ell \in C^0(]0, T[)$, then $t \longmapsto u(t) + \int_0^t \ell(s)ds$ is nondecreasing if and only if $u' + \ell \geq 0$ or $-u' - \ell \leq 0$ in the viscosity sense.*

Moreover, it can be proved similarly that u is nonincreasing if and only if $u' \leq 0$ in the viscosity sense. Hence, $u' = 0$ in the viscosity sense is equivalent to u being a constant.

In order to generalize Lemma A.1.3 to a higher dimension, let $u \in C^0(\Omega \times [0, T])$. For a fixed $z = (x_2, \ldots, x_d, t) \in \mathbb{R}^{d-1} \times [0, T]$, we set

$$\Omega_z := \{x_1 \in \mathbb{R} : x = (x_1, z) \in \Omega \times [0, T]\}$$

and

$$u_z : \Omega_z \longrightarrow \mathbb{R}, \qquad u_z(x_1) := u(x_1, z).$$

Lemma A.1.4. *Let $u, \ell \in C^0(\Omega \times [0,T])$. Then the following statements are equivalent:*

(*i*) *for each $z \in \mathbb{R}^{d-1} \times [0,T]$, u_z is a viscosity supersolution of*

$$u_z'(x_1) \geq \ell_z(x_1) \quad \textit{in } \Omega_z;$$

(*ii*) *u is a viscosity supersolution of*

$$\begin{aligned}\frac{\partial u}{\partial t}(x_1, x_2, \ldots, x_d, t) + \frac{\partial u}{\partial x_1}&(x_1, x_2, \ldots, x_d, t) \\ &\geq \ell(x_1, x_2, \ldots, x_d, t) \quad \textit{in } \Omega \times [0, T[.\end{aligned}$$

Proof. At first we prove that (*ii*) implies (*i*). Let $z^0 \in \mathbb{R}^{d-1} \times [0,T]$ such that $\Omega_{z^0} \neq \emptyset$ and assume that x_1^0 is a strict local minimum for $u_{z^0} - \eta$ with $\eta \in C^1$. It is not restrictive to assume that $\eta \leq -1$ in $B(x_1^0, \delta)$ for some $\delta > 0$. Consider now

$$\varphi_\varepsilon(x_1, \ldots, x_d, t) := \eta(x_1)\left(1 + \frac{\|z - z^0\|^2}{\varepsilon}\right), \quad \varepsilon > 0.$$

If $x^\varepsilon = (x_1^\varepsilon, z^\varepsilon)$ is a minimum point for $u - \varphi_\varepsilon$ in $\overline{B(x^0, \delta)}$ ($x^0 = (x_1^0, z^0)$), then

$$\begin{aligned}u(x^\varepsilon) - \varphi_\varepsilon(x^\varepsilon) = u(x^\varepsilon) - \eta(x_1^\varepsilon) - \eta(x_1^\varepsilon)&\frac{\|z^\varepsilon - z^0\|^2}{\varepsilon} \\ &\leq u(x^0) - \varphi_\varepsilon(x^0) = u_{z^0}(x_1^0) - \eta(x_1^0). \quad \text{(A.1.2)}\end{aligned}$$

Since $\eta \leq -1$ in $B(x_1^0, \delta)$, it follows that

$$\frac{\|z^\varepsilon - z^0\|^2}{\varepsilon} \leq u_{z^0}(x_1^0) - \eta(x_1^0) + \eta(x_1^\varepsilon) - u(x^\varepsilon).$$

Therefore

$$\frac{\|z^\varepsilon - z^0\|^2}{\varepsilon} \leq C, \quad z^\varepsilon \to z^0, \quad \text{as } \varepsilon \to 0^+.$$

Then, at least for a subsequence,

$$x_1^\varepsilon \to \bar{x}_1, \quad \frac{\|z^\varepsilon - z^0\|^2}{\varepsilon} \to \alpha \geq 0, \quad \text{as } \varepsilon \to 0^+.$$

Letting $\varepsilon \to 0^+$ in (A.1.2) we obtain

$$u_{z^0}(x_1^0) - \eta(x_1^0) \geq u(\bar{x}_1, z^0) - \eta(\bar{x}_1) - \eta(\bar{x}_1)\alpha \geq u_{z^0}(\bar{x}_1) - \eta(\bar{x}_1).$$

Since x_1^0 was a local strict minimum for $u_{z^0} - \eta$, the above implies $\bar{x}_1 = x_1^0$ and $\alpha = 0$. Now, assuming the validity of (ii), we have

$$\frac{\partial \varphi_\varepsilon}{\partial t}(x^\varepsilon) + \frac{\partial \varphi_\varepsilon}{\partial x_1}(x^\varepsilon) = \eta(x_1)\frac{2(t^\varepsilon - t^0)}{\varepsilon} + \eta'(x_1^\varepsilon)\left(1 + \frac{\|z^\varepsilon - z^0\|^2}{\varepsilon}\right) \geq \ell(x^\varepsilon).$$

If we let $\varepsilon \to 0^+$ in the above inequality, we conclude

$$\eta'(x_1^0) \geq \ell(x_1^0, z^0) = \ell_{z^0}(x_1^0),$$

which shows that (i) holds.

The proof of the reverse implication is straightforward. It is enough to observe that if $\bar{x} = (\bar{x}_1, \bar{z})$ is a local minimum for $u - \varphi$, $\varphi \in C^1(\Omega \times [0,T])$, then $\bar{x}_1$ is a local minimum for $u_{\bar{z}}(x_1) - \varphi(x_1, \bar{z})$. $\square$

In the following, we use the notation

$$\tau_{(x,t)} := \min\{\inf\{\tau \geq t : y_{(x,t)}(\tau) \notin \Omega\}, T\}, \quad (x,t) \in \Omega \times [0,T],$$

where $y_{(x,t)}(\cdot)$ is the solution of (A.1.1).

Lemma A.1.5. *Let us assume $u, \ell \in C^0(\Omega \times [0,T])$, $\lambda \in \mathbb{R}$. Then the following statements are equivalent:*

(i) *for all $x \in \Omega$ and $t \leq s \leq \tau < \tau_{(x,t)}$,*

$$e^{-\lambda(s-t)}u(y_{(x,t)}(s), s) - e^{-\lambda(\tau-t)}u(y_{(x,t)}(\tau), \tau) \leq \int_s^\tau e^{-\lambda(\zeta-t)}\ell(y_{(x,t)}(\zeta), \zeta)d\zeta;$$

(ii) $u_t(x,t) - \lambda u(x,t) + f(x) \cdot D_x u(x,t) + \ell(x,t) \geq 0$, $(x,t) \in \Omega \times [0,T[$, *in the viscosity sense;*

(iii) $-u_t(x,t) + \lambda u(x,t) - f(x) \cdot D_x u(x,t) - \ell(x,t) \leq 0$, $(x,t) \in \Omega \times [0,T[$, *in the viscosity sense.*

Proof. At first let us observe that it is not restrictive to assume $\lambda = 0$. Indeed, u satisfies (ii) if and only if $\hat{u}(x, x_{d+1}, t) := x_{d+1}u(x,t)$ is a viscosity supersolution of

$$\hat{f} \cdot D_x \hat{u} + \hat{\ell} \geq 0 \quad \text{in } \Omega \times \mathbb{R}_+ \times [0,T],$$

where $\hat{f}(x, x_{d+1}) := (f(x), -\lambda x_{d+1})$ and $\hat{\ell}(x, x_{d+1}, t) := x_{d+1}\ell(x,t)$ (see Lemma A.1.2, (*ii*), and Remark A.1.1). On the other hand, it is easy to check that (*i*) is equivalent to

$$\hat{u}(\hat{y}(s), s) - \hat{u}(\hat{y}(\tau), \tau) \leq \int_s^\tau \hat{\ell}(\hat{y}(\zeta), \zeta)d\zeta, \quad t \leq s \leq \tau,$$

where $\hat{y}$ is the solution of

$$\begin{cases} \hat{y}'(\tau) = \hat{f}(\hat{y}(\tau)) = (f(y(\tau)), -\lambda y_{d+1}(\tau)) \\ \hat{y}(t) = (x, 1) \end{cases}.$$

Let us prove that (*i*) implies (*ii*). Assume then that (*i*) holds with $\lambda = 0$. It is not hard to show that, for $(x,t) \in \Omega \times [0, T[$ and $s < t$, $|s - t|$ small enough,

$$u(y_{(x,t)}(s), s) - u(x,t) \leq \int_s^t \ell(y_{(x,t)}(\zeta), \zeta)d\zeta. \tag{A.1.3}$$

Now, if $(x,t) \in \Omega \times [0,T[$ is a local minimum for $u - \varphi$, $\varphi \in C^1(\Omega \times [0,T])$, then

$$\varphi(y_{(x,t)}(s), s) - \varphi(x,t) \leq u(y_{(x,t)}(s), s) - u(x,t) \tag{A.1.4}$$

for $|s - t|$ small enough. Combining (A.1.3) and (A.1.4) we obtain

$$\varphi(y_{(x,t)}(s), s) - \varphi(x,t) \leq \int_s^t \ell(y_{(x,t)}(\zeta), \zeta)d\zeta.$$

Dividing this by $s - t$ and letting $s \to t$ we conclude that $-\varphi_t(x,t) - f(x) \cdot D_x\varphi(x,t) \leq \ell(x,t)$ and (*ii*) is proved.

To prove the reverse implication, let us assume first that $(x_0, t_0) \in \Omega \times [0, T[$ is a local minimum for $u - \varphi$ with $\varphi \in C^1(\Omega \times [0,T])$. Then (*ii*) gives

$$u_t(x_0, t_0) + f(x_0) \cdot D_x u(x_0, t_0) + \ell(x_0, t_0) \geq 0 \tag{A.1.5}$$

in the viscosity sense. Now if $f(x_0) = 0$, then $y_{(x_0,t_0)}(\tau) \equiv x_0$. Inequality (*i*) reduces in this case to

$$u(x_0, t_0) - u(x_0, \tau) \leq \int_{t_0}^\tau \ell(x_0, \zeta)d\zeta.$$

Dividing both members by $\tau - t_0$ and letting $\tau \to t_0$, we get

$$-u_t(x_0, t_0) \leq \ell(x_0, t_0),$$

which is clearly implied by (A.1.5).

Consider now the case $f(x_0) \neq 0$. Under the hypotheses on f, by classical results on ordinary differential equations, there exists a local diffeomorphism Φ such that, in the new coordinates, $\xi = \Phi(x)$ system (A.1.1) becomes

$$\begin{cases} \xi'(\tau) = e_1 = (1, 0, \dots, 0) \\ \xi(t) = x_0 \end{cases} . \tag{A.1.6}$$

The change of coordinates in (ii) implies that $u(\Phi^{-1}(\xi), t)$ satisfies

$$u_t(\Phi^{-1}(\xi), t) + f(\Phi^{-1}(\xi)) \cdot J\Phi(\Phi^{-1}(\xi)) D_x u(\Phi^{-1}(\xi), t) \\ + \ell(\Phi^{-1}(\xi), t) \geq 0$$

in the viscosity sense. Due to (A.1.6), this gives

$$u_t(\Phi^{-1}(\xi), t) + \frac{\partial u}{\partial \xi_1}(\Phi^{-1}(\xi), t) + \ell(\Phi^{-1}(\xi), t) \geq 0.$$

Using now Lemmas A.1.3 and A.1.4 (see also Remark A.1.2) we conclude that

$$u(\Phi^{-1}(s, 0, \dots, 0), s) - u(\Phi^{-1}(\tau, 0, \dots, 0), \tau) \\ \leq \int_s^\tau \ell(\Phi^{-1}(\zeta, 0, \dots, 0), \zeta) d\zeta$$

for $t \leq s \leq \tau$, with s, τ sufficiently close to t.

By definition of Φ, this is the same as

$$u(y_{(x_0,t)}(s), s) - u(y_{(x_0,t)}(\tau), \tau) \leq \int_s^\tau \ell(y_{(x_0,t)}(\zeta), \zeta) d\zeta,$$

for $t \leq s \leq \tau$, s, τ sufficiently close to t. A simple continuation argument shows the validity of (i) for any $t \leq s \leq \tau < \tau_{(x,t)}$.

Up to now we have proved that (i) holds for any $(x, t) \in A^- = \{(x, t) \in \Omega \times [0, T[: D^- u(x, t) \neq \emptyset\}$ (recall Lemma A.1.1). Since A^- is dense in $\Omega \times [0, T[$ (see Lemma A.1.2, (i)), we conclude the validity of (i) for all $(x, t) \in \Omega \times [0, T[$ using the continuous dependence of the solution of (A.1.1) with respect to the initial datum.

The equivalence between (i) and (iii) can be proved similarly. $\square$

Let us now consider the controlled system (1.1.7). We set, for $(x, t) \in \Omega \in [0, T]$ and $\alpha \in \mathcal{A}$,

$$\tau_{(x,t)}(\alpha) := \min\{\inf\{\tau \geq t : y_{(x,t)}(\tau; \alpha) \notin \Omega\}, T\}.$$

If we use constant controls $\alpha(\tau) \equiv a$ in (1.1.7), then the equivalence between

$$e^{-\lambda(s-t)}u(y_{(x,t)}(s;a)) - e^{-\lambda(\tau-t)}u(y_{(x,t)}(\tau;a)) \leq \int_s^\tau e^{-\lambda(\zeta-t)}\ell(y_{(x,t)}(\zeta;a),a,\zeta)d\zeta$$

for every $x \in \Omega$, $a \in A$ and $t \leq s \leq \tau < \tau_{(x,t)}(a)$,

$$-u_t(x,t) + \lambda u(x,t) + H(x,t,D_x u(x,t)) \leq 0 \tag{A.1.7}$$

for every $(x,t) \in \Omega \times [0,T[$, in the viscosity sense, and

$$u_t(x,t) - \lambda u(x,t) - H(x,t,D_x u(x,t)) \geq 0 \tag{A.1.8}$$

for every $(x,t) \in \Omega \times [0,T[$, in the viscosity sense, is a straightforward consequence of Lemma A.1.5. A repeated application of Lemma A.1.5 shows that (A.1.7) and (A.1.8) are equivalent to

$$e^{-\lambda(s-t)}u(y_{(x,t)}(s;\alpha),s) - e^{-\lambda(\tau-t)}u(y_{(x,t)}(\tau;\alpha),\tau) \leq \int_s^\tau e^{-\lambda(\zeta-t)}\ell(y_{(x,t)}(\zeta;\alpha),\alpha(\zeta),\zeta)d\zeta \tag{A.1.9}$$

for all $x \in \Omega$, $t \leq s \leq \tau < \tau_{(x,t)}(\alpha)$, $\alpha \in \mathcal{PC}$, where $\mathcal{PC} \subset \mathcal{A}$ is the class of piecewise constant controls. To prove the equivalence between (A.1.7), (A.1.8), (A.1.9) for general controls $\alpha \in \mathcal{A}$, that is Lemma 1.1.1 in §1.1.3 and Lemma A.1.7 below, we need the following

Lemma A.1.6. *Let $\alpha \in \mathcal{A}$, $(x,t) \in \Omega \times [0,T]$ and $y(\tau) = y_{(x,t)}(\tau;\alpha)$ be the corresponding solution of* (1.1.7). *Under the hypotheses on f in §1.1.3, for every $T > 0$ there exists a sequence $\{\alpha_n\} \subset \mathcal{A}$ such that*

$$\begin{cases} \alpha_n \textit{ is piecewise constant on } [t,T], \\ |\alpha_n(\tau) - \alpha(\tau)| < 1/n \quad \textit{for every } \ \tau \in E_n \subset [t,T], \\ \qquad\qquad E_n \textit{ compact and } \mathcal{L}([t,T] \setminus E_n) < 1/n \end{cases}, \tag{A.1.10}$$

$$y_n \longrightarrow y \quad \textit{uniformly in } [t,T], \tag{A.1.11}$$

where $y_n(\tau) = y_{(x,t)}(\tau;\alpha_n)$.

Proof. Assertion (A.1.10) is a consequence of Lusin's Theorem (see Theorem B.2.1). To prove (A.1.11) observe that

$$|y_n(\tau) - y(\tau)| \leq L\int_t^\tau |y_n(s) - y(s)|ds + \int_t^\tau |B_n(s)|ds,$$

where L is as in (1.1.8) and

$$B_n(s) = f(y(s), \alpha_n(s)) - f(y(s), \alpha(s)).$$

Hence, by Gronwall's inequality, for $\tau \in [t, T]$

$$|y_n(\tau) - y(\tau)| \leq \int_t^T |B_n(s)|ds + Le^{LT} \int_t^T \int_s^T |B_n(\zeta)|d\zeta ds. \qquad \text{(A.1.12)}$$

From (A.1.10), $\alpha_n(\tau) \longrightarrow \alpha(\tau)$ a.e. in $[t, T]$ (at least for a subsequence). Hence by continuity

$$B_n(\zeta) \longrightarrow 0 \quad \text{a.e. in } [t, T].$$

Moreover, by the continuity and the boundedness of f on $\Omega \times A$,

$$|B_n(\zeta)| \leq C \quad \text{for every } \zeta \in [t, T],\ n \in \mathbb{N}\,.$$

Assertion (A.1.11) now follows from (A.1.12) and the Dominated Convergence Theorem. □

Lemma A.1.7. *Let us assume $\ell \in C^0(\Omega \times A \times [0, T])$, ℓ bounded, $\lambda \in \mathbb{R}$ and $u \in C^0(\Omega \times [0, T])$. Then the following statements are equivalent:*

(*i*) *for all $x \in \Omega$, $\alpha \in \mathcal{A}$ and $t \leq s \leq \tau < \tau_{(x,t)}(\alpha)$,*

$$e^{-\lambda(s-t)}u(y_{(x,t)}(s; \alpha), s) - e^{-\lambda(\tau-t)}u(y_{(x,t)}(\tau; \alpha), \tau) \leq \int_s^\tau e^{-\lambda(\zeta-t)}\ell(y_{(x,t)}(\zeta; \alpha), \alpha(\zeta), \zeta)d\zeta,$$

(*ii*) *$u_t(x, t) - \lambda u(x, t) - H(x, t, D_x u(x, t)) \geq 0$, $(x, t) \in \Omega \times [0, T[$, in the viscosity sense,*

(*iii*) *$-u_t(x, t) + H(x, t, D_x u(x, t)) \leq 0$, $(x, t) \in \Omega \times [0, T[$, in the viscosity sense,*

where H is as in §1.1.3.

Proof. The discussion before Lemma A.1.6 shows that in order to prove Lemma A.1.7 is sufficient to show that (A.1.9) implies (*i*) for any $u \in C^0(\Omega \times [0, T])$ (the reverse implication being trivial). To this aim, let $\alpha \in \mathcal{A}$ and take α_n and y_n as in Lemma A.1.6. By (A.1.9),

$$e^{-\lambda(s-t)}u(y_n(s), s) - e^{-\lambda(\tau-t)}u(y_n(\tau), \tau) \leq \int_s^\tau e^{-\lambda(\zeta-t)}\ell(y_n(\zeta), \alpha_n(\zeta), \zeta)d\zeta$$

for $t \leq s \leq \tau < \tau_{(x,t)}(\alpha_n)$. Since $\alpha_n \longrightarrow \alpha$ almost everywhere and $y_n \longrightarrow y_{(x,t)}(\cdot; \alpha)$ uniformly on compact intervals by Lemma A.1.6, letting $n \to +\infty$ in the previous inequality, we get the desired result. □

A.2 On the uniqueness of the equilibrium in Ch. 3, §3.4

In this section, we show two examples that do not necessarily meet in all their aspects the model studied in Ch. 3. They are just inspiring examples about the possible use of a monotonicity property in order to prove the uniqueness of the equilibrium.

We first recall that, as in §3.4, a mean-field equilibrium is a function $\rho \in L^2$ such that $\rho \in \psi(\rho)$, which means that ρ is a juxtaposed convex combination of the extremal evolutions generated by ρ itself via the optimization functions P (3.1.6).

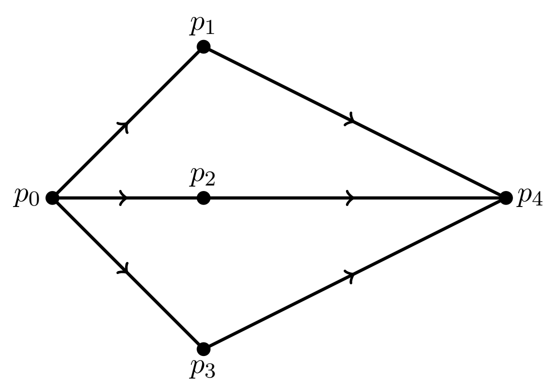


Figure A.1: The network of Example A.2.1

Example A.2.1. *Consider the network in Figure A.1, where the goal is to start from p_0 and to arrive to p_4, along the three possible paths: $p_0 \to p_1 \to p_4$, $p_0 \to p_2 \to p_4$ and $p_0 \to p_3 \to p_4$. Moreover, we suppose that all the agents at the time $t = 0$ are on p_0, that at the time $t = 1$ they are all forced to switch to one of the three nodes p_1, p_2 and p_3, and that at the time $t = T = 2$ they are all forced to switch to p_4, ending the game. Since the switching instants are fixed and the significant nodes are just p_1, p_2 and p_3, we only give the cost of stay on such nodes respectively, independently of time: $C_1(\rho_1) = \rho_1, C_2(\rho_2) = 2\rho_2, C_3(\rho_3) = 3\rho_3$, where ρ_i is the mass in the node p_i. In this case a mean-field equilibrium is given by $(\lambda_1, \lambda_2, \lambda_3) = (6/11, 3/11, 2/11)$, which means that, denoted by ρ_0 the initial distribution in p_0, at time $t = 1$ the fraction $\lambda_i\rho_0$ switches to the node p_i, $i = 1, 2, 3$. Indeed, with these fractions all the costs C_1, C_2, C_3 are equal to $(6/11)\rho_0$. Hence if all the agents in p_0 conjecture such a distribution,* *then all the possible generated extremal distributions are the following ones:* $(\rho_0, 0, 0), (0, \rho_0, 0), (0, 0, \rho_0)$, *that is all the switches are optimal. The actual mass* $(\lambda_1\rho_0, \lambda_2\rho_0, \lambda_3\rho_0)$ *is then a convex combination of the generated extremal distributions with convex coefficients* $(\lambda_1, \lambda_2, \lambda_3)$, *and hence it is a mean field equilibrium. By linearity of the costs, the coefficients* λ_i *are easily calculated by imposing* $C_1(\lambda_1) = C_2(\lambda_2) = C_3(\lambda_3)$ *with the constraint* $\lambda_i \in [0, 1]$ *and* $\lambda_1 + \lambda_2 + \lambda_3 = 1$, *and they are the only ones satisfying the system and the constraint. Note that if, for example, we are looking for a possible equilibrium using just the nodes* p_1 *and* p_2, *that is we look for* $\lambda_1, \lambda_2 \geq 0$, $\lambda_1 + \lambda_2 = 1$ *and* $C_1(\lambda_1) = C_2(\lambda_2)$, *we find* $\lambda_1 = 2/3, \lambda_2 = 1/3$ *and then we have the distribution* $(\lambda_1\rho_0, \lambda_2\rho_0, 0)$. *But such a distribution is not an equilibrium because it gives the costs* $((2/3)\rho_0, (2/3)\rho_0, 0)$, *which generates the only extremal distribution* $(0, 0, \rho_0)$: *all agents switch to* p_3. *And* $(\lambda_1\rho_0, \lambda_2\rho_0, 0)$ *is not a convex combination of (i.e., is not equal to) the singleton* $\{(0, 0, \rho_0)\}$. *The problem then has a unique equilibrium which is given by* $((6/11)\rho_0, (3/11)\rho_0, (2/11)\rho_0)$.

Note that, whenever we find a triple of convex coefficients $(\lambda_1, \lambda_2, \lambda_3)$ *such that* $C_1(\lambda_1) = C_2(\lambda_2) = C_3(\lambda_3)$, *then the corresponding distribution* $(\lambda_1\rho_0, \lambda_2\rho_0, \lambda_3\rho_0)$ *is an equilibrium because it gives the same costs along any path, and then generates all the extremal distributions* $(\rho_0, 0, 0)$, $(0, \rho_0, 0)$, $(0, 0, \rho_0)$ *of which it is a convex combination. The question about uniqueness is then: given three functions* $C_i : [0, 1] \to \mathbb{R}$, $i = 1, 2, 3$, *under which condition there exists at most one triple of convex coefficients* $(\lambda_1, \lambda_2, \lambda_3)$ *such that*

$$C_1(\lambda_1\rho_0) = C_2(\lambda_2\rho_0) = C_3(\lambda_3\rho_0)? \tag{A.2.13}$$

A condition that guarantees such a uniqueness is the following monotonicity property which is, in our discrete case, the condition in [31]:

$$\begin{gathered}\sum_{i=1}^{3} \left(C_i(\lambda_i'\rho_0) - C_i(\lambda_i''\rho_0)\right)(\lambda_i' - \lambda_i'') > 0 \\ \text{for all } (\lambda_1', \lambda_2', \lambda_3') \neq (\lambda_1'', \lambda_2'', \lambda_3'') \text{ convex triples and for any } \rho_0 > 0.\end{gathered} \tag{A.2.14}$$

Indeed, let us suppose that there are two convex triples

$$(\lambda_1', \lambda_2', \lambda_3') = (\lambda_1', \lambda_2', 1 - \lambda_1' - \lambda_2'),\ (\lambda_1'', \lambda_2'', \lambda_3'') = (\lambda_1'', \lambda_2'', 1 - \lambda_1'' - \lambda_2'')$$

satisfying (A.2.13), *and denoting by* C', C'' *the common costs, for the single triple respectively, we obtain*

$$\sum_{i=1}^{3}(C_i' - C_i')(\lambda_i' - \lambda_i'')$$
$$= (C' - C'')\sum_{i=1}^{2}(\lambda_i' - \lambda_i'') + (C' - C'')(1 - \lambda_1' - \lambda_2' - 1 + \lambda_1'' + \lambda_2'') = 0$$

and hence, by (A.2.14), $(\lambda_1', \lambda_2', \lambda_3') = (\lambda_1'', \lambda_2'', \lambda_3'')$.

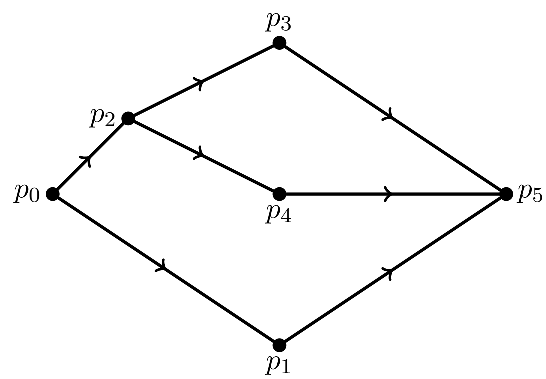


Figure A.2: The network of Example A.2.2

Example A.2.2. *Consider the network in Figure A.2. The goal is to start from* p_0 *and to reach* p_5 *among one of the possible paths* $p_0 \to p_1 \to p_5$, $p_0 \to p_2 \to p_3 \to p_5$ *and* $p_0 \to p_2 \to p_4 \to p_5$*. Again, the agents at* $t = 0$ *are all on* p_0*, with distribution* ρ_0*, at time* $t = 1$ *they are forced to switch to* p_1 *or to* p_2*, at time* $t = 3/2$ *the agents on* p_2 *are forced to switch to* p_3 *or* p_4 *and at the time* $t = T = 2$ *they are all forced to switch to* p_5*. The costs are* $C_1(\rho_1) = \rho_1$, $C_2(\rho_2) = 4\rho_2$, $C_3(\rho_3) = 3\rho_3$, $C_4(\rho_4) = 2\rho_4$*. Moreover, the costs are also multiplied by the amount of the time spent on the node. We denote by* $(\lambda_1, \lambda_2, \lambda_{2,3}, \lambda_{2,4})$ *the coefficients of a possible equilibrium, that is: at* $t = 1$ *the fraction given by* $\lambda_1\rho_0$ *switches to* p_1 *and the fraction given by* $\lambda_2\rho_0$ *switches to* p_2*; at time* $t = 3/2$*, the fraction* $\lambda_2\lambda_{2,3}\rho_0$ *switches from* p_2 *to* p_3 *and the fraction* $\lambda_2\lambda_{2,4}\rho_0$ *switches from* p_2 *to* p_4*. Still by linearity of the costs, such coefficients are founded by solving*

$$\begin{cases} 2\lambda_2 + \frac{3}{2}\lambda_2\lambda_{2,3} = \lambda_1 \\ 2\lambda_2 + \lambda_2\lambda_{2,4} = \lambda_1 \\ \lambda_1 + \lambda_2 = \lambda_{2,3} + \lambda_{2,4} = 1 \end{cases}, \tag{A.2.15}$$

which corresponds to, taking also account of the time spent on the node,

$$\begin{cases} \frac{C_2(\lambda_2\rho_0)}{2} + \frac{C_3(\lambda_2\lambda_{2,3}\rho_0)}{2} = C_1(\lambda_1\rho_0) \\ \frac{C_2(\lambda_2\rho_0)}{2} + \frac{C_4(\lambda_2\lambda_{2,4}\rho_0)}{2} = C_1(\lambda_1\rho_0) \\ \lambda_1 + \lambda_2 = \lambda_{2,3} + \lambda_{2,4} = 1 \end{cases} . \tag{A.2.16}$$

From (A.2.15)*, we obtain the unique solution*

$$(\lambda_1, \lambda_2, \lambda_{2,3}, \lambda_{2,4}) = \left(\frac{13}{18}, \frac{5}{18}, \frac{2}{5}, \frac{3}{5}\right).$$

This is an equilibrium because it generates the distribution

$$\left(\frac{13}{18}\rho_0, \frac{5}{18}\rho_0, \frac{1}{9}\rho_0, \frac{1}{6}\rho_0\right), \tag{A.2.17}$$

which gives the cost, for each one of the three paths, equal to 13/18*. Hence all the paths are equivalent and the distribution generates all the possible extremal evolutions* $(\rho_0, 0, 0, 0)$*,* $(0, \rho_0, \rho_0, 0)$*,* $(0, \rho_0, 0, \rho_0)$ *of which* (A.2.17) *is a juxtaposed convex combination.*

Similarly as in (A.2.14)*, the uniqueness of the solution of* (A.2.16) *is guaranteed by the following monotonicity conditions*

$$\begin{cases} \sum_{i=3}^{4} \Big(C_i(\lambda\lambda'_{2,i}\rho_0) - C_i(\lambda\lambda''_{2,i}\rho_0)\Big)(\lambda'_{2,i} - \lambda''_{2,i}) > 0 \quad \textit{for every } \lambda > 0, \\ \\ \end{cases}$$
$$\begin{cases} (C_1(\lambda'_1\rho_0) - C_1(\lambda''_1\rho_0))(\lambda'_1 - \lambda''_1) + \frac{1}{2}\Big(C_2(\lambda'_2\rho_0) + C_3(\lambda'_2\lambda'_{2,3}\rho_0) \\ -C_2(\lambda''_2\rho_0) - C_3(\lambda''_2\lambda''_{2,3}\rho_0)\Big)(\lambda'_2 - \lambda''_2) > 0 \end{cases}$$
$$\forall (\lambda'_1, \lambda'_2) \neq (\lambda''_1, \lambda''_2),\ (\lambda'_{2,3}, \lambda'_{2,4}) \neq (\lambda''_{2,3}, \lambda''_{2,4}) \textit{ convex pairs and } \rho_0 > 0. \tag{A.2.18}$$

Indeed, by the second inequality we have the uniqueness of the pair of convex coefficients (λ_1, λ_2)*, which, putting* $\lambda = \lambda_2$ *in the first inequality, gives the uniqueness of the pair* $(\lambda_{2,3}, \lambda_{2,4})$*.*

Remark A.2.1. *Similarly as in Example A.2.1 (see Figure A.1) when the number of the nodes is* n *instead of* 3*, the uniqueness of the* n*-string of convex coefficients satisfying* $C_i(\lambda_i\rho_0) = C_j(\lambda_j\rho_0)$ *for all* $i, j = 1, \dots, n$ *is guaranteed by the monotonicity conditions as* (A.2.14)*, replacing* $n = 3$ *by the generic* n*. As seen in Example A.2.2 (see Figure A.2), in the case of more complex networks, the conditions are much more involved and less* *treatable, because of the peculiar characteristics of the problem. The topology of the network in fact strongly affects the monotonicity property, the way of representing it and, ultimately, its applicability. However, we point out that if all the single costs C_i are strictly monotone, then they will certainly satisfy the corresponding monotonicity property.*

Remark A.2.2. *In the two examples here presented, the switching instants are a priori fixed for all agents, and hence they do not enter in the optimization process performed by the single agent. In the model in Ch. 3, we instead consider also the switching time as well as the decisional time as part of the control for the agents, and the costs also depend on them. This fact obviously makes the situation much more complicated in order to establish a reasonable condition for the uniqueness of the mean-field game.*

Remark A.2.3. *The monotonicity conditions* (A.2.14) *and* (A.2.18) *and their possible generalization to more complicated networks, guarantee only the uniqueness of the possible n-string of convex coefficients but not, in general, its existence. Note that, if the (unique) solution presents some $\lambda_i = 0$, then it means that the corresponding node will be not reached by the equilibrium, but anyway, even with zero mass, that node produces the same cost as the others. Moreover, we may not have existence of the n-string convex solution. Looking at Example A.2.1 (generalized to n intermediate nodes), this means that we do not have a n-string which gives the fraction of mass switching to the n nodes. That is there is at least a node which must be not considered in the game from the beginning. For example, a node p_i such that $C_i(\lambda_i\rho_0) > C_j(\lambda_j\rho_0)$ for all $j \neq i$ and λ_i, λ_j: it is a too expensive node, no one will switch to it. In this situation, the actual game is with just $n-1$ nodes and not with n. Hence, one must look for a possible unique $(n-1)$-string of convex combination solving the corresponding problem without that node. Proceeding in this way, one can find a possible unique m-string, and will set the other components to 0: no flow through such nodes. However note that, in the model in Ch. 3, we have also the time spent on the node at our disposal, which possibly modulate the paid cost, and hence the situation is more flexible but less prone to have a good condition for uniqueness.*

The points and the questions of these last remarks are certainly worth investigating and may be the argument of future studies.

A.3 On the convexity of V and single-valued feature of φ in Ch. 3, §3.1.1 and §3.4

In this section, we show an example of a possible cost which guarantees the convexity of the value function V and the single-valued feature of the map φ in Ch. 3, §3.1.1 and §3.4.

Let us assume

$$C(p,p',t,\tau,\rho)=\frac{\bar{C}(p,p',\rho)}{\tau-t},\quad \tau\longmapsto\tilde{C}(\bar{p},\tau)\text{ strictly decreasing.}$$

In particular, $\bar{C}$ does not explicitly depend on t and τ, for example

$$\bar{C}(p,p',\rho)=\frac{a(p)}{T}\int_0^T\rho_p(s)ds+\frac{a(p')}{T}\int_0^T\rho_{p'}(s)ds$$

for some weight $p\longmapsto a(p)$. A possible strictly non-decreasing $\tilde{C}$ is $\tilde{C}(\bar{p},\tau)=T-\tau$.

Let p_1 be a node directly linked to $\bar{p}$, i.e. $\sum_i p_1^i=N-1$, and let $t<T$. Hence we have

$$V(p_1,t)=\inf_{\tau\in]t,T]}\left\{\frac{\bar{C}(p_1,\bar{p},\rho)}{\tau-t}+\tilde{C}(\bar{p},\tau)\right\}=\frac{\bar{C}(p_1,\bar{p},\rho)}{T-t}+\tilde{C}(\bar{p},T).$$

Therefore, $t\longmapsto V(p_1,t)$ is strictly convex and $\varphi(t)=T$ is single-valued.

Now, let p_2 be a node linked to $\bar{p}$ with two switches, i.e. $\sum_i p_2^i=N-2$, and let $p_1\in\mathcal{I}_{p_2}$ and $t<T$. We consider the function

$$\begin{aligned}\psi_{p_2,p_1}:]t,T[\ni\tau\longmapsto V(p_1,\tau)+\frac{\bar{C}(p_2,p_1,\rho)}{\tau-t}&\\ =\frac{\bar{C}(p_1,\bar{p},\rho)}{T-\tau}+\tilde{C}(\bar{p},T)+\frac{\bar{C}(p_2,p_1,\rho)}{\tau-t}&.\end{aligned}$$

Note that $\lim_{\tau\to t^+}\psi_{p_2,p_1}(\tau)=\lim_{\tau\to T^-}\psi_{p_2,p_1}(\tau)=+\infty$. Hence, the minimization problem

$$\inf_{\tau\in]t,T[}\psi_{p_2,p_1}(\tau)$$

has a solution $\varphi_{p_2,p_1}(t)\in]t,T[$ and it must be

$$\frac{\bar{C}(p_1,\bar{p},\rho)}{(T-\varphi_{p_2,p_1}(t))^2}-\frac{\bar{C}(p_2,p_1,\rho)}{(\varphi_{p_2,p_1}(t)-t)^2}=0,\tag{A.3.19}$$

which gives a unique possible point of minimum

$$\varphi_{p_2,p_1}(t) = \frac{\sqrt{\frac{\bar{C}(p_2,p_1,\rho)}{\bar{C}(p_1,\bar{p},\rho)}}T + t}{\sqrt{\frac{\bar{C}(p_2,p_1,\rho)}{\bar{C}(p_1,\bar{p},\rho)}} + 1} \in]t, T[,$$

and note that φ is strictly increasing and linear and hence derivable. Moreover, its derivative satisfies

$$0 < \varphi'_{p_2,p_1}(t) < 1. \tag{A.3.20}$$

We now consider the function

$$V_{p_2,p_1} : t \longmapsto \psi_{p_2,p_1}(\varphi_{p_2,p_1}(t)) = \frac{\bar{C}(p_1,\bar{p},\rho)}{T - \varphi_{p_2,p_1}(t)} + \tilde{C}(\bar{p},T) + \frac{\bar{C}(p_2,p_1,\rho)}{\varphi_{p_2,p_1}(t) - t},$$

which represents the optimum when, being on p_2 at time t, the agent decides that it will switch to p_1 before T, that is it will perform the path $p_2 \to p_1 \to \bar{p}$. Such a function is then twice derivable and it is strictly convex in $]0, T[$. Indeed, taking account of (A.3.19) and (A.3.20), it is

$$V''_{p_2,p_1}(t) = \frac{2\bar{C}(p_2,p_1,\rho)(\varphi_{p_2,p_1}(t) - t)(1 - \varphi'_{p_2,p_1}(t))}{(\varphi_{p_2,p_1}(t) - t)^4} > 0.$$

Note that we do not need the second derivative of φ_{p_2,p_3} (even if it exists, in our example) because in the calculation of V'_{p_2,p_1} it cancels in view of (A.3.19). Finally, note that $\lim_{t\to T^-} V_{p_2,p_1}(t) = +\infty$.

Now, we take p_3 such that $p_2 \in \mathcal{I}_{p_3}$ and consider the function

$$V_{p_3,p_1,p_1} : t \longmapsto \inf_{\tau\in]t,T[} \left\{ V_{p_2,p_1}(\tau) + \frac{\bar{C}(p_3,p_2,\rho)}{\tau - t} \right\},$$

which represents the optimum when, being on p_3 at time t, the agent decides that it will perform the path $p_3 \to p_2 \to p_1 \to \bar{p}$. Note that the function $]t,T[\ni \tau \longmapsto \psi_{p_3,p_2,p_1}$, inside the minimization, is twice derivable and satisfies $\lim_{\tau\to t^+} \psi_{p_3,p_2,p_1}(\tau) = \lim_{\tau\to T^+} \psi_{p_3,p_2.p_4}(\tau) = +\infty$. Hence the minimization process has a solution $\varphi_{p_3,p_2,p_1}(t) \in]t,T[$, and such a solution is unique. Indeed, again, it must be

$$V'_{p_2,p_1}(\varphi_{p_3,p_2,p_1}(t)) = \frac{\bar{C}(p_3,p_2,\rho)}{(\varphi_{p_3,p_2,p_1}(t) - t)^2}. \tag{A.3.21}$$

Whereas $\tau \longmapsto V'_{p_2,p_1}(\tau)$ is strictly increasing (being V_{p_2,p_1} strictly convex) and $\tau \longmapsto \bar{C}(p_3,p_2,\rho)/(\tau-t)^2$ is strictly decreasing, the solution $\varphi_{p_3,p_2,p_1}(t) \in]t,T[$ is unique. Moreover, by the Implicit Function Theorem, φ_{p_3,p_2,p_1} is derivable. Differentiating the equality (A.3.21), we get (we write φ for φ_{p_3,p_2,p_1})

$$\left(V''_{p_2,p_1}(\varphi(t) + \frac{2\bar{C}(p_3,p_1,\rho)(\varphi(t)-t)}{(\varphi(t)-t)^4}\right)\varphi'(t) = \frac{2\bar{C}(p_3,p_2,\rho)(\varphi(t)-t)}{(\varphi(t)-t)^4},$$

from which, being $V''_{p_2,p_1} > 0$ and $\varphi(t) > t$, we get

$$0 < \varphi'_{p_3,p_2,p_1}(t) < 1 \tag{A.3.22}$$

and in particular φ_{p_1,p_2,p_3} is strictly increasing. Now, we prove that (still denoting φ_{p_3,p_2,p_1} by φ)

$$t \longmapsto V_{p_3,p_2,p_1}(t) = V_{p_2,p_1}(\varphi(t)) + \frac{\bar{C}(p_3,p_2,\rho)}{\varphi(t)-t}$$

is strictly convex. Indeed, differentiating two times, taking account of (A.3.21) and (A.3.22), we get again

$$V''_{p_3,p_2,p_1}(t) = \frac{2\bar{C}(p_3,p_2,\rho)(\varphi(t)-t)(1-\varphi'(t))}{(\varphi(t)-t)^4} > 0.$$

Again, note that we do not need the second derivative of φ (even if it exists, in our example) because in the calculation of V'_{p_3,p_2,p_1} it cancels in view of (A.3.21). Finally note that $\lim_{t\to T^-} V_{p_3,p_2,p_1}(t) = +\infty$.

Proceeding in this way we obtain that, for every path $p_n \to p_{n-1} \to \cdots \to p_1 \to \bar{p}$, the function

$$V_{p_n,p_{n-1},\ldots,p_1}(t) = \inf_{\tau\in]t,T[} \left\{ V_{p_{n-1},\ldots,p_1}(t) + \frac{\bar{C}(p_n,p_{n-1},\rho)}{\tau-t} \right\}$$

is realized by a unique $\tau = \varphi_{p_n,\ldots,p_1}(t) \in]t,T[$, it is strictly convex, and $\varphi_{p_n,\ldots,p_1}$ is strictly increasing with derivative less than 1.

We finally obtain that the value function, for all $p \neq \bar{p}$ and $t < T$,

$$V(p,t) = \inf_{\substack{\tau\in]t,T] \\ p'\in\mathcal{I}_p}} \left\{ V(p',\tau) + \frac{\bar{C}(p,p',\rho)}{\tau-t} \right\},$$

is realized by a unique, strictly increasing (for t such that $\varphi(t) > T$) single-valued function $t \longmapsto \tau = \varphi(t) \in]t,T]$, giving the optimal instant $\tau \in]t,T]$ for switching to the optimal node $p' \in \mathcal{I}_p$.

Appendix B

Mathematical tools

B.1 Viscosity solutions: definition and first properties

In this section, we briefly recall the definition and the basic properties of continuous viscosity solutions of the Hamilton-Jacobi equation

$$F(x, u(x), D_x u(x)) = 0, \quad x \in \Omega, \tag{B.1.1}$$

where Ω is an open domain of $\mathbb{R}^d$ and the Hamiltonian $F = F(x, r, p)$ is a continuous real valued function on $\Omega \times \mathbb{R} \times \mathbb{R}^d$.

Definition B.1.1. *A function $u \in C^0(\Omega)$ is a viscosity subsolution of* (B.1.1) *if, for any $\varphi \in C^1(\Omega)$,*

$$F(x_0, u(x_0), D_x \varphi(x_0)) \le 0 \tag{B.1.2}$$

at any local maximum point $x_0 \in \Omega$ of $u - \varphi$. Similarly, $u \in C^0(\Omega)$ is a viscosity supersolution of (B.1.1) *if, for any $\varphi \in C^1(\Omega)$,*

$$F(x_1, u(x_1), D_x \varphi(x_1)) \ge 0 \tag{B.1.3}$$

at any local minimum point $x_1 \in \Omega$ of $u - \varphi$. Finally, u is a viscosity solution of (B.1.1) *if it is simultaneously a viscosity sub- and supersolution.*

The definition applies also to evolutionary Hamilton-Jacobi equation of the form

$$u_t(y, t) + F(y, t, u(y, t), D_x u(y, t)) = 0, \quad (y, t) \in D \times]0, T[, \; T > 0 \tag{B.1.4}$$

In fact, equation (B.1.4) is reduced to the form (B.1.1) by setting

$$x = (y, t) \in \Omega = D \times]0, T[\subseteq \mathbb{R}^{N+1}, \quad \tilde{F}(x, r, q) = q_{N+1} + F(x, r, q_1, \dots, q_N)$$

with

$$q = (q_1, \dots, q_N, q_{N+1}) \in \mathbb{R}^{N+1}.$$

In Ch. 1, §1.1.3 (see also Appendix A, §A.1) and §1.1.5, and Ch. 2, §2.1.3, we focus on evolutionary Hamilton-Jacobi equations of the type (B.1.4). A special attention is dedicated to the case where $\tilde{F}(x, r, q)$ is of the form

$$\begin{aligned}\tilde{F}(y, t, r, q) &= q_{N+1} + r + H(y, t, q_1, \dots, q_N) \\ &= q_{N+1} + r + \sup_{a \in A} \{-f(y, a) \cdot (q_1, \dots, q_N) - \ell(y, a, t)\}.\end{aligned}$$

Remark B.1.1. *In the Definition B.1.1, we can always assume that x_0 is a local strict maximum (minimum) point for $u - \varphi$ (otherwise, replace $\varphi(x)$ by $\varphi(x) + |x - x_0|^2$) for subsolutions (supersolutions). Moreover, since* (B.1.2) *and* (B.1.3) *depend only on the value of $D_x\varphi$ at x_0, it is not restrictive to assume that $u(x_0) = \varphi(x_0)$. Geometrically, this means that the validity of the subsolution condition* (B.1.2) *(supersolution condition* (B.1.3)*) for u is tested on smooth functions "touching from above" ("touching from below") the graph of u at x_0.*

The following result shows the local character of the notion of viscosity solution and its consistency with the classical pointwise definition.

Proposition B.1.1.

(*a*) *If $u \in C^0(\Omega)$ is a viscosity solution of* (B.1.1) *in Ω, then u is a viscosity solution of* (B.1.1) *in Ω' for any open set $\Omega' \subset \Omega$;*

(*b*) *if $u \in C^0(\Omega)$ is a classical solution of* (B.1.1)*, that is, u is differentiable at any $x \in \Omega$ and*

$$F(x, u(x), D_x u(x)) = 0 \quad \text{for every } x \in \Omega,$$

then u is a viscosity solution of (B.1.1)*;*

(*c*) *if $u \in C^1(\Omega)$ is a viscosity solution of* (B.1.1)*, then u is a classical solution of* (B.1.1).

Proof. See for example [9], Ch. II, §1, Proposition 1.3. □

In Appendix A, §A.1, we describe an alternative way of defining viscosity solutions for equation (B.1.1) (in particular, in that case, for equation (B.1.4)) and prove the equivalence of the new definition with the one given here (for the evolutionary case). More precisely, see Lemma A.1.1 and Lemma A.1.2, Remark A.1.1 for other properties.

The next Proposition B.1.2 is on the change of unknown in (B.1.1) and in its evolutionary form (B.1.4).

Proposition B.1.2.

(*a*) *Let* $u \in C^0(\Omega)$ *be a viscosity solution of* (B.1.1) *and* $\Phi \in C^1(\mathbb{R})$ *be such that* $\Phi'(t) > 0$. *Then* $v = \Phi(u)$ *is a viscosity solution of*

$$F(x, \Psi(v(x)), \Psi'(v(x))D_x v(x)) = 0, \quad x \in \Omega,$$

where $\Psi = \Phi^{-1}$.

(*b*) *Let* $u \in C^0(\Omega)$ *be a viscosity solution of* (B.1.1) *and* $\Phi : \Omega \times \mathbb{R} \longrightarrow \mathbb{R}$ *a* C^1 *function such that*

$$\Phi_r(x, r) > 0 \quad \text{for every } (x, r) \in \Omega \times \mathbb{R}.$$

Then the function $v \in C^0(\Omega)$ *defined implicitly by*

$$\Phi(x, v(x)) = u(x),$$

is a viscosity solution of

$$\tilde{F}(x, v(x), D_x v(x)) = 0, \qquad x \in \Omega,$$

where

$$\tilde{F}(x, r, p) = F(x, \Phi(x, r), D_x\Phi(x, r) + \Phi_r(x, r)p).$$

Proof. See for example [9], Ch. II, §2, Proposition 2.5 and 2.6. □

Now, we give a comparison result concerning the evolutionary case. It gives a uniqueness result for the Cauchy problem

$$\begin{cases} u_t(x,t) + H(t, D_x u(x,t)) = 0, & (x,t) \in \mathbb{R}^d \times]0, T[\\ u(x, 0) = u_0(x), & x \in \mathbb{R}^d \end{cases},$$

with initial condition $u_0 \in \mathrm{BUC}(\mathbb{R}^d)$. See Remark B.1.2 for more general Hamiltonians.

Theorem B.1.1. *Assume $H \in C^0(\mathbb{R}^d \times [0,T])$. Let $u_1, u_2 \in \mathrm{BUC}(\mathbb{R}^d \times [0,T])$ be, respectively, viscosity sub- and supersolution of*

$$u_t(x,t) + H(t, D_x u(x,t)) = 0, \quad (x,t) \in \mathbb{R}^d \times]0,T[.$$

Then,

$$\sup_{\mathbb{R}^d \times [0,T]} (u_1 - u_2) \leq \sup_{\mathbb{R}^d} (u_1(0,\cdot) - u_2(0,\cdot)).$$

Proof. See for example [9], Ch. 2, §3, Theorem 3.7. □

Remark B.1.2. *The comparison Theorem B.1.1 can be extended to the equation*

$$u_t + H(x,t,D_x u) = 0$$

if the Hamiltonian H is uniformly continuous in $\mathbb{R}^d \times [0,T] \times B(0,R)$ for every $R > 0$ and satisfies

$$|H(x,p) - H(y,p)| \leq \omega_1(\|x-y\|(1+\|p\|))$$

for $x, y, p \in \mathbb{R}^d$, where $\omega_1 : [0,+\infty[\longrightarrow [0,+\infty[$ is continuous, nondecreasing with $\omega_1(0) = 0$ and independent of $t \in [0,T]$.

Such results (Proposition B.1.2 and Theorem B.1.1) are needed to prove Theorem 2.1.4 in Ch. 2, §2.1.3.

B.2 Some measure theory and fixed-point results

In this section, we state some useful measure theory results and the fixed-point Schauder-Tychonoff and Kakutani-Ky Fan theorems.

The following result is essential to prove Lemma 1.1.1 in Ch. 1, §1.1.3 (see also Lemma A.1.6 in Appendix A, §A.1).

Theorem B.2.1 (Lusin's Theorem)**.** *Let μ be a Borel regular measure on $\mathbb{R}^d$ and $f : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ be μ-measurable. Assume that $A \subset \mathbb{R}^d$ is μ-measurable and $\mu(A) < \infty$. Fix $\varepsilon > 0$. Then there exists a compact set $K \subset A$ such that*

(i) $\mu(A - K) < \varepsilon$;

(ii) $f\big|_K$ *is continuous.*

Proof. See for example [23], Ch. 1, §1.2, Theorem 2. □

The following theorems B.2.2 and B.2.3 are needed in Ch. 2 to prove Lemma 2.1.1 (§2.1.1), Lemma 2.1.2 (§2.1.4) and Theorem 2.1.1 (§2.1.1), Theorem 2.1.5 (§2.1.4) respectively.

Theorem B.2.2 (Coarea Formula). *Let $f : \mathbb{R}^d \longrightarrow \mathbb{R}^m$ be Lipschitz continuous, $d \geq m$. Then for each $\mathcal{L}^d$-measurable set $A \subset \mathbb{R}^d$, we have*

$$\int_A Jf dx = \int_{\mathbb{R}^m} \mathcal{H}^{d-m}(A \cap f^{-1}\{y\})dy.$$

Proof. See for example [23], Ch. 3, §3.4.2, Theorem 1. □

Theorem B.2.3 (Disintegration Theorem). *Let Y and X be metric spaces, μ a Radon measure on Y, $\pi : Y \longrightarrow X$ a Borel map and let $\nu = \pi\sharp\mu$. Then there exists a ν-a.e. uniquely determined measurable family of Radon measures $\{\mu_x\}_{x\in X}$ such that*

$$\mu_x(Y \setminus \pi^{-1}(x)) = 0 \quad \textit{for } \nu\textit{-a.e. } x \in X$$

and

$$\int_Y f(y)d\mu(y) = \int_X \left(\int_{\pi^{-1}(x)} f(y)d\mu_x(y) \right) d\nu(x)$$

for every $f \in C_c^0(Y)$. The family $\{\mu_x\}_{x\in X}$ is called the disintegration of μ with respect to π (and ν).

Proof. See for example [1], Ch. 2, §2.5, Theorem 2.28 and [26], Ch. 45, §452 for a more general discussion. □

The next result is used in Ch. 3, §3.1.1, Remark 3.1.3.

Theorem B.2.4 (Alexandrov's Theorem). *Let $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ be convex. Then f has a second derivative $\mathcal{L}^d$ almost everywhere.*

Proof. See for example [23], Ch. 6, §6.4, Theorem 1. □

The following fixed-point theorems B.2.5 and B.2.6 are required for proving respectively Theorem 2.1.3 in Ch. 2, §2.1.2 and Theorem 3.3.1 in Ch. 3, §3.3.1.

Theorem B.2.5 (Schauder-Tychonoff Theorem). *Let X be a locally convex space, $K \subset X$ be nonempty and convex (not necessarily closed) and $K_0 \subset K$ be a compact set. Given a continuous map $f : K \longrightarrow K_0$, there exists $\bar{x} \in K_0$ such that $f(\bar{x}) = \bar{x}$.*

Proof. See for example [34], Ch. 10, Theorem 10.1. □

Theorem B.2.6 (Kakutani-Ky Fan Theorem)**.** *Let K be a nonempty, compact and convex subset of a locally convex space X. Let $f : K \longrightarrow \mathcal{P}(K)$ be upper semicontinuous such that $f(x)$ is nonempty, convex and closed for every $x \in K$. Then f has a fixed point $\bar{x} \in K$.*

Proof. See for example [34], Ch. 13, Theorem 13.1. □

Bibliography

[1] L. Ambrosio, N. Fusco, D. Pallara, Functions of Bounded Variation and Free Discontinuity Problems, *Oxford Science Publications*, *Clarendon Press*, *Oxford*, 2000.

[2] F. Bagagiolo, S. Faggian, R. Maggistro, R. Pesenti, *Optimal control of the mean field equilibrium for a pedestrian tourists' flow model*, Networks and Spatial Economics, (2019).

[3] F. Bagagiolo, M. Benetton, *About an optimal visiting problem*, Applied Mathematics & Optimization, **65**(1), 31–51 (2012).

[4] F. Bagagiolo, R. Maggistro, R. Pesenti, *Origin-to-destination network flow with path preferences and velocity controls: a mean field game-like approach*, Journal of Dynamics and Games, **8**(4), 359–380 (2021).

[5] F. Bagagiolo, A. Festa, L. Marzufero, *A hybrid control framework for an optimal visiting problem*, IFAC-PapersOnLine, **54**(5), 241–246 (2021).

[6] F. Bagagiolo, A. Festa, L. Marzufero, *Hybrid control for optimal visiting problems for a single player and for a crowd*, Nonlinear Differential Equations and Applications, **29**(1), 4 (2022).

[7] F. Bagagiolo, A. Festa, L. Marzufero, *The orienteering problem: a hybrid control formulation*, IFAC-PapersOnLine, **54**(5), 175–180 (2021).

[8] F. Bagagiolo, L. Marzufero, *A time-dependent switching mean-field game on networks motivated by optimal visiting problems*, arXiv:2201.00260, (2021).

[9] M. Bardi, I. Capuzzo Dolcetta, Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman equations, *Birkhäuser*, *Boston*, 1997.

[10] R. Bellman, *Dynamic programming treatment of the travelling salesman problem*, Journal of the ACM, **9**(1), 61–63 (1962).

[11] A. Bensoussan, J. L. Menaldi, *Hybrid control and dynamic programming*, Dynamics of Continuous, Discrete and Impulsive Systems, **3**(4), 395–442 (1997).

[12] C. Bertucci, *Fokker-Planck equations of jumping particles and mean field games of impulse control*, Annales de l'Institut Henri Poincaré C, Analyse non linéaire, **37**(5), 1211–1244 (2020).

[13] C. Bertucci, *Optimal stopping in mean field games, an obstacle problem approach*, Journal de Mathématiques Pures et Appliquées, **120**, 165–194 (2018).

[14] P. H. L. Bovy, S. P. Hoogendoorn, *Pedestrian Travel Behavior Modeling*, Network and Spatial Economics, **5**(2), 139–216 (2005).

[15] M. S. Branicky, V. S. Borkar, S. K. Mitter, *A unified framework for hybrid control: Model and optimal control theory*, IEEE Transactions on Automatic Control, **43**(1), 31–45 (1998).

[16] F. Camilli, E. Carlini, C. Marchi, *A model problem for Mean Field Games on networks*, Discrete and Continuous Dynamical Systems, **35**(9), 4173–4192 (2015).

[17] F. Camilli, R. De Maio, A. Tosin, *Transport of measures on networks*, Networks and Heterogeneous Media, **12**(2), 191–215 (2017).

[18] P. Cardaliaguet, Notes on Mean Field Games, *Lecture notes* (from P.-L. Lions' lectures at Collége de France), 2013.

[19] P. Cardaliaguet, P. J. Graber, *Mean field games systems of first order*, ESAIM: Control, Optimisation and Calculus of Variations, **21**(3), 690–722 (2015).

[20] P. Centorrino, A. Corbetta, E. Cristiani, E. Onofri, *Managing crowded museums: Visitors flow measurement, analysis, modeling, and optimization*, Journal of Computational Science, **53** (2021).

[21] M. Chowdhurry, S. Yang, S. Khuller, J. You, M. Purohit, *Near Optimal Coflow Scheduling in Networks*, SPAA '19: The 31st ACM Symposium on Parallelism in Algorithms and Architectures, 123–134 (2019).

[22] S. Dharmatti, M. Ramaswamy, *Hybrid control systems and viscosity solutions*, SIAM Journal on Control and Optimization, **44**(4), 1259–1288 (2005).

[23] L. C. Evans, R. F. Gariepy, Measure theory and fine properties of functions, *Revised edition*, *CRC Press*, *USA*, 2015.

[24] A. Festa, S. Göttlich, *A mean field game approach for multi-lane traffic management*, IFAC-PapersOnLine, **51**(32), 793–798 (2018).

[25] A. Festa, R. Guglielmi, R. Hermosilla, A. Picarelli, S. Sahu, A. Sassi, F. J. Silva, *Hamilton-Jacobi-Bellman equations*, Lecture Notes in Mathematics, **2180**, 127–261 (2017).

[26] D. H. Fremlin, Measure Theory, volume 4, *Torres Fremlin*, 2003.

[27] B. Gavish, S.-C. Graves, The travelling salesman problem and related problems, *Massachusetts Institute of Technology*, *Operations Research Center*, 1978.

[28] T. Gonzalez, S. Sahni, *Open Shop Scheduling to Minimize Finish Time*, Journal of the ACM, **23**(4), 665–679 (1976).

[29] D. A. Gomes, J. Saúde, *Mean Field Models–A Brief Survey*, Dynamic Games and Applications, **4**(2), 110–154 (2014).

[30] M. Huang, *Large-population LQG games involving a major player: the Nash certainty equivalence principle*, SIAM Journal on Control and Optimization, **48**(5), 3318–3353 (2009).

[31] J. M. Lasry, P. L. Lions, *Mean field games*, Japanese Journal of Mathematics, **2**(1), 229–260 (2007).

[32] X. Li, S. Song, Impulsive Systems with Delays: Stability and Control, *Springer*, *Singapore*, 2022.

[33] M. Morandotti, F. Solombrino, *Mean-field Analysis of Multipopulation Dynamics with Label Switching*, SIAM Journal of Mathematical Analysis, **52**(2), 1427–1462 (2020).

[34] V. Pata, Fixed Point Theorems and Applications, *Springer*, *Cham*, 2019.

[35] B. Piccoli, F. Rossi, *Generalized Wasserstein distance and its application to transport equations with source*, Archive for Rational Mechanics and Analysis, **211**(1), 335–358 (2014).

[36] B. Piccoli, F. Rossi, *On properties of the generalized Wasserstein distance*, Archive for Rational Mechanics and Analysis, **222**(3), 1339–1365 (2016).

[37] M. L. Pinedo, Scheduling: Theory, Algorithms, and Systems, 5rd Edition, *Springer*, *Boston*, 2012.

[38] A. Visintin, Mathematical models of hysteresis, volume 1, *Springer New York*, *New York*, 2006.