



UNIVERSITÀ DEGLI STUDI DI MILANO
FACOLTÀ DI SCIENZE E TECNOLOGIE

Dipartimento di Matematica “Federigo Enriques”

Dottorato di Ricerca in Scienze Matematiche

Ciclo XXXVIII

Tesi di Dottorato di Ricerca

**Advances in mean-field optimal control
via dynamic programming equations**

Settore scientifico-disciplinare: MATH-03/B

Tutor:
Prof. Marco Fuhrman

Co-Tutor:
Prof. Andrea Cosso

Coordinatore del corso di dottorato:
Prof. Giulio Ciraolo

Candidata:
Silvia Rudà

Matricola: R13875
ORCID ID: 0009-0006-7222-7433

A. A. 2024-2025

Thesis Summary

This Thesis is devoted to the study of several classes of optimal control problems of the McKean-Vlasov type via an analytic approach, which consists in characterizing the value of each problem by means of a suitable partial differential equation, called Hamilton-Jacobi-Bellman (HJB) equation. As the coefficients of a McKean-Vlasov stochastic differential equation depend on the state process also through its law, the associated HJB equation is defined on the Wasserstein space of probability measures with finite second moment $\mathcal{P}_2(\mathbb{R}^d)$. It is our purpose to discuss the well-posedness of this PDE in a suitable viscosity sense directly on the Wasserstein space, without lifting it to a Hilbert space of square-integrable random variables L^2 , in contrast with previous contributions.

The first chapter of this Thesis focuses on the formulation and the analysis of a general optimal control problem of the McKean-Vlasov type with infinite time horizon and discounted gain functional. Under standard assumptions on the coefficients of the state equation, the gain function and the discount rate β , the optimization problem is well-posed. We firstly prove that the value function does not depend on the initial time of the dynamics. This property, which is often considered natural for discounted control problems over an unbounded time interval, is not trivial when mean-field effects occur: we provide an example where the property fails, due to the non-linear dependence of the gain function on the law of the state process. Nonetheless, assuming that the gain function is sufficiently smooth, we extend this time invariance property to our McKean-Vlasov type problem. We then show that the value function depends on the initial condition of the dynamics only through its law. The value of the problem can thus be identified with a time-independent function v defined on $\mathcal{P}_2(\mathbb{R}^d)$. Hence, a suitable elliptic Hamilton-Jacobi-Bellman equation is derived on the same space; the value v is completely characterized as its unique solution in a suitable viscosity sense.

The second Chapter presents a different optimal control problem with the same McKean-Vlasov dynamics, infinite time horizon but gain functional in ergodic (long time average) form. Under suitable dissipativity assumptions, we derive strong stability properties of the solutions to the state equation, called ergodic estimates. We use the vanishing discount method, relying on auxiliary control problems with discounted gain functional, to construct a real value λ and a function $\phi : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$, which provide a viscosity solution to an ergodic HJB equation on the space $\mathcal{P}_2(\mathbb{R}^d)$. Moreover, the pair (λ, ϕ) is proved to characterize the asymptotic behaviour of the value functions of finite horizontal mean-field control problems when the time horizon

tends to $+\infty$ and of infinite horizontal mean-field control problems with discounted gain when the discount rate tends to 0. We eventually identify λ with the value function of the ergodic control problem, thus showing that it is constant. We also present a verification theorem to construct optimal controls for the ergodic optimal control problem by means of a solution to the ergodic HJB equation in closed-loop form. In contrast with previous contributions, we do not derive the ergodic HJB equation using stability properties of solutions to elliptic PDEs, but our main results follow from a fundamental functional relation involving ϕ and λ .

The third chapter presents some preliminary results concerning an optimal control problem of the McKean-Vlasov type with partial observation: the mean-field dynamics is influenced by a continuous-time Markov chain M with finite state space which can not be observed by the representative agent. The analysis of the problem thus requires filtering techniques, relying on a fully observable auxiliary control problem, called separated problem, with a measure-valued controlled state equation (Zakai equation); this SDE can be rewritten as an $\mathbb{R}^N$ -valued state equation as M can only assume N values. It is then proved that the separated control problem is equivalent to the original control problem with partial observation in a weak form. Adopting this approach, due to the presence of mean-field effects in the original state equation, we obtain a non-standard optimal control problem of the McKean-Vlasov type, whose coefficients depend on the joint law of the state equation and the Brownian noise. We thus study its well-posedness and stability properties. We also discuss future research with the aim of adopting the analytic approach, eventually leading to a suitable HJB equation on a Wasserstein space.

The research presented in this Thesis resulted in the following works. The first preprint is currently available on arXiv and submitted for publication, while the second paper has recently been published on SIAM Journal on Control and Optimization:

- Silvia Rudà. Infinite time horizon optimal control of McKean-Vlasov SDEs. *arXiv preprint arXiv:2503.20572*, 2025.
- Marco Fuhrman e Silvia Rudà. Ergodic control of McKean-Vlasov systems on the Wasserstein space. *SIAM Journal on Control and Optimization*, 63(6):4018–4043, 2025.

The model analysed in the third chapter of the thesis, and presented as work in progress when the thesis has been first submitted, has subsequently been extended and generalized in the following preprint, currently available on arXiv and submitted for publication:

M. Fuhrman, H. Pham and S. Rudà. Optimal control of McKean-Vlasov systems under partial observation and hidden Markov switching. *arXiv preprint arXiv:2601.09311*, 2026.

Riassunto della Tesi

In questa tesi di dottorato ci proponiamo di analizzare varie classi di problemi di controllo ottimo di tipo McKean-Vlasov con un approccio analitico, che consiste nella caratterizzazione della funzione valore di ciascun problema mediante un'opportuna equazione alle derivate parziali, detta equazione di Hamilton-Jacobi-Bellman (HJB). Poiché i coefficienti di un'equazione stocastica di McKean-Vlasov dipendono non solo dalle traiettorie della soluzione, ma anche dalla sua legge, l'equazione di HJB associata è definita sullo spazio di Wasserstein delle misure di probabilità di momento secondo finito $\mathcal{P}_2(\mathbb{R}^d)$. In questa tesi, la buona positura di tale equazione in senso viscoso è sempre analizzata direttamente sullo spazio di Wasserstein; non ricorriamo pertanto al lifting dell'equazione su uno spazio di Hilbert di variabili aleatorie di quadrato sommabili, in contrasto con una parte della letteratura su questo argomento.

Nel primo capitolo della tesi, formuliamo e analizziamo un problema di controllo ottimo di tipo McKean-Vlasov con orizzonte temporale infinito e funzionale di guadagno scontato. Il problema è ben posto sotto ipotesi standard per i coefficienti dell'equazione di stato, la funzione di guadagno e il fattore di sconto β . Innanzitutto, dimostriamo che la funzione valore non dipende dal tempo iniziale della dinamica. Questa proprietà, spesso considerata naturale per problemi con orizzonte temporale infinito e fattore di sconto, non è invece banale in presenza di effetti di campo medio: in particolare, esibiamo un esempio in cui la proprietà non vale a causa della dipendenza non lineare della funzione di guadagno dalla legge del processo di stato. Tuttavia, assumendo che la funzione di guadagno sia sufficientemente regolare, estendiamo questa proprietà di invarianza temporale al nostro problema di tipo McKean-Vlasov. Successivamente, mostriamo che la funzione valore dipende dalla condizione iniziale della dinamica solo mediante la sua legge. Il valore del problema può dunque essere identificato con una funzione v indipendente dal tempo e definita su $\mathcal{P}_2(\mathbb{R}^d)$. Infine, sullo stesso spazio deriviamo un'opportuna equazione ellittica di Hamilton-Jacobi-Bellman e caratterizziamo completamente la funzione valore come l'unica soluzione viscosa di tale equazione.

Il secondo capitolo presenta un problema di controllo ottimo con la stessa dinamica di tipo McKean-Vlasov su orizzonte temporale infinito, ma un funzionale di guadagno in forma ergodica. Sotto opportune ipotesi di dissipatività, le soluzioni dell'equazione di stato soddisfano condizioni forti di stabilità rispetto alle condizioni iniziali, dette stime ergodiche. Basandoci su una famiglia di problemi di controllo ausiliari con funzionale di guadagno scontato, costruiamo un valore reale λ e una funzione $\phi : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ che costituiscono una soluzione viscosa di un'equazione ellittica, detta equazione ergodica di HJB, sullo spazio $\mathcal{P}_2(\mathbb{R}^d)$. Inoltre, dimostriamo che la coppia (λ, ϕ) caratterizza il

comportamento asintotico delle funzioni valore associate a problemi di controllo ottimo di tipo McKean-Vlasov ad orizzonte temporale finito quando quest'ultimo tende a $+\infty$ e delle funzioni valore associate a problemi di controllo ottimo di tipo McKean-Vlasov ad orizzonte temporale infinito con fattore di sconto quando esso tende a 0. Infine, identifichiamo λ con il valore del problema di controllo ergodico, dimostrando che esso è costante. Presentiamo poi un teorema di verifica che permette di costruire controlli ottimi per il problema ergodico a partire da soluzioni dell'equazione di HJB ergodica. A differenza di quanto proposto da altri autori, non deriviamo l'equazione ergodica di HJB usando proprietà di stabilità di soluzioni viscosse di PDE ellittiche, ma tutti i nostri risultati principali seguono da una relazione funzionale soddisfatta dalla coppia (λ, ϕ) .

Il terzo capitolo presenta alcuni risultati preliminari riguardanti un problema di controllo ottimo di tipo McKean-Vlasov con osservazione parziale: la dinamica è infatti influenzata da una catena di Markov M a tempo continuo e spazio degli stati finito che non è osservata dall'agente. L'analisi del problema richiede dunque tecniche di filtraggio, basate su un problema di controllo ausiliario, detto problema separato, con un'equazione di stato controllata a valori nello spazio delle misure positive (l'equazione di Zakai). Tale equazione si può poi riscrivere come una SDE a valori in $\mathbb{R}^N$, dal momento che la catena può assumere solo un numero finito N di valori. Adottando questo approccio, a causa della presenza di effetti di campo medio nel problema originale, otteniamo un problema di controllo di tipo McKean-Vlasov non standard, i cui coefficienti dipendono dalla legge congiunta del processo di stato e del moto Browniano. Studiamo dunque la buona positura di tale problema e la stabilità delle soluzioni dell'equazione di stato rispetto alle condizioni iniziali e dimostriamo che esso è equivalente ad una versione debole del problema originale. Presentiamo infine gli sviluppi futuri di questa ricerca, ricordando che intendiamo adottare un approccio di tipo analitico al problema, che porti alla caratterizzazione del valore mediante un'opportuna equazione di HJB su uno spazio di Wasserstein.

La ricerca presentata in questa tesi è risultata nei seguenti articoli; il primo lavoro è un preprint, attualmente disponibile su arXiv e sottomesso per la pubblicazione, mentre il secondo è recentemente stato pubblicato su SIAM Journal on Control and Optimization:

- Silvia Rudà. Infinite time horizon optimal control of McKean-Vlasov SDEs. *arXiv preprint arXiv:2503.20572*, 2025.
- Marco Fuhrman e Silvia Rudà. Ergodic control of McKean-Vlasov systems on the Wasserstein space. *SIAM Journal on Control and Optimization*, 63(6):4018–4043, 2025.

Il modello analizzato nel terzo capitolo, presentato come work in progress all'atto della sottomissione della tesi, è poi stato ampliato e generalizzato nel seguente preprint, attualmente disponibile su arXiv e sottomesso per la pubblicazione:

M. Fuhrman, H. Pham and S. Rudà. Optimal control of McKean-Vlasov systems under partial observation and hidden Markov switching. *arXiv preprint arXiv:2601.09311*, 2026.

Acknowledgements

I would like to express my deepest gratitude to my Thesis supervisor, Prof. Marco Fuhrman, for his constant patience, encouragement and support, as well as for our precious discussions which led to the research in this Thesis.

I would also like to thank my Co-Tutor, Prof. Andrea Cosso, for his interest in my work and for our useful discussions about past contributions and future developments. It is a pleasure to thank Prof. Huyền Pham for our fruitful collaboration.

I thank the two anonymous Referees for their positive feedbacks on the thesis and for their valuable suggestions and remarks.

I also wish to express my gratitude to all the students and researchers I met during all the scientific activities related to the PhD for the opportunity of engaging in interesting debates. Finally, I would like to thank all the people who supported me throughout this academic journey.

Contents

Thesis Summary	i
Riassunto della Tesi	iii
Introduction	1
1 Optimal control of McKean-Vlasov SDEs with discounted gain	11
1.1 Introduction	11
1.2 McKean-Vlasov SDEs and optimal control	16
1.2.1 Notation and assumptions	16
1.2.2 Probabilistic setting	16
1.2.3 Dynamics of the controlled system	18
1.2.4 Gain functional and optimization problem	20
1.3 The time invariance property	25
1.3.1 Preliminary considerations	25
1.3.2 A counterexample to the time invariance property	27
1.3.3 Time invariance of the value function	34
Disintegration of McKean-Vlasov systems	34
Proof of the time invariance property	39
1.4 Continuity of the value function and law invariance property	43
1.5 Hamilton-Jacobi-Bellman equation	45
1.5.1 Lions derivative and viscosity solutions to PDEs	45
1.5.2 Hamilton-Jacobi-Bellman equation and value function	50
1.5.3 Uniqueness of the solution	50
2 Ergodic control of McKean-Vlasov systems on the Wasserstein space	57
2.1 Introduction	57
2.2 Notation and preliminaries	62
2.3 Ergodicity	65
2.4 Ergodic problems on the Wasserstein space	67
2.4.1 The pair (λ, ϕ) and its first properties	68
2.4.2 Applications to limit behaviour of some value functions	72
2.5 Ergodic and parabolic HJB equations on the Wasserstein space	74
2.6 Ergodic control problem	79
3 Optimal control of McKean-Vlasov SDEs with partial observation	85

3.1	Introduction	85
3.2	Motivation	89
3.3	The separated control problem	92
3.3.1	Assumptions and formulation	92
3.3.2	Equivalence with the partially observed control problem	95
3.4	Dynamic programming approach	98
3.4.1	A larger class of control problems	98
3.4.2	Further properties of the separated problem and the value function	102
3.5	Future developments	104
3.6	Proofs and examples	104
3.6.1	Proof of Theorem 3.4.1	104
3.6.2	Proof of Theorem 3.4.3	108
3.6.3	Examples	111
A	Estimates and properties of solutions to McKean-Vlasov SDEs	113
A.1	Continuity of the law of the state process	113
A.2	Proof of estimate (1.4)	115
A.3	Picard iterates for McKean-Vlasov SDEs	118
B	Time invariance property: technical results	127
B.1	Translation invariance property of solutions to SDEs	127
B.2	Well-posedness of the disintegration formula	142
B.3	Proof of the continuity of the gain functional	144
C	Ergodic HJB equations: proofs	151
C.1	Proof of Theorem 2.5.5	151
C.2	Proof of Proposition 2.5.6	153

Introduction

This thesis is devoted to the study of several classes of mean-field optimal control problems via an analytic approach, which consists in characterizing the value function of each control problem by means of a suitable partial differential equation (PDE) on a Wasserstein space of probability measures.

This section introduces McKean-Vlasov stochastic differential equations and some associated control problems. We start with a brief historical introduction to these models. We then discuss finite horizon optimal control problems of the mean-field type, which have been widely studied in the literature, in order to present the main features and the main difficulties related to these models. We especially focus on the analytical approach and the associated Hamilton-Jacobi-Bellman equation, considering possible ways to derive it and to solve it. Finally, we present the content of this Thesis, giving a detailed overview of the main results of each chapter. We will not describe the precise outline of the Thesis, as the organization of each Chapter is provided in its introduction. However, precise references to the main results are provided.

This dissertation focuses on dynamical systems whose dynamics is described by McKean-Vlasov stochastic differential equations (SDEs) of the form

$$dX_s = b(X_s, \mathbb{P}_{X_s}, \alpha_s)ds + \sigma(X_s, \mathbb{P}_{X_s}, \alpha_s)dB_s, \quad (\text{I.1})$$

where B is a m -dimensional Brownian motion on an underlying filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ and X takes values in a finite dimensional space $\mathbb{R}^d$. The stochastic process α belongs to a suitable space of control processes and models the influence of the decisions taken by a representative agent on the evolution of the system. The main feature of these equations is the dependence of the coefficients on the solution X also through the probability distribution of its time-marginals, denoted by $\mathbb{P}_{X_s}$. Distribution-dependent differential equations were first introduced in statistical mechanics to describe the limit behaviour of interacting particle systems. In particular, starting from a model where the interaction between N independent and identically distributed particles satisfies suitable properties and letting the number of such particles tend to $+\infty$, we obtain infinitely many independent and identically distributed processes, each of which representing the dynamics of a single “tagged” particle; the coefficients of the limit equation of a representative particle depend on the law of the solution. This kind of limit result was first called *propagation of chaos* in [61], where it was used to derive Boltzmann equations for hard spheres; this procedure

was subsequently extended in [72] to other transport problems emerging in statistical mechanics. We refer the reader to [88] for an introduction to this subject.

In the context of controlled system, McKean-Vlasov stochastic differential equations first appeared in connection with optimization problems of large-scale populations, where the agents are assumed to be symmetric and interacting only through the empirical measure of the population, representing its average state distribution. As in the uncontrolled case, the limit dynamics of these systems can be represented by the state equation of a single representative agent, whose coefficients depend on the probability distribution of the state at present time. In the case of competitive agents, this leads to the theory of mean-field games, initiated independently in [64] and [58]. On the other hand, if the agents are cooperative (or subject to a central planner aiming at maximizing the collective reward), the limit model will be a mean-field control problem, as rigorously proved in [63]. We refer the reader to [32] and [33] for comprehensive monographs on this subject and to [34] for an overview of the differences between mean-field games and mean-field control problems. Notice that the initial condition of (I.1) is a square integrable random variable on the same filtered probability space underlying the dynamics: as proved in [26], if we restrict to the set of deterministic initial conditions $x \in \mathbb{R}^d$, the solution to equation (I.1) loses some fundamental properties of solutions to SDEs, such as the flow property, which may be restored on a suitable L^2 space, instead.

The choice of the control α in (I.1) is based on an optimization criterion: the representative agent aims at maximizing (or minimizing) a gain (or cost) functional. In the finite horizon case, where the dynamics (I.1) evolves on a bounded time interval $[0, T]$, the gain functional usually takes the form

$$J(t, \xi, \alpha) := \mathbb{E} \left[\int_t^T f(X_s^{t, \xi, \alpha}, \mathbb{P}_{X_s^{t, \xi, \alpha}}, \alpha_s) ds + g(X_T^{t, \xi, \alpha}, \mathbb{P}_{X_T^{t, \xi, \alpha}}) \right],$$

where $X^{t, \xi, \alpha} = (X_s^{t, \xi, \alpha})_{s \in [t, T]}$ denotes the solution to (I.1) starting from the random variable ξ at time t . The functions f and g , representing the current and the terminal gain, respectively, may also depend on the law of the state process. The supremum of the attainable gains is encoded in the so-called value function

$$V(t, \xi) := \sup_{\alpha \in \mathcal{A}} J(t, \xi, \alpha), \tag{I.2}$$

where $\mathcal{A}$ is a suitable class of admissible controls. The main difficulties for this class of optimal control problems are related to the non-linear dependence of the coefficients of the state equation and of the gain functions on the law of the state process, which may lead to unexpected phenomena which can never emerge in absence of mean-field effects.

In the finite horizon framework, this problem has been widely analysed through different methodologies. A probabilistic approach, relying on the Stochastic Maximum Principle, has been adopted to provide necessary (and, in some cases, also sufficient) conditions for optimality; see e.g. in [66], [3], [24], [32] and references therein. On the other hand, the analytic approach, based on the characterization of the value function

by means of a suitable partial differential equation (Hamilton-Jacobi-Bellman, HJB, equation), has attracted the interest of many authors. Under suitable conditions, depending on the specific formulation of the problem, it is possible to prove that the value function of the problem is law invariant, i.e. that it depends on the initial condition only through its law: if two different initial conditions ξ and ξ' have the same law, then $V(t, \xi) = V(t, \xi')$ for all $t \in [0, T]$. The value of the problem can thus be identified with the function

$$v(t, \mu) := V(t, \xi) \quad (\text{I.3})$$

for all $t \geq 0$ and ξ square integrable random variable having law μ and satisfying suitable measurability conditions. v is defined on a space of probability measures, the Wasserstein space of order 2, denoted by $\mathcal{P}_2(\mathbb{R}^d)$. Correspondingly, the associated Hamilton-Jacobi-Bellman equation is a parabolic backward partial differential equation on $\mathcal{P}_2(\mathbb{R}^d)$. It is thus necessary to introduce a notion of derivative with respect to a probability measure: the majority of the contributions adopt the notion of derivative in the lifted sense of Lions, introduced in [70] and detailed in [30]. Accordingly, a notion of viscosity solution to a PDE on this space has to be provided, as it is well known that the value function of an optimal control problem is often not smooth enough to be a classical solution to a PDE. Notice that one of the main difficulties of this approach is related to the lack of a general theory of viscosity solutions to PDEs on Wasserstein spaces. As a consequence, in some papers following this research line (see e.g. [78] and [79]), the HJB equation is in fact lifted to a suitable L^2 space, where it is possible to exploit the theory of viscosity solutions on Hilbert spaces (presented in [67], [68], [69] and subsequently in [51]). Unfortunately, the relation between viscosity solutions to PDEs on Wasserstein spaces and to their lifted version is not yet completely clear (see Remark 3.6 in [43]).

More recently, the authors of [42] derived the following parabolic PDE directly on the Wasserstein space:

$$\begin{cases} -\partial_t v^T(t, \mu) - \int_{\mathbb{R}^d} \left[\sup_{a \in A} \{f(x, \mu, a) + \langle b(x, \mu, a), \partial_\mu \varphi(\mu)(x) \rangle \right. \\ \left. + \frac{1}{2} \text{tr}(\sigma \sigma^T(x, \mu, a) \partial_x \partial_\mu \varphi(\mu)(x)) \right] \mu(dx) = 0, & t \in [0, T], \mu \in \mathcal{P}_2(\mathbb{R}^d), \\ v^T(T, \mu) = \int_{\mathbb{R}^d} g(x, \mu) \mu(dx). \end{cases} \quad (\text{I.4})$$

Moreover, they proved that the value function v is a solution to this equation in a suitable viscosity sense, defined as a generalization of the notion of viscosity solution to a PDE in the classical Crandall-Lions sense ([45]). Uniqueness of viscosity solutions to this equations (and to other second order parabolic PDEs associated to different classes of optimal control problems of the McKean-Vlasov type) is a difficult problem and it is presently the object of thorough investigation: see e.g. [43], [36], [14], [86], [16]. In all the aforementioned references, the coefficients are required to satisfy stronger assumptions and the definition of viscosity solution sometimes needs to be modified.

This thesis aims at extending the existing theory of mean-field control; in particular, it focuses on problems where the system evolves with infinite time horizon, both with gain functional in discounted form (in Chapter 1) and in ergodic form (in Chapter

2), and on a model where the dynamics depends on a source of randomness which can not be observed by the agent controlling the system (in Chapter 3). In the whole dissertation we adopt the analytic approach: we associate to each problem a Hamilton-Jacobi-Bellman equation directly on the Wasserstein space and we discuss its relation with the value of the problem.

In the first chapter, we present a theory of optimal control of McKean-Vlasov SDEs with infinite time horizon and gain in discounted form. We will not focus on a specific model, but on a general dynamical system of the form

$$\begin{cases} dX_s = b(X_s, \mathbb{P}_{X_s}, \alpha_s)ds + \sigma(X_s, \mathbb{P}_{X_s}, \alpha_s)dB_s, & s \geq t \geq 0 \\ X_t = \xi, \end{cases} \quad (\text{I.5})$$

where the coefficients b and σ satisfy suitable Lipschitz continuity and growth conditions guaranteeing that the state equation admits a unique solution denoted by $X^{t,\xi,\alpha}$, ξ is a square summable random variable and α belongs to a set of processes being predictable with respect to an initial enlargement of the completed Brownian filtration. The gain functional can also involve a dependence on the law of the state process: it is defined as

$$J(t, \xi, \alpha) = \mathbb{E} \left[\int_t^{+\infty} e^{-\beta(s-t)} f(X_s^{t,\xi,\alpha}, \mathbb{P}_{X_s^{t,\xi,\alpha}}, \alpha_s) ds \right], \quad (\text{I.6})$$

where the gain function f is subject to growth conditions and the discount rate β is sufficiently large, in order that J is actually well defined. Apparently, the value function

$$V(t, \xi) := \sup_{\alpha \in \mathcal{A}} J(t, \xi, \alpha) \quad (\text{I.7})$$

thus depends on the initial time and condition (t, ξ) . The main novelty of this chapter is the proof of the so-called time invariance property, stating that, for all the initial conditions ξ being admissible from time $t = 0$,

$$V(t, \xi) = V(0, \xi). \quad (\text{I.8})$$

This result is in fact proved to be implied by (or, more precisely, equivalent to) an invariance property of the gain with respect to the class of controls. As a matter of fact, assume for the moment that, starting from a positive time $t > 0$, it is possible to restrict the class of admissible controls, “forgetting” the past of the Brownian motion before t , without modifying the value. Then the time invariance property can be derived following the same reasoning that one can adopt in classical control theory, without mean field effects, using appropriate shifting techniques detailed in Appendix B.1. Therefore, the key property is the equality between the values defined maximizing the reward over different classes of controls, i.e.

$$\sup_{\alpha \in \mathcal{A}} V(t, \xi) = \sup_{\alpha \in \mathcal{A}_t} V(t, \xi), \quad \text{for all } t \geq 0 \text{ and } \xi \quad (\text{I.9})$$

where $\mathcal{A}_t \subseteq \mathcal{A}$ is a smaller class of controls being predictable with respect to an initial enlargement of the filtration generated only by the increments of the Brownian motion

from time t . This equality, which is natural in classical contexts, is not trivial when the coefficients of the problem exhibit a non-linear dependence on the law of the state process: in particular, we present an example (Example 1.3.3) where it fails. On the other hand, we also provide a second example (Example 1.3.6) suggesting that it is possible to restore this essential property under mild continuity assumptions on the gain function f . We thus prove equality (I.9) rigorously for our general model in Theorem 1.3.8; the proof relies on the idea that, under suitable continuity conditions of the reward functional, randomizing the controls by means of the additional random source given by the past of the Brownian trajectory does not improve the value. This result is based on a disintegration formula of solutions to McKean-Vlasov SDEs with respect to the first part of the Brownian trajectory, presented in Proposition 1.3.11: this is a tool that we introduce to apply conditional independence, as we can not directly condition the law of the state process, which is a deterministic object, with respect to the σ -algebra containing the information up to the initial time. Once we have solved this basic question, we can easily extend the aforementioned law invariance property to this infinite horizon setting (Theorem 1.4.3) and then identify the value of the problem with a function $v : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ such that

$$v(\mu) = V(t, \xi)$$

for any $t \geq 0$ and any admissible initial condition ξ having law μ . We prove that this function satisfies a Dynamic Programming Principle (Corollary 1.4.5), from which we derive the following elliptic Hamilton-Jacobi-Bellman equation directly on the Wasserstein space of order 2:

$$\begin{aligned} \beta\varphi(\mu) = \int_{\mathbb{R}^d} \left[\sup_{a \in A} \left\{ f(x, \mu, a) + \langle b(x, \mu, a), \partial_\mu \varphi(\mu)(x) \rangle \right. \right. \\ \left. \left. + \frac{1}{2} \text{tr} (\sigma \sigma^T(x, \mu, a) \partial_x \partial_\mu \varphi(\mu)(x)) \right\} \right] \mu(dx), \quad \mu \in \mathcal{P}_2(\mathbb{R}^d), \end{aligned} \quad (\text{I.10})$$

where we adopt the notion of derivative in the lifted sense of Lions. We define viscosity solutions to this equation in a suitable Crandall-Lions sense, inspired by [36], and we prove that the value function v actually is a viscosity solution to (I.10). We eventually focus on the problem of uniqueness of viscosity solutions to (I.10). In particular, as we could not find a reference for comparison results for elliptic PDEs on the Wasserstein space, we derive our Theorem 1.5.15 from the uniqueness results available for parabolic PDEs of the Hamilton-Jacobi-Bellman type on the Wasserstein space. The proof is based on an approximation argument: starting from any sub (or super) solution u to our equation (I.10), we construct a sequence of optimal control problems with finite time horizon T and terminal gain $g(\mu) = e^{-\beta T} u(\mu)$, defined in (1.67). We prove that the associated value functions v^T satisfy $\lim_{T \rightarrow +\infty} v^T(0, \mu) = v(\mu)$. We subsequently introduce for each time T a sub (or super) solution U^T to the parabolic backward PDE (I.4) such that $U^T(0, \mu) = u(\mu)$; by the comparison theorems available for this equation, we deduce that $U^T \leq v^T$ ($U^T \geq v^T$, respectively) on $[0, T] \times \mathcal{P}_2(\mathbb{R}^d)$. Passing to the limit for $T \rightarrow +\infty$, we conclude that a comparison theorem holds also for our elliptic HJB equation (I.10). Notice that this strategy only relies on the existence of a comparison result for parabolic PDEs on the Wasserstein space and not on the

specific choice of the result; however, we give a possible set of stronger hypotheses under which our uniqueness result holds. The Chapter presents an extended version of the contents of [84], submitted to a journal and currently under review.

Chapter 2 is devoted to a different optimal control problem of the McKean-Vlasov type with infinite time horizon, where the gain functional is in ergodic (long term average) form. The dynamics is still described by equation (I.5), whose solution from time $t = 0$ is denoted by $X^{\xi, \alpha}$, while the gain functional has the form

$$J_{\infty}(\xi, \alpha) = \limsup_{T \rightarrow \infty} \mathbb{E} \left[\frac{1}{T} \int_0^T f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right], \quad (\text{I.11})$$

where f is assumed to be bounded. In classical cases, without mean-field effects, there are two possibilities: if sufficiently strong conditions, such as dissipativity or periodicity conditions, are satisfied by the coefficients of the state equation, then the value function associated with a long term average gain functional is in general constant, i.e. it does not depend on the initial condition (see e.g. [6]). On the other hand, under weaker assumptions (such as the non-expansivity condition, see [25]) the value will not be identically equal to a constant and the system will be called “non-ergodic”. In our mean-field case, we aim at ensuring that the value is in fact constant. In order to do so, we introduce an ad hoc dissipativity assumption (Assumption **(H₂)** in Chapter 2), already considered in the literature (see [11]), guaranteeing that the solutions to the state equation satisfy the following ergodic estimates: for any two solutions $X^{\xi, \alpha}$, $X^{\xi', \alpha}$ to the state equation, we have

$$\begin{aligned} \mathbb{E} \left[|X_t^{\xi, \alpha} - X_t^{\xi', \alpha}|^2 \right] &\leq \mathbb{E} \left[|\xi - \xi'|^2 \right] e^{-2\eta t}, \\ \mathbb{E} \left[|X_t^{\xi, \alpha}|^2 \right] &\leq \mathbb{E} \left[|\xi|^2 \right] e^{-\eta t} + K, \end{aligned}$$

for suitable η and K . These estimates are stronger than the usual stability properties of solutions to SDEs and will play a crucial role in the chapter. Notice that, since it may be difficult to verify Assumption **(H₂)** in concrete examples, we also provide an alternative condition and prove that it implies Assumption **(H₂)** (see Proposition 2.3.2). We subsequently focus on the construction of the value of the problem, adopting the so-called vanishing discount method: under the same dynamics, we construct a family of optimal control problems with discounted gain functional defined as in (I.6), indexed by $\beta > 0$. Applying the results of Chapter 1, we deduce that the associated value functions, denoted here by v^{β} , are defined on the space $\mathcal{P}_2(\mathbb{R}^d)$. Moreover, by the ergodic estimates we can deduce that, for each fixed $\beta > 0$, the function v^{β} is Lipschitz with Lipschitz constant L independent of β (Proposition 2.31). Therefore, letting β tend to 0 along a subsequence, by Ascoli-Arzelà theorem we construct a real value λ as the limit of $\beta v^{\beta}(\mu)$ for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and a function $\phi : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ as the limit (also along a subsequence) of $v^{\beta} - v^{\beta}(\delta_0)$. The second convergence is uniform on any compact subset of $\mathcal{P}_2(\mathbb{R}^d)$. The key result in our dissertation is the following functional relation, involving the function ϕ and the number λ : for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $T > 0$, for every $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ such that $\mathbb{P}_{\xi} = \mu$, we have

$$\phi(\mu) + \lambda T = \sup_{\alpha \in \mathcal{A}} \left\{ \mathbb{E} \left[\int_0^T f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right] + \phi(\mathbb{P}_{X_T^{\xi, \alpha}}) \right\}, \quad (\text{I.12})$$

see Proposition 2.4.2. This equality identifies $\phi + \lambda T$ with the value associated with an optimal control problem with dynamics (I.5), finite time horizon T and terminal gain $g = \phi$. All the main results of the chapter will be derived from this relation (or from an extension for a positive initial time $t \geq 0$, see Proposition 2.4.4). In particular, this identification provides information about the limit behaviour of value functions associated with different optimal control problems under the same McKean-Vlasov dynamics (I.5). First of all, considering the value function of the finite horizon control problem defined in (I.3) for g satisfying suitable growth conditions, we study its limit behaviour when we let the time horizon tend to $+\infty$ in Theorem 2.4.5; in particular, we identify λ with $\lim_{T \rightarrow \infty} \frac{v^T(0, \mu)}{T}$. We subsequently apply this result to prove Theorem 2.4.6, stating that, while λ was originally defined as $\lim_{\beta \rightarrow 0} \beta v^\beta$ along a subsequence $\{\beta_n\}_{n \in \mathbb{N}}$, it coincides with the limit on the entire sequence. Therefore, we can identify the following limits:

$$\begin{aligned} \lambda &= \lim_{\beta \rightarrow 0} \sup_{\alpha \in \mathcal{A}} \mathbb{E} \left[\beta \int_0^\infty e^{-\beta t} f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right] \\ &= \lim_{T \rightarrow \infty} \sup_{\alpha \in \mathcal{A}} \mathbb{E} \left[\frac{1}{T} \int_0^T f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right], \end{aligned} \quad (\text{I.13})$$

leading to a so-called Abelian-Tauberian theorem. Finally, we go back to the original ergodic control problem with gain functional (I.11), proving that its value $V(\xi) := \sup_{\alpha \in \mathcal{A}} J_\infty(\xi, \alpha)$ is equal to λ for all the admissible initial conditions ξ (Proposition 2.6.1). One inequality follows immediately from equality (I.13), as the right hand side is clearly larger than the supremum of (I.11) by definition, while for the opposite inequality we provide an inductive reasoning based on the fundamental equality (I.12). In Proposition 2.6.3, we therefore easily derive optimal controls for the ergodic control problem from controls realizing the supremum in (I.12), provided that the latter exist. Notice that we did not rely on a PDE to identify λ with the value, but we still want to associate a Hamilton-Jacobi-Bellman equation to this problem. In particular, in Theorem 2.5.5, we start again from an extension of the fundamental equality (I.12), identifying the value v^T (I.3) of the finite horizon optimal control problem at time t with $\phi(\mu) + \lambda(T - t)$, and we recall that v^T is a viscosity solution to the backward parabolic PDE (I.10) on the space $\mathcal{P}_2(\mathbb{R}^d)$; we thus deduce that the pair (λ, ϕ) is a viscosity solution to the following ergodic HJB equation on the same space $\mathcal{P}_2(\mathbb{R}^d)$:

$$\begin{aligned} \lambda &= \int_{\mathbb{R}^d} \left[\sup_{a \in A} \left\{ f(x, \mu, a) + \langle b(x, \mu, a), \partial_\mu \varphi(\mu)(x) \rangle \right. \right. \\ &\quad \left. \left. + \frac{1}{2} \text{tr} (\sigma \sigma^T(x, \mu, a) \partial_x \partial_\mu \varphi(\mu)(x)) \right\} \right] \mu(dx). \end{aligned} \quad (\text{I.14})$$

We also discuss uniqueness of viscosity solutions: under the assumption that one of the aforementioned comparison principles for solutions to the backward parabolic PDE (I.4) holds, we can deduce that, if $(\bar{\lambda}, \bar{\phi})$ is another solution to equation (I.14), then $\lambda = \bar{\lambda}$ (Theorem 2.5.7). Uniqueness of ϕ fails because $\phi + c$ is another solution to (I.14) for any $c \in \mathbb{R}$; moreover, due to the presence of mean-field effects, uniqueness of ϕ up to an additive constant is not granted, even if stronger hypotheses are satisfied, see Remark 2.6.7. We eventually provide a verification theorem (Theorem 2.6.6) which

states that we can construct optimal controls in explicit form for the ergodic control problem relying on the ergodic HJB equation (I.14). Notice that this is not the first work dealing with an optimal control problem of the McKean-Vlasov type in ergodic form: in particular, in [11] a similar problem, with a different class of controls and stronger assumptions, is considered. In that paper, the pair (λ, ϕ) is also constructed by means of the vanishing discount method. The main novelty of our approach is that our results are not deduced from stability properties of viscosity solutions to PDEs, which we would use letting $\beta \rightarrow 0$ in the elliptic HJB equation (I.10), as this theory is not yet available and would require a detailed study on the Wasserstein space $\mathcal{P}_2(\mathbb{R}^d)$; we rather derive all the main results from the fundamental equation (I.12). Therefore, we study the ergodic HJB equation (I.14) directly on the Wasserstein space, without lifting it to the Hilbert space L^2 as in [11], and in fact we do not need the HJB equation to identify the value of the ergodic control problem with λ . This Chapter presents an extended version of [56].

In Chapter 3, we analyse a different mean-field optimal control problem, where the dynamics of the controlled system is of the form

$$dZ_s = h(Z_s, \mathbb{Q}_{Z_s}, M_s, \alpha_s)ds + dB_s, \quad s \in [0, T]; \quad (\text{I.15})$$

here, $\mathbb{Q}_{Z_s}$ is the law of Z_s under the probability measure $\mathbb{Q}$, while M is an exogenous and non-observable random source, represented by a continuous-time Markov chain with finite state space. We formulate a partial observation problem, as the information available for the representative agent does not include all the sources of randomness in the system: in particular, the control is assumed to be predictable with respect to the filtration generated by the controlled process Z , which thus plays the role of observed process; the agent can not observe the signal M . Notice that α is not even adapted with respect to the Brownian filtration generated by B , as usually assumed in fully observable control problems: the agent does not even directly observe the Brownian noise. The analysis of the problem thus requires filtering techniques. In particular, we adopt a solution strategy based on the measure-valued Zakai equation. In principle, we construct a fully observable auxiliary control problem, called separated problem, derived from the original control problem by a heuristic reasoning including a change of the underlying probability measure à-la Girsanov. Due to the simplifying assumption that M is a finite-state chain, the dynamics of the separated control problem evolves in a Euclidean space $\mathbb{R}^N$ and is driven by the process Z , which is a Brownian motion starting at $z_0 \in \mathbb{R}^d$ under the new reference probability. The evolution of the system is described by a SDE, called Zakai equation, whose coefficients depend on the joint law of the state process and the Brownian noise in a non-standard way: in particular, the equation is of the form

$$\begin{cases} dX_s &= \Lambda^T X_s ds + \text{diag}(X_s) h\left(Z_s, \mathbb{E}[\langle X_s, 1 \rangle | Z_s = \cdot] \mathbb{P}_{Z_s}, \alpha_s\right) dZ_s, \\ X_0 &= \pi_0. \end{cases} \quad (\text{I.16})$$

The diffusion coefficient thus depends on the conditional law of X with respect to Z and on the law of Z : as explained below, the occurrence of the conditional expectation inside the function h takes into account the passage from the law $\mathbb{Q}_{Z_s}$ under the original probability measure $\mathbb{Q}$ to the law $\mathbb{P}_{Z_s}$ under the new reference probability

measure $\mathbb{P}$. Due to the awkward form of this distribution-dependent equation, we can not apply classical results to guarantee that it admits a unique solution. We thus firstly need to study the well-posedness of this equation under a suitable Lipschitz assumption on the function h (see Theorem 3.3.1): this assumption only concerns the behaviour of h when we consider two different measures which are absolutely continuous with respect to a target measure, as the coefficients of Zakai equation depend on its solution through the term $\mathbb{E} [\langle X_s, 1 \rangle | Z_s = \cdot] \mathbb{P}_{Z_s}$, where Z is fixed. Once we have proved the well-posedness of the problem, we discuss its relation with the original control problem with partial observation: changing the underlying probability measure by means of Girsanov theorem in the separated control problem, we obtain a partially observed control problem which can be interpreted as a weak formulation of the original control problem with partial observation (see Proposition 3.3.2). The aim of this approach, which is often used in optimal control of non-linear systems with partial observation, is to avoid a technical difficulty: if we start from the partial observation control problem and we change the underlying probability measure, as we do heuristically in Section 3.2, in both of these equivalent formulations of the problem the filtration to which the control α is adapted depends on its turn on α . As a matter of fact, α is predictable with respect to the filtration generated by the controlled observation process Z . Changing the probability measure directly in this problem does not solve this issue, as α is still predictable with respect to the filtration generated by Z under the new probability measure, leading to a measurability condition which is difficult to verify in practice. On the other hand, in the separated control problem the Brownian motion Z is fixed and so the filtration that it generates. Admissible controls are required to be predictable with respect to this Brownian filtration, which does not depend on the control process itself; a partially observed control problem emerges as an equivalent formulation and thus provides a motivation to the analysis of the separated control problem, which is in fact the subject of the rest of the Chapter. Notice that, in filtering theory, the dynamics of the separated control problem is described by means of a measure-valued equation, called Zakai equation, whose solution is usually identified (in the equivalent formulation) with the unnormalized conditional distribution of the unobserved source of randomness given the observation; this is true even in our problem but, as the underlying Markov chain only takes a finite number of values, we can easily reduce this equation to a $\mathbb{R}^N$ -valued SDE. We underline here that one of the main difficulties of the filtering technique based on the change of the probability measure in our problem is related to the presence of mean-field effects in equation (I.15): the law of the time-marginal of the state process Z under the original probability measure $\mathbb{Q}$, $\mathbb{Q}_{Z_s}$, will no longer represent the law of Z_s under the new probability measure $\mathbb{P}$, denoted by $\mathbb{P}_{Z_s}$. In fact, the term $\mathbb{E} [\langle X_s, 1 \rangle | Z_s = \cdot] \mathbb{P}_{Z_s}$ in equation (I.16) is introduced exactly to rewrite $\mathbb{Q}_{Z_s}$ in terms of $\mathbb{P}_{Z_s}$, leading to a separated control problem with non-standard McKean-Vlasov dependence. Following the dynamic programming approach, this problem is subsequently embedded in a larger class of optimal control problems, where we allow the dynamics to start from any time $t \in [0, T]$ and the initial conditions of the Brownian noise and of the state process to be random variables. The dynamics will thus be described by means of a system of two SDEs: the first one describes the evolution of the Brownian noise Z starting from a random initial position and the second one coincides with our non-standard

$\mathbb{R}^N$ -valued version of Zakai equation. The state is given by the pair (X, Z) because the coefficients of the second equation depend on the noise Z . As we aim at adopting the analytic approach, we need to deal with a Markovian system and thus to circumvent the presence of a state equation with an exogenous source of randomness, which would lead to the occurrence of stochastic coefficients: we thus include Z in the state. We also rewrite the state equation as a more familiar vector-valued McKean-Vlasov equation, showing that the term containing the conditional expectation can be expressed as a function, denoted by Π , of the joint law of the process (X, Z) . Thus, we obtain an equation of the form

$$\begin{cases} dX_s &= \Lambda^T X_s ds + \text{diag}(X_s) h\left(Z_s, \Pi(\mathbb{P}_{X_s, Z_s}), \alpha_s\right) dW_s, & s \in [t, T] \\ dZ_s &= dW_s, & s \in [t, T] \\ X_t &= \eta, \\ Z_t &= \xi. \end{cases}$$

In Theorem 3.4.1, we prove that such McKean-Vlasov equation is well-posed and that it is invariant with respect to the following set:

$$\mathcal{D} = \left\{ \mu \in \mathcal{P}(\mathbb{R}^N \times \mathbb{R}^d) : \mu_1(\mathbb{R}_+^N) = 1, \mu_1 \in \mathcal{P}_1(\mathbb{R}^N), \langle \bar{\mu}_1, 1 \rangle := \int_{\mathbb{R}^N} \langle x, 1 \rangle \mu_1(dx) = 1 \right\},$$

where μ_1 is the first marginal of μ , meaning that, if the joint law of the initial conditions belongs to $\mathcal{D}$, as it is natural in view of our control-theoretical application, then the joint law $\mathbb{P}_{Z_s, X_s}$ still belongs to it for any time $s \geq 0$. We also discuss stability properties of the solutions to our state equation with respect to the initial condition in Theorem 3.4.3. The introduction of random initial conditions restores a suitable flow property for this McKean-Vlasov SDE (see Proposition 3.4.4); we then prove a dynamic programming principle for our optimal control problem. We conclude by providing some insights about future work with the aim of adopting the analytic approach, characterizing the value function by means of a suitable HJB equation. This Chapter presents preliminary results, which have subsequently been extended and generalized in [55].

Chapter 1

Optimal control of McKean-Vlasov SDEs with discounted gain

1.1 Introduction

This chapter presents an expanded and detailed version of the contents of [84], where we analyse a general optimal control problem of a McKean-Vlasov stochastic differential equation with infinite time horizon and discounted gain functional through the approach based on the dynamic programming principle and Hamilton-Jacobi-Bellman equation. Fix a m -dimensional Brownian motion B and let us consider the following stochastic differential equation (SDE) over the unbounded interval $(t, +\infty)$:

$$dX_s = b(X_s, \mathbb{P}_{X_s}, \alpha_s)ds + \sigma(X_s, \mathbb{P}_{X_s}, \alpha_s)dB_s,$$

whose drift and diffusion coefficients (valued in $\mathbb{R}^d$ and $\mathbb{R}^{d \times m}$, respectively) depend not only on the d -dimensional state process X itself, but also on its probability distribution and on an input control process α . This equation is referred to as controlled McKean-Vlasov SDE with infinite time horizon. The choice of the control α is based on an optimization criterion: the representative agent aims at maximizing over a fixed set of feasible controls a gain functional of the form

$$J(t, \xi, \alpha) = \mathbb{E} \left[\int_t^{+\infty} e^{-\beta(s-t)} f(X_s, \mathbb{P}_{X_s}, \alpha_s) ds \right],$$

where X is the state process starting from ξ at time t . Here f is a suitable running gain, which may also depend on the distribution of the state process, and β is a discount rate. This chapter especially focuses on the properties of the value of the problem and on its characterization as the unique solution to a partial differential equation (Hamilton-Jacobi-Bellman equation) on the space of probability measures with finite second moment, $\mathcal{P}_2(\mathbb{R}^d)$, endowed with the 2-Wasserstein metric.

In classical control theory, the analysis of the evolution of the value relies on an essential tool, the dynamic programming principle (DPP), guaranteeing the so-called time consistency of the problem: if an optimal control for an initial couple of data

(t, x) exists, then its restriction to any later time interval $[s, T]$ must be optimal for the couple $(s, X_s^{t,x})$ (see [93] and references therein for a comprehensive discussion on this notion). Mean-field optimal control problems over a bounded time interval $[0, T]$ were initially thought to be time-inconsistent, due to the possibly non-linear dependence of the gain or cost functions on the law of the state process; see e.g. [3] and [24], where the problem was therefore approached by means of the Pontryagin Maximum Principle, as well as in [66]. The DPP was subsequently proved with feedback controls in [78] and [79], thanks to an appropriate version of the flow property of solutions to McKean-Vlasov SDEs and of their laws. [49] mainly focuses on weak and strong formulations of the problem with associated DPPs. We also mention [15], where a randomized version of the DPP was presented. In this framework, the dynamics is in fact described by a pair of SDEs: a McKean-Vlasov SDE with random initial condition and a classical SDE with deterministic initial condition and coefficients depending on the law of the solution to the first equation, as proposed in [26] in an uncontrolled setting. The value function which satisfies the DPP is associated with an optimization problem involving the solution to the second state equation and thus depends on time and on the couple (x, π) , representing the initial position and law of the system.

Once the DPP is established, a suitable partial differential equation (Hamilton-Jacobi-Bellman, HJB, equation) can usually be derived from it; the goal of this approach is to give an explicit characterization of the value of the problem as the unique solution to this equation, so that we can compute it directly for any initial time and condition, provided that we can solve the HJB equation. Notice that in the mean-field case, due to the presence of a measure-valued state variable, an appropriate definition of derivative with respect to probability measures needs to be provided: in this setting, where the state at time s is in fact given by a couple of a random variable X_s and its law $\mathbb{P}_{X_s}$, it is natural to adopt the notion of Lions derivative, introduced in [70] and detailed in [30]. It is well known that, except for specific cases, the value function is not regular enough to solve a PDE in the classical sense: for this reason, the value function is usually proved to solve the HJB equation in the weaker viscosity sense, defined by Crandall and Lions [45]. For mean-field control problems with finite time horizon, this approach has been studied in [79], where a verification theorem is also stated, and in [78], where the dynamics is subject to a common noise. In both papers, the value function depends on the initial random condition only through its law (i.e. it is law invariant) due to the specific formulation, leading to a parabolic HJB equation in the Wasserstein space of order 2; nonetheless, the equation is lifted to a suitable L^2 space, in order that the theory of viscosity solutions in Hilbert spaces (established in [67], [68] and then in [51]) can be exploited. Unfortunately, the relation between solutions to a PDE on the Wasserstein space and solutions to its lifted version is still unclear, see Remark 3.6 in [43]. As a consequence, the authors of [42] presented a more general formulation of the problem, without common noise but with open loop controls, allowing even dependence of the coefficients on the entire path of the state process, valued in an infinite dimensional Hilbert space. They showed that the value function is law invariant under mild conditions, i.e. that it depends on the initial condition only through its law. They thus defined the parabolic HJB equation directly on the 2-Wasserstein space, rewriting it in different equivalent forms and proving that it admits a viscosity solution, corresponding to the value function. The same authors

analysed uniqueness of solutions to the HJB equation in the 2-Wasserstein space in [43] by means of a N -player smooth approximation of the original problem, based on a propagation of chaos result, and Borwein-Preiss variational principle with a suitable gauge function. Their result was subsequently generalized and corrected in [36] and [14] for a second-order parabolic PDE derived from a McKean-Vlasov control problem with common noise, showing that a stronger regularity assumption on the coefficients and a different definition of viscosity supersolution are needed in order to guarantee uniqueness, see Remark 2.5 and Remark 6.1 in [36]. Different approaches to prove uniqueness of viscosity solutions to a first or second order PDE on 2-Wasserstein spaces have recently been proposed and this is still a very active area of research, as witnessed by [86], [17], [16], [46] and references therein. Notice that all the aforementioned result concern a parabolic PDE associated with a finite horizon optimal control problem.

For the infinite time horizon case, ergodic mean-field control has attracted considerable interest, especially in discrete time, where connections with reinforcement learning are possible (see e.g. [77], [18] and references therein), but also in continuous time (see e.g. [29], where a mean-field ergodic control problem is analysed via Lagrange multipliers). [19] analyses a mean-field control problem with infinite time horizon and discounted gain, similar to the one presented in this chapter, but with slightly different assumptions (e.g. additive noise, a convex set of control actions A). Instead of our analytic approach, the authors adopt a probabilistic approach: in particular, they state and prove a sufficient condition for optimality in the form of an appropriate Pontryagin maximum principle as an application to their well-posedness results for infinite horizon FBSDEs. On the other hand, the mean-field game counterpart of discounted optimization problems has been studied in [35] via probabilistic techniques and in [31] by PDE techniques on a torus. We also mention [11], where a mean-field ergodic control problem is approached by means of an approximation by exponentially weighted infinite horizon optimal control problems, similar to the one considered in this chapter. Nonetheless, as in [78] for the finite horizon case, the elliptic HJB equations associated with the discounted control problems are lifted on a suitable L^2 space, so that the Hilbert setting helps providing a uniqueness result for viscosity solutions. Moreover, only closed loop controls in feedback form are considered.

Our contribution: In this chapter, we address a discounted optimal control problem of the McKean-Vlasov type with infinite time horizon under mild assumptions (adapted from the ones in [42]) with an analytic approach, i.e. through the associated HJB equation on the 2-Wasserstein space. Our framework is also very general due to the class of admissible controls, which are open loop controls. By restricting to closed-loop controls in feedback form, we would avoid many technical difficulties, but stronger assumptions would also be required (for any admissible control α , the measurable function $\tilde{\alpha}$ such that $\alpha_t = \tilde{\alpha}(t, X_t, \mathbb{P}_{X_t})$ has to be Lipschitz and square-integrable, see e.g. [78] and [11]). Moreover, it is not clear if the value over the class of closed loop controls equals the value over the larger and more general class of open loop controls, except for some specific form of the optimization problem, see Section 6 in [78] and [63]. We therefore prefer this larger class, as in [42]: we refer the reader to the discussion in [34], stating that McKean-Vlasov control problems may be formulated as limit problem of N -player models even with open loop controls.

One of our main technical contributions is the proof of the time invariance property, stating that the value function does not depend on the initial time of the dynamics. This is not a trivial result when mean-field effects occur. Notice that, if the optimization problem starts at time $t > 0$, there are two possible classes of open loop controls which can be considered: we can take the class $\mathcal{A}$ of controls depending on the initial condition and on the entire Brownian noise from time 0 or the smaller class, denoted by $\mathcal{A}_t$, of controls depending only on the initial condition and on the increments of the Brownian motion from time t . We choose the larger class $\mathcal{A}$ because the associated value function V satisfies a suitable dynamic programming principle. Notice that optimizing over the different classes of admissible controls leads to the definition of two value functions, which may in principle be different: in particular, our value V , obtained maximizing the gain over the larger class, may be strictly larger. For the value associated to the class $\mathcal{A}_t$, the time invariance is easier to prove and relies on rather standard arguments based on time-shifting techniques applied to the controls and to the state process. Therefore, the time invariance property follows if we prove that the value at time t over $\mathcal{A}$ equals the value over $\mathcal{A}_t$. However, in general, this is not the case. Indeed, we will provide a motivating example showing that, in a McKean-Vlasov setting with rather singular coefficients, the value may worsen when we optimize over the smaller class of controls; this phenomenon is strictly related to the possibility to have a non-linear dependence of the gain function with respect to the law, which can never emerge in classical control problems. We subsequently present a second example, suggesting that the equality of the values may be restored even in a mean-field case, provided that the gain function is sufficiently smooth. We thus conclude that, under suitable regularity hypotheses, the invariance of the value with respect to the class of controls is guaranteed for our control problem. In order to prove this result, we will also provide an interesting disintegration formula of a McKean-Vlasov SDE with respect to the past of the Brownian trajectory. This formula is introduced to circumvent the impossibility to apply conditional independence directly on the probability distribution, which is a deterministic object. Notice that the technical issues described above are due to our choice of open loop controls; the value functions of discounted problems considered in [11], defined over a class of feedback controls, are easily proved to be time invariant.

After extending the law invariance property of the value function presented in [42] to our infinite horizon setting, guaranteeing that the value depends on the initial condition only through its law, we identify the value with a time-independent function on the Wasserstein space of order 2. We subsequently derive an elliptic partial differential equation on the same space and show that it is solved by our value function in a suitable viscosity sense. In order to prove uniqueness of solutions to this PDE, we construct a sequence of finite horizon approximations of our control problem and we rely on the aforementioned uniqueness results for viscosity solutions to parabolic PDEs: the stronger assumptions required for the parabolic result will thus be needed even for our theorem.

The chapter is organized as follows. In Section 1.2, we present the notation and the assumptions, together with the dynamics and the formulation of the optimal control problem, followed by the dynamic programming principle. Section 1.3 is devoted to

the time invariance property. We first show that, for an alternative formulation of the control problem with the smaller class of admissible controls $\mathcal{A}_t$, the desired result holds. We then introduce an example where the two formulations lead to different values and a second example which suggests that the continuity of the gain function may play a crucial role in establishing the equality. Finally, we rigorously prove that the two value functions are equal under suitable regularity assumptions, relying on the aforementioned disintegration formula for the McKean-Vlasov dynamics with respect to the first part of the Brownian trajectory; we postpone the proof of all the technical results to Appendix A and Appendix B. In Section 1.4, we present the law invariance property and other essential properties of the value function in this infinite time horizon framework. We deduce that it can in fact be redefined as a function on the 2-Wasserstein space $\mathcal{P}_2(\mathbb{R}^d)$. Finally, Section 1.5 is concerned with the elliptic Hamilton-Jacobi-Bellman equation on $\mathcal{P}_2(\mathbb{R}^d)$ and the proof that the value function is its unique viscosity solution.

1.2 McKean-Vlasov SDEs and optimal control

1.2.1 Notation and assumptions

Space state

Fix two integers $d \geq 1$, representing the dimension of the space state, and $m \geq 1$. We denote by $\langle \cdot, \cdot \rangle$ the standard inner product in $\mathbb{R}^d$ and by $|\cdot|$ the associated Euclidean norm of a vector. Given a matrix $M \in \mathbb{R}^{d \times m}$, we define its Frobenius norm $|M| := \sqrt{\sum_{i=1}^d \sum_{j=1}^m |M_{ij}|^2} = \sqrt{\text{tr}(MM^T)}$, where M^T is the transpose of M and $\text{tr}(\cdot)$ is the trace of a square matrix.

Spaces of probability measures

Denote by $\mathcal{P}(\mathbb{R}^d)$ the space of probability measures on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, where $\mathcal{B}(\mathbb{R}^d)$ is the Borel σ -algebra of $\mathbb{R}^d$. For any probability measure $\mu \in \mathcal{P}(\mathbb{R}^d)$, we denote by $\|\mu\|_2$ the second moment of the measure, namely

$$\|\mu\|_2 := \left(\int_{\mathbb{R}^d} |x|^2 \mu(dx) \right)^{\frac{1}{2}}.$$

Define the space $\mathcal{P}_2(\mathbb{R}^d) := \{\mu \in \mathcal{P}(\mathbb{R}^d) : \|\mu\|_2 < +\infty\}$. We endow it with the 2-Wasserstein metric defined as

$$W_2(\mu, \nu) := \inf \left\{ \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2 \pi(dx, dy) : \pi \in \mathcal{P}_2(\mathbb{R}^d \times \mathbb{R}^d) \right. \\ \left. \text{such that } \pi(\mathbb{R}^d \times \cdot) = \nu, \pi(\cdot \times \mathbb{R}^d) = \mu \right\}^{\frac{1}{2}}$$

for every $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$. The space $(\mathcal{P}_2(\mathbb{R}^d), W_2)$ is a Polish space (see e.g. Theorem 6.18 in [90]).

1.2.2 Probabilistic setting

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space, on which an m -dimensional Brownian motion $B = (B_t)_{t \geq 0}$ is defined. For every random variable η on $(\Omega, \mathbb{F}, \mathbb{P})$, we denote by $\mathbb{P}_\eta$ or $\mathcal{L}(\eta)$ the law of η . Call $\mathbb{F}^B = (\mathcal{F}_t^B)_{t \geq 0}$ the $\mathbb{P}$ -completion of the filtration generated by B . Fix a sub- σ -algebra $\mathcal{G}$ of $\mathcal{F}$ satisfying the following conditions:

Assumptions $(H_{\mathcal{G}})$.

- (i) $\mathcal{G}$ is the σ -algebra generated by a random variable U with uniform distribution on $[0, 1]$, $U \sim \text{Unif}([0, 1])$.
- (ii) U and B are independent, i.e. $\mathcal{F}_\infty^B = \bigcup_{t \geq 0} \mathcal{F}_t^B$ and $\mathcal{G}$ are independent.

We will require initial conditions to be $\mathcal{G}$ -measurable. Notice that condition (i) guarantees that for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ there exists a $\mathcal{G}$ -measurable, square-integrable

random variable $\xi : \Omega \rightarrow \mathbb{R}^d$ such that $\mu = \mathbb{P}_\xi$, as shown in Lemma 2.1 in [42]. Thus, the σ -algebra $\mathcal{G}$ is “sufficiently rich”. Define the filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ given by $\mathcal{F}_t = \sigma(\mathcal{G} \cup \mathcal{F}_t^B)$. In the following sections, L^2 stands for the Hilbert space of square-integrable random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ endowed with the norm $\|\xi\|_{L^2} := (\mathbb{E}[|\xi|^2])^{\frac{1}{2}}$.

Let A be a compact subset of a Euclidean space $\mathbb{R}^N$, representing the space of control actions. Define the following space of admissible controls:

$$\mathcal{A} = \{\alpha : [0, +\infty) \times \Omega \rightarrow A : \alpha \text{ is } \mathbb{F}\text{-predictable}\}.$$

Remark 1.2.1. For the formulation of the problem and the properties of the value function we could relax the hypotheses on A , assuming that it is a Polish space. We will need A to have at least a Hilbert space structure in Section 1.5 to introduce the Lions derivative on $\mathcal{P}_2(\mathbb{R}^d \times A)$ for the definition of viscosity supersolution; on the other hand, the set A is required to be compact and finite dimensional for the result of uniqueness of viscosity solutions on the Wasserstein space of probability measures (see e.g. Remark 2.1 in [43]). $\square$

Remark 1.2.2. We could equivalently define $\mathcal{A}$ as the class of A -valued progressively measurable processes. Indeed, combining Lemma 3.5.(ii) and Theorem 3.7 in [37], for any progressive process α there exists a predictable process $\tilde{\alpha}$ such that $\tilde{\alpha}(s, \omega) = \alpha(s, \omega)$ for $ds \otimes d\mathbb{P}$ -a.e. (s, ω) , generating the same trajectories as α up to null sets and, thus, the same gain. $\square$

Notice that the predictable σ -algebra has an explicit class of generators, used in the following representation of the admissible controls. The following lemma is inspired by Proposition 10 in [38], adapted to the present framework, where the filtration corresponds to an initial enlargement of a Brownian filtration and the probability space is not necessarily canonical. We recall here that the canonical σ -algebra on the space of continuous functions $\mathcal{C}([0, +\infty))$ is equal to the Borel σ -algebra of the topology of uniform convergence on bounded and closed intervals.

Lemma 1.2.3. Let α be a $\mathbb{F}$ -predictable process valued in A . Then, there exists a Borel measurable function $\underline{\alpha} : [0, +\infty) \times \mathcal{C}([0, +\infty); \mathbb{R}^m) \times [0, 1] \rightarrow A$ such that, for $\mathbb{P}$ -a.e. $\omega \in \Omega$,

$$\alpha_s(\omega) = \underline{\alpha}(s, B_{\cdot \wedge s}(\omega), U(\omega)) \quad \forall s \geq 0, \quad (1.1)$$

where $B_{\cdot \wedge s}(\omega)$ denotes the trajectory of the Brownian motion B associated with ω and stopped at time s .

Proof. Step 1: Denote by $\mathbb{F}^{B,0}$ the filtration generated by the Brownian motion B and not completed and by $\mathbb{F}^0 = (\mathcal{F}_t^0)_{t \geq 0}$ the filtration defined as $\mathcal{F}_t^0 = \sigma(\mathcal{F}_t^{B,0} \cup \mathcal{G})$ and not completed. As the thesis holds up to indistinguishability, we can restrict to $\mathbb{F}^0$ -predictable controls without loss of generality. Indeed, by Theorem 4.37 in [57], any stochastic process predictable with respect to $\mathbb{F}$ is indistinguishable from a $\mathbb{F}^0$ -predictable process. Moreover, we can modify the Brownian motion and the control on a null set, assuming that $B(\omega) \equiv 0$ and $\alpha(\omega) \equiv 0$ for any ω such that the associated trajectory of the Brownian motion is not continuous: we thus obtain a process being indistinguishable from the original control α for which we want to prove the thesis.

Step 2: The result is obtained by a monotone class reasoning. Firstly, we consider the class of predictable processes of the form $\alpha = \mathbb{1}_C$ for some C in the π system $\mathcal{C} := \mathcal{C}_1 \cup \mathcal{C}_2$, where $\mathcal{C}_1 = \{[0, +\infty) \times G : G \in \mathcal{G}\}$ and $\mathcal{C}_2 = \{(s, +\infty) \times (S \cap G) : s \geq 0, S \in \mathcal{F}_s^{B,0}, G \in \mathcal{G}\}$. The processes belonging to this set can be written in the form (1.1), where $\underline{\alpha}$ is a product of indicator functions, by an application of Doob's measurability lemma (Lemma 1.14 [62]). Using the Monotone Class Theorem (see e.g. Theorem 2.1, Chapter 0, in [81]), we subsequently extend the class of processes admitting the representation (1.1), showing that it contains all the indicator processes of the form $\mathbb{1}_K$ for $K \in \mathcal{K} := \{[0, +\infty) \times G : G \in \mathcal{G}\} \cup \{(t, +\infty) \times S : t \geq 0, S \in \mathcal{F}_t^0\}$. Finally, notice that, by Lemma 10.2 in [62], the π system $\mathcal{K}$ generates the predictable σ -algebra $\mathcal{P}(\mathbb{F}^0)$: using the Functional version of the Monotone Class Theorem (see e.g. Lemma 1.6 by [94]), we conclude that all the $\mathbb{F}^0$ -predictable processes satisfy the thesis. $\square$

Therefore, in what follows we will identify (with a slight abuse of notation) the class of admissible controls with the class of processes admitting the representation (1.1) for a suitable measurable function $\underline{\alpha}$. Notice that a similar class of controls has been adopted in [47].

1.2.3 Dynamics of the controlled system

Consider the coefficients

$$b : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times A \rightarrow \mathbb{R}^d \text{ and } \sigma : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times A \rightarrow \mathbb{R}^{d \times m}.$$

Notice that they are both independent from time; this assumption will be essential to obtain the independence of the value function from the initial time.

Suppose that they satisfy the following assumptions:

Assumptions $(H_{b,\sigma})$.

- (H1) b and σ are Borel functions and, for each fixed $(x, \mu) \in \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, the functions $b(x, \mu, \cdot)$ and $\sigma(x, \mu, \cdot)$ are continuous;
- (H2) b and σ are Lipschitz in (x, μ) uniformly with respect to a , i.e. there exists $L > 0$ such that

$$|b(x, \mu, a) - b(x', \mu', a)| + |\sigma(x, \mu, a) - \sigma(x', \mu', a)| \leq L(|x - x'| + W_2(\mu, \mu'))$$

for all $x, x' \in \mathbb{R}^d$, $\mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$, $a \in A$;

- (H3) b and σ satisfy a linear growth condition in (x, μ) uniformly with respect to a , i.e. there exists $M > 0$ such that

$$|b(x, \mu, a)| + |\sigma(x, \mu, a)| \leq M(1 + |x| + \|\mu\|_2)$$

for all $x \in \mathbb{R}^d$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $a \in A$.

Remark 1.2.4. The continuity in a of the functions $b(x, \mu, \cdot)$ and $\sigma(x, \mu, \cdot)$ is not needed to guarantee existence and uniqueness of solutions to the state equation, but we will use it in Remark 1.5.10 to apply a measurable selection argument (Proposition 7.24 in [22]) and write Hamilton-Jacobi-Bellman equation in the form $\mathbb{E}[\sup_{a \in A} H(x, \mu, a)] = 0$ for a suitable function H . The same remark applies to the gain function f introduced in Section 1.2.4. $\square$

Given $\alpha \in \mathcal{A}$, for all $t \in [0, +\infty)$, $\xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^d)$, consider the state equation

$$\begin{cases} dX_s = b(X_s, \mathbb{P}_{X_s}, \alpha_s)ds + \sigma(X_s, \mathbb{P}_{X_s}, \alpha_s)dB_s, & s \in (t, \infty) \\ X_t = \xi. \end{cases} \quad (1.2)$$

The following result guarantees that this equation is well-posed.

Proposition 1.2.5. Assume that hypotheses $(\mathbf{H}_{b,\sigma})$ hold. Then, for any $\alpha \in \mathcal{A}$ and $t \in [0, +\infty)$, $\xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P}; \mathbb{R}^d)$, there exists a unique (up to indistinguishability) continuous and $\mathbb{F}$ -adapted solution to equation (1.2), denoted by $X^{t,\xi,\alpha}$.

Proof. Given any $0 \leq t < T$, denote by $\mathbb{S}_2^{t,T}$ the set of all continuous and $\mathbb{F}$ -adapted processes X (identified up to indistinguishability) defined on $\Omega \times [t, T]$ such that

$$\|X\|_{\mathbb{S}_2^{t,T}} := \left(\mathbb{E} \left[\sup_{s \in [t, T]} |X_s|^2 \right] \right)^{\frac{1}{2}}$$

is finite. Under hypotheses $(\mathbf{H}_{b,\sigma})$, Theorem 4.21 in [32] guarantees that, for every fixed $T > 0$, there exists a unique solution $X^{t,\xi,\alpha;T} \in \mathbb{S}_2^{t,T}$ to the following equation

$$\begin{cases} dX_s = b(X_s, \mathbb{P}_{X_s}, \alpha_s)ds + \sigma(X_s, \mathbb{P}_{X_s}, \alpha_s)dB_s, & s \in [t, T] \\ X_t = \xi, \end{cases} \quad (1.3)$$

i.e to equation (1.2) “truncated” at time T . Consider $T_1 < T_2$ and notice that, by uniqueness of solutions for McKean-Vlasov equations on a finite horizon interval,

$$X^{t,\xi,\alpha;T_1} = X_{|[t,T_1]}^{t,\xi,\alpha;T_2} \quad \text{in } \mathbb{S}_2^{t,T_1}.$$

Hence, taking an increasing sequence of times $(T_n)_{n \in \mathbb{N}}$ diverging to $+\infty$ and using the fact that $\forall n \in \mathbb{N}$ the solution $X^{t,\xi,\alpha;T_n}$ coincides with $X^{t,\xi,\alpha;T_{n-1}}$ on the interval $[t, T_{n-1}]$, we deduce that there exists a unique solution $X^{t,\xi,\alpha}$ to equation (1.2) on the interval $[t, +\infty)$. Moreover, it is clear that $X^{t,\xi,\alpha}$ is continuous and $\mathbb{F}$ -adapted because $X^{t,\xi,\alpha;T_n}$ is continuous and $\mathbb{F}$ -adapted for every $n \in \mathbb{N}$. $\square$

From the construction of the solution to equation (1.2), it is easily verified that

$$\mathbb{E} \left[\sup_{s \in [t, T]} |X_s^{t,\xi,\alpha}|^2 \right] < +\infty \quad \text{for any } T > 0.$$

Moreover, standard arguments, relying essentially on the Itô formula and the Gronwall Lemma (see for instance Theorem 1 in [74]), lead to the following estimate:

$$\mathbb{E}[|X_s^{t,\xi,\alpha}|^2] \leq \left(\|\xi\|_{L^2}^2 + 6M^2(s-t) \right) e^{(12M^2+1)(s-t)}, \quad (1.4)$$

valid for $0 \leq t \leq s$, $\alpha \in \mathcal{A}$ and $\xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$, where M is the constant appearing in Assumption (H3) in $(\mathbf{H}_{b,\sigma})$. For more details, we refer the reader to Appendix A.2.

1.2.4 Gain functional and optimization problem

Given the dynamics (1.2), we want to introduce an optimization problem. Fix a discount factor $\beta > 0$. Define the gain function:

$$f : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times A \rightarrow \mathbb{R}.$$

Suppose that the following assumptions are satisfied:

Assumptions $(\mathbf{H}_{f,\beta})$.

(H4) f is a Borel measurable function and, for each fixed $(x, \mu) \in \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, the function $f(x, \mu, \cdot)$ is continuous;

(H5) f satisfies a quadratic growth condition in (x, μ) uniformly with respect to a , i.e. there exists $K > 0$ such that

$$|f(x, \mu, a)| \leq K(1 + |x|^2 + \|\mu\|_2^2) \quad \text{for all } x \in \mathbb{R}^d, \mu \in \mathcal{P}_2(\mathbb{R}^d), a \in A.$$

(H6) $\beta > 18M^2 + 2$.

Define the gain functional

$$J(t, \xi, \alpha) := \mathbb{E} \left[\int_t^{+\infty} e^{-\beta(s-t)} f(X_s^{t,\xi,\alpha}, \mathbb{P}_{X_s^{t,\xi,\alpha}}, \alpha_s) ds \right] \quad (1.5)$$

for every $t \geq 0$, $\xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$ and $\alpha \in \mathcal{A}$.

Remark 1.2.6. It can be easily verified that, under hypotheses $(\mathbf{H}_{f,\beta})$, the gain functional is well defined for every $t \geq 0$, $\xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$ and $\alpha \in \mathcal{A}$. As a matter of fact,

$$\begin{aligned} & \mathbb{E} \left[\int_t^{+\infty} e^{-\beta(s-t)} |f(X_s^{t,\xi,\alpha}, \mathbb{P}_{X_s^{t,\xi,\alpha}}, \alpha_s)| ds \right] \quad (\text{by (H5)}) \\ & \leq K \mathbb{E} \left[\int_t^{+\infty} e^{-\beta(s-t)} (1 + |X_s^{t,\xi,\alpha}|^2 + \|\mathbb{P}_{X_s^{t,\xi,\alpha}}\|_2^2) ds \right] \quad (\text{Fubini theorem}) \\ & = K \int_t^{+\infty} e^{-\beta(s-t)} (1 + \mathbb{E}[|X_s^{t,\xi,\alpha}|^2] + \|\mathbb{P}_{X_s^{t,\xi,\alpha}}\|_2^2) ds \quad (\text{by definition of } \|\mu\|_2^2) \\ & = K \int_t^{+\infty} e^{-\beta(s-t)} (1 + 2\mathbb{E}[|X_s^{t,\xi,\alpha}|^2]) ds \quad (\text{using the estimate (1.4)}) \\ & \leq K \int_t^{+\infty} e^{-\beta(s-t)} \left(1 + 2 \left(\|\xi\|_{L^2}^2 + 6M^2(s-t) \right) e^{(12M^2+1)(s-t)} \right) ds \\ & = K \int_t^{+\infty} \left[e^{-\beta(s-t)} + 2 \left(\|\xi\|_{L^2}^2 + 6M^2(s-t) \right) e^{(12M^2+1-\beta)(s-t)} \right] ds < +\infty \end{aligned}$$

for $\beta > 18M^2 + 2$. Notice that a weaker assumption, i.e. $\beta > 12M^2 + 1$, would be sufficient, but we will need a stronger bound in Section 1.3.3, see Remark 1.3.12. $\square$

The following alternative set of assumptions also guarantees that the gain functional is well defined:

Assumptions $(\mathbf{H}'_{f,\beta})$.

(H4) f is a Borel measurable function and, for each fixed $(x, \mu) \in \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, the function $f(x, \mu, \cdot)$ is continuous;

(H5)' f is bounded, i.e. there exists a constant C_f such that $|f(x, \mu, a)| \leq C_f$ for any $x \in \mathbb{R}^d$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $a \in A$.

(H6)' $\beta > 0$.

We introduce an additional continuity assumption on the function f :

Assumptions $(\mathbf{H}_{f_{cont}})$. The function f is locally Hölder continuous in (x, μ) uniformly with respect to a , i.e. there exists constants $H > 0$, $\gamma_1 \in (0, 1]$, $\gamma_2 \in (0, 1]$ such that

$$|f(x, \mu, a) - f(x', \mu', a)| \leq H[|x - x'|^{\gamma_1}(1 + |x| + |x'|)^{2-\gamma_1} + W_2(\mu, \mu')^{\gamma_2}(1 + \|\mu\|_2 + \|\mu'\|_2)^{2-\gamma_2}] \quad (1.6)$$

for every $x, x' \in \mathbb{R}^d$, $\mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$ and $a \in A$.

Remark 1.2.7. Notice that this assumption also appears in [47]; it implies that the function f is locally uniformly continuous in (x, μ) uniformly with respect to a , as required in [42]. $\square$

Our goal is the maximization of the gain functional over the class of all feasible controls. We thus define a real function V called *lifted value function* as follows:

$$V(t, \xi) := \sup_{\alpha \in \mathcal{A}} J(t, \xi, \alpha) \quad (1.7)$$

for every $t \in [0, +\infty)$ and $\xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$.

Remark 1.2.8. Notice that V is not defined on a product space, as the space to which ξ belongs depends on t . We also underline that we will avoid the typical measurability issues emerging in the dynamical programming principle because V is defined directly on L^2 . $\square$

We now state the flow property and the dynamic programming principle for the lifted value function defined in (1.7) in this infinite horizon setting.

Lemma 1.2.9 (Flow Property). For every $0 \leq t \leq s \leq r < +\infty$, $\xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$, and for every fixed $\alpha \in \mathcal{A}$,

$$X_r^{t,\xi,\alpha} = X_r^{s,X_s^{t,\xi,\alpha},\alpha}. \quad (1.8)$$

Proof. Fix any $0 \leq t \leq s$. Call $Y = (Y_r)_{r \geq s}$ the process such that $Y_r := X_r^{t,\xi,\alpha}$ for any $r \geq s$. Hence, for any $r \geq s$,

$$Y_r = X_r^{t,\xi,\alpha} = \xi + \int_t^r b(X_u^{t,\xi,\alpha}, \mathbb{P}_{X_u^{t,\xi,\alpha}}, \alpha_u) du + \int_t^r \sigma(X_u^{t,\xi,\alpha}, \mathbb{P}_{X_u^{t,\xi,\alpha}}, \alpha_u) dB_u$$

$$\begin{aligned}
&= \xi + \int_t^s b(X_u^{t,\xi,\alpha}, \mathbb{P}_{X_u^{t,\xi,\alpha}}, \alpha_u) du + \int_t^s \sigma(X_u^{t,\xi,\alpha}, \mathbb{P}_{X_u^{t,\xi,\alpha}}, \alpha_u) dB_u \\
&+ \int_s^r b(X_u^{t,\xi,\alpha}, \mathbb{P}_{X_u^{t,\xi,\alpha}}, \alpha_u) du + \int_s^r \sigma(X_u^{t,\xi,\alpha}, \mathbb{P}_{X_u^{t,\xi,\alpha}}, \alpha_u) dB_u \\
&= X_s^{t,\xi,\alpha} + \int_s^r b(X_u^{t,\xi,\alpha}, \mathbb{P}_{X_u^{t,\xi,\alpha}}, \alpha_u) du \\
&+ \int_s^r \sigma(X_u^{t,\xi,\alpha}, \mathbb{P}_{X_u^{t,\xi,\alpha}}, \alpha_u) dB_u \quad (\text{as } Y = X^{t,\xi,\alpha} \text{ on } [s, +\infty)) \\
&= X_s^{t,\xi,\alpha} + \int_s^r b(Y_u, \mathbb{P}_{Y_u}, \alpha_u) du + \int_s^r \sigma(Y_u, \mathbb{P}_{Y_u}, \alpha_u) dB_u.
\end{aligned}$$

As a consequence, Y solves the following equation on the interval $[s, +\infty)$:

$$\begin{cases} dY_r = b(Y_r, \mathbb{P}_{Y_r}, \alpha_r) dr + \sigma(Y_r, \mathbb{P}_{Y_r}, \alpha_r) dB_r \\ Y_s = X_s^{t,\xi,\alpha}, \end{cases}$$

which corresponds to the state equation (1.2) with initial time s , initial condition $X_s^{t,\xi,\alpha}$ and control α . By uniqueness of solutions to McKean-Vlasov SDEs (Proposition 1.2.5), we conclude that $Y_r = X_r^{s, X_s^{t,\xi,\alpha}, \alpha}$ for every $r \geq s$. $\square$

Theorem 1.2.10 (Dynamic programming principle). Suppose that assumptions $(\mathbf{H}_{b,\sigma})$ and $(\mathbf{H}_{f,\beta})$ (or $(\mathbf{H}'_{f,\beta})$) hold. Then, the lifted value function satisfies the following **dynamic programming principle (DPP)**: for $0 \leq t \leq s < +\infty$ and for any $\xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$

$$\begin{aligned} V(t, \xi) = \sup_{\alpha \in \mathcal{A}} \Big\{ &\mathbb{E} \left[\int_t^s e^{-\beta(r-t)} f(X_r^{t,\xi,\alpha}, \mathbb{P}_{X_r^{t,\xi,\alpha}}, \alpha_r) dr \right] \\ &+ e^{-\beta(s-t)} V(s, X_s^{t,\xi,\alpha}) \Big\}. \end{aligned} \tag{1.9}$$

The proof follows the lines of the proof of Theorem 3.4 in [42], with slight modifications due to the presence of the discount rate. We write it here for completeness.

Proof. For every fixed $\alpha \in \mathcal{A}$, $t \geq 0$ and $\xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$, we call

$$\gamma(t, \xi, \alpha) := \mathbb{E} \left[\int_t^s e^{-\beta(r-t)} f(X_r^{t,\xi,\alpha}, \mathbb{P}_{X_r^{t,\xi,\alpha}}, \alpha_r) dr \right] + e^{-\beta(s-t)} V(s, X_s^{t,\xi,\alpha})$$

and

$$\Lambda(t, \xi) := \sup_{\alpha \in \mathcal{A}} \gamma(t, \xi, \alpha).$$

We want to show that

$$V(t, \xi) = \Lambda(t, \xi).$$

First step: We prove that

$$V(t, \xi) \leq \Lambda(t, \xi).$$

Firstly, notice that, by definition of value function and of supremum, for every fixed $\alpha \in \mathcal{A}$,

$$\begin{aligned} V(s, X_s^{t,\xi,\alpha}) &= \sup_{\delta \in \mathcal{A}} \mathbb{E} \left[\int_s^{+\infty} f(X_r^{s,X_s^{t,\xi,\alpha},\delta}, \mathbb{P}_{X_r^{s,X_s^{t,\xi,\alpha},\delta}}, \delta_r) e^{-\beta(r-s)} dr \right] \quad (\text{choosing } \delta = \alpha) \\ &\geq \mathbb{E} \left[\int_s^{+\infty} f(X_r^{s,X_s^{t,\xi,\alpha},\alpha}, \mathbb{P}_{X_r^{s,X_s^{t,\xi,\alpha},\alpha}}, \alpha_r) e^{-\beta(r-s)} dr \right] \\ &= \mathbb{E} \left[\int_s^{+\infty} f(X_r^{t,\xi,\alpha}, \mathbb{P}_{X_r^{t,\xi,\alpha}}, \alpha_r) e^{-\beta(r-s)} dr \right] \end{aligned}$$

thanks to the flow property (1.8). Thus,

$$\begin{aligned} \gamma(t, \xi, \alpha) &= \mathbb{E} \left[\int_t^s e^{-\beta(r-t)} f(X_r^{t,\xi,\alpha}, \mathbb{P}_{X_r^{t,\xi,\alpha}}, \alpha_r) dr \right] \\ &\quad + e^{-\beta(s-t)} V(s, X_s^{t,\xi,\alpha}) \\ &\geq \mathbb{E} \left[\int_t^s e^{-\beta(r-t)} f(X_r^{t,\xi,\alpha}, \mathbb{P}_{X_r^{t,\xi,\alpha}}, \alpha_r) dr \right] \\ &\quad + e^{-\beta(s-t)} \mathbb{E} \left[\int_s^{+\infty} f(X_r^{t,\xi,\alpha}, \mathbb{P}_{X_r^{t,\xi,\alpha}}, \alpha_r) e^{-\beta(r-s)} dr \right] \\ &= \mathbb{E} \left[\int_t^s e^{-\beta(r-t)} f(X_r^{t,\xi,\alpha}, \mathbb{P}_{X_r^{t,\xi,\alpha}}, \alpha_r) dr \right] \\ &\quad + \mathbb{E} \left[\int_s^{+\infty} e^{-\beta(r-t)} f(X_r^{t,\xi,\alpha}, \mathbb{P}_{X_r^{t,\xi,\alpha}}, \alpha_r) dr \right] \\ &= \mathbb{E} \left[\int_t^{+\infty} e^{-\beta(r-t)} f(X_r^{t,\xi,\alpha}, \mathbb{P}_{X_r^{t,\xi,\alpha}}, \alpha_r) dr \right]. \end{aligned}$$

Thus, taking the supremum over $\mathcal{A}$ in both members, we conclude that

$$\Lambda(t, \xi) \geq V(t, \xi).$$

Second step: We prove that

$$\Lambda(t, \xi) \leq V(t, \xi).$$

By definition of supremum, for all $\epsilon > 0$ there exists a control $\alpha^\epsilon \in \mathcal{A}$:

$$\begin{aligned} \Lambda(t, \xi) - \epsilon &\leq \gamma(t, \xi, \alpha^\epsilon) \\ &= \mathbb{E} \left[\int_t^s e^{-\beta(r-t)} f(X_r^{t,\xi,\alpha^\epsilon}, \mathbb{P}_{X_r^{t,\xi,\alpha^\epsilon}}, \alpha_r^\epsilon) dr \right] \\ &\quad + e^{-\beta(s-t)} V(s, X_s^{t,\xi,\alpha^\epsilon}). \end{aligned} \tag{1.10}$$

Once again, by definition of supremum and value function, for all $\epsilon > 0$ there exists a control $\delta^\epsilon \in \mathcal{A}$:

$$V(s, X_s^{t,\xi,\alpha^\epsilon}) - \epsilon \leq J(s, X_s^{t,\xi,\alpha^\epsilon}, \delta^\epsilon)$$

$$= \mathbb{E} \left[\int_s^{+\infty} e^{-\beta(r-s)} f(X_r^{s, X_s^{t, \xi, \alpha^\epsilon}}, \delta^\epsilon, \mathbb{P}_{X_r^{s, X_s^{t, \xi, \alpha^\epsilon}}, \delta^\epsilon}, \delta_r^\epsilon) dr \right]. \quad (1.11)$$

Now we fix $\epsilon > 0$ and we define a control glueing the two associated controls:

$$\alpha^* := \alpha^\epsilon \mathbb{1}_{[t, s]} + \delta^\epsilon \mathbb{1}_{(s, +\infty)} \in \mathcal{A}. \quad (1.12)$$

Putting together equations (1.10) and (1.11), we get

$$\begin{aligned} \Lambda(t, \xi) &\leq \epsilon + \mathbb{E} \left[\int_t^s e^{-\beta(r-t)} f(X_r^{t, \xi, \alpha^\epsilon}, \mathbb{P}_{X_r^{t, \xi, \alpha^\epsilon}, \alpha_r^\epsilon}) dr \right] \\ &\quad + e^{-\beta(s-t)} V(s, X_s^{t, \xi, \alpha^\epsilon}) \quad (\alpha^\epsilon = \alpha^* \text{ on } [t, s]) \\ &= \epsilon + \mathbb{E} \left[\int_t^s e^{-\beta(r-t)} f(X_r^{t, \xi, \alpha^*}, \mathbb{P}_{X_r^{t, \xi, \alpha^*}, \alpha_r^*}) dr \right] \\ &\quad + e^{-\beta(s-t)} V(s, X_s^{t, \xi, \alpha^*}) \quad (\text{by inequality (1.11)}) \\ &\leq \epsilon + e^{-\beta(s-t)} \epsilon + \mathbb{E} \left[\int_t^s e^{-\beta(r-t)} f(X_r^{t, \xi, \alpha^*}, \mathbb{P}_{X_r^{t, \xi, \alpha^*}, \alpha_r^*}) dr \right] \\ &\quad + e^{-\beta(s-t)} \mathbb{E} \left[\int_s^{+\infty} e^{-\beta(r-s)} f(X_r^{s, X_s^{t, \xi, \alpha^\epsilon}}, \delta^\epsilon, \mathbb{P}_{X_r^{s, X_s^{t, \xi, \alpha^\epsilon}}, \delta^\epsilon}, \delta_r^\epsilon) dr \right] \\ &\leq \epsilon + e^{-\beta(s-t)} \epsilon + \mathbb{E} \left[\int_t^s e^{-\beta(r-t)} f(X_r^{t, \xi, \alpha^*}, \mathbb{P}_{X_r^{t, \xi, \alpha^*}, \alpha_r^*}) dr \right] \\ &\quad + \mathbb{E} \left[\int_s^{+\infty} e^{-\beta(r-t)} f(X_r^{s, X_s^{t, \xi, \alpha^\epsilon}}, \delta^\epsilon, \mathbb{P}_{X_r^{s, X_s^{t, \xi, \alpha^\epsilon}}, \delta^\epsilon}, \delta_r^\epsilon) dr \right] \quad (\text{by (1.12)}) \\ &= \epsilon + e^{-\beta(s-t)} \epsilon + \mathbb{E} \left[\int_t^s e^{-\beta(r-t)} f(X_r^{t, \xi, \alpha^*}, \mathbb{P}_{X_r^{t, \xi, \alpha^*}, \alpha_r^*}) dr \right] \\ &\quad + \mathbb{E} \left[\int_s^{+\infty} e^{-\beta(r-t)} f(X_r^{s, X_s^{t, \xi, \alpha^*}}, \delta^\epsilon, \mathbb{P}_{X_r^{s, X_s^{t, \xi, \alpha^*}}, \delta^\epsilon}, \delta_r^\epsilon) dr \right] \quad (\text{by (1.12)}) \\ &= \epsilon + e^{-\beta(s-t)} \epsilon + \mathbb{E} \left[\int_t^s e^{-\beta(r-t)} f(X_r^{t, \xi, \alpha^*}, \mathbb{P}_{X_r^{t, \xi, \alpha^*}, \alpha_r^*}) dr \right] \\ &\quad + \mathbb{E} \left[\int_s^{+\infty} e^{-\beta(r-t)} f(X_r^{s, X_s^{t, \xi, \alpha^*}}, \alpha^*, \mathbb{P}_{X_r^{s, X_s^{t, \xi, \alpha^*}}, \alpha^*}, \alpha_r^*) dr \right] \quad (\text{by (1.8)}) \\ &= \epsilon + e^{-\beta(s-t)} \epsilon + \mathbb{E} \left[\int_t^s e^{-\beta(r-t)} f(X_r^{t, \xi, \alpha^*}, \mathbb{P}_{X_r^{t, \xi, \alpha^*}, \alpha_r^*}) dr \right] \\ &\quad + \mathbb{E} \left[\int_s^{+\infty} e^{-\beta(r-t)} f(X_r^{t, \xi, \alpha^*}, \mathbb{P}_{X_r^{t, \xi, \alpha^*}, \alpha_r^*}) dr \right] \\ &= \epsilon(1 + e^{-\beta(s-t)}) + \mathbb{E} \left[\int_t^{+\infty} e^{-\beta(r-t)} f(X_r^{t, \xi, \alpha^*}, \mathbb{P}_{X_r^{t, \xi, \alpha^*}, \alpha_r^*}) dr \right] \\ &\leq 2\epsilon + J(t, \xi, \alpha^*) \leq 2\epsilon + V(t, \xi), \end{aligned}$$

where the last line immediately follows from the fact that $e^{-\beta(s-t)} \leq 1$. Thus, by the arbitrariness of $\epsilon > 0$, we deduce that

$$\Lambda(t, \xi) \leq V(t, \xi).$$

□

1.3 The time invariance property

1.3.1 Preliminary considerations

In general, when the time horizon is not finite and the coefficients are autonomous, the value function of a discounted optimal control problem does not depend on time: intuitively, if the time available to realise the gain is not bounded, the maximum attainable gain should not be affected by the initial time instant. In this section we prove that this property holds even in the McKean-Vlasov case for the function $V(t, \xi)$.

As a preliminary step, for any starting time $t \geq 0$ we introduce an alternative class of controls, denoted by $\mathcal{A}_t$, and the corresponding value function $V_t(t, \xi)$. Denote by $B^{(t)} = (B_s^{(t)})_{s \geq t}$ the process given by the increments of B from time t , i.e. $B_s^{(t)} = B_s - B_t$ for every $s \geq t$. Notice that $B^{(t)}$ is a Brownian motion on $[t, +\infty)$. We set $\mathcal{A}_t := \{\alpha : [t, +\infty) \times \Omega \rightarrow A : \alpha \text{ is } \mathbb{F}^t\text{-predictable}\}$, where $\mathbb{F}^t$ is the filtration $\mathbb{F}^t = (\mathcal{F}_s^t)_{s \geq t}$ such that $\mathcal{F}_s^t = \sigma(\mathcal{F}_s^{B,t} \cup \mathcal{G})$ and $\mathbb{F}^{B,t} = (\mathcal{F}_s^{B,t})_{s \geq t}$ is the completed Brownian filtration of $B^{(t)}$. Intuitively, $\mathcal{A}_t$ can be interpreted as the class of controls from time t that “forgets” the past of the Brownian motion. Choosing $\mathcal{A}_t$ instead of $\mathcal{A}$ from the beginning is more natural from a modelling viewpoint, but may generate difficulties in the proof of the dynamic programming principle. In particular, repeating the proof of Theorem 1.2.10 with the class of controls $\mathcal{A}_t$, in the first step it is not possible to take $\delta = \alpha$ as they belong to different classes of controls ($\alpha \in \mathcal{A}_t, \delta \in \mathcal{A}_s \subset \mathcal{A}_t$); moreover, the flow property holds with the same control α for two different initial times $s \geq t$, but α may only belong to $\mathcal{A}_t$ and not to $\mathcal{A}_s$.

Notice that $\mathcal{A}_t \subseteq \mathcal{A}$ for any $t \geq 0$ and $\mathcal{A}_0 = \mathcal{A}$. Define

$$V_t(t, \xi) := \sup_{\alpha \in \mathcal{A}_t} J(t, \xi, \alpha). \quad (1.13)$$

It is thus clear that $V_t(t, \xi) \leq V(t, \xi)$ and $V_0(0, \xi) = V(0, \xi)$. For this value function we can easily prove the following result.

Proposition 1.3.1. Suppose that assumptions $(\mathbf{H}_{b,\sigma})$ and $(\mathbf{H}_{f,\beta})$ or $(\mathbf{H}'_{f,\beta})$ are satisfied. Let $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$. Then,

$$V(0, \xi) = V_t(t, \xi) \quad \forall t \geq 0. \quad (1.14)$$

The proof relies on the following lemma.

Lemma 1.3.2. Consider two solutions of the same controlled stochastic equation of the McKean-Vlasov type (1.2) starting at different initial times from the same initial condition and under different controls. More precisely, we denote by $X^{(0)} = (X_s^{(0)})_{s \geq 0}$ the unique solution to equation

$$\begin{cases} dX_s^{(0)} = b(X_s^{(0)}, \mathbb{P}_{X_s^{(0)}}, \alpha_s)ds + \sigma(X_s^{(0)}, \mathbb{P}_{X_s^{(0)}}, \alpha_s)dB_s & \forall s \geq 0 \\ X_0^{(0)} = \xi \end{cases} \quad (1.15)$$

and by $X^{(t)} = (X_s^{(t)})_{s \geq t}$ the unique solution to

$$\begin{cases} dX_s^{(t)} = b(X_s^{(t)}, \mathbb{P}_{X_s^{(t)}}, \gamma_s)ds + \sigma(X_s^{(t)}, \mathbb{P}_{X_s^{(t)}}, \gamma_s)dB_s & \forall s \geq t \\ X_t^{(t)} = \xi, \end{cases}$$

where the coefficients satisfy Assumptions $(\mathbf{H}_{b,\sigma})$ and $\alpha \in \mathcal{A}_0$, $\gamma \in \mathcal{A}_t$. Assume furthermore that

$$\mathcal{L}(\xi, (B_s)_{s \geq 0}, (\alpha_s)_{s \geq 0}) = \mathcal{L}(\xi, (B_{s+t}^{(t)})_{s \geq 0}, (\gamma_{s+t})_{s \geq 0}). \quad (1.16)$$

Then,

$$\mathcal{L}((X_s^{(0)})_{s \geq 0}, (\alpha_s)_{s \geq 0}) = \mathcal{L}((X_{s+t}^{(t)})_{s \geq 0}, (\gamma_{s+t})_{s \geq 0}).$$

The proof of this Lemma is based on a deterministic change of time in the stochastic integral. The complete proof is provided in Appendix B.1.

Proof of Proposition 1.3.1. Intuitively, the proof relies on the idea to “shift” the controls forward and backward in time using $B^{(t)}$ instead of B (and vice versa) in expression (1.1), leading to a bijection from $\mathcal{A}$ to $\mathcal{A}_t$.

Fix a $\mathbb{F}$ -predictable control $\alpha_s = \underline{\alpha}(s, B_{\cdot \wedge s}, U)$, $s \geq 0$. Choose $t \geq 0$. Consider a measurable function $\bar{\alpha}^t$ such that

$$\bar{\alpha}^t(z, w', u) = \underline{\alpha}(z - t, w'_{\cdot \wedge z - t}, u) \quad \text{for } z \in [t, +\infty), w' \in \mathcal{C}([0, +\infty); \mathbb{R}^m), u \in [0, 1],$$

and construct the $\mathbb{F}^t$ -predictable control $\bar{\alpha} := \bar{\alpha}^t(\cdot, (B_{s+t}^{(t)})_{s \geq 0}, U)$. Therefore, we can denote by $\bar{X} = X^{t, \xi, \bar{\alpha}}$ the unique solution to the state equation with initial time t and initial state ξ under the feasible control $\bar{\alpha}$, i.e. equation

$$\begin{cases} d\bar{X}_s = b(\bar{X}_s, \mathbb{P}_{\bar{X}_s}, \bar{\alpha}^t(s, B_{\cdot \wedge s+t}^{(t)}, U))ds \\ \quad + \sigma(\bar{X}_s, \mathbb{P}_{\bar{X}_s}, \bar{\alpha}^t(s, B_{\cdot \wedge s+t}^{(t)}, U))dB_s \\ \bar{X}_t = \xi. \end{cases} \quad (1.17)$$

Notice that, by construction, $\mathcal{L}((B_s)_{s \geq 0}, U, (\alpha_s)_{s \geq 0}) = \mathcal{L}((B_{t+s}^{(t)})_{s \geq 0}, U, (\bar{\alpha}_{s+t})_{s \geq 0})$, which implies $\mathcal{L}((B_s)_{s \geq 0}, \xi, (\alpha_s)_{s \geq 0}) = \mathcal{L}((B_{t+s}^{(t)})_{s \geq 0}, \xi, (\bar{\alpha}_{s+t})_{s \geq 0})$ for a $\sigma(U)$ -measurable ξ . As a consequence, applying Lemma 1.3.2, we can affirm that

$$\mathcal{L}((X_s^{0, \xi, \alpha})_{s \geq 0}, (\alpha_s)_{s \geq 0}) = \mathcal{L}((\bar{X}_{s+t})_{s \geq 0}, (\bar{\alpha}_{s+t})_{s \geq 0}).$$

Computing the gain, by the previous equality in law and a simple change of variable $z = s + t$,

$$J(0, \xi, \alpha) = \mathbb{E} \left[\int_t^{+\infty} e^{-\beta(z-t)} f(\bar{X}_z, \mathbb{P}_{\bar{X}_z}, \bar{\alpha}_z) dz \right] = J(t, \xi, \bar{\alpha}) \leq V_t(t, \xi).$$

Taking the supremum over all feasible controls, we conclude that $V(0, \xi) \leq V_t(t, \xi)$.

We now want to prove the converse inequality. Fix $\gamma \in \mathcal{A}_t$. By Lemma 1.2.3 on $[t, +\infty)$, γ is indistinguishable from $(\underline{\gamma}(s, B_{\cdot \wedge s}^{(t)}, U))_{s \geq t}$ for a suitable measurable map $\underline{\gamma} : [t, +\infty) \times \mathcal{C}([t, +\infty)) \times [0, 1]$. Defining a control $\alpha \in \mathcal{A}$ setting $\alpha_s := \underline{\gamma}(s+t, B_{\cdot \wedge s}, U)$ for all $s \geq 0$, similar arguments lead to

$$J(t, \xi, \gamma) = J(0, \xi, \alpha) \leq V(0, \xi).$$

Taking the supremum over all $\gamma \in \mathcal{A}_t$, we conclude that $V_t(t, \xi) \leq V(0, \xi)$. $\square$

1.3.2 A counterexample to the time invariance property

Contrary to the case of classical control theory, time invariance for the value function $V(t, \xi)$ of a McKean-Vlasov control problem may fail. Thanks to Proposition 1.3.1, it is enough to exhibit an example where $V(t, \xi) \neq V_t(t, \xi)$ for some $t > 0$. We provide here an example of a finite horizon optimal control problem of the McKean-Vlasov type where restricting the class of controls from $\mathcal{A}$ to $\mathcal{A}_t$ worsens the optimal value.

Example 1.3.3. Fix a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which a scalar Brownian motion B is defined. Consider a deterministic initial condition $\xi = 0$. Fix the dimension $d = 1$, the time horizon $T = 1$ and the action space $A = [-1, 1]$. Choose an initial time $t \in (0, \frac{1}{2})$. Consider the following state equation:

$$dX_s = \alpha_s ds, \quad s \geq t, \quad X_t = 0. \quad (1.18)$$

Remark 1.3.4. Except for the adaptedness condition on α , equation (1.18) looks like a deterministic controlled equation. Nonetheless, we may regard it as a stochastic system as follows. Choose R large enough to guarantee that $\sup_{s \in [t, T]} |X_s^{t, 0, \alpha}| \leq 1 - t < R$

and take $b : \mathbb{R} \rightarrow \mathbb{R}$ and $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ Lipschitz with $b(x) = \sigma(x) = 0$ for $|x| < R$. The trajectories of (1.18) thus coincide with the trajectories of the stochastic controlled system

$$dX_s = (\alpha_s + b(X_s))ds + \sigma(X_s)dB_s, \quad s \geq t, \quad X_t = 0.$$

□

Define the following cost functional of the McKean-Vlasov type, which we want to minimize over a suitable class of controls :

$$J(t, 0, \alpha) := H(\mathbb{P}_{X_1^{t, 0, \alpha}} | \hat{\nu}), \quad (1.19)$$

where $\hat{\nu} = \frac{1}{2}\delta_{\{1-t\}} + \frac{1}{2}\delta_{\{t-1\}}$. The function H is the modified relative entropy of a measure with respect to the target measure $\hat{\nu}$, defined as in Definition 15.1 in [1] and in Definition 9.4.1 in [2]:

$$H(\mu | \nu) := \begin{cases} Q & \text{if } \mu \not\ll \nu \\ \int_{\mathbb{R}^d} h\left(\frac{d\mu}{d\nu}(x)\right) \nu(dx) & \text{if } \mu \ll \nu \end{cases} \quad (1.20)$$

where Q is a sufficiently large real value, $h : [0, +\infty) \rightarrow [0, +\infty)$ is a continuous, strictly convex, non-negative function having a unique minimum at $s = 1$ (e.g. $h(y) = y \log(y) + (1 - y)$ with $h(0) = 1$) and $\frac{d\mu}{d\nu}$ denotes the density of μ with respect to ν when it exists. Notice that H is constantly equal to Q on the set of probability measures which are not absolutely continuous with respect to $\hat{\nu}$. As a consequence, we look for a minimum of this function over the complementary set. If $\mu \ll \hat{\nu}$, then for any $I \in \mathcal{B}(\mathbb{R})$ we have

$$\mu(I) = \int_I \frac{d\mu}{d\hat{\nu}}(x) \hat{\nu}(dx)$$

$$\begin{aligned}
&= \int_I \frac{d\mu}{d\hat{\nu}}(x) \left(\frac{1}{2} \delta_{\{1-t\}}(dx) + \frac{1}{2} \delta_{\{t-1\}}(dx) \right) \\
&= \frac{1}{2} \int_I \frac{d\mu}{d\hat{\nu}}(x) \delta_{\{1-t\}}(dx) + \frac{1}{2} \int_I \frac{d\mu}{d\hat{\nu}}(x) \delta_{\{t-1\}}(dx) \\
&= \frac{1}{2} \left(\frac{d\mu}{d\hat{\nu}}(1-t) \mathbb{1}_I(1-t) + \frac{d\mu}{d\hat{\nu}}(t-1) \mathbb{1}_I(t-1) \right). \tag{1.21}
\end{aligned}$$

Assume that $\text{supp}(\mu) \neq \{1-t, t-1\}$. Then, there exists a Borel set $I \in \mathcal{B}(\mathbb{R})$ such that $I \cap \{1-t, t-1\} = \emptyset$ and $\mu(I) > 0$. By equation (1.21),

$$0 < \mu(I) = \frac{1}{2} \left(\frac{d\mu}{d\hat{\nu}}(1-t) \mathbb{1}_I(1-t) + \frac{d\mu}{d\hat{\nu}}(t-1) \mathbb{1}_I(t-1) \right) = 0,$$

which leads to a contradiction. As a consequence, we can say that, if $\mu \ll \hat{\nu}$, then $\text{supp}(\mu) \subseteq \{1-t, t-1\}$, i.e.

$$\mu = \lambda \delta_{\{1-t\}} + (1-\lambda) \delta_{\{t-1\}} \tag{1.22}$$

for some $\lambda \in [0, 1]$. If $\mu \ll \hat{\nu}$, with μ of the form (1.22), then by (1.21) we deduce that, for any $I \in \mathcal{B}(\mathbb{R})$,

$$\mu(I) = \frac{1}{2} \left(\frac{d\mu}{d\hat{\nu}}(1-t) \mathbb{1}_I(1-t) + \frac{d\mu}{d\hat{\nu}}(t-1) \mathbb{1}_I(t-1) \right);$$

on the other hand,

$$\mu(I) = (\lambda \delta_{\{1-t\}} + (1-\lambda) \delta_{\{t-1\}})(I) = \lambda \mathbb{1}_I(1-t) + (1-\lambda) \mathbb{1}_I(t-1).$$

Comparing the two expressions, we deduce that

$$\frac{d\mu}{d\hat{\nu}}(y) = 2\lambda \mathbb{1}_{\{1-t\}}(y) + 2(1-\lambda) \mathbb{1}_{\{t-1\}}(y). \tag{1.23}$$

We conclude that, when $\mu \ll \hat{\nu}$, then μ is supported in $\{1-t, t-1\}$ with density

$$\frac{d\mu}{d\hat{\nu}}(y) = 2\lambda \mathbb{1}_{\{1-t\}}(y) + 2(1-\lambda) \mathbb{1}_{\{t-1\}}(y)$$

for some $\lambda \in [0, 1]$. Hence, notice that H has a unique strict minimum at $\mu = \hat{\nu}$. As a matter of fact, by the convexity of h ,

$$\begin{aligned}
H(\mu|\hat{\nu}) &= \int_{\mathbb{R}} h\left(\frac{d\mu}{d\hat{\nu}}(y)\right) \hat{\nu}(dy) \quad \text{by (1.23)} \\
&= \int_{\mathbb{R}} h\left(2\lambda \mathbb{1}_{\{1-t\}}(y) + 2(1-\lambda) \mathbb{1}_{\{t-1\}}(y)\right) \left(\frac{1}{2} \delta_{\{1-t\}} + \frac{1}{2} \delta_{\{t-1\}}\right)(dy) \\
&= \frac{1}{2} h\left(2\lambda \mathbb{1}_{\{1-t\}}(1-t) + 2(1-\lambda) \mathbb{1}_{\{t-1\}}(1-t)\right) \\
&\quad + \frac{1}{2} h\left(2\lambda \mathbb{1}_{\{1-t\}}(t-1) + 2(1-\lambda) \mathbb{1}_{\{t-1\}}(t-1)\right)
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2}h(2\lambda) + \frac{1}{2}h(2(1-\lambda)) \\
&= \frac{1}{2}[h(2\lambda) + h(2(1-\lambda))] \\
&\geq h\left(\frac{1}{2}(2\lambda) + \frac{1}{2}(2(1-\lambda))\right) = h(1) = H(\hat{\nu}|\hat{\nu}), \quad \lambda \in [0, 1]. \tag{1.24}
\end{aligned}$$

As h is strictly convex, the inequality is strict for any $\lambda \neq \frac{1}{2}$, proving that the global minimum of H is reached only at $\hat{\nu}$.

1) We first minimize the cost functional over the class $\mathcal{A}_t$.

Call α_1 and α_2 the controls which are identically equal to 1 and -1 , respectively. As $T = 1$, $X_1^{t,0,\alpha} = \int_t^1 \alpha_s ds$ with $\alpha_s \in [-1, 1]$ for every $s \in [0, 1]$. Then, $X_1^{t,0,\alpha_1} = 1 - t$, $X_1^{t,0,\alpha_2} = -1 + t$ and the support of $\mathbb{P}_{X_1^{t,0,\alpha}}$ does not coincide with $\{t-1, 1-t\}$ for any other $\alpha \in \mathcal{A}_t$. As a matter of fact, assume that

$$\mathbb{P}(X_1^{t,0,\alpha} \in \{t-1, 1-t\}) = 1 \quad \text{for some } \alpha \in \mathcal{A}_t.$$

Then,

$$\begin{aligned}
1 &= \mathbb{P}(X_1^{t,0,\alpha} \in \{t-1, 1-t\}) \\
&= \mathbb{P}\left(\int_t^1 \alpha_s ds \in \{t-1, 1-t\}\right) \\
&= \mathbb{P}\left(\int_t^1 \alpha_s ds = t-1\right) + \mathbb{P}\left(\int_t^1 \alpha_s ds = 1-t\right). \tag{1.25}
\end{aligned}$$

Call $A_1 := \{\omega \in \Omega : \int_t^1 \alpha_s(\omega) ds = t-1\}$ and $A_2 := \{\omega \in \Omega : \int_t^1 \alpha_s(\omega) ds = 1-t\}$. Notice that

$$\begin{aligned}
\omega \in A_1 &\iff \int_t^1 \alpha_s(\omega) ds = 1-t \\
&\iff \int_t^1 (\alpha_s(\omega) - 1) ds = 0 \quad \text{with } (\alpha_s(\omega) - 1) \leq 0 \text{ for all } s \geq 0 \\
&\iff \alpha_s(\omega) - 1 = 0 \quad \lambda_{[t,1]} - \text{a.e.},
\end{aligned}$$

where $\lambda_{[t,1]}$ is the Lebesgue measure of the interval $[t, 1]$, i.e. $\alpha_s = (\alpha_1)_s$ for $\lambda_{[t,1]}$ -a.e. $s \in [t, 1]$. Analogously,

$$\omega \in A_2 \iff \alpha_s(\omega) + 1 = 0 \quad \lambda_{[t,1]} - \text{a.e.},$$

i.e. $\alpha = \alpha_2$ $\lambda_{[t,1]}$ -a.e. We want to prove that either $\mathbb{P}(A_1) = 1$ and $\mathbb{P}(A_2) = 0$ or $\mathbb{P}(A_1) = 0$ and $\mathbb{P}(A_2) = 1$. Let $\omega \in A_1$. Then, $\alpha(s, \omega) = 1$ for $\lambda_{[t,1]}$ -a.e. $s \in [t, 1]$. Consider the process $\bar{\alpha} : [t, 1] \rightarrow A$ such that $\bar{\alpha}_s = \int_t^s \alpha_r dr$. Since α is $\mathbb{F}^{B,t}$ -predictable, $\bar{\alpha}$ is $\mathbb{F}^{B,t}$ -adapted; moreover, it is continuous. We conclude that $\bar{\alpha}$ is $\mathbb{F}^{B,t}$ -predictable. Now, repeating the reasoning above we can easily deduce that

$$\omega \in A_1 \iff \alpha_s(\omega) = 1 \text{ for } \lambda_{[t,1]} - \text{a.e. } s \in [t, 1]$$

$$\iff \bar{\alpha}_s(\omega) = 1 \text{ for all } s \in (t, 1].$$

We conclude that

$$\begin{aligned} A_1 &= \{\omega \in \Omega : \bar{\alpha}_s(\omega) = 1 \text{ for all } s \in (t, 1]\} \\ &= \cap_{t < s \leq 1} \{\omega \in \Omega : \bar{\alpha}_s(\omega) = 1\} \in \cap_{t < s \leq 1} \mathcal{F}_s^{B,t} \end{aligned}$$

as $\bar{\alpha}$ is $\mathbb{F}^{B,t}$ -predictable. Now, $\mathbb{F}^{B,t}$ is the (completed) Brownian filtration from time t , which is right continuous; therefore,

$$A_1 \in \cap_{t < s \leq 1} \mathcal{F}_s^{B,t} = \mathcal{F}_t^{B,t},$$

corresponding to the trivial σ -algebra. We thus know that either $A_1 = \emptyset$ or $A_1 = \Omega$, i.e. $\mathbb{P}(A_1) = 0$ (which, by (1.25), implies that $\mathbb{P}(A_2) = 1$) or $\mathbb{P}(A_1) = 1$ (which, by (1.25), implies that $\mathbb{P}(A_2) = 0$). In conclusion, $\mathbb{P}(X_1^{t,0,\alpha} \in \{t-1, 1-t\}) = 1$ if and only if $\mathbb{P}$ -a.s. $\alpha = \alpha_1$ $\lambda_{[t,1]}$ -a.e., leading to the same value of α_1 , or $\mathbb{P}$ -a.s. $\alpha = \alpha_2$ $\lambda_{[t,1]}$ -a.e., leading to the same value of α_2 . In other words, α_1 and α_2 are the only existing optimal controls, up to modifications.

Therefore, $\mathbb{P}_{X_1^{t,0,\alpha_1}} = \delta_{\{1-t\}}$, $\mathbb{P}_{X_1^{t,0,\alpha_2}} = \delta_{\{t-1\}}$ and we can prove that for any other $\alpha \in \mathcal{A}_t$ the support of $\mathbb{P}_{X_1^{0,\xi,\alpha}}$ does not coincide with $\{1-t, t-1\}$. This implies that both α_1 and α_2 are optimal with value $(h(0) + h(2))/2$ by (1.24).

2) We now minimize the cost functional over the class $\mathcal{A}$.

Consider the control α^* defined as follows:

$$\alpha_s^*(\omega) := \begin{cases} 1 & \text{if } B_{\frac{t}{2}}(\omega) \geq 0 \\ -1 & \text{if } B_{\frac{t}{2}}(\omega) < 0 \end{cases} \quad (1.26)$$

for every $s \geq t, \omega \in \Omega$. Clearly, $\alpha^* \in \mathcal{A} \setminus \mathcal{A}_t$ because it depends on the source of randomness $B_{\frac{t}{2}}$, independent of $B^{(t)}$. Moreover, we can rewrite it as $\alpha^* = \underline{\alpha}^*(B_{\cdot \wedge t})$, where $\underline{\alpha}^*(w) := \alpha_1$ if $w_{\frac{t}{2}} \geq 0$ and $\underline{\alpha}^*(w) := \alpha_2$ if $w_{\frac{t}{2}} < 0$. As a consequence, for any $\phi \in \mathcal{C}_b(\mathbb{R}^d; \mathbb{R})$,

$$\begin{aligned} \int_{\mathbb{R}^d} \phi(y) \mathbb{P}_{X_1^{t,0,\alpha^*}}(dy) &= \mathbb{E}[\phi(X_1^{t,0,\alpha^*})] \\ &= \mathbb{P}(\underline{\alpha}^*(B_{\cdot \wedge t}) = \alpha_1) \mathbb{E}[\phi(X_1^{t,0,\alpha_1})] + \mathbb{P}(\underline{\alpha}^*(B_{\cdot \wedge t}) = \alpha_2) \mathbb{E}[\phi(X_1^{t,0,\alpha_2})] \end{aligned}$$

by conditional independence, from which we deduce that

$$\mathbb{P}_{X_1^{t,0,\alpha^*}} = \lambda \mathbb{P}_{X_1^{t,0,\alpha_1}} + (1-\lambda) \mathbb{P}_{X_1^{t,0,\alpha_2}} = \lambda \delta_{\{1-t\}} + (1-\lambda) \delta_{\{t-1\}}, \quad (1.27)$$

where $\lambda = \mathbb{P}(\underline{\alpha}^*(B_{\cdot \wedge t}) = \alpha_1) = \mathbb{P}(B_{\frac{t}{2}} \geq 0) = \frac{1}{2}$. Therefore $\mathbb{P}_{X_1^{t,0,\alpha^*}} = \hat{\nu}$ and from (1.24) we deduce that $H(\mathbb{P}_{X_1^{t,0,\alpha^*}} | \hat{\nu}) = h(1)$, which is the unique global minimum of H .

By **1)** and **2)** we conclude that

$$\inf_{\alpha \in \mathcal{A}_t} H(\mathbb{P}_{X_T^{t,0,\alpha}} | \hat{\nu}) = \frac{h(2) + h(0)}{2} > h(1) = \inf_{\alpha \in \mathcal{A}} H(\mathbb{P}_{X_T^{t,0,\alpha}} | \hat{\nu}) :$$

the value at time t is strictly larger if we optimize over the smaller class of controls $\mathcal{A}_t$.

Remark 1.3.5. The conclusion above follows from the strict convexity of h , used in (1.24), i.e. from the non-linear dependence of the gain functional on the terminal law of the state process. Therefore, it is not possible to construct an analogous example in classical control theory, where the dependence of the terminal gain upon the law is always linear. $\square$

Notice that in the previous example H is not continuous. We present a second example, where the continuity of the gain functional guarantees that the randomization procedure (1.26) does not improve the value of the problem.

Example 1.3.6. Consider the same dynamics of the previous example, described by equation (1.18), with $d = 1$, $T = 1$, $A = [-1, 1]$ and $t \in (0, \frac{1}{2})$. Consider the following optimal control problem of the McKean-Vlasov type: we want to maximize the gain functional

$$F(\mathbb{P}_{X_1^{t,0,\alpha}}) \quad (1.28)$$

over the classes of controls $\mathcal{A}_t$ and $\mathcal{A}$, respectively. The functional F is the variance of its argument, i.e.

$$F(\mu) := \int_{\mathbb{R}} y^2 \mu(dy) - \left(\int_{\mathbb{R}} y \mu(dy) \right)^2$$

for any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Notice that, as $A = [-1, 1]$, $X_1^{t,0,\alpha} \in [-1+t, t-1]$, i.e. $\text{supp}(\mathbb{P}_{X_1^{t,0,\alpha}}) = [-1+t, t-1]$. Notice that, for any fixed $x \in \mathbb{R}$,

$$\arg \max_{\substack{\mu \in \mathcal{P}_2(\mathbb{R}^d): \\ \text{supp}(\mu) = [-x, x]}} F(\mu) = \hat{\mu}_x := \frac{1}{2} \delta_{\{x\}} + \frac{1}{2} \delta_{\{-x\}}.$$

As a matter of fact,

$$F(\hat{\mu}_x) = \frac{1}{2}x^2 + \frac{1}{2}x^2 - \left(\frac{1}{2}x - \frac{1}{2}x \right)^2 = x^2.$$

If we assume that there exists a measure $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ with support $[-x, x]$ such that $F(\mu) \geq x^2$, then

$$\int_{\mathbb{R}} y^2 \mu(dy) - \left(\int_{\mathbb{R}} y \mu(dy) \right)^2 \geq x^2,$$

from which we deduce that

$$\int_{\mathbb{R}} y^2 \mu(dy) \geq x^2 + \left(\int_{\mathbb{R}} y \mu(dy) \right)^2 \geq x^2.$$

As a consequence,

$$\int_{\mathbb{R}} (y^2 - x^2) \mu(dy) \geq 0$$

with a nonpositive integrand; we conclude that

$$y^2 = x^2 \quad \mu - a.s.,$$

i.e.

$$\mu(y) := \lambda \delta_x(y) + (1 - \lambda) \delta_{-x}(y).$$

Then,

$$\int_{\mathbb{R}} y^2 \mu(dy) = x^2$$

and we must have

$$\int_{\mathbb{R}} y \mu(dy) = 0$$

(if not, $F(\mu) < x^2$ and we would contradict the assumption). But

$$\int_{\mathbb{R}} y \mu(dy) = (2\lambda - 1)x,$$

which is null if and only if $\lambda = \frac{1}{2}$. We conclude that $F(\mu) \geq x^2$ if and only if $\mu = \hat{\mu}_x$, which thus realizes the maximum of the variance over all the probability measures of order 2 valued in an interval $[-x, x]$.

Under the dynamics we are considering, $x = 1 - t$.

1) We optimize over the larger class $\mathcal{A}$. Take the same control that we considered in the previous example, i.e.

$$\alpha_s^*(\omega) := \begin{cases} 1 & \text{if } B_{\frac{t}{2}}(\omega) \geq 0 \\ -1 & \text{if } B_{\frac{t}{2}}(\omega) < 0, \end{cases}$$

belonging to $\mathcal{A}$ and not to $\mathcal{A}_t$. Then, by (1.27), we know that

$$\mathbb{P}_{X_1^{t,0,\alpha^*}} = \arg \max_{\substack{\mu \in \mathcal{P}_2(\mathbb{R}^d): \\ \text{supp}(\mu) = [-1+t, 1-t]}} F(\mu),$$

which clearly implies that

$$\alpha^* = \arg \max_{\alpha \in \mathcal{A}} F(\mathbb{P}_{X_1^{t,0,\alpha}})$$

with associated value $F(\mathbb{P}_{X_1^{t,0,\alpha^*}}) = (1 - t)^2$.

2) We now optimize over the smaller class $\mathcal{A}_t$. As already noticed in the previous example, for every $\alpha \in \mathcal{A}_t$, $\mathbb{P}_{X_1^{t,0,\alpha}} \neq \frac{1}{2}\delta_{1-t} + \frac{1}{2}\delta_{t-1}$. On the other hand, notice that for any $\epsilon > 0$ we can construct a control $\alpha^\epsilon \in \mathcal{A}_t$,

$$\alpha_s^\epsilon := \begin{cases} 0 & \text{for } s \in [t, t + \frac{\epsilon}{2}] \\ \alpha_s^1 = 1 & \text{for } s \in (t + \frac{\epsilon}{2}, T) \text{ if } B_{t+\frac{\epsilon}{4}} \geq 0 \\ \alpha_s^2 = -1 & \text{for } s \in (t + \frac{\epsilon}{2}, T) \text{ if } B_{t+\frac{\epsilon}{4}} < 0 \end{cases}$$

such that

$$F(\mathbb{P}_{X^{t,0,\alpha^\epsilon}}) + \epsilon \geq F(\hat{\mu}) = \sup_{\substack{\mu \in \mathcal{P}_2(\mathbb{R}^d): \\ \text{supp}(\mu) = [-1+t, 1-t]}} F(\mu). \quad (1.29)$$

As a matter of fact,

$$\begin{aligned}
X_1^{t,0,\alpha^\epsilon}(\omega) &= \int_t^1 \alpha_s^\epsilon(\omega) ds \\
&= \begin{cases} \int_{t+\frac{\epsilon}{2}}^1 1 ds & \text{if } B_{t+\frac{\epsilon}{4}}(\omega) \geq 0 \\ -\int_{t+\frac{\epsilon}{2}}^1 1 ds & \text{if } B_{t+\frac{\epsilon}{4}}(\omega) < 0 \end{cases} \\
&= \left(1 - t - \frac{\epsilon}{2}\right) \mathbb{1}_{\{B_{t+\frac{\epsilon}{4}} \geq 0\}}(\omega) + \left(t + \frac{\epsilon}{2} - 1\right) \mathbb{1}_{\{B_{t+\frac{\epsilon}{4}} < 0\}}(\omega),
\end{aligned}$$

which implies that, for any $I \in \mathcal{B}(\mathbb{R})$,

$$\begin{aligned}
\mathbb{P}_{X_1^{t,0,\alpha^\epsilon}}(I) &= \begin{cases} 1 & \text{if } \{1 - t - \frac{\epsilon}{2}, t + \frac{\epsilon}{2} - 1\} \subseteq I \\ \mathbb{P}(B_{t+\frac{\epsilon}{4}} \geq 0) & \text{if } \{1 - t - \frac{\epsilon}{2}, t + \frac{\epsilon}{2} - 1\} \cap I = \{1 - t - \frac{\epsilon}{2}\} \\ \mathbb{P}(B_{t+\frac{\epsilon}{4}} < 0) & \text{if } \{1 - t - \frac{\epsilon}{2}, t + \frac{\epsilon}{2} - 1\} \cap I = \{-1 + t + \frac{\epsilon}{2}\} \\ 0 & \text{otherwise} \end{cases} \\
&= \begin{cases} 1 & \text{if } \{1 - t - \frac{\epsilon}{2}, t + \frac{\epsilon}{2} - 1\} \subseteq I \\ \frac{1}{2} & \text{if } \{1 - t - \frac{\epsilon}{2}, t + \frac{\epsilon}{2} - 1\} \cap I = \{1 - t - \frac{\epsilon}{2}\} \\ \frac{1}{2} & \text{if } \{1 - t - \frac{\epsilon}{2}, t + \frac{\epsilon}{2} - 1\} \cap I = \{-1 + t + \frac{\epsilon}{2}\} \\ 0 & \text{otherwise,} \end{cases}
\end{aligned}$$

i.e. $\mathbb{P}_{X_1^{t,0,\alpha^\epsilon}} = \frac{1}{2}\delta_{\{1-t-\frac{\epsilon}{2}\}} + \frac{1}{2}\delta_{\{t+\frac{\epsilon}{2}-1\}}$. As a consequence,

$$\begin{aligned}
F(\mathbb{P}_{X_1^{t,0,\alpha^\epsilon}}) &= \int_{\mathbb{R}} y^2 \mathbb{P}_{X_1^{t,0,\alpha^\epsilon}}(dy) - \left(\int_{\mathbb{R}} y \mathbb{P}_{X_1^{t,0,\alpha^\epsilon}}(dy) \right)^2 \\
&= \frac{1}{2} \int_{\mathbb{R}} y^2 \delta_{\{1-t-\frac{\epsilon}{2}\}}(dy) + \frac{1}{2} \int_{\mathbb{R}} y^2 \delta_{\{t+\frac{\epsilon}{2}-1\}}(dy) \\
&\quad - \left(\frac{1}{2} \int_{\mathbb{R}} y \delta_{\{1-t-\frac{\epsilon}{2}\}}(dy) + \frac{1}{2} \int_{\mathbb{R}} y \delta_{\{t+\frac{\epsilon}{2}-1\}}(dy) \right)^2 \\
&= \frac{1}{2} \left[\left(1 - t - \frac{\epsilon}{2}\right)^2 + \left(t + \frac{\epsilon}{2} - 1\right)^2 \right] - \frac{1}{4} \left(1 - t - \frac{\epsilon}{2} + t + \frac{\epsilon}{2} - 1\right)^2 \\
&= \left(1 - t - \frac{\epsilon}{2}\right)^2.
\end{aligned}$$

Thus,

$$\begin{aligned}
F(\hat{\mu}) - \epsilon &= (1 - t)^2 - \epsilon \\
&< (1 - t)^2 - \epsilon + \frac{\epsilon^2}{4} + \epsilon t \\
&= \left(1 - t - \frac{\epsilon}{2}\right)^2 \\
&= F(\mathbb{P}_{X_1^{t,0,\alpha^\epsilon}})
\end{aligned}$$

and property (1.29) holds. Then, by definition of supremum, we conclude that

$$\sup_{\alpha \in \mathcal{A}_t} F(\mathbb{P}_{X_1^{t,0,\alpha}}) = (1 - t)^2 = \sup_{\substack{\mu \in \mathcal{P}_2(\mathbb{R}^d): \\ \text{supp}(\mu) = [-1+t, 1-t]}} F(\mu) = \sup_{\alpha \in \mathcal{A}} F(\mathbb{P}_{X_1^{t,0,\alpha}}). \quad (1.30)$$

Remark 1.3.7. In this example, the main difference between the two formulations of the problem does not concern the value of the supremum of the gain functional, but the existence of optimal controls: in the larger class of feasible controls $\mathcal{A}$, there exists an optimal control α^* realizing the maximum of $F(\mathbb{P}_{X_1^{t,0,\alpha^*}})$, while in the smaller class $\mathcal{A}_t$ we can only construct ϵ -optimal controls α^ϵ . Notice that the main result of this section focuses only on the value of the problem, and not on existence of optimal controls. $\square$

1.3.3 Time invariance of the value function

The previous examples show that appropriate conditions are required to guarantee the equality $V(t, \xi) = V_t(t, \xi)$ and hence the time invariance property. We will prove the following:

Theorem 1.3.8. Assume that $(\mathbf{H}_{b,\sigma})$, $(\mathbf{H}_{f,\beta})$ (or $(\mathbf{H}'_{f,\beta})$) and $(\mathbf{H}_{f_{cont}})$ are satisfied. Then, for every choice of $t \geq 0$ and $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$, it holds that

$$V_t(t, \xi) = \sup_{\alpha \in \mathcal{A}_t} J(t, \xi, \alpha) = \sup_{\alpha \in \mathcal{A}} J(t, \xi, \alpha) = V(t, \xi). \quad (1.31)$$

As suggested by Example 1.3.6, the continuity assumption $(\mathbf{H}_{f_{cont}})$, which was missing in Example 1.3.3, plays a crucial role here. In the proof of Theorem 1.3.8, it guarantees the continuity of the gain functional with respect to time, but it also qualifies as a key technical assumption for the rest of the chapter (especially for the law invariance property and the continuity of the value function). In order to prove Theorem 1.3.8, we need to introduce the notations and technical results contained in the following subsection.

Disintegration of McKean-Vlasov systems

We first recall the definition of concatenation of paths.

Definition 1.3.9. Fix a time instant $t \geq 0$. Let w_1, w_2 be two $\mathbb{R}^m$ -valued functions with domain $[0, +\infty)$ and $[t, +\infty)$, respectively. Define the concatenation of w_1 and w_2 as the function $w_1 \otimes_t w_2$ such that

$$(w_1 \otimes_t w_2)(s) = \begin{cases} w_1(s) & \text{if } 0 \leq s \leq t \\ w_1(t) + w_2(s) - w_2(t) & \text{if } s \geq t. \end{cases}$$

In particular, if w_1 and w_2 are Brownian trajectories, we deduce that, for every $t \geq 0$,

$$(B \otimes_t B)(\omega) = (B \otimes_t B^{(t)})(\omega) = B(\omega).$$

As a consequence, by Lemma 1.2.3, every $\mathbb{F}$ -predictable process α valued in A is indistinguishable both from $(\underline{\alpha}(s, (B \otimes_t B)_{\cdot \wedge s}, U))_{s \geq 0}$ and from $(\underline{\alpha}(s, (B \otimes_t B^{(t)})_{\cdot \wedge s}, U))_{s \geq 0}$ for a suitable measurable function $\underline{\alpha} : [0, +\infty) \times \mathcal{C}([0, +\infty); \mathbb{R}^m) \times [0, 1] \rightarrow A$, for any choice of $t \geq 0$.

We now want to state a “disintegration property” of the McKean-Vlasov equation with respect to the first part of the Brownian trajectory. Assume that hypotheses $(\mathbf{H}_{b,\sigma})$ and $(\mathbf{H}_{f,\beta})$ (or $(\mathbf{H}'_{f,\beta})$) are satisfied. Fix an initial time $r > 0$ and a control $\alpha = (\underline{\alpha}(s, B_{\cdot \wedge (r \wedge s)} \otimes_r B_{\cdot \wedge s}^{(r)}, U))_{s \geq 0} \in \mathcal{A}$. Call $X := X^{r,\xi,\alpha}$ the unique solution to the McKean-Vlasov SDE (1.2) starting from ξ at time r under the control α , i.e.

$$dX_s = b(X_s, \mathbb{P}_{X_s}, \alpha_s)ds + \sigma(X_s, \mathbb{P}_{X_s}, \alpha_s)dB_s, \quad s \geq r, \quad X_r = \xi. \quad (1.32)$$

Now associate to each path $w \in \mathcal{C}([0, r])$ a control $\alpha^r(w) \in \mathcal{A}_r$ defined as follows:

$$\alpha_s^r(w) := \underline{\alpha}(s, w \otimes_r B_{\cdot \wedge s}^{(r)}, U) \quad \text{for all } s \geq r. \quad (1.33)$$

Consider the following system of SDEs:

$$\begin{aligned} dX_s^w = & b\left(X_s^w, \int_{\mathcal{C}([0, r])} \mathbb{P}_{X_s^{w'}} W_{[0, r]}(dw'), \alpha_s^r(w)\right) ds \\ & + \sigma\left(X_s^w, \int_{\mathcal{C}([0, r])} \mathbb{P}_{X_s^{w'}} W_{[0, r]}(dw'), \alpha_s^r(w)\right) dB_s, \quad s \geq r, \quad X_r^w = \xi \end{aligned} \quad (1.34)$$

parametrized by $w \in \mathcal{C}([0, r])$, where $W_{[0, r]}$ denotes the Wiener measure on the space $\mathcal{C}([0, r])$ and $\int_{\mathcal{C}([0, r])} \mathbb{P}_{X_s^{w'}} W_{[0, r]}(dw')$ denotes the probability measure $\tilde{\nu}$ such that, for all $E \in \mathcal{B}(\mathcal{C}([0, r]))$, $\tilde{\nu}(E) := \int_{\mathcal{C}([0, r])} \mathbb{P}_{X_s^{w'}}(E) W_{[0, r]}(dw')$.

Lemma 1.3.10. Under hypotheses $(\mathbf{H}_{b,\sigma})$, the system of SDEs (1.34) is well-posed, i.e. it admits a unique solution $(X^w)_{w \in \mathcal{C}([0, r])}$.

We postpone the proof to Appendix B.2. Letting $X = (X^w)_{w \in \mathcal{C}([0, r])}$ denote the solution to (1.34), call $X^B := X_{|w=B_{\cdot \wedge r}^w}$ the stochastic process defined by composition with the Brownian trajectory on $[0, r]$, namely

$$X_s^B(\omega) = X_s^{B_{[0, r]}(\omega)}(\omega)$$

for every $s \geq r$, $\omega \in \Omega$. Then we have the following disintegration result.

Proposition 1.3.11. Under hypotheses $(\mathbf{H}_{b,\sigma})$, up to null sets, X^B is equal to the unique solution $X^{r,\xi,\alpha}$ to the state equation (1.32) starting from ξ at time r under the control α .

Proof. We begin by observing that X^B is well defined because, by Section 12 in [87], the map $(s, \omega, w) \rightarrow X_s^w(\omega)$ is measurable with respect to $\mathcal{P}(\mathbb{F}^r) \otimes \mathcal{B}(\mathcal{C}([0, r]))$, where $\mathcal{P}(\mathbb{F}^r)$ denotes the predictable σ -algebra with respect to $\mathbb{F}^r$. Moreover, notice that the random variable $\omega \rightarrow B_{[0, r]}(\omega)$ is independent of X^w , as all the sources of randomness appearing in equation (1.34) (ξ , U and $B^{(r)}$) are independent of $B_{[0, r]}$. Then, as $X = (X^w)_{w \in \mathcal{C}([0, r])}$ is the solution to (1.34), X^B solves

$$\begin{aligned} dX_s^B = & b\left(X_s^B, \int_{\mathcal{C}([0, r])} \mathbb{P}_{X_s^{w'}} W_{[0, r]}(dw'), \alpha_s^r(w)_{|w=B_{[0, r]}}\right) ds \\ & + \sigma\left(X_s^B, \int_{\mathcal{C}([0, r])} \mathbb{P}_{X_s^{w'}} W_{[0, r]}(dw'), \alpha_s^r(w)_{|w=B_{[0, r]}}\right) dB_s, \quad s \geq r \end{aligned}$$

$$X_r^B = \xi. \quad (1.35)$$

We now want to prove that equation (1.35) corresponds to equation (1.32); in this case, the thesis follows from uniqueness of solutions to McKean-Vlasov equations.

Notice that, by definition (1.33) of $\alpha^r(w)$, $\alpha_s^r(w)|_{w=B_{[0,r]}} = \alpha_s$ for every $s \geq r$. It is thus sufficient to prove that

$$\mathbb{P}_{X_s^B} = \int_{\mathcal{C}([0,r])} \mathbb{P}_{X_s^{w'}} W_{[0,r]}(dw') \quad \text{for every } s \geq r.$$

Now, by Fubini theorem, for every $s \geq r$ and every $\phi \in \mathcal{C}_b(\mathbb{R}^d)$,

$$\begin{aligned} \int_{\mathbb{R}^d} \phi(z) \left(\int_{\mathcal{C}([0,r])} \mathbb{P}_{X_s^{w'}} W_{[0,r]}(dw') \right) (dz) &= \int_{\mathcal{C}([0,r])} \mathbb{E}[\phi(X_s^{w'})] W_{[0,r]}(dw') \\ &= \mathbb{E} \left[\phi(X_s^B) \right] = \int_{\mathbb{R}^d} \phi(z) \mathbb{P}_{X_s^B}(dz) \end{aligned}$$

by conditional independence (as X^w is independent of $B_{[0,r]}$) and the thesis is satisfied. $\square$

Remark 1.3.12. Notice that, under the assumptions of Lemma 1.3.10, for every fixed $w \in \mathcal{C}([0,r])$ we can introduce the following gain functional:

$$J^w(r, \xi, \alpha) := \mathbb{E} \left[\int_r^{+\infty} e^{-\beta(s-r)} f(X_s^w, \int_{\mathcal{C}([0,r])} \mathbb{P}_{X_s^{w'}} W_{[0,r]}(dw'), \alpha_s^r(w)) dr \right], \quad (1.36)$$

which is well defined by standard estimates and Proposition 1.3.11. Firstly, by Fubini Theorem

$$\begin{aligned} \left\| \int_{\mathcal{C}([0,r])} \mathbb{P}_{X_v^{w'}} W_{[0,r]}(dw') \right\|_2^2 &= \int_{\mathbb{R}^d} |z|^2 \left(\int_{\mathcal{C}([0,r])} \mathbb{P}_{X_v^{w'}}(dz) W_{[0,r]}(dw') \right) \\ &= \int_{\mathcal{C}([0,r])} \int_{\mathbb{R}^d} |z|^2 \mathbb{P}_{X_v^{w'}}(dz) W_{[0,r]}(dw') \\ &= \int_{\mathcal{C}([0,r])} \|\mathbb{P}_{X_v^{w'}}\|_2^2 W_{[0,r]}(dw') \\ &= \int_{\mathcal{C}([0,r])} \mathbb{E}[|X_v^{w'}|^2] W_{[0,r]}(dw'). \end{aligned} \quad (1.37)$$

As a consequence, using the Itô formula for functions of probability measures (see e.g. Theorem 4.16 in [42]) and Assumptions $(\mathbf{H}_{b,\sigma})$,

$$\begin{aligned} \mathbb{E}[|X_s^w|^2] &= \mathbb{E}[|\xi|^2] + \int_r^s \mathbb{E} \left[\langle b \left(X_v^w, \int_{\mathcal{C}([0,r])} \mathbb{P}_{X_v^{w'}} W_{[0,r]}(dw'), \alpha_v^r(w) \right), 2X_v^w \rangle \right] dv \\ &\quad + \int_r^s \mathbb{E} [tr(\sigma \sigma^T(X_v^w, \int_{\mathcal{C}([0,r])} \mathbb{P}_{X_v^{w'}} W_{[0,r]}(dw'), \alpha_v^r(w)))] dv \\ &\leq \|\xi\|_{L^2}^2 + 6M^2 \int_r^s \mathbb{E} \left[1 + |X_v^w|^2 + \left\| \int_{\mathcal{C}([0,r])} \mathbb{P}_{X_v^{w'}} W_{[0,r]}(dw') \right\|_2^2 \right] dv \end{aligned}$$

$$\begin{aligned}
& + \int_r^s \mathbb{E}[|X_v^w|^2] dv \quad (\text{by (1.37)}) \\
& \leq \|\xi\|_{L^2}^2 + 6M^2 \int_r^s \mathbb{E} \left[1 + |X_v^w|^2 + \int_{\mathcal{C}([0,r])} \mathbb{E}[|X_v^{w'}|^2] W_{[0,r]}(dw') \right] dv \\
& + \int_r^s \mathbb{E}[|X_v^w|^2] dv \\
& = \|\xi\|_{L^2}^2 + 6M^2(s-r) + (6M^2+1) \int_r^s \mathbb{E} \left[|X_v^w|^2 \right] dv \\
& + 6M^2 \int_r^s \int_{\mathcal{C}([0,r])} \mathbb{E}[|X_v^{w'}|^2] W_{[0,r]}(dw') dv \quad (\text{by conditional independence}) \\
& = \|\xi\|_{L^2}^2 + 6M^2(s-r) + (6M^2+1) \int_r^s \mathbb{E} \left[|X_v^w|^2 \right] dv \\
& + 6M^2 \int_r^s \mathbb{E}[|X_v^{B_{[0,r]}}|^2] dv \quad (\text{by Proposition 1.3.11}) \\
& = \|\xi\|_{L^2}^2 + 6M^2(s-r) + (6M^2+1) \int_r^s \mathbb{E} \left[|X_v^w|^2 \right] dv \\
& + 6M^2 \int_r^s \mathbb{E}[|X_v^{r,\xi,\alpha}|^2] dv \quad (\text{by estimate (1.4)}) \\
& = \|\xi\|_{L^2}^2 + 6M^2(s-r) + (6M^2+1) \int_r^s \mathbb{E} \left[|X_v^w|^2 \right] dv \\
& + 6M^2 \int_r^s \left(\|\xi\|_{L^2}^2 + 6M^2(v-r) \right) e^{(12M^2+1)(v-r)} dv \\
& = \|\xi\|_{L^2}^2 + h(s) + C \int_r^s \mathbb{E} \left[|X_v^w|^2 \right] dv,
\end{aligned}$$

where $C = (6M^2 + 1)$ and

$$h(s) = 6M^2(s-r) + 6M^2 \int_r^s \left(\|\xi\|_{L^2}^2 + 6M^2(v-r) \right) e^{(12M^2+1)(v-r)} dv \quad (1.38)$$

is an increasing function. Thus, by Gronwall Lemma,

$$\begin{aligned}
\mathbb{E}[|X_s^w|^2] & \leq \left(\|\xi\|_{L^2}^2 + h(s) \right) \left(1 + C \int_r^s e^{C(s-v)} dv \right) \\
& \leq \left(\|\xi\|_{L^2}^2 + h(s) \right) e^{C(s-r)}.
\end{aligned} \quad (1.39)$$

If f and β satisfy $(\mathbf{H}_{f,\beta})$, we can thus prove that J^w is well defined. As a matter of fact, using again $(\mathbf{H}_{f,\beta})$, (1.37), Proposition 1.3.11 and conditional independence,

$$\begin{aligned}
& \mathbb{E} \left[\int_r^{+\infty} e^{-\beta(s-r)} \left| f(X_s^w, \int_{\mathcal{C}([0,r])} \mathbb{P}_{X_s^{w'}} W_{[0,r]}(dw'), \alpha_s^r(w)) \right| ds \right] \\
& \leq \mathbb{E} \left[\int_r^{+\infty} e^{-\beta(s-r)} K \left(1 + |X_s^w|^2 + \int_{\mathcal{C}([0,r])} \|\mathbb{P}_{X_s^{w'}}\|_2^2 W_{[0,r]}(dw') \right) ds \right]
\end{aligned}$$

$$\begin{aligned}
&\leq \int_r^{+\infty} e^{-\beta(s-r)} K \left(1 + \mathbb{E} \left[|X_s^w|^2 \right] + \int_{\mathcal{C}([0,r])} \|\mathbb{P}_{X_s^{w'}}\|_2^2 W_{[0,r]}(dw') \right) ds \\
&= \int_r^{+\infty} e^{-\beta(s-r)} K \left(1 + \mathbb{E} \left[|X_s^w|^2 \right] + \mathbb{E}[|X_s^{r,\xi,\alpha}|^2] \right) ds \quad (\text{by (1.4) and (1.39)}) \\
&\leq \int_r^{+\infty} e^{-\beta(s-r)} K \left[1 + \left(\|\xi\|_{L^2}^2 + h(s) \right) e^{(6M^2+1)(s-r)} \right. \\
&\quad \left. + \left(\|\xi\|_{L^2}^2 + 6M^2(s-r) \right) e^{(12M^2+1)(s-r)} \right] ds.
\end{aligned}$$

This integral converges if and only if the following integrals are all convergent:

$$I_1 := \int_r^{+\infty} e^{-\beta(s-r)} K ds,$$

$$I_2 := \int_r^{+\infty} e^{-\beta(s-r)} K \left(\|\xi\|_{L^2}^2 + h(s) \right) e^{(6M^2+1)(s-r)} ds$$

and

$$I_3 := \int_r^{+\infty} e^{(12M^2+1-\beta)(s-r)} K \left(\|\xi\|_{L^2}^2 + 6M^2(s-r) \right) ds.$$

We already know that $I_1 < +\infty$ for any $\beta > 0$ and that $I_3 < +\infty$ for $\beta > 12M^2 + 1$. Using the explicit expression of the function h (1.38),

$$\begin{aligned}
I_2 &= \int_r^{+\infty} K e^{(6M^2+1-\beta)(s-r)} \left[\|\xi\|_{L^2}^2 + 6M^2(s-r) \right. \\
&\quad \left. + 6M^2 \int_r^s \left(\|\xi\|_{L^2}^2 + 6M^2(v-r) \right) e^{(12M^2+1)(v-r)} dv \right] ds \quad (\text{the integrand is increasing}) \\
&\leq \int_r^{+\infty} K e^{(6M^2+1-\beta)(s-r)} \left[\|\xi\|_{L^2}^2 + 6M^2(s-r) \right. \\
&\quad \left. + 6M^2(s-r) \left(\|\xi\|_{L^2}^2 + 6M^2(s-r) \right) e^{(12M^2+1)(s-r)} \right] ds,
\end{aligned}$$

which on its turn converges if and only if

$$\int_r^{+\infty} K e^{(6M^2+1-\beta)(s-r)} \left[\|\xi\|_{L^2}^2 + 6M^2(s-r) \right] ds < +\infty,$$

which is true for $\beta > 12M^2 + 1 > 6M^2 + 1$, and

$$\int_r^{+\infty} K e^{(6M^2+1+12M^2+1-\beta)(s-r)} 6M^2(s-r) \left(\|\xi\|_{L^2}^2 + 6M^2(s-r) \right) < +\infty,$$

which holds provided that we require the stronger bound $\beta > 18M^2 + 2$, as in $(\mathbf{H}_{f,\beta})$.

Obviously, if f satisfies $(\mathbf{H}'_{f,\beta})$, then J^w is well defined for any $\beta > 0$. $\square$

Proof of the time invariance property

In order to prove Theorem 1.3.8 we need to establish some continuity property of the gain functional J .

Proposition 1.3.13 (Right Continuity of J). Assume that hypotheses $(\mathbf{H}_{b,\sigma})$, $(\mathbf{H}_{f,\beta})$ (or $(\mathbf{H}'_{f,\beta})$) and $(\mathbf{H}_{f_{cont}})$ are satisfied. Fix an initial condition $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ and a time $t \geq 0$. Then, the function $(u, \alpha) \rightarrow J(u, \xi, \alpha)$ is right continuous at t in the variable u uniformly with respect to $\alpha \in \mathcal{A}$, i.e. for every $\epsilon > 0$ there exists a $\zeta = \zeta(t, \epsilon, \xi)$ such that, if $u \in (t, t + \zeta)$, then

$$|J(t, \xi, \alpha) - J(u, \xi, \alpha)| < \epsilon.$$

We postpone the proof to Appendix B.3 and we present here the proof of the main result of this section.

Proof of Theorem 1.3.8. Fix an initial condition $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ and a time $t \geq 0$. First notice that $\mathcal{A}_t \subseteq \mathcal{A}$; therefore,

$$V_t(t, \xi) \leq V(t, \xi).$$

For the converse inequality, consider a time $u > t$. By Proposition 1.3.13, we can say that for any $\epsilon > 0$ there exists a $\delta > 0$ such that

$$|J(t, \xi, \alpha') - J(u, \xi, \alpha')| < \epsilon \quad \forall \alpha' \in \mathcal{A}, \quad (1.40)$$

provided that $u \in (t, t + \delta)$. As $\mathcal{A}_t \subseteq \mathcal{A}$, (1.40) clearly holds even if we substitute $\mathcal{A}$ with $\mathcal{A}_t$. Assume for the moment that the following claim holds:

$$\begin{aligned} &\text{For every } u > t \text{ and every fixed } \alpha \in \mathcal{A} \text{ there exists a control } \alpha' = \alpha'_{\alpha, u} \in \mathcal{A}_t \\ &\text{such that } J(u, \xi, \alpha') = J(u, \xi, \alpha). \end{aligned} \quad (1.41)$$

Now fix a time instant $u \in (t, t + \delta)$ (with δ as in equation (1.40)); for a fixed control $\alpha \in \mathcal{A}$ choose a control $\alpha' = \alpha'_{\alpha, u} \in \mathcal{A}_t$ as in equation (1.41) and compute

$$\begin{aligned} |J(t, \xi, \alpha) - J(t, \xi, \alpha')| &\leq |J(t, \xi, \alpha) - J(u, \xi, \alpha)| + |J(u, \xi, \alpha) - J(u, \xi, \alpha')| \\ &\quad + |J(u, \xi, \alpha') - J(t, \xi, \alpha')| < 2\epsilon \quad (\text{by (1.40) and (1.41)}). \end{aligned}$$

Therefore, for any $\alpha \in \mathcal{A}$ there exists a control $\alpha' \in \mathcal{A}_t$ such that

$$J(t, \xi, \alpha) \leq 2\epsilon + J(t, \xi, \alpha') \leq 2\epsilon + V_t(t, \xi).$$

Then, taking the supremum over the set of feasible controls $\mathcal{A}$ and by the arbitrariness of $\epsilon > 0$, we conclude that

$$V(t, \xi) \leq V_t(t, \xi).$$

It is thus sufficient to prove that (1.41) holds true. The proof is based on the idea that, with a time u “close enough” to t , it is possible to map any trajectory on the time interval $[0, u]$ into a trajectory on the “small interval” $[t, u]$; as a consequence,

the additional source of randomness given by $B_{[0,u]}$ can be reproduced by means of $(B_s^{(t)})_{s \in [t,u]}$ and does not improve the value obtained optimizing over $\mathcal{A}$.

We introduce the following “scaling” function, which maps every continuous path defined on $[t, u]$ to a continuous path defined on $[0, u]$:

$$\begin{aligned} \phi : \mathcal{C}([t, u]) &\rightarrow \mathcal{C}([0, u]) \\ v = (v(s))_{s \in [t, u]} &\rightarrow \phi(v) = w = (w(s))_{s \in [0, u]} : \\ w(s) &= v\left(t + s \frac{u-t}{u}\right) \sqrt{\frac{u}{u-t}} \quad \text{for every } s \in [0, u]. \end{aligned} \quad (1.42)$$

For any $\tilde{B}$ standard Brownian motion on $[t, u]$, we can easily verify that the process $\tilde{B}' := \phi(\tilde{B})$ is a standard Brownian motion on $[0, u]$. Therefore, the push-forward of $W_{[t,u]}$ through ϕ corresponds to $W_{[0,u]}$. Now consider equation (1.34) with initial condition ξ at time u , i.e.

$$\begin{aligned} dX_s^w &= b\left(X_s^w, \int_{\mathcal{C}([0,u])} \mathbb{P}_{X_s^{w'}} W_{[0,u]}(dw'), \alpha_s^u(w)\right) ds \\ &+ \sigma\left(X_s^w, \int_{\mathcal{C}([0,u])} \mathbb{P}_{X_s^{w'}} W_{[0,u]}(dw'), \alpha_s^u(w)\right) dB_s, \quad X_u^w = \xi \end{aligned} \quad (1.43)$$

for every $s \geq u$ and $w \in \mathcal{C}([0, u])$. For every $w \in \mathcal{C}([0, u])$, consider $v \in \mathcal{C}([t, u])$ such that $\phi(v) = w$; such v exists because ϕ is invertible. We introduce the notation

$$X^v := X^{\phi(v)} \quad (1.44)$$

to denote the solution to equation (1.43) labelled by $\phi(v)$ and

$$\gamma_s^u(v) := \alpha_s^u(\phi(v)), \quad s \geq u \quad (1.45)$$

for the associated control. By the previous observation about the push-forward of the Wiener measure, we can see that

$$\int_{\mathcal{C}([0,u])} \mathbb{P}_{X_s^{w'}} W_{[0,u]}(dw') = \int_{\mathcal{C}([t,u])} \mathbb{P}_{X_s^{v'}} W_{[t,u]}(dv'). \quad (1.46)$$

We deduce that $(X^v)_{v \in \mathcal{C}([t,u])}$ solves the system of equations

$$\begin{aligned} dX_s^v &= b\left(X_s^v, \int_{\mathcal{C}([t,u])} \mathbb{P}_{X_s^{v'}} W_{[t,u]}(dv'), \gamma_s^u(v)\right) ds \\ &+ \sigma\left(X_s^v, \int_{\mathcal{C}([t,u])} \mathbb{P}_{X_s^{v'}} W_{[t,u]}(dv'), \gamma_s^u(v)\right) dB_s, \quad X_u^v = \xi \end{aligned} \quad (1.47)$$

for every $s \geq u$ and for every $v \in \mathcal{C}([t, u])$. We now compose X^v with $B_{[t,u]} := (B_s^{(t)})_{t \leq s \leq u}$, which is independent of X^v . Call $\bar{X} := X_{|v=B_{[t,u]}}^v$ and $\bar{\gamma} := \gamma^u(B_{[t,u]})$; by the definition (1.45) of γ^u , we have $\bar{\gamma}_s = \underline{\gamma}(s, \phi(B_{[t,u]}) \otimes_u B^{(u)}, U)$ for a suitable measurable function $\underline{\gamma}$. As a consequence, our control $\bar{\gamma} \in \mathcal{A}_t$ because it is a measurable function

of $B_{[t,u]}$, $B^{(u)}$ and U . We also notice that, repeating the same reasoning as in the proof of Proposition 1.3.11,

$$\int_{\mathcal{C}([t,u])} \mathbb{P}_{X_s^v} W_{[t,u]}(dv) = \mathbb{P}_{\bar{X}_s}, \quad (1.48)$$

i.e. $\bar{X}$ is the unique solution to equation

$$d\bar{X}_s = b\left(\bar{X}_s, \mathbb{P}_{\bar{X}_s}, \bar{\gamma}_s\right)ds + \sigma\left(\bar{X}_s, \mathbb{P}_{\bar{X}_s}, \bar{\gamma}_s\right)dB_s, \quad \bar{X}_u = \xi, \quad (1.49)$$

which is a McKean-Vlasov controlled equation of the form of equation (1.32) with initial time u ; by the property of uniqueness of solutions to McKean-Vlasov SDEs, $\bar{X} = X^{u,\xi,\bar{\gamma}}$. We now compute the gain associated with this dynamics: recalling (1.48), (1.44), (1.45), (1.46) and the definition of push-forward,

$$\begin{aligned} J(u, \xi, \bar{\gamma}) &= \mathbb{E}\left[\int_u^{+\infty} e^{-\beta(s-u)} f(X_s^{u,\xi,\bar{\gamma}}, \mathbb{P}_{X_s^{u,\xi,\bar{\gamma}}}, \bar{\gamma}_s) ds\right] \\ &= \mathbb{E}\left[\int_u^{+\infty} e^{-\beta(s-u)} f(\bar{X}_s, \mathbb{P}_{\bar{X}_s}, \bar{\gamma}_s) ds\right] \text{ (by conditional independence)} \\ &= \int_{\mathcal{C}([t,u])} \mathbb{E}\left[\int_u^{+\infty} e^{-\beta(s-u)} f(X_s^v, \int_{\mathcal{C}([t,u])} \mathbb{P}_{X_s^v} W_{[t,u]}(dv), \gamma_s^u(v)) ds\right] W_{[t,u]}(dv) \\ &= \int_{\mathcal{C}([t,u])} \mathbb{E}\left[\int_u^{+\infty} e^{-\beta(s-u)} f(X_s^{\phi(v)}, \int_{\mathcal{C}([t,u])} \mathbb{P}_{X_s^{\phi(v)}} W_{[t,u]}(dv), \alpha_s^u(\phi(v))) ds\right] W_{[t,u]}(dv) \\ &= \int_{\mathcal{C}([t,u])} \mathbb{E}\left[\int_u^{+\infty} e^{-\beta(s-u)} f(X_s^{\phi(v)}, \int_{\mathcal{C}([0,u])} \mathbb{P}_{X_s^{w'}} W_{[0,u]}(dw'), \alpha_s^u(\phi(v))) ds\right] W_{[t,u]}(dv) \\ &= \int_{\mathcal{C}([0,u])} \mathbb{E}\left[\int_u^{+\infty} e^{-\beta(s-u)} f(X_s^w, \int_{\mathcal{C}([0,u])} \mathbb{P}_{X_s^w} W_{[0,u]}(dw'), \alpha_s^u(w)) ds\right] W_{[0,u]}(dw) \\ &= \mathbb{E}\left[\int_u^{+\infty} e^{-\beta(s-u)} f(X^{u,\xi,\alpha}, \mathbb{P}_{X^{u,\xi,\alpha}}, \alpha_s) ds\right] \\ &= J(u, \xi, \alpha), \end{aligned}$$

where in the second-to-last equality we used again conditional independence and Proposition 1.3.11 applied to equation (1.43). $\square$

Remark 1.3.14. The basic arguments in the previous proof as well as in Example 1.3.3 are based on a randomization procedure, where the source of noise was derived from the trajectory of the Brownian motion preceding the initial time. One could randomize in a similar way by means of the independent random variable U . The previous proof has the advantage that it applies to situations where the initial condition is deterministic and the σ -algebra $\mathcal{G}$ can be trivial. $\square$

By Proposition 1.3.1 and Theorem 1.3.8 we immediately conclude the following.

Theorem 1.3.15 (Time invariance property). Suppose that assumptions $(\mathbf{H}_{b,\sigma})$, $(\mathbf{H}_{f,\beta})$ (or $(\mathbf{H}'_{f,\beta})$) and $(\mathbf{H}_{f,cont})$ are satisfied. Let $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$. Then,

$$V(0, \xi) = V(t, \xi) \quad \forall t \geq 0, \quad (1.50)$$

i.e. the value associated to a $\mathcal{G}$ -measurable initial condition does not depend on the initial time.

Remark 1.3.16. In this section, we proved that, under assumptions $(\mathbf{H}_{b,\sigma})$, $(\mathbf{H}_{f,\beta})$ or $(\mathbf{H}'_{f,\beta})$ and $(\mathbf{H}_{f_{cont}})$, the equivalence of the value functions

$$V_t(t, \xi) = V(t, \xi) \quad \text{for all } t \geq 0, \xi \in L^2(\Omega, \mathcal{G}, \mathbb{P}) \quad (1.51)$$

implies the time invariance property. The two properties are in fact equivalent. Indeed, assume that the Time Invariance Property holds, i.e. that for every $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ it holds that

$$V(0, \xi) = V(t, \xi) \quad \forall t \geq 0.$$

Assume by contradiction that property (1.51) does not hold. As $\mathcal{A}_t \subseteq \mathcal{A}$, it is clear that for every $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ and $t \geq 0$

$$\sup_{\alpha \in \mathcal{A}_t} J(t, \xi, \alpha) \leq \sup_{\alpha \in \mathcal{A}} J(t, \xi, \alpha).$$

Therefore, if (1.51) does not hold, there exist a fixed time $t \geq 0$ and a random variable $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ such that

$$V_t(t, \xi) = \sup_{\alpha \in \mathcal{A}_t} J(t, \xi, \alpha) < V(t, \xi). \quad (1.52)$$

Now fix $\alpha \in \mathcal{A}$; by Lemma 1.2.3, α is indistinguishable from

$$((\underline{\alpha}(s, B_{\cdot \wedge s}, U)))_{s \geq 0}$$

for some Borelian function $\underline{\alpha}$. Construct a control γ as follows:

$$\gamma_s(\omega) := \underline{\alpha}(s - t, B_{\cdot \wedge s}^{(t)}(\omega), U(\omega)) \quad \text{for every } s \geq t, \omega \in \Omega.$$

Then, clearly

$$\mathcal{L}((\alpha_s)_{s \geq 0}, (B_s)_{s \geq 0}) = \mathcal{L}((\gamma_{s+t})_{s \geq 0}, ((B_{s+t}^{(t)})_{s \geq 0}))$$

and γ belongs to $\mathcal{A}_t$. Now, by Lemma 1.3.2,

$$\begin{aligned} J(0, \xi, \alpha) &= \mathbb{E} \left[\int_0^{+\infty} e^{-\beta s} f(X_s^{0, \xi, \alpha}, \mathbb{P}_{X_s^{0, \xi, \alpha}}, \alpha_s) ds \right] \\ &= \mathbb{E} \left[\int_0^{+\infty} e^{-\beta s} f(X_{s+t}^{t, \xi, \gamma}, \mathbb{P}_{X_{s+t}^{t, \xi, \gamma}}, \gamma_{s+t}) ds \right] \quad (z = s + t) \\ &= \mathbb{E} \left[\int_t^{+\infty} e^{-\beta(z-t)} f(X_z^{t, \xi, \gamma}, \mathbb{P}_{X_z^{t, \xi, \gamma}}, \gamma_z) dz \right] \\ &= J(t, \xi, \gamma) \leq V_t(t, \xi) \end{aligned}$$

because $\gamma \in \mathcal{A}_t$. Taking the supremum over all $\alpha \in \mathcal{A}$, we conclude that

$$V(0, \xi) \leq V_t(t, \xi).$$

Now, as the Time Invariance Property holds and ξ is a $\mathcal{G}$ -measurable random variable, this inequality implies that

$$V(t, \xi) = V(0, \xi) \leq V_t(t, \xi),$$

which contradicts (1.52). $\square$

1.4 Continuity of the value function and law invariance property

In this section, we discuss the properties of the lifted value function.

Proposition 1.4.1. Under assumptions $(\mathbf{H}_{b,\sigma})$ and $(\mathbf{H}_{f,\beta})$, for any $t \geq 0$ the lifted value function V at time t satisfies a quadratic growth condition in ξ , i.e. there exists a positive constant C such that

$$|V(t, \xi)| \leq C(1 + \|\xi\|_{L^2}^2) \quad (1.53)$$

for every $\xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$. On the other hand, under assumptions $(\mathbf{H}_{b,\sigma})$ and $(\mathbf{H}'_{f,\beta})$, the lifted value function V is bounded, i.e. there exists a constant $C > 0$ such that $|V(t, \xi)| \leq C$ for every $t \geq 0$ and $\xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$.

Proof. Firstly, notice that $|V(t, \xi)| \leq \sup_{\alpha \in \mathcal{A}} |J(t, \xi, \alpha)|$; it is thus sufficient to prove the desired bounds for the functional J uniformly with respect to α . If assumptions $(\mathbf{H}_{b,\sigma})$ and $(\mathbf{H}_{f,\beta})$ are satisfied,

$$\begin{aligned} |J(t, \xi, \alpha)| &\leq \mathbb{E} \left[\int_t^{+\infty} K e^{-\beta(s-t)} (1 + |X_s^{t,\xi,\alpha}|^2 + \|\mathbb{P}_{X_s^{t,\xi,\alpha}}\|_2^2) ds \right] \\ &\leq \int_t^{+\infty} K e^{-\beta(s-t)} ds + \int_t^{+\infty} 2K e^{-\beta(s-t)} \mathbb{E} \left[|X_s^{t,\xi,\alpha}|^2 \right] ds \quad (\text{by (1.4)}) \\ &\leq \frac{K}{\beta} + \int_t^{+\infty} 2K e^{(-\beta+12M^2+1)(s-t)} \left(\|\xi\|_{L^2}^2 + 6M^2(s-t) \right) ds \\ &\leq \frac{K}{\beta} + \tilde{C} \|\xi\|_{L^2}^2 + \bar{C} \\ &\leq C(1 + \|\xi\|_{L^2}^2) \end{aligned}$$

where

$$\tilde{C} := 2K \int_t^{+\infty} e^{(12M^2+1-\beta)(s-t)} ds$$

and

$$\bar{C} = 2K \int_t^{+\infty} 6M^2(s-t) e^{(12M^2+1-\beta)(s-t)} ds$$

are finite constants due to hypotheses $(\mathbf{H}_{f,\beta})$, so that

$$C := \max \left\{ \tilde{C}, \frac{K}{\beta} + \bar{C} \right\}$$

is a positive and finite constant depending only on β , K and M , but not on α .

On the other hand, if Assumptions $(\mathbf{H}_{b,\sigma})$ and $(\mathbf{H}'_{f,\beta})$ hold, then

$$|J(t, \xi, \alpha)| \leq \mathbb{E} \left[\int_t^{+\infty} e^{-\beta(s-t)} |f(X_s^{t,\xi,\alpha}, \mathbb{P}_{X_s^{t,\xi,\alpha}}, \alpha_s)| ds \right]$$

$$\leq C_f \int_t^{+\infty} e^{-\beta(s-t)} ds = \frac{C_f}{\beta}$$

for every $\alpha \in \mathcal{A}$; therefore, the thesis follows with $C = \frac{C_f}{\beta}$. $\square$

From now on, we always assume that $(\mathbf{H}_{b,\sigma})$, $(\mathbf{H}_{f,\beta})$ (or $(\mathbf{H}'_{f,\beta})$) and $(\mathbf{H}_{f_{cont}})$ are satisfied.

Proposition 1.4.2. The lifted value function V is right continuous with respect to time, i.e. for any $t \geq 0$ and $\xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$, $V(u, \xi) \rightarrow V(t, \xi)$ when $u \rightarrow t^+$. Moreover, V is continuous with respect to the initial condition, i.e. for every $t \geq 0$, $\xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$, $V(t, \eta) \rightarrow V(t, \xi)$ when $\eta \rightarrow \xi$ in $L^2(\Omega, \mathcal{F}_t, \mathbb{P})$.

Proof. For both properties, it is sufficient to prove the corresponding continuity of the gain functional J uniformly with respect to α . Therefore, the right continuity of V easily follows from Proposition 1.3.13, as the right continuity of J with respect to time is uniform with respect to α . The continuity of J with respect to the initial condition uniformly in α can be proved with similar techniques; we refer the reader to Proposition B.3.2 in Appendix B.3 for more details. $\square$

In analogy with the finite horizon case, presented in [42], we want to redefine the *value function* of the problem as a real-valued function on $\mathcal{P}_2(\mathbb{R}^d)$. This requires the so-called law invariance property, stating that the lifted value function V depends on the initial condition only through its probability distribution (which motivates the choice of the expression “lifted value function”).

Theorem 1.4.3 (Law Invariance Property). Fix any $t \geq 0$ and any $\xi, \eta \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$ such that $\mathbb{P}_\eta = \mathbb{P}_\xi$. Then, $V(t, \eta) = V(t, \xi)$

Proof. The result can be proved following the lines of the proof of Theorem 3.6 in [42] with minor changes. In particular, our initial conditions $\xi, \eta \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$ are random variables instead of stochastic processes; the result can be proved approximating $\xi, \eta \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$ by means of suitable sequences of discrete random variables, for which the thesis holds as in Theorem 3.6 in [42], and then exploiting the continuity of the lifted value function with respect to the initial condition (Proposition 1.4.2). $\square$

Therefore, thanks to Theorem 1.3.15 and Theorem 1.4.3, we can define the *value function* $v : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ as

$$v(\mu) = V(0, \xi) \quad \text{for any } \mu \in \mathcal{P}_2(\mathbb{R}^d)$$

for every $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ such that $\mathbb{P}_\xi = \mu$.

Remark 1.4.4. Notice that $v(\mu)$ represents the value of the problem for any initial time $t \geq 0$ and any initial condition $\eta \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$ such that $\mathbb{P}_\eta = \mu$. As a matter of fact, as $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $\mathcal{G} = \sigma(U)$, there exists another random variable $\eta' \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ such that $\mathbb{P}_{\eta'} = \mathbb{P}_\eta = \mu$. Notice that, if η' is $\mathcal{G}$ -measurable, it is also $\mathcal{F}_t$ -measurable for any $t \geq 0$. Then, $V(t, \eta) = V(t, \eta')$ by Theorem 1.4.3. Furthermore, by Theorem 1.3.15 applied to the $\mathcal{G}$ -measurable random variable η' , $V(t, \eta') = V(0, \eta')$. Applying

again Theorem 1.4.3, $V(0, \eta') = V(0, \xi)$ for any $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ such that $\mathbb{P}_\xi = \mu$, which also equals $v(\mu)$. We conclude that, for any $t \geq 0$ and any $\eta \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$ having law μ , it holds that $V(t, \eta) = v(\mu)$.

Consequently, the main properties of v can be deduced from the properties of the lifted value function. We start by stating the dynamic programming principle for the value function v , which follows from Theorem 1.2.10 with $t = 0$.

Corollary 1.4.5 (DPP). For any $s \geq 0$ and for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$,

$$v(\mu) = \sup_{\alpha \in \mathcal{A}} \left\{ \mathbb{E} \left[\int_0^s e^{-\beta r} f(X_r^{0, \xi, \alpha}, \mathbb{P}_{X_r^{0, \xi, \alpha}}, \alpha_r) dr \right] + e^{-\beta s} v(\mathbb{P}_{X_s^{0, \xi, \alpha}}) \right\} \quad (1.54)$$

for any $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ such that $\mathbb{P}_\xi = \mu$.

We then deduce the following result from Proposition 1.4.1.

Corollary 1.4.6. Under assumptions $(\mathbf{H}_{f, \beta})$, the value function v satisfies a quadratic growth condition in μ , i.e. there exists a positive constant C such that $|v(\mu)| \leq C(1 + \|\mu\|_2^2)$ for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. On the other hand, if assumptions $(\mathbf{H}'_{f, \beta})$ hold, the value function v is bounded.

We conclude the section studying the continuity of the value function, which will be a key result for its viscosity property.

Proposition 1.4.7. The value function v is continuous, i.e. for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and any sequence $(\mu_n)_{n \in \mathbb{N}} \subseteq \mathcal{P}_2(\mathbb{R}^d)$ converging to μ , it holds that $\lim_{n \rightarrow +\infty} v(\mu_n) = v(\mu)$.

Proof. Fix any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and any sequence $(\mu_n)_{n \in \mathbb{N}} \subseteq \mathcal{P}_2(\mathbb{R}^d)$ converging to μ . By Lemma 2.1 in [10], as $\mathcal{G} = \sigma(U)$, there exists a sequence $(\xi_n)_{n \in \mathbb{N}} \subseteq L^2(\Omega, \mathcal{G}, \mathbb{P})$ with $\mathbb{P}_{\xi_n} = \mu_n$ for any $n \in \mathbb{N}$ converging $\mathbb{P}$ -a.s. and in $L^2(\Omega, \mathcal{G}, \mathbb{P})$ to some random variable $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ with $\mathbb{P}_\xi = \mu$. As the lifted value function at time $t = 0$ is continuous with respect to the initial condition (Proposition 1.4.2), we conclude that $\{V(0, \xi_n)\}_{n \in \mathbb{N}} = \{v(\mu_n)\}_{n \in \mathbb{N}}$ converges to $V(0, \xi) = v(\mu)$. $\square$

1.5 Hamilton-Jacobi-Bellman equation

1.5.1 Lions derivative and viscosity solutions to PDEs

In this section, we are studying our optimal control problem by means of the classical approach based on Hamilton-Jacobi-Bellman equation: we aim at deriving from the dynamic programming principle, i.e. Corollary 1.4.5, a partial differential equation on the domain of the value function, in order to characterize v as its unique solution. As in many applications of control theory it is not easy or even possible to prove the regularity of the value function, we expect it to solve a PDE in the *viscosity sense*. We first need to introduce an appropriate notion of derivative on the space $\mathcal{P}_2(\mathbb{R}^d)$; we are using the derivative in the lifted sense of Lions ([70], see also [30]). We recall here that, given a map $u : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$, we call *lifting of u* any function $U : L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d) \rightarrow \mathbb{R}$ such that $U(\xi) = u(\mathbb{P}_\xi)$ for any $\xi \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$.

Definition 1.5.1. A function $u : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ is said to be first-order L -differentiable if its lifting U admits a continuous Fréchet derivative $D_\xi U : L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d) \rightarrow L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$.

Remark 1.5.2. By Theorem 6.5 in [30], if u is first-order differentiable, then there exists, for any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, a measurable function $\partial_\mu u(\mu) : \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that $\mathbb{P}$ -a.s. $D_\xi U(\xi) = \partial_\mu u(\mu)(\xi)$ for any $\xi \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ with $\mathbb{P}_\xi = \mu$. The function $\partial_\mu u(\mu)$ (defined μ -a.s.) is called first-order Lions derivative of u at μ . $\square$

Definition 1.5.3. We denote by $\mathcal{C}^2(\mathcal{P}_2(\mathbb{R}^d))$ the space of continuous functions $\phi : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ such that the following conditions hold:

1. ϕ is first-order L -differentiable;
2. The map $(\mu, x) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \rightarrow \partial_\mu \phi(\mu)(x) \in \mathbb{R}^d$ is jointly continuous.
3. The derivative $\partial_x \partial_\mu \phi : (\mu, x) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \rightarrow \partial_x \partial_\mu \phi(\mu)(x) \in \mathbb{R}^{d \times d}$ exists and is jointly continuous.

Remark 1.5.4. This is not the only available notion of derivative with respect to a measure; we choose the derivative in the lifted sense because we defined the function v starting exactly from its lifting, the function V . Moreover, this choice is consistent with a part of the existing literature (see e.g. [42], [43], [78], [79]) from which this work was originally inspired. We refer to Section 5 in [32] for the definition of other notions of differentiability, which are equivalent to our definition under suitable regularity conditions (see e.g. Proposition 5.51 and Proposition 5.48 in [32]). $\square$

We now introduce the set of test functions used in this section.

Definition 1.5.5. We denote by $\mathcal{TF}(\mathbb{R}^d)$ the subset of $\mathcal{C}^2(\mathcal{P}_2(\mathbb{R}^d))$ of functions $\phi : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ such that, for some constant $C_\phi \geq 0$,

$$|\partial_\mu \phi(\mu)(x)| \leq C_\phi(1 + |x|) \quad \text{and} \quad |\partial_x \partial_\mu \phi(\mu)(x)| \leq C_\phi$$

for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$.

On the space $\mathcal{P}_2(\mathbb{R}^d)$ endowed with the 2-Wasserstein space, we introduce the following elliptic partial differential equation, which we will refer to as Hamilton-Jacobi-Bellman (HJB) equation:

$$\begin{aligned} \beta v(\mu) = \mathbb{E} \left[\sup_{a \in A} \left\{ f(\xi, \mu, a) + \langle b(\xi, \mu, a), \partial_\mu v(\mu)(\xi) \rangle \right. \right. \\ \left. \left. + \frac{1}{2} \text{tr} \left(\sigma(\xi, \mu, a) \sigma^T(\xi, \mu, a) \partial_x \partial_\mu v(\mu)(\xi) \right) \right\} \right] \end{aligned} \quad (1.55)$$

for $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and any $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ with $\mathbb{P}_\xi = \mu$.

We want to define a suitable notion of viscosity solution to HJB equation.

Definition 1.5.6. A continuous function $u : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ is a viscosity subsolution to equation (1.55) if, for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and every $\phi \in \mathcal{TF}(\mathbb{R}^d)$ such that $u - \phi$ attains a global maximum at μ with value 0, the following inequality is satisfied:

$$0 \leq -\beta \phi(\mu) + \mathbb{E} \left[\sup_{a \in A} \left\{ f(\xi, \mu, a) + \langle b(\xi, \mu, a), \partial_\mu \phi(\mu)(\xi) \rangle \right. \right.$$

$$+ \frac{1}{2} \text{tr} \left(\sigma(\xi, \mu, a) \sigma^T(\xi, \mu, a) \partial_x \partial_\mu \phi(\mu)(\xi) \right) \Big] \quad (1.56)$$

for every $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ with law μ .

In order to define supersolutions, we underline that any function $u : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ can be extended to the space $\mathcal{P}_2(\mathbb{R}^d \times A)$ in a canonical way: for any $\nu \in \mathcal{P}_2(\mathbb{R}^d \times A)$, it is sufficient to define $u(\nu) := u(\nu_1)$ where ν_1 is the marginal of ν such that $\nu(\cdot \times A) = \nu_1(\cdot)$. We also introduce the projection on the first d components $\pi_d : \mathbb{R}^d \times A \rightarrow \mathbb{R}^d$ where, for any $y = (y_1, \dots, y_d, y_{d+1}, \dots) \in \mathbb{R}^d \times A$, $\pi_d(y) = (y_1, \dots, y_d)$.

Definition 1.5.7. A continuous function $u : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ is a viscosity supersolution to equation (1.55) if, for any $\nu \in \mathcal{P}_2(\mathbb{R}^d \times A)$ and for any $\psi \in \mathcal{TF}(\mathbb{R}^d \times A)$ such that $u - \psi$ attains a global minimum with value 0 at ν over the space $\mathcal{P}_2(\mathbb{R}^d \times A)$, the following inequality holds:

$$0 \geq -\beta \psi(\nu) + \int_{\mathbb{R}^d \times A} \left\{ f(x, \mu, a) + \langle b(x, \mu, a), \partial_\mu \psi(\nu)(x, a) \rangle + \frac{1}{2} \text{tr} \left(\sigma(x, \mu, a) \sigma^T(x, \mu, a) \partial_x \partial_\mu \psi(\nu)(x, a) \right) \right\} \nu(dx, da), \quad (1.57)$$

where μ is the marginal distribution of ν on $\mathbb{R}^d$ and $\partial_\mu \psi(\nu)(\cdot, \cdot) := \pi_d(\partial_\nu \psi(\nu)(\cdot, \cdot))$.

Definition 1.5.8. We call u a viscosity solution to equation (1.55) if it is both a supersolution and a subsolution to it.

Remark 1.5.9. Notice that the requirement that A is a Euclidean space enables us to define Lions derivatives on the space $\mathcal{P}_2(\mathbb{R}^d \times A)$. In fact, these derivatives will reduce to Lions derivatives on the space $\mathcal{P}_2(\mathbb{R}^d)$ for a suitable class of functions. Indeed, if $\psi : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ is first-order L -differentiable at μ and ν is such that $\nu(\cdot \times A) = \mu(\cdot)$, then $\partial_\mu \psi(\nu)(\cdot, a) = \partial_\mu \psi(\mu)(\cdot)$ for any $a \in A$. As a matter of fact, by definition,

$$\partial_\nu \psi(\nu)(\eta) = D_\eta \Psi(\eta) \quad \mathbb{P}\text{-a.s.} \quad (1.58)$$

for any Ψ lifting on ψ and $\eta \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d \times A)$ having law ν , where, with a slight abuse of notation, we still write ψ for its canonical extension on $\mathbb{R}^d \times A$. If ψ is in fact defined on $\mathcal{P}_2(\mathbb{R}^d)$, then

$$\Psi(\eta) = \psi(\nu) = \psi(\mu) = \tilde{\Psi}(\pi_d(\eta)), \quad (1.59)$$

where $\tilde{\Psi} : L^2(\Omega, \mathcal{G}, \mathbb{P}; \mathbb{R}^d) \rightarrow \mathbb{R}$ is the lifting of ψ on $\mathcal{P}_2(\mathbb{R}^d)$. By definition of Fréchet derivative, for every $\eta \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d \times A)$,

$$\Psi(\eta + h) = \Psi(\eta) + \langle D_\eta \Psi(\eta), h \rangle_{L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d \times A)} + o(h) \quad (1.60)$$

and, for every $\xi \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$,

$$\tilde{\Psi}(\xi + k) = \tilde{\Psi}(\xi) + \langle D_\xi \tilde{\Psi}(\xi), k \rangle_{L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)} + o(k). \quad (1.61)$$

Since, for any $\eta \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d \times A)$, $\Psi(\eta) = \tilde{\Psi}(\pi_d(\eta))$, if we take $k \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ and call h the vector in $L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d \times A)$ such that

$$h = (k_1, \dots, k_d, 0, 0, \dots),$$

which also implies $\|h\|_{L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d \times A)} = \|k\|_{L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)}$, we can compute

$$\begin{aligned}
\tilde{\Psi}(\pi_d(\eta) + k) &= \Psi(\eta + h) \quad \text{by (1.60)} \\
&= \Psi(\eta) + \langle D_\eta \Psi(\eta), h \rangle_{L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d \times A)} + o(h) \\
&= \tilde{\Psi}(\pi_d(\eta)) + \langle D_\eta \Psi(\eta), h \rangle_{L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d \times A)} + o(h) \\
&= \tilde{\Psi}(\pi_d(\eta)) + \langle D_\eta \Psi(\eta), h \rangle_{L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d \times A)} + o(k) \\
&= \tilde{\Psi}(\pi_d(\eta)) + \langle \pi_d(D_\eta \Psi(\eta)), k \rangle_{L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)} + o(k),
\end{aligned}$$

where the last equality follows from the fact that, for any $F \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d \times A)$,

$$\begin{aligned}
\langle F, h \rangle_{L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d \times A)} &= \mathbb{E}[\langle F, h \rangle_{\mathbb{R}^d \times A}] \\
&= \mathbb{E}\left[\sum_{i=1}^{d+N} F_i h_i\right] \quad (h_i = 0 \text{ for } i > d, h_i = k_i \text{ for } i = 1, \dots, d) \\
&= \mathbb{E}\left[\sum_{i=1}^d F_i k_i\right] \\
&= \mathbb{E}[\langle \pi_d(F), k \rangle_{\mathbb{R}^d}] \\
&= \langle \pi_d(F), k \rangle_{L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)}.
\end{aligned}$$

By (1.61) and by uniqueness of Fréchet derivative, we conclude that

$$D_\xi \tilde{\Psi}(\pi_d(\eta)) = \pi_d(D_\eta \Psi(\eta))$$

in $L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$. Now recall that, by definition of Fréchet derivative,

$$D_\xi \tilde{\Psi}(\pi_d(\eta)) = \partial_\mu \psi(\mu)(\pi_d(\eta)) \quad \mathbb{P}\text{-a.s.},$$

where $\eta \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d \times A)$ has law ν with $\nu(\cdot \times A) = \mu(\cdot)$. By (1.58), we thus conclude that

$$\partial_\mu \psi(\mu)(\pi_d(\eta)) = \pi_d(\partial_\nu \psi(\nu)(\eta)) \quad \mathbb{P}\text{-a.s.},$$

which implies the desired equality.

Remark 1.5.10. A more natural definition of supersolutions can be introduced: indeed, we may define supersolutions “à la Crandall-Lions” to equation (1.55) as continuous functions $u : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ such that, for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and every $\phi \in \mathcal{TF}(\mathbb{R}^d)$ such that $u - \phi$ attains a global minimum at μ with value 0, inequality (1.56) is satisfied with sign $\geq$ instead of $\leq$. As pointed out in [36], Remark 6.1, and in [14], Remark 2.5, this definition of viscosity supersolution is not sufficient to prove uniqueness of viscosity solutions to Hamilton-Jacobi-Bellman equations. We thus prefer to define supersolutions as in Definition 1.5.7, adapting the notion presented in [36] and [14] to an elliptic equation. Nonetheless, in analogy with Remark 6.1 in [36], we prove that inequality (1.57) implies that u is also a supersolution “à la Crandall-Lions” under Assumptions $(\mathbf{H}_{b, \sigma})$, $(\mathbf{H}_{f, \beta})$ or $(\mathbf{H}'_{f, \beta})$ and $(\mathbf{H}_{f_{cont}})$. As a matter of fact, assume u is a viscosity supersolution to equation (1.55) in the sense of Definition 1.5.7; if we fix $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $\phi \in \mathcal{C}_b^2(\mathbb{R}^d)$ such that $u - \phi$ attains a global minimum with value 0 at μ , we can extend both u and ϕ to $\mathcal{P}_2(\mathbb{R}^d \times A)$. Choosing $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ with

law μ and any α_0 $\mathcal{G}$ -measurable and A -valued random variable, call ν the joint law $\mathcal{L}(\xi, \alpha_0) = \nu$. Then, $u - \phi$ attains a global minimum with value 0 at ν over $\mathcal{P}_2(\mathbb{R}^d \times A)$; therefore, by (1.57),

$$0 \geq -\beta\phi(\nu) + \mathbb{E} \left[f(\xi, \mu, \alpha_0) + \langle b(\xi, \mu, \alpha_0), \partial_\mu \phi(\nu)(\xi, \alpha_0) \rangle + \frac{1}{2} \text{tr} \left(\sigma \sigma^T(\xi, \mu, \alpha_0) \partial_x \partial_\mu \phi(\nu)(\xi, \alpha_0) \right) \right].$$

By the arbitrariness of α_0 ,

$$0 \geq -\beta\phi(\nu) + \sup_{\substack{\alpha_0: \Omega \rightarrow A \\ \alpha_0 \text{ } \mathcal{G}\text{-measurable}}} \mathbb{E} \left[f(\xi, \mu, \alpha_0) + \langle b(\xi, \mu, \alpha_0), \partial_\mu \phi(\nu)(\xi, \alpha_0) \rangle + \frac{1}{2} \text{tr} \left(\sigma \sigma^T(\xi, \mu, \alpha_0) \partial_x \partial_\mu \phi(\nu)(\xi, \alpha_0) \right) \right].$$

We claim that the above is equivalent to

$$0 \geq -\beta\phi(\mu) + \mathbb{E} \left[\sup_{a \in A} \left\{ f(\xi, \mu, a) + \langle b(\xi, \mu, a), \partial_\mu \phi(\mu)(\xi) \rangle + \frac{1}{2} \text{tr} \left(\sigma \sigma^T(\xi, \mu, a) \partial_x \partial_\mu \phi(\mu)(\xi) \right) \right\} \right].$$

Call

$$H(x, \mu, a) := f(\xi, \mu, a) + \langle b(\xi, \mu, a), \partial_\mu \phi(\mu)(\xi) \rangle + \frac{1}{2} \text{tr} \left(\sigma \sigma^T(\xi, \mu, a) \partial_x \partial_\mu \phi(\mu)(\xi) \right);$$

thanks to Remark 1.5.9, $\partial_\mu \phi(\mu)(\xi) = \partial_\mu \phi(\nu)(\xi, \alpha_0)$ and our claim reduces to

$$\sup_{\substack{\alpha_0: \Omega \rightarrow A \\ \alpha_0 \text{ } \mathcal{G}\text{-measurable}}} \mathbb{E}[H(\xi, \mu, \alpha_0)] = \int_{\mathbb{R}^d} [\sup_{a \in A} H(x, \mu, a)] \mu(dx).$$

Notice that, for any fixed α_0 $\mathcal{G}$ -measurable random variable,

$$H(\xi, \mu, \alpha_0) \leq \sup_{a \in A} H(\xi, \mu, a) \quad \mathbb{P}\text{-a.s.},$$

which implies that $\sup_{\substack{\alpha_0: \Omega \rightarrow A \\ \alpha_0 \text{ } \mathcal{G}\text{-measurable}}} \mathbb{E}[H(\xi, \mu, \alpha_0)] \leq \int_{\mathbb{R}^d} [\sup_{a \in A} H(x, \mu, a)] \mu(dx)$. For the

converse inequality, since $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$ is a metric space, A is a compact subset of an Euclidean space and H is continuous under $(\mathbf{H}_{f, \beta})$, $(\mathbf{H}_{b, \sigma})$ and $(\mathbf{H}_{f_{cont}})$, we can apply Proposition 7.34 in [22], stating that for any $\epsilon > 0$ there exists a Borel measurable map $\phi^\epsilon : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow A$ such that

$$\sup_{a \in A} H(x, \mu, a) - \epsilon \leq H(x, \mu, \phi^\epsilon(x, \mu))$$

for any $x, \mu \in \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$. Composing H with $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$, the inequality becomes

$$\sup_{a \in A} H(\xi, \mu, a) - \epsilon \leq H(\xi, \mu, \phi^\epsilon(\xi, \mu)) \quad \mathbb{P}\text{-a.s.}$$

Now, $\phi^\epsilon(\xi, \mu)$ becomes a $\mathcal{G}$ -measurable random variable; we conclude that, for any $\epsilon > 0$ there exists a $\mathcal{G}$ -measurable random variable α^ϵ such that

$$\sup_{a \in A} H(\xi, \mu, a) - \epsilon \leq H(\xi, \mu, \alpha^\epsilon) \quad \mathbb{P}\text{-a.s.};$$

therefore,

$$\int_{\mathbb{R}^d} \sup_{a \in A} H(x, \mu, a) \mu(dx) - \epsilon \leq \mathbb{E}[H(\xi, \mu, \alpha^\epsilon)] \leq \sup_{\substack{\alpha_0: \Omega \rightarrow A \\ \alpha_0 \text{ } \mathcal{G}\text{-measurable}}} \mathbb{E}[H(\xi, \mu, \alpha_0)].$$

As ϵ is arbitrary, our claim is proved. Notice that the additional regularity of b, σ, f with respect to a is required to use the measurable selection argument guaranteeing the existence of ϕ^ϵ .

We furthermore underline that the stronger definition of supersolution is only needed to prove uniqueness of solutions; we could prove that v is a supersolution to equation (1.55) in the weaker Crandall-Lions sense, but we prove it directly with the stronger Definition 1.5.7 for consistency with the uniqueness result. $\square$

1.5.2 Hamilton-Jacobi-Bellman equation and value function

We now prove that v solves the Hamilton-Jacobi-Bellman equation associated with our optimal control problem.

Theorem 1.5.11. Under hypotheses $(\mathbf{H}_{b,\sigma})$, $(\mathbf{H}'_{f,\beta})$ and $(\mathbf{H}_{f_{cont}})$, the value function v is a viscosity solution to Hamilton Jacobi Bellman equation (1.55).

Proof. The result follows from an adaptation of Theorem 6.2 in [36] and Theorem 4.2 in [14], with slight modifications due to the elliptic structure of the equation. $\square$

Remark 1.5.12. The result is still valid if we substitute assumptions $(\mathbf{H}'_{f,\beta})$ with $(\mathbf{H}_{f,\beta})$, with minor changes in the proof; it is sufficient to follow the lines of Theorem 5.3 in [42]. Notice that the boundedness of the test functions is not required in our result, while both [14] and [36] require the boundedness of the derivative of test functions with respect to time in the parabolic setting: the difference is due to the fact that they both deal with the stochastic conditional law of the state process with respect to the common noise and thus need to apply the Dominated Convergence Theorem to a term containing this derivative, while we consider a deterministic law and only need to apply Lebesgue Differentiation Theorem, which requires the integrand to be in L^1_{loc} . $\square$

1.5.3 Uniqueness of the solution

We now discuss uniqueness of solutions to equation (1.55). To the best of our knowledge, a comparison principle for viscosity solutions to elliptic partial differential equations on the space $\mathcal{P}_2(\mathbb{R}^d)$ is not directly available. We thus proceed by an approximation argument: starting from a sub (super) solution u to the elliptic equation (1.55), for each time $T > 0$ we construct a suitable finite horizon problem on $[0, T]$ with

associated value function v_T , solving (in the viscosity sense) a parabolic Hamilton-Jacobi-Bellman equation on $[0, T] \times \mathcal{P}_2(\mathbb{R}^d)$, as stated e.g. in [42], and a sub/super solution to the same parabolic equation U_T such that $U_T(0, \cdot) = u(\cdot)$. For this class of Hamilton-Jacobi-Bellman parabolic PDEs, there exist comparison results, ensuring that $v_T \geq U_T$ (respectively, $\leq U_T$) for U_T sub (super) solution to the parabolic HJB equation. Passing to the limit $T \rightarrow +\infty$, we show that $v_T(0, \cdot)$ converges to our value function v and conclude that (1.55) also admits a unique viscosity solution, corresponding to v . The proof of our uniqueness result will thus heavily rely on a comparison principle for parabolic PDEs on $\mathcal{P}_2(\mathbb{R}^d)$; therefore, in this section we will need to impose stronger assumptions, guaranteeing that the comparison principle holds for all the finite horizontal approximations of our infinite horizon problem. Following [14], assume that:

Assumptions (H_{un}).

- ($H_{un}(1)$) The Polish space A is a compact subset of an Euclidean space;
- ($H_{un}(2)$) The function σ is independent of the measure argument, i.e. $\sigma : \mathbb{R}^d \times A \rightarrow \mathbb{R}^{d \times m}$;
- ($H_{un}(3)$) The functions b, σ, f are jointly continuous in all their arguments (with respect to the W_1 distance in the measure argument);
- ($H_{un}(4)$) The functions b, σ and f are Lipschitz in (x, μ) uniformly with respect to a in the W_1 Wasserstein distance, i.e. there exists a constant $L' > 0$ such that

$$|b(x, \mu, a) - b(x', \mu', a)| + |\sigma(x, a) - \sigma(x', a)| \\ + |f(x, \mu, a) - f(x', \mu', a)| \leq L'(|x - x'| + W_1(\mu, \mu'))$$

for all $x, x' \in \mathbb{R}^d$, $\mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$ and $a \in A$.

- ($H_{un}(5)$) The functions b, σ and f are bounded, i.e. there exists $C' > 0$ such that

$$|b(x, \mu, a)| + |\sigma(x, a)| + |f(x, \mu, a)| \leq C'$$

for all $x \in \mathbb{R}^d$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $a \in A$;

- ($H_{un}(6)$) For any fixed $a \in A$, the function $\sigma(\cdot, a)$ belongs to $C^2(\mathbb{R}^d)$; moreover, there exists a constant $\Sigma > 0$ such that

$$|\partial_{x_i} \sigma(x, a)| + |\partial_{x_i x_j}^2 \sigma(x, a)| \leq \Sigma$$

for all $(x, a) \in \mathbb{R}^d \times A$ and all $i = 1, \dots, d$, $j = 1, \dots, m$;

- ($H_{uc}(7)$) The discount rate $\beta > 0$ is sufficiently large.

Remark 1.5.13. Notice that the W_1 -Lipschitz continuity of all the coefficients in Assumption ($H_{un}(4)$) is the main novelty introduced by [36] in order to correct the uniqueness result presented in [43], as explained in Remark 5.1 [36]. The lower bound for the admissible value of β will be made explicit in the proof of the following Lemma. $\square$

Lemma 1.5.14. Under assumptions $(\mathbf{H}_{un})$, the value function v is W_1 -Lipschitz, i.e. there exists a constant $\tilde{L} > 0$ such that, for any $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$,

$$|v(\mu) - v(\nu)| \leq \tilde{L}W_1(\mu, \nu).$$

Proof. The proof is based on the following claim: there exists a constant $\tilde{L} \geq 0$ such that, for any $\xi, \eta \in L^2(\Omega, \mathcal{G}, \mathbb{P})$,

$$|J(0, \xi, \alpha) - J(0, \eta, \alpha)| \leq \tilde{L} \|\xi - \eta\|_{L^1} \text{ for any } \alpha \in \mathcal{A}.$$

As a matter of fact, from this inequality we easily deduce that, for any $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$,

$$\begin{aligned} |v(\mu) - v(\nu)| &= \inf\{|V(0, \xi) - V(0, \eta)| : \xi, \eta \in L^2(\Omega, \mathcal{G}, \mathbb{P}) \text{ with } \mathbb{P}_\xi = \mu, \mathbb{P}_\eta = \nu\} \\ &\leq \inf\{\tilde{L} \|\xi - \eta\|_{L^1} : \xi, \eta \in L^2(\Omega, \mathcal{G}, \mathbb{P}) \text{ with } \mathbb{P}_\xi = \mu, \mathbb{P}_\eta = \nu\} \\ &= \tilde{L}W_1(\mu, \nu) \end{aligned}$$

as $V(0, \cdot)$ is law invariant and $\mathcal{G}$ is a sufficiently rich σ -algebra, and the thesis holds.

In order to prove the claim, fix any two initial conditions $\xi, \eta \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ and denote the solution to the state equation (1.2) starting at time $t = 0$ from ξ and η , respectively, by X^ξ and X^η . By standard estimates relying on the Burkholder-Davis-Gundy inequality (see e.g. Theorem 5.16 in [65]) and $(\mathbf{H}_{un})$,

$$\mathbb{E} \left[\sup_{0 \leq r \leq s} |X_r^\xi - X_r^\eta| \right] \leq 2 \|\xi - \eta\|_{L^1}. \quad (1.62)$$

for $0 \leq s \leq s_0 := s_0(L', C_1)$, where C_1 is the constant appearing in Burkholder-Davis-Gundy inequality for $p = 1$. Iterating (1.62), we deduce that

$$\mathbb{E} \left[\sup_{0 \leq r \leq ks_0} |X_r^\xi - X_r^\eta| \right] \leq 2^k \|\xi - \eta\|_{L^1}. \quad (1.63)$$

Now, for fixed $s \geq 0$, we can always find an integer k such that $(k-1)s_0 \leq s \leq ks_0$. As a consequence,

$$\mathbb{E} \left[\sup_{0 \leq r \leq s} |X_r^\xi - X_r^\eta| \right] \leq e^{\bar{C}_0(s+1)} \|\xi - \eta\|_{L^1} \quad (1.64)$$

with $\bar{C}_0 = \bar{C}_0(s_0)$. Therefore, by $(\mathbf{H}_{un})$,

$$\begin{aligned} |J(0, \xi, \alpha) - J(0, \eta, \alpha)| &\leq \mathbb{E} \left[\int_0^{+\infty} e^{-\beta s} L' \left(|X_s^\xi - X_s^\eta| + W_1(\mathbb{P}_{X_s^\xi}, \mathbb{P}_{X_s^\eta}) \right) ds \right] \quad (\text{by (1.64)}) \\ &\leq 2L' e^{\bar{C}_0} \int_0^{+\infty} e^{(-\beta + \bar{C}_0)s} ds \|\xi - \eta\|_{L^1} \quad (\text{if } \beta > \bar{C}_0) \\ &= \tilde{L} \|\xi - \eta\|_{L^1}, \end{aligned}$$

where $\tilde{L}$ is a finite constant independent of α , provided that $\beta > \bar{C}_0$. $\square$

We now state and prove the main result of this section, which characterizes the value function v as the unique viscosity solution to equation (1.55) in a specific class of functions.

Theorem 1.5.15. Under assumptions $(\mathbf{H}_{un})$, the value function v of our control problem is the unique bounded and W_1 -Lipschitz viscosity solution to equation (1.55).

Proof. Notice that under Assumptions $(\mathbf{H}_{un})$ v is W_2 -continuous, bounded and also W_1 -Lipschitz, see Corollary 1.4.6, Proposition 1.4.7, and Lemma 1.5.14. Furthermore, by Theorem 1.5.11, it is sufficient to prove the following: for any bounded and W_1 -Lipschitz viscosity subsolution u to Hamilton-Jacobi-Bellman equation (1.55), it holds that

$$v(\mu) \geq u(\mu) \quad \forall \mu \in \mathcal{P}_2(\mathbb{R}^d) \quad (1.65)$$

and, for any bounded and W_1 -Lipschitz viscosity supersolution w to Hamilton-Jacobi-Bellman equation (1.55), it holds that

$$w(\mu) \geq v(\mu) \quad \forall \mu \in \mathcal{P}_2(\mathbb{R}^d). \quad (1.66)$$

Fix any bounded and W_1 -Lipschitz subsolution u to equation (1.55). We first approximate v with a suitable sequence of value functions of finite horizon problems. For any $T > 0$, we aim at maximizing the gain functional

$$J(t, \xi, \alpha) := \mathbb{E} \left[\int_t^T e^{-\beta s} f(X_s, \mathbb{P}_{X_s}, \alpha_s) ds \right] + e^{-\beta T} u(\mathbb{P}_{X_T}) \quad (1.67)$$

over the feasible controls belonging to $\mathcal{A}^T := \{\alpha : [0, T] \times \Omega \rightarrow A : \alpha \text{ is } \mathbb{F}\text{-predictable}\}$ under the dynamics

$$\begin{cases} dX_s = b(X_s, \mathbb{P}_{X_s}, \alpha_s) ds + \sigma(X_s, \alpha_s) dB_s & s \in [t, T] \\ X_t = \xi \end{cases} \quad (1.68)$$

for $t \in [0, T]$ and $\xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$. We define the lifted value function for our finite horizon control problem as $V_T(t, \xi) := \sup_{\alpha \in \mathcal{A}^T} J(t, \xi, \alpha)$; by Theorem 3.6 in [42], the function V_T is law invariant and we can thus define the true value function of the problem as

$$v_T(t, \mu) := V_T(t, \xi) \quad \text{for any } \xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P}) \text{ having law } \mu.$$

It is clear that, for any $\alpha \in \mathcal{A}$, its restriction $\alpha|_{[0, T]}$ belongs to $\mathcal{A}^T$. On the other hand, by a monotone class reasoning, it can be proved that for any $\tilde{\alpha} \in \mathcal{A}^T$ there exists an extension $\text{ext}(\tilde{\alpha}) : [0, +\infty) \times \Omega \rightarrow A \in \mathcal{A}$; with a slight abuse, we can thus identify $\mathcal{A}^T$ with the set $\{\alpha|_{[0, T]} : \alpha \in \mathcal{A}\}$, which implies

$$\begin{aligned} |v(\mu) - v_T(0, \mu)| &\leq \sup_{\alpha \in \mathcal{A}} \left\{ \mathbb{E} \left[\int_T^{+\infty} e^{-\beta s} |f(X_s, \mathbb{P}_{X_s}, \alpha_s)| ds \right] + \left| e^{-\beta T} u(\mathbb{P}_{X_T}) \right| \right\} \\ &\leq \tilde{C} \left(\frac{1}{\beta} + 1 \right) e^{-\beta T} \xrightarrow{T \rightarrow +\infty} 0 \quad \forall \mu \in \mathcal{P}_2(\mathbb{R}^d), \end{aligned}$$

for some constant $\tilde{C}$, where we used the boundedness of u and of f . Thus, the sequence $v_T(0, \cdot)$ approximates v .

We now recall that, by Theorem 4.2 in [14], v_T solves in the viscosity sense the following parabolic Hamilton-Jacobi-Bellman equation:

$$\begin{cases} -\partial_t w(t, \mu) = \int_{\mathbb{R}^d} \left[\sup_{a \in A} \left\{ e^{-\beta t} f(x, \mu, a) + \langle b(x, \mu, a), \partial_\mu w(t, \mu)(x) \rangle \right. \right. \\ \left. \left. + \frac{1}{2} \text{tr}(\sigma(x, a) \sigma^T(x, a) \partial_x \partial_\mu w(t, \mu)(x)) \right\} \right] \mu(dx) & (t, \mu) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \\ w(T, \mu) = e^{-\beta T} u(\mu) & \mu \in \mathcal{P}_2(\mathbb{R}^d) \end{cases} \quad (1.69)$$

Moreover, as $e^{-\beta t} f(x, \mu, a)$ and $g(\mu) := e^{-\beta T} u(\mu)$ satisfy Assumptions **(A)** in [14], by the comparison principle in the parabolic case, i.e. Theorem 4.3 in [14],

$$v_T(t, \mu) \geq U_T(t, \mu) \quad \text{for any } t \in [0, T], \mu \in \mathcal{P}_2(\mathbb{R}^d) \quad (1.70)$$

for any W_1 -continuous and bounded subsolution U_T to equation (1.69).

We now want to construct an appropriate subsolution to (1.69). Consider

$$U_T(t, \mu) := e^{-\beta t} u(\mu)$$

for all $t \in [0, T]$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Notice that $U_T(0, \mu) = u(\mu)$ for any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Moreover, for any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $U_T(\cdot, \mu) \in C^\infty(0, T)$ with $\partial_t U_T(t, \mu) = -\beta U_T(t, \mu)$. We prove that U_T is a viscosity subsolution to the parabolic equation (1.69) (where the definition is analogous to the definition in the elliptic case, see Definition 2.5 in [14]). Firstly, U_T is bounded, $([0, T], |\cdot|) \times (\mathcal{P}_2(\mathbb{R}^d), W_1)$ -continuous and the boundary condition is verified as $U_T(T, \mu) = e^{-\beta T} u(\mu)$ for any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Fix a W_2 -continuous function ψ such that the derivatives $\partial_t \psi$, $\partial_\mu \psi$, $\partial_x \partial_\mu \psi$ exist and are jointly continuous and there exists a constant C_ψ such that

$$|\partial_t \psi(t, \mu)| + |\partial_x \partial_\mu \psi(t, \mu)(x)| \leq C_\psi \quad \text{and} \quad |\partial_\mu \psi(t, \mu)(x)| \leq C_\psi(1 + |x|)$$

for any $(t, \mu, x) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$. Now fix a point $(\bar{t}, \bar{\mu}) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^d)$ and assume that $(U_T - \psi)(\bar{t}, \bar{\mu}) = 0$ and $(\bar{t}, \bar{\mu})$ is a point of maximum for the function $U_T - \psi$, i.e.

$$U_T(\bar{t}, \bar{\mu}) - \psi(\bar{t}, \bar{\mu}) \geq U_T(t, \mu) - \psi(t, \mu) \quad \forall t \in (0, T), \mu \in \mathcal{P}_2(\mathbb{R}^d). \quad (1.71)$$

Then, ψ is a test function for the subsolution property of the parabolic HJB equation (1.69). Multiplying (1.71) by $e^{\beta \bar{t}}$ and choosing $t = \bar{t}$, we deduce that

$$u(\bar{\mu}) - e^{\beta \bar{t}} \psi(\bar{t}, \bar{\mu}) \geq u(\mu) - e^{\beta \bar{t}} \psi(\bar{t}, \mu) \quad \forall \mu \in \mathcal{P}_2(\mathbb{R}^d).$$

Call $\phi(\mu) := e^{\beta \bar{t}} \psi(\bar{t}, \mu)$; notice that it is a function belonging to $\mathcal{TF}(\mathbb{R}^d)$ such that $u - \phi$ has a point of maximum in $\bar{\mu}$ and $(u - \phi)(\bar{\mu}) = 0$. Thus, ϕ is a test function for the subsolution property of the elliptic Hamilton-Jacobi-Bellman equation (1.55).

By hypothesis, u is a subsolution to the elliptic Hamilton-Jacobi-Bellman equation (1.55), which means that the following inequality is satisfied:

$$-\beta u(\bar{\mu}) + \int_{\mathbb{R}^d} \sup_{a \in A} \left(f(x, \bar{\mu}, a) + \langle b(x, \bar{\mu}, a), \partial_\mu \phi(\bar{\mu})(x) \rangle \right. \\ \left. + \frac{1}{2} \text{tr}(\sigma(x, a) \sigma^T(x, a) \partial_x \partial_\mu \phi(\bar{\mu})(x)) \right) \mu(dx) \geq 0$$

$$+ \frac{1}{2} \text{tr}(\sigma(x, a) \sigma^T(x, a) \partial_x \partial_\mu \phi(\bar{\mu})(x)) \Big) \bar{\mu}(dx) \geq 0.$$

If we multiply by $e^{-\beta \bar{t}}$, by the definition of ϕ we deduce that

$$\begin{aligned} 0 &\leq -\beta e^{-\beta \bar{t}} u(\bar{\mu}) + \int_{\mathbb{R}^d} \sup_{a \in A} \left(f(x, \bar{\mu}, a) e^{-\beta \bar{t}} + \langle b(x, \bar{\mu}, a), e^{-\beta \bar{t}} \partial_\mu \phi(\bar{\mu})(x) \rangle \right. \\ &\quad \left. + \frac{1}{2} \text{tr}(\sigma(x, a) \sigma^T(x, a) \partial_x \partial_\mu e^{-\beta \bar{t}} \phi(\bar{\mu})(x)) \right) \bar{\mu}(dx) \\ &= -\beta U_T(\bar{t}, \bar{\mu}) + \int_{\mathbb{R}^d} \sup_{a \in A} \left(f(x, \bar{\mu}, a) e^{-\beta \bar{t}} + \langle b(x, \bar{\mu}, a), \partial_\mu \psi(\bar{t}, \bar{\mu})(x) \rangle \right. \\ &\quad \left. + \frac{1}{2} \text{tr}(\sigma(x, a) \sigma^T(x, a) \partial_x \partial_\mu \psi(\bar{t}, \bar{\mu})(x)) \right) \bar{\mu}(dx) \\ &= \partial_t U_T(\bar{t}, \bar{\mu}) + \int_{\mathbb{R}^d} \sup_{a \in A} \left(f(x, \bar{\mu}, a) e^{-\beta \bar{t}} + \langle b(x, \bar{\mu}, a), \partial_\mu \psi(\bar{t}, \bar{\mu})(x) \rangle \right. \\ &\quad \left. + \frac{1}{2} \text{tr}(\sigma(x, a) \sigma^T(x, a) \partial_x \partial_\mu \psi(\bar{t}, \bar{\mu})(x)) \right) \bar{\mu}(dx) \\ &= \partial_t \psi(\bar{t}, \bar{\mu}) + \int_{\mathbb{R}^d} \sup_{a \in A} \left(f(x, \bar{\mu}, a) e^{-\beta \bar{t}} + \langle b(x, \bar{\mu}, a), \partial_\mu \psi(\bar{t}, \bar{\mu})(x) \rangle \right. \\ &\quad \left. + \frac{1}{2} \text{tr}(\sigma(x, a) \sigma^T(x, a) \partial_x \partial_\mu \psi(\bar{t}, \bar{\mu})(x)) \right) \bar{\mu}(dx), \end{aligned}$$

where in the last equality we used the fact that $(\bar{t}, \bar{\mu})$ is a point of maximum for the function $U_T - \psi$, which is of class $\mathcal{C}^1$ with respect to the time variable. By the arbitrariness of $(\bar{t}, \bar{\mu})$ and ψ , U_T is a viscosity subsolution to the parabolic Hamilton Jacobi Bellman equation (1.69). As a consequence, by (1.70), $U_T(t, \mu) \leq v_T(t, \mu)$ for all $t \in [0, T]$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$.

Choosing $t = 0$, we obtain

$$u(\mu) = U_T(0, \mu) \leq v_T(0, \mu) \quad \forall \mu \in \mathcal{P}_2(\mathbb{R}^d);$$

taking the limit for $T \rightarrow +\infty$ and using the convergence of $v_T(0, \cdot)$ to v , we conclude that

$$u(\mu) \leq v(\mu) \quad \forall \mu \in \mathcal{P}_2(\mathbb{R}^d).$$

Repeating a similar reasoning for supersolutions, where some modifications are needed as the class of test functions is defined over $\mathcal{P}_2(\mathbb{R}^d \times A)$, we conclude that the thesis holds. $\square$

Remark 1.5.16. The proof of our uniqueness result relies on Theorem 4.3 in [14]. Nonetheless, the same procedure could be applied in any case where a comparison principle for parabolic Hamilton-Jacobi-Bellman equations on the space $\mathcal{P}_2(\mathbb{R}^d)$ is available. Moreover, the reasoning does not depend on the specific choice of the definition of viscosity supersolution: it would still hold for a definition à la Crandall-Lions, provided that a uniqueness result applies for the analogue definition of supersolution to a parabolic HJB. $\square$

Chapter 2

Ergodic control of McKean-Vlasov systems on the Wasserstein space

2.1 Introduction

This chapter is mainly devoted to the study of a different optimal control problem of the McKean-Vlasov type with infinite time horizon: here, the gain functional will not be in discounted form, as in the previous chapter, but in ergodic (long time average) form. In fact, we will also provide interesting connections with the limit behaviour of the value functions associated with other classes of optimal control problems of the McKean-Vlasov type. The chapter presents the content of the paper [56] (joint work with Marco Fuhrman).

The state equations that we will consider are of the form

$$\begin{cases} dX_s = b(X_s, \mathbb{P}_{X_s}, \alpha_s) ds + \sigma(X_s, \mathbb{P}_{X_s}, \alpha_s) dB_s, & s \in [t, T] \subset [0, T], \\ X_t = \xi \end{cases} \quad (2.1)$$

in the case of a finite time horizon $T > 0$, or

$$\begin{cases} dX_t = b(X_t, \mathbb{P}_{X_t}, \alpha_t) dt + \sigma(X_t, \mathbb{P}_{X_t}, \alpha_t) dB_t, & t \geq 0, \\ X_0 = \xi, \end{cases} \quad (2.2)$$

in the case of infinite time horizon. In these equations, B is a Brownian motion in $\mathbb{R}^d$, ξ a random initial condition and α is a control process, taking values in a space A and chosen in a suitable class $\mathcal{A}$. On the coefficients b, σ , we first impose standard Lipschitz and linear growth conditions, to ensure well-posedness of the state equation.

We briefly recall that, as already mentioned in the previous chapter, optimal control problems for McKean-Vlasov systems have been studied by many authors in the finite horizon case, namely for a reward functional of the form

$$J^T(t, \xi, \alpha) = \mathbb{E} \left[\int_t^T f(X_s^{t, \xi, \alpha}, \mathbb{P}_{X_s^{t, \xi, \alpha}}, \alpha_s) ds + g(X_T^{t, \xi, \alpha}, \mathbb{P}_{X_T^{t, \xi, \alpha}}) \right],$$

where $X^{t,\xi,\alpha}$ denotes the solution to (2.1) and f and g both satisfy some suitable growth conditions.

The value function of this optimal control problem can be initially defined setting

$$V^T(t, \xi) = \sup_{\alpha \in \mathcal{A}} J^T(t, \xi, \alpha); \quad (2.3)$$

namely, V^T is a function of $t \in [0, T]$ and $\xi \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ for a suitable underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$. However, this value function satisfies the so-called law invariance property, presented in [42] under suitable general assumptions and already mentioned in the previous chapter: whenever ξ and ξ' are initial conditions with the same law, then one has $V^T(t, \xi) = V^T(t, \xi')$. As a consequence, one can define a value function v^T on $[0, T] \times \mathcal{P}_2(\mathbb{R}^d)$ setting

$$v^T(t, \mu) = V^T(t, \xi), \quad \text{when} \quad \mathbb{P}_\xi = \mu. \quad (2.4)$$

In contrast, the function defined in (2.3) will be called lifted value function. It turns out that v^T is a solution to a HJB equation on $[0, T] \times \mathcal{P}_2(\mathbb{R}^d)$ of the form

$$\begin{cases} -\partial_t v^T(t, \mu) - F(\mu, \partial_\mu v^T(t, \mu), \partial_x \partial_\mu v^T(t, \mu)) = 0, & t \in [0, T], \mu \in \mathcal{P}_2(\mathbb{R}^d), \\ v^T(T, \mu) = \int_{\mathbb{R}^d} g(x, \mu) \mu(dx), \end{cases} \quad (2.5)$$

where, for suitable $\varphi : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$, we set

$$\begin{aligned} F(\mu, \partial_\mu \varphi(\mu), \partial_x \partial_\mu \varphi(\mu)) &:= \int_{\mathbb{R}^d} \left[\sup_{a \in A} \{f(x, \mu, a) + \langle b(x, \mu, a), \partial_\mu \varphi(\mu)(x) \rangle \right. \\ &\quad \left. + \frac{1}{2} \text{tr}(\sigma \sigma^T(x, \mu, a) \partial_x \partial_\mu \varphi(\mu)(x)) \} \right] \mu(dx). \end{aligned}$$

In (2.5) the symbol ∂_μ still denotes the so-called L -derivative in the sense of Lions: see the previous chapter or Section 2.5 for some reminders. As already mentioned, this equation has been studied by many authors and is presently the object of thorough investigation: see e.g. [42], [43], [17], [14].

The main goal of this chapter is to study the (infinite horizon) ergodic optimal control problem, where the reward functional to be maximized is

$$J_\infty(\xi, \alpha) = \limsup_{T \rightarrow \infty} \mathbb{E} \left[\frac{1}{T} \int_0^T f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right], \quad (2.6)$$

and $X^{\xi, \alpha}$ denotes the solution to the state equation (2.2); we assume that f is bounded in order to simplify the proofs.

Classical ergodic control problems of stochastic dynamics have been studied since long: we will only mention [20] and [6] as early contributions and we refer to the book [4] for further reading. In recent years, methods based on Backward Stochastic Differential Equations have also been used, see [54], [48], [41], [44], but we will not pursue this methodology. In the aforementioned references, the coefficients of the dynamics are usually required to satisfy suitable dissipativity or periodicity conditions, guaranteeing

that the value of the problem is equal to some constant λ and does not depend on ξ . Notice that the long-time average of reward/cost functionals has also been treated in non-ergodic cases, where the value varies with the initial condition: for instance, this class of problems is the subject of [25], where a non-expansivity condition is adopted in place of the stronger dissipativity condition; see also [27] for more references on this topic.

In the classical PDE approach to stochastic ergodic control problems, the identification of the value λ is usually coupled with the construction of a suitable function, denoted by ϕ in the sequel, which is expected to be a solution to a HJB equation of elliptic type. The pair (λ, ϕ) can be furthermore used to characterize various asymptotic properties related to optimal control problems.

In the context of McKean-Vlasov systems, this program has first been carried out in [11]. In that paper the authors construct the pair (λ, ϕ) , where the function $\phi(\xi)$ is mostly considered as a real function defined for $\xi \in L^2$, and they use the theory of viscosity solutions on a Hilbert space in order to obtain well-posedness results and investigate the ergodic and other related control problems. We also mention that some ergodic control problems with mean field effects have been considered in continuous time with specific forms of the state equation and reward functions (see e.g. [29] and the references therein) and in the context of time-discrete systems (see for instance [18] and the references therein).

Our contribution: We complement the results in [11] choosing an alternative approach and constructing the pair (λ, ϕ) where ϕ is treated as a function $\mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$. Correspondingly, we prove that ϕ is a solution in the viscosity sense to an elliptic PDE on the Wasserstein space (called ergodic HJB equation); we do not use stability properties of viscosity solutions to PDEs on Hilbert spaces but we entirely rely on a suitable functional relation (equation (2.10)). Furthermore, starting from this basic equation, we are also able to connect the original control problem with the limit behaviour of different classes of optimal control problems of the McKean-Vlasov type, both with finite and infinite time horizon. In comparison to [11], we should mention that our results are mainly complementary, since they refer to value functions and HJB equations defined on the Wasserstein space rather than on the Hilbert space L^2 . Nonetheless, our work contributes to a meaningful extension of the existing theory, since it is not yet clear if well-posedness results for viscosity solutions to a second-order PDE on the space $\mathcal{P}_2(\mathbb{R}^d)$ are immediately implied by analogous results for the lifted version of the same PDE on a suitable L^2 space (as pointed out in Remark 3.6 in [43] as well as in Remark 3.2 in [11]). Accordingly, several technical aspects are different from the ones proposed in [11]. First of all, while choosing the same formulation of the dissipativity condition, we introduce a stronger condition which helps in verifying the hypothesis directly. Furthermore, we avoid some boundedness and regularity assumptions on b, σ and we consider open-loop controls rather than Lipschitz controls in feedback form (except for our verification theorem). We now briefly describe our main results.

Firstly, we assume that the coefficients of the state equation satisfy the following dissipativity condition, already introduced in [11]: we require that there exists a

constant $\eta > 0$ such that

$$\mathbb{E} \left[\langle b(\xi, \mathbb{P}_\xi, a) - b(\xi', \mathbb{P}_{\xi'}, a), \xi - \xi' \rangle + \frac{1}{2} |\sigma(\xi, \mathbb{P}_\xi, a) - \sigma(\xi', \mathbb{P}_{\xi'}, a)|^2 \right] \leq -\eta \mathbb{E} [|\xi - \xi'|^2], \quad (2.7)$$

for every $a \in A$ and every pair of square summable $\mathbb{R}^d$ -valued random variables ξ, ξ' . As far as the assumptions are concerned, the main novelty of our work is the introduction of a second condition, which implies our assumption (2.7) but is easier to verify: we require that the inequality

$$\langle b(x, \mu, a) - b(x', \mu, a), x - x' \rangle + \frac{1}{2} |\sigma(x, \mu, a) - \sigma(x', \mu, a)|^2 \leq -\gamma |x - x'|^2, \quad (2.8)$$

holds for all $x, x' \in \mathbb{R}^d$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $a \in A$, where the constant $\gamma > 0$ has to be sufficiently large, depending on the Lipschitz constants of the coefficients. After proving that the dissipativity assumption (2.7) implies suitable ergodicity properties for the solutions to the state equation, we construct the pair (λ, ϕ) using the so-called vanishing discount method: firstly, for any $\beta > 0$ we consider the infinite horizon optimal control problem

$$J_\beta(\xi, \alpha) = \mathbb{E} \left[\int_0^\infty e^{-\beta t} f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right]. \quad (2.9)$$

It was proved in Theorem 4.3 in [84] (i.e. in Theorem 1.4.3 in the previous chapter) that the corresponding value only depends on ξ via its law, so that one can in fact define a value function $v^\beta : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ setting

$$v^\beta(\mu) = \sup_{\alpha \in A} J_\beta(\xi, \alpha), \quad \text{where} \quad \mathbb{P}_\xi = \mu.$$

By compactness arguments we prove that for a sequence $\beta_n \rightarrow 0$ the limits

$$\lambda = \lim_{n \rightarrow \infty} \beta_n v^{\beta_n}(\delta_0), \quad \phi(\mu) = \lim_{n \rightarrow \infty} (v^{\beta_n}(\mu) - v^{\beta_n}(\delta_0)),$$

exist, and the function $\phi : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ is Lipschitz. Many of our results are based on the noteworthy equality

$$\phi(\mu) + \lambda T = \sup_{\alpha \in A} \left\{ \mathbb{E} \left[\int_0^T f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right] + \phi(\mathbb{P}_{X_T^{\xi, \alpha}}) \right\}, \quad (2.10)$$

where $\mu = \mathbb{P}_\xi$, derived starting from the dynamic programming principle for the discounted infinite horizon optimal control problem (Corollary 4.3.1 in [84] and Corollary 1.4.5 in this thesis), by an argument which is perhaps new. Up to the term λT , this identifies the function ϕ with the value function of an optimal control problem where the terminal reward is ϕ itself. The use of this formula, or some related ones, simplifies many arguments and, in particular, it allows to study the so-called ergodic HJB equation on the Wasserstein space

$$\lambda = \int_{\mathbb{R}^d} \left[\sup_{a \in A} \left\{ f(x, \mu, a) + \langle b(x, \mu, a), \partial_\mu \varphi(\mu)(x) \rangle + \frac{1}{2} \text{tr} (\sigma \sigma^T(x, \mu, a) \partial_x \partial_\mu \varphi(\mu)(x)) \right\} \right] \mu(dx), \quad (2.11)$$

by a direct application of known results on HJB for the finite horizon case. Adopting this approach, we avoid the usual deduction of existence of a solution to the ergodic HJB based on stability properties of viscosity solutions to the elliptic HJB, which would require new techniques and a detailed study in the case of equations on the Wasserstein space. The pair (λ, ϕ) is subsequently proved to be connected to various optimization problems: we first show that (λ, ϕ) can be used to characterize the long time behavior, as the time horizon T tends to infinity, of the value function v^T introduced in (2.4), namely

$$|v^T(0, \mu) - \phi(\mu) - \lambda T| \leq C (1 + W_2(\mu, \delta_0)^2), \quad \mu \in \mathcal{P}_2(\mathbb{R}^d), \quad (2.12)$$

where W_2 denotes the 2-Wasserstein distance, which implies

$$\lim_{T \rightarrow \infty} \frac{v^T(0, \mu)}{T} = \lambda, \quad \mu \in \mathcal{P}_2(\mathbb{R}^d). \quad (2.13)$$

Another result concerns the study of the limits

$$\lim_{\beta \rightarrow 0} \sup_{\alpha \in \mathcal{A}} \mathbb{E} \left[\beta \int_0^\infty e^{-\beta t} f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right], \quad \lim_{T \rightarrow \infty} \sup_{\alpha \in \mathcal{A}} \mathbb{E} \left[\frac{1}{T} \int_0^T f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right]. \quad (2.14)$$

where $X^{\xi, \alpha}$ is again the solution to (2.2). Note that both functionals can be written in the form

$$\mathbb{E} \left[\int_0^\infty k(t) f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right]$$

where k is the probability density $k(t) = \beta e^{-\beta t}$ or $k(t) = T^{-1} 1_{[0, T]}(t)$ respectively, and in both cases $k(t)$ tends to zero uniformly on $[0, \infty)$. Our main result states that both limits in (2.14) exist and that they are both equal to the constant λ (the left limit is close to the definition of λ , but it is not computed along a subsequence β_n). In particular, they do not depend on the initial condition ξ . This kind of result is sometimes called a (nonlinear) Abelian-Tauberian theorem. Finally, we show that λ coincides with the value of the ergodic control problem with reward functional (2.6). In order to prove this result we do not need to assume that an optimal control exists. Existence of an optimal control is proved under additional assumptions in the form of a verification theorem applied to the ergodic HJB equation.

This chapter is organized as follows. After some preliminaries in Section 2.2, in Section 2.3 we present the aforementioned dissipativity assumption (2.7), see Assumptions **(H₂)**. Since the direct verification of (2.7) may be difficult, in Proposition 2.3.2 we show that it is a consequence of the stronger requirement (2.8). These dissipativity conditions imply suitable ergodicity properties for the solutions to (2.2): see Proposition 2.3.1. In Section 2.4 we construct a pair (λ, ϕ) using the vanishing discount method, i.e. setting

$$\lambda = \lim_{n \rightarrow \infty} \beta_n v^{\beta_n}(\delta_0), \quad \phi(\mu) = \lim_{n \rightarrow \infty} (v^{\beta_n}(\mu) - v^{\beta_n}(\delta_0)),$$

for a suitable sequence $\beta_n \rightarrow 0$. We then prove the fundamental equation (2.10) and a generalization of it starting from a time $t > 0$ in Propositions 2.4.2 and 2.4.4,

respectively. As a first application of these results, we study the connection of the pair (λ, ϕ) with the long-time behaviour of the value function associated with a finite horizon optimal control problem of the McKean-Vlasov type when the time horizon tends to infinity (Theorem 2.4.5) and, in parallel, with the limit behaviour of the value function associated with an infinite time horizon optimal control problem of the McKean-Vlasov type when the discount factor tends to 0, leading to the Abelian-Tauberian Theorem 2.4.6. In Section 2.5, after some reminders on L -derivatives and viscosity solutions on the Wasserstein space, we prove the result, announced above, stating that (λ, ϕ) is a viscosity solution to the ergodic HJB equation on the Wasserstein space, see Theorem 2.5.5. Moreover, in Theorem 2.5.7, we prove that the value of λ is uniquely determined. In Proposition 2.5.6 we also show that the properties (2.12)-(2.13) can be recast into a result on asymptotic behavior of viscosity solutions to a parabolic PDE of HJB type written forward in time. We underline that adding an arbitrary constant to the function ϕ yields another solution to (2.11), but even in classical cases it is known that additional assumptions are required in order to obtain uniqueness up to an additive constant: see the discussion at the end of Section 2.5 and also Remark 2.6.7 for more details. Finally, in Section 2.6 we turn to the ergodic control problem with reward functional (2.6) and we prove that λ coincides with its value, see Proposition 2.6.1. We also discuss optimality of controls (Proposition 2.6.3) and we present a verification theorem applied to the ergodic HJB equation, see Theorem 2.6.6.

We mention that some papers considering stochastic control of McKean-Vlasov systems on finite horizon deal with a more general case, also including dependence on the marginal laws of the control process in the coefficients of the controlled equation and the reward functional: for instance, $b(X_s, \mathbb{P}_{X_s}, \alpha_s)$ becomes $b(X_s, \mathbb{P}_{X_s}, \alpha_s, \mathbb{P}_{\alpha_s})$. We believe that many of our results extend to this more general case, under natural assumptions for the extended coefficients. However, since for our construction of the basic pair (λ, ϕ) we rely on the vanishing discount method, we need corresponding extended results for the optimal control problem on infinite horizon. However, we could not find any such references in the literature. Developing such preliminary results would take a long time and may constitute the object of a separate research project. For this reason we prefer to concentrate on a simpler basic case where all the essential arguments are presented.

2.2 Notation and preliminaries

In the Euclidean space $\mathbb{R}^d$ the standard norm and scalar product are denoted by $|\cdot|$, $\langle \cdot, \cdot \rangle$. For a matrix $\sigma \in \mathbb{R}^{d \times m}$ we set $|\sigma| = (tr(\sigma\sigma^T))^{1/2}$ where tr denotes the trace and σ^T the transpose matrix. For $q \geq 1$ we denote by $\mathcal{P}_q(\mathbb{R}^d)$ the space of Borel probability measures on $\mathbb{R}^d$ with finite moment of order q equipped with the Wasserstein metric W_q (see for instance [32] or [90]).

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space, where an m -dimensional Brownian motion $B = (B_t)_{t \geq 0}$ is defined. We assume that there exists a random variable U independent of B with uniform distribution on $(0, 1)$. We denote by $\mathcal{G}$ the σ -algebra generated by U . We will require initial conditions of controlled equations to

be $\mathcal{G}$ -measurable. Recall that the previous assumptions imply that for every Borel probability μ on $\mathbb{R}^d$ there exists a $\mathcal{G}$ -measurable random variable ξ such that $\mu = \mathbb{P}_\xi$, where the latter symbol denotes the law of ξ : a proof of this classical result can be found in [42], Lemma 2.1.

We denote by $(\mathcal{F}_t^B)_{t \geq 0}$ the $\mathbb{P}$ -completion of the filtration generated by B and we set $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ where $\mathcal{F}_t = \sigma(\mathcal{G} \cup \mathcal{F}_t^B)$. Let A be a Polish space, endowed with its Borel σ -algebra, which we identify with the space of control actions. We define the set $\mathcal{A}$ of admissible controls as the set of A -valued $\mathbb{F}$ -predictable processes $(\alpha_t)_{t \geq 0}$.

Dynamics of the system

From now on, we will assume that the coefficients b, σ of the state equation satisfy the following assumptions:

Assumptions ($\mathbf{H}_1$). (i) The functions

$$b : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times A \rightarrow \mathbb{R}^d, \quad \sigma : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times A \rightarrow \mathbb{R}^{d \times m},$$

are Borel measurable and there exist positive constants $L_{bx}, L_{b\mu}, L_{\sigma x}, L_{\sigma\mu}, M$ such that

$$\begin{aligned} |b(x, \mu, a) - b(x', \mu, a)| &\leq L_{bx} |x - x'|, \\ \text{and } |b(x, \mu, a) - b(x, \mu', a)| &\leq L_{b\mu} W_2(\mu, \mu'), \\ |\sigma(x, \mu, a) - \sigma(x', \mu, a)| &\leq L_{\sigma x} |x - x'|, \\ \text{and } |\sigma(x, \mu, a) - \sigma(x, \mu', a)| &\leq L_{\sigma\mu} W_2(\mu, \mu'), \end{aligned} \tag{2.15}$$

$$|b(0, \delta_0, a)| + |\sigma(0, \delta_0, a)| \leq M \tag{2.16}$$

for all $x, x' \in \mathbb{R}^d, \mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d), a \in A$.

(ii) For all fixed $(x, \mu) \in \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, $b(x, \mu, \cdot)$ and $\sigma(x, \mu, \cdot)$ are continuous.

For $t \in [0, T]$, $\xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$, $\alpha \in \mathcal{A}$ let us consider the state equation

$$\begin{cases} dX_s = b(X_s, \mathbb{P}_{X_s}, \alpha_s) ds + \sigma(X_s, \mathbb{P}_{X_s}, \alpha_s) dB_s, & s \in [t, T], \\ X_t = \xi. \end{cases} \tag{2.17}$$

We have the following standard existence and uniqueness result. The case $q = 2$ is well known (see e.g. [32] Theorem 4.21), but we could not find a reference for the general case $q \geq 2$, for which we sketch a proof.

Proposition 2.2.1. For $q \geq 2$, $0 \leq t \leq T$, $\xi \in L^q(\Omega, \mathcal{F}_t, \mathbb{P})$, $\alpha \in \mathcal{A}$, there exists a unique (up to indistinguishability) continuous adapted solution to (2.17) satisfying $\mathbb{E}[\sup_{s \in [t, T]} |X_s|^q] < \infty$. Moreover, there exists a constant C , depending only on $T - t, q$ and the constants in Assumptions ($\mathbf{H}_1$), such that

$$\mathbb{E} \left[\sup_{s \in [t, T]} |X_s|^q \right] \leq C (1 + \mathbb{E}[|\xi|^q]).$$

Proof. We introduce the Banach space $\mathcal{S}^q$ of continuous adapted processes on $[t, T]$ such that the norm $\|X\| = (\mathbb{E} [\sup_{s \in [t, T]} |X_s|^q])^{1/q}$ is finite. Solutions to (2.17) coincide with fixed points of the mapping $\Phi : \mathcal{S}^q \rightarrow \mathcal{S}^q$ defined as

$$\Phi(X)_s = \xi + \int_t^s b(X_r, \mathbb{P}_{X_r}, \alpha_r) dr + \int_t^s \sigma(X_r, \mathbb{P}_{X_r}, \alpha_r) dB_r, \quad s \in [t, T].$$

We prove that Φ is a contraction in $\mathcal{S}^q$, leaving aside the verification that Φ is well defined. For $X, X' \in \mathcal{S}^q$, by the Burkholder-Davis-Gundy inequality, we have

$$\begin{aligned} \mathbb{E} \left[\sup_{s \in [t, T]} |\Phi(X)_s - \Phi(X')_s|^q \right] &\leq C \mathbb{E} \left[\left(\int_t^T |b(X_s, \mathbb{P}_{X_s}, \alpha_s) - b(X'_s, \mathbb{P}_{X'_s}, \alpha_s)| ds \right)^q \right] \\ &\quad + C \mathbb{E} \left[\left(\int_t^T |\sigma(X_s, \mathbb{P}_{X_s}, \alpha_s) - \sigma(X'_s, \mathbb{P}_{X'_s}, \alpha_s)|^2 ds \right)^{q/2} \right] \\ &=: I_1 + I_2. \end{aligned}$$

for a constant C depending only on q . Letting $\bar{X} = X - X'$ and using the Lipschitz property of σ we have, for some constant C that may differ from line to line,

$$\begin{aligned} I_2 &\leq C \mathbb{E} \left[\left(\int_t^T (|\bar{X}_s|^2 + W_2(\mathbb{P}_{X_s}, \mathbb{P}_{X'_s})^2) ds \right)^{q/2} \right] \\ &\leq C \mathbb{E} \left[\left(\int_t^T |\bar{X}_s|^2 ds \right)^{q/2} \right] + C \left(\int_t^T \mathbb{E} [|\bar{X}_s|^2] ds \right)^{q/2} \\ &\leq C (T - t)^{q/2} \mathbb{E} \left[\sup_{s \in [t, T]} |\bar{X}_s|^q \right] \\ &= C (T - t)^{q/2} \|\bar{X}\|^q. \end{aligned}$$

In a similar way one proves that $I_1 \leq C (T - t)^q \|\bar{X}\|^q$ and the required contraction property holds provided $T - t \leq \delta$ for a suitable δ . The general case follows partitioning the interval $[t, T]$ into subintervals of length $\leq \delta$. $\square$

We thus denote by $X^{t, \xi, \alpha} = (X_s^{t, \xi, \alpha})_{s \in [t, T]}$ the solution to (2.17). Notice that, with analogous computations, it is also easy to verify that, given two solutions $X = X^{t, \xi, \alpha}, X' = X^{t, \xi', \alpha}$ corresponding to the initial conditions $\xi, \xi' \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$, respectively, we have

$$\mathbb{E} \left[\sup_{s \in [t, T]} |X_s - X'_s|^2 \right] \leq C \mathbb{E} [|\xi - \xi'|^2], \quad (2.18)$$

for some constant C that does not depend on α .

We will often consider the state equation on the time horizon $[0, \infty)$:

$$\begin{cases} dX_t = b(X_t, \mathbb{P}_{X_t}, \alpha_t) dt + \sigma(X_t, \mathbb{P}_{X_t}, \alpha_t) dB_t, & t \geq 0 \\ X_0 = \xi, \end{cases} \quad (2.19)$$

for $q \geq 2$, $\xi \in L^q(\Omega, \mathcal{G}, \mathbb{P})$, $\alpha \in \mathcal{A}$. The solution is well defined on every bounded interval $[0, T]$ and thus for all $t \geq 0$. We will denote by $X^{\xi, \alpha}$ the solution; notice that, for every fixed $T > 0$, this process coincides with the solution $X^{0, \xi, \alpha}$ to the state equation (2.17) on $[0, T]$. In what follows, we will thus use both notations.

2.3 Ergodicity

We start by stating the dissipativity condition which was presented in the Introduction.

Assumptions ($\mathbf{H}_2$). There exists a constant $\eta > 0$ such that

$$\mathbb{E} \left[\langle b(\xi, \mathbb{P}_\xi, a) - b(\xi', \mathbb{P}_{\xi'}, a), \xi - \xi' \rangle + \frac{1}{2} |\sigma(\xi, \mathbb{P}_\xi, a) - \sigma(\xi', \mathbb{P}_{\xi'}, a)|^2 \right] \leq -\eta \mathbb{E} [|\xi - \xi'|^2], \quad (2.20)$$

for every $a \in A$ and every pair of square summable $\mathbb{R}^d$ -valued random variables ξ, ξ' .

Later in this section we comment on the applicability of (2.20). The main consequence of these assumptions is the following proposition. The statement and proof of (2.21) is the same as Lemma 2.2 in [11], but we report the short proof for convenience. The inequality (2.22) is similar to the conclusion of Lemma 2.1 in [11], but we dispense with the boundedness conditions on b, σ imposed there.

Proposition 2.3.1. Under Assumptions ($\mathbf{H}_1$)-($\mathbf{H}_2$), for any two solutions X, X' to the state equation (2.19) we have

$$\mathbb{E} [|X_t - X'_t|^2] \leq \mathbb{E} [|X_0 - X'_0|^2] e^{-2\eta t}, \quad (2.21)$$

$$\mathbb{E} [|X_t|^2] \leq \mathbb{E} [|X_0|^2] e^{-\eta t} + K, \quad (2.22)$$

for every $t \geq 0$, and for a constant K that only depends on the constant η in (2.20) and on the Lipschitz and growth constants $L_{\sigma x}, L_{\sigma \mu}, M$ in (2.15) and (2.16).

Proof. By the Itô formula, for any $t \geq 0$,

$$\begin{aligned} |X_t - X'_t|^2 &= |X_0 - X'_0|^2 + \int_0^t 2 \langle X_s - X'_s, b(X_s, \mathbb{P}_{X_s}, \alpha_s) - b(X'_s, \mathbb{P}_{X'_s}, \alpha_s) \rangle ds \\ &\quad + \int_0^t 2 \langle X_s - X'_s, \sigma(X_s, \mathbb{P}_{X_s}, \alpha_s) dB_s - \sigma(X'_s, \mathbb{P}_{X'_s}, \alpha_s) dB_s \rangle \\ &\quad + \int_0^t |\sigma(X_s, \mathbb{P}_{X_s}, \alpha_s) - \sigma(X'_s, \mathbb{P}_{X'_s}, \alpha_s)|^2 ds. \end{aligned}$$

It can be easily verified that the stochastic integrals have zero expectation. Taking expectation we thus see that the function $t \mapsto \mathbb{E}[|X_t - X'_t|^2]$ is absolutely continuous and it follows that

$$\begin{aligned} \frac{d}{dt} \mathbb{E} |X_t - X'_t|^2 &= 2 \mathbb{E} \langle X_t - X'_t, b(X_t, \mathbb{P}_{X_t}, \alpha_t) - b(X'_t, \mathbb{P}_{X'_t}, \alpha_t) \rangle \\ &\quad + \mathbb{E} |\sigma(X_t, \mathbb{P}_{X_t}, \alpha_t) - \sigma(X'_t, \mathbb{P}_{X'_t}, \alpha_t)|^2 \\ &\leq -2\eta \mathbb{E} |X_t - X'_t|^2, \end{aligned} \quad (2.23)$$

where we use (2.20) in the last inequality. Therefore, (2.21) follows immediately.

Arguing as we did for (2.23) we obtain

$$\frac{d}{dt} \mathbb{E} |X_t|^2 = 2 \mathbb{E} \langle X_t, b(X_t, \mathbb{P}_{X_t}, \alpha_t) \rangle + \mathbb{E} |\sigma(X_t, \mathbb{P}_{X_t}, \alpha_t)|^2. \quad (2.24)$$

By the ergodicity condition (2.20),

$$\begin{aligned}
& 2\langle b(X_t, \mathbb{P}_{X_t}, \alpha_t), X_t \rangle + |\sigma(X_t, \mathbb{P}_{X_t}, \alpha_t)|^2 \\
&= 2\langle b(X_t, \mathbb{P}_{X_t}, \alpha_t) - b(0, \delta_0, \alpha_t), X_t \rangle + 2\langle b(0, \delta_0, \alpha_t), X_t \rangle \\
&\quad + |\sigma(X_t, \mathbb{P}_{X_t}, \alpha_t) - \sigma(0, \delta_0, \alpha_t) + \sigma(0, \delta_0, \alpha_t)|^2 \\
&\leq -2\eta|X_t|^2 + 2\langle b(0, \delta_0, \alpha_t), X_t \rangle + |\sigma(0, \delta_0, \alpha_t)|^2 + 2\langle \sigma(X_t, \mathbb{P}_{X_t}, \alpha_t) - \sigma(0, \delta_0, \alpha_t), \sigma(0, \delta_0, \alpha_t) \rangle.
\end{aligned}$$

From the Lipschitz conditions (2.15) we obtain

$$\begin{aligned}
2\langle b(X_t, \mathbb{P}_{X_t}, \alpha_t), X_t \rangle + |\sigma(X_t, \mathbb{P}_{X_t}, \alpha_t)|^2 &\leq -2\eta|X_t|^2 + 2|b(0, \delta_0, \alpha_t)| |X_t| + |\sigma(0, \delta_0, \alpha_t)|^2 \\
&\quad + 2|\sigma(0, \delta_0, \alpha_t)|(L_{\sigma x} |X_t| + L_{\sigma\mu} W_2(\mathbb{P}_{X_t}, \delta_0)).
\end{aligned}$$

Taking into account the growth condition (2.16) we have $|b(0, \delta_0, \alpha_t)| \leq M$ and $|\sigma(0, \delta_0, \alpha_t)| \leq M$. It follows that

$$\begin{aligned}
& 2\langle b(X_t, \mathbb{P}_{X_t}, \alpha_t), X_t \rangle + |\sigma(X_t, \mathbb{P}_{X_t}, \alpha_t)|^2 \\
&\leq -2\eta|X_t|^2 + 2M(1 + L_{\sigma x}) |X_t| + 2ML_{\sigma\mu} W_2(\mathbb{P}_{X_t}, \delta_0) + M^2.
\end{aligned}$$

Replacing in (2.24) and recalling that $W_2(\mathbb{P}_{X_t}, \delta_0) = (\mathbb{E}|X_t|^2)^{\frac{1}{2}}$ by definition and $\mathbb{E}|X_t| \leq (\mathbb{E}|X_t|^2)^{1/2}$ by Hölder inequality, we obtain

$$\frac{d}{dt} \mathbb{E}|X_t|^2 \leq -2\eta \mathbb{E}|X_t|^2 + C(1 + 2(\mathbb{E}|X_t|^2)^{1/2}), \quad (2.25)$$

where C only depends on $M, L_{\sigma x}, L_{\sigma\mu}$. Using the inequality $2C(\mathbb{E}|X_t|^2)^{1/2} \leq \eta \mathbb{E}|X_t|^2 + C'$, for a suitable constant C' , the required conclusion (2.22) is easily attained. $\square$

We conclude this section with some comments on the dissipativity conditions. Since Assumption **(H₂)** may be difficult to verify directly, in the following proposition we state a sufficient condition for it to hold.

Proposition 2.3.2. Suppose that Assumptions **(H₁)** hold and that there exists a constant γ such that

$$\langle b(x, \mu, a) - b(x', \mu, a), x - x' \rangle + \frac{1}{2} |\sigma(x, \mu, a) - \sigma(x', \mu, a)|^2 \leq -\gamma|x - x'|^2, \quad (2.26)$$

for all $x, x' \in \mathbb{R}^d$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $a \in A$, and

$$\eta := \gamma - \left(L_{b\mu} + L_{\sigma x} L_{\sigma\mu} + \frac{1}{2} L_{\sigma\mu}^2 \right) > 0. \quad (2.27)$$

Then Assumption **(H₂)** holds true.

We postpone the proof and we first note that (2.26)-(2.27) may be easily checked in specific cases. For instance, if b, σ satisfy Assumptions **(H₁)**, then the coefficients

$$b(x, \mu, a) - Kx, \quad \sigma(x, \mu, a),$$

satisfy (2.26)-(2.27), and hence Assumption **(H₂)**, provided that K is sufficiently large.

Remark 2.3.3. More generally, when dealing with particular models, it may be possible to verify directly that (2.21) and (2.22) hold true for some constants K and $\eta > 0$. In this case the dissipativity condition (2.20) is no longer needed. In other words, in the rest of this chapter, one may replace Assumption **(H₂)** by the inequalities (2.21) and (2.22). $\square$

We end this section with the

Proof of Proposition 2.3.2. From (2.26) it follows that, for all $x, x' \in \mathbb{R}^d$, $\mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$, $a \in A$,

$$\begin{aligned} & 2\langle b(x, \mu, a) - b(x', \mu', a), x - x' \rangle + |\sigma(x, \mu, a) - \sigma(x', \mu', a)|^2 \\ &= 2\langle b(x, \mu, a) - b(x', \mu, a), x - x' \rangle + 2\langle b(x', \mu, a) - b(x', \mu', a), x - x' \rangle \\ &\quad + |\sigma(x, \mu, a) - \sigma(x', \mu, a) + \sigma(x', \mu, a) - \sigma(x', \mu', a)|^2 \\ &\leq -2\gamma|x - x'|^2 + 2\langle b(x', \mu, a) - b(x', \mu', a), x - x' \rangle \\ &\quad + |\sigma(x', \mu, a) - \sigma(x', \mu', a)|^2 + 2\langle \sigma(x, \mu, a) - \sigma(x', \mu, a), \sigma(x', \mu, a) - \sigma(x', \mu', a) \rangle. \end{aligned}$$

From the Lipschitz conditions (2.15) we obtain

$$\begin{aligned} & 2\langle b(x, \mu, a) - b(x', \mu', a), x - x' \rangle + |\sigma(x, \mu, a) - \sigma(x', \mu', a)|^2 \\ &\leq -2\gamma|x - x'|^2 + 2L_{b\mu} W_2(\mu, \mu') |x - x'| + L_{\sigma\mu}^2 W_2(\mu, \mu')^2 + 2L_{\sigma x} L_{\sigma\mu} W_2(\mu, \mu') |x - x'|. \end{aligned}$$

Then, given $a \in A$ and $\xi, \xi' \in L^2(\Omega, \mathcal{G}, \mathbb{P})$, we deduce that

$$\begin{aligned} & \mathbb{E} \left[\langle b(\xi, \mathbb{P}_\xi, a) - b(\xi', \mathbb{P}_{\xi'}, a), \xi - \xi' \rangle + \frac{1}{2} |\sigma(\xi, \mathbb{P}_\xi, a) - \sigma(\xi', \mathbb{P}_{\xi'}, a)|^2 \right] \\ &\leq -\gamma \mathbb{E} |\xi - \xi'|^2 + L_{b\mu} W_2(\mathbb{P}_\xi, \mathbb{P}_{\xi'}) \mathbb{E} |\xi - \xi'| \\ &\quad + \frac{1}{2} L_{\sigma\mu}^2 W_2(\mathbb{P}_\xi, \mathbb{P}_{\xi'})^2 + L_{\sigma x} L_{\sigma\mu} W_2(\mathbb{P}_\xi, \mathbb{P}_{\xi'}) \mathbb{E} |\xi - \xi'|. \end{aligned}$$

Now, as $W_2(\mathbb{P}_\xi, \mathbb{P}_{\xi'}) \leq (\mathbb{E} |\xi - \xi'|^2)^{1/2}$ by definition and $\mathbb{E} |\xi - \xi'| \leq (\mathbb{E} |\xi - \xi'|^2)^{1/2}$ by Hölder inequality, we have

$$\begin{aligned} & \mathbb{E} \left[\langle b(\xi, \mathbb{P}_\xi, a) - b(\xi', \mathbb{P}_{\xi'}, a), \xi - \xi' \rangle + \frac{1}{2} |\sigma(\xi, \mathbb{P}_\xi, a) - \sigma(\xi', \mathbb{P}_{\xi'}, a)|^2 \right] \\ &\leq (-\gamma + L_{b\mu} + \frac{1}{2} L_{\sigma\mu}^2 + L_{\sigma x} L_{\sigma\mu}) \mathbb{E} |\xi - \xi'|^2 = -\eta \mathbb{E} |\xi - \xi'|^2, \end{aligned}$$

where $\eta > 0$ was defined in (2.27). $\square$

2.4 Ergodic problems on the Wasserstein space

We first introduce assumptions for the running reward function f , which will be used in connection with various optimization problems.

Assumptions ($\mathbf{H}_3$). (i) The function $f : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times A \rightarrow \mathbb{R}$ is Borel measurable and satisfies for some constants L_f, M_f ,

$$\begin{aligned} |f(x, \mu, a)| &\leq M_f, \\ |f(x, \mu, a) - f(x', \mu', a)| &\leq L_f (|x - x'| + W_2(\mu, \mu')), \end{aligned} \quad (2.28)$$

for every $x, x' \in \mathbb{R}^d$, $\mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$, $a \in A$.

(ii) For every fixed $(x, \mu) \in \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, $f(x, \mu, \cdot)$ is a continuous function.

In the rest of this Section we assume that Assumptions ($\mathbf{H}_1$)-($\mathbf{H}_3$) hold true.

In the following paragraph we introduce the pair (λ, ϕ) described in the Introduction, consisting of a real number λ and a function $\phi : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$. We then prove the basic Proposition 2.4.2 and its variant Proposition 2.4.4, which will play a crucial role in all the sequel. In the final paragraph we state and prove two results illustrating the relevance of (λ, ϕ) to the limit behavior of several optimization problems.

2.4.1 The pair (λ, ϕ) and its first properties

We will use the so-called vanishing discount approach and to this end we introduce some notation for the corresponding infinite horizon discounted expected reward given by

$$J_\beta(\xi, \alpha) = \mathbb{E} \left[\int_0^\infty e^{-\beta t} f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right] \quad (2.29)$$

for $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$, $\alpha \in \mathcal{A}$ and $\beta > 0$, where $X^{\xi, \alpha}$ is the solution to the state equation (2.19). It can be proved (Theorem 4.3 in [84], i.e. Theorem 1.4.3) that the corresponding value $\sup_{\alpha \in \mathcal{A}} J_\beta(\xi, \alpha)$ only depends on ξ via its law, so that one can in fact define a value function $v^\beta : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ setting

$$v^\beta(\mu) = \sup_{\alpha \in \mathcal{A}} J_\beta(\xi, \alpha), \quad \mu \in \mathcal{P}_2(\mathbb{R}^d), \quad (2.30)$$

for any $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ such that $\mathbb{P}_\xi = \mu$.

Proposition 2.4.1. We have, for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $\beta > 0$,

$$\beta |v^\beta(\mu)| \leq M_f. \quad (2.31)$$

Moreover, there exists a constant L , depending only on the constants appearing in Assumptions ($\mathbf{H}_1$)-($\mathbf{H}_3$), such that

$$|v^\beta(\mu) - v^\beta(\nu)| \leq L W_2(\mu, \nu), \quad (2.32)$$

for every $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$ and $\beta > 0$.

Proof. Since f is bounded by M_f we have $|J_\beta(\xi, \alpha)| \leq M_f/\beta$ and (2.31) follows immediately.

Now let us consider two solutions X, X' to the state equation (2.19) such that the joint law of $(X_0, X'_0) = (\xi, \xi')$ is an optimal coupling for W_2 , namely such that $\mathbb{E}[|\xi - \xi'|^2] =$

$W_2(\mu, \nu)^2$ with $\mathbb{P}_{X_0} = \mu$, $\mathbb{P}_{X'_0} = \nu$. The existence of such a pair follows from the existence of the variable U introduced at the beginning of Section 2.2. From (2.21) we have $W_2(\mathbb{P}_{X_t}, \mathbb{P}_{X'_t})^2 \leq \mathbb{E}[|X_t - X'_t|^2] \leq \mathbb{E}[|X_0 - X'_0|^2] e^{-2\eta t} = W_2(\mu, \nu)^2 e^{-2\eta t}$. By the Lipschitz continuity of f (see (2.28)) we have, for every $\alpha \in \mathcal{A}$,

$$\begin{aligned} |J_\beta(X_0, \alpha) - J_\beta(X'_0, \alpha)| &\leq L_f \int_0^\infty e^{-\beta t} \left(\mathbb{E}|X_t - X'_t| + W_2(\mathbb{P}_{X_t}, \mathbb{P}_{X'_t}) \right) dt \\ &\leq L_f W_2(\mu, \nu) \mathbb{E} \int_0^\infty 2e^{-\beta t} e^{-\eta t} dt \end{aligned}$$

and (2.32) follows. $\square$

Since the family $(\beta v^\beta(\delta_0))_{\beta>0}$ is bounded, we can find a subsequence $\beta_n \rightarrow 0$ and a $\lambda \in \mathbb{R}$ such that

$$\beta_n v^{\beta_n}(\delta_0) \rightarrow \lambda, \quad n \rightarrow \infty. \quad (2.33)$$

Then we have $\beta_n v^{\beta_n}(\mu) \rightarrow \lambda$ for any other $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, since, by (2.32),

$$|\beta_n v^{\beta_n}(\mu) - \lambda| \leq L \beta_n W_2(\mu, \delta_0) + |\beta_n v^{\beta_n}(\delta_0) - \lambda| \rightarrow 0, \quad n \rightarrow +\infty.$$

Now consider the sequence $(w^\beta(\cdot) := v^\beta(\cdot) - v^\beta(\delta_0))_{\beta>0}$. From (2.32) it follows that this sequence is equicontinuous and bounded on every W_2 -ball of $\mathcal{P}_2(\mathbb{R}^d)$. Moreover, we recall that $\mathcal{P}_2(\mathbb{R}^d)$ is separable (see [90] Theorem 6.18, and [23] for a short direct proof); we can thus fix a sequence of probability measures $\{\mu_i\}_{i \in \mathbb{N}}$ which is dense in $\mathcal{P}_2(\mathbb{R}^d)$: for all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $\epsilon > 0$ there exists an index $\bar{i} \in \mathbb{N}$ such that

$$W_2(\mu, \mu_{\bar{i}}) < \frac{\epsilon}{4}. \quad (2.34)$$

We then adopt a diagonal argument:

- The sequence $\{w^{\beta_n}(\mu_1)\}_{\beta>0}$ is bounded by $L\|\mu_1\|_2$: there thus exists a subsequence $\{\beta_{1,n}\}_{n \in \mathbb{N}} \subseteq \{\beta_n\}_{n \in \mathbb{N}}$ such that $\{w^{\beta_{1,n}}(\mu_1)\}_{n \in \mathbb{N}}$ converges to a certain value $\phi(\mu_1)$;
- $\{w^{\beta_{1,n}}(\mu_2)\}_{n \in \mathbb{N}}$ is bounded by $L\|\mu_2\|_2$: there exist a subsequence $\{\beta_{2,n}\}_{n \in \mathbb{N}} \subseteq \{\beta_{1,n}\}_{n \in \mathbb{N}}$ and a value $\phi(\mu_2)$ such that

$$w^{\beta_{2,n}}(\mu_2) \xrightarrow{n \rightarrow +\infty} \phi(\mu_2);$$

iterating the above reasoning, for each $i \in \mathbb{N}$ we can construct a sequence $\beta_{i,n} \subseteq \beta_{i-1,n}$ such that

$$w^{\beta_{i,n}}(\mu_i) \xrightarrow{n \rightarrow +\infty} \phi(\mu_i).$$

Notice that, as $\beta_{j,n} \subseteq \beta_{i,n}$ for all $j \geq i$, then

$$w^{\beta_{j,n}}(\mu_i) \xrightarrow{n \rightarrow +\infty} \phi(\mu_i)$$

for all $j \geq i$. We therefore deduce that

$$w^{\beta_{n,n}}(\mu_i) \xrightarrow{n \rightarrow +\infty} \phi(\mu_i) \quad \text{for all } i \in \mathbb{N}$$

Moreover,

$$\begin{aligned}
|\phi(\mu_i) - \phi(\mu_j)| &= \lim_{n \rightarrow +\infty} |w^{\beta_{n,n}}(\mu_i) - w^{\beta_{n,n}}(\mu_j)| \\
&= \lim_{n \rightarrow +\infty} |v^{\beta_{n,n}}(\mu_i) - v^{\beta_{n,n}}(\delta_0) - v^{\beta_{n,n}}(\mu_j) + v^{\beta_{n,n}}(\delta_0)| \quad (\text{by (2.32)}) \\
&\leq LW_2(\mu_i, \mu_j).
\end{aligned} \tag{2.35}$$

Therefore, by Theorem 1 in [73], there exists a unique Lipschitz extension of Φ , i.e. a Lipschitz function $\Phi : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ with Lipschitz constant L . Furthermore, for any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$,

$$\begin{aligned}
|w^{\beta_{n,n}}(\mu) - \phi(\mu)| &= |v^{\beta_{n,n}}(\mu) - v^{\beta_{n,n}}(\delta_0) - \phi(\mu)| \\
&= |v^{\beta_{n,n}}(\mu) - v^{\beta_{n,n}}(\mu_i) + v^{\beta_{n,n}}(\mu_i) - v^{\beta_{n,n}}(\delta_0) - \phi(\mu) + \phi(\mu_i) - \phi(\mu_i)| \\
&\leq |v^{\beta_{n,n}}(\mu) - v^{\beta_{n,n}}(\mu_i)| + |w^{\beta_{n,n}}(\mu_i) - \phi(\mu_i)| + |\phi(\mu) - \phi(\mu_i)| \quad (\text{by (2.32) and (2.35)}) \\
&\leq 2LW_2(\mu, \mu_i) + |w^{\beta_{n,n}}(\mu_i) - \phi(\mu_i)| < \epsilon
\end{aligned}$$

for a sufficiently large n and a suitable choice of the index $i \in \mathbb{N}$. We deduce that

$$\lim_{n \rightarrow +\infty} w^{\beta_{n,n}}(\mu) = \phi(\mu)$$

for any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. As the sequence $\{w^{\beta_{n,n}}\}_{n \in \mathbb{N}}$ is equicontinuous and equibounded on every W_2 -ball of $\mathcal{P}_2(\mathbb{R}^d)$, by Ascoli-Arzelà theorem it converges uniformly on any compact subset K of $\mathcal{P}_2(\mathbb{R}^d)$; as we already know that it converges pointwise to $\phi(\cdot)$, we deduce that

$$v^{\beta_n}(\mu) - v^{\beta_n}(\delta_0) \rightarrow \phi(\mu), \quad n \rightarrow \infty, \tag{2.36}$$

uniformly for $\mu \in K$, where K is any compact subset of $\mathcal{P}_2(\mathbb{R}^d)$ and we denote $\beta_{n,n}$ by β_n for simplicity. Moreover,

$$|\phi(\mu)| \leq L W_2(\mu, \delta_0), \quad |\phi(\mu) - \phi(\nu)| \leq L W_2(\mu, \nu), \tag{2.37}$$

for every $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$.

Proposition 2.4.2. For every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $T > 0$, for every $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ such that $\mathbb{P}_\xi = \mu$, we have

$$\phi(\mu) + \lambda T = \sup_{\alpha \in \mathcal{A}} \left\{ \mathbb{E} \left[\int_0^T f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right] + \phi(\mathbb{P}_{X_T^{\xi, \alpha}}) \right\} \tag{2.38}$$

We prepare the proof with a Lemma.

Lemma 2.4.3. If $\mu = \mathbb{P}_\xi$ belongs to $\mathcal{P}_q(\mathbb{R}^d)$ for some $q > 2$, then the family of terminal laws $\left(\mathbb{P}_{X_T^{\xi, \alpha}} \right)_{\alpha \in \mathcal{A}}$ lies in a compact set of $\mathcal{P}_2(\mathbb{R}^d)$.

Proof. By Proposition 2.2.1, $\mathbb{E}[|X_T^{\xi, \alpha}|^q]$ is bounded by a constant which does not depend on α . It follows that $\left(\mathbb{P}_{X_T^{\xi, \alpha}} \right)_{\alpha \in \mathcal{A}}$ lies in a bounded set of $\mathcal{P}_q(\mathbb{R}^d)$ and it is therefore relatively compact in $\mathcal{P}_2(\mathbb{R}^d)$ by a known result (see e.g. Lemma 2.1 in [14]). $\square$

Proof of Proposition 2.4.2. We start from the dynamic programming principle for the infinite horizon optimal control problem (Corollary 4.3.1 in [84], i.e. Corollary 1.4.5 in this thesis):

$$v^\beta(\mu) = \sup_{\alpha \in \mathcal{A}} \left\{ \mathbb{E} \left[\int_0^T e^{-\beta t} f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right] + e^{-\beta T} v^\beta(\mathbb{P}_{X_T^{\xi, \alpha}}) \right\}, \quad (2.39)$$

which gives

$$\begin{aligned} v^\beta(\mu) - v^\beta(\delta_0) + \frac{1 - e^{-\beta T}}{\beta T} \beta v^\beta(\delta_0) T \\ = \sup_{\alpha \in \mathcal{A}} \left\{ \mathbb{E} \left[\int_0^T e^{-\beta t} f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right] + e^{-\beta T} \left[v^\beta(\mathbb{P}_{X_T^{\xi, \alpha}}) - v^\beta(\delta_0) \right] \right\}. \end{aligned}$$

Now we set $\beta = \beta_n$ and we let $n \rightarrow \infty$. The left-hand side tends to $\phi(\mu) + \lambda T$. It is clear that

$$\mathbb{E} \left[\int_0^T e^{-\beta t} f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right] \rightarrow \mathbb{E} \left[\int_0^T f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right]$$

uniformly in $\alpha \in \mathcal{A}$ by the boundedness of f . We will show that

$$\sup_{\alpha \in \mathcal{A}} \left| v^\beta(\mathbb{P}_{X_T^{\xi, \alpha}}) - v^\beta(\delta_0) - \phi(\mathbb{P}_{X_T^{\xi, \alpha}}) \right| \rightarrow 0 \quad (2.40)$$

when $\mu \in \mathcal{P}_q(\mathbb{R}^d)$ for some $q > 2$. Indeed, in this case, by Lemma 2.4.3, the family $(\mathbb{P}_{X_T^{\xi, \alpha}})_{\alpha \in \mathcal{A}}$ lies in a compact subset K of $\mathcal{P}_2(\mathbb{R}^d)$, so the left-hand side of (2.40) is bounded by $\sup_{\mu \in K} |v^\beta(\mu) - v^\beta(\delta_0) - \phi(\mu)|$ which tends to 0 by (2.36).

So letting $n \rightarrow \infty$ it follows that (2.38) holds for $\mu \in \mathcal{P}_q(\mathbb{R}^d)$ for every $q > 2$. The general case can be proved by an approximation argument. If $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ we take a sequence $\xi^n \in L^q(\Omega, \mathcal{G}, \mathbb{P})$ such that $\mathbb{E}[|\xi^n - \xi|^2] \rightarrow 0$, so that $\mu^n = \mathbb{P}_{\xi^n} \rightarrow \mu = \mathbb{P}_\xi$ in the W_2 distance. We write (2.38) for ξ^n and μ^n and we let $n \rightarrow \infty$. In the left-hand side we have $\phi(\mu^n) \rightarrow \phi(\mu)$ because of (2.37). Concerning the right-hand side we have

$$\begin{aligned} & \left| \sup_{\alpha \in \mathcal{A}} \left\{ \mathbb{E} \left[\int_0^T f(X_t^{\xi^n, \alpha}, \mathbb{P}_{X_t^{\xi^n, \alpha}}, \alpha_t) dt \right] + \phi(\mathbb{P}_{X_T^{\xi^n, \alpha}}) \right\} \right. \\ & \quad \left. - \sup_{\alpha \in \mathcal{A}} \left\{ \mathbb{E} \left[\int_0^T f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right] + \phi(\mathbb{P}_{X_T^{\xi, \alpha}}) \right\} \right| \\ & \leq \sup_{\alpha \in \mathcal{A}} \mathbb{E} \left[\int_0^T |f(X_t^{\xi^n, \alpha}, \mathbb{P}_{X_t^{\xi^n, \alpha}}, \alpha_t) - f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t)| dt \right] + \sup_{\alpha \in \mathcal{A}} \left| \phi(\mathbb{P}_{X_T^{\xi^n, \alpha}}) - \phi(\mathbb{P}_{X_T^{\xi, \alpha}}) \right| \\ & \leq L_f \sup_{\alpha \in \mathcal{A}} \int_0^T \left[\mathbb{E} |X_t^{\xi^n, \alpha} - X_t^{\xi, \alpha}| + W_2(\mathbb{P}_{X_t^{\xi^n, \alpha}}, \mathbb{P}_{X_t^{\xi, \alpha}}) \right] dt + L \sup_{\alpha \in \mathcal{A}} W_2(\mathbb{P}_{X_T^{\xi^n, \alpha}}, \mathbb{P}_{X_T^{\xi, \alpha}}), \end{aligned}$$

where we used Assumptions **(H₃)** and (2.37) in the last inequality. By Hölder inequality, this can be further estimated by

$$L_f \sup_{\alpha \in \mathcal{A}} \int_0^T 2 \left(\mathbb{E} |X_t^{\xi^n, \alpha} - X_t^{\xi, \alpha}|^2 \right)^{1/2} dt + L \sup_{\alpha \in \mathcal{A}} \left(\mathbb{E} |X_T^{\xi^n, \alpha} - X_T^{\xi, \alpha}|^2 \right)^{1/2},$$

which tends to 0 by (2.18). Thus, we can pass to the limit and conclude that (2.38) holds for $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ and $\mu = \mathbb{P}_\xi \in \mathcal{P}_2(\mathbb{R}^d)$. $\square$

In the sequel we will need the following slight extension of the previous result. Recall that $X^{t,\xi,\alpha} = (X_s^{t,\xi,\alpha})_{s \in [t,T]}$ denotes the solution to the state equation (2.17) starting from ξ at time t .

Proposition 2.4.4. For every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $0 \leq t \leq T$ and every $\xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$ such that $\mathbb{P}_\xi = \mu$, we have

$$\phi(\mu) + \lambda(T - t) = \sup_{\alpha \in \mathcal{A}} \left\{ \mathbb{E} \left[\int_t^T f(X_s^{t,\xi,\alpha}, \mathbb{P}_{X_s^{t,\xi,\alpha}}, \alpha_s) ds \right] + \phi(\mathbb{P}_{X_T^{t,\xi,\alpha}}) \right\}.$$

Proof. Instead of (2.39), one starts from the following form of the dynamic programming principle for the infinite horizon optimal control problem: by Theorem 2.9 in [84], i.e. Theorem 1.2.10 in this thesis,

$$V(t, \xi) = \sup_{\alpha \in \mathcal{A}} \left\{ \mathbb{E} \left[\int_t^T e^{-\beta(s-t)} f(X_s^{t,\xi,\alpha}, \mathbb{P}_{X_s^{t,\xi,\alpha}}, \alpha_s) ds \right] + e^{-\beta(T-t)} V(T, X_T^{t,\xi,\alpha}) \right\}.$$

By the time invariance and the law invariance property of the value function associated to a discounted infinite horizon optimal control problem of the McKean-Vlasov type, i.e. Theorem 1.3.15 and 1.4.3 in the previous chapter, we deduce that

$$v^\beta(\mu) = \sup_{\alpha \in \mathcal{A}} \left\{ \mathbb{E} \left[\int_t^T e^{-\beta(s-t)} f(X_s^{t,\xi,\alpha}, \mathbb{P}_{X_s^{t,\xi,\alpha}}, \alpha_s) ds \right] + e^{-\beta(T-t)} v^\beta(\mathbb{P}_{X_T^{t,\xi,\alpha}}) \right\}$$

and the rest of the proof is completely analogous to the proof of Proposition 2.4.2. $\square$

2.4.2 Applications to limit behaviour of some value functions

We use equation (2.38) to study the asymptotic behaviour of the value functions of different optimization problems associated with our McKean-Vlasov dynamics (2.17) and (2.19).

As a first application of the previous constructions and properties, we consider a class of optimal control problems with finite horizon $T > 0$ and we study the behaviour of their value functions when $T \rightarrow \infty$. We introduce a measurable terminal reward function $g : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ satisfying, for some constant M_g ,

$$|g(x, \mu)| \leq M_g (1 + |x| + W_2(\mu, \delta_0))^2, \quad x \in \mathbb{R}^d, \mu \in \mathcal{P}_2(\mathbb{R}^d). \quad (2.41)$$

We assume that g is uniformly continuous on bounded sets of $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$. We consider the reward functional

$$J^T(t, \xi, \alpha) = \mathbb{E} \left[\int_t^T f(X_s^{t,\xi,\alpha}, \mathbb{P}_{X_s^{t,\xi,\alpha}}, \alpha_s) ds + g(X_T^{t,\xi,\alpha}, \mathbb{P}_{X_T^{t,\xi,\alpha}}) \right]. \quad (2.42)$$

It can be proved (Theorem 3.6 in [42]) that the corresponding value $\sup_{\alpha \in \mathcal{A}} J^T(t, \xi, \alpha)$ only depends on ξ via its law, so that one can in fact define a value function $v^T : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ setting

$$v^T(t, \mu) = \sup_{\alpha \in \mathcal{A}} J^T(t, \xi, \alpha), \quad t \in [0, T], \mu \in \mathcal{P}_2(\mathbb{R}^d), \quad (2.43)$$

for any $\xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$ such that $\mathbb{P}_\xi = \mu$. We note that Proposition 2.4.2 shows that $\phi(\mu) + \lambda T$ is the value function at time $t = 0$ corresponding to the terminal reward $g = \phi$. We are interested in studying the asymptotic behaviour of $v^T(0, \mu)$ as $T \rightarrow \infty$, which is described in the following result.

Theorem 2.4.5. We have

$$|v^T(0, \mu) - \phi(\mu) - \lambda T| \leq C (1 + W_2(\mu, \delta_0)^2), \quad \mu \in \mathcal{P}_2(\mathbb{R}^d),$$

where the constant C only depends on the constant M_g in (2.41) and on the constants appearing in Assumptions $(\mathbf{H}_1)$ – $(\mathbf{H}_3)$. In particular,

$$\lim_{T \rightarrow \infty} \frac{v^T(0, \mu)}{T} = \lambda, \quad \mu \in \mathcal{P}_2(\mathbb{R}^d).$$

Proof. Take $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ such that $\mathbb{P}_\xi = \mu$. Comparing (2.38) with

$$v^T(0, \mu) = \sup_{\alpha \in \mathcal{A}} \mathbb{E} \left[\int_0^T f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_s) ds + g(X_T^{\xi, \alpha}, \mathbb{P}_{X_T^{\xi, \alpha}}) \right],$$

we deduce

$$|v^T(0, \mu) - \phi(\mu) - \lambda T| \leq \sup_{\alpha \in \mathcal{A}} \mathbb{E} \left| \phi(\mathbb{P}_{X_T^{\xi, \alpha}}) - g(X_T^{\xi, \alpha}, \mathbb{P}_{X_T^{\xi, \alpha}}) \right|.$$

By the linear growth property of ϕ (see (2.37)) and by the growth condition (2.41) on g we have

$$\mathbb{E} \left| \phi(\mathbb{P}_{X_T^{\xi, \alpha}}) - g(X_T^{\xi, \alpha}, \mathbb{P}_{X_T^{\xi, \alpha}}) \right| \leq C (1 + \mathbb{E} [|X_T^{\xi, \alpha}|^2])$$

and the result follows from (2.22), which gives $\mathbb{E} [|X_T^{\xi, \alpha}|^2] \leq W_2(\mu, \delta_0)^2 e^{-\eta T} + K$. $\square$

Our next result concerns the limits (2.14) mentioned in the Introduction; it is in fact a nonlinear Abelian-Tauberian theorem.

Theorem 2.4.6. The number $\lambda \in \mathbb{R}$ defined in (2.33) satisfies

$$\lim_{\beta \rightarrow 0} \beta v^\beta(\mu) = \lambda, \quad \mu \in \mathcal{P}_2(\mathbb{R}^d). \quad (2.44)$$

Moreover,

$$\lambda = \lim_{\beta \rightarrow 0} \sup_{\alpha \in \mathcal{A}} \mathbb{E} \left[\beta \int_0^\infty e^{-\beta t} f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right] = \lim_{T \rightarrow \infty} \sup_{\alpha \in \mathcal{A}} \mathbb{E} \left[\frac{1}{T} \int_0^T f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right], \quad (2.45)$$

where $X^{\xi, \alpha}$ is the solution to the state equation (2.19) and $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ is any random variable with $\mathbb{P}_\xi = \mu$.

Proof. Recall that $(\beta v^\beta(\delta_0))_{\beta > 0}$ is bounded and that λ was defined in (2.33) as the limit of $\beta_n v^{\beta_n}(\delta_0)$ for a suitable subsequence $\beta_n \rightarrow 0$. Now consider the finite horizon control problem (2.43) with terminal reward $g = 0$ and call $v^T(t, \mu)$ the associated

value function. Note that the right limit in (2.45) equals $\lim_{T \rightarrow \infty} T^{-1}v^T(0, \mu)$ and coincides with λ by Theorem 2.4.5. If there existed another subsequence $(\beta_{n'})_{n' \in \mathbb{N}}$ such that $\beta_{n'} v^{\beta_{n'}}(\delta_0) \rightarrow \lambda' \neq \lambda$ then, proceeding as before, we could extract further subsequences, always denoted by n' , and construct a limit function ϕ' such that

$$v^{\beta_{n'}}(\cdot) - v^{\beta_{n'}}(\delta_0) \xrightarrow{n' \rightarrow +\infty} \phi'(\cdot)$$

uniformly on any compact subset of $\mathcal{P}_2(\mathbb{R}^d)$. We could thus derive the corresponding version of (2.38):

$$\phi'(\mu) + \lambda'T = \sup_{\alpha \in \mathcal{A}} \left\{ \int_0^T f(X_s^{\xi, \alpha}, \mathbb{P}_{X_s^{\xi, \alpha}}, \alpha_s) ds + \phi'(\mathbb{P}_{X_T^{\xi, \alpha}}) \right\};$$

repeating the proof of Theorem 2.4.5, we would thus conclude that $\lim_{T \rightarrow \infty} T^{-1}v^T(0, \mu) = \lambda'$, leading to a contradiction. Therefore, every convergent subsequence of the bounded sequence $\beta v^\beta(\delta_0)$ converges to the same value λ . It follows that the limit (2.44) exists and it is given by the first equality in (2.45). $\square$

2.5 Ergodic and parabolic HJB equations on the Wasserstein space

We still suppose that Assumptions (H_1) – (H_3) hold true. In this Section we will show that the pair (λ, ϕ) constructed in the previous Section is a solution to a fully nonlinear elliptic equation on the space $\mathcal{P}_2(\mathbb{R}^d)$, called the ergodic HJB equation. We will also show that (λ, ϕ) describes the long time behavior of solutions to parabolic HJB equations written forward in time. While these results follow fairly easily from the results presented above, precise statements require a rather lengthy preparation.

Firstly, we need to recall a suitable notion of derivative of real functions on the space $\mathcal{P}_2(\mathbb{R}^d)$. Following the approach developed for instance in [42], [43], [78] for the finite horizon case, we are using the derivative in the lifted sense of Lions ([70], see also [30]).

Given a map $u : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$, we call *lifting of u* any function $U : L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d) \rightarrow \mathbb{R}$ such that $U(\xi) = u(\mathbb{P}_\xi)$ for any $\xi \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$. We say that u is *L -differentiable* if its lifting U admits a continuous Fréchet derivative $D_\xi U : L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d) \rightarrow L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$. In this case, by Theorem 6.5 in [30], there exists, for any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, a measurable function $\partial_\mu u(\mu) : \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that $\mathbb{P}$ -a.s. $D_\xi U(\xi) = \partial_\mu u(\mu)(\xi)$ for any $\xi \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ with $\mathbb{P}_\xi = \mu$. The function $\partial_\mu u(\mu)$, uniquely defined up to μ -null sets, is called the *Lions derivative of u at μ* .

Definition 2.5.1. We denote by $\tilde{C}^2(\mathcal{P}_2(\mathbb{R}^d))$ the space of continuous functions $\varphi : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ such that the following conditions hold:

1. φ is L -differentiable;
2. the map $(\mu, x) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \rightarrow \partial_\mu \varphi(\mu)(x) \in \mathbb{R}^d$ is jointly continuous;
3. the derivative $\partial_x \partial_\mu \varphi : (\mu, x) \in \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \rightarrow \partial_x \partial_\mu \varphi(\mu)(x) \in \mathbb{R}^{d \times d}$ exists and is jointly continuous;

4. there exists a constant C_φ such that

$$|\partial_\mu \varphi(\mu)(x)| \leq C_\varphi(1 + |x|), \quad |\partial_x \partial_\mu \varphi(\mu)(x)| \leq C_\varphi, \quad \mu \in \mathcal{P}_2(\mathbb{R}^d), x \in \mathbb{R}^d. \quad (2.46)$$

For functions φ also depending on a time parameter, we define the space $\tilde{C}^{1,2}([0, T] \times \mathcal{P}_2(\mathbb{R}^d))$ in a similar way.

Definition 2.5.2. We denote by $\tilde{C}^{1,2}([0, T] \times \mathcal{P}_2(\mathbb{R}^d))$ the space of continuous functions $\varphi : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ such that the following conditions hold:

1. φ is L -differentiable, i.e. its lifting $\Phi : [0, T] \times L^2(\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}$, $\Phi(t, \xi) = \varphi(t, \mathbb{P}_\xi)$ for all $\xi \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, admits a continuous Fréchet derivative;
2. the map $(t, \mu, x) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \rightarrow \partial_\mu \varphi(t, \mu)(x) \in \mathbb{R}^d$ is jointly continuous;
3. the derivatives $\partial_x \partial_\mu \varphi : (t, \mu, x) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \rightarrow \partial_x \partial_\mu \varphi(t, \mu)(x) \in \mathbb{R}^{d \times d}$ and $\partial_t \varphi : (t, \mu, x) \in [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \rightarrow \partial_t \varphi(t, \mu)(x) \in \mathbb{R}$ exist and are jointly continuous;

4. there exists a constant C_φ such that

$$|\partial_\mu \varphi(t, \mu)(x)| \leq C_\varphi(1 + |x|), \quad |\partial_x \partial_\mu \varphi(t, \mu)(x)| + |\partial_t \varphi(t, \mu)(x)| \leq C_\varphi \quad (2.47)$$

for all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $t \in [0, T]$, $x \in \mathbb{R}^d$.

For $\varphi \in \tilde{C}^2(\mathcal{P}_2(\mathbb{R}^d))$ we set

$$F(\mu, \partial_\mu \varphi(\mu), \partial_x \partial_\mu \varphi(\mu)) := \int_{\mathbb{R}^d} \left[\sup_{a \in A} \left\{ f(x, \mu, a) + \langle b(x, \mu, a), \partial_\mu \varphi(\mu)(x) \rangle + \frac{1}{2} \text{tr}(\sigma \sigma^T(x, \mu, a) \partial_x \partial_\mu \varphi(\mu)(x)) \right\} \right] \mu(dx).$$

We will be interested both in elliptic and parabolic equations containing the nonlinearity F . Given an initial condition $G : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$, we first consider the following equation for an unknown function $w : [0, \infty) \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$:

$$\begin{cases} \partial_t w(t, \mu) - F(\mu, \partial_\mu w(t, \mu), \partial_x \partial_\mu w(t, \mu)) = 0, & t > 0, \mu \in \mathcal{P}_2(\mathbb{R}^d), \\ w(0, \mu) = G(\mu), \end{cases} \quad (2.48)$$

Definition 2.5.3. A continuous function $w : [0, \infty) \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ is called a viscosity subsolution (respectively, supersolution) to equation (2.48) if

$$w(0, \mu) \leq G(\mu), \quad \mu \in \mathcal{P}_2(\mathbb{R}^d), \quad (2.49)$$

(respectively, $w(0, \mu) \geq G(\mu)$ for $\mu \in \mathcal{P}_2(\mathbb{R}^d)$) and, for every $T > 0$, $(t, \mu) \in (0, T) \times \mathcal{P}_2(\mathbb{R}^d)$ and every $\varphi \in \tilde{C}^{1,2}([0, T] \times \mathcal{P}_2(\mathbb{R}^d))$ such that $w - \varphi$ attains a global maximum (respectively, a global minimum) over $[0, T]$ with value 0 at (t, μ) , we have

$$\partial_t \varphi(t, \mu) - F(\mu, \partial_\mu \varphi(t, \mu), \partial_x \partial_\mu \varphi(t, \mu)) \leq 0$$

(respectively, $\partial_t \varphi(t, \mu) - F(\mu, \partial_\mu \varphi(t, \mu), \partial_x \partial_\mu \varphi(t, \mu)) \geq 0$). A viscosity solution to equation (2.48) is a function which is both a supersolution and a subsolution to it.

Below we will also consider parabolic equations which are written backward in time on an interval $[0, T]$ and with a terminal condition at time $T > 0$, as well as elliptic equations for functions depending on $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ only. The corresponding definitions of subsolutions and supersolutions require minor changes and should be clear. We underline that the definition is slightly different for an equation of the form

$$c = F(\mu, \partial_\mu w(\mu), \partial_x \partial_\mu w(\mu)), \quad (2.50)$$

which we will refer to as *ergodic Hamilton-Jacobi-Bellman equation*: a viscosity subsolution (respectively, supersolution) to (2.50) is in fact a pair (c, w) given by a constant $c \in \mathbb{R}$ and a continuous function $w : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ such that, for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and for every $\varphi \in \tilde{C}^2(\mathcal{P}_2(\mathbb{R}^d))$ such that $w - \varphi$ attains a global maximum (respectively, global minimum) at μ with value 0, we have

$$c - F(\mu, \partial_\mu \varphi(\mu), \partial_x \partial_\mu \varphi(\mu)) \leq 0 \quad (2.51)$$

(respectively, $c - F(\mu, \partial_\mu \varphi(\mu), \partial_x \partial_\mu \varphi(\mu)) \geq 0$). A viscosity solution to equation (2.50) is a pair which is both a supersolution and a subsolution to it.

We now recall the following result, proved in [42] (Theorem 5.5), which extends a classical result on viscosity solutions to McKean-Vlasov optimal control problems.

Theorem 2.5.4. Consider again the finite horizon control problem (2.42)-(2.43) under the same assumptions (2.41) on g . Then the value function v^T defined in (2.43) is a viscosity solution to the HJB equation

$$\begin{cases} -\partial_t v^T(t, \mu) - F(\mu, \partial_\mu v^T(t, \mu), \partial_x \partial_\mu v^T(t, \mu)) = 0, & t \in [0, T], \mu \in \mathcal{P}_2(\mathbb{R}^d), \\ v^T(T, \mu) = \int_{\mathbb{R}^d} g(x, \mu) \mu(dx). \end{cases} \quad (2.52)$$

From this theorem we can easily derive the result on the ergodic HJB equation announced at the beginning of this Section.

Theorem 2.5.5. The pair (λ, ϕ) is a viscosity solution to

$$\lambda - F(\mu, \partial_\mu \phi(\mu), \partial_x \partial_\mu \phi(\mu)) = 0, \quad \mu \in \mathcal{P}_2(\mathbb{R}^d). \quad (2.53)$$

Proof. Proposition 2.4.4 states that, when the terminal reward is chosen as $g(x, \mu) = \phi(\mu)$, the value function is given by $v^T(t, \mu) = \phi(\mu) + \lambda(T - t)$. Writing down equation (2.52) for this function gives (2.53): this easy verification is given for convenience in Appendix C.1. $\square$

We now address for a moment the issue of uniqueness for viscosity solutions to equation (2.52), which we write in the form

$$-\partial_t v(t, \mu) - F(\mu, \partial_\mu v(t, \mu), \partial_x \partial_\mu v(t, \mu)) = 0, \quad (2.54)$$

for a generic unknown function $v : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$. Uniqueness results are often deduced as a consequence of the following comparison principle:

(CP) Let $v_1, v_2 : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ be, respectively, a viscosity subsolution and a viscosity supersolution to (2.54) on $[0, T] \times \mathcal{P}_2(\mathbb{R}^d)$, satisfying $v_1(T, \mu) \leq v_2(T, \mu)$ for $\mu \in \mathcal{P}_2(\mathbb{R}^d)$; then $v_1 \leq v_2$ on $[0, T] \times \mathcal{P}_2(\mathbb{R}^d)$.

As already mentioned in Section 1.5, comparison results of this form are presented in many recent papers, see e.g. [43], [36], [14], for classes of Lipschitz continuous functions and [16] for a class of semi continuous functions. The aforementioned results hold under stronger conditions on the coefficients (for instance, independence of the diffusion coefficient σ from the measure argument, boundedness and Lipschitz regularity in (x, μ) uniformly with respect to a of the coefficients b, σ and f , where the Lipschitz continuity with respect to the measure argument is intended in the W_1 distance, possibly higher regularity for the diffusion coefficient). For our next result we will assume that **(CP)** holds for (2.54) in the class of functions which are Lipschitz on $\mathcal{P}_2(\mathbb{R}^d)$ uniformly in t , namely when there exists a constant C such that

$$|v_i(t, \mu) - v_i(t, \mu')| \leq C W_2(\mu, \mu'), \quad \mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d), \quad t \in [0, T], \quad i = 1, 2.$$

However, we avoid repeating precise technical conditions in the above references, which would also involve modifications of the concept of viscosity solution, since our results are not related to specific requirements but only to the uniqueness property.

We are now ready for the result on well-posedness and long time behaviour of solutions to forward parabolic HJB equations announced above.

Proposition 2.5.6. Assume that $g : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ satisfies (2.41) and that it is Lipschitz, i.e. that there exists a constant $L_g \geq 0$ such that

$$|g(x, \mu) - g(x', \mu')| \leq L_g(|x - x'| + W_2(\mu, \mu')).$$

Then,

1. there exists a viscosity solution $w : [0, \infty) \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ to the following parabolic partial differential equation of HJB type (written forward in time):

$$\begin{cases} \partial_t w(t, \mu) - F(\mu, \partial_\mu w(t, \mu), \partial_x \partial_\mu w(t, \mu)) = 0, & t \geq 0, \mu \in \mathcal{P}_2(\mathbb{R}^d), \\ w(0, \mu) = \int_{\mathbb{R}^d} g(x, \mu) \mu(dx). \end{cases} \quad (2.55)$$

Moreover we have

$$|w(t, \mu) - \phi(\mu) - \lambda t| \leq C (1 + W_2(\mu, \delta_0)^2), \quad \mu \in \mathcal{P}_2(\mathbb{R}^d),$$

where the constant C only depends on the constant M_g in (2.41) and on the constants appearing in Assumptions **(H₁)**-**(H₃)**. In particular, we obtain

$$\lim_{t \rightarrow \infty} \frac{w(t, \mu)}{t} = \lambda, \quad \mu \in \mathcal{P}_2(\mathbb{R}^d).$$

2. Assume in addition that, for every fixed time horizon $T > 0$, **(CP)** holds for (2.54) in the class of functions that are Lipschitz on $\mathcal{P}_2(\mathbb{R}^d)$ uniformly in t . Then, the viscosity solution to (2.55) is unique in the class of functions $[0, +\infty) \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ which are Lipschitz in μ uniformly with respect to t on every bounded interval $[0, T]$.

Proof. The proof is based on the idea to establish a bijection between viscosity solutions to equations (2.55) and (2.52), in the class of functions that are Lipschitz on $\mathcal{P}_2(\mathbb{R}^d)$ uniformly in $t \in [0, T]$, using the following equality: on any time interval $[0, T]$, define

$$v^T(t, \mu) = w(T - t, \mu), \quad \mu \in \mathcal{P}_2(\mathbb{R}^d), t \in [0, T],$$

The complete proof is given in Appendix C.2. $\square$

We end this Section with a uniqueness result for the ergodic HJB (2.53).

Theorem 2.5.7. Assume that **(CP)** holds and suppose that a pair $(\bar{\lambda}, \bar{\phi})$, with $\bar{\lambda} \in \mathbb{R}$ and $\bar{\phi} : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ Lipschitz continuous, is a viscosity solution to (2.53). Then we have $\lambda = \bar{\lambda}$.

Proof. For any $T > 0$ define $\bar{v}^T(t, \mu) := \bar{\phi}(\mu) + \bar{\lambda}(T - t)$. If $(\bar{\lambda}, \bar{\phi})$ is a viscosity solution to (2.53), then $\bar{v}^T$ is a viscosity solution to the HJB equation

$$\begin{cases} -\partial_t \bar{v}^T(t, \mu) - F(\mu, \partial_\mu \bar{v}^T(t, \mu), \partial_x \partial_\mu \bar{v}^T(t, \mu)) = 0, & t \in [0, T], \mu \in \mathcal{P}_2(\mathbb{R}^d), \\ \bar{v}^T(T, \mu) = \bar{\phi}(\mu); \end{cases} \quad (2.56)$$

this fact is trivial for classical solutions, but it can also be easily verified for viscosity solutions, following the same steps of the proof in Appendix C.1. Since we assume **(CP)**, we have uniqueness for viscosity solutions to this equation. Therefore, by Theorem 2.5.4, $\bar{v}^T$ coincides with the value function for the finite horizon control problem (2.42)-(2.43) with terminal condition $g(x, \mu) = \bar{\phi}(\mu)$, namely

$$\bar{v}^T(t, \mu) = \sup_{\alpha \in \mathcal{A}} \left\{ \mathbb{E} \left[\int_t^T f(X_s^{t, \xi, \alpha}, \mathbb{P}_{X_s^{t, \xi, \alpha}}, \alpha_s) ds \right] + \bar{\phi}(\mathbb{P}_{X_T^{t, \xi, \alpha}}) \right\}, \quad t \in [0, T], \mu \in \mathcal{P}_2(\mathbb{R}^d), \quad (2.57)$$

for any $\xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$ such that $\mathbb{P}_\xi = \mu$. Now recall that, by Proposition 2.4.2, we have

$$\phi(\mu) + \lambda T = \sup_{\alpha \in \mathcal{A}} \left\{ \mathbb{E} \left[\int_0^T f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right] + \phi(\mathbb{P}_{X_T^{\xi, \alpha}}) \right\}.$$

Setting $t = 0$ in (2.57) and subtracting we deduce

$$|\bar{\phi}(\mu) + \bar{\lambda}T - \phi(\mu) - \lambda T| = |v^T(0, \mu) - \phi(\mu) - \lambda T| \leq \sup_{\alpha \in \mathcal{A}} \left| \bar{\phi}(\mathbb{P}_{X_T^{\xi, \alpha}}) - \phi(\mathbb{P}_{X_T^{\xi, \alpha}}) \right|. \quad (2.58)$$

By the Lipschitz property of ϕ and $\bar{\phi}$ we have, for some constant C ,

$$\left| \bar{\phi}(\mathbb{P}_{X_T^{\xi, \alpha}}) - \phi(\mathbb{P}_{X_T^{\xi, \alpha}}) \right| \leq C \left(1 + W_2(\delta_0, \mathbb{P}_{X_T^{\xi, \alpha}}) \right) \leq C \left(1 + [\mathbb{E}|X_T^{\xi, \alpha}|^2]^{1/2} \right),$$

which is bounded by a constant independent of T , by (2.22). Dividing (2.58) by T and letting $T \rightarrow \infty$ we obtain $\lambda = \bar{\lambda}$. $\square$

This result states that for any pair (λ, ϕ) solution to the ergodic HJB, the number λ is uniquely determined. Uniqueness for ϕ clearly fails, since $\phi + C$ is also a solution for any constant C . Even in simpler situations, other solutions may also exist: see for instance Remark 1.2 in Chapter VII of [12]. We come back to the uniqueness of ϕ (up to a constant) in Remark 2.6.7 below.

2.6 Ergodic control problem

We still suppose that Assumptions $(\mathbf{H}_1)$ – $(\mathbf{H}_3)$ hold true and we denote (λ, ϕ) the pair constructed in Section 2.4. In this Section we consider the ergodic reward

$$J_\infty(\xi, \alpha) = \limsup_{T \rightarrow \infty} \mathbb{E} \left[\frac{1}{T} \int_0^T f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right], \quad (2.59)$$

where $X^{\xi, \alpha}$ is the solution to the state equation (2.19), and we maximize $J_\infty(\xi, \alpha)$ over all $\alpha \in \mathcal{A}$. If the maximum is attained by some $\hat{\alpha}$ we say that $\hat{\alpha}$ is optimal for the ergodic problem.

Our first result shows that the value $\sup_{\alpha \in \mathcal{A}} J_\infty(\xi, \alpha)$ is constant, i.e. it does not depend on ξ , and that it is equal to λ .

Proposition 2.6.1. For every $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ we have

$$\lambda = \sup_{\alpha \in \mathcal{A}} \limsup_{T \rightarrow \infty} \mathbb{E} \left[\frac{1}{T} \int_0^T f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right]. \quad (2.60)$$

Proof. We first prove that

$$\lambda \geq \sup_{\alpha \in \mathcal{A}} \limsup_{T \rightarrow \infty} \mathbb{E} \left[\frac{1}{T} \int_0^T f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right]. \quad (2.61)$$

For any numerical function $q(\alpha, T)$ the inequality

$$\limsup_{T \rightarrow \infty} \sup_{\alpha \in \mathcal{A}} q(\alpha, T) \geq \sup_{\alpha \in \mathcal{A}} \limsup_{T \rightarrow \infty} q(\alpha, T)$$

is easily verified. Choosing $q(\alpha, T) = \mathbb{E} \left[\frac{1}{T} \int_0^T f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right]$, the second equality in (2.45) gives $\lambda = \lim_{T \rightarrow \infty} \sup_{\alpha \in \mathcal{A}} q(\alpha, T)$ and (2.61) follows.

Next, we prove the opposite inequality. Let us fix $T > 0$, $\epsilon > 0$, and $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ and set $\mu = \mathbb{P}_\xi$. By Proposition 2.4.4, for every integer $n \geq 0$, and every $\zeta \in L^2(\Omega, \mathcal{F}_{nT}, \mathbb{P})$ we have

$$\phi(\mathbb{P}_\zeta) = \sup_{\alpha \in \mathcal{A}} \left\{ \mathbb{E} \int_{nT}^{(n+1)T} \left[f(X_s^{nT, \zeta, \alpha}, \mathbb{P}_{X_s^{nT, \zeta, \alpha}}, \alpha_s) - \lambda \right] ds + \phi \left(\mathbb{P}_{X_{(n+1)T}^{nT, \zeta, \alpha}} \right) \right\}. \quad (2.62)$$

Setting $n = 0$, $\zeta = \xi$, we see that there exists $\alpha^0 \in \mathcal{A}$ such that

$$\phi(\mu) - \epsilon T = \phi(\mathbb{P}_\xi) - \epsilon T \leq \mathbb{E} \int_0^T \left[f(X_s^{0, \xi, \alpha^0}, \mathbb{P}_{X_s^{0, \xi, \alpha^0}}, \alpha_s^0) - \lambda \right] ds + \phi \left(\mathbb{P}_{X_T^{0, \xi, \alpha^0}} \right).$$

Suppose that for some $n \geq 1$ we have constructed a control process $\alpha^{n-1} \in \mathcal{A}$ such that

$$\phi(\mu) - n \epsilon T \leq \mathbb{E} \int_0^{nT} \left[f(X_s^{0,\xi,\alpha^{n-1}}, \mathbb{P}_{X_s^{0,\xi,\alpha^{n-1}}}, \alpha_s^{n-1}) - \lambda \right] ds + \phi\left(\mathbb{P}_{X_{nT}^{0,\xi,\alpha^{n-1}}}\right). \quad (2.63)$$

By (2.62) with $\zeta = X_{nT}^{0,\xi,\alpha^{n-1}} \in L^2(\Omega, \mathcal{F}_{nT}, \mathbb{P})$ we find $\tilde{\alpha}^n \in \mathcal{A}$ such that

$$\phi(\mathbb{P}_\zeta) - \epsilon T \leq \mathbb{E} \int_{nT}^{(n+1)T} \left[f(X_s^{nT,\zeta,\tilde{\alpha}^n}, \mathbb{P}_{X_s^{nT,\zeta,\tilde{\alpha}^n}}, \tilde{\alpha}_s^n) - \lambda \right] ds + \phi\left(\mathbb{P}_{X_{(n+1)T}^{nT,\zeta,\tilde{\alpha}^n}}\right). \quad (2.64)$$

We define $\alpha_s^n = \alpha_s^{n-1} 1_{[0,nT)}(s) + \tilde{\alpha}_s^n 1_{[nT,\infty)}(s)$ and we note that $X_s^{0,\xi,\alpha^n} = X_s^{0,\xi,\alpha^{n-1}}$ for $s \in [0, nT]$ and $X_s^{0,\xi,\alpha^n} = X_s^{nT,\zeta,\tilde{\alpha}^n}$ for $s \geq nT$ by the flow property of solutions to McKean-Vlasov SDEs. Summing (2.63) and (2.64), we obtain

$$\phi(\mu) - (n+1) \epsilon T \leq \mathbb{E} \int_0^{(n+1)T} \left[f(X_s^{0,\xi,\alpha^n}, \mathbb{P}_{X_s^{0,\xi,\alpha^n}}, \alpha_s^n) - \lambda \right] ds + \phi\left(\mathbb{P}_{X_{(n+1)T}^{0,\xi,\alpha^n}}\right)$$

and we conclude by induction that (2.63) holds for every n . We can also consistently define $\hat{\alpha} \in \mathcal{A}$ setting $\hat{\alpha}_s = \alpha_s^{n-1}$ for $s \in [0, nT)$ so that $X_s^{0,\xi,\hat{\alpha}} = X_s^{0,\xi,\alpha^{n-1}}$ for $s \in [0, nT)$. Rewriting (2.63) in terms of $\hat{\alpha}$ and $X_s^{0,\xi,\hat{\alpha}}$ and dividing by nT , we obtain

$$\frac{1}{nT} \phi(\mu) - \frac{1}{nT} \phi\left(\mathbb{P}_{X_{nT}^{0,\xi,\hat{\alpha}}}\right) + \lambda - \epsilon \leq \frac{1}{nT} \mathbb{E} \int_0^{nT} f(X_s^{0,\xi,\hat{\alpha}}, \mathbb{P}_{X_s^{0,\xi,\hat{\alpha}}}, \hat{\alpha}_s) ds.$$

Since, by (2.37) and (2.22),

$$\left| \phi\left(\mathbb{P}_{X_{nT}^{0,\xi,\hat{\alpha}}}\right) \right| \leq L W_2\left(\mathbb{P}_{X_{nT}^{0,\xi,\hat{\alpha}}}, \delta_0\right) \leq L \left(\mathbb{E} \left[|X_{nT}^{0,\xi,\hat{\alpha}}|^2 \right] \right)^{1/2} \leq L \left(\mathbb{E} [|\xi|^2] e^{-\eta nT} + K \right)^{1/2}, \quad (2.65)$$

we obtain

$$\begin{aligned} \lambda - \epsilon &\leq \limsup_{n \rightarrow \infty} \frac{1}{nT} \mathbb{E} \int_0^{nT} f(X_s^{0,\xi,\hat{\alpha}}, \mathbb{P}_{X_s^{0,\xi,\hat{\alpha}}}, \hat{\alpha}_s) ds \\ &\leq \limsup_{T' \rightarrow \infty} \frac{1}{T'} \mathbb{E} \int_0^{T'} f(X_s^{0,\xi,\hat{\alpha}}, \mathbb{P}_{X_s^{0,\xi,\hat{\alpha}}}, \hat{\alpha}_s) ds. \end{aligned}$$

It follows that

$$\lambda - \epsilon \leq \sup_{\alpha \in \mathcal{A}} \limsup_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \int_0^T f(X_t^{0,\xi,\alpha}, \mathbb{P}_{X_t^{0,\xi,\alpha}}, \alpha_t) dt$$

and the required inequality is obtained letting $\epsilon \rightarrow 0$. $\square$

Remark 2.6.2. For the reader's convenience we sketch another alternative proof of the inequality (2.61), based on an argument which dates back at least to [5] (Theorem 5) in the case of deterministic control; see also [11] Theorem 3.4 for a recent application to stochastic control problems with mean-field effects. We start again from (2.39) which gives, for every $\alpha \in \mathcal{A}$ and setting $\mu = \mathbb{P}_\xi$,

$$v^\beta(\mu) \geq \mathbb{E} \left[\int_0^T e^{-\beta t} f(X_t^{\xi,\alpha}, \mathbb{P}_{X_t^{\xi,\alpha}}, \alpha_t) dt \right] + e^{-\beta T} v^\beta\left(\mathbb{P}_{X_T^{\xi,\alpha}}\right).$$

Multiplying by β ,

$$\begin{aligned} \beta v^\beta(\mu) - \beta e^{-\beta T} v^\beta(\mathbb{P}_{X_T^{\xi, \alpha}}) &\geq \mathbb{E} \left[\int_0^T \beta e^{-\beta t} f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right] \\ &= \mathbb{E} \left[\int_0^T \beta f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right] + \mathbb{E} \left[\int_0^T \beta (e^{-\beta t} - 1) f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right]. \end{aligned}$$

The last expectation is bounded by $M_f \int_0^T \beta |e^{-\beta t} - 1| dt = M_f (\beta T + e^{-\beta T} - 1)$ by $(\mathbf{H}_3)$. Setting $\delta = \beta T$,

$$\beta v^\beta(\mu) - e^{-\delta} \beta v^\beta(\mathbb{P}_{X_T^{\xi, \alpha}}) \geq \delta \mathbb{E} \left[\frac{1}{T} \int_0^T f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right] - M_f (\delta + e^{-\delta} - 1).$$

Now we let $\beta \rightarrow 0$ (possibly just along a subsequence) and $T \rightarrow \infty$, keeping δ fixed. We recall that $\beta v^\beta(\mu) \rightarrow \lambda$. Since, by (2.32) and (2.22), $\beta |v^\beta(\mu) - v^\beta(\mathbb{P}_{X_T^{\xi, \alpha}})| \leq \beta L W_2(\mathbb{P}_{X_T^{\xi, \alpha}}, \mu) \rightarrow 0$, we conclude that $\beta v^\beta(\mathbb{P}_{X_T^{\xi, \alpha}}) \rightarrow \lambda$. It follows that

$$\lambda (1 - e^{-\delta}) \geq \delta \limsup_{T \rightarrow \infty} \mathbb{E} \left[\frac{1}{T} \int_0^T f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right] - M_f (\delta + e^{-\delta} - 1).$$

Dividing by δ and letting $\delta \rightarrow 0$ we obtain $\lambda \geq \limsup_{T \rightarrow \infty} \mathbb{E} \left[\frac{1}{T} \int_0^T f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right]$ and (2.61) follows.

One may wonder whether this argument can be adapted to yield the opposite inequality as well, according to some results which can be found in the literature. Given $\delta > 0$ one can find a control $\tilde{\alpha}$ such that

$$v^\beta(\mu) - \delta^2 \leq \mathbb{E} \left[\int_0^T e^{-\beta t} f(X_t^{\xi, \tilde{\alpha}}, \mathbb{P}_{X_t^{\xi, \tilde{\alpha}}}, \tilde{\alpha}_t) dt \right] + e^{-\beta T} v^\beta(\mathbb{P}_{X_T^{\xi, \tilde{\alpha}}})$$

and, after similar rearrangements and estimates, one can let $T \rightarrow \infty$ and $\beta \rightarrow 0$ keeping $\delta = \beta T$ fixed as before. However, to finish the proof by this argument, one needs to find the control $\tilde{\alpha}$ independent of T and β . While dependence on T can be avoided (using a more general version of the dynamic programming principle, see e.g. [12], Chapter VII Proposition 1.3, for deterministic control), the dependence on β seems unavoidable. In order to circumvent this difficulty we have devised the previous proof. $\square$

We end this Section with some results on the existence of optimal controls for the ergodic problem.

Proposition 2.6.3. Let $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ and denote $\mu = \mathbb{P}_\xi$. Suppose that $\hat{\alpha} \in \mathcal{A}$ satisfies one of the following equivalent conditions:

i) for every $T > 0$,

$$\phi(\mu) + \lambda T = \mathbb{E} \left[\int_0^T f(X_t^{\xi, \hat{\alpha}}, \mathbb{P}_{X_t^{\xi, \hat{\alpha}}}, \hat{\alpha}_t) dt \right] + \phi(\mathbb{P}_{X_T^{\xi, \hat{\alpha}}}); \quad (2.66)$$

ii) for every $T > 0$, $\hat{\alpha}$ is optimal for the finite horizon problem

$$\sup_{\alpha \in \mathcal{A}} \left\{ \mathbb{E} \left[\int_0^T f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right] + \phi(\mathbb{P}_{X_T^{\xi, \alpha}}) \right\}. \quad (2.67)$$

Then $\hat{\alpha}$ is optimal for the ergodic problem.

Proof. The equivalence of i) and ii) follows from Proposition 2.4.2. If (2.66) holds, then dividing by T and letting $T \rightarrow \infty$, using again (2.37) and (2.22) as before (see for instance (2.65)), we obtain

$$\lambda = \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\int_0^T f(X_t^{\xi, \hat{\alpha}}, \mathbb{P}_{X_t^{\xi, \hat{\alpha}}}, \hat{\alpha}_t) dt \right]$$

which proves the optimality of $\hat{\alpha}$ by Proposition 2.6.1. $\square$

This Proposition leads to a sufficient condition for optimality in the form of a martingale optimality principle, as proposed in [44] Proposition 6.1-(ii):

Corollary 2.6.4. Suppose that there exists $\hat{\alpha} \in \mathcal{A}$ such that

$$\phi(\mathbb{P}_{X_T^{\xi, \hat{\alpha}}}) + \int_0^T f(X_t^{\xi, \hat{\alpha}}, \mathbb{P}_{X_t^{\xi, \hat{\alpha}}}, \hat{\alpha}_t) dt - \lambda T, \quad T \geq 0,$$

is a martingale. Then $\hat{\alpha}$ is optimal for the ergodic control problem.

The proof is obvious, as taking expectation we obtain (2.66).

Remark 2.6.5. Proposition 2.6.3 admits the following extension. Suppose that there exists a pair (λ', ϕ') such that the function $\phi' : \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ satisfies the quadratic growth condition $|\phi'(\mu)| \leq C(1 + W_2(\mu, \delta_0)^2)$ for some constant C and every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, and such that the analogue of (2.38) holds, namely for every $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $T > 0$ and $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ such that $\mathbb{P}_\xi = \mu$ we have

$$\phi'(\mu) + \lambda' T = \sup_{\alpha \in \mathcal{A}} \left\{ \mathbb{E} \left[\int_0^T f(X_t^{\xi, \alpha}, \mathbb{P}_{X_t^{\xi, \alpha}}, \alpha_t) dt \right] + \phi'(\mathbb{P}_{X_T^{\xi, \alpha}}) \right\}. \quad (2.68)$$

If we consider the finite horizon control problem (2.42)-(2.43) with arbitrary terminal value g (for instance $g = 0$) and we call v^T the corresponding value function, then arguing as in Theorem 2.4.5 we first deduce that

$$\lambda' = \lim_{T \rightarrow \infty} \frac{v^T(0, \mu)}{T} = \lambda, \quad \mu \in \mathcal{P}_2(\mathbb{R}^d).$$

Next also suppose that there exists $\hat{\alpha} \in \mathcal{A}$ such that for every $T > 0$ we have

$$\phi'(\mu) + \lambda T = \mathbb{E} \left[\int_0^T f(X_t^{\xi, \hat{\alpha}}, \mathbb{P}_{X_t^{\xi, \hat{\alpha}}}, \hat{\alpha}_t) dt \right] + \phi'(\mathbb{P}_{X_T^{\xi, \hat{\alpha}}}), \quad (2.69)$$

which is the analogue to (2.66). Then $\hat{\alpha}$ is optimal for the ergodic problem. The proof remains the same as in Proposition 2.6.3. $\square$

Another possibility to construct an optimal control by checking the conditions in Proposition 2.6.3 makes use of the ergodic HJB equation for ϕ . It can be stated in the form of a verification theorem. Recall the definition of the space $\tilde{C}^2(\mathcal{P}_2(\mathbb{R}^d))$ given in the previous Section.

Theorem 2.6.6. Suppose that $\phi \in \tilde{C}^2(\mathcal{P}_2(\mathbb{R}^d))$ and that the pair (λ, ϕ) is a classical solution to the ergodic equation (2.53). Suppose that there exists a Borel measurable function $\underline{\alpha} : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow A$ which realizes the supremum in the equation, namely satisfying

$$\lambda = \int_{\mathbb{R}^d} \left\{ f(x, \mu, \underline{\alpha}(x, \mu)) + \langle b(x, \mu, \underline{\alpha}(x, \mu)), \partial_\mu \phi(\mu)(x) \rangle + \frac{1}{2} \text{tr} \left(\sigma \sigma^T(x, \mu, \underline{\alpha}(x, \mu)) \partial_x \partial_\mu \phi(\mu)(x) \right) \right\} \mu(dx), \quad x \in \mathbb{R}^d, \mu \in \mathcal{P}_2(\mathbb{R}^d). \quad (2.70)$$

Finally assume that the closed-loop equation

$$\begin{cases} dX_t = b(X_t, \mathbb{P}_{X_t}, \underline{\alpha}(X_t, \mathbb{P}_{X_t})) dt + \sigma(X_t, \mathbb{P}_{X_t}, \underline{\alpha}(X_t, \mathbb{P}_{X_t})) dB_t, & t \geq 0 \\ X_0 = \xi, \end{cases} \quad (2.71)$$

admits a solution X for some $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ and that the control

$$\hat{\alpha}_t := \underline{\alpha}(X_t, \mathbb{P}_{X_t}), \quad t \geq 0,$$

belongs to $\mathcal{A}$. Then $\hat{\alpha}$ is optimal for the ergodic problem.

Proof. We fix $T > 0$ arbitrary and we apply the Itô formula on the interval $[0, T]$ (see e.g. Theorem 4.16 and Remark 4.17 in [42]) to the composition of φ with the solution of the closed-loop equation, obtaining:

$$\begin{aligned} \phi(\mathbb{P}_{X_T^{\xi, \hat{\alpha}}}) - \phi(\mu) &= \mathbb{E} \int_0^T \left\{ \langle b(X_t, \mathbb{P}_{X_t}, \underline{\alpha}(X_t, \mathbb{P}_{X_t})), \partial_\mu \phi(\mathbb{P}_{X_t})(X_t) \rangle \right. \\ &\quad \left. + \frac{1}{2} \text{tr} \left(\sigma \sigma^T(X_t, \mathbb{P}_{X_t}, \underline{\alpha}(X_t, \mathbb{P}_{X_t})) \partial_x \partial_\mu \phi(\mu)(X_t) \right) \right\} dt. \end{aligned}$$

From (2.70) it follows that

$$\phi(\mathbb{P}_{X_T^{\xi, \hat{\alpha}}}) - \phi(\mu) = \mathbb{E} \int_0^T \left\{ \lambda - f(X_t, \mathbb{P}_{X_t}, \underline{\alpha}(X_t, \mathbb{P}_{X_t})) \right\} dt = \lambda T - \mathbb{E} \int_0^T f(X_t, \mathbb{P}_{X_t}, \hat{\alpha}_t) dt.$$

Therefore (2.66) holds and the conclusion follows from Proposition 2.6.3. $\square$

Remark 2.6.7. We recall that for a pair (λ, ϕ) which is a viscosity solution to the ergodic HJB equation (2.53), uniqueness holds for the number λ . In the classical framework of ergodic control, assuming more regularity on ϕ and some of the assumptions of the previous theorem (and suitable additional conditions) one can prove that the function ϕ is also uniquely determined, up to a constant. This conclusion is often

reached as a consequence of results on long time behavior of the solution w to the parabolic equation (2.55): setting

$$G(t, \mu) = w(t, \mu) - \phi(\mu) - \lambda t, \quad t \geq 0, \mu \in \mathcal{P}_2(\mathbb{R}^d),$$

besides the conclusion that $\sup_{t \geq 0} |G(t, \mu)| < \infty$, reached in Proposition 2.5.6, one may conjecture that $G(t, \mu)$ tends to a constant as $t \rightarrow \infty$, thus proving the required uniqueness for ϕ . We refer for instance to [59], [60] and the references therein on the large literature on asymptotic behavior for HJB equations on $\mathbb{R}^d$, as well as [41] for a probabilistic and control-theoretic approach to the same issue. It seems that there is no immediate extension of these results to our framework. Moreover, the proofs in the classical case also rely on existence and properties of an invariant measure for the solution $X^{\xi, \hat{\alpha}}$ to the closed-loop equation. Since the latter is non-Markovian, because of the occurrence of mean-field effects, even this part of the proof would require a detailed investigation and probably more restrictive assumptions. In particular, existence, uniqueness and (exponential) convergence to invariant measures for uncontrolled McKean-Vlasov systems have been studied in many contributions, see e.g. in [71], [89] and [28] for the case of additive noise, [91] for time-homogeneous coefficients and [92], [80] and references therein for distribution-dependent drift term; the results in the aforementioned references hold under additional assumptions on the coefficients of the state equation (e.g. monotonicity and Lyapunov conditions, boundedness, non-degeneracy,...) and should be adapted to the case of a SDE in closed-loop form. For these reasons we will not present results on uniqueness for the function ϕ in this dissertation.

Chapter 3

Optimal control of McKean-Vlasov SDEs with partial observation

3.1 Introduction

In this Chapter we present some preliminary results for a work which is currently in progress, in collaboration with M. Fuhrman and H. Pham. Our aim is to develop a theory of optimal control for a partially observed control problem, where the dynamics of the system is of the McKean-Vlasov type and also depends on a non-observable, exogenous source of randomness M ; it is thus described by an equation of the form

$$dZ_s = h(Z_s, \mathbb{Q}_{Z_s}, \alpha_s, M_s) ds + dB_s, \quad s \in [0, T].$$

In our model, M is a continuous-time Markov chain taking values in a set with finite cardinality. The agent only observes the process Z ; the control process α will thus be $\mathbb{F}^Z$ -predictable, where $\mathbb{F}^Z$ is the filtration generated by Z . The analysis of this problem requires filtering techniques.

Filtering theory aims at describing the evolution in time of the conditional distribution of the signal process, i.e. the unknown source of randomness, given the filtration generated by the observation process: this distribution is called filter and corresponds to the best approximation of the unknown underlying source of randomness given the observation. It has firstly been developed for uncontrolled systems: for this case we refer the reader to many classical references, see e.g. [8] and [75]. More recently, the theory of optimal control of partially observed problems has attracted considerable interest, due to its many possible applications: for classical models, without mean-field effects, we refer the reader to [21] and [50] as general references. For the specific case where the unobserved random source is represented by an uncontrolled Markov chain, we mention e.g. [52], [53], [82] and [85], which are mainly devoted to specific models of interest in Mathematical Finance, without mean-field effects. A mean-field control problem with Markov switching has been studied through the analytic approach in [13], where the authors also provide a finite agent approximation and a propagation of chaos result. However, our approach differs because that paper presents a fully observable control problem, where the feedback admissible controls also depend on

the Markov chain; moreover, the Markov chain is treated as a common noise and the coefficients of the state equation thus depend on the conditional law of the state process with respect to the Markov chain.

In our problem, we adopt the strategy based on Zakai equation: after a change of the underlying probability measure relying on Girsanov's theorem, we derive a measure-valued SDE, called Zakai equation, driven by the process Z , which is a Brownian motion under the new probability measure $\mathbb{P}$. The solution to Zakai equation will be identified with the unnormalized conditional distribution of the signal given the observation process, namely

$$\rho_s(\phi) = \mathbb{E}[\phi(M_s) L_s | \mathcal{F}_s^Z], \quad s \in [0, T] \quad \text{for all } \phi : S \rightarrow \mathbb{R},$$

where L is the density process of $\mathbb{Q}$ with respect to $\mathbb{P}$. The main difficulty of this approach in a McKean-Vlasov setting is precisely due to the change of the underlying probability measure: as a matter of fact, if we define two different probabilities $\mathbb{P}$ and $\mathbb{Q}$ over the same measurable space $(\Omega, \mathcal{F})$, the law of any random variable Y under $\mathbb{Q}$, given by the push-forward of Y through $\mathbb{Q}$, does no longer represent the law of the same random variable under $\mathbb{P}$. In our case, as the two probability measures are equivalent and the density of one measure with respect to the other is a time-marginal of an exponential martingale L due to Girsanov's theorem, we provide an explicit characterization of the law of the time-marginals of the state process under the original probability measure $\mathbb{Q}$ in terms of the new probability measure $\mathbb{P}$. More precisely, we obtain

$$\mathbb{Q}_{Z_s}(dx) = \mu_s(x) \mathbb{P}_{Z_s}(dx),$$

where $\mu_s : \mathbb{R}^d \rightarrow \mathbb{R}$ is a measurable function such that

$$\mathbb{E}[\rho_s(1) | Z_s] = \mu_s(Z_s).$$

We then rewrite Zakai equation and the reward functional entirely under the new probability measure $\mathbb{P}$: the term $\mu_s(x) \mathbb{P}_{Z_s}(dx)$ will thus appear in place of $\mathbb{Q}_{Z_s}$. Since the Markov chain M takes value in a finite set $S := \{s_1, \dots, s_N\}$, any solution ρ to Zakai equation, which is a positive measure on S , can easily be identified with a vector X with N nonnegative components, where $X^i := \rho(\mathbb{1}_{\{s_i\}})$. Correspondingly, our control problem can be written in terms of X : in particular, the dynamics evolves in $\mathbb{R}^N$ and can be written with a suitable vector notation. The control problem is now of the following form: maximize

$$J(\alpha) = \mathbb{E} \left[\int_0^T \langle X_s, f(Z_s, \mu_s \mathbb{P}_{Z_s}, \alpha_s) \rangle ds + \langle X_T, g(Z_T, \mu_T \mathbb{P}_{Z_T}) \rangle \right],$$

under the dynamics

$$\begin{cases} dX_s = \Lambda^T X_s ds + \text{diag}(X_s) h(Z_s, \mu_s \mathbb{P}_{Z_s}, \alpha_s) dW_s, \\ \mu_s(x) = \mathbb{E}[\langle X_s, 1 \rangle | Z_s = x], \\ X_0 = \pi_0, \end{cases} \quad (3.1)$$

where Λ is the transition matrix of the Markov chain M , $\text{diag}(x)$ is the $N \times N$ matrix having the entries of the vector x on its diagonal and 0 in all the other entries and

1 is a N -dimensional vector with all the entries equal to 1. Finally, $W_s := Z_s - z_0$ is a standard Brownian motion under $\mathbb{P}$ and the measure $\mu_s(x) \mathbb{P}_{Z_s}(dx)$ is denoted by $\mu_s \mathbb{P}_{Z_s}$.

In optimal control theory of partially observed systems, it is well known that there is a basic problem: the class of admissible control processes α is difficult to characterize, since the control should be predictable with respect to the filtration generated by the process Z , which in its turn depends on α itself. The same difficulty persists in the fully observed control problem obtained after changing the probability measure, as both the Wiener process driving Zakai equation $W = Z - z_0$ and the associated filtration depend on the control α . As this measurability condition is difficult to handle, we proceed in the opposite direction, as often done in filtering theory: we start from a fully observed control problem, called separated problem, on a reference probability space where the controlled Zakai equation (3.1) (rewritten in vector form) is the state equation and the gain function depends on the state process X . In this formulation, the Brownian motion W driving the dynamics is fixed. We then change the underlying probability measure, using again Girsanov's theorem, and we prove that the reformulation of the control problem under the new probability measure is a control problem with partial observation having the same structure of our original control problem (see Proposition 3.3.2). The separated problem is thus equivalent to a weak form of the original partial observation control problem.

The rest of the chapter focuses on the separated control problem. We firstly study its well-posedness: as the state equation is a controlled $\mathbb{R}^N$ -valued SDE with non Lipschitz coefficients depending on the law of Z and on the conditional distribution of X given Z , we need to show that there exists a unique solution to the state equation, satisfying suitable conditions. This is true under an appropriate Lipschitz continuity condition on the function h of the form

$$|h(z, \nu_1, a, i) - h(z, \nu_2, a, i)| \leq L \left(\int_{\mathbb{R}^d} |\phi_1(x) - \phi_2(x)|^2 \gamma_s^{z_0}(dx) \right)^{1/2}$$

for every $z \in \mathbb{R}^d$, $a \in A$, $i \in S$, whenever $s \geq 0$ and $\nu_1, \nu_2 \in \mathcal{P}_q(\mathbb{R}^d)$ are absolutely continuous with respect to a normal distribution with mean z_0 and variance sI , denoted by $\gamma_s^{z_0}$, with densities ϕ_1, ϕ_2 , see Assumption **(A.z₀)**. Intuitively, if we consider two different solutions to equation (3.1), as Z is fixed, we only need to handle the difference between the densities with respect to $\mathbb{P}_{Z_s}$. We subsequently adopt the dynamic programming approach and embed this problem in a larger class of control problems of the McKean-Vlasov type, where the starting time is variable and the initial conditions of the state process and of the Brownian motion may be random, belonging to a suitable L^p space. Notice that, if the state were only described by the solution X to Zakai equation (3.1), its coefficients would depend also on the process Z and they would thus be stochastic. As our aim is to adopt the analytic approach, we want to avoid dealing with stochastic coefficients. We thus describe the state of the system by a pair (Z, X) , where Z is a Wiener process with random initial condition and X is the solution to Zakai equation (3.1), see equation (3.25). We rewrite the state equation proving that it is in fact a McKean-Vlasov equation, whose coefficients depend on the joint law of (Z, X) (see equation (3.35)). In spite of the Lipschitz condition imposed

on h , the coefficients of the state equation are not globally Lipschitz. Therefore, we can not apply standard well-posedness results: hence, we need to state and prove Theorem 3.4.1, guaranteeing that our new state equation admits a unique solution, provided that h satisfies Assumption **(A.β)**, which is the extension of Assumption **(A.z₀)** for random initial conditions. We also identify a suitable set $\mathcal{D}$ of probability measures on $\mathbb{R}^N \times \mathbb{R}^d$ with the property that the law of the initial condition (η, ξ) naturally belongs to $\mathcal{D}$ and we prove that $\mathcal{D}$ is left invariant for any choice of the control (namely, the joint law of (X_s, Z_s) remains in $\mathcal{D}$ for any $s \in [t, T]$). Under additional continuity assumptions on the function h (see Assumption **(A.3)**), we also derive stability properties of the solutions to the equation with respect to the initial conditions. We subsequently prove that, allowing the initial conditions to be random, we restore the flow property. As a consequence, we derive a dynamic programming principle in this setting. Although this is left for future work, it is our purpose to derive a suitable HJB equation on the domain $\mathcal{D}$.

This Chapter is organised as follows. In Section 3.2, we present the original problem with partial observation and we explain how to apply filtering techniques, rewriting it heuristically under a different probability measure. Notice that this procedure only provides a motivation to this work because the subject of our analysis will in fact be the so-called separated control problem, presented in Section 3.3. In Subsection 3.3.1 we discuss its well-posedness, while subsection 3.3.2 is devoted to the analysis of the precise relationship between the original partially observed control problem and the separated control problem. In Section 3.4, we present the aforementioned larger class of optimal control problems; we study the well-posedness of the new problem and the stability of the solutions to the state equation with respect to the initial conditions. Subsection 3.4.2 presents further properties of this new optimal control problem: in particular, we rewrite the equation in a more standard McKean-Vlasov form and we state the flow property and the dynamic programming principle. In Section 3.5 we discuss future developments of this research project with the aim to adopt the analytic approach. The proofs of the main results are postponed to Section 3.6, while in subsection 3.6.3 we provide some examples of functions h satisfying Assumptions **(A.β)** in order to show that this is not a restrictive hypothesis. Indeed, the main restriction to our results is the presence of an identity diffusion coefficient in the original partially observed control problem. Moreover, we also require that h is a bounded function. While slightly more general diffusions could be considered, the boundedness condition on h seems difficult to avoid and is common to many results on similar topics.

Note added in proof: After the first submission of the Thesis, the model presented in this Chapter has been generalized and analysed in the preprint [55], submitted for publication to a peer-reviewed journal. In that paper we have covered the case of a non-constant diffusion and we have finalized the project, developing the results which have been left for future work in this chapter.

3.2 Motivation

Let us consider a controlled state equation

$$\begin{cases} dZ_s = h(Z_s, \mathbb{Q}_{Z_s}, \alpha_s, M_s) ds + dB_s, & s \in [0, T], \\ Z_0 = z_0, \end{cases} \quad (3.2)$$

and a reward functional

$$J(\alpha) = \mathbb{E}_{\mathbb{Q}} \left[\int_0^T f(Z_s, \mathbb{Q}_{Z_s}, \alpha_s, M_s) ds + g(Z_T, \mathbb{Q}_{Z_T}, M_T) \right], \quad (3.3)$$

where $\mathbb{E}_{\mathbb{Q}}$ denotes the expectation under $\mathbb{Q}$ and $\mathbb{Q}_{Z_s}$ the law of Z_s under $\mathbb{Q}$. $z_0 \in \mathbb{R}^d$ is a given starting state, B is a d -dimensional Brownian motion defined on a probability space $(\Omega, \mathcal{F}, \mathbb{Q})$, α is a suitable control process. As the coefficients depend on the marginal laws $\mathbb{Q}_{Z_s}$ of the controlled state, this is a McKean-Vlasov optimal control problem. The coefficients are also subject to random switching, namely they also depend on a piecewise-constant process M . We assume that M is a continuous-time Markov chain on a set S with finite cardinality N , with transition rates $\lambda_{ij} \geq 0$ ($i, j \in S, i \neq j$) and initial distribution π_0 . We also require that it is independent of B .

Assume that the coefficients h, f, g satisfy suitable regularity and growth conditions; in particular, for technical reasons, h is required to be bounded.

Our goal is to model a problem with partial observation, where the source of randomness M is not observed. In this case, the control process α will be predictable with respect to the filtration generated by Z , which represents all the information available to the agent.

Define the exponential process L^{-1} setting

$$L_s^{-1} = \exp \left(- \int_0^s h(Z_r, \mathbb{Q}_{Z_r}, \alpha_r, M_r) dB_r - \frac{1}{2} \int_0^s |h(Z_r, \mathbb{Q}_{Z_r}, \alpha_r, M_r)|^2 dr \right), \quad s \in [0, T]. \quad (3.4)$$

As h is bounded, L^{-1} is an exponential martingale; we can thus introduce the so-called reference probability $\mathbb{P}$ setting $d\mathbb{P} = L_T^{-1} d\mathbb{Q}$. By Girsanov Theorem, Z is a Brownian motion starting from z_0 under $\mathbb{P}$; moreover, applying the techniques presented in [8] and [21], it can be proved that the processes Z and M will still be independent under $\mathbb{P}$ and M remains a Markov chain with transition rates λ_{ij} . We define

$$W_s = Z_s - z_0, \quad s \geq 0,$$

so that, under $\mathbb{P}$, W is a standard Brownian motion in $\mathbb{R}^d$, independent of M , and the filtration generated by Z is the same as the filtration generated by W , denoted by $\mathbb{F}^W = (\mathcal{F}_t^W)$.

Let us introduce the process ρ of unnormalized conditional laws setting, for every $\phi : S \rightarrow \mathbb{R}$,

$$\rho_s(\phi) = \mathbb{E} [\phi(M_s) L_s | \mathcal{F}_s^W], \quad s \in [0, T], \quad (3.5)$$

where $\mathbb{E}$ denotes the expectation under $\mathbb{P}$. This process represents the best approximation that one can get of $\phi(M)$ under the new probability $\mathbb{P}$ given the information available for the agent at present time. We note that $\rho_0 = \pi_0$: indeed, since $L_0 = 1$,

$$\rho_0(\phi) = \mathbb{E}[\phi(M_0) | \mathcal{F}_0^W] = \mathbb{E}[\phi(M_0)] = \mathbb{E}_{\mathbb{Q}}[\phi(M_0)] = \pi_0(\phi).$$

Define as usual $\lambda_{ii} = -\sum_{j \neq i} \lambda_{ij}$, and let Λ denote the matrix $(\lambda_{ij})_{i,j \in S}$. Then by standard arguments in filtering theory the process ρ satisfies the following Zakai equation:

$$d\rho_s(\phi) = \rho_s(\Lambda\phi) ds + \rho_s\left(\phi(\cdot) h(Z_s, \mathbb{Q}_{Z_s}, \alpha_s, \cdot)\right) dW_s, \quad (3.6)$$

for every $\phi : S \rightarrow \mathbb{R}$, where $\Lambda\phi(i) = \sum_{j \in S} \lambda_{ij}\phi(j)$, $i \in S$. Moreover, by an easy computation, the objective functional can be expressed as an expectation under $\mathbb{P}$ in the form

$$J(\alpha) = \mathbb{E} \left[\int_0^T \rho_s \left(f(Z_s, \mathbb{Q}_{Z_s}, \alpha_s, \cdot) \right) ds + \rho_T \left(g(Z_T, \mathbb{Q}_{Z_T}, \cdot) \right) \right]. \quad (3.7)$$

Notice that, in the state equation and in the reformulation of the gain functional, we still see a dependence from $\mathbb{Q}_{Z_s}$, which is a deterministic function of time, but apparently has no explicit expression in terms of the new underlying probability measure $\mathbb{P}$. In fact, it is possible to express it in terms of $\mathbb{P}_{Z_s}$, in order to arrive at a problem entirely formulated under $\mathbb{P}$. For every bounded measurable $\psi : \mathbb{R}^d \rightarrow \mathbb{R}$,

$$\begin{aligned} \int_{\mathbb{R}^d} \psi(x) \mathbb{Q}_{Z_s}(dx) &= \mathbb{E}_{\mathbb{Q}}[\psi(Z_s)] = \mathbb{E}[\psi(Z_s) L_s] = \mathbb{E}[\psi(Z_s) \mathbb{E}[L_s | \mathcal{F}_s^W]] \\ &= \mathbb{E}[\psi(Z_s) \rho_s(1)] = \mathbb{E}[\psi(Z_s) \mathbb{E}[\rho_s(1) | Z_s]]. \end{aligned}$$

Let $\mu_s : \mathbb{R}^d \rightarrow \mathbb{R}$ be a measurable function such that

$$\mathbb{E}[\rho_s(1) | Z_s] = \mu_s(Z_s).$$

Symbolically, $\mu_s(x) = \mathbb{E}[\rho_s(1) | Z_s = x]$. Then

$$\int_{\mathbb{R}^d} \psi(x) \mathbb{Q}_{Z_s}(dx) = \mathbb{E}[\psi(Z_s) \mu_s(Z_s)] = \int_{\mathbb{R}^d} \psi(x) \mu_s(x) \mathbb{P}_{Z_s}(dx).$$

So we conclude that

$$\mathbb{Q}_{Z_s}(dx) = \mu_s(x) \mathbb{P}_{Z_s}(dx).$$

We denote shortly by $\mu_s \mathbb{P}_{Z_s}$ the latter probability measure. Recalling that Z is a Brownian motion starting at z_0 , $\mathbb{P}_{Z_s}$ is simply the Gaussian law in $\mathbb{R}^d$ with mean z_0 and variance sI .

Then the Zakai equation and the reward functional can be written in the form

$$\begin{cases} d\rho_s(\phi) &= \rho_s(\Lambda\phi) ds + \rho_s\left(\phi(\cdot) h(Z_s, \mu_s \mathbb{P}_{Z_s}, \alpha_s, \cdot)\right) dW_s, \\ \mu_s(x) &= \mathbb{E}[\rho_s(1) | Z_s = x], \\ \rho_0 &= \pi_0, \end{cases} \quad (3.8)$$

$$J(\alpha) = \mathbb{E} \left[\int_0^T \rho_s \left(f(Z_s, \mu_s \mathbb{P}_{Z_s}, \alpha_s, \cdot) \right) ds + \rho_T \left(g(Z_T, \mu_T \mathbb{P}_{Z_T}, \cdot) \right) \right]. \quad (3.9)$$

The latter equations show that the original optimization problem is now reduced to a problem with full observation, driven by the Brownian motion W , where the control α is predictable with respect to the corresponding filtration $\mathbb{F}^W$, the state equation being given by (3.8) and the reward functional by (3.9).

Before proceeding further we exploit the fact that the number N of states of the Markov chain M is finite, which allows to rewrite (3.8) as an equation with values in $\mathbb{R}^N$. We identify the set of positive measures on S with the set $\mathbb{R}_+^N$. Under this identification, the positive measure-valued process ρ corresponds to the $\mathbb{R}_+^N$ -valued process X defined by

$$X_s^i := \rho_s(\mathbb{1}_{\{i\}}),$$

where $\mathbb{1}_{\{i\}}$ denotes the indicator function of the state $i \in S$.

Then the Zakai equation takes the following form: for every $i \in S$,

$$\begin{cases} dX_s^i &= \sum_{j \in S} \lambda_{ji} X_s^j ds + X_s^i h(Z_s, \mu_s \mathbb{P}_{Z_s}, \alpha_s, i) dW_s, \\ \mu_s(x) &= \mathbb{E} \left[\sum_{j \in S} X_s^j \mid Z_s = x \right], \\ \rho_0 &= \pi_0. \end{cases} \quad (3.10)$$

and we look for a solution X taking values in $\mathbb{R}_+^N$. The reward functional becomes

$$J(\alpha) = \mathbb{E} \left[\int_0^T \sum_{j \in S} X_s^j f(Z_s, \mu_s \mathbb{P}_{Z_s}, \alpha_s, j) ds + \sum_{j \in S} X_T^j g(Z_T, \mu_T \mathbb{P}_{Z_T}, j) \right]. \quad (3.11)$$

It is convenient to use a vector notation. For $x \in \mathbb{R}^d$, $a \in A$ and any probability μ in $\mathbb{R}^d$, we define $\mathbb{R}^N$ -valued functions setting

$$X_s = (X_s^i)_{i \in S}, \quad f(x, \mu, a) = (f(x, \mu, a, i))_{i \in S}, \quad g(x, \mu) = (g(x, \mu, i))_{i \in S},$$

and, denoting h^k ($k = 1, \dots, d$) the components of the function h , we also define the $N \times d$ matrix

$$h(x, \mu, a) = (h^k(x, \mu, a, i))_{i \in S; k=1, \dots, d}.$$

We can rewrite the controlled state equation and the reward functional in the form

$$\begin{cases} dX_s &= \Lambda^T X_s ds + \text{diag}(X_s) h(Z_s, \mu_s \mathbb{P}_{Z_s}, \alpha_s) dW_s, \\ \mu_s(x) &= \mathbb{E} [\langle X_s, 1 \rangle \mid Z_s = x], \\ X_0 &= \pi_0, \end{cases} \quad (3.12)$$

$$J(\alpha) = \mathbb{E} \left[\int_0^T \langle X_s, f(Z_s, \mu_s \mathbb{P}_{Z_s}, \alpha_s) \rangle ds + \langle X_T, g(Z_T, \mu_T \mathbb{P}_{Z_T}) \rangle \right], \quad (3.13)$$

where Λ^T is the transpose matrix of Λ , $\text{diag}(X_s)$ denotes the $N \times N$ matrix with the vector X_s on the diagonal and the other entries equal to zero, 1 denotes the vector in $\mathbb{R}^N$ with unit entries, $\langle \cdot, \cdot \rangle$ denotes the scalar product in $\mathbb{R}^N$.

The reformulated optimization problem (3.12)-(3.13) retains a specific difficulty. Although this is not explicit in the notation we have used, the process W and the filtration $\mathbb{F}^W$ depend on the control $\alpha \in \mathcal{A}^W$, which is in its turn required to be adapted to $\mathbb{F}^W$. Thus, the description of admissible controls is rather involved and difficult to handle. In order to circumvent this difficulty, using a standard methodology, in the following section we will formulate an optimization problem, called the separated problem, starting from the reference probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and the equations (3.12)-(3.13) as the controlled equation and the reward functional. Under suitable conditions on the data we will prove that the separated problem is equivalent to the original partially observed optimization problem, provided that the latter is formulated in a suitable weak form.

3.3 The separated control problem

3.3.1 Assumptions and formulation

Assumptions (A.1). (i) $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space.

(ii) M is a continuous-time Markov chain on a finite set S with cardinality N . For simplicity, we identify S with the set $\{1, \dots, N\}$. Moreover, M has transition rates $\lambda(i, j)$ ($i, j \in S, i \neq j$) and initial distribution π_0 . We define $\lambda(i, i) = -\sum_{j \neq i} \lambda(i, j)$, and denote by Λ the $N \times N$ matrix with entries $\lambda(i, j)$.

(iii) W is a standard Brownian motion in $\mathbb{R}^d$, independent of M .

(iv) A , the set of control actions, is a Polish space. $T > 0$ denotes the time horizon of the control problem.

(v) For some $q \geq 1$, the function $h : \mathbb{R}^d \times \mathcal{P}_q(\mathbb{R}^d) \times A \times S \rightarrow \mathbb{R}^d$ is measurable and bounded.

(vi) The functions

$$f : \mathbb{R}^d \times \mathcal{P}_q(\mathbb{R}^d) \times A \times S \rightarrow \mathbb{R}, \quad g : \mathbb{R}^d \times \mathcal{P}_q(\mathbb{R}^d) \times S \rightarrow \mathbb{R},$$

are measurable and satisfy a polynomial growth condition in (x, μ) uniformly with respect to a , i.e. there exist a constant $K > 0$ and an integer $r \geq 1$ such that

$$|f(x, \mu, a, i)| + |g(x, \mu, i)| \leq K(1 + |x|^r + \|\mu\|_q^r) \quad (3.14)$$

for all $x \in \mathbb{R}^d$, $\mu \in \mathcal{P}_q(\mathbb{R}^d)$, $a \in A$ and $i \in S$, where we define as usual

$$\|\mu\|_q := \left(\int_{\mathbb{R}^d} |x|^q \mu(dx) \right)^{\frac{1}{q}} \quad \text{for all } \mu \in \mathcal{P}_q(\mathbb{R}^d).$$

(vii) $z_0 \in \mathbb{R}^d$ is a (deterministic) initial state.

(viii) p is a positive exponent such that $p \geq 2$.

We define the set $\mathcal{A}^W$ of admissible control processes as the collection of all processes $\alpha : \Omega \times [0, T] \rightarrow A$ which are predictable with respect to $\mathbb{F}^W = (\mathcal{F}_t^W)_{t \geq 0}$, the completed Brownian filtration of W . We also define

$$Z_s = z_0 + W_s, \quad s \geq 0. \quad (3.15)$$

We rewrite f and g as $\mathbb{R}^N$ -valued functions setting, for $x \in \mathbb{R}^d$, $a \in A$ and $\mu \in \mathcal{P}_q(\mathbb{R}^d)$,

$$f(x, \mu, a) = (f(x, \mu, a, i))_{i=1}^N, \quad g(x, \mu) = (g(x, \mu, i))_{i=1}^N.$$

Furthermore, we identify h with a $\mathbb{R}^{N \times d}$ -valued function

$$h(x, \mu, a) = (h^k(x, \mu, a, i))_{\substack{i=1, \dots, N, \\ k=1, \dots, d}},$$

where h^k ($k = 1, \dots, d$) are the components of h .

Call $\mathbf{1}$ the vector in $\mathbb{R}^N$ with all entries equal to 1 and let $\langle \cdot, \cdot \rangle$ denote the scalar product in $\mathbb{R}^N$. For any random variable $Y : \Omega \rightarrow \mathbb{R}^d$, $Y \in L^1(\Omega, \mathbb{F}, \mathbb{P})$ such that, for all $s \in [0, T]$, $\mathbb{E}[\langle Y, \mathbf{1} \rangle | Z_s] \geq 0$ $\mathbb{P}$ -a.s. and $\mathbb{E}[\langle Y, \mathbf{1} \rangle] = 1$, define

$$x \mapsto \mathbb{E}[\langle Y, \mathbf{1} \rangle | Z_s = x] \quad (3.16)$$

as any $\sigma(Z_s)$ -measurable function $\mu(x)$ such that $\mathbb{E}[\langle Y, \mathbf{1} \rangle | Z_s] = \mu(Z_s)$; μ is in fact defined up to $\mathbb{P}_{Z_s}$ -null sets. We also denote by

$$\mathbb{E}[\langle Y, \mathbf{1} \rangle | Z_s = x] \mathbb{P}_{Z_s}(dx) \quad \text{or simply} \quad \mathbb{E}[\langle Y, \mathbf{1} \rangle | Z_s = \cdot] \mathbb{P}_{Z_s}$$

the probability measure on $\mathbb{R}^d$ being absolutely continuous with respect to $\mathbb{P}_{Z_s}$ and admitting the density $\mathbb{E}[\langle Y, \mathbf{1} \rangle | Z_s = x]$ with respect to $\mathbb{P}_{Z_s}(dx)$.

We are ready to state the separated optimal control problem: for $\alpha \in \mathcal{A}^W$ the state equation is

$$\begin{cases} dX_s &= \Lambda^T X_s ds + \text{diag}(X_s) h(Z_s, \mathbb{E}[\langle X_s, \mathbf{1} \rangle | Z_s = \cdot] \mathbb{P}_{Z_s}, \alpha_s) dW_s, \\ X_0 &= \pi_0, \end{cases} \quad (3.17)$$

and the reward functional is

$$\begin{aligned} J(\alpha) = \mathbb{E} \bigg[& \int_0^T \langle X_s, f(Z_s, \mathbb{E}[\langle X_s, \mathbf{1} \rangle | Z_s = \cdot] \mathbb{P}_{Z_s}, \alpha_s) \rangle ds \\ & + \langle X_T, g(Z_T, \mathbb{E}[\langle X_T, \mathbf{1} \rangle | Z_T = \cdot] \mathbb{P}_{Z_T}) \rangle \bigg]. \end{aligned} \quad (3.18)$$

As before, Λ^T is the transpose matrix of Λ , $\text{diag}(X_s)$ denotes the $N \times N$ matrix with the vector X_s on the diagonal and the other entries equal to zero. We note that in (3.17) the initial distribution π_0 is identified with a vector in $\mathbb{R}^N$ defined by $\pi_0(i) := \pi_0(\{i\})$. As π_0 is a probability measure on S , the corresponding vector must satisfy the following conditions:

$$\pi_0(i) \geq 0 \quad \text{for all } i = 1, \dots, N, \quad \sum_{i=1}^N \pi_0(i) = 1,$$

i.e. π_0 is a vector in $\mathbb{R}_+^N$ satisfying $\langle \pi_0, 1 \rangle = 1$.

Notice that equation (3.17) is not a standard McKean-Vlasov equation, for which we can rely on classical results of existence and uniqueness of solutions; we thus firstly need to discuss its well-posedness. In what follows, we will need to impose some form of Lipschitz condition on the function h . For this section we will consider Assumption **(A.z₀)**; it may be useful to recall that $\mathbb{P}_{Z_s}$ is simply the Gaussian law in $\mathbb{R}^d$ with mean z_0 and variance sI , which will be denoted by $\gamma_s^{z_0}$.

Assumption (A.z₀). There exists a constant L such that, whenever $s \geq 0$ and $\nu_1, \nu_2 \in \mathcal{P}_q(\mathbb{R}^d)$ are absolutely continuous with respect to $\gamma_s^{z_0}$ with densities ϕ_1, ϕ_2 , respectively, the inequality

$$|h(z, \nu_1, a, i) - h(z, \nu_2, a, i)| \leq L \left(\int_{\mathbb{R}^d} |\phi_1(x) - \phi_2(x)|^2 \gamma_s^{z_0}(dx) \right)^{1/2}$$

holds for every $z \in \mathbb{R}^d$, $a \in A$, $i \in S$.

We will provide some examples of functions satisfying Assumption **(A.z₀)** in Section 3.6.3.

We now state a well-posedness result. We will prove it in greater generality in Section 3.6; we will thus assume it for the moment.

Theorem 3.3.1. Under Assumptions **(A.1)** and **(A.z₀)**, for every $\alpha \in \mathcal{A}^W$ there exists a unique solution X to equation (3.17), $\mathbb{F}^W$ -adapted with continuous trajectories, such that $\langle X, 1 \rangle$ is a martingale and

$$\sup_{s \in [0, T]} \mathbb{E} [|X_s|^p] < \infty, \quad \mathbb{P}(X_s \in \mathbb{R}_+^N) = 1, \quad \mathbb{E} [\langle X_s, 1 \rangle] = 1, \quad s \in [0, T].$$

We underline that, if X is a process satisfying

$$\sup_{s \in [0, T]} \mathbb{E} [|X_s|^p] < \infty, \quad \mathbb{P}(X_s \in \mathbb{R}_+^N) = 1, \quad \mathbb{E} [\langle X_s, 1 \rangle] = 1, \quad s \in [0, T],$$

the measure $\mathbb{E} [\langle X_s, 1 \rangle | Z_s = \cdot] \mathbb{P}_{Z_s}$ is in $\mathcal{P}_q(\mathbb{R}^d)$ for any $q \geq 1$. As a matter of fact, it is a probability measure because

$$\begin{aligned} \int_{\mathbb{R}^d} \mathbb{E} [\langle X_s, 1 \rangle | Z_s = z] \mathbb{P}_{Z_s}(dz) &= \mathbb{E} [\mathbb{E} [\langle X_s, 1 \rangle | Z_s]] \\ &= \mathbb{E} [\langle X_s, 1 \rangle] = 1 \end{aligned}$$

by the tower property, as $\langle X, 1 \rangle$ is a martingale and $\langle \pi_0, 1 \rangle = 1$. Moreover,

$$\begin{aligned} \int_{\mathbb{R}^d} |z|^q \mathbb{E} [\langle X_s, 1 \rangle | Z_s = z] \mathbb{P}_{Z_s}(dz) &= \mathbb{E} [|Z_s|^q \mathbb{E} [\langle X_s, 1 \rangle | Z_s]] = \mathbb{E} [|Z_s|^q \langle X_s, 1 \rangle] \\ &\leq C_q \mathbb{E} [|z_0|^q \langle X_s, 1 \rangle] + C_q \mathbb{E} [|W_s|^q \langle X_s, 1 \rangle], \end{aligned}$$

which is finite by the Hölder and Cauchy-Schwartz inequalities since $\mathbb{E} [|X_s|^p] < \infty$ and $\mathbb{E} [|W_s|^n] < \infty$ for all $n \geq 1$ and $\mathbb{E} [\langle X_s, 1 \rangle] = 1$. Therefore, as f, g satisfy polynomial growth conditions detailed in Assumptions **(A.1)**, we can prove by standard estimates that the separated problem is well defined and its value can be introduced:

$$V_0 = \sup_{\alpha \in \mathcal{A}^W} J(\alpha). \quad (3.19)$$

3.3.2 Equivalence with the partially observed control problem

In this paragraph we clarify the connection of the separated control problem and the original partially observed optimal control problem (3.2)-(3.3). Throughout, we suppose that Assumptions **(A.1)** and **(A.z₀)** hold true.

Denote by $\mathbb{F}^M = (\mathcal{F}_t^M)_{t \geq 0}$ the filtration generated by the Markov chain M and by $\mathbb{G} = (\mathcal{G}_t)_{t \geq 0}$ the filtration given by $\mathcal{G}_t = \mathcal{F}_t^W \vee \mathcal{F}_t^M$. Using the independence of W and M it is possible to verify that W is a Wiener process with respect to $\mathbb{G}$. For every $\alpha \in \mathcal{A}^W$, we denote by X^α the solution to the state equation (3.17) given by Theorem 3.3.1. Let us also define the processes

$$h_s^\alpha = h\left(Z_s, \mathbb{E}[\langle X_s^\alpha, 1 \rangle \mid Z_s = \cdot] \mathbb{P}_{Z_s, \alpha_s, M_s}\right), \quad s \in [0, T],$$

$$L_s^\alpha = \exp\left(\int_0^s h_r^\alpha dW_r - \frac{1}{2} \int_0^s |h_r^\alpha|^2 dr\right), \quad B_s^\alpha = W_s - \int_0^s h_r^\alpha dr \quad s \in [0, T].$$

The process h^α is progressively measurable with respect to $\mathbb{G}$, so that the stochastic integral in the definition of L^α is well defined. Since h^α is also bounded, L^α is a $\mathbb{G}$ -martingale and we can define a probability $\mathbb{Q}^\alpha$ setting $d\mathbb{Q}^\alpha = L_T^\alpha d\mathbb{P}$. With this notation we have the following result.

Proposition 3.3.2. (i) Under $\mathbb{Q}^\alpha$, the process B^α is a $\mathbb{G}$ -Brownian motion.

(ii) Under $\mathbb{Q}^\alpha$, the process M remains a Markov chain with transition rates $\lambda(i, j)$ and initial distribution π_0 .

(iii) M and B^α are independent under $\mathbb{Q}^\alpha$.

(iv) For $i \in S$ and $s \in [0, T]$ we have

$$X_s^i = \mathbb{E}[\mathbb{1}_{\{M_s=i\}} L_s^\alpha \mid \mathcal{F}_s^W]. \quad (3.20)$$

(v) The process Z satisfies

$$\begin{cases} dZ_s = h(Z_s, \mathbb{Q}_{Z_s}^\alpha, \alpha_s, M_s) ds + dB_s^\alpha, & s \in [0, T], \\ Z_0 = z_0, \end{cases} \quad (3.21)$$

(vi) The reward functional $J(\alpha)$ defined in (3.18) satisfies

$$J(\alpha) = \mathbb{E}_{\mathbb{Q}^\alpha} \left[\int_0^T f(Z_s, \mathbb{Q}_{Z_s}^\alpha, \alpha_s, M_s) ds + g(Z_T, \mathbb{Q}_{Z_T}^\alpha, M_T) \right]. \quad (3.22)$$

Proof. (i) The thesis follows immediately from the Girsanov theorem.

(ii) Since $L_0^\alpha = 1$, $\mathbb{P}$ and $\mathbb{Q}^\alpha$ agree on $\mathcal{G}_0$. It follows that the initial distribution of M under $\mathbb{P}$ and $\mathbb{Q}^\alpha$ is the same and it is equal to π_0 .

For any $\phi : S \rightarrow \mathbb{R}$, we define the process

$$M_t^\phi := \phi(M_t) - \phi(M_0) - \int_0^t \Lambda \phi(M_s) ds.$$

By Theorem 20.6 in [83], given the initial distribution π_0 of M , the probability measure $\mathbb{P}$ is the unique probability measure under which, for all $\phi : S \rightarrow \mathbb{R}$, M^ϕ is a $\mathbb{F}^M$ -martingale. If we prove that M^ϕ are also $(\mathbb{F}^M, \mathbb{Q}^\alpha)$ -martingales for all $\phi : S \rightarrow \mathbb{R}$, we conclude that $\mathbb{P} = \mathbb{Q}^\alpha$ on $\mathbb{F}^M$ and the law of M is thus the same under the two probability measures.

Fix any $\phi : S \rightarrow \mathbb{R}$. By the independence of M and W , we can easily deduce that M^ϕ is also a $(\mathbb{G}, \mathbb{P})$ -martingale. Since L^α is a continuous $\mathbb{G}$ -martingale while M^ϕ has finite variation trajectories, by Theorem 10.2.21 in [40] they are orthogonal, meaning that their product $M^\phi L^\alpha$ is a $\mathbb{G}$ -martingale under $\mathbb{P}$. Then, by Lemma 12.1 in [9], M^ϕ is a $\mathbb{G}$ -martingale under $\mathbb{Q}^\alpha$. Since M^ϕ is adapted to $\mathbb{F}^M$, we easily conclude that M^ϕ is a $\mathbb{F}^M$ -martingale under $\mathbb{Q}^\alpha$ because, for all $0 \leq s \leq t$, by the tower property

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}^\alpha} \left[M_t^\phi | \mathcal{F}_s^M \right] &= \mathbb{E}_{\mathbb{Q}^\alpha} \left[\mathbb{E}_{\mathbb{Q}^\alpha} \left[M_t^\phi | \mathcal{G}_s \right] | \mathcal{F}_s^M \right] \\ &= \mathbb{E}_{\mathbb{Q}^\alpha} \left[M_s^\phi | \mathcal{F}_s^M \right] \\ &= M_s^\phi \end{aligned}$$

since M^ϕ is $\mathbb{F}^M$ -adapted. By the characterization property mentioned above, M thus is a Markov chain with transition rates $\lambda(i, j)$ under $\mathbb{Q}^\alpha$.

- (iii) Let us define another filtration $\mathbb{H} = (\mathcal{H}_t)_{t \geq 0}$ setting $\mathcal{H}_t = \mathcal{F}_t^W \vee \mathcal{F}_\infty^M$. Since $\mathbb{H}$ is an enlargement of $\mathbb{G}$, the process h^α is also progressively measurable with respect to $\mathbb{H}$. Moreover, since W is independent of M , W is a Wiener process under $\mathbb{P}$ with respect to $\mathbb{H}$ as well. We may now apply the Girsanov theorem with the filtration $\mathbb{H}$ and conclude that B^α is a Wiener process with respect to $\mathbb{H}$ under $\mathbb{Q}^\alpha$. In particular, it is independent of $\mathcal{H}_0$ under $\mathbb{Q}^\alpha$. Since $\mathcal{F}_\infty^M \subset \mathcal{H}_0$ (in fact, they are equal up to null sets), we conclude that B^α is independent of M under $\mathbb{Q}^\alpha$.

- (iv) By the definition of B^α we have $dZ_s = h_s^\alpha ds + dB_s^\alpha$, namely Z satisfies

$$\begin{cases} dZ_s &= h\left(Z_s, \mathbb{E}[\langle X_s^\alpha, 1 \rangle | Z_s = \cdot] \mathbb{P}_{Z_s, \alpha_s, M_s}\right) ds + dB_s^\alpha, \\ Z_0 &= z_0. \end{cases} \quad (3.23)$$

We look at the processes M, Z from the point of view of filtering theory: in the probability space $(\Omega, \mathcal{F}, \mathbb{Q}^\alpha)$, B^α is the Brownian motion driving the dynamics of Z , the control α is adapted to the filtration generated by Z , which thus corresponds to the observation process, and M is the (unobserved) signal process. This is thus a partially observed controlled dynamical system. Notice that the processes M, Z, α, B^α are $\mathbb{G}$ -progressive and recall that, by (i)-(iii) in this Proposition, M is a Markov chain and B^α is an independent Brownian motion. Introduce the process

$$(L_s^\alpha)^{-1} = \exp \left(- \int_0^s h_r^\alpha dB_r^\alpha - \frac{1}{2} \int_0^s |h_r^\alpha|^2 dr \right), \quad s \in [0, T],$$

which is a $\mathbb{G}$ -martingale under $\mathbb{Q}^\alpha$, and notice that $d\mathbb{P} = (L_T^\alpha)^{-1} d\mathbb{Q}^\alpha$. For $i \in S$, let us denote by ρ_s^i the unnormalized conditional law of M_s given the present observation, namely $\rho_s^i = \mathbb{E}[\mathbb{1}_{\{M_s=i\}} L_s^\alpha | \mathcal{F}_s^W]$. More precisely, we define ρ^i to be the optional projection of $\mathbb{1}_{\{M=i\}} L^\alpha$ on the filtration $\mathbb{F}^W$. Noting that α is $\mathbb{F}^W$ -predictable, we can deduce by standard arguments (see e.g. [8]) that the Zakai filtering equation holds: in particular, for every $i \in S$,

$$\begin{cases} d\rho_s^i &= \sum_{j \in S} \lambda_{ji} \rho_s^j ds + \rho_s^i h\left(Z_s, \mathbb{E}[\langle X_s^\alpha, 1 \rangle | Z_s = \cdot] \mathbb{P}_{Z_s, \alpha_s, i}\right) dW_s, \\ \rho_0 &= \pi_0 \end{cases}$$

under $\mathbb{P}$. Since this equation coincides with the equation (3.17) satisfied by X^α under $\mathbb{P}$, we conclude that ρ^i coincides with $(X^\alpha)^i$ for $i = 1, \dots, N$ and point (iv) is proved.

(v) In view of (3.23) it is enough to prove the equality

$$\mathbb{E}[\langle X_s^\alpha, 1 \rangle | Z_s = \cdot] \mathbb{P}_{Z_s} = \mathbb{Q}_{Z_s}^\alpha. \quad (3.24)$$

For any bounded measurable $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$,

$$\int_{\mathbb{R}^d} \phi(z) \mathbb{Q}_{Z_s}^\alpha(dz) = \mathbb{E}_{\mathbb{Q}^\alpha}[\phi(Z_s)] = \mathbb{E}[\phi(Z_s) L_s^\alpha] = \mathbb{E}[\phi(Z_s) \mathbb{E}[L_s^\alpha | \mathcal{F}_s^W]].$$

From point (iv) it follows that $\mathbb{E}[L_s^\alpha | \mathcal{F}_s^W] = \langle X_s^\alpha, 1 \rangle$ and we obtain

$$\begin{aligned} \int_{\mathbb{R}^d} \phi(z) \mathbb{Q}_{Z_s}^\alpha(dz) &= \mathbb{E}[\phi(Z_s) \langle X_s^\alpha, 1 \rangle] = \mathbb{E}[\phi(Z_s) \mathbb{E}[\langle X_s^\alpha, 1 \rangle | Z_s]] \\ &= \int_{\mathbb{R}^d} \phi(z) \mathbb{E}[\langle X_s^\alpha, 1 \rangle | Z_s = z] \mathbb{P}_{Z_s}(dz), \end{aligned}$$

which proves (3.24).

(vi) The thesis follows by standard arguments. Compute

$$\begin{aligned} &\mathbb{E}_{\mathbb{Q}^\alpha} \left[\int_0^T f(Z_s, \mathbb{Q}_{Z_s}^\alpha, \alpha_s, M_s) ds \right. \\ &\quad \left. + g(Z_T, \mathbb{Q}_{Z_T}^\alpha, M_T) \right] \quad (d\mathbb{Q}^\alpha = L_T^\alpha d\mathbb{P}) \\ &= \mathbb{E} \left[\int_0^T f(Z_s, \mathbb{Q}_{Z_s}^\alpha, \alpha_s, M_s) L_s^\alpha ds \right. \\ &\quad \left. + g(Z_T, \mathbb{Q}_{Z_T}^\alpha, M_T) L_T^\alpha \right] \quad (\text{tower property and Fubini theorem}) \\ &= \mathbb{E} \left[\int_0^T \mathbb{E} \left[f(Z_s, \mathbb{Q}_{Z_s}^\alpha, \alpha_s, M_s) L_s^\alpha | \mathcal{F}_s^W \right] ds \right] \\ &\quad + \mathbb{E} \left[\mathbb{E} \left[g(Z_T, \mathbb{Q}_{Z_T}^\alpha, M_T) L_T^\alpha | \mathcal{F}_T^W \right] \right] \\ &= \mathbb{E} \left[\int_0^T \langle X_s^\alpha, f(Z_s, \mathbb{Q}_{Z_s}^\alpha, \alpha_s) \rangle ds \right] \end{aligned}$$

$$+\mathbb{E}\left[\langle X_T^\alpha, g(Z_T, \mathbb{Q}_{Z_T}^\alpha) \rangle\right],$$

where we use the fact that, for all $\phi : S \rightarrow \mathbb{R}$,

$$\begin{aligned}\mathbb{E}[\phi(M_s)L_s^\alpha|\mathcal{F}_s^W] &= \sum_{i=1}^N \mathbb{E}[\mathbb{1}_{\{M_s=i\}}\phi(i)L_s^\alpha|\mathcal{F}_s^W] \\ &= \sum_{i=1}^N \phi(i)(X_s^\alpha)^i \\ &= \langle X_s^\alpha, \phi \rangle\end{aligned}$$

by point (iv). The thesis then follows from equality (3.24). □

Remark 3.3.3. The proofs of (i) – (iii) also essentially follow from Proposition 3.13 in the book [8].

This proposition shows that the separated problem introduced above is equivalent to a weak version of the partially observed control problem. Indeed, for every admissible control α , the reward functional $J(\alpha)$ of the separated problem can be written via equations (3.21)-(3.22) which coincide with (3.2)-(3.3), except for the fact that the Brownian motion B^α , as well as the probability $\mathbb{Q}^\alpha$, depend on α . On the contrary, in the separated problem the process W and the filtration $\mathbb{F}^W$ do not depend on α (which was not the case for the original partially observed control problem).

In the rest of this paper we will only deal with the separated control problem (3.17)-(3.18) or its extensions, leaving aside the original partially observed problem.

3.4 Dynamic programming approach

3.4.1 A larger class of control problems

In this section we start the study of the separated control problem (3.17)-(3.18) by the dynamic programming approach. The first step consists in embedding the separated problem into a larger family of optimization problems. Firstly, we will consider a dynamical system starting from time $t \in [0, T]$ instead of time 0. Furthermore, in order to deal with the dependence of the coefficients on the marginal laws of the controlled process, we also need to replace the deterministic initial conditions z_0 and π_0 (for the processes Z and X) by random variables ξ and η . It is convenient to assume that they depend on an additional source of noise and for this reason, beside Assumption (A.1), we will assume the following.

Assumption (A.2). On the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ there exists a real random variable U having uniform distribution on $(0, 1)$ and independent of M, W . We introduce the filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$, where $\mathcal{F}_t = \mathcal{F}_t^W \vee \sigma(U)$.

Finally, it is also convenient to consider a larger class of admissible control processes: we denote by $\mathcal{A}$ the collection of all processes $\alpha : \Omega \times [0, T] \rightarrow A$ which are predictable with respect to $\mathbb{F}$.

We start by rewriting the state equation (3.17) incorporating those extensions and in a slightly different form: for any time $t \in [0, T]$ and any $\alpha \in \mathcal{A}$ we consider the following system for a controlled pair $(X_s, Z_s)_{s \in [t, T]}$ with values in $\mathbb{R}^N \times \mathbb{R}^d$:

$$\begin{cases} dX_s &= \Lambda^T X_s ds + \text{diag}(X_s) h\left(Z_s, \mathbb{E}[\langle X_s, 1 \rangle | Z_s = z] \mathbb{P}_{Z_s}(dz), \alpha_s\right) dW_s, \\ dZ_s &= dW_s, \\ X_t &= \eta, \\ Z_t &= \xi, \end{cases} \quad (3.25)$$

and the reward functional

$$\begin{aligned} J(t, \eta, \xi, \alpha) &= \mathbb{E} \left[\int_t^T \langle X_s, f\left(Z_s, \mathbb{E}[\langle X_s, 1 \rangle | Z_s = z] \mathbb{P}_{Z_s}(dz), \alpha_s\right) \rangle ds \right. \\ &\quad \left. + \langle X_T, g\left(Z_T, \mathbb{E}[\langle X_T, 1 \rangle | Z_T = z] \mathbb{P}_{Z_T}(dz)\right) \rangle \right]. \end{aligned} \quad (3.26)$$

The equations (3.25)-(3.26) provide the form of the separated problem required for dynamic programming arguments. Note that, in order to reduce to a Markovian setting without stochastic coefficients, we introduce the process Z as another component of the state, although it is not affected by the control and it is given explicitly by the formula

$$Z_s = \xi + W_s - W_t, \quad s \in [t, T]. \quad (3.27)$$

The random initial condition (ξ, η) , valued in $\mathbb{R}^N \times \mathbb{R}^d$, is assumed to be $\mathcal{F}_t$ -measurable and to satisfy appropriate integrability conditions stated below.

Our first result is well-posedness of the state equation (3.25). To this end we need a Lipschitz condition which generalizes Assumption **(A.z₀)**. Recall that the numbers $p \geq 2$ and $q \geq 1$ were introduced in Assumption **(A.1)**.

Let us denote by β the law of ξ . Since ξ is $\mathcal{F}_t$ -measurable, it is independent of $W_s - W_t$ for $s \geq t$ and therefore, recalling (3.27), we see that $\mathbb{P}_{Z_s} = \gamma_{s-t}^0 * \beta$, where γ_{s-t}^0 denotes the centered Gaussian law in $\mathbb{R}^d$ with variance $(s-t)I$ and $*$ denotes convolution.

Assumption (A.β). There exists a constant L such that, whenever $s \geq 0$ and $\nu_1, \nu_2 \in \mathcal{P}_q(\mathbb{R}^d)$ are absolutely continuous with respect to $\gamma_s^0 * \beta$ with densities ϕ_1, ϕ_2 respectively, the inequality

$$|h(z, \nu_1, a, i) - h(z, \nu_2, a, i)| \leq L \left(\int_{\mathbb{R}^d} |\phi_1(x) - \phi_2(x)|^2 (\gamma_s^0 * \beta)(dx) \right)^{1/2}$$

holds for every $z \in \mathbb{R}^d$, $a \in A$, $i \in S$.

Notice that this assumption makes sense for every probability measure β on $\mathbb{R}^d$ and that Assumption **(A.z₀)** coincides with Assumption **(A.β)** when $\beta = \delta_{z_0}$. We also

underline that, in spite of the boundedness of h and Assumption **(A.β)**, the controlled state equation does not satisfy a global Lipschitz condition. We refer the reader to Section 3.6.3 for some examples of functions satisfying Assumption **(A.β)**.

Theorem 3.4.1. Suppose that Assumptions **(A.1)**, **(A.2)** hold true. Let η be $\mathbb{R}^N$ -valued and $\mathcal{F}_t$ -measurable satisfying

$$\mathbb{E} [|\eta|^p] < \infty, \quad \mathbb{P}(\eta \in \mathbb{R}_+^N) = 1, \quad \mathbb{E} [\langle \eta, 1 \rangle] = 1.$$

Let ξ be $\mathbb{R}^d$ -valued and $\mathcal{F}_t$ -measurable satisfying

$$\mathbb{E} [|\xi|^q \langle \eta, 1 \rangle] < \infty. \quad (3.28)$$

Finally suppose that Assumption **(A.β)** holds for the probability law $\beta = \mathbb{P}_\xi$.

Then, for every $\alpha \in \mathcal{A}$, there exists a unique solution (X, Z) to equation (3.25), where Z is given by (3.27) and X is $\mathbb{F}$ -adapted with continuous trajectories, $\langle X, 1 \rangle$ is a martingale, and

$$\sup_{s \in [t, T]} \mathbb{E} [|X_s|^p] < \infty, \quad \mathbb{P}(X_s \in \mathbb{R}_+^N) = 1, \quad \mathbb{E} [\langle X_s, 1 \rangle] = 1, \quad s \in [t, T].$$

Moreover there exists a constant C_0 , depending only on $p, T, \Lambda, \sup |h|$, such that

$$\sup_{s \in [t, T]} \mathbb{E} [|X_s|^p] \leq C_0 \mathbb{E} [|\eta|^p]. \quad (3.29)$$

The proof is postponed to section 3.6.

Remark 3.4.2. (i) Let ξ be $\mathbb{R}^d$ -valued and $\mathcal{F}_t$ -measurable and assume

$$\mathbb{E} [|\xi|^p] < \infty, \quad p \geq q + 1. \quad (3.30)$$

Then (3.28) holds true. Indeed, by the Hölder inequality, denoting p' the conjugate exponent of p ,

$$\mathbb{E} [|\xi|^q \langle \eta, 1 \rangle] \leq \mathbb{E} [|\eta|^p]^{1/p} \mathbb{E} [|\xi|^{p'q}]^{1/p'} < \infty,$$

since $p'q = pq/(p-1) \leq p$ due to the condition $p \geq q + 1$.

- (ii) The requirement that $X_s \in \mathbb{R}_+^N$ is justified since we continue to identify $\mathbb{R}_+^N$ with the set of nonnegative measures on S . The technical condition $\mathbb{E} [\langle X_s, 1 \rangle] = 1$ plays a role that will be explained below.
- (iii) We note an invariance property for the flow $s \mapsto \mathbb{P}_{(X_s, Z_s)}$ of the joint laws of X_s and Z_s . For $\mu \in \mathcal{P}(\mathbb{R}^N \times \mathbb{R}^d)$, denote by $\mu_1 \in \mathcal{P}(\mathbb{R}^N)$ the first marginal of μ . Define the set

$$\mathcal{D} = \left\{ \mu \in \mathcal{P}(\mathbb{R}^N \times \mathbb{R}^d) : \mu_1(\mathbb{R}_+^N) = 1, \mu_1 \in \mathcal{P}_1(\mathbb{R}^N), \langle \bar{\mu}_1, 1 \rangle := \int_{\mathbb{R}^N} \langle x, 1 \rangle \mu_1(dx) = 1 \right\},$$

where $\bar{\mu}_1$ denotes the mean of $\mu_1 \in \mathcal{P}(\mathbb{R}^N)$. Our assumptions imply that $\mathbb{P}_{(\eta, \xi)}$ belongs to $\mathcal{D}$; Theorem 3.4.1 guarantees that $\mathbb{P}_{(X_s, Z_s)} \in \mathcal{D}$. Thus, for any control process $\alpha \in \mathcal{A}$, $\mathcal{D}$ is left invariant by the flow $s \mapsto \mathbb{P}_{(X_s, Z_s)}$.

It follows from the previous result that the optimal control problem (3.25)-(3.26) is well posed under the growth conditions on f, g detailed in Assumptions (A.1). We can thus define the lifted value function

$$V(t, \eta, \xi) = \sup_{\alpha \in \mathcal{A}} J(t, \eta, \xi, \alpha).$$

We note that we are taking the supremum over $\alpha \in \mathcal{A}$, whereas in the separated problem of the previous section we had $\alpha \in \mathcal{A}^W$. In the latter case the control is only adapted to the observation process W , whereas $\alpha \in \mathcal{A}$ also depends on the auxiliary uniform random variable U . However, we observe that, if the sup in the definition of $V(0, \pi_0, z_0)$ with deterministic initial conditions π_0, z_0 is attained by a control $\alpha \in \mathcal{A}$ which also belongs to $\mathcal{A}^W$, then α is optimal for the problem (3.17)-(3.18) and for the partially observed control problem. In our program of future research, we aim at constructing exactly an optimal control of this form, as explained in Section 3.5; as a consequence, the analysis of this larger class of optimal control problems will also provide an optimal control (and, thus, the optimal value) for our separated control problem.

We end this paragraph with a stability result for the state equation of the separate problem. This is a complement to Theorem 3.4.1: it is not needed for the well-posedness of the problem, but it is usually required to prove continuity properties, which we want to prove in the future in order to adopt the analytic approach. We need the following additional assumptions:

Assumptions (A.3). (i) The function $z \mapsto h(z, \nu, a, i)$ is uniformly continuous, uniformly with respect to $a \in A, \nu \in \mathcal{P}_q(\mathbb{R}^d)$: there exists a function $\omega_0 : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, bounded and continuous in 0, satisfying $\omega_0(0) = 0$ and

$$|h(z, \nu, a, i) - h(z', \nu, a, i)| \leq \omega_0(|z - z'|), \quad (3.31)$$

for every $z, z' \in \mathbb{R}^d, a \in A, \nu \in \mathcal{P}_q(\mathbb{R}^d), i \in S$.

(ii) The function $\nu \mapsto h(z, \nu, a, i)$ is also uniformly continuous, uniformly with respect to $z \in \mathbb{R}^d, a \in A$: there exists a function $\omega_q : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, continuous in 0 and satisfying $\omega_q(0) = 0$ such that

$$|h(z, \nu, a, i) - h(z, \nu', a, i)| \leq \omega_q(W_q(\nu, \nu')), \quad (3.32)$$

for every $\nu, \nu' \in \mathcal{P}_q(\mathbb{R}^d), z \in \mathbb{R}^d, a \in A, i \in S$.

Note that the assumption of boundedness of ω_0 does not involve any loss of generality. Indeed, since h is bounded, we can replace ω_0 by $\omega_0 \wedge \sup |h|$.

Theorem 3.4.3. Suppose that Assumptions (A.1), (A.2), (A.3) hold true. Suppose that η, η' are $\mathbb{R}^N$ -valued and $\mathcal{F}_t$ -measurable random variables satisfying

$$\begin{aligned} \mathbb{E}[|\eta|^p] < \infty, \quad \mathbb{E}[|\eta'|^p] < \infty, \quad \mathbb{P}(\eta \in \mathbb{R}_+^N) = \mathbb{P}(\eta' \in \mathbb{R}_+^N) = 1, \\ \mathbb{E}[\langle \eta, 1 \rangle] = \mathbb{E}[\langle \eta', 1 \rangle] = 1. \end{aligned}$$

Let ξ, ξ' be $\mathbb{R}^d$ -valued and $\mathcal{F}_t$ -measurable satisfying

$$\mathbb{E}[|\xi|^p] < \infty, \quad \mathbb{E}[|\xi'|^p] < \infty. \quad (3.33)$$

For any $\alpha \in \mathcal{A}$, let $(X, Z), (X', Z')$ denote the corresponding solutions to equation (3.25) given by Theorem 3.4.1. Finally, suppose that $p \geq q + 1$ and that Assumption **(A.β)** holds for every probability law $\beta \in \mathcal{P}_p(\mathbb{R}^d)$. Then, there exists a constant C such that

$$\begin{aligned} \mathbb{E} \left[\sup_{s \in [t, T]} |X_s - X'_s|^p \right] &\leq C e^{C(1 + \mathbb{E}|\eta'|^p)} \left\{ \mathbb{E}[|\eta'|^p] \omega_q \left((\mathbb{E}|\eta'|^p)^{1/pq} (\mathbb{E}|\xi - \xi'|^{qp'})^{1/qp'} \right)^p \right. \\ &\quad \left. + \mathbb{E} \left[\omega_0(|\xi - \xi'|)^p \int_t^T |X'_s|^p ds \right] + \mathbb{E}[|\eta - \eta'|^p] \right\}, \end{aligned} \quad (3.34)$$

where p' denotes the conjugate exponent of p and C only depends on $p, q, d, T, \Lambda, \sup |h|, L$ occurring in the Assumptions.

In particular, denoting by (X, Z) and (X^n, Z^n) the solutions corresponding to data (η, ξ) and (η^n, ξ^n) satisfying the assumptions above, if

$$\mathbb{E}[|\xi - \xi^n|^p] + \mathbb{E}[|\eta - \eta^n|^p] \rightarrow 0, \quad n \rightarrow \infty,$$

then

$$\mathbb{E} \left[\sup_{s \in [t, T]} |X_s - X_s^n|^p \right] + \mathbb{E} \left[\sup_{s \in [t, T]} |Z_s - Z_s^n|^p \right] \rightarrow 0.$$

The proof is postponed to Section 3.6.

3.4.2 Further properties of the separated problem and the value function

It is convenient to rewrite the separated problem (3.25)-(3.26) in a more standard McKean-Vlasov form, by rewriting the probability measure $\mathbb{E} [\langle X_s, 1 \rangle | Z_s = z] \mathbb{P}_{Z_s}(dz)$ (on $\mathbb{R}^d$) as a function of the joint law of X_s and Z_s . To this end, recall the set $\mathcal{D}$ introduced in Remark 3.4.2-(iii) and define a function $\Pi : \mathcal{D} \rightarrow \mathcal{P}(\mathbb{R}^d)$ setting $\nu = \Pi(\mu)$ where

$$\nu(dz) = \int_{\mathbb{R}_+^N} \langle x, 1 \rangle \mu(dx dz).$$

Note that ν is a nonnegative measure due to the condition $\mu_1(\mathbb{R}_+^N) = 1$ and it is a probability since $\nu(\mathbb{R}^d) = \langle \bar{\mu}_1, 1 \rangle = 1$. It is also absolutely continuous with respect to $\mu_2 \in \mathcal{P}(\mathbb{R}^d)$, the second marginal of μ ; indeed, disintegrating μ in the form $\mu(dx dz) = \mu^z(dx) \mu_2(dz)$ for a suitable kernel $\mu^z(dx)$, we have

$$\nu(dz) = \left(\int_{\mathbb{R}_+^N} \langle x, 1 \rangle \mu^z(dx) \right) \mu_2(dz).$$

Now, if $\mu = \mathbb{P}_{(X_s, Z_s)}$, for any bounded measurable $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$,

$$\begin{aligned} \int_{\mathbb{R}^d} \phi(z) \Pi(\mu)(dz) &= \int_{\mathbb{R}^N \times \mathbb{R}^d} \phi(z) \langle x, 1 \rangle \mu(dx dz) = \mathbb{E} [\phi(Z_s) \langle X_s, 1 \rangle] \\ &= \mathbb{E} [\phi(Z_s) \mathbb{E} [\langle X_s, 1 \rangle | Z_s]] = \int_{\mathbb{R}^d} \phi(z) \mathbb{E} [\langle X_s, 1 \rangle | Z_s = z] \mathbb{P}_{Z_s}(dz). \end{aligned}$$

This proves that $\Pi(\mu)(dz) = \mathbb{E} [\langle X_s, 1 \rangle | Z_s = z] \mathbb{P}_{Z_s}(dz)$ and so, with this notation, equation (3.25) becomes

$$\begin{cases} dX_s &= \Lambda^T X_s ds + \text{diag}(X_s) h(Z_s, \Pi(\mathbb{P}_{(X_s, Z_s)}), \alpha_s) dW_s, \\ dZ_s &= dW_s, \\ X_t &= \eta, \\ Z_t &= \xi, \end{cases} \quad (3.35)$$

and the reward functional (3.26) can also be rewritten as

$$J(t, \eta, \xi, \alpha) = \mathbb{E} \left[\int_t^T \langle X_s, f(Z_s, \Pi(\mathbb{P}_{(X_s, Z_s)}), \alpha_s) \rangle ds + \langle X_T, g(Z_T, \Pi(\mathbb{P}_{X_T, Z_T})) \rangle \right]. \quad (3.36)$$

We have now rewritten the separated problem in a more familiar form. However, due to the occurrence of the mapping Π , we can not directly apply known results on optimal control of McKean-Vlasov systems.

Notice that we aim at applying the analytic approach. We thus need two essential technical tools: the flow property and the dynamic programming principle. In what follows, for the processes Z and X we use the notation $Z_s^{t, \xi}$ and $X_s^{t, \eta, \xi, \alpha}$, noting that Z does not depend either on η or on α .

Proposition 3.4.4 (Flow property). For any $0 \leq t \leq s \leq r \leq T$, (η, ξ) $\mathcal{F}_t$ -measurable satisfying the Assumptions of Theorem 3.4.1 and for any $\alpha \in \mathcal{A}$, we have, $\mathbb{P}$ -a.s.,

$$(Z_r^{t, \xi}, X_r^{t, \eta, \xi, \alpha}) = (Z_r^{s, Z_s^{t, \xi}}, X_r^{s, X_s^{t, \eta, \xi, \alpha}, Z_s^{t, \xi}, \alpha}). \quad (3.37)$$

Proposition 3.4.5 (Dynamic programming principle). The lifted value function

$$V(t, \eta, \xi) = \sup_{\alpha \in \mathcal{A}} J(t, \eta, \xi, \alpha),$$

satisfies the following DPP: for $0 \leq t \leq s \leq T$, (η, ξ) $\mathcal{F}_t$ -measurable satisfying the Assumptions of Theorem 3.4.1, we have

$$V(t, \eta, \xi) = \sup_{\alpha \in \mathcal{A}} \left\{ \mathbb{E} \left[\int_t^s \langle X_r, f(Z_r, \Pi(\mathbb{P}_{X_r, Z_r}), \alpha_r) \rangle dr \right] + V(s, X_s, Z_s) \right\}, \quad (3.38)$$

where $Z_r = Z_r^{t, \xi}$ and $X_r = X_r^{t, \eta, \xi, \alpha}$ for $r \in [t, T]$.

The proofs of these results follow the lines of the proofs of classical analogous results: see for instance the proof of Lemma 1.2.9, with slight modifications due to the presence of the conditional expectation, and of Theorem 1.2.10 in Chapter 1.

3.5 Future developments

As explained in the Introduction, the results of the present chapter are still a work in progress. Here we outline our future developments, which are likely to be completed.

1. We plan to check some continuity property of the lifted value function V .
2. Using those continuity results, we want to verify that the law invariance property holds, i.e. if (ξ, η) , (ξ', η') are two admissible initial conditions and $\mathbb{P}_{\xi, \eta} = \mathbb{P}_{\xi', \eta'}$, then $V(t, \eta, \xi) = V(t, \eta', \xi')$. As usual, this allows to define the value function

$$v(t, \mu) = V(t, \eta, \xi), \quad \text{if } \mu = \mathbb{P}_{(\eta, \xi)}.$$

3. We then want to rewrite the dynamic programming principle for the value function $v(t, \mu)$, in order to derive a parabolic Hamilton-Jacobi-Bellman equation on a suitable Wasserstein space. Many of the needed technical results require some variations from known cases. We also plan to derive a viscosity solution property for the value function introduced above, after an appropriate formulation.
4. We aim at proving a verification theorem based on this equation. In particular, we want to consider the case where the initial condition is given by a pair of deterministic objects. In this case, any feedback control realizing the supremum in the HJB equation will only be adapted to the filtration $\mathbb{F}^W$ and will thus belong to $\mathcal{A}^W$. As a consequence, it will also be optimal for our original separated control problem.

3.6 Proofs and examples

3.6.1 Proof of Theorem 3.4.1

Proof of Theorem 3.4.1. Let Z be given by (3.27), namely $Z_s = \xi + W_s - W_t$. Let us consider the space $\mathcal{S}^p$ of $\mathbb{F}$ -adapted processes X with continuous trajectories satisfying

$$\|X\| := \left(\mathbb{E} \left[\sup_{s \in [t, T]} |X_s|^p \right] \right)^{1/p} < \infty,$$

identified up to indistinguishability. This is a Banach space with the indicated norm. Call $\mathcal{S}_0^p$ its closed subset consisting of processes such that $\langle X, 1 \rangle$ is a martingale and

$$\mathbb{P}(X_s \in \mathbb{R}_+^N) = 1, \quad \mathbb{E}[\langle X_s, 1 \rangle] = 1, \quad s \in [t, T]. \quad (3.39)$$

We need to prove that there exists a unique $X \in \mathcal{S}_0^p$ satisfying

$$dX_s = \Lambda^T X_s ds + \text{diag}(X_s) h \left(Z_s, \mathbb{E}[\langle X_s, 1 \rangle | Z_s = z] \mathbb{P}_{Z_s}(dz), \alpha_s \right) dW_s, \quad X_t = \eta. \quad (3.40)$$

We firstly verify that the measure $\mathbb{E}[\langle X_s, 1 \rangle | Z_s = z] \mathbb{P}_{Z_s}(dz)$ belongs to $\mathcal{P}_q(\mathbb{R}^d)$ whenever $X \in \mathcal{S}_0^p$ with $X_t = \eta$. Indeed, it is a probability measure by (3.39); moreover we

have

$$\begin{aligned} \int_{\mathbb{R}^d} |z|^q \mathbb{E}[\langle X_s, 1 \rangle | Z_s = z] \mathbb{P}_{Z_s}(dz) &= \mathbb{E}[|Z_s|^q \mathbb{E}[\langle X_s, 1 \rangle | Z_s]] = \mathbb{E}[|Z_s|^q \langle X_s, 1 \rangle] \\ &\leq C_q \mathbb{E}[|\xi|^q \langle X_s, 1 \rangle] + C_q \mathbb{E}[|W_s - W_t|^q \langle X_s, 1 \rangle] \end{aligned}$$

for some constant C_q depending on q . The second summand is finite by Hölder inequality, since $\mathbb{E}[|X_s|^p] < \infty$ and $\mathbb{E}[|W_s - W_t|^r] < \infty$ for all $r \geq 1$. For the first summand we note that ξ is $\mathcal{F}_t$ -measurable and $\langle X_s, 1 \rangle$ is a positive martingale, so for every $s \geq t$

$$\mathbb{E}[|\xi|^q \langle X_s, 1 \rangle] = \mathbb{E}[\mathbb{E}[|\xi|^q \langle X_s, 1 \rangle | \mathcal{F}_t]] = \mathbb{E}[|\xi|^q \langle X_t, 1 \rangle] = \mathbb{E}[|\xi|^q \langle \eta, 1 \rangle]$$

which is finite by our assumptions.

For $X \in \mathcal{S}_0^p$ let us define the process $Y = \Phi(X)$ as the solution to the linear equation

$$Y_s = \eta + \int_t^s \Lambda^T Y_r dr + \int_t^s \text{diag}(Y_r) h\left(Z_r, \mathbb{E}[\langle X_r, 1 \rangle | Z_r = z] \mathbb{P}_{Z_r}(dz), \alpha_r\right) dW_r, \quad s \in [t, T]. \quad (3.41)$$

Recalling that h is bounded, it is easily verified by standard theory of solutions to SDEs that there exists a unique solution $Y \in \mathcal{S}^p$. Noting that $\langle \Lambda^T Y_r, 1 \rangle = \langle Y_r, \Lambda 1 \rangle = 0$, since $\Lambda 1 = 0$ because Λ is a transition matrix of a Markov chain, it can be checked that $\langle Y, 1 \rangle$ is a martingale with expectation equal to 1. We now state the following Lemma, whose proof is postponed:

Lemma 3.6.1. $\mathbb{P}(Y_s \in \mathbb{R}_+^N) = 1$.

By Lemma 3.6.1, we conclude that $Y \in \mathcal{S}_0^p$, so that Φ maps this space into itself. Moreover, X is a solution to (3.40) if and only if it is a fixed point of $\Phi : \mathcal{S}_0^p \rightarrow \mathcal{S}_0^p$. We also have the following result, whose proof is again postponed:

Lemma 3.6.2. There exists a constant C_0 , depending only on $p, T, \Lambda, \sup |h|$, such that

$$\sup_{s \in [t, T]} \mathbb{E}[|Y_s|^p] \leq C_0 \mathbb{E}[|\eta|^p]. \quad (3.42)$$

It is thus enough to show that Φ has a unique fixed point. Indeed, in this case the estimate (3.29) follows from the previous lemma.

We wish to prove that a suitable iteration Φ^n of the map Φ is a contraction for the norm $\|\cdot\|$. To this end, take $X, X' \in \mathcal{S}_0^p$, define $Y = \Phi(X)$, $Y' = \Phi(X')$, $\bar{X} = X - X'$, $\bar{Y} = Y - Y'$. Then $\bar{Y}_t = 0$ and subtracting the equations which define Y and Y' we have

$$\begin{aligned} d\bar{Y}_s &= \Lambda^T \bar{Y}_s ds + \text{diag}(\bar{Y}_s) h\left(Z_s, \mathbb{E}[\langle X_s, 1 \rangle | Z_s = z] \mathbb{P}_{Z_s}(dz), \alpha_s\right) dW_s \\ &\quad + \text{diag}(Y'_s) \left[h\left(Z_s, \mathbb{E}[\langle X_s, 1 \rangle | Z_s = z] \mathbb{P}_{Z_s}(dz), \alpha_s\right) \right. \\ &\quad \left. - h\left(Z_s, \mathbb{E}[\langle X_s, 1 \rangle | Z_s = z] \mathbb{P}_{Z_s}(dz), \alpha_s\right) \right] \end{aligned}$$

$$-h\left(Z_s, \mathbb{E}[\langle X'_s, 1 \rangle | Z_s = z] \mathbb{P}_{Z_s}(dz), \alpha_s\right) \Big] dW_s$$

By standard estimates, relying on the Burkholder-Davis-Gundy and the Hölder inequalities and the boundedness of h , we obtain

$$\begin{aligned} \mathbb{E} \sup_{u \in [t, s]} |\bar{Y}_u|^p &\leq C \int_t^s \mathbb{E} |\bar{Y}_r|^p dr \\ &\quad + C \mathbb{E} \left(\int_t^s |Y'_r|^2 \left| h\left(Z_r, \mathbb{E}[\langle X_r, 1 \rangle | Z_r = z] \mathbb{P}_{Z_r}(dz), \alpha_r\right) \right. \right. \\ &\quad \left. \left. - h\left(Z_r, \mathbb{E}[\langle X'_r, 1 \rangle | Z_r = z] \mathbb{P}_{Z_r}(dz), \alpha_r\right) \right|^2 dr \right)^{p/2}, \end{aligned} \quad (3.43)$$

where we denote by C any constant that only depends on $p, T, \Lambda, \sup |h|$, possibly changing from line to line. Recall that $\beta = \mathbb{P}_\xi$ and that ξ is $\mathcal{F}_t$ -measurable, therefore independent of $W_s - W_t$. Since $Z_s = \xi + W_s - W_t$, we have $\mathbb{P}_{Z_s} = \gamma_{s-t}^0 * \beta$. We can apply Assumption **(A.β)** with the densities

$$\phi_1(z) = \mathbb{E}[\langle X_r, 1 \rangle | Z_r = z], \quad \phi_2(z) = \mathbb{E}[\langle X'_r, 1 \rangle | Z_r = z],$$

and we obtain the following inequality:

$$\begin{aligned} &\left| h\left(Z_r, \mathbb{E}[\langle X_r, 1 \rangle | Z_r = z] \mathbb{P}_{Z_r}(dz), \alpha_r\right) - h\left(Z_r, \mathbb{E}[\langle X'_r, 1 \rangle | Z_r = z] \mathbb{P}_{Z_r}(dz), \alpha_r\right) \right|^2 \\ &\leq L^2 \int_{\mathbb{R}^d} |\mathbb{E}[\langle X_r, 1 \rangle | Z_r = z] - \mathbb{E}[\langle X'_r, 1 \rangle | Z_r = z]|^2 \mathbb{P}_{Z_s}(dz). \end{aligned}$$

The latter integral can be rewritten and estimated as follows:

$$\begin{aligned} &\int_{\mathbb{R}^d} |\mathbb{E}[\langle \bar{X}_r, 1 \rangle | Z_r = z]|^2 \mathbb{P}_{Z_s}(dz) \leq \int_{\mathbb{R}^d} \mathbb{E}[|\langle \bar{X}_r, 1 \rangle|^2 | Z_r = z] \mathbb{P}_{Z_s}(dz) \\ &= \mathbb{E}[|\langle \bar{X}_r, 1 \rangle|^2] \leq \mathbb{E}[|\bar{X}_r|^2]. \end{aligned}$$

Replacing in (3.43), by Jensen inequality with $\frac{p}{2} \geq 1$, we have

$$\begin{aligned} \mathbb{E} \sup_{u \in [t, s]} |\bar{Y}_u|^p &\leq C \int_t^s \mathbb{E} |\bar{Y}_r|^p dr + CL^p \mathbb{E} \left(\int_t^s |Y'_r|^2 \mathbb{E}[|\bar{X}_r|^2] dr \right)^{p/2} \\ &\leq C \int_t^s \mathbb{E} |\bar{Y}_r|^p dr + CL^p \int_t^s \mathbb{E}[|Y'_r|^p] \mathbb{E}[|\bar{X}_r|^p] dr. \end{aligned}$$

From Lemma 3.6.2 it follows that $\mathbb{E}[|Y'_r|^p] \leq C_0 \mathbb{E}[|\eta|^p]$ and so

$$\mathbb{E} \sup_{u \in [t, s]} |\bar{Y}_u|^p \leq C \int_t^s \mathbb{E} |\bar{Y}_r|^p dr + CL^p \mathbb{E}[|\eta|^p] \int_t^s \mathbb{E}[|\bar{X}_r|^p] dr.$$

By the Gronwall lemma, there exists a constant K , that also depends on L and $\mathbb{E}[|\eta|^p]$, such that

$$\mathbb{E} \sup_{u \in [t, s]} |\bar{Y}_u|^p \leq K \int_t^s \mathbb{E}[|\bar{X}_r|^p] dr.$$

Define, for $n \geq 0$,

$$\Delta_s^0 = \mathbb{E} \sup_{u \in [t, s]} |X_u - X'_u|^p, \quad \Delta_s^n = \mathbb{E} \sup_{u \in [t, s]} |\Phi^n(X)_u - \Phi^n(X')_u|^p.$$

We have proved that $\Delta_s^1 \leq K \int_t^s \Delta_r^0 dr$. Iterating this inequality in a standard way we obtain

$$\Delta_s^n \leq \frac{s^{n+1} K^n}{n!} \Delta_s^0.$$

Choosing n large enough to guarantee that $T^{n+1} K^n < n!$, we conclude that Φ^n is a contraction and therefore has a unique fixed point. $\square$

Proof of Lemma 3.6.1. We write equation (3.41), which defines Y , in scalar form as follows:

$$dY_s^i = \sum_{j \in S} \lambda_{ji} Y_s^j ds + Y_s^i h_s^i dW_s, \quad s \in [t, T], \quad (3.44)$$

where we set

$$h_s^i := h\left(Z_s, \mathbb{E}[\langle X_s, 1 \rangle \mid Z_s = z] \mathbb{P}_{Z_s}(dz), \alpha_s, i\right) \in \mathbb{R}^d.$$

Next we use the so-called robust resolution of the Zakai equation: we define

$$E_s^i = \exp\left(-\int_t^s h_u^i dW_u\right), \quad N_s^i = Y_s^i E_s^i, \quad s \in [t, T].$$

By the Itô rule one finds $dE_s^i = -E_s^i h_s^i dW_s + \frac{1}{2} E_s^i |h_s^i|^2 ds$ and

$$\begin{aligned} dN_s^i &= \sum_{j \in S} \lambda_{ji} Y_s^j E_s^i ds + Y_s^i E_s^i h_s^i dW_s - Y_s^i E_s^i h_s^i dW_s + \frac{1}{2} Y_s^i E_s^i |h_s^i|^2 ds - Y_s^i E_s^i |h_s^i|^2 ds \\ &= \sum_{j \in S} \lambda_{ji} Y_s^j E_s^i ds - \frac{1}{2} Y_s^i E_s^i |h_s^i|^2 ds \\ &= \sum_{j \in S} \lambda_{ji} N_s^j E_s^i (E_s^j)^{-1} ds - \frac{1}{2} N_s^i |h_s^i|^2 ds. \end{aligned}$$

We conclude that N satisfies an ordinary differential equation with stochastic coefficients of the form $(d/ds)N_s = A(s)N_s$, where the matrix $A(s)$ has entries

$$A_{ij}(s) = \begin{cases} \lambda_{ji} E_s^i (E_s^j)^{-1}, & i \neq j, \\ \lambda_{ii} - \frac{1}{2} |h_s^i|^2, & i = j. \end{cases}$$

Notice that, as $E_s^i > 0$ for all $i = 1, \dots, N$, $\mathbb{P}(Y_s \in \mathbb{R}_+^N) = 1$ if and only if $\mathbb{P}(N_s \in \mathbb{R}_+^N) = 1$. Since $N_0 = Y_0 = \eta \in \mathbb{R}_+^N$, it is enough to show that the set $\mathbb{R}_+^N$ is invariant for the flow of N . To this purpose we apply the Nagumo invariance theorem. The tangent cone $T(x) \in \mathbb{R}^N$ at each point $x \in \mathbb{R}_+^N$ is easily computed as follows: setting

$$J(x) = \{j \in \{1, \dots, N\} : x_j = 0\},$$

we have

$$T(x) = \{v \in \mathbb{R}^N : v_i \geq 0 \text{ for every } i \in J(x)\}.$$

To apply Nagumo's theorem we have to check that $A(s)x \in T(x)$ for every $x \in \mathbb{R}_+^N$. We have

$$(A(s)x)_i = \sum_{j \in S} A_{ij}(s)x_j = \sum_{j \notin J(x)} A_{ij}(s)x_j.$$

If $i \in J(x)$, for any $j \notin J(x)$ it is clear that $i \neq j$; as a consequence, $A_{ij}(s) \geq 0$ for $i \neq j$ and $x_j > 0$ for every $j \notin J(x)$. Therefore, we see that $(A(s)x)_i \geq 0$ whenever $i \in J(x)$, as required. It follows that $\mathbb{R}_+^N$ is invariant with respect to the flow of N and the proof is finished.

A statement of Nagumo's theorem can be found in Theorem 1.2 in [7]. One can also consult Theorem 2.3 in Chapter 0 of [39]. In our case, we need the extension to nonautonomous systems (i.e. with time-dependent coefficients), which can be treated as indicated in Section 3 of Chapter 4 in the same book. Finally, the continuity requirement with respect to t can easily be achieved by mollification of the coefficients $t \mapsto h_t^i$. $\square$

Proof of Lemma 3.6.2. Denoting by C any constant that only depends on p, T, Λ , $\sup |h|$ - not necessarily the same at any occurrence - we have

$$\begin{aligned} \sup_{u \in [t, s]} |Y_u|^p &\leq C |\eta|^p + C \left(\int_t^s |\Lambda^T Y_r| dr \right)^p \\ &\quad + C \sup_{u \in [t, s]} \left| \int_t^u \text{diag}(Y_r) h \left(Z_r, \mathbb{E}[\langle X_r, 1 \rangle | Z_r = z] \mathbb{P}_{Z_r}(dz), \alpha_r \right) dW_r \right|^p, \end{aligned}$$

and, by the Burkholder-Davis-Gundy and the Hölder inequalities and the boundedness of h ,

$$\begin{aligned} \mathbb{E} \sup_{u \in [t, s]} |Y_u|^p &\leq C \mathbb{E} |\eta|^p + C \int_t^s \mathbb{E} |Y_r|^p dr \\ &\quad + C \mathbb{E} \left(\int_t^s \left| \text{diag}(Y_r) h \left(Z_r, \mathbb{E}[\langle X_r, 1 \rangle | Z_r = z] \mathbb{P}_{Z_r}(dz), \alpha_r \right) \right|^2 dr \right)^{p/2} \\ &\leq C \mathbb{E} |\eta|^p + C \int_t^s \mathbb{E} |Y_r|^p dr + C \mathbb{E} \left(\int_t^s |Y_r|^2 dr \right)^{p/2} \\ &\leq C \mathbb{E} |\eta|^p + C \int_t^s \mathbb{E} |Y_r|^p dr. \end{aligned}$$

The conclusion follows from the Gronwall Lemma. $\square$

3.6.2 Proof of Theorem 3.4.3

Proof of Theorem 3.4.3. Let us denote

$$\bar{\eta} = \eta - \eta', \quad \bar{\xi} = \xi - \xi', \quad \bar{X}_s = X_s - X'_s, \quad \bar{Z}_s = Z_s - Z'_s.$$

We note that $Z_s - Z'_s = \bar{\xi}$ for every $s \in [t, T]$. Subtracting the equations for X and X' , we obtain

$$d\bar{X}_s = \Lambda^T \bar{X}_s ds + I_s dW_s, \quad \bar{X}_t = \bar{\eta}. \quad (3.45)$$

where we set

$$\begin{aligned} I_s &= \text{diag}(X_s) h\left(Z_s, \mathbb{E}[\langle X_s, 1 \rangle \mid Z_s = z] \mathbb{P}_{Z_s}(dz), \alpha_s\right) \\ &\quad - \text{diag}(X'_s) h\left(Z'_s, \mathbb{E}[\langle X'_s, 1 \rangle \mid Z'_s = z] \mathbb{P}_{Z'_s}(dz), \alpha_s\right). \end{aligned}$$

By the Burkholder-Davis-Gundy and the Hölder and the Jensen inequalities, we have

$$\begin{aligned} \mathbb{E} \sup_{r \in [t, s]} |\bar{X}_r|^p &\leq C \mathbb{E}[|\bar{\eta}|^p] + C \mathbb{E} \left[\left(\int_t^s |\Lambda^T \bar{X}_r| dr \right)^p \right] + C \mathbb{E} \left[\left(\int_t^s |I_r|^2 dr \right)^{p/2} \right] \\ &\leq C \mathbb{E}[|\bar{\eta}|^p] + C \mathbb{E} \int_t^s |\bar{X}_r|^p dr + C \mathbb{E} \int_t^s |I_r|^p dr. \end{aligned} \quad (3.46)$$

Adding and subtracting terms we have $I_s = I_s^1 + I_s^2 + I_s^3 + I_s^4$, where

$$\begin{aligned} I_s^1 &= \text{diag}(\bar{X}_s) h\left(Z_s, \mathbb{E}[\langle X_s, 1 \rangle \mid Z_s = z] \mathbb{P}_{Z_s}(dz), \alpha_s\right), \\ I_s^2 &= \text{diag}(X'_s) \left\{ h\left(Z_s, \mathbb{E}[\langle X_s, 1 \rangle \mid Z_s = z] \mathbb{P}_{Z_s}(dz), \alpha_s\right) - h\left(Z'_s, \mathbb{E}[\langle X_s, 1 \rangle \mid Z_s = z] \mathbb{P}_{Z_s}(dz), \alpha_s\right) \right\}, \\ I_s^3 &= \text{diag}(X'_s) \left\{ h\left(Z'_s, \mathbb{E}[\langle X_s, 1 \rangle \mid Z_s = z] \mathbb{P}_{Z_s}(dz), \alpha_s\right) - h\left(Z'_s, \mathbb{E}[\langle X'_s, 1 \rangle \mid Z_s = z] \mathbb{P}_{Z_s}(dz), \alpha_s\right) \right\}, \\ I_s^4 &= \text{diag}(X'_s) \left\{ h\left(Z'_s, \mathbb{E}[\langle X'_s, 1 \rangle \mid Z_s = z] \mathbb{P}_{Z_s}(dz), \alpha_s\right) - h\left(Z'_s, \mathbb{E}[\langle X'_s, 1 \rangle \mid Z'_s = z] \mathbb{P}_{Z'_s}(dz), \alpha_s\right) \right\} \end{aligned}$$

The first two terms are estimated as follows, using (3.31) and the fact that h is bounded:

$$|I_s^1| \leq C |\bar{X}_s|, \quad |I_s^2| \leq |X'_s| \omega_0(|Z_s - Z'_s|) = |X'_s| \omega_0(|\bar{\xi}|).$$

Using Assumption **(A.β)**, we have

$$\begin{aligned} |I_s^3|^2 &\leq C |X'_s|^2 \int_{\mathbb{R}^d} \left| \mathbb{E}[\langle X_s, 1 \rangle \mid Z_s = z] - \mathbb{E}[\langle X'_s, 1 \rangle \mid Z_s = z] \right|^2 \mathbb{P}_{Z_s}(dz) \\ &= C |X'_s|^2 \int_{\mathbb{R}^d} \left| \mathbb{E}[\langle \bar{X}_s, 1 \rangle \mid Z_s = z] \right|^2 \mathbb{P}_{Z_s}(dz) \\ &\leq C |X'_s|^2 \int_{\mathbb{R}^d} \mathbb{E}[|\langle \bar{X}_s, 1 \rangle|^2 \mid Z_s = z] \mathbb{P}_{Z_s}(dz) \\ &= C |X'_s|^2 \mathbb{E}[|\langle \bar{X}_s, 1 \rangle|^2] \leq C |X'_s|^2 \mathbb{E}[|\bar{X}_s|^2]. \end{aligned}$$

By (3.32) we also have $|I_s^4| \leq C |X'_s| \omega_q(W_q(\mu, \nu))$, where we set for short

$$\mu(dz) = \mathbb{E}[\langle X'_s, 1 \rangle \mid Z_s = z] \mathbb{P}_{Z_s}(dz), \quad \nu(dz) = \mathbb{E}[\langle X'_s, 1 \rangle \mid Z'_s = z] \mathbb{P}_{Z'_s}(dz).$$

In order to estimate $W_q(\mu, \nu)$ we use the Kantorovich duality (see Theorem 5.10 in [90]): we take $f, g : \mathbb{R}^d \rightarrow \mathbb{R}$ bounded Lipschitz and satisfying

$$f(z_1) + g(z_2) \leq |z_1 - z_2|^q, \quad z_1, z_2 \in \mathbb{R}^d,$$

and we compute

$$\begin{aligned}
& \int_{\mathbb{R}^d} f(z_1) \mu(dz_1) + \int_{\mathbb{R}^d} g(z_2) \nu(dz_2) \\
&= \int_{\mathbb{R}^d} f(z_1) \mathbb{E}[\langle X'_s, 1 \rangle \mid Z_s = z_1] \mathbb{P}_{Z_s}(dz_1) + \int_{\mathbb{R}^d} g(z_2) \mathbb{E}[\langle X'_s, 1 \rangle \mid Z'_s = z_2] \mathbb{P}_{Z'_s}(dz_s) \\
&= \mathbb{E}[f(Z_s) \mathbb{E}[\langle X'_s, 1 \rangle \mid Z_s]] + \mathbb{E}[g(Z'_s) \mathbb{E}[\langle X'_s, 1 \rangle \mid Z'_s]] \\
&= \mathbb{E}[(f(Z_s) + g(Z'_s)) \langle X'_s, 1 \rangle] \\
&\leq \mathbb{E}[|Z_s - Z'_s|^q \langle X'_s, 1 \rangle] = \mathbb{E}[|\bar{\xi}|^q \langle X'_s, 1 \rangle] = \mathbb{E}[|\bar{\xi}|^q \langle \eta', 1 \rangle],
\end{aligned}$$

where the last equality follows from the fact that $\langle X'_s, 1 \rangle$ is a positive martingale. It follows that $W_q(\mu, \nu)^q \leq \mathbb{E}[|\bar{\xi}|^q \langle \eta', 1 \rangle]$ and finally that

$$|I_s^4| \leq C |X'_s| \omega_q \left(\left(\mathbb{E}[|\bar{\xi}|^q \langle \eta', 1 \rangle] \right)^{1/q} \right).$$

Replacing the obtained estimates in (3.46) we have

$$\begin{aligned}
\mathbb{E} \sup_{r \in [t, s]} |\bar{X}_r|^p &\leq C \mathbb{E}[|\bar{\eta}|^p] + C \mathbb{E} \int_t^s |\bar{X}_r|^p dr + C \mathbb{E} \int_t^s |X'_r|^p \omega_0(|\bar{\xi}|)^p dr \\
&\quad + C \mathbb{E} \int_t^s |X'_r|^p \left(\mathbb{E}[|\bar{X}_r|^2] \right)^{p/2} dr + C \mathbb{E} \int_t^s |X'_r|^p \omega_q \left(\left(\mathbb{E}[|\bar{\xi}|^q \langle \eta', 1 \rangle] \right)^{1/q} \right)^p dr \\
&\leq C \mathbb{E}[|\bar{\eta}|^p] + C \int_t^s \mathbb{E}[|\bar{X}_r|^p] dr + C \mathbb{E} \left[\omega_0(|\bar{\xi}|)^p \int_t^s |X'_r|^p dr \right] \\
&\quad + C \int_t^s \mathbb{E}[|X'_r|^p] \mathbb{E}[|\bar{X}_r|^p] dr + C \omega_q \left(\left(\mathbb{E}[|\bar{\xi}|^q \langle \eta', 1 \rangle] \right)^{1/q} \right)^p \int_t^s \mathbb{E}[|X'_r|^p] dr.
\end{aligned}$$

From the a priori estimate $\mathbb{E}[|X'_s|^p] \leq C_0 \mathbb{E}[|\eta'|^p]$ in Theorem 3.4.1 (i.e. inequality (3.29)) we obtain

$$\begin{aligned}
\mathbb{E} \sup_{r \in [t, s]} |\bar{X}_r|^p &\leq C \mathbb{E}[|\bar{\eta}|^p] + C (1 + \mathbb{E}[|\eta'|^p]) \int_t^s \mathbb{E}[|\bar{X}_r|^p] dr + \\
&\quad + C \mathbb{E} \left[\omega_0(|\bar{\xi}|)^p \int_t^s |X'_r|^p dr \right] + C \mathbb{E}[|\eta'|^p] \omega_q \left(\left(\mathbb{E}[|\bar{\xi}|^q \langle \eta', 1 \rangle] \right)^{1/q} \right)^p.
\end{aligned}$$

By the Gronwall lemma,

$$\begin{aligned}
\mathbb{E} \sup_{s \in [t, T]} |\bar{X}_s|^p &\leq C e^{C(1 + \mathbb{E}[|\eta'|^p])} \left\{ \mathbb{E}[|\bar{\eta}|^p] + \mathbb{E} \left[\omega_0(|\bar{\xi}|)^p \int_t^T |X'_s|^p ds \right] \right. \\
&\quad \left. + \mathbb{E}[|\eta'|^p] \omega_q \left(\left(\mathbb{E}[|\bar{\xi}|^q \langle \eta', 1 \rangle] \right)^{1/q} \right)^p \right\}.
\end{aligned}$$

The condition $p \geq q + 1$ implies that $p'q = pq/(p-1) \leq p$ and so, by the Hölder inequality, $\mathbb{E}[|\bar{\xi}|^q \langle \eta', 1 \rangle] \leq \mathbb{E}[|\eta'|^p]^{1/p} \mathbb{E}[|\bar{\xi}|^{p'q}]^{1/p'}$ and (3.34) is proved.

To prove the final convergence result we first note that, since $Z_s - Z'_s = \xi - \xi'$, we clearly have $\mathbb{E}[\sup_{s \in [t, T]} |Z_s - Z'_s|^p] = \mathbb{E}[|\xi - \xi'|^p] \rightarrow 0$. Next we apply (3.34) to the

pairs (X^n, Z^n) and (X, Z) and we obtain

$$\mathbb{E} \left[\sup_{s \in [t, T]} |X_s^n - X_s|^p \right] \leq C e^{C(1+\mathbb{E}|\eta|^p)} \left\{ \mathbb{E}[|\eta|^p] \omega_q \left((\mathbb{E}|\eta|^p)^{1/pq} (\mathbb{E}|\xi^n - \xi|^{qp'})^{1/qp'} \right)^p \right. \\ \left. + \mathbb{E} \left[\omega_0(|\xi^n - \xi|)^p \int_t^T |X_s|^p ds \right] + \mathbb{E}[|\eta^n - \eta|^p] \right\}.$$

Since $qp' \leq p$ we have $\mathbb{E}|\xi^n - \xi|^{qp'} \rightarrow 0$. It remains to prove that the second term in the last displayed inequality vanishes. Suppose on the contrary that for some $\delta > 0$ and some subsequence $\{n_k\}$ we had

$$\mathbb{E} \left[\omega_0(|\xi^{n_k} - \xi|)^p \int_t^T |X_s|^p ds \right] \geq \delta.$$

Then we may extract a subsubsequence $\{n_{k_j}\}$ such that $\xi^{n_{k_j}} \rightarrow \xi$ a.s. for $j \rightarrow \infty$. Since ω_0 is continuous at zero and bounded, and since $\mathbb{E} \int_t^T |X_s|^p ds < \infty$, we get a contradiction from the dominated convergence theorem. $\square$

3.6.3 Examples

In this section, we provide several examples of classes of functions satisfying Assumption $(\mathbf{A}, \mathbf{z}_0)$, or more generally $(\mathbf{A}, \boldsymbol{\beta})$, thus showing that this hypothesis is not too restrictive.

We consider cylindrical functions of the form

$$h(z, \nu, a, i) = F \left(z, \int_{\mathbb{R}^d} \psi_1(x) \nu(dx), \dots, \int_{\mathbb{R}^d} \psi_k(x) \nu(dx), a, i \right), \quad (3.47)$$

for $z \in \mathbb{R}^d$, $\nu \in \mathcal{P}_q(\mathbb{R}^d)$, $a \in A$, $i \in S$, where the functions $\psi_i : \mathbb{R}^d \rightarrow \mathbb{R}$ satisfy a polynomial growth condition of order $r \geq 0$, and $F : \mathbb{R}^d \times \mathbb{R}^k \times A \times S \rightarrow \mathbb{R}^d$ is bounded and Lipschitz on $\mathbb{R}^k$ uniformly with respect to z, a, i . It is obvious that h is well defined (i.e., its definition makes sense for every $\nu \in \mathcal{P}_q(\mathbb{R}^d)$) if $q \geq r$. Let us verify $(\mathbf{A}, \boldsymbol{\beta})$ for some $\beta \in \mathcal{P}_q(\mathbb{R}^d)$. Let us take $\nu_1, \nu_2 \in \mathcal{P}_q(\mathbb{R}^d)$ which are absolutely continuous with respect to $\gamma_s^0 * \beta$ with densities ϕ_1, ϕ_2 respectively. By the Lipschitz character of F ,

$$|h(z, \nu_1, a, i) - h(z, \nu_2, a, i)| \\ \leq C \sum_{i=1}^k \left| \int_{\mathbb{R}^d} \psi_i(x) (\nu_1(dx) - \nu_2(dx)) \right| \\ \leq C \sum_{i=1}^k \int_{\mathbb{R}^d} |\psi_i(x)| |\phi_1(x) - \phi_2(x)| (\gamma_s^0 * \beta)(dx) \\ \leq C \left(\int_{\mathbb{R}^d} |\phi_1(x) - \phi_2(x)|^2 (\gamma_s^0 * \beta)(dx) \right)^{\frac{1}{2}} \sum_{i=1}^k \left(\int_{\mathbb{R}^d} |\psi_i(x)|^2 (\gamma_s^0 * \beta)(dx) \right)^{\frac{1}{2}}.$$

Recalling that γ_s^0 is a Gaussian measure, we see that the latter term is finite provided $q \geq 2r$ and we conclude that **(A.β)** holds for every $\beta \in \mathcal{P}_q(\mathbb{R}^d)$ when $q \geq 2r$.

In particular, this example covers the case

$$h(z, \nu, a, i) = F(z, \bar{\nu}, a, i),$$

where $\bar{\nu} \in \mathbb{R}^d$ denotes the mean of ν , provided that $\beta \in \mathcal{P}_2(\mathbb{R}^d)$. It also holds for functions depending on moments of ν :

$$h(z, \nu, a, i) = F\left(z, \int_{\mathbb{R}^d} x_1^{m_1} \cdots x_d^{m_d} \nu(dx), a, i\right),$$

for a multi-index (m_i) of length m , provided that $\beta \in \mathcal{P}_{2m}(\mathbb{R}^d)$.

Finally, adding the requirement that F is uniformly continuous in z , uniformly with respect to the other arguments, and that the functions ψ_i are Lipschitz, it is easy to see that the function h in (3.47) also satisfies Assumption **(A.3)**.

Several variations of these examples are possible, for instance the case

$$h(z, \nu, a, i) = F\left(z, \int_{\mathbb{R}^d} xx^T \nu(dx), a, i\right),$$

for a suitable function $F : \mathbb{R}^d \times \mathbb{R}^{d \times d} \times A \times S \rightarrow \mathbb{R}^d$; in this case, h satisfies **(A.β)** provided that $q \geq 4$.

Appendix A

Estimates and properties of solutions to McKean-Vlasov SDEs

In this Appendix we provide results which are complementary to the content of Section 1.2 in Chapter 1; as a consequence, we always assume that Assumptions $(\mathbf{H}_g)$ and $(\mathbf{H}_{b,\sigma})$ are satisfied. In Section A.1, we prove that the law of a solution to our McKean-Vlasov dynamics is continuous with respect to time, while in Section A.2 we report the complete proof of estimate (1.4), which is a stability result for the solutions to the state equation. We conclude the Appendix with Section A.3, where we provide an approximation of the solution to the state equation via Picard iterates. Even though these results are standard in classical theory of SDEs, we could not find a reference for McKean-Vlasov SDEs and we thus provided directly the proofs.

A.1 Continuity of the law of the state process

In this Section we prove that the law of the state process is a continuous function. In order to do so, we recall here some basic estimates, presented in Proposition 2.8 and Remark 2.9 in [42] in the finite horizon case. Notice that these estimates are still valid in our setting on bounded intervals because, by construction, the solution $X^{t,\xi,\alpha}$ to equation (1.2) coincides with the solution to equation (1.3) $X^{t,\xi,\alpha;T}$ up to time T .

Proposition A.1.1. Assume that the assumptions $(\mathbf{H}_{b,\sigma})$ are satisfied. Fix a finite time horizon $T > 0$. The following estimates hold:

1. There exists a constant $C(T) > 0$ such that, for any initial time $r \in [0, T]$, any two initial conditions, $\xi, \eta \in L^2(\Omega, \mathcal{F}_r, \mathbb{P})$, and any feasible control $\alpha \in \mathcal{A}$,

$$\mathbb{E} \left[|X_s^{r,\xi,\alpha} - X_s^{r,\eta,\alpha}|^2 \right] \leq 3C(T) \mathbb{E} \left[|\xi - \eta|^2 \right] \quad (\text{A.1})$$

for all $s \in [r, T]$.

2. There exists a constant $K(T) > 0$ such that, for any initial time $t \geq 0$, any

initial condition $\xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$, and any control $\alpha \in \mathcal{A}$,

$$\mathbb{E} \left[|X_s^{t,\xi,\alpha} - \xi|^2 \right] \leq K(T)(s-t)[1 + \|\xi\|_{L^2}^2] \quad (\text{A.2})$$

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

3. There exists a constant $K_0(T)$ such that, for any $0 \leq t < u \leq T$, $\xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$, and $\alpha \in \mathcal{A}$,

$$\mathbb{E} \left[|X_s^{t,\xi,\alpha} - X_s^{u,\xi,\alpha}|^2 \right] \leq K_0(T)(u-t)(1 + \|\xi\|_{L^2}^2) \quad (\text{A.3})$$

for all $s \in [u, T]$.

Proposition A.1.2 (Continuity of the law of the state process). Assume that assumptions $(\mathbf{H}_{b,\sigma})$ are satisfied. Fix $\alpha \in \mathcal{A}$ and any $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$. Let $X = X^{0,\xi,\alpha}$ be the unique solution to the state equation (1.2) under the control $\alpha \in \mathcal{A}$ with initial condition ξ at time 0. Then, the function

$$s \rightarrow \mathbb{P}_{X_s}$$

is continuous.

Proof. We want to show that: for any $t \geq 0$ and for any $\epsilon > 0$ there exists a $\delta = \delta(\epsilon, t, \xi)$ such that

$$W_2(\mathbb{P}_{X_t}, \mathbb{P}_{X_s}) < \epsilon$$

provided that $s \in (t - \delta, t + \delta)$.

Fix $\epsilon > 0$ and $t \geq 0$. Notice that, for any $t, s \geq 0$,

$$W_2^2(\mathbb{P}_{X_t}, \mathbb{P}_{X_s}) \leq \mathbb{E} \left[|X_s - X_t|^2 \right]. \quad (\text{A.4})$$

It is thus sufficient to estimate $\|X_s - X_t\|_{L^2}$. First assume $s \leq t$. Then,

$$\begin{aligned} & \mathbb{E} \left[|X_s - X_t|^2 \right] \quad (\text{by the flow property (1.8)}) \\ & \leq \mathbb{E} \left[|X_s - X_t^{s,X_s,\alpha}|^2 \right] \quad (\text{by (A.2) on } [0, t]) \\ & \leq K(t)(t-s)[1 + \|X_s\|_{L^2}^2] \quad (\text{by the analogue of estimate (1.4) on } [0, T]) \\ & \leq K(t)(t-s) \left[1 + \left(\|\xi\|_{L^2}^2 + 6M^2s \right) e^{(12M^2+1)s} \right] \quad (s \leq t) \\ & \leq K(t) \left[1 + \left(\|\xi\|_{L^2}^2 + 6M^2t \right) e^{(12M^2+1)t} \right] (t-s) \\ & = \tilde{K}(t, \|\xi\|_{L^2})(t-s) < \epsilon^2 \end{aligned}$$

provided that $t-s < \frac{\epsilon^2}{\tilde{K}(t, \|\xi\|_{L^2})}$, where $\tilde{K}(t, \|\xi\|_{L^2}) := K(t) \left[1 + \left(\|\xi\|_{L^2}^2 + 6M^2t \right) e^{(12M^2+1)t} \right] > 0$.

Now assume that $s \geq t$. As we want to have s “close enough” to t , we can assume without loss of generality that $s \leq t + 1$. Then, again by the flow property,

$$\begin{aligned}\mathbb{E}\left[|X_s - X_t|^2\right] &= \mathbb{E}\left[|X_s^{t, X_t, \alpha} - X_t|^2\right] \quad (\text{by (A.2) on } [0, t + 1]) \\ &\leq K(t + 1)(s - t) \left[1 + \|X_t\|_{L^2}^2\right] \quad (\text{by (1.4)}) \\ &\leq K(t + 1)(s - t) \left[1 + \left(\|\xi\|_{L^2}^2 + 6M^2t\right)e^{(12M^2+1)t}\right] \\ &= \hat{K}(t, \|\xi\|_{L^2}^2)(s - t) < \epsilon^2\end{aligned}$$

provided that $s - t \leq \min\left(1, \frac{\epsilon^2}{\hat{K}(t, \|\xi\|_{L^2}^2)}\right)$ with $\hat{K}(t, \|\xi\|_{L^2}^2) > 0$. Substituting and taking the square root in both sides in (A.4), we conclude that there exists a $\delta = \min\{1, \frac{\epsilon^2}{\hat{K}(t)}, \frac{\epsilon^2}{\tilde{K}(t)}\}$ (where both constants have an explicit form) such that

$$W_2(\mathbb{P}_{X_t}, \mathbb{P}_{X_s}) < \epsilon$$

for $s \in (t - \delta, t + \delta)$. □

Remark A.1.3. The same result holds if X is the solution to a McKean-Vlasov SDE starting from ξ at any time $r > 0$; the proof is the same, but the constants $\tilde{K}$ and $\hat{K}$ will also depend on the fixed initial time r .

A.2 Proof of estimate (1.4)

In this section, we rigorously derive estimate (1.4). We want to study the dependence of the solution to (1.2) on the initial state; we especially look for an upper bound for $\mathbb{E}[|X_r^{t, \xi, \alpha}|^2]$ by means of the L^2 -norm of the initial condition ξ . Notice that estimate (2.5) of Proposition 2.8 in [42] is not sufficient in our setting because the constant C of that inequality depends on the finite time horizon T .

The following computations will rely on the classical result of Gronwall Lemma, that we recall here in the formulation of Theorem 1 in [74].

Lemma A.2.1 (Gronwall Lemma). Let $g, h, k : [0, +\infty) \rightarrow \mathbb{R}$ be nonnegative continuous functions and $n : [0, +\infty) \rightarrow \mathbb{R}$ be a positive, monotonic, nondecreasing and continuous function for which the following inequality holds:

$$g(t) \leq n(t) + k(t) \left(\int_0^t h(s)g(s)ds \right) \quad \forall t \geq 0.$$

Then,

$$g(t) \leq n(t) \left[1 + k(t) \int_0^t h(s) \exp \left(\int_s^t k(\tau)h(\tau)d\tau \right) ds \right] \quad \forall t \geq 0.$$

Notice that

$$\mathbb{E}[|X_r^{t,\xi,\alpha}|^2] = \int_{\mathbb{R}^d} |x|^2 \mathbb{P}_{X_r^{t,\xi,\alpha}}(dx) = \|\mathbb{P}_{X_r^{t,\xi,\alpha}}\|_2^2. \quad (\text{A.5})$$

We can thus compute the differential of this function applying the Itô formula for functions of probability measures (see Theorem 4.16 and Remark 4.17 in [42], where we underline that the boundedness of ϕ is not exploited in the proof) to $u(\mathbb{P}_{X_s^{t,\xi,\alpha}})$ for $u(\mu) = \|\mu\|_2^2$. We obtain that, for every $r \geq t$,

$$\begin{aligned} \mathbb{E}[|X_r^{t,\xi,\alpha}|^2] &= \mathbb{E}[|\xi|^2] + \int_t^r \mathbb{E}[\langle b(X_s^{t,\xi,\alpha}, \mathbb{P}_{X_s^{t,\xi,\alpha}}, \alpha_s), 2X_s^{t,\xi,\alpha} \rangle] ds \\ &\quad + \frac{1}{2} \int_t^r \mathbb{E}[\text{tr}(2\sigma(X_s^{t,\xi,\alpha}, \mathbb{P}_{X_s^{t,\xi,\alpha}}, \alpha_s)\sigma^T(X_s^{t,\xi,\alpha}, \mathbb{P}_{X_s^{t,\xi,\alpha}}, \alpha_s))] ds \end{aligned}$$

Notice that $2 \langle z, w \rangle \leq |z|^2 + |w|^2$ for any $z, w \in \mathbb{R}^d$ and $\text{tr}(MM^T) = \|M\|^2$ by definition of Frobenius norm. Thus, for every $r \geq t$,

$$\begin{aligned} \mathbb{E}[|X_r^{t,\xi,\alpha}|^2] &\leq \|\xi\|_{L^2}^2 + \int_t^r \mathbb{E}[|b(X_s^{t,\xi,\alpha}, \mathbb{P}_{X_s^{t,\xi,\alpha}}, \alpha_s)|^2 + |X_s^{t,\xi,\alpha}|^2] ds \\ &\quad + \int_t^r \mathbb{E}[|\sigma(X_s^{t,\xi,\alpha}, \mathbb{P}_{X_s^{t,\xi,\alpha}}, \alpha_s)|^2] ds \quad (\text{by (H3) in } (\mathbf{H}_{b,\sigma})) \\ &\leq \|\xi\|_{L^2}^2 + \int_t^r \mathbb{E}[M^2(1 + |X_s^{t,\xi,\alpha}| + \|\mathbb{P}_{X_s^{t,\xi,\alpha}}\|_2)^2 + |X_s^{t,\xi,\alpha}|^2] ds \\ &\quad + \int_t^r \mathbb{E}[M^2(1 + |X_s^{t,\xi,\alpha}| + \|\mathbb{P}_{X_s^{t,\xi,\alpha}}\|_2)^2] ds \quad (\text{A.6}) \\ &\leq \|\xi\|_{L^2}^2 + 3M^2 \int_t^r \mathbb{E}[1 + |X_s^{t,\xi,\alpha}|^2 + \|\mathbb{P}_{X_s^{t,\xi,\alpha}}\|_2^2] ds \\ &\quad + \int_t^r \mathbb{E}[|X_s^{t,\xi,\alpha}|^2] ds \\ &\quad + 3M^2 \int_t^r \mathbb{E}[1 + |X_s^{t,\xi,\alpha}|^2 + \|\mathbb{P}_{X_s^{t,\xi,\alpha}}\|_2^2] ds \quad (\text{by equation (A.5)}) \\ &= \|\xi\|_{L^2}^2 + 6M^2 \int_t^r \mathbb{E}[1 + |X_s^{t,\xi,\alpha}|^2 + \mathbb{E}[|X_s^{t,\xi,\alpha}|^2]] ds \\ &\quad + \int_t^r \mathbb{E}[|X_s^{t,\xi,\alpha}|^2] ds \\ &= \|\xi\|_{L^2}^2 + 6M^2(r-t) + (12M^2 + 1) \int_t^r \mathbb{E}[|X_s^{t,\xi,\alpha}|^2] ds; \end{aligned}$$

for inequality (A.6) we used, once more, the fact that $2 \langle z, w \rangle \leq |z|^2 + |w|^2$ in $\mathbb{R}^d$ for any dimension d : as a matter of fact, for any $z, w, y \in \mathbb{R}$

$$\begin{aligned} (z + w + y)^2 &= z^2 + w^2 + y^2 + 2zw + 2zy + 2wy \\ &\leq z^2 + w^2 + y^2 + z^2 + w^2 + z^2 + y^2 + w^2 + y^2 \\ &= 3z^2 + 3w^2 + 3y^2. \end{aligned}$$

If we call $g(r) = \mathbb{E}[|X_r^{t,\xi,\alpha}|^2]$, which is continuous and nonnegative, we can rewrite the previous inequality as

$$g(r) \leq \|\xi\|_{L^2}^2 + 6M^2(r-t) + (12M^2 + 1) \int_t^r g(s) ds.$$

Using the notation of Gronwall lemma, $h(s) = 1$ for all $s \geq 0$,

$$k(s) := C = (12M^2 + 1)$$

and

$$n(s) := \|\xi\|_{L^2}^2 + 6M^2(s - t).$$

The functions g , h and k are nonnegative and continuous, while n is continuous, positive and monotone increasing. Applying Gronwall lemma, we obtain

$$\begin{aligned} \mathbb{E}[|X_r^{t,\xi,\alpha}|^2] &\leq \left(\|\xi\|_{L^2}^2 + 6M^2(r - t) \right) \left[1 + (12M^2 + 1) \int_t^r \exp \left(\int_s^r (12M^2 + 1) d\tau \right) ds \right] \\ &= \left(\|\xi\|_{L^2}^2 + 6M^2(r - t) \right) \left[1 + (12M^2 + 1) \int_t^r e^{(12M^2+1)(r-s)} ds \right] \\ &\leq \|\xi\|_{L^2}^2 + 6M^2(r - t) \\ &\quad + (12M^2 + 1)(\|\xi\|_{L^2}^2 + 6M^2(r - t)) \int_t^r e^{(12M^2+1)(r-s)} ds \\ &= \|\xi\|_{L^2}^2 + 6M^2(r - t) \\ &\quad + (12M^2 + 1)\|\xi\|_{L^2}^2 \int_t^r e^{(12M^2+1)(r-s)} ds \\ &\quad + (12M^2 + 1)6M^2(r - t) \int_t^r e^{(12M^2+1)(r-s)} ds. \end{aligned} \tag{A.7}$$

Compute explicitly the following integral:

$$\begin{aligned} \int_t^r e^{(12M^2+1)(r-s)} ds &= -\frac{1}{12M^2 + 1} e^{(12M^2+1)(r-s)} \Big|_{s=t}^{s=r} \\ &= -\frac{1}{12M^2 + 1} + \frac{1}{12M^2 + 1} e^{(12M^2+1)(r-t)} \\ &= \frac{1}{12M^2 + 1} \left(e^{(12M^2+1)(r-t)} - 1 \right). \end{aligned}$$

Substituting the explicit value of the integral in (A.7), we obtain the following estimate:

$$\begin{aligned} \mathbb{E}[|X_r^{t,\xi,\alpha}|^2] &\leq \|\xi\|_{L^2}^2 + 6M^2(r - t) \\ &\quad + (12M^2 + 1)\|\xi\|_{L^2}^2 \cdot \frac{1}{12M^2 + 1} \left(e^{(12M^2+1)(r-t)} - 1 \right) \\ &\quad + 6M^2(r - t)(12M^2 + 1) \left[\frac{1}{12M^2 + 1} \left(e^{(12M^2+1)(r-t)} - 1 \right) \right] \\ &= \|\xi\|_{L^2}^2 + 6M^2(r - t) + \|\xi\|_{L^2}^2 \left(e^{(12M^2+1)(r-t)} - 1 \right) \\ &\quad + 6M^2(r - t) \left[e^{(12M^2+1)(r-t)} - 1 \right] \\ &= \|\xi\|_{L^2}^2 + 6M^2(r - t) + \|\xi\|_{L^2}^2 e^{(12M^2+1)(r-t)} - \|\xi\|_{L^2}^2 \\ &\quad + 6M^2(r - t) e^{(12M^2+1)(r-t)} - 6M^2(r - t) \end{aligned}$$

$$\begin{aligned}
&= \|\xi\|_{L^2}^2 e^{(12M^2+1)(r-t)} + 6M^2(r-t)e^{(12M^2+1)(r-t)} \\
&= \left(\|\xi\|_{L^2}^2 + 6M^2(r-t) \right) e^{(12M^2+1)(r-t)},
\end{aligned}$$

which is a positive and monotone increasing function for every $r > t$ and is equal to $\|\xi\|_{L^2}^2$ for $r = t$. We conclude that the following estimate holds:

$$\mathbb{E}[|X_r^{t,\xi,\alpha}|^2] \leq \left(\|\xi\|_{L^2}^2 + 6M^2(r-t) \right) e^{(12M^2+1)(r-t)}.$$

A.3 Picard iterates for McKean-Vlasov SDEs

In this section we will provide an approximation of solutions to McKean-Vlasov SDEs via Picard iterates, as we could not find a reference for this basic extension of the classical approximation technique to solutions to SDEs.

Theorem A.3.1. Suppose that assumptions $(\mathbf{H}_{b,\sigma})$ hold. Call X the unique solution to the McKean-Vlasov SDE (1.2) with initial condition $X(0) = \xi$ and $\mathbb{P}_\xi = \mu$. Construct the sequence of the Picard iterates $\{X_n\}_{n \in \mathbb{N}}$ as follows: call $X_0 = \xi$ and

$$X_{n+1}(t) = \xi + \int_0^t b(X_n(s), \mathbb{P}_{X_n(s)}, \alpha_s) ds + \int_0^t \sigma(X_n(s), \mathbb{P}_{X_n(s)}, \alpha_s) dB_s.$$

Then, the sequence of the Picard iterates converges to X in the following sense:

$$\lim_{n \rightarrow +\infty} \sup_{0 \leq t \leq T} |X_n(t) - X(t)| = 0$$

Proof. This proof follows the lines of the proofs of Theorem 2.1 Chapter IX in [81] and of Theorem 9.2 in [9]. We prove the result on any interval $[0, T]$: if the thesis holds on a bounded interval, it is sufficient to “glue” the Picard iterates on each bounded interval in order to get the processes approximating X on the entire real line.

Fix the control and consider the corresponding solution X to equation (1.2). Firstly, as usual, define the space $S^2([0, T])$ as the space of $\mathbb{F}$ -progressively measurable, continuous processes such that $\mathbb{E} \left[\sup_{s \in [0, T]} |Y_s|^2 \right] < +\infty$, which is a Banach space when endowed with the norm

$$\|Y\|_{S^2} := \left(\mathbb{E} \left[\sup_{s \in [0, T]} |Y_s|^2 \right] \right)^{\frac{1}{2}}.$$

In order to ease the notation, define a map

$$S : S^2([0, T]) \rightarrow S^2([0, T])$$

such that

$$SY_t := \xi + \int_0^t b(Y_s, \mathbb{P}_{Y_s}, \alpha_s) ds + \int_0^t \sigma(Y_s, \mathbb{P}_{Y_s}, \alpha_s) dB_s$$

for every $Y \in S^2([0, T])$. The map is well defined because, for every $Y \in S^2([0, T])$, SY is still progressively measurable and continuous. Moreover, by Jensen's and Burkholder-Davis-Gundy's (BDG) inequalities,

$$\begin{aligned} \mathbb{E} \left[\sup_{s \in [0, T]} |SY_s|^2 \right] &\leq 3 \left(\mathbb{E} \left[|\xi|^2 \right] + \mathbb{E} \left[T \int_0^T |b(Y_u, \mathbb{P}_{Y_u}, \alpha_u)|^2 du \right] \right. \\ &\quad \left. + C_2 \mathbb{E} \left[\int_0^T |\sigma(Y_u, \mathbb{P}_{Y_u}, \alpha_u)|^2 du \right] \right) \quad (\text{by } (\mathbf{H}_{\mathbf{b}, \sigma})) \\ &\leq 3 \left(\mathbb{E} \left[|\xi|^2 \right] + 3M^2(T + C_2) \mathbb{E} \left[\int_0^T (1 + |Y_u|^2 + \|\mathbb{P}_{Y_u}\|_2^2) ds \right] \right) \\ &\leq 3\|\xi\|_{L^2}^2 + 9M^2(T + C_2)T \left(1 + 2\|Y\|_{S^2}^2 \right) < +\infty. \end{aligned}$$

Denote by $\Phi = (\Phi_t)_{t \in [0, T]}$ the flow of squared distances

$$\Phi_t(Y, Z) = \mathbb{E} \left[\sup_{0 \leq u \leq t} |Y_u - Z_u|^2 \right].$$

Compute

$$\begin{aligned} \Phi_t(SY, SZ) &= \mathbb{E} \left[\sup_{0 \leq u \leq t} |SY_u - SZ_u|^2 \right] \\ &= \mathbb{E} \left[\sup_{0 \leq u \leq t} \left| \int_0^u b(Y_s, \mathbb{P}_{Y_s}, \alpha_s) ds + \int_0^u \sigma(Y_s, \mathbb{P}_{Y_s}, \alpha_s) dB_s \right. \right. \\ &\quad \left. \left. - \int_0^u b(Z_s, \mathbb{P}_{Z_s}, \alpha_s) ds - \int_0^u \sigma(Z_s, \mathbb{P}_{Z_s}, \alpha_s) dB_s \right|^2 \right] \quad (\text{as } (a - b)^2 \leq 2a^2 + 2b^2) \\ &\leq 2\mathbb{E} \left[\sup_{0 \leq u \leq t} \left| \int_0^u b(Y_s, \mathbb{P}_{Y_s}, \alpha_s) ds - \int_0^u b(Z_s, \mathbb{P}_{Z_s}, \alpha_s) ds \right|^2 \right] \\ &\quad + 2\mathbb{E} \left[\sup_{0 \leq u \leq t} \left| \int_0^u \sigma(Y_s, \mathbb{P}_{Y_s}, \alpha_s) dB_s - \int_0^u \sigma(Z_s, \mathbb{P}_{Z_s}, \alpha_s) dB_s \right|^2 \right] \\ &\leq 2\mathbb{E} \left[\sup_{0 \leq u \leq t} u \int_0^u \left| b(Y_s, \mathbb{P}_{Y_s}, \alpha_s) - b(Z_s, \mathbb{P}_{Z_s}, \alpha_s) \right|^2 ds \right] \\ &\quad + 8\mathbb{E} \left[\int_0^t \left| \sigma(Y_s, \mathbb{P}_{Y_s}, \alpha_s) - \sigma(Z_s, \mathbb{P}_{Z_s}, \alpha_s) \right|^2 ds \right] \\ &\leq 2t\mathbb{E} \left[\int_0^t \left| b(Y_s, \mathbb{P}_{Y_s}, \alpha_s) - b(Z_s, \mathbb{P}_{Z_s}, \alpha_s) \right|^2 ds \right] \\ &\quad + 8\mathbb{E} \left[\int_0^t \left| \sigma(Y_s, \mathbb{P}_{Y_s}, \alpha_s) - \sigma(Z_s, \mathbb{P}_{Z_s}, \alpha_s) \right|^2 ds \right], \end{aligned}$$

where we used again the Burkholder-Davis-Gundy and Jensen inequalities. Then, using the hypotheses $(\mathbf{H}_{\mathbf{b}, \sigma})$,

$$\Phi_t(SY, SZ) \leq 2t\mathbb{E} \left[\int_0^t \left| b(Y_s, \mathbb{P}_{Y_s}, \alpha_s) - b(Z_s, \mathbb{P}_{Z_s}, \alpha_s) \right|^2 ds \right]$$

$$\begin{aligned}
& + 8\mathbb{E}\left[\int_0^t \left|\sigma(Y_s, \mathbb{P}_{Y_s}, \alpha_s) - \sigma(Z_s, \mathbb{P}_{Z_s}, \alpha_s)\right|^2 ds\right] \\
& \leq 2t\mathbb{E}\left[\int_0^t L^2(|Y_s - Z_s| + W_2(\mathbb{P}_{Y_s}, \mathbb{P}_{Z_s}))^2 ds\right] \\
& + 8\mathbb{E}\left[\int_0^t L^2(|Y_s - Z_s| + W_2(\mathbb{P}_{Y_s}, \mathbb{P}_{Z_s}))^2 ds\right] \\
& \leq 4(t+4)L^2\mathbb{E}\left[\int_0^t |Y_s - Z_s|^2 + W_2(\mathbb{P}_{Y_s}, \mathbb{P}_{Z_s})^2 ds\right] \\
& \leq 4(t+4)L^2\left(\mathbb{E}\left[\int_0^t |Y_s - Z_s|^2 + \mathbb{E}[|Y_s - Z_s|^2]ds\right)\right] \quad (\text{by Fubini's theorem}) \\
& \leq 4(t+4)L^2\left(\int_0^t \mathbb{E}[|Y_s - Z_s|^2]ds + \int_0^t \mathbb{E}[\sup_{u \in [0,s]} |Y_u - Z_u|^2]ds\right) \\
& = 8(t+4)L^2 \int_0^t \mathbb{E}\left[\sup_{u \in [0,s]} |Y_u - Z_u|^2\right]ds \\
& = 8(t+4)L^2 \int_0^t \Phi_s(Y, Z)ds \leq C \int_0^t \Phi_s(Y, Z)ds, \tag{A.8}
\end{aligned}$$

where $C := 8L^2(T+4)$. Notice that with this notation we have $X_0 = \xi$ and

$$X_{n+1} = SX_n \quad \forall n \in \mathbb{N}.$$

Using again the Jensen and the Burkholder-Davis-Gundy inequalities,

$$\begin{aligned}
\Phi_t(X_0, X_1) &= \mathbb{E}\left[\sup_{0 \leq r \leq t} |X_1(r) - \xi|^2\right] \\
&= \mathbb{E}\left[\sup_{0 \leq r \leq t} \left|\int_0^r b(\xi, \mu, \alpha_s)ds + \int_0^r \sigma(\xi, \mu, \alpha_s)dB_s\right|^2\right] \quad (\text{as above}) \\
&\leq 2\mathbb{E}\left[\sup_{0 \leq r \leq t} \left|\int_0^r b(\xi, \mu, \alpha_s)ds\right|^2 + \left|\int_0^r \sigma(\xi, \mu, \alpha_s)dB_s\right|^2\right] \\
&\leq 2t\mathbb{E}\left[\int_0^t |b(\xi, \mu, \alpha_s)|^2 ds\right] \\
&+ 8\mathbb{E}\left[\int_0^t |\sigma(\xi, \mu, \alpha_s)|^2 ds\right] \quad (\text{by hypothesis (H3) in } (\mathbf{H}_{b,\sigma})) \\
&\leq 6tM^2\mathbb{E}\left[\int_0^t \left(1 + |\xi|^2 + \|\mu\|_2^2\right)ds\right] \\
&+ 24M^2\mathbb{E}\left[\int_0^t \left(1 + |\xi|^2 + \|\mu\|_2^2\right)ds\right] \\
&= (6t + 24)M^2t\left(1 + 2\mathbb{E}[|\xi|^2]\right) \\
&\leq (6T + 24)M^2\left(1 + 2\mathbb{E}[|\xi|^2]\right)t \\
&= CD\left(1 + 2\mathbb{E}[|\xi|^2]\right)t < +\infty,
\end{aligned}$$

where we call $D := 3\frac{M^2}{4L^2}$. By (A.8), we deduce that, for any $t \leq T$

$$\begin{aligned}\Phi_t(X_1, X_2) &\leq C \int_0^t \Phi_s(X_0, X_1) ds \quad (\text{as above}) \\ &\leq C \int_0^t CD \left(1 + 2\mathbb{E}[\|\xi\|^2]\right) s ds \\ &\leq C^2 D \left(1 + 2\mathbb{E}[\|\xi\|^2]\right) \frac{t^2}{2} \\ &\leq \tilde{D} C^2 \frac{t^2}{2},\end{aligned}$$

where $\tilde{D} = D(1 + 2\mathbb{E}[\|\xi\|^2])$. Proceeding by induction, assume that

$$\Phi_t(X_{n-1}, X_n) \leq C^n \tilde{D} \frac{t^n}{n!} \quad \text{for any } t \geq 0.$$

Then, for any $t \geq 0$ and any choice of $n \in \mathbb{N}$,

$$\begin{aligned}\Phi_t(X_n, X_{n+1}) &= \Phi_t(SX_{n-1}, SX_n) \quad (\text{by (A.8)}) \\ &\leq C \int_0^t \Phi_s(X_{n-1}, X_n) ds \quad (\text{by the inductive hypothesis}) \\ &\leq C C^n \tilde{D} \int_0^t \frac{s^n}{n!} ds \\ &= \tilde{D} C^{n+1} \frac{t^{n+1}}{(n+1)!}.\end{aligned}$$

By induction, we conclude that the following estimate holds for any $n \in \mathbb{N}$:

$$\Phi_t(X_n, X_{n+1}) \leq \tilde{D} C^{n+1} \frac{t^{n+1}}{(n+1)!}. \quad (\text{A.9})$$

As a consequence, for any $n \in \mathbb{N}$ we can define the event

$$E_n := \left\{ \sup_{0 \leq t \leq T} |X_{n+1}(t) - X_n(t)| > \frac{1}{2^n} \right\}$$

and compute its probability as follows: by the Markov's inequality,

$$\begin{aligned}\mathbb{P}(E_n) &\leq \frac{\mathbb{E} \left[\sup_{0 \leq t \leq T} |X_{n+1}(t) - X_n(t)|^2 \right]}{\left(\frac{1}{2^n}\right)^2} \\ &= 2^{2n} \mathbb{E} \left[\sup_{0 \leq t \leq T} |X_{n+1}(t) - X_n(t)|^2 \right] \\ &\leq 2^{2n} \tilde{D} C^{n+1} \frac{T^{n+1}}{(n+1)!}.\end{aligned}$$

Therefore,

$$\sum_{n=1}^{+\infty} \mathbb{P}(E_n) \leq \sum_{n=1}^{+\infty} 2^{2n} \tilde{D} C^{n+1} \frac{T^{n+1}}{(n+1)!}$$

$$\begin{aligned}
&= \sum_{n=1}^{+\infty} \tilde{D} \frac{1}{4} \frac{(4CT)^{n+1}}{(n+1)!} \\
&= \sum_{n=0}^{+\infty} \frac{\tilde{D}}{4} \frac{(4CT)^n}{n!} \\
&= \frac{\tilde{D}}{4} e^{4CT} < +\infty.
\end{aligned}$$

Now call

$$E = \bigcap_{m=1}^{+\infty} \bigcup_{n \geq m} E_n = \{\omega : \omega \in E_n \text{ for infinitely many } n \in \mathbb{N}\}.$$

By Borel-Cantelli lemma (see e.g. Proposition 1.7 in [9]), we conclude that $\mathbb{P}(E) = 0$, which means that for a.e. $\omega \in \Omega$ there exists an index $m \in \mathbb{N}$ such that $\omega \in E_n^C$ for every $n \geq m$, i.e. for a.e. $\omega \in \Omega$ there exists an index $m = m(\omega) \in \mathbb{N}$ such that

$$\sup_{0 \leq t \leq T} |X_{n+1}(t) - X_n(t)| \leq \frac{1}{2^n} \quad (\text{A.10})$$

for every $n \geq m$. Now observe that

$$X_m(t) = \sum_{k=0}^{m-1} [X_{k+1}(t) - X_k(t)] + \xi;$$

for a.e. $\omega \in \Omega$, the series $\sum_{m=1}^{+\infty} X_m(t)$ converges uniformly in time because, for all fixed $\epsilon > 0$, there exists $m \geq m(\omega)$ such that

$$\begin{aligned}
&\sum_{k=0}^{+\infty} |X_{k+1}(\omega, t) - X_k(\omega, t)| - \sum_{k=0}^m |X_{k+1}(\omega, t) - X_k(\omega, t)| \\
&= \sum_{k=m+1}^{+\infty} |X_{k+1}(\omega, t) - X_k(\omega, t)| \\
&\leq \sum_{k=m+1}^{+\infty} \sup_{t \in [0, T]} |X_{k+1}(\omega, t) - X_k(\omega, t)| \\
&\leq \sum_{k=m+1}^{+\infty} \frac{1}{2^k} < \epsilon
\end{aligned}$$

provided that m is sufficiently large because the last term is the remainder of a convergent (geometric) series. We thus conclude that, for a.e. fixed $\omega \in \Omega$, the limit

$$Y(t) := \lim_{m \rightarrow +\infty} X_m(t) = \xi + \sum_{k=1}^{+\infty} [X_{k+1}(t) - X_k(t)]$$

exists and is finite; in particular, the convergence of X_m to Y is uniform in time.

Now, the process Y is continuous (it is the uniform limit of continuous processes) and it belongs to $M_{loc}^2([0, T])$ because it inherits the progressive measurability of the processes X_n and

$$\int_0^T |Y_s|^2 ds < +\infty$$

by Weierstrass theorem. Now it is sufficient to prove that Y solves equation (1.2): by pathwise uniqueness of solutions to McKean-Vlasov equations, this implies that $X = Y$ in $S^2([0, T])$.

Recall that

$$X_{n+1}(u) = \xi + \int_t^u b(X_n(s), \mathbb{P}_{X_n(s)}, \alpha_s) ds + \int_t^u \sigma(X_n(s), \mathbb{P}_{X_n(s)}, \alpha_s) dB_s. \quad (\text{A.11})$$

The left hand side converges to Y_u uniformly with respect to time. For the right hand side, notice that, by the hypotheses $(\mathbf{H}_{b,\sigma})$

$$\begin{aligned} & \sup_{0 \leq t \leq T} |b(X_n(t), \mathbb{P}_{X_n(t)}, \alpha_t) - b(Y(t), \mathbb{P}_{Y(t)}, \alpha_t)| \\ & \leq L \sup_{0 \leq t \leq T} \left(|X_n(t) - Y(t)| + W_2(\mathbb{P}_{X_n(t)}, \mathbb{P}_{Y(t)}) \right) \end{aligned} \quad (\text{A.12})$$

and

$$\begin{aligned} & \sup_{0 \leq t \leq T} |\sigma(X_n(t), \mathbb{P}_{X_n(t)}, \alpha_t) - \sigma(Y(t), \mathbb{P}_{Y(t)}, \alpha_t)| \\ & \leq L \sup_{0 \leq t \leq T} \left(|X_n(t) - Y(t)| + W_2(\mathbb{P}_{X_n(t)}, \mathbb{P}_{Y(t)}) \right). \end{aligned} \quad (\text{A.13})$$

We already know that

$$\lim_{n \rightarrow +\infty} \sup_{0 \leq t \leq T} |X_n(t) - Y(t)| = 0$$

due to uniform convergence of $\{X_n\}$ with respect to time. Now notice that

$$\begin{aligned} \sup_{0 \leq t \leq T} W_2^2(\mathbb{P}_{X_n(t)}, \mathbb{P}_{X_{n+1}(t)}) & \leq \sup_{0 \leq t \leq T} \mathbb{E} \left[|X_n(t) - X_{n+1}(t)|^2 \right] \\ & \leq \mathbb{E} \left[\sup_{0 \leq t \leq T} |X_n(t) - X_{n+1}(t)|^2 \right] \\ & = \Phi_T(X_n, X_{n+1}) \leq DC^{m+1} \frac{T^{n+1}}{(n+1)!} \end{aligned}$$

by the estimate (A.9). Fix $\epsilon > 0$ and choose $n \leq m \in \mathbb{N}$; then,

$$\begin{aligned} \sup_{0 \leq t \leq T} W_2^2(\mathbb{P}_{X_n(t)}, \mathbb{P}_{X_m(t)}) & \leq \sup_{0 \leq t \leq T} \sum_{k=n}^{m-1} 2^{k-n+1} W_2^2(\mathbb{P}_{X_k(t)}, \mathbb{P}_{X_{k+1}(t)}) \\ & \leq 2^{-n} \sum_{k=n}^{m-1} D \left((2C)^{k+1} \frac{T^{k+1}}{(k+1)!} \right) < \epsilon \end{aligned}$$

for $n \geq N_\epsilon$ for a sufficiently large N_ϵ (because the last term is the remainder of a series converging to e^{2TC}). This means that $\{(\mathbb{P}_{X_n(t)})_{t \in [0, T]}\}_{n \in \mathbb{N}}$ is a Cauchy sequence (uniformly with respect to time) in the complete metric space $\mathcal{P}(\mathcal{C}([0, T]; \mathbb{R}^d))$ endowed with the uniform-in-time 2-Wasserstein metric

$$d(\mu, \nu) := \sup_{t \in [0, T]} W_2(\mu_t, \nu_t) \quad \text{for all } \mu, \nu \in \mathcal{C}([0, T]; \mathbb{R}^d).$$

We thus conclude that there exists a limit flow of probability measures $\mu = (\mu_t)_{t \geq 0}$ such that

$$\lim_{n \rightarrow +\infty} \sup_{0 \leq t \leq T} W_2(\mathbb{P}_{X_n(t)}, \mu_t) = 0$$

for every $t \geq 0$. Now,

1. The convergence in the 2-Wasserstein metric implies weak convergence, i.e.

$$\mathbb{P}_{X_n(t)} \xrightarrow{w} \mu_t$$

for every $t \geq 0$ uniformly with respect to time;

2. The $\mathbb{P}$ -a.s. convergence of $\{X_n(t)\}_{n \in \mathbb{N}}$ to $Y(t)$ implies weak convergence of the laws:

$$\mathbb{P}_{X_n(t)} \xrightarrow{w} \mathbb{P}_{Y(t)}$$

for every $t \geq 0$ uniformly with respect to time.

By uniqueness of the weak limit,

$$\mathbb{P}_{Y(t)} = \mu_t \quad \forall t \in [0, T].$$

Then, we conclude that

$$\lim_{n \rightarrow +\infty} \sup_{0 \leq t \leq T} W_2(\mathbb{P}_{X_n(t)}, \mathbb{P}_{Y(t)}) = \lim_{n \rightarrow +\infty} \sup_{0 \leq t \leq T} W_2(\mathbb{P}_{X_n(t)}, \mu_t) = 0.$$

Substituting in (A.12) and (A.13), we deduce that

$$b(X_n(t), \mathbb{P}_{X_n(t)}, \alpha_t) \xrightarrow{n \rightarrow +\infty} b(Y(t), \mathbb{P}_{Y(t)}, \alpha_t)$$

and

$$\sigma(X_n(t), \mathbb{P}_{X_n(t)}, \alpha_t) \xrightarrow{n \rightarrow +\infty} \sigma(Y(t), \mathbb{P}_{Y(t)}, \alpha_t)$$

uniformly with respect to time, which implies that

$$\int_0^t b(X_n(s), \mathbb{P}_{X_n(s)}, \alpha_s) ds \xrightarrow{n \rightarrow +\infty} \int_0^t b(Y(s), \mathbb{P}_{Y(s)}, \alpha_s) ds \quad \text{a.s.}$$

and

$$\int_0^t |\sigma(X_n(s), \mathbb{P}_{X_n(s)}, \alpha_s) - \sigma(Y(s), \mathbb{P}_{Y(s)}, \alpha_s)|^2 ds \xrightarrow{n \rightarrow +\infty} 0 \quad \text{a.s.}$$

(by the Dominated Convergence Theorem with uniform convergence). As a.s. convergence implies convergence in probability, from the second limit we deduce that

$$\int_0^t \sigma(X_n(s), \mathbb{P}_{X_n(s)}, \alpha_s) dB_s \xrightarrow{n \rightarrow +\infty} \int_0^t \sigma(Y(s), \mathbb{P}_{Y(s)}, \alpha_s) dB_s$$

in probability by Proposition 7.3 in [9]. Now, convergence in probability implies a.s. convergence up to subsequencies. Therefore, taking the limit in equation (A.11), we obtain that Y is a solution to the state equation (1.2) and it thus coincides with X . $\square$

Appendix B

Time invariance property: technical results

In this Appendix, we present the detailed proofs of the technical results needed for the time invariance property, presented in Section 1.3 in Chapter 1. Under the assumptions of that Section, we firstly provide all the technical lemmata which are required to prove Lemma 1.3.2, which guarantees that we can “shift” in time the solutions to a controlled McKean-Vlasov SDE and the associated control in a suitable sense: this is the subject of Section B.1. In Section B.2, we prove the well-posedness of the “disintegration system” of the McKean-Vlasov equation with respect to the first part of the Brownian trajectory, presented in Lemma 1.3.10. Finally, in Section B.3 we present the proof of the continuity property of the gain functional with respect to time (Proposition 1.3.13). Moreover, we also state and prove the continuity of the gain functional with respect to the initial condition (Proposition B.3.2). The precise statements for the theorems proved in this Appendix can be found in Section 1.3; the Appendix should thus be read in parallel with that Section.

B.1 Translation invariance property of solutions to SDEs

In this section, we present a deterministic change of variable formula for the law of stochastic integrals and the solutions to stochastic differential equations.

The first result guarantees that the law of a stochastic integral with respect to a m -dimensional Brownian motion is uniquely determined by the joint law of the integrand, the initial condition and the Brownian motion.

Remark B.1.1. Notice that this result generalizes Exercise 5.16, Chapter IV by [81], which holds in dimension 1. It is expected that a similar result holds for general semimartingales; nonetheless, the proof may need modifications, since the class of feasible integrands with respect to a general m -dimensional semimartingale does not correspond in general to the set of vector-valued processes such that each component is integrable with respect to the corresponding component of the m -dimensional semi-

martingale (see e.g. [76], Section 1.2). A broader extension of this result is beyond the scope of this thesis.

Lemma B.1.2 (The law of the integrand characterizes the law of the integral). Let $(\Omega^1, \mathcal{F}^1, \mathbb{F}^1, \mathbb{P}^1)$ and $(\Omega^2, \mathcal{F}^2, \mathbb{F}^2, \mathbb{P}^2)$ be two filtered probability spaces, on which we define two Brownian motions, denoted by B^1 and B^2 , respectively. Let X^i be a $\mathbb{F}^i$ -predictable process for $i = 1, 2$ such that

$$\int_0^T |X_s^i|^2 ds < +\infty \quad \forall T > 0,$$

i.e. $X^i \in M_{loc}^2([0, +\infty))$ for the filtration $\mathbb{F}^i$, for $i = 1, 2$. Let ξ^i be a random variable belonging to $L^2(\Omega^i, \mathcal{F}_0^i, \mathbb{P}^i)$ for $i = 1, 2$, where we assume that $\mathcal{F}_0^i = \sigma(U^i)$ for U^i uniform random variable on $[0, 1]$ independent of B^i . Denote by I^i the stochastic integral of X^i with respect to B^i summed to the initial condition ξ^i , i.e.

$$I_t^i = \xi^i + \int_0^t X_s^i dB_s^i \quad \forall t \geq 0$$

for $i = 1, 2$. Suppose furthermore that the joint probability distribution of ξ^1, X^1, B^1 and ξ^2, X^2, B^2 are equal, i.e.

$$\mathcal{L}((\xi^1, X^1, B^1)) = \mathcal{L}((\xi^2, X^2, B^2)).$$

Then,

$$\mathcal{L}((\xi^1, X^1, B^1, I^1)) = \mathcal{L}((\xi^2, X^2, B^2, I^2)).$$

Proof. Notice that it is sufficient to prove that the thesis holds for any fixed time horizon T , as

$$M_{loc}^2([0, +\infty)) = \{X : [0, +\infty) \times \Omega \rightarrow \mathbb{R}^d : X \in \bigcap_{T>0} M_{loc}^2([0, T])\}.$$

We thus fix $T > 0$. We proceed by an approximation argument, starting with elementary processes.

Step 1: Assume that X^i is an elementary process for $i = 1, 2$. In order to prove the thesis, we aim at showing that the vectors (ξ^1, X^1, B^1, I^1) and (ξ^2, X^2, B^2, I^2) have the same finite-dimensional distributions.

Without loss of generality, we can suppose that the two elementary processes have the same jumping times. We can thus assume that there exists a finite number k of times $\{s_n\}_{n=1}^k$ such that $s_0 = 0 \leq s_1 \leq \dots \leq s_k = T$ for which we can write

$$\begin{aligned} X^1(s, \omega) &= \mathbb{1}_{\{0\}}(s) X_{-1}^1 + \sum_{n=0}^{k-1} \mathbb{1}_{(s_n, s_{n+1}]}(s) X_n^1(\omega) \\ X^2(s, \omega) &= \mathbb{1}_{\{0\}}(s) X_{-1}^2 + \sum_{n=0}^{k-1} \mathbb{1}_{(s_n, s_{n+1}]}(s) X_n^2(\omega), \end{aligned}$$

where X_{-1}^i is $\mathcal{F}_0^i$ -measurable and X_n^i is a $\mathcal{F}_{s_n}^i$ -measurable random variable for $i = 1, 2$ and for all $n = 0, \dots, k-1$. Recall that the stochastic integral of processes of this form is defined as

$$\int_0^t X^i(s) dB_s^i = \sum_{n=0}^{k-1} X_n^i [B_{s_{n+1} \wedge t}^i - B_{s_n \wedge t}^i]$$

for all $t \geq 0$ for $i = 1, 2$. As a consequence,

$$\begin{aligned} & (I_{s_0}^i, I_{s_1}^i, I_{s_2}^i, I_{s_3}^i, \dots, I_{s_k}^i) \\ &= \left(\xi^i, \xi^i + \int_0^{s_1} X^i(s) dB_s^i, \xi^i + \int_0^{s_2} X^i(s) dB_s^i, \xi^i + \int_0^{s_3} X^i(s) dB_s^i, \dots, \xi^i + \int_0^{s_k} X^i(s) dB_s^i \right) \\ &= \left(\xi^i, \xi^i + X_0^i B_{s_1}^i, \xi^i + X_0^i B_{s_1}^i + X_1^i [B_{s_2}^i - B_{s_1}^i], \right. \\ & \quad \xi^i + X_0^i B_{s_1}^i + X_1^i [B_{s_2}^i - B_{s_1}^i] + X_2^i [B_{s_3}^i - B_{s_2}^i], \\ & \quad \left. \dots, \xi^i + \sum_{n=1}^k X_{n-1}^i [B_{s_n}^i - B_{s_{n-1}}^i] \right) \\ &= \Phi((\xi^i, X_{s_1}^i, X_{s_2}^i, X_{s_3}^i, \dots, X_{s_k}^i, B_{s_1}^i, B_{s_2}^i, B_{s_3}^i, \dots, B_{s_k}^i)) \end{aligned}$$

for $i = 1, 2$, where Φ is a Borel measurable function independent of the index i . Using the hypothesis, which guarantees that

$$\mathcal{L}(\xi^1, X_{s_1}^1, X_{s_2}^1, \dots, X_{s_n}^1, B_{s_1}^1, B_{s_2}^1, \dots, B_{s_n}^1) = \mathcal{L}(\xi^2, X_{s_1}^2, X_{s_2}^2, \dots, X_{s_n}^2, B_{s_1}^2, B_{s_2}^2, \dots, B_{s_n}^2),$$

we deduce that

$$\mathcal{L}(I_{s_0}^1, I_{s_1}^1, I_{s_2}^1, I_{s_3}^1, \dots, I_{s_n}^1) = \mathcal{L}(I_{s_0}^2, I_{s_1}^2, I_{s_2}^2, I_{s_3}^2, \dots, I_{s_n}^2).$$

because an equality in law of random vector implies the equality of measurable functions of the same random vectors. As a matter of fact, if Φ is any measurable function and Z_1, Z_2 are two D -dimensional random variables having the same law, then for any $\psi \in C_b(\mathbb{R}^D; \mathbb{R})$ we can compute

$$\begin{aligned} \int_{\mathbb{R}^D} \psi(y) \mathbb{P}_{\Phi(Z_1)}(dy) &= \mathbb{E}[\psi(\Phi(Z_1))] \\ &= \int_{\mathbb{R}^D} \psi(\Phi)(z) \mathbb{P}_{Z_1}(dz) \\ &= \int_{\mathbb{R}^D} \psi(\Phi)(z) \mathbb{P}_{Z_2}(dz) \\ &= \mathbb{E}[\psi(\Phi(Z_2))] \\ &= \int_{\mathbb{R}^D} \psi(y) \mathbb{P}_{\Phi(Z_2)}(dy), \end{aligned}$$

which implies that

$$\mathbb{P}_{\Phi(Z_1)} = \mathbb{P}_{\Phi(Z_2)}.$$

Fix $m \in \mathbb{N}$ and a sequence of time instants $0 \leq t_1 \leq t_2 \leq \dots t_m$. Once again, by definition of stochastic integral,

$$I_{t_j}^i = \xi^i + \sum_{n=1}^k X_{n-1}^i [B_{s_n \wedge t_j}^i - B_{s_{n-1} \wedge t_j}^i] \text{ for all } j = 1, \dots, m,$$

which implies that

$$(I_{t_1}^i, \dots, I_{t_m}^i) = \Psi((\xi^i, X_{s_1}^i, X_{s_2}^i, \dots, X_{s_k}^i, B_{t_1}^i, \dots, B_{t_m}^i, B_{s_1}^i, B_{s_2}^i, \dots, B_{s_k}^i))$$

for some Borel measurable Ψ and for $i = 1, 2$. As a consequence,

$$\begin{aligned} & \mathcal{L}\left(\xi^1, X_{s_1}^1, X_{s_2}^1, \dots, X_{s_k}^1, B_{s_1}^1, B_{s_2}^1, B_{s_3}^1, \dots, B_{s_k}^1, I_{s_0}^1, I_{s_1}^1, I_{s_2}^1, I_{s_3}^1, \dots, I_{s_k}^1, \right. \\ & \quad \left. X_{t_1}^1, \dots, X_{t_m}^1, B_{t_1}^1, \dots, B_{t_m}^1, I_{t_1}^1, \dots, I_{t_m}^1\right) \\ &= \mathcal{L}\left(\xi^1, X_{s_1}^1, X_{s_2}^1, \dots, X_{s_k}^1, B_{s_1}^1, B_{s_2}^1, B_{s_3}^1, \dots, B_{s_k}^1, \right. \\ & \quad \Phi((\xi^1, X_{s_1}^1, X_{s_2}^1, X_{s_3}^1, \dots, X_{s_k}^1, B_{s_1}^1, B_{s_2}^1, B_{s_3}^1, \dots, B_{s_k}^1)), \\ & \quad X_{t_1}^1, \dots, X_{t_m}^1, B_{t_1}^1, \dots, B_{t_m}^1, \\ & \quad \left. \Psi((\xi^1, X_{s_1}^1, X_{s_2}^1, \dots, X_{s_k}^1, B_{t_1}^1, \dots, B_{t_m}^1, B_{s_1}^1, B_{s_2}^1, \dots, B_{s_k}^1))\right) \\ &= \mathcal{L}\left(\xi^2, X_{s_1}^2, X_{s_2}^2, \dots, X_{s_k}^2, B_{s_1}^2, B_{s_2}^2, B_{s_3}^2, \dots, B_{s_k}^2, \right. \\ & \quad \Phi((\xi^2, X_{s_1}^2, X_{s_2}^2, X_{s_3}^2, \dots, X_{s_k}^2, B_{s_1}^2, B_{s_2}^2, B_{s_3}^2, \dots, B_{s_k}^2)), \\ & \quad X_{t_1}^2, \dots, X_{t_m}^2, B_{t_1}^2, \dots, B_{t_m}^2, \\ & \quad \left. \Psi((\xi^2, X_{s_1}^2, X_{s_2}^2, \dots, X_{s_k}^2, B_{t_1}^2, \dots, B_{t_m}^2, B_{s_1}^2, B_{s_2}^2, \dots, B_{s_k}^2))\right) \\ &= \mathcal{L}\left(\xi^2, X_{s_1}^2, X_{s_2}^2, \dots, X_{s_k}^2, B_{s_1}^2, B_{s_2}^2, B_{s_3}^2, \dots, B_{s_k}^2, I_{s_0}^2, I_{s_1}^2, I_{s_2}^2, I_{s_3}^2, \dots, I_{s_k}^2, \right. \\ & \quad \left. X_{t_1}^2, \dots, X_{t_m}^2, B_{t_1}^2, \dots, B_{t_m}^2, I_{t_1}^2, \dots, I_{t_m}^2\right). \end{aligned}$$

The second to last equality follows from the hypothesis

$$\mathcal{L}(\xi^1, X^1, B^1) = \mathcal{L}(\xi^2, X^2, B^2)$$

as before. Now, the equality in law that we have just proved implies the equality in law of the following finite-dimensional marginals:

$$\begin{aligned} & \mathcal{L}(\xi^1, X_{t_1}^1, \dots, X_{t_m}^1, B_{t_1}^1, \dots, B_{t_m}^1, I_{t_1}^1, \dots, I_{t_m}^1) \\ &= \mathcal{L}(\xi^2, X_{t_1}^2, \dots, X_{t_m}^2, B_{t_1}^2, \dots, B_{t_m}^2, I_{t_1}^2, \dots, I_{t_m}^2). \end{aligned}$$

Recall that $m \in \mathbb{N}$ and $t_1, \dots, t_m$ were selected arbitrarily. As a consequence, all the finite-dimensional distributions of the vector-valued processes (ξ^1, X^1, B^1, I^1) and (ξ^2, X^2, B^2, I^2) are equal, which means that they are equivalent. We conclude that

$$\mathcal{L}(\xi^1, X^1, B^1, I^1) = \mathcal{L}(\xi^2, X^2, B^2, I^2)$$

if X^1, X^2 are elementary processes.

Step 2: Now suppose that X^1, X^2 are in $M^2([0, T])$, i.e. that

$$\mathbb{E} \left[\int_0^T |X_s^i|^2 ds \right] < +\infty$$

for $i = 1, 2$. Construct two sequences of processes $\{X_n^1\}_{n \in \mathbb{N}}$ and $\{X_n^2\}_{n \in \mathbb{N}}$ with a construction inspired by [9]: consider a sequence $\{\Delta^n\}_{n \in \mathbb{N}}$ of equispaced partitions of $[0, T]$ such that $\Delta^n = \{t_j^n\}_{j=0}^{k_n} \subseteq \Delta^{n+1}$ for every $n \in \mathbb{N}$ and $|\Delta^n| = \frac{T}{n}$. Define

$$\begin{aligned} X_n^1(s, \omega) &= \sum_{t_j^n \in \Delta^n \setminus \{t_0^n\}} \left(\frac{1}{t_j^n - t_{j-1}^n} \int_{t_{j-1}^n}^{t_j^n} X^1(r, \omega) dr \right) \mathbb{1}_{(t_j^n, t_{j+1}^n]}(s) \\ X_n^2(s, \omega) &= \sum_{t_j^n \in \Delta^n \setminus \{t_0^n\}} \left(\frac{1}{t_j^n - t_{j-1}^n} \int_{t_{j-1}^n}^{t_j^n} X^2(r, \omega) dr \right) \mathbb{1}_{(t_j^n, t_{j+1}^n]}(s). \end{aligned}$$

Then,

- X_n^i is an elementary process for each $n \in \mathbb{N}$ and $i = 1, 2$.
- By Lebesgue differentiation Theorem, for $i = 1, 2$ and for every $s \geq 0$

$$\lim_{n \rightarrow +\infty} X_n^i(s, \omega) = X^i(s, \omega) \quad \mathbb{P}\text{-a.s.}$$

- For $i = 1, 2$

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left[\int_t^T |X^i(s) - X_n^i(s)|^2 ds \right] = 0 \quad \forall T \geq t$$

by Lemma 7.2 in [9].

- If $\mathcal{L}(\xi^1, B^1, X^1) = \mathcal{L}(\xi^2, B^2, X^2)$, then it is clear that, for any fixed $n \in \mathbb{N}$,

$$\mathcal{L}(\xi^1, B^1, X_n^1) = \mathcal{L}(\xi^2, B^2, X_n^2)$$

because the law of X_n^i only depends on the law of X^i .

By step 1, we know that this implies

$$\mathcal{L}(\xi^1, B^1, X_n^1, I_n^1) = \mathcal{L}(\xi^2, B^2, X_n^2, I_n^2) \quad \text{for all } n \in \mathbb{N}, \quad (\text{B.1})$$

where

$$I_n^i(s) = \int_0^s X_n^i(r) dB_r^i \quad \text{for } i = 1, 2.$$

By Corollary 7.1 in [9], the stochastic integral of X^i with respect to B^i is defined as

$$I^i(s) = L^2 - \lim_{n \rightarrow +\infty} I_n^i(s) \quad \forall s \geq 0.$$

Fix any $m \in \mathbb{N}$ and any $t \leq s_1 \leq \dots \leq s_m \leq T$. Now, as $X_n^i(s_j)$ converges to $X^i(s_j)$ pointwise and $\{I_n^i(s_j)\}$ converges to $\{I^i(s_j)\}$ in the $L^2(\Omega)$ -sense for all $j \in \{1, \dots, m\}$, we can always extract a subsequence (which we will not rename) such that the vector

$$(\xi^i, B^i(s_1), \dots, B^i(s_m), X_n^i(s_1), \dots, X_n^i(s_m), I_n^i(s_1), \dots, I_n^i(s_m))$$

converges in distribution to

$$(\xi^i, B^i(s_1), \dots, B^i(s_m), X^i(s_1), \dots, X^i(s_m), I^i(s_1), \dots, I^i(s_m))$$

for $i = 1, 2$. Weak convergence implies pointwise convergence of the characteristic functions:

$$\begin{aligned} \lim_{n \rightarrow +\infty} \phi_{(\xi^i, B^i(s_1), \dots, B^i(s_m), X_n^i(s_1), \dots, X_n^i(s_m), I_n^i(s_1), \dots, I_n^i(s_m))}(\theta) \\ = \phi_{(\xi^i, B^i(s_1), \dots, B^i(s_m), X^i(s_1), \dots, X^i(s_m), I^i(s_1), \dots, I^i(s_m))}(\theta), \quad i = 1, 2 \end{aligned}$$

for every $\theta \in \mathbb{R}^{3m+1}$, where ϕ denotes the characteristic function. By equation (B.1), as two random vectors have the same law if and only if they have the same characteristic function, we deduce that

$$\begin{aligned} \phi_{(\xi^1, B^1(s_1), \dots, B^1(s_m), X_n^1(s_1), \dots, X_n^1(s_m), I_n^1(s_1), \dots, I_n^1(s_m))}(\theta) = \\ \phi_{(\xi^2, B^2(s_1), \dots, B^2(s_m), X_n^2(s_1), \dots, X_n^2(s_m), I_n^2(s_1), \dots, I_n^2(s_m))}(\theta) \quad \text{for all } \theta \in \mathbb{R}^{3m+1} \text{ and all } n \in \mathbb{N}. \end{aligned}$$

As a consequence, the two sequences must have the same limit:

$$\begin{aligned} \phi_{(\xi^1, B^1(s_1), \dots, B^1(s_m), X^1(s_1), \dots, X^1(s_m), I^1(s_1), \dots, I^1(s_m))}(\theta) \\ = \phi_{(\xi^2, B^2(s_1), \dots, B^2(s_m), X^2(s_1), \dots, X^2(s_m), I^2(s_1), \dots, I^2(s_m))}(\theta) \quad \text{for all } \theta \in \mathbb{R}^{3m+1}, \end{aligned}$$

from which we deduce that

$$\begin{aligned} \mathcal{L}(\xi^1, B^1(s_1), \dots, B^1(s_m), X^1(s_1), \dots, X^1(s_m), I^1(s_1), \dots, I^1(s_m)) \\ = \mathcal{L}(\xi^2, B^2(s_1), \dots, B^2(s_m), X^2(s_1), \dots, X^2(s_m), I^2(s_1), \dots, I^2(s_m)) \end{aligned}$$

because the characteristic function characterizes uniquely the law of random vectors. By the arbitrariness of m and $s_1, \dots, s_m \in [t, T]$, the random vectors

$$(\xi^1, B^1, X^1, I^1) \text{ and } (\xi^2, B^2, X^2, I^2)$$

have the same finite dimensional distributions, i.e. they are equivalent. We conclude that

$$\mathcal{L}(\xi^1, B^1, X^1, I^1) = \mathcal{L}(\xi^2, B^2, X^2, I^2).$$

Step 3: Assume $X \in M_{loc}^2([t, T])$. It is sufficient to repeat the approximation reasoning above with the approximating processes

$$X_n(s) = X(s) \mathbb{1}_{\{\tau_n \leq s\}},$$

with $\tau_n = \inf\{s : \int_0^s |X_u|^2 du > n\}$ for every $n \in \mathbb{N}$ to conclude that the thesis holds in this general case. $\square$

We now prove a deterministic change-of-time formula for stochastic integrals. In order to ease the notation, we will introduce an appropriate family of operators, called *shift operators*, following [81], Chapter 3.

Definition B.1.3. Let $(E, \mathcal{E})$ be any measurable space. Denote by π the coordinate process

$$\pi : [0, +\infty) \times E^{[0, +\infty)} \rightarrow E : \pi_t(w) = w(t).$$

We define the family of shift operators $(\theta_t)_{t \geq 0}$ on the canonical space $E^{[0, +\infty)}$ as follows:

$$\pi_s(\theta_t(w)) := \pi_{t+s}(w) \quad \forall s \geq 0,$$

i.e. the effect of θ_t on a path is to cut off the part of the path before t and to shift the remaining part in time.

Remark B.1.4. As explained in [81], $(\theta_t)_{t \geq 0}$ is a semigroup and θ_t is measurable with respect to $\sigma(\pi_s : s \geq t)$ for every $t \geq 0$. Notice that, if a path w is defined only on the interval $[t, +\infty)$, then $\theta_t(w)(s) = w(s + t)$ for any $s \geq 0$: the introduction of the shift operator makes it possible to define the path from $s = 0$ just by a translation in time. In what follows, we will apply these operators on any trajectory of a stochastic process X and, with a slight abuse of notation, we will denote by $\theta_t(X)$ the stochastic process whose paths are given by the shifted operator θ_t applied to the corresponding paths of the process X . If we consider a trajectory of the Brownian motion $B^{(t)}$, we deduce that

$$\begin{aligned} B^{(t)}(s)(\omega) &= [B_s - B_t](\omega) \quad \text{for } s \geq t \\ \implies \theta_t(B^{(t)})(\omega)(s) &= (B_{s+t} - B_t)(\omega) \quad \text{for } s \geq 0. \end{aligned}$$

It is clear that $\theta_t(B^{(t)})$ is a Brownian motion on $[0, +\infty)$, while $B^{(t)}$ is a Brownian motion on $[t, +\infty)$.

Lemma B.1.5. Let $(\Omega, \mathcal{F}, P)$ be a filtered probability space on which a Brownian motion B is defined. Fix a time $t > 0$ and denote by $\mathbb{F}^{B,t}$ the natural filtration generated by the increments of B from time t and completed with the $\mathbb{P}$ -null sets. Let U be a uniform random variable independent of B . Call $\mathbb{F}^t$ the filtration such that $\mathcal{F}_s^t = \sigma(\mathcal{F}_s^{B,t} \cup \sigma(U))$. Let H be any $\mathbb{F}^t$ -predictable process such that

$$\int_t^T |H_s|^2 ds < +\infty \quad \forall T \geq t,$$

i.e. $H \in M_{loc}^2([t, +\infty))$ and let $\xi \in L^2(\Omega, \sigma(U), \mathbb{P})$. Denote by $I(t, H, B)$ the stochastic integral of H with respect to the Brownian motion starting at time t and summed to ξ , i.e.

$$I_s(t, H, B) = \xi + \int_t^s H_r dB_r.$$

Then, the processes $\theta_t(I(t, H, B))$ and $I(0, \theta_t(H), \theta_t(B^{(t)}))$ are indistinguishable, i.e. $\mathbb{P}$ -a.s.

$$\xi + \int_t^{t+s} H_r dB_r = \xi + \int_0^s H_{r+t} d\theta_t(B^{(t)})_r \quad \forall s \geq 0. \quad (\text{B.2})$$

Proof. Step 1: We first need to prove that the second stochastic integral in (B.2) is well defined. We already mentioned that $\theta_t(B^{(t)})$ is a Brownian motion, thus having quadratic variation at time r equal to r . As

$$\int_t^T |H_s|^2 ds < +\infty \quad \forall T \geq t,$$

by a change of variables in the deterministic integral it is clear that

$$\int_0^{\tilde{T}} |H_{r+t}|^2 dr < +\infty \quad \forall \tilde{T} \geq 0.$$

If we want to ensure the integrability of $\theta_t(H)$ with respect to $\theta_t(B^{(t)})$, it is thus sufficient to prove the following statement: for any H $\mathbb{F}^t$ -predictable on $[t, +\infty)$, the process $\theta_t(H)$ is Θ -predictable on $[0, +\infty)$, where

$$\Theta = (\Theta_s^t)_{s \geq 0} \text{ such that } \Theta_s^t = \sigma(\mathcal{F}_s^{\theta_t(B^{(t)})} \cup \sigma(U))$$

and $\mathbb{F}^{\theta_t(B^{(t)})}$ is the completed Brownian filtration associated with $\theta_t(B^{(t)})$.

Substep 1.1: Let

$$\mathcal{LC} := \{H : [t, +\infty) \times \Omega \rightarrow \mathbb{R}^d : H \text{ is left continuous and } \mathbb{F}^t\text{-adapted}\}.$$

Thanks to Lemma 10.2 in [62], we know that the set $\mathcal{LC}$ generates the predictable σ -field $\mathcal{P}(\mathbb{F}^t)$ with respect to the filtration $\mathbb{F}^t$ on $[t, +\infty)$. Now, suppose $H \in \mathcal{LC}$. Then it is clear that $\theta_t(H)$ is left continuous. Moreover, $\theta_t(H)$ is Θ -adapted. As a matter of fact, notice that

$$\begin{aligned} \mathcal{F}_s^{\theta_t(B^{(t)})} &= \sigma(\theta_t(B^{(t)})(r) : 0 \leq r \leq s) \cup \mathcal{N} \\ &= \sigma(B_{t+r}^{(t)} : 0 \leq r \leq s) \cup \mathcal{N} \\ &= \sigma(B_{t+r} - B_t : 0 \leq r \leq s) \cup \mathcal{N} = \mathcal{F}_{s+t}^{B,t} \quad \forall s \geq 0, \end{aligned}$$

which implies $\Theta_s = \mathcal{F}_s^{\theta_t(B^{(t)})} \cup \sigma(U) = \mathcal{F}_{s+t}^{B,t} \cup \sigma(U) = \mathcal{F}_{s+t}^t$ for every $s \geq 0$. Now, for any fixed $s \geq 0$, $\theta_t(H)(s) = H_{s+t}$ is $\mathcal{F}_{s+t}^t$ -measurable because H is $\mathbb{F}^t$ -adapted. We conclude that, for any fixed $s \geq 0$, $\theta_t(H)_s$ is measurable with respect to Θ_s and $\theta_t(H)$ thus is Θ -adapted.

Substep 1.2: Suppose that the dimension d is equal to 1. We show that the thesis holds by a monotone class reasoning (following e.g. Theorem 2.2 in Chapter 0 of [81], which can easily be extended for unbounded functions by an approximation argument). Call

$$\mathcal{B} := \{H : [t, +\infty) \times \Omega \rightarrow \mathbb{R} : \theta_t(H) \text{ is } \Theta\text{-predictable on } [0, +\infty)\}.$$

By substep 1.1, $\mathcal{LC} \subseteq \mathcal{B}$. Moreover,

1. $\mathcal{LC}$ is stable under pointwise multiplication. As a matter of fact, if $H, K \in \mathcal{LC}$, then $H \cdot K$ is still left continuous and $\mathbb{F}^t$ -adapted.

2. $\mathcal{B}$ is a vector space because, if we take $H, K \in \mathcal{B}$ and $\lambda_1, \lambda_2 \in \mathbb{R}$, we get

$$\theta_t(\lambda_1 H + \lambda_2 K) = \lambda_1 \theta_t(H) + \lambda_2 \theta_t(K),$$

which inherits the measurability from $\theta_t(H)$ and $\theta_t(K)$ and is thus Θ -predictable.

3. All the constant processes are in $\mathcal{B}$. As a matter of fact, if $H_s(\omega) = h \in \mathbb{R}$ for any $s \geq t, \omega \in \Omega$, then also $\theta_t(H)$ is constantly equal to h , and constant processes are predictable with respect to any filtration.
4. If $\{H_n\}_{n \in \mathbb{N}}$ is an increasing sequence of positive processes belonging to $\mathcal{B}$ and converging pointwise to a finite limit $H := \sup_{n \in \mathbb{N}} H_n$, then $H \in \mathcal{B}$. As a matter of fact, it is clear that the sequence of positive elements $\{\theta_t(H_n)\}_{n \in \mathbb{N}}$ is increasing and converging to $\theta_t(H)$ because, for any $s \geq 0$,

$$\theta_t(H)(s) = H_{s+t} = \sup_{n \in \mathbb{N}} H_n(s+t) = \sup_{n \in \mathbb{N}} \theta_t(H_n)(s).$$

As $\{H_n\}_{n \in \mathbb{N}} \subseteq \mathcal{B}$, all the processes $\theta_t(H_n)$ are Θ -predictable, their supremum $\theta_t(H)$ inherits their measurability with respect to the predictable σ -algebra and thus is Θ -predictable. We conclude that $H \in \mathcal{B}$.

Now, by the monotone class theorem, $\mathcal{B}$ contains all the $\sigma(\mathcal{LC})$ -measurable processes, corresponding to all the $\mathbb{F}^t$ -predictable processes.

Substep 1.3: Consider the case of a general dimension $d \geq 1$. By proposition 1.5 in [94], we know that $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ is Borel isomorphic to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, i.e. there exists a bijection

$$f : \mathbb{R}^d \rightarrow \mathbb{R}$$

such that both f and f^{-1} are Borel measurable functions and

$$f(\mathcal{B}(\mathbb{R}^d)) = \mathcal{B}(\mathbb{R}).$$

Fix the Brownian motion B (which can also be multidimensional) and a $\mathbb{F}^t$ -predictable process

$$H : [t, +\infty) \times \Omega \rightarrow \mathbb{R}^d.$$

Define the process

$$\begin{aligned} \tilde{H} &: [t, +\infty) \times \Omega \rightarrow \mathbb{R} \\ \tilde{H}(s, \omega) &= f(H(s, \omega)). \end{aligned}$$

Notice that $\tilde{H}$ is $\mathbb{F}^t$ -predictable. By the previous substep, $\theta_t(\tilde{H})$ is Θ -predictable. Notice that

$$\theta_t(\tilde{H})(s) = \tilde{H}_{s+t} = f(H(s+t)) = f(\theta_t(H)(s)) \quad \forall s \geq 0,$$

from which we deduce that $\theta_t(\tilde{H}) = f \circ \theta_t(H)$. Recall that f is a bijection with Borel measurable inverse. Now, for any $E \in \mathcal{B}(\mathbb{R}^d)$, there exists a Borel set $\tilde{E} \in \mathcal{B}(\mathbb{R})$ such that $E = f^{-1}(\tilde{E})$. We thus have

$$(\theta_t(H))^{-1}(E) = (\theta_t(H))^{-1}(f^{-1}(\tilde{E})) = (f \circ \theta_t(H))^{-1}(\tilde{E}) = (\theta_t(\tilde{H}))^{-1}(\tilde{E}) \in \mathcal{P}(\Theta),$$

which means that $\theta_t(H)$ is Θ -predictable.

We conclude that the second integral in (B.2) is well defined under the hypotheses of the proposition.

Step 2: We want to prove the thesis by an approximation argument. We start by supposing that H is an elementary process. Thus, there exists an increasing sequence of times $\{s_i\}_{i \in \mathbb{N}}$ with $0 = s_0 < s_1 < s_2 < \dots$ and $\lim_{i \rightarrow +\infty} s_i = +\infty$ such that

$$H(s, \omega) = H_{-1}(\omega) \mathbb{1}_{\{t\}}(s) + \sum_{i \in \mathbb{N}} H_i(\omega) \mathbb{1}_{(s_i+t, s_{i+1}+t]}(s) \quad \forall s \geq t, \omega \in \Omega, \quad (\text{B.3})$$

where H_{-1} is a $\mathcal{F}_t^t$ -measurable random variable and H_i is a $\mathcal{F}_{s_i+t}^t$ -measurable random variable for every $i \in \mathbb{N}$. Notice that, if $H(s, \omega)$ has the form of equality (B.3), then $K := \theta_t(H)$ is an elementary process and for every $\omega \in \Omega$

$$\begin{aligned} K(s, \omega) &= H(s+t, \omega) = \sum_{i \in \mathbb{N}} H_i(\omega) \mathbb{1}_{(s_i+t, s_{i+1}+t]}(s+t) \\ &= \sum_{i \in \mathbb{N}} H_i(\omega) \mathbb{1}_{(s_i, s_{i+1}]}(s) \quad \forall s \geq 0, . \end{aligned} \quad (\text{B.4})$$

In this case, by definition of stochastic integral for elementary processes, for any $s \geq 0$

$$\begin{aligned} \xi + \int_t^{t+s} H_r dB_r &= \xi + \sum_{i \in \mathbb{N}} H_i \left[B_{(s_{i+1}+t) \wedge (s+t)} - B_{(s_i+t) \wedge (s+t)} \right] \\ &= \xi + \sum_{i \in \mathbb{N}} H_i \left[B_{(s_{i+1} \wedge s) + t} - B_{(s_i \wedge s) + t} \right]. \end{aligned}$$

Proceeding with the computations

$$\begin{aligned} \xi + \int_t^{t+s} H_r dB_r &= \xi + \sum_{i \in \mathbb{N}} H_i \left[B_{(s_{i+1} \wedge s) + t} - B_{(s_i \wedge s) + t} \right] \\ &= \xi + \sum_{i \in \mathbb{N}} H_i \left[B_{(s_{i+1} \wedge s) + t} - B_t - \left(B_{(s_i \wedge s) + t} - B_t \right) \right] \\ &= \xi + \sum_{i \in \mathbb{N}} H_i \left[B_{(s_{i+1} \wedge s) + t}^{(t)} - B_{(s_i \wedge s) + t}^{(t)} \right] \\ &= \xi + \sum_{i \in \mathbb{N}} H_i \left[\theta_t(B_{(s_{i+1} \wedge s)}^{(t)}) - \theta_t(B_{(s_i \wedge s)}^{(t)}) \right] \quad (\text{by (B.4)}) \\ &= \xi + \int_0^s H_{r+t} d(\theta_t(B^{(t)}))_r. \end{aligned}$$

We conclude that $\mathbb{P}$ -a.s.

$$\theta_t(I(t, H, B))(s) = I(0, \theta_t(H), \theta_t(B^{(t)}))(s) \quad \text{for every } s \geq 0.$$

Step 3: Suppose that $H \in M^2([t, +\infty))$, i.e. that

$$\mathbb{E} \left[\int_t^T |H_r|^2 dr \right] < +\infty \quad \forall T \geq t.$$

Then, by Lemma 7.2 in [9], there exists a sequence of elementary processes $\{H^n\}_{n \in \mathbb{N}}$ such that

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left[\int_t^T |H_r - H_r^n|^2 dr \right] = 0; \quad (\text{B.5})$$

by definition of stochastic integral, we thus know that

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left[\left| I(t, H^n, B) - I(t, H, B) \right|^2 \right] = 0.$$

Now, with a simple change of variables in (B.5), we can say that

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left[\int_0^{T'} |H_{r'+t} - H_{r'+t}^n|^2 dr \right] = 0 \quad \forall T' \geq 0,$$

which means that $\{\theta_t(H^n)\}_{n \in \mathbb{N}}$ converges to $\theta_t(H)$ in $M^2([0, T'])$ for all $T' > 0$. Moreover, the sequence of processes and their limit are predictable with respect to Θ . Thus, by definition of stochastic integral, the following equation holds by Corollary 7.1 in [9]:

$$\lim_{n \rightarrow +\infty} \mathbb{E} \left[\left| \int_0^T \theta_t(H)(r) d(\theta_t(B)(r)) - \int_t^T \theta_t(H^n)(r) d(\theta_t(B)(r)) \right|^2 \right] = 0 \quad \forall T > t,$$

meaning that

$$L^2 - \lim_{n \rightarrow +\infty} I(0, \theta_t(H^n), \theta_t(B^{(t)})) = I(0, \theta_t(H), \theta_t(B^{(t)})).$$

Notice that convergence in L^2 implies pointwise convergence (up to subsequencies) and that the application of the shift operator preserves $\mathbb{P}$ -a.s. convergence, from which we deduce that

$$\lim_{n \rightarrow +\infty} \theta_t(I(t, H^n, B))(s) = \theta_t(I(t, H, B))(s) \quad \forall s \geq 0$$

and

$$\lim_{n \rightarrow +\infty} I(0, \theta_t(H^n), \theta_t(B^{(t)}))(s) = I(0, \theta_t(H), \theta_t(B^{(t)}))(s) \quad \forall s \geq 0.$$

As, for all $n \in \mathbb{N}$, H^n is an elementary process, by the previous step we can affirm that

$$\theta_t(I(t, H^n, B))(s) = I(0, \theta_t(H^n), \theta_t(B^{(t)}))(s) \quad \text{for every } s \geq 0, \text{ for every } n \in \mathbb{N}.$$

By uniqueness of the pointwise limit, we conclude that

$$\theta_t(I(t, H, B))(s) = I(0, \theta_t(H), \theta_t(B^{(t)}))(s) \quad \forall s \geq 0.$$

Step 4: Consider the most general case, when $H \in M_{loc}^2([t, +\infty))$, i.e.

$$\int_t^T |H_s|^2 ds < +\infty \quad \forall T \geq t \quad \mathbb{P}\text{-a.s.}$$

Then, there exists a sequence of processes $\{H^n\}_{n \in \mathbb{N}} \subseteq M^2([t, +\infty))$ such that

$$\lim_{n \rightarrow +\infty} H^n = H \quad \mathbb{P}\text{-a.s.}$$

and, by Definition 7.2 in [9], the stochastic integral is defined as

$$I(t, B, H)(r) = \int_t^r H_s dB_s = \lim_{n \rightarrow +\infty} \int_t^r H_s^n dB_s \quad \mathbb{P}\text{-a.s.}$$

Repeating the reasoning in the previous step, we can say that, as $\theta_t(H^n)$ still is in $M^2([0, +\infty))$ for all $n \in \mathbb{N}$ (by a simple change of the integration variable), the integral

$$\int_0^s \theta_t(H)(r) d\theta_t(B)(r)$$

still is defined as the $\mathbb{P}$ -a.s limit of

$$\int_0^s \theta_t(H^n)(r) d\theta_t(B)(r).$$

The convergence $\mathbb{P}$ -a.s. of $\{H^n\}_{n \in \mathbb{N}}$ to H clearly implies $\mathbb{P}$ -a.s. convergence of $\{\theta_t(H^n)\}_{n \in \mathbb{N}}$ to $\theta_t(H)$, and all the convergences in probability imply $\mathbb{P}$ -a.s. convergence up to subsequences. By the previous step, we know that

$$\theta_t(I(t, H^n, B)) = I(0, \theta_t(H^n), \theta_t(B^{(t)}));$$

passing to the a.s. limit in both sides, we conclude that

$$\theta_t(I(t, H, B)) = I(0, \theta_t(H), \theta_t(B^{(t)}))$$

and the thesis is verified. $\square$

We will apply the previous results to show the following Translation Invariance Property of solutions to SDEs: if the joint law of the initial conditions, the controls and the Brownian motions are the same, then the solutions to the associated SDEs have the same joint law with the associated controls. The proof of the following result is inspired by Theorem 9.5 in [9].

Proposition B.1.6. Suppose that hypotheses $(\mathbf{H}_{b,\sigma})$ hold for the coefficients b, σ . Suppose that B^1, B^2 are two Brownian motions and U^1, U^2 are two uniform random variables on $[0, 1]$ such that U^i is independent of B^i for $i = 1, 2$. Denote by $\mathbb{F}^i$ the initial enlargement of the completed Brownian filtration of B^i by means of $\sigma(U^i)$, $i = 1, 2$. Choose two initial conditions $\xi^i \in L^2(\Omega, \sigma(U^i), \mathbb{P})$ for $i = 1, 2$. Fix two controls α^1, α^2 predictable, respectively, with respect to $\mathbb{F}^1$ and $\mathbb{F}^2$. Call $X^i := X^{t, \xi^i, B^i, \alpha^i}$ the unique solution to the following controlled stochastic differential equation of the McKean-Vlasov type:

$$\begin{aligned} dX_s^i &= b(X_s^i, \mathbb{P}_{X_s^i}, \alpha_s^i) ds + \sigma(X_s^i, \mathbb{P}_{X_s^i}, \alpha_s^i) dB_s^i \\ X_t &= \xi^i \quad i = 1, 2. \end{aligned} \tag{B.6}$$

Furthermore, assume that

$$\mathcal{L}(\xi^1, B^1, \alpha^1) = \mathcal{L}(\xi^2, B^2, \alpha^2);$$

then

$$\mathcal{L}(X^1, \alpha^1, B^1) = \mathcal{L}(X^2, \alpha^2, B^2)$$

(the equality in law “propagates” in time).

Proof. Step 1: We proceed by induction on Picard’s iterates. We denote by $X^{i,n}$ the n -th Picard iterate associated to the process X^i , defined as in Theorem A.3.1.

Start from the basis, for $n = 0$: by definition,

$$X_s^{i,0} = \xi^i \quad \text{for every } s \geq t, i = 1, 2.$$

By hypothesis,

$$\mathcal{L}(\xi^1, B^1, \alpha^1) = \mathcal{L}(\xi^2, B^2, \alpha^2),$$

which obviously implies that

$$\mathcal{L}(X^{1,0}, B^1, \alpha^1) = \mathcal{L}(X^{2,0}, B^2, \alpha^2).$$

We now consider the inductive step. Suppose that

$$\mathcal{L}(X^{1,n}, B^1, \alpha^1) = \mathcal{L}(X^{2,n}, B^2, \alpha^2) \tag{B.7}$$

for the n -th Picard iterate of X^i . We show that the thesis holds for the $(n + 1)$ -th Picard iterate of X^i . By definition,

$$X^{i,n+1} = \xi^i + \int_t^\cdot b(X_r^{i,n}, \mathbb{P}_{X_r^{i,n}}, \alpha_r^i) dr + \int_t^\cdot \sigma(X_r^{i,n}, \mathbb{P}_{X_r^{i,n}}, \alpha_r^i) dB_r^i \quad \forall n \in \mathbb{N}.$$

Notice that equation (B.7) immediately implies that

$$\begin{aligned} &\mathcal{L}(\xi^1, b(X^{1,n}, \mathbb{P}_{X^{1,n}}, \alpha^1), \sigma(X^{1,n}, \mathbb{P}_{X^{1,n}}, \alpha^1), B^1, \alpha^1) = \\ &\mathcal{L}(\xi^2, b(X^{2,n}, \mathbb{P}_{X^{2,n}}, \alpha^2), \sigma(X^{2,n}, \mathbb{P}_{X^{2,n}}, \alpha^2), B^2, \alpha^2) \end{aligned}$$

by the measurability of b and σ . Then, we can repeat the proof of Theorem B.1.2 including the controls in the vector of elements whose law is compared, observing that we can start with a sequence of elementary approximations $\{H^{i,n}\}_{n \in \mathbb{N}}$ of $H^i = \sigma(X^{i,n}, \mathbb{P}_{X^{i,n}}, \alpha^i)$ and get

$$\begin{aligned} &\mathcal{L}\left(\xi^1 + \int_t^\cdot b(X_r^{1,n}, \mathbb{P}_{X_r^{1,n}}, \alpha_r^1) dr + \int_t^\cdot \sigma(X_r^{1,n}, \mathbb{P}_{X_r^{1,n}}, \alpha_r^1) dB_r^1, B^1, \alpha^1\right) \\ &= \mathcal{L}\left(\xi^2 + \int_t^\cdot b(X_r^{2,n}, \mathbb{P}_{X_r^{2,n}}, \alpha_r^2) dr + \int_t^\cdot \sigma(X_r^{2,n}, \mathbb{P}_{X_r^{2,n}}, \alpha_r^2) dB_r^2, B^2, \alpha^2\right). \end{aligned}$$

We conclude that

$$\mathcal{L}(X^{1,(n+1)}, B^1, \alpha^1) = \mathcal{L}(X^{2,(n+1)}, B^2, \alpha^2),$$

which means that the thesis holds for all the Picard iterates of the processes X^i , $i = 1, 2$.

Step 2: Notice that, as showed in Appendix A.3, a solution X to the McKean-Vlasov SDE (B.6) is defined as the $\mathbb{P}$ -a.s. limit (uniform with respect to time) of the sequence of Picard iterates $\{X^{i,n}\}_{n \in \mathbb{N}}$ for $i = 1, 2$. Then, for any choice of $m \in \mathbb{N}$ and $0 \leq s_1 \leq \dots \leq s_m$ we can always extract a subsequence (that we will not rename) such that

$$\mathcal{L}(X^{1,(n+1)}(s_1), \dots, X^{1,(n+1)}(s_m), B^1(s_1), \dots, B^1(s_m), \alpha^1(s_1), \dots, \alpha^1(s_m))$$

converges weakly to

$$\mathcal{L}(X^1(s_1), \dots, X^1(s_m), B^1(s_1), \dots, B^1(s_m), \alpha^1(s_1), \dots, \alpha^1(s_m))$$

and

$$\mathcal{L}(X^{2,(n+1)}(s_1), \dots, X^{2,(n+1)}(s_m), B^2(s_1), \dots, B^2(s_m), \alpha^2(s_1), \dots, \alpha^2(s_m))$$

converges weakly to

$$\mathcal{L}(X^2(s_1), \dots, X^2(s_m), B^2(s_1), \dots, B^2(s_m), \alpha^2(s_1), \dots, \alpha^2(s_m)).$$

On the other hand, by the previous step

$$\begin{aligned} & \mathcal{L}(X^{1,(n+1)}(s_1), \dots, X^{1,(n+1)}(s_m), B^1(s_1), \dots, B^1(s_m), \alpha^1(s_1), \dots, \alpha^1(s_m)) \\ &= \mathcal{L}(X^{2,(n+1)}(s_1), \dots, X^{2,(n+1)}(s_m), B^2(s_1), \dots, B^2(s_m), \alpha^2(s_1), \dots, \alpha^2(s_m)); \end{aligned}$$

by uniqueness of the weak limit, we deduce that

$$\begin{aligned} & \mathcal{L}(X^1(s_1), \dots, X^1(s_m), B^1(s_1), \dots, B^1(s_m), \alpha^1(s_1), \dots, \alpha^1(s_m)) \\ &= \mathcal{L}(X^2(s_1), \dots, X^2(s_m), B^2(s_1), \dots, B^2(s_m), \alpha^2(s_1), \dots, \alpha^2(s_m)), \end{aligned}$$

which implies the thesis by the arbitrariness of $m \in \mathbb{N}$ and $s_1, \dots, s_m \in [0, +\infty)$. $\square$

We now get to the fundamental Lemma 1.3.2, telling us that we can “translate” our equation in time keeping the same joint law of the solution and the associated control. We recall here the statement of the Lemma using the notation introduced in this Appendix.

Lemma B.1.7. Consider two solutions of the same controlled stochastic equation of McKean-Vlasov type starting at different initial times from the same initial condition and under different controls. More precisely, assume that $(\Omega, \mathbb{F}, \mathbb{P})$ is a complete probability space, on which we define a Brownian motion B and an independent uniform random variable U . As usual, denote by $B^{(t)}$ the process such that $B_s^{(t)} = B_s - B_t$ for all $s \geq t$. Call $\mathbb{F}^B$ the completed Brownian filtration on $[0, +\infty)$ associated with B and $\mathbb{F}^{B,t}$ the completed Brownian filtration on $[t, +\infty)$ associated with $B^{(t)}$. Let $\mathbb{F} = (\mathcal{F}_s)_{s \geq 0}$ and $\mathbb{F}^t = (\mathcal{F}_s^t)_{s \geq t}$ be the filtrations such that $\mathcal{F}_s = \mathcal{F}_s^B \cup \mathcal{G}$ and $\mathcal{F}_s^t = \mathcal{F}_s^{B,t} \cup \mathcal{G}$, where $\mathcal{G} = \sigma(U)$. Consider two controls α and γ being $\mathbb{F}$ and

$\mathbb{F}^t$ -predictable, respectively. We denote by $X^{(0)} = (X_s^{(0)})_{s \geq 0}$ the unique solution to equation

$$\begin{cases} dX_s = b(X_s, \mathbb{P}_{X_s}, \alpha_s)ds + \sigma(X_s, \mathbb{P}_{X_s}, \alpha_s)dB_s & \forall s \geq 0 \\ X_0 = \xi \end{cases}$$

and by $X^{(t)} = (X_s^{(t)})_{s \geq 0}$ the unique solution to

$$\begin{cases} dX_s = b(X_s, \mathbb{P}_{X_s}, \gamma_s)ds + \sigma(X_s, \mathbb{P}_{X_s}, \gamma_s)dB_s & \forall s \geq t \\ X_s = \xi, \end{cases}$$

where the coefficients satisfy Assumptions $(\mathbf{H}_{b,\sigma})$. Assume furthermore that

$$\mathcal{L}(\xi, B, \alpha) = \mathcal{L}(\xi, \theta_t(B^{(t)}), \theta_t(\gamma)).$$

Then,

$$\mathcal{L}(X^{(0)}, \alpha) = \mathcal{L}(\theta_t(X^{(t)}), \theta_t(\gamma)).$$

Remark B.1.8. Notice that we defined $X^{(t)}$ on the interval $[t, +\infty)$. The process $\theta_t(X^{(t)})$ will thus be defined on $[0, +\infty)$ and its initial value will be

$$(\theta_t(X^{(t)}))_0 = \xi.$$

When we write that $\mathcal{L}(X^{(0)}, \alpha) = \mathcal{L}(\theta_t(X^{(t)}), \theta_t(\gamma))$, we mean equivalently that, for any choice of $n \in \mathbb{N}$ and $t_1, \dots, t_n \geq 0$, we have

$$\begin{aligned} \mathcal{L}\left(X_{t_1}^{(0)}, \alpha_{t_1}, \dots, X_{t_n}^{(0)}, \alpha_{t_n}\right) &= \mathcal{L}\left(\theta_t(X^{(t)})_{t_1}, \theta_t(\gamma)_{t_1}, \dots, \theta_t(X^{(t)})_{t_n}, \theta_t(\gamma)_{t_n}\right) \\ &= \mathcal{L}\left(X_{t_1+t}^{(t)}, \gamma_{t+t_1}, \dots, X_{t_n+t}^{(t)}, \gamma_{t+t_n}\right). \end{aligned}$$

Proof. By Proposition B.1.6, it is sufficient to prove that the process $\theta_t(X^{(t)})$ is a solution to the same SDE as $X^{(0)}$ under the control $\theta_t(\gamma)$ and with respect to the Brownian motion $\theta_t(B^{(t)})$ instead of α and B , respectively. Recall that

$$X_r^{(0)} = \xi + \int_0^r b(X_s^{(0)}, \mathbb{P}_{X_s^{(0)}}, \alpha_s)ds + \int_0^r \sigma(X_s^{(0)}, \mathbb{P}_{X_s^{(0)}}, \alpha_s)dB_s \quad \text{for all } r \geq 0. \quad (\text{B.8})$$

On the other hand, compute

$$\theta_t(X^{(t)})_r = X_{t+r}^{(t)} = \xi + \int_t^{t+r} b(X_s^{(t)}, \mathbb{P}_{X_s^{(t)}}, \gamma_s)ds + \int_t^{t+r} \sigma(X_s^{(t)}, \mathbb{P}_{X_s^{(t)}}, \gamma_s)dB_s \quad \text{for all } r \geq 0. \quad (\text{B.9})$$

In order to apply Proposition B.1.6, it is sufficient to prove that

$$\begin{aligned} \theta_t(X^{(t)})_r &= \xi + \int_0^r b(\theta_t(X^{(t)})_s, \mathbb{P}_{\theta_t(X^{(t)})_s}, \theta_t(\gamma)_s)ds \\ &\quad + \int_0^r \sigma(\theta_t(X^{(t)})_s, \mathbb{P}_{\theta_t(X^{(t)})_s}, \theta_t(\gamma)_s)d\theta_t(B^{(t)})_s \quad \text{for all } r \geq 0. \end{aligned} \quad (\text{B.10})$$

For simplicity, we will restrict to the case where $b = 0$. Notice that the proof can be easily extended to the case of a non-null drift term because, by a simple change of variable $z = s - t$ in the Lebesgue integral, we deduce that

$$\begin{aligned} \int_t^{t+r} b(X_s^{(t)}, \mathbb{P}_{X_s^{(t)}}, \gamma_s) ds &= \int_0^r b(X_{z+t}^{(t)}, \mathbb{P}_{X_{z+t}^{(t)}}, \gamma_{z+t}) dz \\ &= \int_0^r b(\theta_t(X^{(t)})_z, \mathbb{P}_{\theta_t(X^{(t)})_z}, \theta_t(\gamma)_z) dz. \end{aligned}$$

If $b = 0$, it is sufficient to notice that

$$\sigma(\theta_t(X^{(t)})_z, \mathbb{P}_{\theta_t(X^{(t)})_z}, \theta_t(\gamma)_z) = \sigma(X_{z+t}^{(t)}, \mathbb{P}_{X_{z+t}^{(t)}}, \gamma_{z+t}) = \theta_t(\sigma(X^{(t)}, \mathbb{P}_{X^{(t)}}, \gamma))(z).$$

Using Lemma B.1.5, we can say that $\mathbb{P}$ -a.s.

$$\xi + \int_t^{t+} \sigma(X_s^{(t)}, \mathbb{P}_{X_s^{(t)}}, \gamma_s) dB_s = \xi + \int_0^+ \theta_t(\sigma(X^{(t)}, \mathbb{P}_{X^{(t)}}, \gamma)(s)) d\theta_t(B^{(t)})_s,$$

thus implying (B.10) as equation (B.9) holds. $\square$

B.2 Well-posedness of the disintegration formula

In this appendix, we prove the well-posedness of the “disintegration system” of the McKean-Vlasov equation with respect to the first part of the Brownian trajectory, i.e. Lemma 1.3.10. Notice that the system (1.34) is not a system of classical SDEs or a system of McKean-Vlasov SDEs; it is a system of SDEs parametrised by $w \in \mathcal{C}([0, r])$, which does not belong to a finite or countable set of indices, where the “particles” X^w interact through the term

$$\int_{\mathcal{C}([0, r])} \mathbb{P}_{X_s^{w'}} W_{[0, r]}(dw').$$

Proof of Lemma 1.3.10. It is sufficient to prove the result on a bounded interval $[r, T]$ for any $T > r$ and then to “glue” the solutions to construct the unique solution on the interval $[0, +\infty)$ with a standard procedure.

Fix any $T > r$. For any flow of probability measures $\nu = (\nu_s)_{s \in [r, T]} \in \mathcal{C}([r, T]; \mathcal{P}_2(\mathbb{R}^d))$, we can define the system of SDEs

$$\begin{cases} dX_s^{w, \nu} = b(X_s^{w, \nu}, \nu_s, \alpha_s^r(w)) ds + \sigma(X_s^{w, \nu}, \nu_s, \alpha_s^r(w)) dB_s, & s \in [r, T] \\ X_r^{w, \nu} = \xi \end{cases} \quad (\text{B.11})$$

for every $w \in \mathcal{C}([0, r])$. For every fixed ν , (B.11) is a system of decoupled SDEs with random initial condition, each of which indexed by a different parameter $w \in \mathcal{C}([0, r])$. Under assumptions $(\mathbf{H}_{b, \sigma})$, there thus exists a unique solution to (B.11), denoted by $X^{w, \nu}$ and such that $\mathbb{P}$ -a.s.

$$X_s^{w, \nu} = \xi + \int_r^s b(X_u^{w, \nu}, \nu_u, \alpha_u^r(w)) du + \int_r^s \sigma(X_u^{w, \nu}, \nu_u, \alpha_u^r(w)) dB_u$$

for every $s \geq r$ and for every $w \in \mathcal{C}([0, r])$.

We now define a map

$$\begin{aligned}\Phi : \mathcal{C}([r, T]; \mathcal{P}_2(\mathbb{R}^d)) &\rightarrow \mathcal{C}([r, T]; \mathcal{P}_2(\mathbb{R}^d)) \\ \nu = (\nu_s)_{s \geq 0} &\rightarrow (\Phi(\nu)(s))_{s \geq 0}\end{aligned}$$

such that

$$\Phi(\nu)(s)(E) := \int_{\mathcal{C}([0, r])} \mathbb{P}_{X_s^{w, \nu}}(E) W_{[0, r]}(dw) \quad \text{for all } E \in \mathcal{B}(\mathbb{R}^d)$$

for every $s \geq r$. Notice that Φ is well defined because, for every fixed $E \in \mathcal{B}(\mathbb{R}^d)$ and $s \geq r$, the map $w \rightarrow \mathbb{P}_{X_s^{w, \nu}}(E)$ is in $L^1(\mathcal{C}([0, r]), \mathcal{B}(\mathcal{C}([0, r])), W_{[0, r]})$ by Fubini theorem and Section 12 in [87], which guarantees the joint measurability of $X^{w, \nu}$ with respect to (s, ω, w) . Clearly, there exists a unique solution to the system (1.34) if and only if the map Φ admits a unique fixed point $\hat{\nu}$.

We endow the space $\mathcal{C}([r, T]; \mathcal{P}_2(\mathbb{R}^d))$ with the uniform-in-time Wasserstein distance, i.e. for any $\mu, \nu \in \mathcal{C}([r, T]; \mathcal{P}_2(\mathbb{R}^d))$

$$d(\mu, \nu) := \sup_{r \leq s \leq T} W_2(\mu_s, \nu_s).$$

Consider two flows of probability measures $\mu, \nu \in \mathcal{C}([r, T]; \mathcal{P}_2(\mathbb{R}^d))$. Compute

$$\begin{aligned}&\sup_{r \leq s \leq T} W_2^2(\Phi(\mu)(s), \Phi(\nu)(s)) \\ &= \sup_{r \leq s \leq T} W_2^2\left(\int_{\mathcal{C}([0, r])} \mathbb{P}_{X_s^{w, \mu}} W_{[0, r]}(dw), \int_{\mathcal{C}([0, r])} \mathbb{P}_{X_s^{w, \nu}} W_{[0, r]}(dw)\right).\end{aligned}\tag{B.12}$$

We estimate the right hand side of equation (B.12) by means of the Kantorovich Duality Theorem (see e.g. Theorem 5.10 in [90]):

$$\begin{aligned}&W_2^2\left(\int_{\mathcal{C}([0, r])} \mathbb{P}_{X_s^{w, \mu}} W_{[0, r]}(dw), \int_{\mathcal{C}([0, r])} \mathbb{P}_{X_s^{w, \nu}} W_{[0, r]}(dw)\right) \\ &= \sup_{\substack{\phi, \psi \in C_b(\mathbb{R}^d; \mathbb{R}): \\ \phi(x) + \psi(y) \leq |x - y|^2}} \left[\int_{\mathbb{R}^d} \phi(x) \left(\int_{\mathcal{C}([0, r])} \mathbb{P}_{X_s^{w, \mu}} W_{[0, r]}(dw) \right) (dx) + \right. \\ &\quad \left. \int_{\mathbb{R}^d} \psi(y) \left(\int_{\mathcal{C}([0, r])} \mathbb{P}_{X_s^{w, \nu}} W_{[0, r]}(dw) \right) (dy) \right] \quad (\text{by Fubini theorem}) \\ &= \sup_{\substack{\phi, \psi \in C_b(\mathbb{R}^d; \mathbb{R}): \\ \phi(x) + \psi(y) \leq |x - y|^2}} \int_{\mathcal{C}([0, r])} \left[\int_{\mathbb{R}^d} \phi(x) \mathbb{P}_{X_s^{w, \mu}}(dx) + \int_{\mathbb{R}^d} \psi(y) \mathbb{P}_{X_s^{w, \nu}}(dy) \right] W_{[0, r]}(dw) \\ &\leq \int_{\mathcal{C}([0, r])} W_2^2(\mathbb{P}_{X_s^{w, \mu}}, \mathbb{P}_{X_s^{w, \nu}}) W_{[0, r]}(dw) \\ &\leq \int_{\mathcal{C}([0, r])} \mathbb{E} \left[|X_s^{w, \mu} - X_s^{w, \nu}|^2 \right] W_{[0, r]}(dw),\end{aligned}\tag{B.13}$$

where the second-to-last inequality follows applying again Kantorovich duality theorem. Putting equation (B.12) and inequality (B.13) together, we can say that

$$\sup_{r \leq s \leq T} W_2^2(\Phi(\mu)(s), \Phi(\nu)(s)) \leq \sup_{r \leq s \leq T} \int_{\mathcal{C}([0, r])} \mathbb{E} \left[|X_s^{w, \mu} - X_s^{w, \nu}|^2 \right] W_{[0, r]}(dw) \quad (\text{B.14})$$

Proceeding as in the proof of Theorem 4.21 in [32], we can show that, for $C = 4L^2(T + 4)$,

$$\mathbb{E} \left[|X_s^{w, \mu} - X_s^{w, \nu}|^2 \right] \leq C e^{C(s-r)} \int_r^s W_2^2(\nu_t, \mu_t) dt \quad \text{for all } s \in [r, T]. \quad (\text{B.15})$$

Pulling inequality (B.15) back into (B.14), we deduce that there exists a constant $K(T) > 0$ such that

$$\sup_{r \leq s \leq T} W_2^2(\Phi(\mu)(s), \Phi(\nu)(s)) \leq K(T) \int_r^T W_2^2(\nu_t, \mu_t) dt. \quad (\text{B.16})$$

Now denote by Φ^n the composition of n times the map Φ . Iterating (B.16), we conclude that

$$\begin{aligned} \sup_{t \leq s \leq T} W_2^2(\Phi^n(\mu)(s), \Phi^n(\nu)(s)) &\leq K^n(T) \int_r^T \frac{(T-s)^{n-1}}{(n-1)!} W_2^2(\mu_s, \nu_s) ds \\ &\leq K^n(T) \frac{(T-r)^n}{n!} \sup_{r \leq s \leq T} W_2^2(\mu_s, \nu_s). \end{aligned}$$

For sufficiently large $n \in \mathbb{N}$, $K^n(T) \frac{(T-r)^n}{n!} < 1$ and Φ^n is a contraction.

$\mathcal{C}([r, T]; \mathcal{P}_2(\mathbb{R}^d))$ endowed with the uniform-in-time Wasserstein distance is a non-empty complete metric space; hence, by Banach-Caccioppoli theorem, there exists a unique fixed point for the map Φ^n . Therefore, Φ admits a unique fixed point and the thesis holds. $\square$

B.3 Proof of the continuity of the gain functional

We conclude the Appendix with the proof of the continuity of J with respect to time uniformly with respect to the control parameter (Proposition 1.3.13).

Proof of Proposition 1.3.13. Fix any $\alpha \in \mathcal{A}$, $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ and $\epsilon > 0$. Compute, for $u > t$,

$$|J(u, \xi, \alpha) - J(t, \xi, \alpha)| \leq I_1 + I_2 \quad (\text{B.17})$$

with

$$I_1 := \mathbb{E} \left[\int_u^{+\infty} |e^{-\beta(s-u)} f(X_s^{u, \xi, \alpha}, \mathbb{P}_{X_s^{u, \xi, \alpha}}, \alpha_s) - e^{-\beta(s-t)} f(X_s^{t, \xi, \alpha}, \mathbb{P}_{X_s^{t, \xi, \alpha}}, \alpha_s)| ds \right] \quad (\text{B.18})$$

$$I_2 := \mathbb{E} \left[\int_t^u e^{-\beta(s-t)} |f(X_s^{t, \xi, \alpha}, \mathbb{P}_{X_s^{t, \xi, \alpha}}, \alpha_s)| ds \right]. \quad (\text{B.19})$$

We firstly consider I_2 . Under hypotheses $(\mathbf{H}_{f,\beta})$, using the estimate (1.4)

$$\begin{aligned} I_2 &\leq K \int_t^u e^{-\beta(s-t)} \mathbb{E}[1 + |X_s^{t,\xi,\alpha}|^2 + \|\mathbb{P}_{X_s^{t,\xi,\alpha}}\|_2^2] ds \\ &\leq \frac{K}{\beta} [1 - e^{-\beta(u-t)}] + 2K \left(\|\xi\|_{L^2}^2 + 6M^2(u-t) \right) (u-t) e^{(12M^2+1-\beta)(u-t)} < \frac{\epsilon}{2} \end{aligned} \quad (\text{B.20})$$

for $u-t < \zeta_1(\epsilon, M, K, \beta, t, \xi)$ independent of α , provided that $\beta > 12M^2 + 1$, which is guaranteed by $(\mathbf{H}_{f,\beta})$. We now study I_1 . Notice that, for any choice of $T > 0$, we can write

$$I_1 = I_T + R_T, \quad (\text{B.21})$$

where

$$I_T = \mathbb{E} \left[\int_u^T |e^{-\beta(s-u)} f(X_s^{u,\xi,\alpha}, \mathbb{P}_{X_s^{u,\xi,\alpha}}, \alpha_s) - e^{-\beta(s-t)} f(X_s^{t,\xi,\alpha}, \mathbb{P}_{X_s^{t,\xi,\alpha}}, \alpha_s)| ds \right]$$

and

$$R_T = \mathbb{E} \left[\int_T^{+\infty} |e^{-\beta(s-u)} f(X_s^{u,\xi,\alpha}, \mathbb{P}_{X_s^{u,\xi,\alpha}}, \alpha_s) - e^{-\beta(s-t)} f(X_s^{t,\xi,\alpha}, \mathbb{P}_{X_s^{t,\xi,\alpha}}, \alpha_s)| ds \right].$$

Consider the remainder R_T . Notice that, under hypotheses $(\mathbf{H}_{f,\beta})$, using again estimate (1.4)

$$\begin{aligned} R_T &\leq K \int_T^{+\infty} e^{-\beta(s-u)} (1 + 2\mathbb{E}[|X_s^{u,\xi,\alpha}|^2]) ds + K \int_T^{+\infty} e^{-\beta(s-t)} (1 + 2\mathbb{E}[|X_s^{t,\xi,\alpha}|^2]) \\ &\leq K \int_T^{+\infty} e^{-\beta(s-u)} ds + 2K \int_T^{+\infty} e^{(12M^2+1-\beta)(s-u)} \left(\|\xi\|_{L^2}^2 + 6M^2(s-u) \right) ds \\ &\quad + K \int_T^{+\infty} e^{-\beta(s-t)} ds + 2K \int_T^{+\infty} e^{(12M^2+1-\beta)(s-t)} \left(\|\xi\|_{L^2}^2 + 6M^2(s-t) \right) ds \\ &\leq 2Ke^{\beta(t+1)} \int_T^{+\infty} e^{-\beta s} ds + 4Ke^{\beta(t+1)} \int_T^{+\infty} e^{(12M^2+1-\beta)s} \left(\|\xi\|_{L^2}^2 + 6M^2(s-t) \right) ds \\ &< \frac{\epsilon}{4} \end{aligned} \quad (\text{B.22})$$

provided that $u-t \leq 1$ and $T = T(\epsilon)$ (independent of α) is sufficiently large, as all the terms appearing in the previous inequality are remainders of convergent integrals. As equality (B.21) holds for any choice of T , we fix $T = T(\epsilon)$.

We now study I_T . Firstly, notice that, under hypotheses $(\mathbf{H}_{f_{cont}})$, the function

$$\psi : (r, x, \mu, a) \rightarrow e^{\beta r} f(x, \mu, a)$$

is locally Hölder continuous in (x, μ) and Lipschitz in r with Lipschitz constant depending on (x, μ) , uniformly with respect to a ; namely, there exists a constant $H' > 0$ such that, for any $r, r' \in [u, T]$, $x, x' \in \mathbb{R}^d$, $\mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$ and $a \in A$

$$|\psi(r, x, \mu, a) - \psi(r', x', \mu', a)| \leq H' [|x - x'|^{\gamma_1} (1 + |x| + |x'|)^{2-\gamma_1}]$$

$$\begin{aligned}
& + W_2(\mu, \mu')^{\gamma_2} (1 + \|\mu\|_2 + \|\mu'\|_2)^{2-\gamma_2} \\
& + |r - r'| (1 + |x'|^2 + \|\mu'\|_2^2). \tag{B.23}
\end{aligned}$$

By estimate (1.4) restricted to the interval $[0, T]$, we can easily deduce that

$$\mathbb{E}[|X_r^{t,\xi,\alpha}|^2] \leq \left(\|\xi\|_{L^2}^2 + 6M^2T \right) e^{(12M^2+1)T} := h(T, M, \xi), \tag{B.24}$$

and that the same estimate holds for $\mathbb{E}[|X_r^{u,\xi,\alpha}|^2]$. Furthermore, we recall the continuity condition (A.3) of the solution to a McKean-Vlasov SDE with respect to the initial time on a bounded time interval: there exists a constant $K_0(T)$ such that, for any $0 \leq t < u \leq T$, $\xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$ and $\alpha \in \mathcal{A}$,

$$\mathbb{E} \left[|X_s^{t,\xi,\alpha} - X_s^{u,\xi,\alpha}|^2 \right] \leq K_0(T)(u-t)(1 + \|\xi\|_{L^2}^2)$$

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

$$\begin{aligned}
I_T &= \mathbb{E} \left[\int_u^T e^{-\beta s} |\psi(u, X_s^{u,\xi,\alpha}, \mathbb{P}_{X_s^{u,\xi,\alpha}}, \alpha_s) - \psi(t, X_s^{t,\xi,\alpha}, \mathbb{P}_{X_s^{t,\xi,\alpha}}, \alpha_s)| ds \right] \\
&\leq \mathbb{E} \left[\int_u^T H' \left(|X_s^{u,\xi,\alpha} - X_s^{t,\xi,\alpha}|^{\gamma_1} (1 + |X_s^{u,\xi,\alpha}| + |X_s^{t,\xi,\alpha}|)^{2-\gamma_1} \right) ds \right] \\
&+ H' \mathbb{E} \left[\int_u^T \left(W_2(\mathbb{P}_{X_s^{u,\xi,\alpha}}, \mathbb{P}_{X_s^{t,\xi,\alpha}})^{\gamma_2} (1 + \|\mathbb{P}_{X_s^{u,\xi,\alpha}}\|_2 + \|\mathbb{P}_{X_s^{t,\xi,\alpha}}\|_2)^{2-\gamma_2} \right) ds \right] \\
&+ H' \mathbb{E} \left[\int_u^T (u-t)(1 + |X_s^{t,\xi,\alpha}|^2 + \|\mathbb{P}_{X_s^{t,\xi,\alpha}}\|_2^2) ds \right] \quad (\text{H\"older inequality}) \\
&\leq \int_u^T H' \left(\mathbb{E} \left[|X_s^{u,\xi,\alpha} - X_s^{t,\xi,\alpha}|^2 \right] \right)^{\frac{\gamma_1}{2}} \left(\mathbb{E} \left[(1 + |X_s^{u,\xi,\alpha}| + |X_s^{t,\xi,\alpha}|)^2 \right] \right)^{\frac{2-\gamma_1}{2}} ds \\
&+ H' \int_u^T \left(\mathbb{E} \left[|X_s^{u,\xi,\alpha} - X_s^{t,\xi,\alpha}|^2 \right] \right)^{\gamma_2} (1 + \mathbb{E}^{\frac{1}{2}}[|X_s^{u,\xi,\alpha}|^2] + \mathbb{E}^{\frac{1}{2}}[|X_s^{t,\xi,\alpha}|^2])^{2-\gamma_2} ds \\
&+ H'(u-t)(T-u) + H'(u-t) \int_u^T 2\mathbb{E} \left[|X_s^{t,\xi,\alpha}|^2 \right] ds \quad (\text{by (A.3) and (B.24)}) \\
&\leq H' \left(K_0(T)(1 + \|\xi\|_{L^2}^2)(u-t) \right)^{\frac{\gamma_1}{2}} (3 + 6h(T, M, \xi))^{\frac{2-\gamma_1}{2}} (T-u) \\
&+ H' \left(K_0(T)(1 + \|\xi\|_{L^2}^2)(u-t) \right)^{\gamma_2} (1 + 2h(T, M, \xi)^{\frac{1}{2}})^{2-\gamma_2} (T-u) \\
&+ H'(u-t)(T + 2h(T, M, \xi)) < \frac{\epsilon}{4} \tag{B.25}
\end{aligned}$$

for $u-t < \zeta_2(\epsilon, t, \xi, M, T, H')$ independent of α .

The thesis then follows plugging (B.22), (B.25) into (B.21) and then (B.21) and (B.20) into (B.17) with $u \in (t, t+\zeta)$ and $\zeta = \zeta_{\epsilon, t, \xi} = \min\{1, \zeta_1, \zeta_2\}$ independent of $\alpha \in \mathcal{A}$. $\square$

Remark B.3.1. Notice that the proof can be repeated with f and β satisfying $(\mathbf{H}'_{f,\beta})$ instead of $(\mathbf{H}_{f,\beta})$. In this case, we would have the following bounds. For the integral

I_2 ,

$$I_2 \leq C_f \int_t^u e^{-\beta(s-t)} ds \leq C_f(u-t) < \frac{\epsilon}{2} \quad (\text{B.26})$$

provided that $t \leq u \leq t + \frac{\epsilon}{2C_f}$. For the remainder R_T ,

$$R_T \leq 2C_f e^{\beta(1+t)} \int_T^{+\infty} e^{-\beta s} ds < \frac{\epsilon}{4} \quad (\text{B.27})$$

provided that $u - t < \min\{1, \frac{\epsilon}{2C_f}\}$ and $T = T(\epsilon)$ is sufficiently large. Furthermore, the function ψ is locally Hölder continuous in (x, μ) and Lipschitz in r uniformly with respect to a , i.e. there exists a constant $H' > 0$ such that

$$|\psi(r, x, \mu, a) - \psi(r', x', \mu', a)| \leq H' [|x - x'|^{\gamma_1} (1 + |x| + |x'|)^{2-\gamma_1} + W_2(\mu, \mu')^{\gamma_2} (1 + \|\mu\|_2 + \|\mu'\|_2)^{2-\gamma_2} + |r - r'|].$$

The bound for I_T thus becomes

$$\begin{aligned} I_T &= \mathbb{E} \left[\int_u^T e^{-\beta s} |\psi(u, X_s^{u, \xi, \alpha}, \mathbb{P}_{X_s^{u, \xi, \alpha}}, \alpha_s) - \psi(t, X_s^{t, \xi, \alpha}, \mathbb{P}_{X_s^{t, \xi, \alpha}}, \alpha_s)| ds \right] \\ &\leq \mathbb{E} \left[\int_u^T H' \left(|X_s^{u, \xi, \alpha} - X_s^{t, \xi, \alpha}|^{\gamma_1} (1 + |X_s^{u, \xi, \alpha}| + |X_s^{t, \xi, \alpha}|)^{2-\gamma_1} \right) ds \right] \\ &\quad + H' \mathbb{E} \left[\int_u^T \left(W_2(\mathbb{P}_{X_s^{u, \xi, \alpha}}, \mathbb{P}_{X_s^{t, \xi, \alpha}})^{\gamma_2} (1 + \|\mathbb{P}_{X_s^{u, \xi, \alpha}}\|_2 + \|\mathbb{P}_{X_s^{t, \xi, \alpha}}\|_2)^{2-\gamma_2} \right) ds \right] \\ &\quad + H' \mathbb{E} \left[\int_u^T (u - t) ds \right] \quad (\text{Hölder inequality and (A.3) and (B.24)}) \\ &\leq \int_u^T H' \left(K_0(T) (1 + \|\xi\|_{L^2}^2) (u - t) \right)^{\frac{\gamma_1}{2}} (3 + 6h(T, M, \xi))^{\frac{2-\gamma_1}{2}} ds \\ &\quad + \int_u^T H' \left(K_0(T) (1 + \|\xi\|_{L^2}^2) (u - t) \right)^{\gamma_2} (1 + 2h(T, M, \xi)^{\frac{1}{2}})^{2-\gamma_2} ds \\ &\quad + H' T (u - t) < \frac{\epsilon}{4} \end{aligned}$$

provided that $u - t < \zeta_3(\epsilon, T, M, \xi, t, H', K_0)$ independent of α . Putting the three bounds together, $|J(u, \xi, \alpha) - J(t, \xi, \alpha)| < \epsilon$ for $u \in (t, t + \zeta)$ and $\zeta = \zeta_{\epsilon, t, \xi} = \min\{\zeta_3, 1, \frac{\epsilon}{2C_f}\}$ independent of α . $\square$

We now prove by similar techniques that the functional J is also continuous with respect to the initial condition uniformly in α .

Proposition B.3.2. Assume that $(\mathbf{H}_{b, \sigma})$, $(\mathbf{H}_{f, \beta})$ (or $(\mathbf{H}'_{f, \beta})$) and $(\mathbf{H}_{f, cont})$ hold. Let $t \geq 0$. For all $\epsilon > 0$ and all $\xi \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$ there exists a $\delta = \delta(\epsilon, t, \mathbb{P}_\xi)$ such that, if $\|\xi - \eta\|_{L^2} < \delta$, then $|J(t, \xi, \alpha) - J(t, \eta, \alpha)| < \epsilon$.

Proof. Let $t \geq 0$ and consider $\xi, \eta \in L^2(\Omega, \mathcal{F}_t, \mathbb{P})$. In order to ease the notation, we will denote by X^ξ and X^η the processes $X^{t, \xi, \alpha}$ and $X^{t, \eta, \alpha}$, respectively. Fix $\epsilon > 0$. For all fixed $T > 0$, we can compute

$$|J(t, \xi, \alpha) - J(t, \eta, \alpha)| \leq \mathbb{E} \left[\int_t^{+\infty} e^{-\beta(s-t)} |f(X_s^\xi, \mathbb{P}_{X_s^\xi}, \alpha_s) - f(X_s^\eta, \mathbb{P}_{X_s^\eta}, \alpha_s)| ds \right]$$

$$= I_T + R_T, \quad (\text{B.28})$$

where

$$I_T := \mathbb{E} \left[\int_t^T e^{-\beta(s-t)} |f(X_s^\xi, \mathbb{P}_{X_s^\xi}, \alpha_s) - f(X_s^\eta, \mathbb{P}_{X_s^\eta}, \alpha_s)| ds \right]$$

and

$$R_T := \mathbb{E} \left[\int_T^{+\infty} e^{-\beta(s-t)} |f(X_s^\xi, \mathbb{P}_{X_s^\xi}, \alpha_s) - f(X_s^\eta, \mathbb{P}_{X_s^\eta}, \alpha_s)| ds \right].$$

Notice that, as we want to take η in a L^2 -neighbourhood of ξ , we can assume $\|\eta\|_{L^2}^2 - \|\xi\|_{L^2}^2 \leq 1$. First suppose that $(\mathbf{H}_{f,\beta})$ hold. Therefore,

$$\begin{aligned} R_T &\leq K \int_T^{+\infty} e^{-\beta(s-t)} \left(1 + \mathbb{E}[|X_s^\xi|^2] + \|\mathbb{P}_{X_s^\xi}\|_2^2 \right) ds \\ &\quad + K \int_T^{+\infty} e^{-\beta(s-t)} \left(1 + \mathbb{E}[|X_s^\eta|^2] + \|\mathbb{P}_{X_s^\eta}\|_2^2 \right) ds \\ &= K \int_T^{+\infty} e^{-\beta(s-t)} \left(1 + 2\mathbb{E}[|X_s^\xi|^2] \right) ds + K \int_T^{+\infty} e^{-\beta(s-t)} \left(1 + 2\mathbb{E}[|X_s^\eta|^2] \right) ds \quad (\text{by (1.4)}) \\ &\leq 2K e^{\beta t} \int_T^{+\infty} e^{-\beta s} ds + 2K \int_T^{+\infty} e^{(12M^2+1-\beta)(s-t)} \left(\|\xi\|_{L^2}^2 + 6M^2(s-t) \right) ds \\ &\quad + 2K \int_T^{+\infty} e^{(12M^2+1-\beta)(s-t)} \left(\|\eta\|_{L^2}^2 + 6M^2(s-t) \right) ds \quad (\text{as } \|\eta\|_{L^2}^2 \leq \|\xi\|_{L^2}^2 + 1) \\ &\leq 2K e^{\beta t} \int_T^{+\infty} e^{-\beta s} ds + 4K \int_T^{+\infty} e^{(12M^2+1-\beta)(s-t)} \left(\|\xi\|_{L^2}^2 + 6M^2(s-t) \right) ds \\ &\quad + 2K e^{-(12M^2+1-\beta)t} \int_T^{+\infty} e^{(12M^2+1-\beta)s} ds < \frac{\epsilon}{2} \end{aligned} \quad (\text{B.29})$$

provided that $T \geq T(\epsilon, t, \beta, M, K, \|\xi\|_{L^2})$ because all the integrals above are remainders of convergent integrals. As (B.28) holds for all fixed $T > 0$, we fix $T \geq T(\epsilon, t, \beta, M, K, \|\xi\|_2)$. Now study I_T : by $(\mathbf{H}_{f_{cont}})$,

$$\begin{aligned} I_T &= \mathbb{E} \left[\int_t^T e^{-\beta(s-t)} |f(X_s^\xi, \mathbb{P}_{X_s^\xi}, \alpha_s) - f(X_s^\eta, \mathbb{P}_{X_s^\eta}, \alpha_s)| ds \right] \\ &\leq H \int_t^T e^{-\beta(s-t)} \left(\mathbb{E}[|X_s^\xi - X_s^\eta|^{\gamma_1} (1 + |X_s^\xi| + |X_s^\eta|)^{2-\gamma_1}] \right. \\ &\quad \left. + W_2(\mathbb{P}_{X_s^\xi}, \mathbb{P}_{X_s^\eta})^{\gamma_2} (1 + \|\mathbb{P}_{X_s^\xi}\|_2 + \|\mathbb{P}_{X_s^\eta}\|_2)^{2-\gamma_2} ds \right) \quad (\text{by Hölder inequality}) \\ &\leq H \int_t^T e^{-\beta(s-t)} \left(\mathbb{E}[|X_s^\xi - X_s^\eta|^2] \right)^{\frac{\gamma_1}{2}} \left(\mathbb{E}[(1 + |X_s^\xi| + |X_s^\eta|)^2] \right)^{\frac{2-\gamma_1}{2}} \\ &\quad + W_2(\mathbb{P}_{X_s^\xi}, \mathbb{P}_{X_s^\eta})^{\gamma_2} (1 + \|\mathbb{P}_{X_s^\xi}\|_2 + \|\mathbb{P}_{X_s^\eta}\|_2)^{2-\gamma_2} ds \quad (\text{by (A.1)}) \\ &\leq H \int_t^T e^{-\beta(s-t)} (3C(T) \|\xi - \eta\|_{L^2}^2)^{\frac{\gamma_1}{2}} \left(\mathbb{E}[(1 + |X_s^\xi| + |X_s^\eta|)^2] \right)^{\frac{2-\gamma_1}{2}} \\ &\quad + (3C(T) \|\xi - \eta\|_{L^2}^2)^{\gamma_2} (1 + \|\mathbb{P}_{X_s^\xi}\|_2 + \|\mathbb{P}_{X_s^\eta}\|_2)^{2-\gamma_2} ds \end{aligned}$$

$$\begin{aligned}
&\leq H(3C(T))^{\frac{\gamma_1}{2}} \|\xi - \eta\|_{L^2}^{\gamma_1} \int_t^T e^{-\beta(s-t)} \left(3(1 + \mathbb{E}[|X_s^\xi|^2] + \mathbb{E}[|X_s^\eta|^2]) \right)^{\frac{2-\gamma_1}{2}} ds \\
&+ (3C(T)\|\xi - \eta\|_{L^2}^2)^{\gamma_2} \int_t^T e^{-\beta(s-t)} \left(3(1 + \mathbb{E}[|X_s^\xi|^2] + \mathbb{E}[|X_s^\eta|^2]) \right)^{2-\gamma_2} ds \quad (\text{by (B.24)}) \\
&\leq H(3C(T))^{\frac{\gamma_1}{2}} \|\xi - \eta\|_{L^2}^{\gamma_1} (6 + 6h(T, M, \|\xi\|_{L^2}))^{\frac{2-\gamma_1}{2}} (T - t) \\
&+ (3C(T)\|\xi - \eta\|_{L^2}^2)^{\gamma_2} (6 + 6h(T, M, \|\xi\|_{L^2}))^{2-\gamma_2} (T - t) < \frac{\epsilon}{2} \tag{B.30}
\end{aligned}$$

provided that $\|\xi - \eta\|_{L^2} < \min(\delta, 1)$ with $\delta = \delta(\epsilon, \|\xi\|_{L^2}, t, M, H, \gamma_1, \gamma_2)$ independent of $\alpha \in \mathcal{A}$, as $h(T, M, \xi) = \left(\|\xi\|_{L^2}^2 + 6M^2T \right) e^{(12M^2+1)T}$ is an increasing function in $\|\xi\|_{L^2}$. Notice that δ depends on ξ only through its L^2 -norm, i.e. only through its law. The thesis then follows pulling (B.30) and (B.29) back into (B.28).

On the other hand, if f satisfies $(\mathbf{H}'_{f,\beta})$, then (B.28) still holds; moreover, we can compute

$$\begin{aligned}
R_T &= \mathbb{E} \left[\int_T^{+\infty} e^{-\beta(s-t)} |f(X_s^\xi, \mathbb{P}_{X_s^\xi}, \alpha_s) - f(X_s^\eta, \mathbb{P}_{X_s^\eta}, \alpha_s)| ds \right] \\
&\leq 2C_f e^{\beta t} \int_T^{+\infty} e^{-\beta s} ds < \frac{\epsilon}{2}
\end{aligned}$$

for $T \geq T(\epsilon, C_f, \beta, t)$ because this is the remainder of a convergent integral, while the computations in (B.30) can be repeated exactly in this case, leading to the same conclusion. $\square$

Appendix C

Ergodic HJB equations: proofs

C.1 Proof of Theorem 2.5.5

In this Appendix, we briefly complete the proof of Theorem 2.5.5. As already mentioned, if we call v^T the value function of an optimal control problem with finite time horizon T defined as in (2.43) and terminal gain $g = \phi$, then, by Proposition 2.4.4, we have

$$v^T(t, \mu) = \phi(\mu) + \lambda(T - t). \quad (\text{C.1})$$

Recall that, by Theorem 2.5.4, v^T is a viscosity solution to the backward parabolic Hamilton-Jacobi-Bellman equation

$$\begin{cases} -\partial_t v^T(t, \mu) - F(\mu, \partial_\mu v^T(t, \mu), \partial_x \partial_\mu v^T(t, \mu)) = 0, & t \in [0, T], \mu \in \mathcal{P}_2(\mathbb{R}^d), \\ v^T(T, \mu) = \phi(\mu) \end{cases} \quad (\text{C.2})$$

with

$$\begin{aligned} & F(\mu, \partial_\mu \varphi(\mu), \partial_x \partial_\mu \varphi(\mu)) \\ &= \int_{\mathbb{R}^d} \left[\sup_{a \in A} \left\{ f(x, \mu, a) + \langle b(x, \mu, a), \partial_\mu \varphi(\mu)(x) \rangle + \frac{1}{2} \text{tr}(\sigma \sigma^T(x, \mu, a) \partial_x \partial_\mu \varphi(\mu)(x)) \right\} \right] \mu(dx). \end{aligned}$$

There are two possible cases:

1. We assume in addition that v^T is sufficiently regular (it belongs to $\tilde{C}^{1,2}([0, T] \times \mathcal{P}_2(\mathbb{R}^d))$) and it thus is a classical solution to (C.2). Then, by (C.1), we can compute explicitly its derivatives:

$$\begin{aligned} \partial_t v^T(t, \mu) &= -\lambda \\ \partial_\mu v^T(t, \mu) &= \partial_\mu \phi(\mu) \\ \partial_x \partial_\mu v^T(t, \mu) &= \partial_x \partial_\mu \phi(\mu). \end{aligned}$$

Substituting in (C.2), we obtain

$$\lambda = F(\mu, \partial_\mu \phi(\mu), \partial_x \partial_\mu \phi(\mu)),$$

from which we conclude that (λ, ϕ) is a classical solution to the ergodic Hamilton-Jacobi-Bellman equation (2.53).

2. We now extend this result to the case where v^T is a viscosity solution to (C.2). We will only prove that ϕ is a viscosity subsolution to (2.53), as its supersolution property is completely analogous.

Let $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $\psi \in \tilde{C}^2(\mathcal{P}_2(\mathbb{R}^d))$ such that $\phi - \psi$ has a maximum with value 0 at μ . Then,

$$0 = \phi(\mu) - \psi(\mu) \geq \phi(\nu) - \psi(\nu)$$

for any $\nu \in \mathcal{P}_2(\mathbb{R}^d)$. Fix any $T > 0$ and $t \in [0, T]$; then, adding and subtracting the value $\lambda(T - t)$ to the previous equality, we deduce that

$$\lambda(T - t) = \phi(\mu) - \psi(\mu) + \lambda(T - t) \geq \phi(\nu) - \psi(\nu) + \lambda(T - t)$$

for all $\nu \in \mathcal{P}_2(\mathbb{R}^d)$. Then, by (C.1),

$$\lambda(T - t) = v^T(t, \mu) - \psi(\mu) \geq v^T(t, \nu) - \psi(\nu) \quad (\text{C.3})$$

for all $\nu \in \mathcal{P}_2(\mathbb{R}^d)$. Consider the function

$$\psi^T(t, \mu) = \psi(\mu) + \lambda(T - t),$$

which belongs to the class $\tilde{C}^{1,2}([0, T] \times \mathcal{P}_2(\mathbb{R}^d))$ by the regularity of ψ and because

$$|\partial_t \psi^T(t, \mu)| = |\lambda| \leq C := \lambda + C_\psi.$$

By (C.3), we deduce that

$$\begin{aligned} 0 &= v^T(t, \mu) - \psi^T(t, \mu) \geq v^T(t, \nu) - \psi^T(t, \nu) \\ &= \phi(\nu) + \lambda(T - t) - \psi(\mu) - \lambda(T - t) \\ &= \phi(\nu) - \psi(\mu) \\ &= \phi(\nu) + \lambda(T - s) - \psi(\mu) - \lambda(T - s) \\ &= v^T(s, \nu) - \psi^T(s, \nu) \end{aligned}$$

for all $\nu \in \mathcal{P}_2(\mathbb{R}^d)$ and $s \in [0, T]$. Then, $v^T - \psi^T$ has a maximum with value 0 at (t, μ) and ψ^T is a test function for the subsolution property of v^T . As v^T is a viscosity solution to (C.2), we deduce that

$$-\partial_t \psi^T(t, \mu) - F(\mu, \partial_\mu \psi^T(t, \mu), \partial_x \partial_\mu \psi^T(t, \mu)) \leq 0. \quad (\text{C.4})$$

Computing explicitly the derivatives of the test function ψ^T , we obtain:

$$\begin{aligned} \partial_t \psi^T(t, \mu) &= -\lambda; \\ \partial_\mu \psi^T(t, \mu) &= \partial_\mu \psi(\mu); \\ \partial_x \partial_\mu \psi^T(t, \mu) &= \partial_x \partial_\mu \psi(\mu). \end{aligned}$$

Substituting these expressions in (C.4), we deduce that

$$\lambda - F(\mu, \partial_\mu \psi(\mu), \partial_x \partial_\mu \psi(\mu)) \leq 0.$$

By the arbitrariness of $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and of ψ in the class of test functions for the subsolution property of the ergodic HJB equation, we conclude that (λ, ϕ) is a viscosity subsolution to equation (2.53) and the proof is concluded.

C.2 Proof of Proposition 2.5.6

In this section, we prove Proposition 2.5.6.

Proof of Proposition 2.5.6. 1) Fix any $T > 0$. Recall that, by Theorem 5.5 in [42], the value function

$$v^T(t, \mu) := \sup_{\alpha \in \mathcal{A}} \mathbb{E} \left[\int_t^T f(X_s^{t, \xi, \alpha}, \mathbb{P}_{X_s^{t, \xi, \alpha}}, \alpha_s) ds + g(X_T^{t, \xi, \alpha}, \mathbb{P}_{X_T^{t, \xi, \alpha}}) \right]$$

of the finite horizon optimal control problem (2.43) with time horizon T is a viscosity solution to the backward parabolic HJB equation

$$\begin{cases} -\partial_t v^T(t, \mu) - F(\mu, \partial_\mu v^T(t, \mu), \partial_x \partial_\mu v^T(t, \mu)) = 0, & t \in [0, T], \mu \in \mathcal{P}_2(\mathbb{R}^d), \\ v^T(T, \mu) = \int_{\mathbb{R}^d} g(x, \mu) \mu(dx), \end{cases} \quad (\text{C.5})$$

where

$$\begin{aligned} & F(\mu, \partial_\mu \varphi(\mu), \partial_x \partial_\mu \varphi(\mu)) \\ &= \int_{\mathbb{R}^d} \left[\sup_{a \in A} \left\{ f(x, \mu, a) + \langle b(x, \mu, a), \partial_\mu \varphi(\mu)(x) \rangle + \frac{1}{2} \text{tr}(\sigma \sigma^T(x, \mu, a) \partial_x \partial_\mu \varphi(\mu)(x)) \right\} \right] \mu(dx). \end{aligned}$$

For any $T > 0$, define the function $w^T : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ by the equality

$$w^T(T - t, \mu) = v^T(t, \mu), \quad \mu \in \mathcal{P}_2(\mathbb{R}^d), t \in [0, T]. \quad (\text{C.6})$$

Now set $w(t, \mu) := w^T(t, \mu)$ for all $0 \leq t \leq T$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Then, w is well-defined on $[0, +\infty) \times \mathcal{P}_2(\mathbb{R}^d)$. As a matter of fact, take any pair of times $0 \leq T_1 < T_2$ and choose $t \in [0, T_1]$. We need to prove that $w(t, \mu)$ is uniquely defined, i.e. that

$$v^{T_1}(T_1 - t, \mu) = v^{T_2}(T_2 - t, \mu) \quad \text{for any } t \in [0, T_1]. \quad (\text{C.7})$$

Consider, for any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $\xi \in L^2(\Omega, \mathcal{G}, \mathbb{P})$ with $\mathbb{P}_\xi = \mu$, any $u \in [0, T_2]$, and any $\alpha \in \mathcal{A}$, the gain functional

$$J^{T_2}(u, \xi, \alpha) = \mathbb{E} \left[\int_u^{T_2} f(X_s^{u, \xi, \alpha}, \mathbb{P}_{X_s^{u, \xi, \alpha}}, \alpha_s) ds + g(X_{T_2}^{u, \xi, \alpha}, \mathbb{P}_{X_{T_2}^{u, \xi, \alpha}}) \right].$$

The functional J^{T_1} is defined analogously for initial times $r \in [0, T_1]$. Then, repeating the proof of Theorem 1.3.8 on a bounded time interval (as the gain functional J still is right-continuous in time uniformly with respect to α under the hypotheses of this Proposition, see Proposition 1.3.13), we have

$$\begin{aligned} v^{T_2}(T_2 - t, \mu) &= \sup_{\alpha \in \mathcal{A}} J(T_2 - t, \xi, \alpha) \\ &= \sup_{\alpha \in \mathcal{A}} \mathbb{E} \left[\int_{T_2 - t}^{T_2} f(X_s^{T_2 - t, \xi, \alpha}, \mathbb{P}_{X_s^{T_2 - t, \xi, \alpha}}, \alpha_s) ds + g(X_{T_2}^{T_2 - t, \xi, \alpha}, \mathbb{P}_{X_{T_2}^{T_2 - t, \xi, \alpha}}) \right]. \end{aligned}$$

If $\alpha \in \mathcal{A}^{T_2-t}$, by an extension of Lemma 1.2.3 on any interval $[s, +\infty)$ for $s \geq 0$, it is indistinguishable from

$$(\underline{\alpha}(s, (B^{(T_2-t)})_{\cdot \wedge s}, U))_{s \geq T_2-t}$$

for a suitable measurable function $\underline{\alpha}$. Call $h := T_2 - T_1$. Construct $\gamma \in \mathcal{A}^{T_1-t}$ such that $\gamma_s := \underline{\alpha}(s + h, (B^{(T_1-t)})_{\cdot \wedge s}, U)$ for all $s \geq T_1 - t$. Then, it can be easily verified that $\gamma \in \mathcal{A}^{T_1-t}$ and

$$\mathcal{L}\left(\xi, \gamma, B^{(T_1-t)}\right) = \mathcal{L}\left(\xi, \alpha_{\cdot+h}, B_{\cdot+h}^{(T_2-t)}\right) \quad \text{on } [T_1 - t, T_1].$$

Using the same “shifting” techniques as in the proof of Lemma B.1.5 with initial time $T_1 - t > 0$, this implies that

$$\mathcal{L}\left(X_{\cdot+h}^{T_2-t, \xi, \alpha}, \alpha_{\cdot+h}\right) = \mathcal{L}\left(X^{T_1-t, \xi, \gamma}, \gamma\right) \quad \text{on } [T_1 - t, T_1]. \quad (\text{C.8})$$

As a consequence,

$$\begin{aligned} & J^{T_2}(T_2 - t, \xi, \alpha) \\ &= \mathbb{E} \left[\int_{T_2-t}^{T_2} f(X_s^{T_2-t, \xi, \alpha}, \mathbb{P}_{X_s^{T_2-t, \xi, \alpha}}, \alpha_s) ds + g(X_{T_2}^{T_2-t, \xi, \alpha}, \mathbb{P}_{X_{T_2}^{T_2-t, \xi, \alpha}}) \right] \quad (z=s-h) \\ &= \mathbb{E} \left[\int_{T_1-t}^{T_1} f(X_{z+h}^{T_2-t, \xi, \alpha}, \mathbb{P}_{X_{z+h}^{T_2-t, \xi, \alpha}}, \alpha_{z+h}) dz + g(X_{T_2}^{T_2-t, \xi, \alpha}, \mathbb{P}_{X_{T_2}^{T_2-t, \xi, \alpha}}) \right] \quad (\text{by (C.8) with } T_2 = T_1 + h) \\ &= \mathbb{E} \left[\int_{T_1-t}^{T_1} f(X_z^{T_1-t, \xi, \gamma}, \mathbb{P}_{X_z^{T_1-t, \xi, \gamma}}, \gamma_z) dz + g(X_{T_1}^{T_1-t, \xi, \gamma}, \mathbb{P}_{X_{T_1}^{T_1-t, \xi, \gamma}}) \right] \\ &= J^{T_1}(T_1 - t, \xi, \gamma) \\ &\leq v^{T_1}(T_1 - t, \xi). \end{aligned}$$

Taking the supremum over $\mathcal{A}^{T_2-t}$, we conclude that

$$v^{T_2}(T_2 - t, \mu) \leq v^{T_1}(T_1 - t, \mu).$$

With analogous computations, shifting forward instead of backward in time, we can prove the opposite inequality, leading to the equality

$$v^{T_2}(T_2 - t, \mu) = v^{T_1}(T_1 - t, \mu) \quad \text{for all } \mu \in \mathcal{P}_2(\mathbb{R}^d), t \in [0, T_1].$$

The function w is thus well defined on the entire interval $[0, +\infty)$.

As explained at the beginning of the proof, v^T is a viscosity solution to the parabolic backward Hamilton-Jacobi-Bellman equation (C.5). We want to prove that this implies that w is a viscosity solution to the parabolic forward HJB equation

$$\begin{cases} \partial_t w(t, \mu) - F(\mu, \partial_\mu w(t, \mu), \partial_x \partial_\mu w(t, \mu)) = 0, & t \geq 0, \mu \in \mathcal{P}_2(\mathbb{R}^d), \\ w(0, \mu) = \int_{\mathbb{R}^d} g(x, \mu) \mu(dx). \end{cases} \quad (\text{C.9})$$

First of all, notice that, by definition of w , $w(0, \mu) = v^T(T, \mu) = \int_{\mathbb{R}^d} g(x, \mu) \mu(dx)$ for all $T \geq 0$: the boundary condition is thus satisfied. We prove now the subsolution

property of w . Take any $T > 0$, $(t, \mu) \in (0, T) \times \mathcal{P}_2(\mathbb{R}^d)$ and any $\psi \in \tilde{C}^{1,2}([0, T] \times \mathcal{P}_2(\mathbb{R}^d))$ such that $w - \psi$ has a maximum over $[0, T]$ with value 0 at (t, μ) . Then,

$$0 = w(t, \mu) - \psi(t, \mu) \geq w(s, \nu) - \psi(s, \nu) \quad \text{for all } s \in [0, T], \nu \in \mathcal{P}_2(\mathbb{R}^d).$$

By definition of w , we can rewrite this inequality as

$$0 = v^T(T - t, \mu) - \psi(t, \mu) \geq v^T(T - s, \mu) - \psi(s, \nu) \quad \text{for all } s \in [0, T], \nu \in \mathcal{P}_2(\mathbb{R}^d). \quad (\text{C.10})$$

Define the function $\chi^T : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ such that $\chi^T(s, \nu) := \psi(T - s, \nu)$ for all $s \in [0, T], \nu \in \mathcal{P}_2(\mathbb{R}^d)$. Then clearly $\chi^T \in \tilde{C}^{1,2}([0, T] \times \mathcal{P}_2(\mathbb{R}^d))$ by the regularity of ψ because

$$\begin{aligned} \partial_t \chi^T(s, \nu) &= -\partial_t \psi(T - s, \nu); \\ \partial_\mu \chi^T(s, \nu) &= \partial_\mu \psi(T - s, \nu); \\ \partial_x \partial_\mu \chi^T(s, \nu) &= \partial_x \partial_\mu \psi(T - s, \nu). \end{aligned} \quad (\text{C.11})$$

Therefore, inequality (C.10) reads as

$$0 = v^T(T - t, \mu) - \chi^T(T - t, \mu) \geq v^T(T - s, \mu) - \chi^T(T - s, \nu) \quad \text{for all } s \in [0, T], \nu \in \mathcal{P}_2(\mathbb{R}^d).$$

Then, the function $v^T - \chi^T$ has a maximum with value 0 at $(T - t, \mu)$ over $[0, T] \times \mathcal{P}_2(\mathbb{R}^d)$. As v^T is a viscosity (sub)solution to (C.5), we have

$$-\partial_t \chi^T(T - t, \mu) - F(\mu, \partial_\mu \chi^T(T - t, \mu), \partial_x \partial_\mu \chi^T(T - t, \mu)) \leq 0.$$

Using the explicit expression of the derivatives of χ^T (C.11), this implies

$$\partial_t \psi(t, \mu) - F(\mu, \partial_\mu \psi(t, \mu), \partial_x \partial_\mu \psi(t, \mu)) \leq 0.$$

By the arbitrariness of (t, μ) in $(0, T) \times \mathcal{P}_2(\mathbb{R}^d)$ and ψ in $\tilde{C}^{1,2}([0, T] \times \mathcal{P}_2(\mathbb{R}^d))$, we conclude that w is a viscosity subsolution to the parabolic forward HJB equation (C.9). The supersolution property is completely analogous. Then, w solves (C.9) in the viscosity sense.

Moreover, by definition, for any $t \in [0, +\infty)$,

$$|w(t, \mu) - \phi(\mu) - \lambda t| = |v^t(0, \mu) - \phi(\mu) - \lambda t| \leq C (1 + W_2(\mu, \delta_0)^2), \quad \mu \in \mathcal{P}_2(\mathbb{R}^d)$$

by Theorem 2.4.5, where C is the same constant that appears in Theorem 2.4.5. We conclude that

$$\lim_{t \rightarrow \infty} \frac{w(t, \mu)}{t} = \lambda.$$

2) We want to prove uniqueness of viscosity solutions to (C.9) under the additional assumptions. As g is also assumed to be $(|\cdot| \times W_2)$ -Lipschitz, by **(H₃)** and Proposition 2.3 in [43], for any fixed $T > 0$ the value function v^T is W_2 -Lipschitz in μ uniformly with respect to t , i.e. there exists a constant $C := C_T$ (which may depend on the time horizon $T > 0$) such that

$$|v^T(t, \mu) - v^T(t, \mu')| \leq C W_2(\mu, \mu') \quad \text{for all } \mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d).$$

As **(CP)** holds for every fixed $T > 0$, we conclude that v^T is the unique viscosity solution to the backward parabolic HJB equation (C.5) on the interval $[0, T]$ in the class of functions which are Lipschitz on $\mathcal{P}_2(\mathbb{R}^d)$ uniformly with respect to t .

Now assume that there exists another viscosity solution w_2 to (C.9) which is Lipschitz on $\mathcal{P}_2(\mathbb{R}^d)$ uniformly with respect to t on every bounded interval $[0, T]$. Fix any $T > 0$ and define $u(t, \mu) = w_2(T-t, \mu)$ for any $t \in [0, T]$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Then, u is Lipschitz on $\mathcal{P}_2(\mathbb{R}^d)$ uniformly with respect to t on $[0, T]$. Moreover, $u(T, \mu) = w_2(0, \mu) = \int_{\mathbb{R}^d} g(x, \mu) \mu(dx)$. Now fix any $(t, \mu) \in (0, T) \times \mathcal{P}_2(\mathbb{R}^d)$ and any $\psi \in \tilde{C}^{1,2}([0, T] \times \mathcal{P}_2(\mathbb{R}^d))$ such that $u - \psi$ has a maximum with value 0 at (t, μ) . Then,

$$0 = u(t, \mu) - \psi(t, \mu) \geq u(s, \nu) - \psi(s, \nu) \quad \text{for all } s \in [0, T], \nu \in \mathcal{P}_2(\mathbb{R}^d).$$

Defining a function $\chi^T : [0, T] \times \mathcal{P}_2(\mathbb{R}^d) \rightarrow \mathbb{R}$ as $\chi^T(s, \nu) := \psi(T-s, \nu)$, which clearly belongs to $\tilde{C}^{1,2}([0, T] \times \mathcal{P}_2(\mathbb{R}^d))$, we rewrite this inequality as

$$0 = w_2(T-t, \mu) - \chi^T(T-t, \mu) \geq w_2(T-s, \nu) - \chi^T(T-s, \nu) \quad \text{for all } s \in [0, T], \nu \in \mathcal{P}_2(\mathbb{R}^d).$$

Then, $w_2 - \chi^T$ admits a maximum over $[0, T]$ with value 0 at $(T-t, \mu)$; as w_2 is a viscosity subsolution to (C.9), we deduce that

$$\partial_t \chi^T(T-t, \mu) - F(\mu, \partial_\mu \chi^T(T-t, \mu), \partial_x \partial_\mu \chi^T(T-t, \mu)) \leq 0;$$

using the definition of χ^T , this inequality corresponds to

$$-\partial_t \psi(t, \mu) - F(\mu, \partial_\mu \psi(t, \mu), \partial_x \partial_\mu \psi(t, \mu)) \leq 0.$$

We conclude that u is a subsolution to the backward parabolic HJB (C.5). The supersolution property is completely analogous; thus, u is a viscosity solution to (C.5). As a consequence, $u(t, \mu) = v^T(t, \mu)$ for all $t \in [0, T]$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, which implies $w_2(t, \mu) = u(T-t, \mu) = v^T(T-t, \mu) = w(t, \mu)$ for all $t \in [0, T]$ by the construction of w . Then, (C.9) admits a unique solution corresponding to w .

□

Bibliography

- [1] L. Ambrosio, E. Brué, and D. Semola. *Lectures on optimal transport*, volume 130 of *Unitext. La Matematica per il 3+2*. Springer, Cham, 2021.
- [2] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2005.
- [3] D. Andersson and B. Djehiche. A maximum principle for SDEs of mean-field type. *Appl. Math. Optim.*, 63(3):341–356, 2011.
- [4] A. Arapostathis, V. S. Borkar, and M. K. Ghosh. *Ergodic control of diffusion processes*, volume 143 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2012.
- [5] M. Arisawa. Ergodic problem for the Hamilton-Jacobi-Bellman equation. II. *Ann. Inst. H. Poincaré C Anal. Non Linéaire*, 15(1):1–24, 1998.
- [6] M. Arisawa and P.-L. Lions. On ergodic stochastic control. *Comm. Partial Differential Equations*, 23(11-12):2187–2217, 1998.
- [7] J.-P. Aubin. *Viability theory*. Systems & Control: Foundations & Applications. Birkhäuser Boston, Inc., Boston, MA, 1991.
- [8] A. Bain and D. Crisan. *Fundamentals of stochastic filtering*, volume 60 of *Stochastic Modelling and Applied Probability*. Springer, New York, 2009.
- [9] P. Baldi. *Stochastic calculus. An introduction through theory and exercises*. Universitext. Springer, Cham, 2017.
- [10] E. Bandini, A. Cosso, M. Fuhrman, and H. Pham. Randomized filtering and Bellman equation in Wasserstein space for partial observation control problem. *Stochastic Process. Appl.*, 129(2):674–711, 2019.
- [11] X. Bao and S. Tang. Ergodic control of McKean-Vlasov SDEs and associated Bellman equation. *J. Math. Anal. Appl.*, 527(1):Paper No. 127404, 2023.
- [12] M. Bardi and I. Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Systems & Control: Foundations & Applications. Birkhäuser Boston, Inc., Boston, MA, 1997.
- [13] E. Bayraktar, A. Cecchin, and P. Chakraborty. Mean field control and finite agent approximation for regime-switching jump diffusions. *Appl. Math. Optim.*, 88(2):Paper No. 36, 2023.

- [14] E. Bayraktar, H. Cheung, I. Ekren, J. Qiu, H. M. Tai, and X. Zhang. Viscosity solutions of fully second-order HJB equations in the Wasserstein space. *arXiv preprint arXiv:2501.01612*, 2025.
- [15] E. Bayraktar, A. Cosso, and H. Pham. Randomized dynamic programming principle and Feynman-Kac representation for optimal control of McKean-Vlasov dynamics. *Trans. Amer. Math. Soc.*, 370(3):2115–2160, 2018.
- [16] E. Bayraktar, I. Ekren, X. He, and X. Zhang. Comparison for semi-continuous viscosity solutions for second order PDEs on the Wasserstein space. *Journal of Differential Equations*, 455:Paper No. 113963, 2026.
- [17] E. Bayraktar, I. Ekren, and X. Zhang. Comparison of viscosity solutions for a class of second-order PDEs on the Wasserstein space. *Communications in Partial Differential Equations*, 50(4):570–613, 2025.
- [18] E. Bayraktar and A. D. Kara. Infinite horizon average cost optimality criteria for mean-field control. *SIAM J. Control Optim.*, 62(5):2776–2806, 2024.
- [19] E. Bayraktar and X. Zhang. Solvability of infinite horizon McKean-Vlasov FBSDEs in mean field control problems and games. *Appl. Math. Optim.*, 87(1):Paper No. 13, 2023.
- [20] A. Bensoussan. Équations paraboliques intervenant en contrôle optimal ergodique. *Mat. Apl. Comput.*, 6(3):211–255, 1987.
- [21] A. Bensoussan. *Stochastic control of partially observable systems*. Cambridge University Press, Cambridge, 1992.
- [22] D. P. Bertsekas and S. E. Shreve. *Stochastic optimal control. The discrete time case*, volume 139 of *Mathematics in Science and Engineering*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1978.
- [23] F. Bolley. Separability and completeness for the Wasserstein distance. In *Séminaire de probabilités XLI*, volume 1934 of *Lecture Notes in Math.*, pages 371–377. Springer, Berlin, 2008.
- [24] R. Buckdahn, B. Djehiche, and J. Li. A general stochastic maximum principle for SDEs of mean-field type. *Appl. Math. Optim.*, 64(2):197–216, 2011.
- [25] R. Buckdahn, D. Goreac, and M. Quincampoix. Existence of asymptotic values for nonexpansive stochastic control systems. *Appl. Math. Optim.*, 70(1):1–28, 2014.
- [26] R. Buckdahn, J. Li, S. Peng, and C. Rainer. Mean-field stochastic differential equations and associated PDEs. *Ann. Probab.*, 45(2):824–878, 2017.
- [27] R. Buckdahn, J. Li, M. Quincampoix, and J. Renault. Representation formulas for limit values of long run stochastic optimal controls. *SIAM J. Control Optim.*, 58(4):1846–1873, 2020.
- [28] O. A. Butkovsky. On ergodic properties of nonlinear Markov chains and stochastic McKean-Vlasov equations. *Theory Probab. Appl.*, 58(4):661–674, 2014.
- [29] F. Cannerozzi and G. Ferrari. Cooperation, correlation and competition in ergodic n -player games and mean-field games of singular controls: A case study. *arXiv preprint arXiv:2404.15079*, 2024.

- [30] P. Cardaliaguet. Notes on mean field games (from P.-L. Lions' lectures at Collège de France). <https://www.ceremade.dauphine.fr/~cardaliaguet/MFG20130420.pdf>, September 2013.
- [31] P. Cardaliaguet and A. Porretta. Long time behavior of the master equation in mean field game theory. *Anal. PDE*, 12(6):1397–1453, 2019.
- [32] R. Carmona and F. Delarue. *Probabilistic theory of mean field games with applications. I. Mean field FBSDEs, control, and games*, volume 83 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2018.
- [33] R. Carmona and F. Delarue. *Probabilistic theory of mean field games with applications. II. Mean field games with common noise and master equations*, volume 84 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2018.
- [34] R. Carmona, F. Delarue, and A. Lachapelle. Control of McKean-Vlasov dynamics versus mean field games. *Math. Financ. Econ.*, 7(2):131–166, 2013.
- [35] R. Carmona, L. Tangpi, and K. Zhang. A probabilistic approach to discounted infinite horizon and invariant mean field games. *arXiv preprint arXiv:2407.03642*, 2024.
- [36] H. Cheung, H. M. Tai, and J. Qiu. Viscosity solutions of a class of second order Hamilton–Jacobi–Bellman equations in the Wasserstein space. *Applied Mathematics & Optimization*, 91(1):23, 2025.
- [37] K. L. Chung and R. J. Williams. *Introduction to stochastic integration*. Probability and its Applications. Birkhäuser Boston, Inc., Boston, MA, second edition, 1990.
- [38] J. Claisse, D. Talay, and X. Tan. A pseudo-Markov property for controlled diffusion processes. *SIAM J. Control Optim.*, 54(2):1017–1029, 2016.
- [39] F. H. Clarke, Y. S. Ledyaev, R. J. Stern, and P. R. Wolenski. *Nonsmooth analysis and control theory*, volume 178 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1998.
- [40] S. N. Cohen and R. J. Elliott. *Stochastic calculus and applications*. Probability and its Applications. Springer, Cham, second edition, 2015.
- [41] A. Cosso, M. Fuhrman, and H. Pham. Long time asymptotics for fully nonlinear Bellman equations: a backward SDE approach. *Stochastic Process. Appl.*, 126(7):1932–1973, 2016.
- [42] A. Cosso, F. Gozzi, I. Kharroubi, H. Pham, and M. Rosestolato. Optimal control of path-dependent McKean-Vlasov SDEs in infinite-dimension. *Ann. Appl. Probab.*, 33(4):2863–2918, 2023.
- [43] A. Cosso, F. Gozzi, I. Kharroubi, H. Pham, and M. Rosestolato. Master Bellman equation in the Wasserstein space: uniqueness of viscosity solutions. *Trans. Amer. Math. Soc.*, 377(1):31–83, 2024.
- [44] A. Cosso, G. Guatteri, and G. Tessitore. Ergodic control of infinite-dimensional stochastic differential equations with degenerate noise. *ESAIM Control Optim. Calc. Var.*, 25:Paper No. 12, 2019.

- [45] M. G. Crandall, H. Ishii, and P.-L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, 27(1):1–67, 1992.
- [46] S. Daudin and B. Seeger. A comparison principle for semilinear Hamilton-Jacobi-Bellman equations in the Wasserstein space. *Calc. Var. Partial Differential Equations*, 63(4):Paper No. 106, 2024.
- [47] De Crescenzo, Anna, Fuhrman, Marco, Kharroubi, Idris, and Pham, Huy  n. Mean-field control of non exchangeable systems. *ESAIM: COCV*, 32:3, 2026.
- [48] A. Debussche, Y. Hu, and G. Tessitore. Ergodic BSDEs under weak dissipative assumptions. *Stochastic Process. Appl.*, 121(3):407–426, 2011.
- [49] M. F. Djete, D. Possama  , and X. Tan. McKean-Vlasov optimal control: the dynamic programming principle. *Ann. Probab.*, 50(2):791–833, 2022.
- [50] R. J. Elliott, L. Aggoun, and J. B. Moore. *Hidden Markov models. Estimation and control*, volume 29 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1995.
- [51] G. Fabbri, F. Gozzi, and A.   wiech. *Stochastic optimal control in infinite dimension. Dynamic programming and HJB equations.*, volume 82 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2017.
- [52] R. Frey, A. Gabih, and R. Wunderlich. Portfolio optimization under partial information with expert opinions. *International Journal of Theoretical and Applied Finance*, 15(01):1250009, 2012.
- [53] R. Frey, A. Gabih, and R. Wunderlich. Portfolio optimization under partial information with expert opinions: a dynamic programming approach. *Commun. Stoch. Anal.*, 8(1):49–79, 2014.
- [54] M. Fuhrman, Y. Hu, and G. Tessitore. Ergodic BSDEs and optimal ergodic control in Banach spaces. *SIAM J. Control Optim.*, 48(3):1542–1566, 2009.
- [55] M. Fuhrman, H. Pham, and S. Rud  . Optimal control of McKean-Vlasov systems under partial observation and hidden Markov switching. *arXiv preprint arXiv:2601.09311*, 2026.
- [56] M. Fuhrman and S. Rud  . Ergodic control of McKean–Vlasov systems on the Wasserstein space. *SIAM Journal on Control and Optimization*, 63(6):4018–4043, 2025.
- [57] S. W. He, J. G. Wang, and J. A. Yan. *Semimartingale theory and stochastic calculus*. Kexue Chubanshe (Science Press), Beijing; CRC Press, Boca Raton, FL, 1992.
- [58] M. Huang, R. P. Malham  , and P. E. Caines. Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Commun. Inf. Syst.*, 6(3):221–251, 2006.
- [59] N. Ichihara. Large time asymptotic problems for optimal stochastic control with super-linear cost. *Stochastic Process. Appl.*, 122(4):1248–1275, 2012.
- [60] N. Ichihara and S.-J. Sheu. Large time behavior of solutions of Hamilton-Jacobi-Bellman equations with quadratic nonlinearity in gradients. *SIAM J. Math. Anal.*, 45(1):279–306, 2013.

- [61] M. Kac. Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III*, pages 171–197. Univ. California Press, Berkeley-Los Angeles, Calif., 1956.
- [62] O. Kallenberg. *Foundations of modern probability*, volume 99 of *Probability Theory and Stochastic Modelling*. Springer, Cham, third edition, 2021.
- [63] D. Lacker. Limit theory for controlled McKean-Vlasov dynamics. *SIAM Journal on Control and Optimization*, 55(3):1641–1672, 2017.
- [64] J.-M. Lasry and P.-L. Lions. Mean field games. *Jpn. J. Math.*, 2(1):229–260, 2007.
- [65] J.-F. Le Gall. *Brownian motion, martingales, and stochastic calculus*, volume 274 of *Graduate Texts in Mathematics*. Springer, Cham, 2016.
- [66] J. Li. Stochastic maximum principle in the mean-field controls. *Automatica J. IFAC*, 48(2):366–373, 2012.
- [67] P.-L. Lions. Viscosity solutions of fully nonlinear second-order equations and optimal stochastic control in infinite dimensions. I. The case of bounded stochastic evolutions. *Acta Math.*, 161(3-4):243–278, 1988.
- [68] P. L. Lions. Viscosity solutions of fully nonlinear second-order equations and optimal stochastic control in infinite dimensions. II: Optimal control of Zakai’s equation. In *Stochastic Partial Differential Equations and Applications II*, pages 147–170. Springer Berlin Heidelberg, 1989.
- [69] P.-L. Lions. Viscosity solutions of fully nonlinear second-order equations and optimal stochastic control in infinite dimensions. III. Uniqueness of viscosity solutions for general second-order equations. *J. Funct. Anal.*, 86(1):1–18, 1989.
- [70] P.-L. Lions. Théorie des jeux de champ moyen et applications. <https://www.college-de-france.fr>, 2006-2012. Cours au Collège de France.
- [71] F. Malrieu. Convergence to equilibrium for granular media equations and their Euler schemes. *Ann. Appl. Probab.*, 13(2):540–560, 2003.
- [72] H. P. McKean, Jr. A class of Markov processes associated with nonlinear parabolic equations. *Proc. Nat. Acad. Sci. U.S.A.*, 56:1907–1911, 1966.
- [73] E. J. McShane. Extension of range of functions. *Bull. Amer. Math. Soc.*, 40(12):837–842, 1934.
- [74] B. G. Pachpatte. On some generalizations of Bellman’s lemma. *J. Math. Anal. Appl.*, 51:141–150, 1975.
- [75] E. Pardoux. Filtrage non lineaire et equations aux derivees partielles stochastiques associees. In *Ecole d’Eté de Probabilités de Saint-Flour XIX — 1989*, pages 68–163, Berlin, Heidelberg, 1991. Springer Berlin Heidelberg.
- [76] H. Pham. *Continuous-time stochastic control and optimization with financial applications*, volume 61 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2009.
- [77] H. Pham and X. Wei. Discrete time McKean-Vlasov control problem: a dynamic programming approach. *Appl. Math. Optim.*, 74(3):487–506, 2016.

- [78] H. Pham and X. Wei. Dynamic programming for optimal control of stochastic McKean-Vlasov dynamics. *SIAM J. Control Optim.*, 55(2):1069–1101, 2017.
- [79] H. Pham and X. Wei. Bellman equation and viscosity solutions for mean-field stochastic control problem. *ESAIM Control Optim. Calc. Var.*, 24(1):437–461, 2018.
- [80] P. Ren and F.-Y. Wang. Exponential convergence in entropy and Wasserstein for McKean-Vlasov SDEs. *Nonlinear Anal.*, 206:Paper No. 112259, 2021.
- [81] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin, third edition, 1999.
- [82] U. Rieder and N. Bäuerle. Portfolio optimization with unobservable Markov-modulated drift process. *Journal of Applied Probability*, 42(2):362–378, 2005.
- [83] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales. Vol. 2. Itô calculus*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000.
- [84] S. Rudà. Infinite time horizon optimal control of McKean-Vlasov SDEs. *arXiv preprint arXiv:2503.20572*, 2025.
- [85] Y. Shen and T. K. Siu. Consumption-portfolio optimization and filtering in a hidden Markov-modulated asset price model. *J. Indust. Manage. Optim.*, 13(1):23–46, 2017.
- [86] H. M. Soner and Q. Yan. Viscosity solutions for McKean-Vlasov control on a torus. *SIAM J. Control Optim.*, 62(2):903–923, 2024.
- [87] C. Stricker and M. Yor. Calcul stochastique dépendant d’un paramètre. *Z. Wahrsch. Verw. Gebiete*, 45(2):109–133, 1978.
- [88] A.-S. Sznitman. Topics in propagation of chaos. In *École d’Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.
- [89] A. Y. Veretennikov. On ergodic measures for McKean-Vlasov stochastic equations. In *Monte Carlo and quasi-Monte Carlo methods 2004*, pages 471–486. Springer, Berlin, 2006.
- [90] C. Villani. *Optimal transport. Old and new*, volume 338 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin, 2009.
- [91] F.-Y. Wang. Distribution dependent SDEs for Landau type equations. *Stochastic Process. Appl.*, 128(2):595–621, 2018.
- [92] F.-Y. Wang. Exponential ergodicity for non-dissipative McKean-Vlasov SDEs. *Bernoulli*, 29(2):1035–1062, 2023.
- [93] J. Yong. Time-inconsistent optimal control problems. In *Proceedings of the International Congress of Mathematicians—Seoul 2014. Vol. IV*, pages 947–969. Kyung Moon Sa, Seoul, 2014.
- [94] J. Yong and X. Y. Zhou. *Stochastic controls. Hamiltonian systems and HJB equations*, volume 43 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1999.