DIPARTIMENTO MATEMATICA
DIPARTIMENTO DI MATEMATICA "TULLIO LEVI-CIVITA"

Università degli Studi di Padova

Dipartimento di Matematica "Tullio Levi-Civita"

Corso di dottorato in Scienze Matematiche
Curriculum Matematica Computazionale
XXXVI ciclo.

Three Essays in Financial Mathematics

Coordinatore
Prof. Giovanni Colombo
Università degli Studi di Padova

Supervisore
Prof. Claudio Fontana
Università degli Studi di Padova

Dottorando
Giacomo Lanaro
Matricola: 2018575

Anno Accademico
2023-2024

To Laura

Acknowledgments

In reaching the end of this three-year journey, it is time to reflect on the efforts, results, and experiences I've encountered throughout my doctoral studies. Even though I am very happy and proud of the results described in this thesis, I believe that the people I have met and the relationships we have built hold equal importance to the outcomes achieved in the last three years. I believe so not only because many of the people I will mention in the following have inspired me from an academic standpoint, but also because all of them have contributed to both my professional and personal growth.

First of all, I am immensely grateful to my supervisor, Prof. Claudio Fontana. I still remember the first time I met him when I was a master student, asking for some suggestions about a possible theme for my master thesis. From that moment he has guided me with patience and has introduced me to the world of the mathematical finance. Before my master's degree, he suggested me to attend the presentations of a workshop in financial mathematics in Padova. That experience, such fascinating to me, was the first step towards the decision to undertake a PhD program. Later, at the beginning of my PhD, I had no doubt to continue my formation with him. I have learned more than I could have imagined, not only in terms of the application of a rigorous mathematical reasoning, but also in the care devoted to the financial intuition underlying the problems we have tackled. I will be always grateful to him both from a personal and professional level.

I would like to thank Professors Peter Tankov and Martin Larsson, who reviewed my thesis, despite the heterogeneity of the topics it contains. I am grateful for all the comments and suggestions I have received from them. They helped me to improve the quality of this research work and they have shown some insights regarding possible future developments.

A huge thank goes to Alekos Cecchin, who has been almost a second supervisor of this thesis. I thank him for spending a lot of time guiding me through the technical aspects of the project described in the second chapter of this thesis. With patience, he taught me how to tackle the most critical points of the strategy we followed to reach the solution of the problem we faced. My gratitude goes also to Prof. Markus Fischer and Agatha Murgoci who contributed to develop and improve the projects described in this thesis.

I want to thank all the research group of probability and mathematical finance here in Padova. I've been part of a dynamic environment that taught me the importance of conducting critical research supported by the organization of events such as seminars and workshops focused on various

aspects of this research area.

A special thank goes also to my PhD mates, who have shared with me ups and downs of this experience. In particular, thank you Karim for the support, the morning coffees before starting work and especially for showing me a different perspective on various topics, be it in mathematics, politics, or everyday life. Thanks Momo for the jokes, the lunches in mensa, the mutual sharing of the problems encountered in our three years of research work, the matches of billiards (unfortunately for me) and basketball (unfortunately for you). A big thank goes to Ofelia, for being not only a living example on how mathematics should be studied by a doctoral student, but also a true friend during all the years of the university and the doctoral program.

I am also grateful to the entire administrative office, especially to Loretta and Marinella, for their patience and for helping me with bureaucratic and technical issues over these three years.

I am infinitely grateful to my family: my mother Cristina, my father Giorgio and my sister Maria. They have encouraged me since the beginning of my academic journey, making today's achievement possible. In particular, my parents have supported me at the cost of many sacrifices, especially when I was a university student. They taught me not to give up and to keep working hard. Their teaching has been crucial in the last three years to complete this thesis. I would also like to thank my aunt Maddalena, for the patience with which she listened to me repeating the presentations just to maintain a tradition we had started back to the times of my bachelor's degree. I will be forever grateful to my lifelong friends for their constant support, Michele, Federico, Giovanni and Eleonora. I cannot avoid to mention Honey and Tequila, always ready for a walk after hours spent on books.

Finally, thank you Laura. We have been together throughout this journey, from the early days of my undergraduate degree to the conclusion of my doctoral studies. When I faced difficult times, whether due to personal circumstances, pandemic-related problems, or the challenges encountered in tackling the issues of the PhD, you were always there. Over these three years, you have patiently supported me when I was stressed about work, using every time the right words to encourage me. In conclusion, you have been fundamental to complete this research activity.

Abstract

The following PhD Thesis consists of three chapters. In these chapters I describe the projects we developed during my PhD. The first chapter, starting from the preliminary results of my Master's Thesis, presents the study of the problems of consistency and existence of finite-dimensional realizations for Heath-Jarrow-Morton multi-curve interest-rate models. We generalize to the multi-curve setting the geometric approach at the basis of the results obtained by T. Björk and coauthors in the single-curve framework. Hence, we propose a calibration algorithm based on the theoretical results we proved, applying it to Euribor market data.

The second chapter deals with the analysis of a financial market populated by agents who access different amount of information. We study the problem of equilibrium price formation, obtained balancing the demand and supply of a single asset. To do so, we adopt a mean-field approach. We prove the existence of a mean-field equilibrium price, showing that it fills the information gap between a more informed major agent and a group of less informed standard agents. Moreover, we prove that, in the context of a market populated by finitely many standard agents and a major agent, the mean-field equilibrium price satisfies a weak version of the balance between demand and supply.

In the third chapter, we provide several results towards a foundamental theorem of asset pricing for statistical arbitrage opportunities. We study a result present in the literature, that establishes a characterization of the absence of statistical arbitrage opportunities for markets defined on finite probability spaces. For this result a counterexample is provided in a recent paper. Actually, since this counterexample does not disprove the original statement, we confirm the latest in a rigorous mathematical setting. Therefore, we focus on market models defined on more general probability spaces and we show that the characterization proved for the finite markets is still guaranteed, under some additional and tailor-made assumptions.

Sommario

Questa tesi di dottorato è formata da tre capitoli. In questi capitoli descrivo i progetti sviluppati durante il mio percorso di dottorato. Il primo capitolo, a partire dai risultati preliminari ottenuti nella mia tesi magistrale, presenta lo studio dei problemi di consistenza ed esistenza delle realizzazioni finito-dimensionali per modelli Heath-Jarrow-Morton in un contesto multi curva, applicati al mercato dei tassi di interesse. In particolare, generalizziamo al contesto multi curva l'approccio geometrico alla base dei risultati di T. Björk e coautori nel contesto di modelli a curva singola. Proponiamo quindi un algoritmo di calibrazione per un modello Hull-White a tre curve basato sui risultati teorici che abbiamo dimostrato, applicandolo al mercato Euribor.

Nel secondo capitolo analizziamo il meccanismo di formazione del prezzo in un mercato popolato da agenti che hanno accesso a quantità di informazione differenti. In particolare, ci concentriamo sul prezzo che bilancia la domanda e l'offerta di un'azione. Per fare ciò, adottiamo un approccio a campo medio, derivando l'esistenza di un processo di prezzo all'equilibrio. Inoltre, mostriamo che un prezzo così costruito annulla la differenza informativa tra un agente maggiore e maggiormente informato e un gruppo di agenti standard e meno informati. Infine, mostriamo che, nel contesto di un mercato popolato da un agente maggiore e un numero finito di agenti standard, il prezzo di equilibrio ottenuto nel limite a campo medio soddisfa una versione debole dell'incontro tra domanda e offerta.

Nel terzo capitolo, proponiamo alcuni risultati nella direzione di un teorema fondamentale per l'assenza di arbitraggi statistici. Analizziamo un risultato presente in letteratura, che stabilisce una caratterizzazione dell'assenza di arbitraggi statistici nel caso di un mercato definito su uno spazio di probabilità finito. Per questo risultato, un contro esempio è proposto in un articolo più recente. Proviamo però che il contro esempio non confuta il risultato originale, che confermiamo adottando un approccio matematicamente rigoroso. Ci focalizziamo quindi sulla generalizzazione di questa caratterizzazione dell'assenza di arbitraggi statistici, nel contesto di mercati definiti su spazi di probabilità più generali. Mostriamo quindi alcuni risultati preliminari, verificando che la caratterizzazione provata nel caso finito può essere estesa anche a contesti più generali, sotto alcune ipotesi piuttosto restrittive.

Contents

Introduction

As stated in the title, this PhD Thesis deals with three essays concerning some issues in financial mathematics.

In Chapter 1, which is based on [FLM24], we present the problem of consistency and existence of finite-dimensional realizations for a class of multi-curve interest-rate models. The multi-curve framework has been adopted to capture the richer market structure that have emerged in the interest-rate market since fifteen years ago. Motivated by the parameters recalibration problem, we study the problem of consistency between an interest-rate model $\mathcal{M}$ and a parameterized family $\mathcal{G}$. We generalize the geometric approach introduced by Björk and co-authors ([Bjö04; BC99] and [BS01]) in the single curve framework, highlighting the differences that appear in the multi-curve setting, such as the presence of spreads and the interdependence between the different curves of the model. We focus on a class of multi-curve interest-rate models determined by $m + 1$ different tenor lengths. We adopt the Heath-Jarrow-Morton approach to describe $m + 1$ forward-rate curves. On the other hand, m spread components are represented by Itô processes. We assume that each forward-rate component of the model lives on a suitable Hilbert space. Therefore, we provide a characterization of the consistency condition between $\mathcal{M}$ and $\mathcal{G}$, generalizing [Bjö04, Theorem 4.2.]. Consequently, we focus on the problem of existence of finite-dimensional realizations for a given multi-curve model. The finite-dimensional realizations are defined by an opportune finite-dimensional stochastic process, from which the realization of the multi-curve model can be determined. We present conditions that guarantee the existence of finite-dimensional realizations for specific classes of interest-rate models. For these results we adopt the approach proposed in [Sli10]. Hence, we study the conditions under which some properties of the finite-dimensional realizations, provided in the single-curve setting, are respected in the multi-curve framework. Moreover, we introduce a new definition of consistency, based on the multi-curve structure of the interest rates. Finally, applying the theoretical results that we provided, we construct an algorithm to calibrate the parameters of a three-curve Hull-White model. We apply this algorithm to a time series of daily Euribor market data, associated with three different tenor lengths (one day, three months, six months). We obtain stable results.

Chapter 2 deals with the mechanism of equilibrium price formation in presence of asymmetry in the information accessible to the agents of the market. We aim at studying the impact of heterogeneous information on the price of an asset traded in a financial market, adopting a mean-field game approach. Mean-field games theory analyses dynamic systems determined by the interaction of infinitely many rational players. The application of mean-field games theory to financial mathematics has been widely developed in the last years. However, financial application involving heterogeneous information frameworks have been studied only in recent works (see for example [CJ20; CJ19; BS24] and [BCR23]). We adopt the setting described in [FT22b; FT22c], and [FT22a], on which the problem of price formation is studied through a mean-field games approach in the context of symmetric information. We focus on a market model populated by N standard agents and a more informed major agent. The asymmetry of information impacts on the strategies and, therefore, in the functional cost of the major agent. Every agent has to solve an optimal control problem, which depends on an a priori exogenously given stochastic process ϖ describing the price. We want to determine the equilibrium price process obtained by imposing the balance between the demand and supply of the asset. This condition is called market clearing. Assuming that the agents solve their optimal control problems, the market clearing condition leads to an equation for ϖ. To overcome the highly recursive structure of the equation for the equilibrium price, we study its mean-field limit. We derive a mean-field equation for the limit ϖ^{mf} of the equilibrium price ϖ. The equation for ϖ^{mf} involves the dependence on its own filtration, which has to be determined endogenously. To overcome this issue, we hinge on an analogy between the structure of the mean-field equation for ϖ^{mf} and the consistency condition of a weak mean-field game equilibrium in the presence of common noise, described in [CDL16; Lac16; CD18b]. We adapt to our purposes the analogous results developed in [CD18b], that refer to the framework of FBSDE theory. Hence, we prove the existence on the canonical space of a stochastic process ϖ^{mf}. Finally, assuming additional conditions on the market structure, we prove that ϖ^{mf} satisfies a weak version of the market clearing condition.

In Chapter 3, we focus on a class of portfolio strategies called *statistical arbitrage opportunities*. In finance literature, the term statistical arbitrage is used to denote different kinds of trading strategies, which yield a net profit with an assessable amount of risk. As a consequence, statistical arbitrage opportunities have been widely applied in the financial industry. However, a formal definition of statistical arbitrage has not been proposed yet and the term is used to denote ambiguously various different trading strategies. In [LBŠ+18], the authors proposed a first attempt to clarify and compare all the different definitions associated with the term statistical arbitrage. We focus on the definition proposed in [Bon03]. In [Bon03], in the context of financial markets defined on a finite probability space, the author defined a statistical arbitrage as a trading strategy for which the expected payoff is positive and the conditional expected payoff in each final state of the economy is nonnegative. Although the simple framework of [Bon03], it is possible to recognize a link between the results proved by Bondarenko and the role of the conditional expectation, interpreted as a linear operator between L^p spaces. More recently, the notion of statistical arbitrage introduced in [Bon03] has been generalized to the context of more general trading strategies yielding a net conditional expected payoff with respect to an augmented information set, defined by a sigma-algebra $\mathcal{G}$. In [RRS21], some preliminary results have been provided. In this chapter, we study a counterexample proposed in [RRS21] to the result provided in [Bon03]. We show that this counterexample does not disprove the original statement and we present, in a rigorous mathematical setting, the characterization of absence of statistical arbitrage opportunities, provided in [Bon03] for finite market models. Hence, we present several results in the direction of a fundamental theorem of asset pricing for this class of trading strategies. In particular, we focus on the conditions that guarantee the absence of statistical arbitrage opportunities in more general cases, such as a discrete-time market model and a semimartingale model. We prove that the characterizations of the absence of statistical arbitrage opportunities in the framework of finite market models can be extended to market models defined on more general probability spaces under restrictive and tailor-made assumptions.

CHAPTER 1

The geometry of multi-curve interest-rate models

1.1 Introduction

In this chapter we study some geometric properties of the interest-rate market, in a multi-curve setting. This chapter is based on a joint project developed together with my supervisor, Claudio Fontana and Agatha Murgoci. The results of this chapter are described in the work [FLM24], available on ArXiv.

We study how the problems of consistency and existence of finite dimensional realizations can be generalized from the single-curve framework, analysed through the geometric approach developed by T. Björk and coauthors, to a setting that is more coherent with the specificities emerged in the interest-rate market after the 2008 crisis and still effective nowadays. The consistency problem is of interest from an applied and a theoretic point of view. In particular, it is linked to the parameter recalibration problem. Indeed, when an interest-rate model has to be recalibrated, the usual strategy is based on the following two steps:

- Fit the term-structure $\Gamma^M = \{\widehat{r}^M(x);\ x \geq 0\}$ to market data. Usually, families G of parameterized functions, such as the one introduced by Nelson Siegel in [NS87], are employed for this purpose.
- The interest rate model $\mathcal{M}$ is calibrated to the term-structure Γ^M in order to obtain future realizations of the interest-curve in accordance with $\mathcal{M}$.

This procedure is motivated by the notion of invariance between a model $\mathcal{M}$ and the term structure determined by a manifold $\mathcal{G} := \mathrm{Im}[G]$, image of a given parameterized family G. We say that a model $\mathcal{M}$ and a manifold $\mathcal{G}$ are invariant, if the realizations of $\mathcal{M}$ (that are the solution to the dynamics determining the model) belong to $\mathcal{G}$ for a positive time interval. Under this definition, the consistency problem is given by conditions that guarantee a model to be invariant with respect to the image of the parameterized family G used to calibrate the initial term structure Γ^M. We are going to prove that the consistency condition is equivalent to ask that the drift and volatility terms of the dynamics of the realizations of $\mathcal{M}$ are tangent vector fields to $\mathcal{G}$.

A second problem, strictly related to the notion of consistency, concerns the existence of finite-dimensional realizations (FDRs) of an interest-rate model $\mathcal{M}$. An n-dimensional realization for $\mathcal{M}$ is a stochastic process Z taking values on $\mathbb{R}^n$, whose image through a parameterized family G describes the realization of $\mathcal{M}$ for a positive time interval. We show that the existence of an n-dimensional realization for a model $\mathcal{M}$ is equivalent to the existence of a manifold $\mathcal{G}$ such that the couple $(\mathcal{M}, \mathcal{G})$ is invariant. We study conditions that guarantee this property for specific classes of interest-rate models.

The problem of consistency was introduced at first for the single-curve framework in [Bjö04] and [BC99]. At the same time, the existence of FDRs was studied under analogous hypotheses on [BG99], [BL02] and [BS01]. The results included in these papers are based on the interpretation of the realization of a forward-rate model as a curve living on suitable Hilbert space. Some of these results have been extended to a more general setting. In [FT03] and [FT04] analogous geometric properties are presented in the context of generic forward rate models living on Fréchét spaces. Moreover, the study of the geometric properties related to the problem of consistency and existence of FDRs in the case of Lévy models is provided in [FT08; FTT10a; FTT10b; Tap10] and [Tap12].

As mentioned, we consider the problem of consistency and existence of FDRs in the setting of multi-curve. The multi-curve framework has been introduced to describe the new features emerged as a consequence of the profound changes that have affected the interest-rate market in the last fifteen years. Indeed, since the 2008 financial crisis, in order to capture the no longer negligible credit and liquidity risk, spreads have emerged between interbank rates (IBOR rates) with different tenor lengths and overnight rate. As a consequence, the interbank rates associated tenors greater than one day have been considered risk sensitive, while the overnight rate have been assumed risk free. In this setting, to capture the dynamics of both the risk free rate and these risk sensitive rates, the adoption of a multi-curve framework has been a convenient solution. Since 2023, LIBOR is no longer the benchmark inter-bank exchange rate and it has been replaced by risk free rates (RFRs in the following), like SOFR (Secured Overnight Financing Rate) for the American market and €STR (Euro Short Term Rate) for the Eurozone market. However, private and commercial banks in the US and the UK have expressed the desire of an interest-rate that captures the credit and liquidity risk present in the market. Therefore, several American financial institutions, like Bloomberg and AFX, have proposed a set of risk-sensitive rates (RSRs in the following), respectively called BSBY and Ameribor. These rates are designed through a multi-curve setting.

Therefore, we consider the problems of consistency and existence of finite-dimensional realizations for the interest-rate market described by a multi-curve structure determined by $m + 1$ different curves. We introduce a RFR and a set of RSRs, each of them that depends on a tenor δ_j, $j = 1, \dots, m$. We adopt a Heath-Jarrow-Morton approach to represent the dynamics of the instantaneous forward-rate. Then, we introduce multiplicative spot spreads S_t^j between the RSR associated with every tenor δ_j and the RFR, in line with [FGGS20]. Thus, we obtain an infinite-dimensional system of stochastic differential equations whose solution, $\widehat{r}_t(x)$ that is determined by a drift and volatility term, respectively denoted by the parameterized vectors $\widehat{\mu}$ and $\widehat{\sigma}$.

In this setting, we exploit an analogy between the multi-curve interest-rate market and a multi-currency market in order to adapt some results of [Sli10] to the interest-rate market. In particular, we interpret the spread processes between different forward-curves as the exchange rate between two currencies. Following an analogous procedure, we provide conditions that guarantee the existence of FDRs for a model $\mathcal{M}$, driven by a d-dimensional Brownian motion and whose volatility depends on the solution $\widehat{r}$ through a scalar factor.

This chapter is organised as follows. In Section 1.3, we describe the modelling framework. We give the details of the system of stochastic differential equations we deal with and we introduce the functional space at the base of the geometric interpretation of the forward-rate curves. In Section 1.4, we present the main result of the paper for what regards the problem of consistency. In particular, we give a characterization of the notion of consistency. Moreover, we present an application to this result to well known market models. In Section 1.5, we describe conditions the guarantee the existence of FDRs in general and we described some specific cases in which the computations can be provided explictly. Then, in Section 1.6, we present an alternative notion of invariance, that exploits the richer structure given by the multi-curve framework. In particular, a definition of consistency in which the spreads of $\mathcal{M}$ are included in the state process that determines the realizations of $\mathcal{M}$ is proposed. Finally, in Section 1.7, we propose an algorithm based on these results to calibrate the parameters of a three-curve Hull-White model $\mathcal{M}(\theta)$, determined by a set of parameters θ. The goal of a calibration algorithm is to provide an estimation of θ, through the observation of an historical time series of market data. The calibration algorithm that we develop is determined by a parameterized family $G(\theta)$ that is consistent with the model $\mathcal{M}(\theta)$. Therefore, we estimate the parameter θ^* that minimizes the l_2 distance between $G(\theta)$ and the multi-curve term structure extracted by the market data. We test this algorithm on a time series of daily market data represented by Euribor rates associated with three different tenors (one day, three months, six months).

1.2 Notation

We introduce the main notation that we are going to adopt in the paper.

- Calligraphic letters $\mathcal{H}$ and $\mathcal{G}$ are used to denote manifolds defined on the Hilbert space on which the forward-rate curves live.

- We introduce the functionals, applied to differentiable functions $f : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}^n$:

$$\mathbf{F}f(t,x) := \frac{\partial}{\partial x} f(t,x), \quad \mathbf{H}f(t,x) := \left(\int_0^x f^i(t, t+u) du \right)_{i=1,\dots,n}, \quad \mathbf{B}f(t,x) := f(t,0).$$

- We denote by $A^\top$ the transpose of a matrix A and by $v \cdot w$ the usual scalar product between two vectors, $v, w \in \mathbb{R}^n$.
- Given a Fréchét differentiable function $f : \mathcal{H}_1 \to \mathcal{H}_2$, its Fréchét derivative at $\widehat{r} \in \mathcal{H}_1$ will be denoted by $\partial_{\widehat{r}} f(\widehat{r})$. If the Fréchét derivative of a function f is null, we denote it by $\partial_{\widehat{r}} f(\widehat{r}) = \mathbb{O}$.
- We denote the identity map defined on a vector space $\mathcal{H}$ by $\mathbb{1}$. Moreover, if $\mathcal{H} = \mathbb{R}^k$, the identity map on $\mathcal{H}$ is defined by $\mathbb{1}_k$ for each $k \in \mathbb{N} \setminus \{0\}$.
- We denote by $||v||$ the Euclidean norm of a vector $v \in \mathbb{R}^d$.

1.3 The modelling framework

1.3.1 Market set up

In this section, we introduce the dynamics of the multi-curve class of models we are going to study. As discussed, the multi-curve setting is necessary to describe the no longer negligible the credit and liquidity risk present the interest-rate market. Indeed, spreads between RSRs and the RFR can be observed and then interest rates associated with different time lengths (tenors) do no more evolve equivalently. In this framework, we adopt as discount curve, the curve associated the RFR (in analogy with [GR15, Section 1.4.4.]). We model separately the RFR and every RSR, each of them associated with a tenor δ in $\Delta := \{\delta_1 < \delta_2 < \cdots < \delta_m\}$. We adopt a Heath-Jarrow-Morton (HJM) approach to describe all interest rate curves, in line with [CFG16]. In particular, we consider an interest-rate market composed by $m + 1$ curves, one curve associated with the RFR and one for every RSR for each given tenor $\delta_j \in \Delta$.

As discussed, RSRs must take into account the risk that is not captured by the RFR. As a consequence, denoting by $L^\delta(t; T, T + \delta)$ the RSR associated with tenor δ and with $L^0(t; T, T + \delta)$ the RFR, at time t for the interval $[T, T+\delta]$, $L^\delta(t; T, T+\delta) > L^0(t; T, T+\delta)$ typically holds. Hence, we introduce a family of multiplicative spread processes. These spreads are linked to the credit and liquidity risk associated with the forward-rate L^δ:

Definition 1.1. The multiplicative spot spread between the risk-sensitive rate associated with every tenor δ and the risk-free is:

$$S_t^\delta := \frac{1 + \delta L^\delta(t; t, t + \delta)}{1 + \delta L^0(t; t, t + \delta)}. \tag{1.1}$$

It is convenient to introduced also a set of fictitious bonds $B_t^\delta(T)$, defined as follows:

$$B_t^\delta(T) := \frac{B_t(T)}{S_t^\delta} \frac{1 + \delta L^\delta(t; T, T + \delta)}{1 + \delta L^0(t; T, T + \delta)}, \quad t \le T. \tag{1.2}$$

We refer to them as fictitious bonds, because the terminal bond equivalence holds: $B_t^\delta(t) = 1$.

We introduce a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{Q})$, where $\mathbb{Q}$ is a risk-neutral measure. On Ω, a $\mathbb{R}^d$-valued $\mathbb{Q}$-Brownian motion $W = (W_t)_{t\geq 0}$ is defined. Then, we consider the following processes:

Risk-free Curve We describe the risk free instantaneous forward-rates, through Musiela parameterization, as the solution to

$$dr_t^0(x) := \alpha_t^0(t+x)dt + \sigma_t^0(t+x)dW_t, \quad 0 \leq t \leq T, \tag{1.3}$$

where $x \geq 0$ is the time-to-maturity. We denote the price of a zero-coupon bond on r^0 by

$$B_t(T) := \exp\left(- \int_0^{T-t} r_t^0(x)dx \right), \quad t \in [0, T].$$

The savings account numéraire associated with the RFR is given by $S^0 := \exp(\int_0^\cdot r_t^0(0)dt)$.

Risk-sensitive Curve The risk-sensitive curve, associated with the tenor δ_j for $j = 1, \dots, m$, is characterized by the dynamics of instantaneous rates associated with the fictitious Bonds (1.2) as the solution to

$$dr_t^j(x) := \alpha_t^j(t+x)dt + \sigma_t^j(t+x)dW_t, \quad j = 1, \dots, m, \tag{1.4}$$

for suitable coefficients $\alpha_t^j(t+x)$ and $\sigma_t^j(t+x)$. As a consequence,

$$B_t^j(T) := \exp\Big(- \int_0^{T-t} r_t^j(x)dx \Big), \quad t \in [0, T].$$

Even if these bonds are not actually traded in the market, it is possible to reconstruct their market value substituting in (1.2) the values of $L^\delta(t; T, T+\delta)$, $L^0(t; T, T+\delta)$ (that can be obtained via bootstrapping techniques from market quotations, see Section 1.7.4 below).

The spreads The spread processes are defined as the solution to an exponential of an Itô process, determined by suitable coefficients γ^j and β^j. In particular, $S_t^j := e^{Y_t^j}$, for $t \geq 0$ where

$$dY_t^j = \gamma_t^j dt + \beta_t^j dW_t, \quad j = 1, \dots, m. \tag{1.5}$$

In the classical single-curve setting, the HJM drift condition implies that the drift term α^0 in (1.3) is determined by the volatility σ^0 ([Bjö04, Proposition 1.1]). In the present multi-curve setup, risk-neutrality of $\mathbb{Q}$ implies that, for each $j = 1, ..., m$, the drift term α^j in (1.4) is determined by the volatility σ^j as well as by the covariation between r^j and the log-spread process Y^j . Moreover, the drift term γ^j in (1.5) turns out to be endogenously determined. This is the content of the following proposition, which follows as a special case of [FGGS20, Theorem 3.7].

Proposition 1.2. *Under a risk-neutral probability measure $\mathbb{Q}$, the RFR, the RSR and the logarithm* *of the spread process associated with every tenor δ_j with $j = 1, \ldots, m$ are respectively determined by the following system of SDEs:*

$$\begin{cases} dr_t^0(x) &= \Big(\mathbf{F}r_t^0(x) + \sigma_t^0(t+x) \cdot \mathbf{H}\sigma_t^0(x)\Big)dt + \sigma_t^0(t+x)dW_t; \\ dr_t^j(x) &= \Big(\mathbf{F}r_t^j(x) + \sigma_t^j(t+x) \cdot \mathbf{H}\sigma_t^j(x) - \beta_t^j \cdot \sigma_t^j(t+x)\Big)dt + \sigma_t^j(t+x)dW_t; \\ dY_t^j &= \Big(\mathbf{B}r_t^0 - \mathbf{B}r_t^j - \frac{1}{2}||\beta_t^j||^2\Big)dt + \beta_t^j dW_t. \end{cases} \tag{1.6}$$

Remark 1.3. The drift condition on the dynamics of (1.6) stated in Proposition 1.2 is equivalent to the local martingale property under $\mathbb{Q}$ of the processes $\frac{B^0(T)}{S^0}$ and $\frac{S^j B^j(T)}{S^0}$ for all $T > 0$ and $j = 1, ..., m$. This property is taken as the defining property of a risk-neutral probability. As clarified in [FGGS20], this suffices to ensure absence of arbitrage in the financial market composed by all risk-free ZCBs and single-period swaps referencing the risk-sensitive rates $L^j(T; T, T + \delta_j)$, for all $T > 0$ and $j = 1, ..., m$.

System (1.6) is made by $2m + 1$ stochastic differential equations, two for each tenor δ and one for the risk-free curve. Moreover, the first $m + 1$ components, associated with the forward-rates, depend on a positive real parameter x (time-to-maturity). Let us notice that the solution to (1.6) is fully determined by σ^0, σ^j, and β^j, for every $j = 1, \ldots, m$, by non-arbitrage constraints.

1.3.2 The modelling framework

Our purpose is to find conditions that guarantee the couple $(\mathcal{M}, \mathcal{G})$ to be consistent, where $\mathcal{M}$ and $\mathcal{G}$ denote respectively a forward-rate model, whose dynamics is determined by system (1.6), and a parameterized family of forward-rates. The concept of consistency can be introduced as follows:

Definition 1.4. An interest rate model $\mathcal{M}$ and a parameterized family of forward-rate curves $\mathcal{G}$ are consistent if $\mathcal{M}$ produces forward-rate curves which belong to $\mathcal{G}$ for a positive time interval.

As discussed in Section 1.3.1, a multi-curve model $\mathcal{M}$ is the solution of a system of SDEs respecting the structure of (1.6). Denoting a single-curve forward-rate by $r := (r_t)_{t \geq 0}$, then, r can be interpreted as a curve evolving on a Hilbert space $\mathcal{H} \subseteq \mathcal{C}^\infty(\mathbb{R}_+, \mathbb{R})$. We suppose that:

Definition 1.5. The solution of each forward-rate component of the system (1.6), denoted at time $t \geq 0$ by r_t^j for $j = 0, \ldots, m$, belongs to the infinite-dimensional space

$$\mathcal{H} := \Big\{ r \in \mathcal{C}^\infty(\mathbb{R}^+, \mathbb{R}) \text{ s.t. } ||r||_\gamma^2 := \sum_{n=0}^{+\infty} 2^{-n} \int_0^{+\infty} \Big(\frac{\partial^n}{\partial x^n} r(x)\Big)^2 e^{-\gamma x} dx < +\infty \Big\}. \tag{1.7}$$

for $\gamma > 0$.

$(\mathcal{H}, || \cdot ||_\gamma)$ is an Hilbert space for each $\gamma > 0$ ([BS01, Proposition 4.2]). The solution to system (1.6), denoted by $\widehat{r} := (r^0, \ldots, r^m, Y^1, \ldots, Y^m)$, is a stochastic process defined on the space $\widehat{\mathcal{H}} := \mathcal{H}^{m+1} \times \mathbb{R}^m$, where $\mathcal{H}^{m+1}$ is the cartesian product of $m + 1$ copies of $\mathcal{H}$. In the following, we may adopt the notation $\widehat{r} = (r, Y) \in \widehat{\mathcal{H}}$, where $r \in \mathcal{H}^{m+1}$ and $Y \in \mathbb{R}^m$, when convenient.

We make some technical assumptions regarding the regularity of the components of the system.

Assumption 1.6. We suppose that:

- The volatility of each component of system (1.6) is defined by $\sigma_t^j(t+\cdot) := \sigma^j(\widehat{r}_t)$ and $\beta_t^h := \beta^h(\widehat{r}_t)$, for each $j = 0, \dots, m$, $h = 1, \dots, m$, where σ^j, β^h are deterministic functions defined on $\widehat{\mathcal{H}}$ and taking values on $\mathcal{H}^d$ and $\mathbb{R}^d$ respectively, for every $j = 0, \dots, m$ and $h = 1, \dots, m$.
- We assume that σ^i and β^j are smooth functions in Fréchét sense, i.e. they admit continuous n^{th} order Fréchét derivatives for each $n \in \mathbb{N} \setminus \{0\}$. Moreover,

$$
\begin{aligned}
A(\widehat{r}) &:= \sigma^0(\widehat{r}) \cdot \mathbf{H}\sigma^0(\widehat{r}) - \frac{1}{2}\partial_{\widehat{r}}\sigma^0(\widehat{r})\widehat{\sigma}(\widehat{r}),\\
B(\widehat{r}) &:= \sigma^j(\widehat{r}) \cdot \mathbf{H}\sigma^j(\widehat{r}) - \frac{1}{2}\partial_{\widehat{r}}\sigma^j(\widehat{r})\widehat{\sigma}(\widehat{r}) - \sigma^j(\widehat{r}) \cdot \beta^j(\widehat{r}), \quad j = 1, \dots, m
\end{aligned}
$$

  are smooth (in the Fréchét sense) too.

We compactly denote (1.6) by

$$
d\widehat{r}_t = \mu(\widehat{r}_t)dt + \widehat{\sigma}(\widehat{r}_t)dW_t,
$$

where μ and $\widehat{\sigma}$ are fully determined by

$$
\widehat{\sigma}(\widehat{r}) = (\sigma^0(\widehat{r}), \dots, \sigma^m(\widehat{r}), \beta^1(\widehat{r}), \dots, \beta^m(\widehat{r}))^\top \in \widehat{\mathcal{H}}^d.
$$

To apply the chain rule, we pass from the formulation of the dynamics of $\widehat{r}$ under the Itô integral to the one given by Stratonovich integral. In analogy with [KS12, Definition 3.3.13], we denote the Stratonovich integral of a process X with respect a process Y, with the symbol $\int X_s \circ dY_s$. Then, the solution of the forward-rate system is rewritten as:

$$
d\widehat{r}_t = \widehat{\mu}(\widehat{r}_t)dt + \widehat{\sigma}(\widehat{r}_t) \circ dW_t, \quad \widehat{\mu}(\widehat{r}_t) := \mu(\widehat{r}_t) - \frac{1}{2}\partial_{\widehat{r}}\widehat{\sigma}(\widehat{r}_t)\widehat{\sigma}(\widehat{r}_t), \tag{1.8}
$$

where

$$
\widehat{\mu}(\widehat{r}_t) = \begin{pmatrix} \mathbf{F}r^0 + \sigma^0(\widehat{r}_t) \cdot \mathbf{H}\sigma^0(\widehat{r}) \\ \mathbf{F}r^1 + \sigma^1(\widehat{r}) \cdot \mathbf{H}\sigma^1(\widehat{r}) - \sigma^1(\widehat{r}) \cdot \beta^1(\widehat{r}) \\ \vdots \\ \mathbf{F}r^m + \sigma^m(\widehat{r}) \cdot \mathbf{H}\sigma^m(\widehat{r}) - \sigma^m(\widehat{r}) \cdot \beta^m(\widehat{r}) \\ \mathbf{B}r^0 - \mathbf{B}r^1 - \frac{1}{2}||\beta^1(\widehat{r})||^2 \\ \vdots \\ \mathbf{B}r^0 - \mathbf{B}r^m - \frac{1}{2}||\beta^m(\widehat{r})||^2 \end{pmatrix} - \frac{1}{2}\partial_{\widehat{r}}\widehat{\sigma}(\widehat{r}) \begin{pmatrix} \sigma^0(\widehat{r}) \\ \sigma^1(\widehat{r}) \\ \vdots \\ \sigma^m(\widehat{r}) \\ \beta^1(\widehat{r}) \\ \vdots \\ \beta^m(\widehat{r}) \end{pmatrix}, \tag{1.9}
$$

$$\partial_{\widehat{r}}\widehat{\sigma}(\widehat{r}) := \begin{pmatrix} \partial_{r^0}\sigma^0(\widehat{r}) & \cdots & \partial_{r^m}\sigma^0(\widehat{r}) & \partial_{Y^1}\sigma^0(\widehat{r}) & \cdots & \partial_{Y^m}\sigma^0(\widehat{r}) \\ \vdots & & \vdots & \vdots & & \vdots \\ \partial_{r^0}\sigma^m(\widehat{r}) & \cdots & \partial_{r^m}\sigma^m(\widehat{r}) & \partial_{Y^1}\sigma^m(\widehat{r}) & \cdots & \partial_{Y^m}\sigma^m(\widehat{r}) \\ \partial_{r^0}\beta^1 & \cdots & \partial_{r^m}\beta^1 & \partial_{Y^1}\beta^1(\widehat{r}) & \cdots & \partial_{Y^m}\beta^1(\widehat{r}) \\ \vdots & & \vdots & \vdots & & \vdots \\ \partial_{r^0}\beta^m & \cdots & \partial_{r^m}\beta^m & \partial_{Y^1}\beta^m(\widehat{r}) & \cdots & \partial_{Y^m}\beta^m(\widehat{r}) \end{pmatrix}.$$

1.4 Consistency problem

In this section we solve the problem of consistency through the geometric approach developed in [Bjö04], [BC99] and [BS01] and generalized to the multi-curve framework introduced in the previous section.

1.4.1 Characterization of consistency condition

We consider a mapping G defined on an open subset of $\mathbb{R}^n$, denoted by $\mathcal{Z}$, that determines a manifold $\mathcal{G} \subset \widehat{\mathcal{H}}$. We assume that:

Assumption 1.7. $G : \mathcal{Z} \subset \mathbb{R}^n \to \widehat{\mathcal{H}}$ is an injective function such that its Fréchét differential, $\partial_z G : \mathbb{R}^n \to \widehat{\mathcal{H}}$, is injective for each $z \in \mathcal{Z}$. As a consequence, $\mathcal{G} := \text{Im}[G]$ is a submanifold of $\widehat{\mathcal{H}}$.

In the following, we refer to a manifold $\mathcal{G}$ as the image of a parametrized family satisfying Assumption 1.7.

The consistency between $\mathcal{M}$ and a sub-manifold $\mathcal{G}$, is defined by the notion of *invariance*:

Definition 1.8. Let us consider a forward-rate dynamics as (1.8), whose solution $\widehat{r}$ describes a model $\mathcal{M}$ and a function G satisfying Assumption 1.7. Then, if $\mathcal{G}$ denotes the image of G, we say that the couple $(\mathcal{M}, \mathcal{G})$ is locally invariant under the action of $\widehat{r}$ if for each $(\widehat{r}_s, s) \in \mathcal{G} \times \mathbb{R}_+$ there exists a stopping time τ, depending on s and $\widehat{r}_s$, such that:

$$\begin{aligned} &\tau(\widehat{r}_s, s) > s, \qquad \mathbb{Q} - a.s.; \\ &\widehat{r}_t \in \mathcal{G}, \qquad \text{for each } t \in [s, \tau(s, r_s)). \end{aligned}$$

Our aim is to find a characterization of the previous definition in terms of conditions on the coefficients of $\widehat{r}$. To this purpose, we exploit the equivalence between the notion of invariance and the one of $\widehat{r}$-invariance:

Definition 1.9. We say that a parameterized family G is locally $\widehat{r}$-invariant under the action of the forward-rate process $\widehat{r}$ if for each $\widehat{r}_0 \in \mathcal{G}$:=Im$[G]$ there exist a $\mathbb{Q}$-a.s. strictly positive stopping time $\tau(\widehat{r}_0)$ and a stochastic process $(Z_t)_t$ taking values in $\mathbb{R}^n$, whose Stratonovich dynamics is $dZ_t = a(Z_t)dt + b(Z_t) \circ dW_t$ such that for each $t \in [0, \tau(\widehat{r}_0))$, $r_t(x) = G(x, Z_t)$ for each $x \geq 0$, $\mathbb{Q}$-a.s..

The equivalence between Definition 1.8 and Definition 1.9 leads to the following characterization of the consistency conditions in terms of $\widehat{\mu}$ and $\widehat{\sigma}$, introduced in equation (1.9):

Theorem 1.10 (Invariance condition)**.** *We consider a forward-curve manifold $\mathcal{G} = Im[G]$ and a model $\mathcal{M}$ given by the solution to equation* (1.6)*. The couple $(\mathcal{M}, \mathcal{G})$ is invariant if and only if the following conditions hold for each $z \in \mathcal{Z}$:*

$$\begin{aligned} \widehat{\mu}(G(z)) &\in Im[\partial_z G(z)] := T_{G(z)}\mathcal{G}, \\ \widehat{\sigma}_i(G(z)) &\in Im[\partial_z G(z)] := T_{G(z)}\mathcal{G}, \quad \forall i = 1, \dots, d, \end{aligned} \tag{1.10}$$

where $T_{G(z)}\mathcal{G}$ denotes the tangent space of the manifold $\mathcal{G}$ at the point $G(z)$, for each $z \in \mathcal{Z}$.

Proof. Direct generalization of the proof of [BC99, Theorem 4.1]). □

Condition $\widehat{\mu}(G(z)), \widehat{\sigma}(G(z)) \in T_{G(z)}\mathcal{G}$ is equivalent to assume that the distribution generated by $\widehat{\mu}$ and $\widehat{\sigma}$ (the subspace of the tangent bundle of $\widehat{\mathcal{H}}$ generated by $\widehat{\mu}$ and $\widehat{\sigma}_i$ for each $i = 1, \dots, d$) is a subset of $T\mathcal{G}$, where $T\mathcal{G}$ is the tangent bundle of $\mathcal{G}$.

Differently from the analogous result proved for the single-curve setting, we must handle the richer structure of the multi-curve framework. Indeed $\widehat{\mu}(\widehat{r})$ and $\widehat{\sigma}(\widehat{r})$, introduced in (1.9), are vector fields defined on $\widehat{\mathcal{H}}$, that is a finite product of Hilbert spaces. As a consequence, relations between the components of function G can be found in order to guarantee the couple $(\mathcal{M}, \mathcal{G})$ to be invariant.

1.4.2 An example: Hull-White model and Nelson-Siegel family

In this section we apply Theorem 1.10 to determine when classical models and well-known parameterized forward-curves are consistent. We focus on the multi-curve Hull-White model and the Nelson-Siegel family (more specifically a suitable modification of this family) to determine conditions that guarantee this couple to be consistent.

Nelson-Siegel family In analogy with [BC99], we consider a forward parameterized family, frequently used in literature, the Nelson-Siegel (NS) family. Introduced in [NS87], NS family is parameterized by $z = (z_1, z_2, z_3, z_4) \in \mathcal{Z} := \mathbb{R}^4$ as:

$$G^{NS}(z; x) := z_1 + z_2 e^{-z_4 x} + z_3 x e^{-z_4 x} = z_1 + e^{-z_4 x}[z_2 + z_3 x]. \tag{1.11}$$

For a detailed description of this family we refer to [Fil99]. We denote the manifold $\mathrm{Im}[G^{NS}] \subset \mathcal{H}$ by $\mathcal{G}^{NS}$.

Hull-White model We consider the Hull-White (HW) model. We focus on the case of a 1-dim Brownian motion W, but the results of this section can be generalized to the case of Hull-White model driven by a general d-dimensional Brownian motion. The volatility of HW model is σe^{-ax}

where $\sigma, a \in \mathbb{R}_+$. Then, the forward-rate equation in Musiela parameterization is:

$$dr_t(x) = \Big[\mathbf{F}r_t(x) + \frac{\sigma^2}{a}e^{-ax}\Big(1 - e^{-ax}\Big)\Big]dt + \sigma e^{-ax}dW_t. \tag{1.12}$$

The multi-curve model We consider a multi-curve version of HW model $\mathcal{M}$, whose dynamics is of the form of (1.8). The forward-rate volatilities of $\mathcal{M}$ are $\sigma^j e^{-a^j x}$, where $a^j, \sigma^j > 0$. Since HW model is a constant volatility model (σe^{-ax} does not depend on t), it is natural to suppose that the volatility of the log-spot spread is a constant $\beta^j > 0$ for each $j = 1, \ldots, m$. As a consequence, the volatility of $\mathcal{M}$ is $\widehat{\sigma} := (\sigma^0 e^{-a^0 x},\ \cdots,\ \sigma^m e^{-a^m x},\ \beta^1,\ \cdots,\ \beta^m)$.

Our purpose is to find a manifold $\mathcal{G}$ in $\widehat{\mathcal{H}}$ such the drift and the volatility of $\mathcal{M}$ belong to the tangent bundle of $\mathcal{G}$. In particular, the conditions of Theorem 1.10 on $\widehat{\mu}$ and $\widehat{\sigma}$ must hold on every component. Since HW is inconsistent with $\mathcal{G}^{NS}$ already in the single-curve ([BC99, Section 5]), we propose the following modification the NS family

$$G_0(z^0; x) := z_1^0 + z_2^0 e^{-a^0 x} + z_3^0 x e^{-a^0 x} + z_4^0 e^{-2a^0 x}, \tag{1.13}$$

to construct a consistent parameterized family, in analogy with [BC99].

Proceeding component-wise, we notice that $\mu^0(G_0(z^0; x)) = \partial_{z^0} G_0(z^0; x)\eta^0(z^0)$, where

$$\eta^0(z^0) := \Big(0,\quad -a^0 z_2^0 + z_3^0 + \tfrac{(\sigma^0)^2}{a^0},\quad -a^0 z_3^0,\quad -\Big(2a^0 z_4^0 + \tfrac{(\sigma^0)^2}{a^0}\Big)\Big) \in \mathbb{R}^4. \tag{1.14}$$

The image of η with respect to the Fréchét derivative of G_0 is μ^0, thus the first component $\widehat{\mu}$ belongs to $\mathrm{Im}[\partial_{z^0} G_0]$. For what regards the first component of the volatility term, we notice that $\sigma^0 e^{-a^0 x} = \partial_{z^0} G_0(z^0; x)\xi^0(z^0)$, where $\xi^0(z^0) := (0,\ \sigma^0,\ 0,\ 0)$. Therefore, we constructed two vector fields on $\mathbb{R}^4$ for which the characterization of the consistency condition (1.10) holds in the component associated with the RFR.

For the risk-sensitive forward-rates the structure of coefficients is similar to μ^0 and σ^0, thus it is possible to find vector fields η^j and ξ^j such that condition (1.10) is satisfied for μ^j and σ^j for every $j = 1, \ldots, m$. In particular, we introduce the functions:

$$G_j(z^j; x) = z_1^j + z_2^j e^{-a^j x} + z_3^j x e^{-a^j x} + z_4^j e^{-2a^j x}, \quad j = 1, \ldots, m, \tag{1.15}$$

where $z^j = (z_1^j, \ldots, z_4^j) \in \mathbb{R}^4$. Then, defining

$$\eta^j(z^j) := \Big(0,\quad -a^j z_2^j + z_3^j + \tfrac{(\sigma^j)^2}{a^j} - \beta^j\sigma^j,\quad -a^j z_3^j,\quad -\Big(2a^j z_4^j + \tfrac{(\sigma^j)^2}{a^j}\Big)\Big). \tag{1.16}$$

and $\xi^j = (0,\ \sigma^j,\ 0,\ 0)$, conditions $\partial_{z^j} G_z(z^j; x)\eta^j(z^j) = \mu^j(x)$ and $\partial_{z^j} G_z(z^j; x)\xi^j(z^j) = \sigma^j e^{-a^j x}$, hold. Thus, we introduce the function

$$G := (G_0, \ldots, G_m, G_{m+1}, \ldots, G_{2m}), \tag{1.17}$$

where G_{m+j} are suitable $\mathbb{R}$-valued functions for $j = 1, \ldots, m$. As a consequence, the vector fields:

$$\begin{aligned} \eta &:= \left(\eta^0, \ldots, \eta^m\right) : \mathbb{R}^{4(m+1)} \to \mathbb{R}^{4(m+1)}, \\ \xi &:= \left(\xi^0, \ldots, \xi^m\right) : \mathbb{R}^{4(m+1)} \to \mathbb{R}^{4(m+1)}, \end{aligned} \tag{1.18}$$

satisfy the consistency on every forward-rate component. In conclusion, we characterized the forward-rate components of HW model via a consistent parameterized family G that is defined on the state space $\mathbb{R}^n$ with $n := 4(m+1)$.

The presence of the spreads in the multi-curve framework has to be handled in a different way. Indeed, there is not a standard way to parameterize the spread components of a consistent family G. Thus, we present two strategies to construct the spread components of the function G in order to guarantee condition (1.10). In the first case, we enlarge the state space to introduce additional parameters that determine the spread components of the model, with no relation between this new parameters and the vectors η and ξ of equation (1.18). On the other hand, we try to determine conditions on the components of the volatility, β^j, σ^j for $j = 1, \ldots, m$ and σ^0, such that $G_{m+1} \ldots, G_{2m}$ are implicitly determined by the vector fields η and ξ introduced in equation (1.18). In particular:

- We can enlarge the domain of definition of the state space $\mathbb{R}^n$. Since the components associated with the spreads are finite-dimensional, we introduce the consistent parameterized family as $G_{m+j}(u^j) := u^j$, $u^j \in \mathbb{R}$, $j = 1, \ldots, m$. To guarantee the consistency condition, we suppose that $\xi^{m+j}(z) = \beta^j$ and we introduce vector fields η^{m+j} on $\mathbb{R}$ as follows:

  $$\eta^{m+j}(z) := \widehat{\mu}^{m+j}(G(z)) = \mathbf{B}G_0(z^0) - \mathbf{B}G_j(z^j) - \frac{1}{2}(\beta^j)^2 = z_1^0 + z_2^0 + z_4^0 - (z_1^j + z_2^j + z_4^j) - \frac{1}{2}(\beta^j)^2.$$

  Proceeding analogously for the volatility term, we conclude that the parameterized family

  $$G(z) := (G_0(z^0), \ldots, G_m(z^m), u^1, \ldots, u^m), \tag{1.19}$$

  defined for any $z := (z^0, \ldots, z^m, u^1, \ldots, u^m) \in \mathbb{R}^{5m+4}$, forms a consistent couple with the HW model.

- Alternatively, one can find relations between the coefficients of the volatility term such (1.10) is satisfied by the vector fields defined in (1.18). In particular, Proposition 1.11 below can be proved.

Proposition 1.11. *We consider the model $\mathcal{M}$ given by the HW model for each forward-rate equation. We suppose that the following constraint is satisfied:*

$$\beta^j = \frac{\sigma^j}{a^j} - \frac{\sigma^0}{a^0} \tag{1.20}$$

We consider a manifold $\mathcal{G}$ image of a function $G : \mathbb{R}^{4(m+1)} \to \widehat{\mathcal{H}}$, of the form introduced in *equation* (1.17)*. In particular, if G is defined by the function G_j introduced in equation* (1.15) *for $j = 0, \dots, m$ and*

$$G_{m+j}(z) = \frac{1}{a^0}\Big[- z_2^0 + \Big(-z_1^0 - \frac{(\sigma^0)^2}{2(a^0)^2} + \frac{1}{2}(\beta^j)^2\Big)\log z_3^0 - \frac{z_3^0}{a^0} - \frac{1}{2}z_4^0\Big] + \frac{1}{a^j}\Big[z_2^j + \Big(z_1^j + \frac{(\sigma^j)^2}{2(a^j)^2} - \frac{\beta^j\sigma^j}{a^j}\Big)\log z_3^j + \frac{z_3^j}{a^j} + \frac{1}{2}z_4^j\Big], \tag{1.21}$$

for $j = 1, \dots, m$, then the couple $(\mathcal{M}, \mathcal{G})$ is consistent.

Proof. We construct a set of functions G_{m+j}, such that the characterization of the consistency condition expressed in Theorem 1.10, for any $j = 1, \dots, m$ is given by:

$$\widehat{\mu}^{m+j}(G(z)) := \mathbf{B}G_0(z^0) - \mathbf{B}G_j(z^j) - \frac{1}{2}(\beta^j)^2 = \partial_z G_{m+j}(z)\eta(z), \tag{1.22}$$

$$\widehat{\sigma}^{m+j}(G(z)) := \beta^j = \partial_z G_{m+j}(z)\xi(z), \tag{1.23}$$

where η and ξ are determined in (1.18). By (1.13) and (1.15), the following equality holds:

$$\mathbf{B}G_0(z^0) - \mathbf{B}G_j(z^j) - \frac{1}{2}(\beta^j)^2 = z_1^0 + z_2^0 + z_4^0 - z_1^j - z_2^j - z_4^j - \frac{1}{2}(\beta^j)^2. \tag{1.24}$$

As a consequence, by (1.22), for any $j = 1, \dots, m$, it seems natural to suppose that $G_{m+j}(z)$ depends only on the variables $(z_1^0, z_2^0, z_3^0, z_4^0, z_1^j, z_2^j, z_3^j, z_4^j)$. Moreover, it is convenient, in order to develop the computations, to assume the following structure for the function G_{m+j}:

$$G_{m+j}(z) := \phi_1^0(z_1^0, z_3^0) + \phi_2^0(z_2^0) + \phi_4^0(z_4^0) + \phi_1^j(z_1^j, z_3^j) + \phi_2^j(z_2^j) + \phi_4^j(z_4^j), \tag{1.25}$$

for suitable functions $\phi_1^0, \phi_2^0, \phi_3^0, \phi_4^0, \phi_1^j, \phi_2^j, \phi_3^j$ and ϕ_4^j. Substituting (1.24) and (1.25) in (1.22), we conclude that:

$$\begin{aligned}\partial_z G_{m+j}(z)\eta(z) = {} & \partial_{z_2^0}\phi_2^0(z_2^0)\Big(-a^0 z_2^0 + z_3^0 + \frac{(\sigma^0)^2}{a^0}\Big) + \partial_{z_3^0}\phi_1^0(z_1^0, z_3^0)\Big(-a^0 z_3^0\Big) \\ & - \partial_{z_4^0}\phi_4^0(z_4^0)\Big(-2a^0 z_4^0 - \frac{(\sigma^0)^2}{a^0}\Big) + \partial_{z_2^j}\phi_2^j(z_2^j)\Big(-a^j z_2^j + z_3^j + \frac{(\sigma^j)^2}{a^j} \\ & - \beta^j\sigma^j\Big) + \partial_{z_3^j}\phi_3^j(z_3^j)\Big(-a^j z_3^j\Big) - \partial_{z_4^j}\phi_4^j(z_4^j)\Big(-2a^j z_4^j - \frac{(\sigma^j)^2}{a^j}\Big) \\ = {} & z_1^0 + z_2^0 + z_4^0 - z_1^j - z_2^j - z_4^j - \frac{1}{2}(\beta^j)^2.\end{aligned} \tag{1.26}$$

We aim at determining the functions ϕ_k^j that determine G_{m+j}. The first step is to separate the terms dependent on the variables $(z_1^0, z_2^0, z_3^0, z_4^0)$ with the ones dependent on $(z_1^j, z_2^j, z_3^j, z_4^j)$. Therefore, assuming $\partial_{z_4^0}\phi_4^0(z_4^0) = -\frac{1}{2a^0}$ we simplify the term dependent on z_4^0. We compute the second order derivative with respect to z_2^0 and z_3^0 to conclude that it is admissible to assume $\partial_{z_2^0}\phi_2^0(z_2^0) = -\frac{1}{a^0}$. Doing so, we simplify the term on z_2^0. Finally, we differentiate with respect to z_1^0, to obtain

$\partial^2_{z_3^0,z_1^0}\phi_1^0(z_1^0,z_3^0) = -\frac{1}{a^0 z_3^0}$. Hence, we must require $z_3^0 > 0$ to integrate this equation and to obtain $\partial_{z_3^0}\phi_1^0(z_1^0,z_3^0) = -\frac{1}{a^0 z_3^0}(z_1^0 + \bar{c}) + \psi(z_3^0)$, for a constant $\bar{c} \in \mathbb{R}$ and a function ψ. Finally, to determine $\bar{c}$ and ψ we substitute those terms in (1.26). We adopt a similar strategy for the variables $(z_1^j, z_2^j, z_3^j, z_4^j)$ and we conclude that G_{m+j} introduced in equation (1.21) can be adopted as a consistent function for the drift term of the HW spread process. To verify the consistency condition for the volatility term of the spread process, we notice that, recalling that ξ is the vector field introduced in (1.18), condition (1.23) becomes:

$$\beta^j = \partial_z G_{m+j}(z)\xi = \partial_{z_2^0 G_{m+j}}(z)\sigma^0 + \partial_{z_2^j} G_{m+j}\xi_2^j = -\frac{\sigma^0}{a^0} + \frac{\sigma^j}{a^j}.$$

This condition is already satisfied by the constraint (1.20) on the volatility term. □

Remark 1.12. The first strategy extends the domain of the function G, introduced in equation (1.17), for which the consistency condition is guaranteed only on the forward-rate components of the HW model. This procedure leads to a consistent parameterized family whose domain is $\mathbb{R}^{n_1}$, where $n_1 := 4(m+1) + m$. Adopting the second strategy, the domain of the consistent function is $\mathbb{R}^{n_2}$, with $n_2 := 4(m+1)$. Then, we obtain a more parsimonious consistent family through the second approach, at the cost of a condition on the coefficients of the volatility term of HW model. It is possible to show that condition $\beta^j = \frac{\sigma^j}{a^j} - \frac{\sigma^0}{a^0}$ implies that the spread process Y^j is a obtained as a deterministic affine transformation of the spot processes $(\mathbf{B}r_t^0, \mathbf{B}r_t^j)$. As a consequence, a parameterized family that is consistent with the $(\mathbf{B}r_t^0, \mathbf{B}r_t^j)$ determines automatically a parameterized family for the spread process Y^j. This justifies why, assuming $\beta^j = \frac{\sigma^j}{a^j} - \frac{\sigma^0}{a^0}$, it is possible not to extend the domain of G.

1.5 Existence of finite-dimensional realizations

In this section, we focus on the following problem: *"Can the solution of the system* (1.6) *be described as the image of a finite-dimensional stochastic process?"*.

This problem concerns the existence of the finite-dimensional realizations (FDRs) for a given model $\mathcal{M}$. We present conditions that guarantee a given model, described by equation (1.6), to possess a finite-dimensional representation. Moreover, if it is the case, we provide a strategy to construct the FDRs. To this effect, we exploit some results of infinite-dimensional geometric theory, applying them to the setting developed in Section 1.4. Then, we present two examples. First, we analyse the case of a constant volatility model, driven by a scalar Brownian motion. Finally, we study a model driven by a d-dimensional Brownian motion whose volatility depends on the solution the system (1.6) via scalar factors.

1.5.1 The strategy

We say that the solution to (1.6) has FDRs if it has an n-dimensional realization, for some $n \in \mathbb{N}$, defined as follows:

Definition 1.13. We say that the solution to (1.8) has a n-dimensional realization if, for each $\widehat{r}_0^M \in \widehat{\mathcal{H}}$, there exist a stopping time $\tau > 0$ a.s., $z_0 \in \mathbb{R}^n$, $d+1$ smooth vector fields $a, b_1 \ldots, b_d$, defined on a neighborhood of z_0, denoted with $\mathcal{Z}$ and a function $G : \mathcal{Z} \to \widehat{\mathcal{H}}$, satisfying Assumption 1.7, such that $\widehat{r}$ satisfies $\widehat{r}_t = G(Z_t)$ for each $t \in [0, \tau)$, where

$$\begin{cases} dZ_t &= a(Z_t)dt + b(Z_t) \circ dW_t, \\ Z_0 &= z_0. \end{cases}$$

Definition 1.13 is strictly related to the concept of $\widehat{r}$-invariance (Definition 1.9). Indeed, the existence of a FDR for a model described by the solution $\widehat{r}$ of system (1.8) is equivalent to the existence of a $\widehat{r}$-invariant parameterized family G for $\widehat{r}$. By Theorem 1.10, given a forward-rate model $\mathcal{M}$, we are looking for a sub-manifold $\mathrm{Im}[G] := \mathcal{G} \subset \widehat{\mathcal{H}}$ such that $\widehat{\mu}(G(z)), \widehat{\sigma}(G(z)) \in T_{G(z)}\mathcal{G}$, for each $G(z) \in U$, where U is a neighborhood of $\widehat{r}^M$ and $\widehat{r}^M \in \mathcal{G}$. In other words, we are looking for the tangential sub-manifold $\mathcal{G}$ of the distribution $F := \mathrm{Span}\{\widehat{\mu}, \widehat{\sigma}_1, \ldots, \widehat{\sigma}_d\}$, where we recall that:

Definition 1.14. Let F be a smooth distribution and let x_0 be a fixed point in $\mathcal{X}$, an $\mathcal{H}$-manifold. A submanifold $\mathcal{G} \subset \mathcal{X}$, with $x_0 \in \mathcal{G}$, is called *tangential manifold* through x_0 for F, if $F(x) \leq T_x\mathcal{G}$, $\forall x \in U$, where U is an open neighborhood of $x_0 \in \mathcal{G}$.

It can be proved that a tangential sub-manifold for a smooth distribution F exists if and only if F is involutive, i.e. if and only if the Lie-brackets between two vector fields contained in F is still in F. The Lie-Bracket of two vector fields v_1 and v_2 on $\widehat{\mathcal{H}}$ is defined by:

$$[v_1, v_2](\widehat{r}) := \partial_{\widehat{r}} v_1(\widehat{r}) v_2(\widehat{r}) - \partial_{\widehat{r}} v_2(\widehat{r}) v_1(\widehat{r}).$$

This result, that is an infinite-dimensional version of the Frobenius Theorem ([Lan12]), can be used to construct a tangential sub-manifold when the distribution F generated by $\widehat{\mu}$ and $\widehat{\sigma}$ is involutive. We state this result in the version proved in [BS01, Theorem 2.1]

Theorem 1.15 (Frobenius)**.** *Let F be a smooth distribution on the open set V on a Banach space $\mathcal{K}$. Let x be an arbitrary point in V . Then, there exists a diffeomorphism $\phi : U \to \mathcal{K}$ on some neighborhood $U \subseteq V$ of x such that the push-forward of F with respect to ϕ, denoted by $\phi_* F$, is constant on $\phi(U)$ if and only if F is involutive.*

We recall that if $\phi : A \to B$ and ξ a vector field on F, then $\phi_*\xi$ is a vector field on B, defined as $\phi_*\xi(\phi(x)) = \partial_x \phi(x)\xi(x)$.

However, in general F is not involutive, therefore we must consider the smallest involutive distribution which contains F. Such distribution is called *Lie algebra of F*. We recall the following result, proved in [BS01, Theorem 2.2]

Proposition 1.16. *We denote the Lie algebra generated by $\widehat{\mu}$ and $\widehat{\sigma}$ by $\mathcal{L} := \{\widehat{\mu}, \widehat{\sigma}_1, \ldots \widehat{\sigma}_d\}_{LA}$. Then, the existence of FDRs is equivalent to the existence of a finite-dimensional tangential sub-*

manifold. This is equivalent to:

$$dim[\mathcal{L}] = dim\{\widehat{\mu}, \widehat{\sigma}_1, \dots, \widehat{\sigma}_d\}_{LA} < +\infty. \tag{1.27}$$

Under condition (1.27), a FDR can be constructed. To this end, we provide a strategy based on [Bjö04, Section 5]. We proceed as follows:

S.1 Choose a finite number of vector fields $\xi_1, \dots \xi_n$, which span $\{\widehat{\mu}, \widehat{\sigma}_1, \dots \widehat{\sigma}_d\}_{LA}$;

S.2 Compute the invariant manifold

$$G(z_1, \dots, z_n) = e^{\xi_n z_n} \cdots e^{\xi_1 z_1} \widehat{r}^M, \tag{1.28}$$

where $e^{\xi_n z_n}$ denotes the integral curve of ξ_n at time z_n;

S.3 Define the state space process $Z \subseteq \mathbb{R}^n$, such that $\widehat{r} = G(Z)$. Z is defined by:

$$dZ_t = a(Z_t)dt + b(Z_t) \circ dW_t, \tag{1.29}$$

where

$$\begin{aligned} \partial_z G(z) a(z) &= \widehat{\mu}(G(z)), \\ \partial_z G(z) b_i(z) &= \widehat{\sigma}_i(G(z)), \quad i = 1, \dots, d. \end{aligned} \tag{1.30}$$

The uniqueness of a and b is guaranteed since G respects Assumption 1.7, then it is a local diffeomorphism. Therefore, there exists a unique vector field defined on $\mathcal{Z}$ a for $\widehat{\mu}$ and b_i for $\widehat{\sigma}_i$ for every $i = 1, \dots, d$ such that (1.30) is satisfied.

1.5.2 Example: constant volatility

We now examine the case in which the volatility vector field $\widehat{\sigma}(\widehat{r})$ is constant in $\widehat{r}$. Equivalently, this assumption means that $\sigma^0, \sigma^1, \dots, \sigma^m$ are constants in $\mathcal{H}^d$ and $\beta^1, \dots, \beta^m$ are constant on $\mathbb{R}^d$ (d is the dimension of the Brownian motion W driving the process $\widehat{r}$). First, in analogy with (1.9), the following holds:

$$\widehat{\mu}(\widehat{r}) = \begin{pmatrix} \mathbf{F}r^0 + \sigma^0 \cdot \mathbf{H}\sigma^0 \\ \mathbf{F}r^1 + \sigma^1 \cdot \mathbf{H}\sigma^1 - \beta^1 \cdot \sigma^1 \\ \vdots \\ \mathbf{F}r^m + \sigma^m \cdot \mathbf{H}\sigma^m - \beta^m \cdot \sigma^m \\ \mathbf{B}r^0 - \mathbf{B}r^1 - \frac{1}{2}||\beta^1||^2 \\ \vdots \\ \mathbf{B}r^0 - \mathbf{B}r^m - \frac{1}{2}||\beta^m||^2 \end{pmatrix}, \quad \widehat{\sigma}(\widehat{r}) = \begin{pmatrix} \sigma^0 \\ \vdots \\ \sigma^m \\ \beta^1 \\ \vdots \\ \beta^m \end{pmatrix}. \tag{1.31}$$

To determine the Lie algebra of $\mathrm{Span}\{\widehat{\mu}, \widehat{\sigma}_i,\ i = 1, \dots, d\}$ we compute the successive Lie brackets between $\widehat{\mu}$ and $\widehat{\sigma}$,

$$\widehat{\mu}, \widehat{\sigma}](\widehat{r}) = \partial_{\widehat{r}}\widehat{\mu}(\widehat{r})(\widehat{\sigma}(\widehat{r})) - \partial_{\widehat{r}}\widehat{\sigma}(\widehat{r})(\widehat{\mu}(\widehat{r})).$$

Then, while $\partial_{\widehat{r}}\widehat{\sigma}(\widehat{r}) = \mathbb{0}$, the Fréchet derivative of $\widehat{\mu}$ is:

$$\frac{\partial}{\partial \widehat{r}}\widehat{\mu}(\widehat{r}) = \begin{pmatrix} \mathbf{F} & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \mathbf{F} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \mathbf{F} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \mathbf{F} & 0 & \cdots & 0 \\ \mathbf{B} & -\mathbf{B} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \mathbf{B} & 0 & -\mathbf{B} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \mathbf{B} & 0 & 0 & \cdots & -\mathbf{B} & 0 & \cdots & 0 \end{pmatrix}. \tag{1.32}$$

The Lie bracket of $\widehat{\mu}$ and $\widehat{\sigma}_i$, for $i \in \{1, \dots, d\}$ is:

$$[\widehat{\mu}, \widehat{\sigma}_i](\widehat{r}) = \partial_{\widehat{r}}\widehat{\mu}(\widehat{r})\widehat{\sigma}_i(\widehat{r}) - \overbrace{\partial_{\widehat{r}}\widehat{\sigma}_i(\widehat{r})}^{=\mathbb{0}}\,\widehat{\mu}(\widehat{r}) = \begin{pmatrix} \mathbf{F}\sigma_i^0 \\ \mathbf{F}\sigma_i^1 \\ \vdots \\ \mathbf{F}\sigma_i^m \\ \mathbf{B}\sigma_i^0 - \mathbf{B}\sigma_i^1 \\ \vdots \\ \mathbf{B}\sigma_i^0 - \mathbf{B}\sigma_i^m \end{pmatrix},$$

that is constant on $\widehat{\mathcal{H}}$. As a consequence, in $\{\widehat{\mu}, \widehat{\sigma}\}_{LA}$, the only vector field which is not constant is $\widehat{\mu}$. Therefore, it is sufficient to compute the Lie bracket between $\widehat{\mu}$ and the successive Lie bracket between $\widehat{\mu}$ and $\widehat{\sigma}$. These computations lead to ([Lan19, Section 3.2.]:

$$\mathcal{L} := \{\widehat{\mu}, \widehat{\sigma}\}_{LA} := \mathrm{Span}(\mathcal{N}), \quad \mathcal{N} := \Big\{\widehat{\mu}, \widehat{\sigma}_1, \dots, \widehat{\sigma}_d, \nu_i^k \mid k \in \mathbb{N} \setminus \{0\}, \quad i = 1, \dots, d\Big\}, \tag{1.33}$$

where

$$\nu_i^k = \Big(\mathbf{F}^k\sigma_i^0, \quad \mathbf{F}^k\sigma_i^1, \quad \cdots, \quad \mathbf{F}^k\sigma_i^m, \quad \mathbf{B}\mathbf{F}^{k-1}\sigma_i^0 - \mathbf{B}\mathbf{F}^{k-1}\sigma_i^1, \quad \cdots, \quad \mathbf{B}\mathbf{F}^{k-1}\sigma_i^0 - \mathbf{B}\mathbf{F}^{k-1}\sigma_i^m\Big)^\top.$$

To find equivalent conditions for $\dim[\mathcal{L}] < +\infty$, we introduce the following concept:

Definition 1.17. A quasi-exponential function (QE) f is any function of the form:

$$f(x) := \sum_i e^{\gamma_i x} + \sum_j e^{\alpha_j x}[p_j(x)\cos\omega_j x + q_j(x)\sin\omega_j x],$$

where γ_i, α_j and ω_j are real numbers and p_i, q_j are real polynomials.

As in [BS01, Lemma 5.1.], QE functions can be charaterised as follows:

Lemma 1.18. *A function f is QE if and only if it is a component of the solution of a vector valued linear ODE with constant coefficients $f^{(n)} = \sum_{i=0}^{n-1} \gamma_i f^{(i)}$, where $f^{(i)}$ denotes the i-th order derivative of f.*

1.5.2.1 Existence of FDRs

We state the following result:

Proposition 1.19. *System* (1.8) *with constant volatility possesses FDRs if and only if $\sigma_i^j(x)$ are QE functions for each $i = 1, \ldots, d$ and $j = 0, \ldots, m$.*

Proof. Straightforward generalization of [Sli10, Proposition 3.2]. □

Remark 1.20. Under the condition of Theorem 1.19, the dimension of $\mathcal{L} := \{\widehat{\mu}, \widehat{\sigma}_1, \ldots, \widehat{\sigma}_d\}_{LA}$ is bounded from above by:

$$n := \dim[\mathcal{L}] \leq 1 + \sum_{i=1}^{d} (1 + n_i). \tag{1.34}$$

where n_i is the dimension of $N_i := \text{Span}\{\nu_i^k,\ k \in \mathbb{N} \setminus \{0\}\}$, for each $i = 1, \ldots, d$. Indeed, every component of ν_i^k is given by combinations of iterated derivatives in x-variable of the functions σ_i^j, for $j = 0, \ldots, m$. Then, if $n_i := \dim(N_i)$, by Lemma 1.18, there exists an annihilator polynomial

$$M_i(\gamma) := \sum_{h=0}^{n_i} \alpha_i^h \gamma^h \tag{1.35}$$

of degree n_i, such that $M_i(\mathbf{F})\sigma_i^j = 0$ for each $j = 0, \ldots, m$. In conclusion, the tangential manifold of dimension n is obtained by the composition of the integral curves of $\widehat{\mu}$, $\widehat{\sigma}_i$, ν_i^k, for $k = 0, \ldots, n_i$, $i = 1, \ldots, d$.

Notation 1.21. It is convenient to introduce the notation:

$$z := (z^0, z_1^0, \ldots, z_1^{n_1}, z_2^0, \ldots, z_2^{n_2}, \ldots, z_d^0, \ldots, z_d^{n_d}) \in \mathbb{R}^n.$$

In order to construct explicitly the FDRs, we apply the strategy outlined at the end of Subsection 1.5.1. By Theorem 1.19, it is sufficient to compute the integral curve of every vector field ξ that generates $\mathcal{N}$, defined in condition (1.33). Then, we compose the integral curves, as in **S.2** and we obtain the tangential manifold. Inverting the consistency condition as in **S.3**, the following proposition can be proved. The proof of this result is based on the direct generalization to the multi-curve setting of [BS01, Proposition 5.2].

Proposition 1.22. *A model $\mathcal{M}$, determined by $\widehat{\sigma}$ introduced in* (1.31)*, admits FDRs if and only if*

$\sigma_i^j(x)$ *are QE functions for each* $i = 1, \ldots, d$ *and* $j = 0, \ldots, m$. *In this case, we introduce the term*

$$S^j(x) := \left(\int_0^x \sigma_i^j(s)ds \right)_{i=1,\ldots,d}, \quad j = 0, \ldots, m \tag{1.36}$$

to describe the consistent parameterized family, that is:

$$\begin{aligned} G^j(x;z) =& r_j^M(x+z^0) + \sum_{i=1}^d \sum_{k=0}^{n_i} \boldsymbol{F}^k \sigma_i^j(x) z_i^k + \frac{1}{2}(||S^j(x+z^0)||^2 - ||S^j(x)||^2) + \\ & - (1-\delta_0^j) \sum_{i=1}^d \int_x^{x+z^0} \beta_i^j \sigma_i^j(s)ds, \quad j = 0, \ldots, m; \end{aligned} \tag{1.37}$$

$$\begin{aligned} G^{m+j}(z) =& \sum_{i=1}^d \sum_{k=1}^{n_i} (\boldsymbol{B}\boldsymbol{F}^{k-1}\sigma_i^0 - \boldsymbol{B}\boldsymbol{F}^{k-1}\sigma_i^j) z_i^k + \sum_{i=1}^d \beta_i^j z_i^0 + y_j^M + \int_0^{z^0} \Big(r_0^M(s) + \\ & - r_j^M(s) \Big) ds + \frac{1}{2} \int_0^{z^0} \Big[||S^0(s)||^2 - ||S^j(s)||^2 \Big] ds + \\ & + \sum_{i=1}^d \int_0^{z^0} \beta_i^j S_i^j(s) ds - \frac{1}{2} ||\beta^j||^2 z^0, \quad j = 1, \ldots, m. \end{aligned}$$

Moreover, the coefficients of the $\mathbb{R}^n$*-valued process introduced in* (1.29) *are:*

$$\begin{cases} a^0 = 1, & \\ a_i^0 = 0, & i = 1, \ldots, d, \\ a_i^k = z_i^{k-1} + z_i^{n_i} \alpha_i^k, & k = 1, \ldots, n_i;\ i = 1, \ldots, d, \\ b_i^0 = 0; & \\ b_{i,i}^0 = 1; & \\ b_{h,i}^k = 0, & h = 1, \ldots, d,\ h \neq i;\ k = 1, \ldots, n_i. \end{cases} \tag{1.38}$$

where we adopted the notation introduced in Notation 1.21. α_i^k *is the* $k-th$ *coefficients of the annihilator polynomial* M_i *introduced in equation* (1.35), *for every* $i = 1, \ldots, d$.

1.5.3 Example: constant direction volatility

In this section we generalize what we proved in Section 1.5.2, through an analogy with the results of [Sli10, Section 4]. In particular, we establish conditions that guarantee the existence of FDRs for a model $\mathcal{M}$ determined by a volatility term, defined for any $i = 1, \ldots, d$, by:

$$\widehat{\sigma}_i(\widehat{r}, x) = \Big(\varphi_i^0(\widehat{r})\lambda_i^0(x), \quad \varphi_i^1(\widehat{r})\lambda_i^1(x), \quad \cdots, \quad \varphi_i^m(\widehat{r})\lambda_i^m(x), \quad \beta_i^1(\widehat{r}), \quad \cdots, \quad \beta_i^m(\widehat{r}) \Big), \tag{1.39}$$

where $\lambda_i^j(x) \in \mathcal{H}$, while φ_i^j and β_i^j are smooth (in Fréchét sense) scalar vector fields defined on $\widehat{\mathcal{H}}$ for each $i = 1, \dots, d$ and $j = 0, \dots, m$. We suppose that:

Assumption 1.23. For each $i = 1, \dots, d$ and $j = 0, \dots, m$, $\varphi_i^j(\widehat{r}) \neq 0$ and $\beta_i^j(\widehat{r}) \neq 0$ for each $\widehat{r} \in \widehat{\mathcal{H}}$.

In what follows we characterize the HJM drift condition (1.9) for $\widehat{\mu}$ when $\widehat{\sigma}$ is like (1.39). First of all, we introduce the following notation for the volatility:

$$\widehat{\sigma}(\widehat{r}) = \begin{pmatrix} \varphi_1^0(\widehat{r})\lambda_1^0(x) & \cdots & \varphi_d^0(\widehat{r})\lambda_d^0(x) \\ \varphi_1^1(\widehat{r})\lambda_1^1(x) & \cdots & \varphi_d^1(\widehat{r})\lambda_d^1(x) \\ \vdots & & \vdots \\ \varphi_1^m(\widehat{r})\lambda_1^m(x) & \cdots & \varphi_d^m(\widehat{r})\lambda_d^m(x) \\ \beta_1^1(\widehat{r}_t) & \cdots & \beta_d^1(\widehat{r}_t) \\ \vdots & & \vdots \\ \beta_1^m(\widehat{r}_t) & \cdots & \beta_d^m(\widehat{r}_t) \end{pmatrix} = \begin{pmatrix} \sigma^0 \\ \vdots \\ \sigma^m \\ \beta^1 \\ \vdots \\ \beta^m \end{pmatrix}. \tag{1.40}$$

To simplify the notation, in the following we omit the x-variable. To reproduce the Stratonovich dynamics of $\widehat{r}$, we notice that, for each $j = 0, \dots, m$:

$$\partial_{\widehat{r}}\widehat{\sigma}(\widehat{r}_t)\widehat{\sigma}(\widehat{r}_t) = \sum_{i=1}^{d}\left(\sum_{h=0}^{m} \lambda_i^j \partial_{r^h}\varphi_i^j(\widehat{r}_t)\varphi_i^h(\widehat{r}_t)\lambda_i^h + \sum_{h=1}^{m}\lambda_i^j\partial_{Y^h}\varphi_i^j(\widehat{r}_t)\beta_i^h(\widehat{r}_t)\right),$$

$$\sigma^j(\widehat{r}_t)\cdot \mathbf{H}\sigma^j(\widehat{r}_t) = (\varphi^j(\widehat{r}_t)\lambda^j)\cdot \int_0^{\cdot}(\varphi^j(\widehat{r}_t)\lambda^j(s))ds = \sum_{i=1}^{d}(\varphi_i^j(\widehat{r}_t))^2\lambda_i^j\int_0^{\cdot}\lambda_i^j(s)ds.$$

We denote the Fréchet derivative of φ_i^j on the variable r^h computed on $\widehat{r}$ acting on the vector λ_i^h by $\partial_{r^h}\varphi_i^j(\widehat{r})[\lambda_i^h]$, for each $h, j = 0, \dots, m$ and $i = 1, \dots, d$. We also define:

$$D_i^j(x) := \lambda_i^j(x)\int_0^x \lambda_i^j(s)ds, \quad j = 0, \dots, m, \; i = 1, \dots, d. \tag{1.41}$$

Therefore, in the Stratonovich form, the dynamics of each equation of system (1.6) is given by:

$$\begin{aligned} dr_t^j = &\left[\mathbf{F}r^j + \sum_{i=1}^{d}(\varphi_i^j(\widehat{r}_t))^2 D_i^j - \frac{1}{2}\sum_{i=1}^{d}\lambda_i^j\left(\sum_{h=0}^{m}\varphi_i^h(\widehat{r}_t)\partial_{r^h}\varphi_i^j(\widehat{r}_t)[\lambda_i^h] + \sum_{h=1}^{m}\partial_{Y^h}\varphi_i^j(\widehat{r}_t)\beta_i^h(\widehat{r}_t) + \right.\right. \\ &\left.\left. + 2(1-\delta_j^0)\varphi_i^j(\widehat{r}_t)\beta_i^j(\widehat{r}_t)\right)\right]dt + \sum_{i=1}^{d}\varphi_i^j(\widehat{r}_t)\lambda_i^j(x)\circ dW_t, \quad j = 0, \dots, m, \\ dY_t^j = &\left\{\mathbf{B}r^0 - \mathbf{B}r^j - \frac{1}{2}\sum_{i=1}^{d}(\beta_i^j(\widehat{r}_t))^2 - \frac{1}{2}\left[\sum_{i=1}^{d}\sum_{h=0}^{m}\partial_{r^h}\beta_i^j(\widehat{r}_t)[\lambda_i^h]\varphi_i^h(\widehat{r}_t) + \sum_{h=1}^{m}\partial_{Y^h}\beta_i^j(\widehat{r}_t)\beta_i^h(\widehat{r}_t)\right]\right\}dt + \\ & + \sum_{i=1}^{d}\beta_h^j(\widehat{r}_t)\circ dW_t, \quad j = 1, \dots, m, \end{aligned}$$

where δ_j^0 denotes the Kronecker delta between the indexes 0 and j. We aim at conditions that guarantee $\dim[\{\widehat{\mu}, \widehat{\sigma}_1, \ldots, \widehat{\sigma}_d\}_{LA}] < \infty$, where:

$$\widehat{\mu}(\widehat{r}) = \begin{pmatrix} \mathbf{F}r^0 + \sum_{i=1}^d (\varphi_i^0(\widehat{r}))^2 D_i^0 - \frac{1}{2}\sum_{i=1}^d \lambda_i^0 \Big(\sum_{h=0}^m \varphi_i^h(\widehat{r}) \partial_{r^h} \varphi_i^0(\widehat{r})[\lambda_i^h] + \\ + \sum_{h=1}^m \partial_{Y^h} \varphi_i^0(\widehat{r}) \beta_i^h(\widehat{r}) \Big) \\ \mathbf{F}r^1 + \sum_{i=1}^d (\varphi_i^1(\widehat{r}))^2 D_i^1 - \frac{1}{2}\sum_{i=1}^d \lambda_i^1 \Big(\sum_{h=0}^m \varphi_i^h(\widehat{r}) \partial_{r^h} \varphi_i^1(\widehat{r})[\lambda_i^h] + \\ + \sum_{j=1}^m \partial_{Y^h} \varphi_i^1(\widehat{r}) \beta_i^h(\widehat{r}) + 2\varphi_i^1(\widehat{r})\beta_i^1(\widehat{r}) \Big) \\ \vdots \\ \mathbf{F}r^m + \sum_{i=1}^d (\varphi_i^m(\widehat{r}))^2 D_i^m - \frac{1}{2}\sum_{i=1}^d \lambda_i^m \Big(\sum_{h=0}^m \varphi_i^h(\widehat{r}) \partial_{r^h} \varphi_i^m(\widehat{r})[\lambda_i^h] + \\ + \sum_{h=1}^m \partial_{Y^h} \varphi_i^m(\widehat{r}) \beta_i^h(\widehat{r}) + 2\varphi_i^m(\widehat{r})\beta_i^m(\widehat{r}) \Big) \\ \mathbf{B}r^0 - \mathbf{B}r^1 - \frac{1}{2}\sum_{i=1}^d (\beta_i^1(\widehat{r}))^2 - \frac{1}{2}\Big[\sum_{i=1}^d \Big(\sum_{h=0}^m \partial_{r^h} \beta_i^1(\widehat{r})[\lambda_i^h] \varphi_i^h(\widehat{r}) + \sum_{h=1}^m \partial_{Y^h} \beta_i^1(\widehat{r}) \beta_i^h(\widehat{r}) \Big) \Big] \\ \vdots \\ \mathbf{B}r^0 - \mathbf{B}r^m - \frac{1}{2}\sum_{i=1}^d (\beta_i^m(\widehat{r}))^2 - \frac{1}{2}\Big[\sum_{i=1}^d \Big(\sum_{h=0}^m \partial_{r^h} \beta_i^m(\widehat{r})[\lambda_i^h] \varphi_i^h(\widehat{r}) + \sum_{h=1}^m \partial_{Y^h} \beta_i^m(\widehat{r}) \beta_i^h(\widehat{r}) \Big) \Big] \end{pmatrix}, \tag{1.42}$$

and $\widehat{\sigma}$ as in equation (1.39).

1.5.3.1 Existence of FDRs

Differently from Section 1.5.2, it is more difficult to compute the integral curve of $\widehat{\mu}$ cause the more complex structure of the drift term. In order to overcome this problem, we provide conditions such that a larger distribution than $\{\widehat{\mu}, \widehat{\sigma}_1, \ldots, \widehat{\sigma}_d\}_{LA}$ is finite-dimensional. As a consequence, we determine a sufficient condition for the existence of FDRs.

Denoting by E_j is the j^{th} element of the canonical basis in $\widehat{\mathcal{H}}$ for $j = 0, \ldots, 2m$, we introduce the following set of vector fields:

$$\mathcal{N} := \{\xi^0, \xi_i^h, \eta_i^h, \gamma_k |\ h = 0, \ldots, m,\ i = 1, \ldots, d,\ k = 1, \ldots, m\}, \tag{1.43}$$

where

$$\xi^0 = \begin{pmatrix} \mathbf{F}r^0 \\ \mathbf{F}r^1 \\ \vdots \\ \mathbf{F}r^m \\ \mathbf{B}r^0 - \mathbf{B}r^1 \\ \vdots \\ \mathbf{B}r^0 - \mathbf{B}r^m \end{pmatrix}, \quad \xi_i^j = \lambda_i^j E_j, \quad \eta_i^j = D_i^j E_j, \quad \gamma_k = E_{m+k}.$$

Then:

$$\widehat{\mu}(\widehat{r}) = \xi^0 + \sum_{h=0}^m \sum_{i=1}^d \Big((\varphi_i^h(\widehat{r}))^2 \eta_i^h - \kappa_i^h \xi_i^h \Big) - \sum_{h=1}^m \zeta^h \gamma_h, \tag{1.44}$$

$$\widehat{\sigma}_i(\widehat{r}) = \sum_{h=0}^{m} \varphi_i^h(\widehat{r})\xi_i^h + \sum_{h=1}^{m} \beta_i^h(\widehat{r})\gamma_h, \quad i = \ldots, d, \tag{1.45}$$

where

$$\kappa_i^j = \frac{1}{2}\Bigg\{\sum_{h=0}^{m} \varphi_i^h(\widehat{r})\partial_{r^h}\varphi_i^j(\widehat{r})[\lambda_i^h] + \sum_{h=1}^{m}\Bigg[\partial_{Y^h}\varphi_i^j(\widehat{r})\beta_i^h(\widehat{r})\Bigg] + 2(1-\delta_j^0)\varphi_i^j(\widehat{r})\beta_i^j(\widehat{r})\Bigg\},$$
$$\zeta^j = \frac{1}{2}\Bigg\{\sum_{i=1}^{d}(\beta_i^j(\widehat{r}))^2 + \sum_{i=1}^{d}\Bigg(\sum_{h=0}^{m}\partial_{r^h}\beta_i^j(\widehat{r})[\lambda_i^h]\varphi_i^h(\widehat{r}) + \sum_{h=1}^{m}\partial_{Y^h}\beta_i^j(\widehat{r})\beta_i^h(\widehat{r})\Bigg)\Bigg].$$

Conditions (1.44) and (1.45) imply that $\mathcal{L} \subseteq \mathcal{L}^1$, where:

$$\begin{aligned}\mathcal{L} &:= \{\widehat{\mu}, \widehat{\sigma}_i,\ i = 1, \ldots, d\}_{LA},\\ \mathcal{L}^1 &:= \{\xi^0, \xi_i^j, \eta_i^j, \gamma_k |\ j = 0, \ldots, m,\ i = 1, \ldots, d,\ k = 1, \ldots, m\}_{LA}.\end{aligned} \tag{1.46}$$

Therefore, if $\mathcal{L}^1$ is finite-dimensional, also $\mathcal{L}$ is finite-dimensional. Hence, the following result holds:

Proposition 1.24. *If $\lambda_i^j(x)$ is a QE function for each $j = 0, \ldots, m$ and $i = 1, \ldots, d$, then the Lie algebra $\mathcal{L}^1$ is finite-dimensional.*

Proof. Straightforward generalization to [Sli10, Proposition 4.2]. □

Then, our purpose is the description of FDRs when the volatility is given by (1.39). To guarantee the existence of FDRs we assume that:

Assumption 1.25. Every function $\lambda_i^j(x)$ is QE for each $i = 1, \ldots, d$ and $j = 0, \ldots, m$.

Under this assumption, it is straightforward to verify that also the functions $D_i^j(x)$, introduced in equation (1.41) are QE. Then, by Lemma 1.18, there exists natural numbers n_i^j and p_i^j for each $i = 1, \ldots, d$ and $j = 0, \ldots, m$ such that

$$\begin{aligned}\mathbf{F}^{n_i^j}\lambda_i^j &= \sum_{k=0}^{n_i^j-1} c_{k,i}^j \mathbf{F}^k \lambda_i^j,\\ \mathbf{F}^{p_i^j} D_i^j &= \sum_{k=0}^{p_i^j-1} d_{k,i}^j \mathbf{F}^k D_i^j\end{aligned} \tag{1.47}$$

are satisfied for suitable real constants $c_{k,i}^j$ and $d_{k,i}^j$. In this case, by the definition of $\mathcal{L}^1$, the dimension of the Lie-algebra $\mathcal{L}^1$ is bounded from above by

$$n := m + 1 + \sum_{i=1}^{d}\sum_{j=0}^{m}(n_i^j + p_i^j). \tag{1.48}$$

In order to build an invariant manifold we introduce the following notation for a vector of the state space $z \in \mathbb{R}^n$:

Notation 1.26. A vector $z \in \mathbb{R}^n$ is denoted by the concatenation of the vectors $(x^j)_j \in \mathbb{R}^{m+1}$, $(z^j_{k,i})_{i,j,k} \in \mathbb{R}^{\sum_{i=1}^d \sum_{j=0}^m n_i^j}$ and $(x^j_{k,i}) \in \mathbb{R}^{\sum_{i=1}^d \sum_{j=0}^m p_i^j}$, i.e

$$z = (x^0, \dots, x^m, z^0_{0,1}, z^0_{1,1}, \dots, z^0_{n^1_0-1,1}, z^1_{0,1}, \dots, z^m_{n^m_1-1,d}, x^0_{0,1}, \dots, x^0_{p^1_0-1,1}, \dots, x^m_{0,1}, \dots, x^m_{p^m_1-1,d}).$$

As in Subsection 1.5.2.1, we construct the tangential manifold of $\mathcal{L}^1$:

$$G(z) := \prod_{i,j,k,h,l} e^{\mathbf{F}^k \lambda_i^j E_j z^j_{k,i}} e^{\mathbf{F}^h D_i^j E_j x^j_{h,i}} e^{\gamma_l x^l} e^{\xi^0 x^0} \widehat{r}^M, \tag{1.49}$$

for an arbitrary point $\widehat{r}^M \in \widehat{\mathcal{H}}$. Every component of such function G are given by:

$$\begin{aligned} G^j(z,x) &= r^M_j(x^0+x) + \sum_{i=1}^d \left\{ \sum_{k=0}^{n_i^j-1} z^j_{k,i} \mathbf{F}^k \lambda_i^j(x) + \sum_{k=0}^{p_i^j-1} x^j_{k,i} \mathbf{F}^k \left(\lambda_i^j(x) \int_0^x \lambda_i^j(s) ds \right) \right\}, \quad j = 0, \dots, m, \\ G^j(z,x^j) &= y^M_j + \int_0^{x^0} (r^M_0(s) - r^M_j(s)) ds + x^j, \quad j = m+1, \dots, 2m. \end{aligned} \tag{1.50}$$

Hence, we determine the coefficients a and b of a finite dimensional process as in equation (1.29) such that $\partial_z G a = \widehat{\mu}$ and $\partial_z G b_i = \widehat{\sigma}_i$ for each $i = 1, \dots, d$. We omit the z variable on the functions a and b and we use for those functions a notation analogous with the one introduced in (1.26):

$$\begin{aligned} a =& (a^0, \dots, a^m, a^0_{0,1}, a^0_{1,1}, \dots, a^m_{n^m_d-1,d}, \widetilde{a}^0_{0,1}, \dots, \widetilde{a}^0_{p^1_0-1,1}, \dots, \widetilde{a}^m_{0,1}, \dots, \widetilde{a}^m_{p^m_1-1,1}), \\ b =& (b_1, \dots, b_d)^\top, \end{aligned}$$

where for any $h = 1, \dots, d$, b_h is given by:

$$b = (b^0_h, \dots, b^m_h, b^0_{0,1,h}, \dots, b^0_{n^1_0-1,1,h}, b^1_{0,1,h}, \dots, b^1_{n^1_1-1,1,h}, \widetilde{b}^0_{0,1,h}, \dots, \widetilde{b}^0_{p^1_0-1,1,h}, \dots, \widetilde{b}^m_{0,1,h}, \dots, \widetilde{b}^m_{p^m_1-1,1,h}).$$

To determine the coefficients a and b, we must invert the consistency condition between the coefficients of the model and tangential manifold, as described in **S.3**. Then, we obtain:

$$\begin{cases} a^0 &= 1, \\ a^j_{0,i} &= z^j_{n^j_i-1,i} c^j_{0,i} - \frac{1}{2}\Big(\sum_{h=0}^m \varphi^h_i(G(z))\partial_{r^h}\varphi^j_i(G(z))[\lambda^j_i] + \\ & \quad + \sum_{h=1}^m \beta^h_i(G(z))\partial_{Y^h}\varphi^j_i(G(z)) + (1-\delta^j_0)2\varphi^j_i(G(z))\beta^j_i(G(z))\Big),\ j=0,\dots,m, \\ a^j_{k,i} &= z^j_{k-1,i} + z^j_{n^j_i-1,i} c^j_{k,i}, \quad k=1,\dots,n^j_i-1, j=0,\dots,m \\ \widetilde{a}^j_{0,i} &= (\varphi^j_i(G(z)))^2 + x^j_{p^j_i-1,i} d^j_{0i},\ j=0,\dots,m, \\ \widetilde{a}^j_{k,i} &= x^j_{k-1,i} + x^j_{p^j_i-1,i} d^j_{k,i} \quad k=1,\dots,p^j_i-1,\ j=0,\dots,m, \\ a^j &= \sum_{i=1}^d \Big[\sum_{k=0}^{n^0_i-1} z^0_{k,i}\mathbf{F}^k\lambda^0_i(0) - \sum_{k=0}^{n^j_i-1} z^j_{k,i}\mathbf{F}^k\lambda^j_i(0) - \frac{1}{2}(\beta^j_i(G(z)))^2 + \\ & \quad -\frac{1}{2}\Big(\sum_{h=0}^m \partial_{r^h}\beta^j_i G(z))[\lambda^h_i]\varphi^h_i(G(z)) + \sum_{h=1}^m \partial_{Y^h}\beta^j_i(G(z))\beta^h_i(G(z))\Big)\Big],\ j=1,\dots,m. \end{cases} \tag{1.51}$$

We follow the same procedure in order to compute the value of b. We get, for any $h=1,\dots,d$:

$$\begin{cases} b^0_h &= 0, \\ b^j_{0,h,h} &= \varphi^j_h(G(z)), \\ b^j_{0,i,h} &= 0, \quad i \neq h, \quad i=1,\dots,d, \\ b^j_{k,i,h} &= 0, \quad k=1,\dots,n^j_i-1, \\ \widetilde{b}^j_{k,i,h} &= 0, \quad k=0,\dots,p^j_i-1, \\ b^0 &= 0, \\ b^j_h &= \beta^j_h(G(z)), \qquad j=1,\dots,m. \end{cases} \tag{1.52}$$

In conclusion, under Assumption 1.25, the FDRs of the model $\mathcal{M}$, defined by $\widehat{\sigma}$ in (1.39), are determined by the immersion G, defined in (1.50) and the finite-dimensional process (1.29) whose drift and volatility terms are expressed by (1.51) and (1.52) respectively.

Remark 1.27. Assumption 1.25 is a sufficient condition, that guarantee the existence of FDRs for the multi-curve $\mathcal{M}$ determined by the volatility term introduced in (1.39). However, if for each $j=0,\dots,m$, $\varphi^j_i = \varphi_i$ for any $i=1,\dots,m$, equivalent conditions for the existence of FDRs can be provided.

1.5.3.1.1 Example: Hull White model with non constant volatility for the spread processes We consider a model $\mathcal{M}$ in which the volatility has constant direction as a vector field in $\widehat{\mathcal{H}}$:

$$\widehat{\sigma}(\widehat{r}) = \begin{pmatrix} \sigma^0 e^{-a^0 x} & 0 & 0 \\ 0 & \sigma^1 e^{-a^1 x} & 0 \\ 0 & 0 & \sigma^2 e^{-a^2 x} \\ \beta^1_1 & \beta^1_2 Y^1_t & 0 \\ \beta^2_1 & 0 & \beta^2_3 Y^2_t \end{pmatrix}, \tag{1.53}$$

where β_i^j, σ^j, a^j are positive constants for $j = 0, 1, 2$ and $i = 1, 2, 3$. $\mathcal{M}$ is then driven by a 3-dimensional Brownian motion $W := (W^1, W^2, W^3)$. The choice of this structure for the volatility term is motivated by the following reasoning: higher values of the spread processes are linked with instability of the market, because the spreads describe the risk captured by RSRs. Therefore, we can suppose that the volatility of Y^j is proportional to its value.

Let us notice that, by $\widehat{\sigma}$ as in (1.53) we establish a correlation between the forward-curves and the spread components.

We aim at computing the FDRs for $\mathcal{M}$. We assume that the model $\mathcal{M}$ starts at an initial point $(r_0^M, r_1^M, r_2^M, y_1^M, y_2^M) \in \widehat{\mathcal{H}}$ such that $y_1^M, y_2^M \neq 0$. Then, by the continuity of Y_t^j, we can assume that $Y_t^j \neq 0$ for each $t \in [0, \tau)$, with $\tau > 0$ a.s.. Under this supposition, the Lie algebra generated by the coefficients of the model $\mathcal{M}$ is finite-dimensional (because Assumption 1.23 holds). As in (1.43), we introduce $\mathcal{N} := \{\xi^0, \xi_i^j, \eta_i^j, \gamma_k | \ j = 0, \ldots, 2, \ i = 1, \ldots, 3, \ k = 1, \ldots, 2\}$. Now, we can apply Proposition 1.24, because the forward-rate components of the volatility term are QE functions. As a consequence,

$$\mathcal{L}^1 = \text{Span}\Big\{\xi^0, \ (\mathbf{F}^n \lambda_i^j) E_j, \ (\mathbf{F}^n D_i^j) E_j, \ \gamma_k \mid j = 0, \ldots, 2, \ i = 1, \ldots, 3 \ k = 1, \ldots, 2, \ n \in \mathbb{N}\Big\}$$

is finite-dimensional and contains $\{\widehat{\mu}, \widehat{\sigma}_1, \widehat{\sigma}_2, \widehat{\sigma}_3\}_{LA}$, with $\widehat{\sigma}$ as (1.53). To construct the FDRs, we first need to compute the dimension of $\text{Span}\{(\mathbf{F}^n \lambda_i^j) E_j, \ (\mathbf{F}^n D_i^j) E_j \mid n \in \mathbb{N}\}$. Due to (1.24), for each $i = 1, 2, 3$ the polynomial annihilator M_i^j of σ_i^j is:

$$\begin{aligned} j \neq i-1: \quad & \sigma_i^j(x) = 0 \Rightarrow M_i^j(\gamma) = \gamma, \\ j = i-1: \quad & \sigma_i^j(x) = \sigma^j e^{-a^j x} \Rightarrow M_i^j(\gamma) = \gamma + a^j, \end{aligned} \tag{1.54}$$

Analogously, we introduce the term D_i^j analogous with (1.41). It is different from zero if and only if $j = i - 1$. In this case, it is given by

$$D_i^j(x) = \lambda_i^j(x) \int_0^x \lambda_i^j(s) ds = e^{-a^j x} \Big(\frac{1 - e^{-a^j x}}{a^j} \Big), \quad j = 0, \ldots, m.$$

Then:

$$\begin{aligned} j \neq i-1: \quad & D_i^j(x) = 0 \\ j = i-1: \quad & D_i^j(x) = \frac{(\sigma^j)^2}{a^j} \Big(e^{-a^j x} - e^{-2a^j x} \Big) \end{aligned} \tag{1.55}$$

Therefore, the minimal annihilator of the function $\lambda_i^{i-1}(x)$ is $M_i^{i-1}(\lambda) := \lambda + a^{i-1}$. In the other cases $M_i^j(\lambda) = 0$. For what regards the minimal annihilator of D_i^{i-1}, we notice that the function

$$f(x) = \frac{e^{-\alpha x} - e^{-2\alpha x}}{\alpha},$$

with the same structure of D_i^{i-1}, satisfies $f'(x) = -e^{-\alpha x} + 2e^{-2\alpha x}$ and $f''(x) = -4\alpha e^{-2\alpha x} + \alpha e^{-\alpha x}$. Hence, the minimal annihilator of $D_i^{i-1}(x)$ has degree 2 for each j. Indeed, if there was a non-zero solution of

$$af(x) + bf'(x) = 0 \quad \Leftrightarrow \quad \Big(\frac{a}{\alpha} - b\Big)e^{-\alpha x} + \Big(-\frac{a}{\alpha} + 2b\Big)e^{-2\alpha x} = 0,$$

the coefficients a, b and α would satisfy

$$\begin{cases} \frac{a}{\alpha} - b = 0, \\ -\frac{a}{\alpha} + 2b = 0. \end{cases}$$

However, this system cannot be solved. Therefore, we consider the equation $af(x) + bf'(x) + cf''(x) = 0$, $\forall x \in \mathbb{R}_+$, that is

$$\Big(\frac{a}{\alpha} - b + \alpha c\Big)e^{-\alpha x} + \Big(-\frac{a}{\alpha} + 2b - 4\alpha c\Big)e^{-2\alpha x} = 0.$$

The solution of this equation is $a = 2\alpha^2 c$ and $b = 3\alpha c$.

In conclusion, the minimal annihilator of $D_i^{i-1}(x)$ is

$$P_i^{i-1}(\lambda) := \lambda^2 + 3a^{i-1}\lambda + 2(a^{i-1})^2, \quad \forall j = 0, \dots, m.$$

Then, for $j \neq i-1$, the polynomial annihilator of D_i^j is $P_i^j(\lambda) = \lambda$, while for each $i = 1, 2, 3$ it is $P_i^{i-1}(\lambda) = \lambda^2 + 3a^{i-1}\lambda + 2(a^{i-1})^2$. In conclusion:

$$\begin{aligned} n_i^j &:= \deg(M_i^j) = 1, \quad j = 0, 1, 2, \ i = 1, 2, 3; \\ p_i^j &:= \deg(P_i^j) = \begin{cases} 2, & j = i-1, \ i = 1, 2, 3; \\ 1, & \text{otherwise} \end{cases} \end{aligned}$$

We denote a vector of the state-space $z \in \mathbb{R}^n$ with $n := 2 + 1 + \sum_{i=1}^3 (n_i^{i-1} + p_i^{i-1}) = 12$ by

$$z = (x^0, x^1, x^2, z^0, z^1, z^2, x_0^0, x_1^0, x_0^1, x_1^1, x_0^2, x_1^2),$$

in analogy with Notation 1.26. In general, the tangential manifold is the image of the function introduced in equation (1.49), that in this case is:

$$G(z) = \prod_{i,h,l} e^{\sigma_i^{i-1} E_{i-1} z^{i-1}} e^{\mathbf{F}^h D_i^{i-1} E_{i-1} x_h^{i-1}} e^{\gamma_l x^l} e^{\xi^0 w^0} \widehat{r}^M,$$

Developing the computations, we obtain that the FDRs are explicitly given by:

$$G^j(z, x) = r_j^M(x^0 + x) + z^j \sigma^j e^{-a^j x} + \frac{(\sigma^j)^2}{a^j} e^{-a^j x}\Big[x_0^j\Big(1 - e^{-a^j x}\Big) + x_1^j a^j \Big(-1 + 2e^{-a^j x}\Big)\Big], \ j = 0, 1, 2,$$

$$G^{2+j}(z) = y_j^M + \int_0^{x^0} (r_0^M(s) - r_j^M(s))ds + x^j, \quad j = 1, 2.$$

The state process $Z_t = (X^0_t, X^1_t, X^2_t, Z^0_t, Z^1_t, Z^2_t, X^0_{0,t}, X^0_{1,t}, X^1_{0,t}, X^1_{1,t}, X^2_{0,t}, X^2_{1,t})$ satisfies (1.29), where:

$$a(Z_t) = \begin{pmatrix} 1 \\ Z^0_t\sigma^0 + (\sigma^0)^2 X^0_{1,t} - Z^1_t\sigma^1 - (\sigma^1)^2 X^1_{1,t} - \frac{1}{2}(\beta^1_2)^2\Big(y^M_1 + \int_0^{X^0_t}(r^M_0(s) - r^M_1(s))ds + \\ +X^1_t\Big)\Big(y^M_1 + \int_0^{X^0_t}(r^M_0(s) - r^M_1(s))ds + X^1_t + 1\Big) - \frac{1}{2}(\beta^1_1)^2 \\ Z^0_t\sigma^0 + (\sigma^0)^2 X^0_{1,t} - Z^2_{0t}\sigma^2 - (\sigma^2)^2 X^2_{1t} - \frac{1}{2}(\beta^2_3)^2\Big(y^M_2 + \int_0^{X^0_t}(r^M_0(s) - r^M_2(s))ds + \\ +X^2_t\Big)\Big(y^M_2 + \int_0^{X^0_t}(r^M_0(s) - r^M_2(s))ds + X^2_t + 1\Big) - \frac{1}{2}(\beta^2_2)^2 \\ -a^0 Z^0_t \\ -Z^1_t a^1 - \beta^1_2\Big(y^M_1 + \int_0^{X^0_t}(r^M_0(s) - r^M_1(s))ds + X^1_t\Big) \\ -Z^2_t a^2 - \beta^2_3\Big(y^M_2 + \int_0^{X^0_t}(r^M_0(s) - r^M_2(s))ds + X^2_t\Big) \\ -2(a^0)^2 X^0_{1t} + 1 \\ X^0_{0,t} - 3a^0 X^0_{1,t} \\ -2(a^1)^2 X^1_{1,t} + 1 \\ X^1_{0,t} - 3a^1 X^1_{1,t} \\ -2(a^2)^2 X^2_{1,t} + 1 \\ X^2_{0,t} - 3a^2 X^2_{1,t} \end{pmatrix} \tag{1.56}$$

and

$$b(Z_t) = \begin{pmatrix} 0 & 0 & 0 \\ \beta^1_1 & \beta^1_2\Big(y^M_1 + \int_0^{X^0_t}(r^M_0(s) - r^M_1(s))ds + X^1_t\Big) & 0 \\ \beta^2_1 & 0 & \beta^2_3\Big(y^M_2 + \int_0^{X^0_t}(r^M_0(s) - r^M_2(s))ds + X^2_t\Big) \\ \sigma^0 & 0 & 0 \\ 0 & \sigma^1 & 0 \\ 0 & 0 & \sigma^2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{1.57}$$

1.5.4 Realizations through a benchmark of maturities

Let us consider a model $\mathcal{M}$ driven by equation (1.8) that admits FDRs of dimension n. This implies that a n-dimensional tangential manifold $\mathcal{G} := \mathrm{Im}[G]$ exists. In this setting, we answer the following question:

"Is it possible to construct another FDRs of the same dimension n, determined by the spread processes Y^j and the forward-rate components applied to a fixed set of benchmark maturities?". In other, words, if we consider a point $\widehat{r} \in \mathcal{G} \subseteq \widehat{\mathcal{H}}$ is it possible to construct the FDRs of $\mathcal{M}$ through

a fixed benchmark of maturities $x = (x_1, \dots, x_n)$ and the linear function:

$$Z^h(r, y) := \sum_{j=0}^{m} a_j^h r^j(x_h) + \sum_{j=1}^{m} b_j^h y^j, \quad h = 1, \dots, n, \tag{1.58}$$

defined for a suitable set of real constants a_j^h, b_j^h. This is equivalent to show that $Z : \widehat{\mathcal{H}} \to \mathbb{R}^n$ is a local system of coordinates for $\mathcal{G}$ for an opportune choice of the constants a_j^h, b_j^h.

To answer this question, we present the following result:

Theorem 1.28. *Let us consider a model $\mathcal{M}$, determined by the dynamics introduced in system* (1.8)*, that admits FDRs of dimension n. Then, for any vector $x := (x_1, \dots, x_n) \in \mathbb{R}^n$, where x_i is chosen freely except for a discrete set of $\mathbb{R}_+$, the FDRs of $\mathcal{M}$ can be described by the inverse of the function introduced in equation* (1.58)*, for suitable constants a_j^h, b_j^h.*

Proof. We must prove that function (1.58) is a diffeomorphism between $\mathcal{G}$ and its image. To prove such result we denote the Fréchet derivative of Z (introduced in equation (1.58)) by

$$\partial_{\widehat{r}} Z : T_{\widehat{r}}\mathcal{G} \to \mathbb{R}^n,$$

where $T_{\widehat{r}}\mathcal{G}$ is the tangent space at $\widehat{r}$ of $\mathcal{G}$ (as in Theorem 1.10). Since $\mathcal{G} \subseteq \widehat{\mathcal{H}}$ has dimension n, $T_{\widehat{r}}\mathcal{G}$ is a n-dimensional subspace of $\widehat{\mathcal{H}}$. We introduce a basis $\widehat{e}_1, \dots, \widehat{e}_n$ of $T_{\widehat{r}}\mathcal{G}$, where we adopt the notation

$$\widehat{e}_h := (e_h^0, \dots, e_h^{2m}) \in \widehat{\mathcal{H}}, \quad h = 1, \dots, n.$$

Then, for each $v \in T_{\widehat{r}}\mathcal{G}$ there exists a suitable set of constants $\gamma := (\gamma_1, \dots, \gamma_n)^\top$ such that $v = \sum_{h=1}^n \gamma_h \widehat{e}_h$. By linearity, the Fréchet derivative of Z is

$$\begin{aligned}
\partial_{\widehat{r}} Z \cdot v &= \begin{pmatrix} \sum_{j=0}^m a_j^1 v^j(x_1) + \sum_{j=1}^m b_j^1 v^{m+j} \\ \vdots \\ \sum_{j=0}^m a_j^n v^j(x_n) + \sum_{j=1}^m b_j^n v^{m+j} \end{pmatrix} \\
&= \sum_{h=1}^n \gamma_h \begin{pmatrix} \sum_{j=0}^m a_j^1 e_h^j(x_1) + \sum_{j=1}^m b_j^1 e_h^{m+j} \\ \vdots \\ \sum_{j=0}^m a_j^n e_h^j(x_n) + \sum_{j=1}^m b_j^n e_h^{m+j} \end{pmatrix} \\
&= \sum_{h=1}^n \gamma_h (T_{\widehat{r}} Z \cdot \widehat{e}_h) =: K_n(x)\gamma
\end{aligned}$$

where

$$K_n(x) := \begin{pmatrix} \alpha^1 \cdot \widehat{e}_1(x_1) & \cdots & \alpha^1 \cdot \widehat{e}_n(x_1) \\ \vdots & \ddots & \vdots \\ \alpha^n \cdot \widehat{e}_1(x_n) & \cdots & \alpha^n \cdot \widehat{e}_n(x_n) \end{pmatrix}, \tag{1.59}$$

and

$$\alpha^h = (a_0^h, \dots, a_m^h, b_1^h, \dots, b_m^h) \in \mathbb{R}^{2m+1}, \quad h = 1, \dots, n, \tag{1.60}$$

$$\widehat{e}_h(x_k) = (e_h^0(x_k), \dots, e_h^m(x_k), e_h^{m+1}, \dots, e_h^{2m}), \quad h = 1, \dots, n, \; k = 1, \dots, n. \tag{1.61}$$

with $a \cdot b$ that denotes the scalar product on $\mathbb{R}^{2m+1}$. Therefore, the function Z introduced in equation (1.58) is a local system of coordinates if $K_n(x)$ is invertible. We prove it by induction on $k \leq n$.

We consider the case $k = 1$:

$$K_1(\overline{x}) = \alpha^1 \cdot \widehat{e}_1(x_1) = \sum_{j=0}^{m} a_j^1 e_1^{r_j}(x_1) + \sum_{j=1}^{m} b_j^1 e_1^{y_j}. \tag{1.62}$$

We can find a vector α^1 and a maturity x_1 such that $K_1(x) \neq 0$. Indeed if not this means that: $K_1(x) = 0$ for every choices of α^1 and x_1. But if we choose $\alpha^1 = e_h$, for $h = 1, \dots, 2m+1$ where e_h is the $h-th$ element of the canonical basis of $\mathbb{R}^{2m+1}$, we will have:

$$\begin{cases} e_1^h(x_1) = 0, & h = 0, \dots, m, \\ e_1^{m+h} = 0, & h = 1, \dots, m. \end{cases}$$

for each $x_1 \in \mathbb{R}_+$. But this implies that $\widehat{e}_1$ is zero and this is a contradiction.

We now consider the inductive step: we suppose that $K_k(x)$ is invertible. Then we consider the matrix:

$$K_{k+1}(x, x_{k+1}) = \begin{pmatrix} \alpha^1 \cdot \widehat{e}_1(x_1) & \cdots & \alpha^1 \cdot \widehat{e}_k(x_1) & \alpha^1 \cdot \widehat{e}_{k+1}(x_1) \\ \vdots & \ddots & \vdots & \vdots \\ \alpha^k \cdot \widehat{e}_1(x_k) & \cdots & \alpha^k \cdot \widehat{e}_k(x_k) & \alpha^k \cdot \widehat{e}_{k+1}(x_k) \\ \alpha^{k+1} \cdot \widehat{e}_1(x_{k+1}) & \cdots & \alpha^{k+1} \cdot \widehat{e}_k(x_{k+1}) & \alpha^{k+1} \cdot \widehat{e}_{k+1}(x_{k+1}) \end{pmatrix}.$$

Let us suppose by contradiction that $K_{k+1}(x, x_{k+1})$ is not invertible for any choice of x_{k+1}, i.e. $\det[K_{k+1}(x, x_{k+1})] = 0$, $\forall x_{k+1} \in \mathbb{R}_+$. Since the first k columns $(C_1, \dots, C_k)$ form a space of rank k, this implies that the last column C_{k+1} has to satisfy:

$$C_{k+1} = \sum_{i=1}^{k} \kappa_i C_i, \tag{1.63}$$

for a suitable set of constants κ_i such that $\kappa := (\kappa_1, \dots, \kappa_k) \neq \overline{0}$. Let us notice that, since $\det[K_k(x)] \neq 0$, if we denote by $K_{k+1}^k(x)$ the first k rows of $K_{k+1}(x, x_{k+1})$, the last column of $K_{k+1}^k(x)$ is determined by the first k columns by relation (1.63). Hence, κ depends only x. As a consequence:

$$\alpha^{k+1} \cdot \widehat{e}_{k+1}(x_{k+1}) = \sum_{i=1}^{k} \kappa_i(x) \alpha^{k+1} \cdot \widehat{e}_i(x_{k+1}),$$

which is equivalent to require that:

$$\alpha^{k+1} \cdot \Big(\overbrace{\widehat{e}_{k+1}(x_{k+1}) - \sum_{i=1}^{k} \kappa_i(x) \widehat{e}_i(x_{k+1})}^{:= v_{k+1}(x, x_{k+1})} \Big) = 0,$$

which implies that $\alpha^{k+1} \in < v_{k+1} >^{\perp}$. Finally, in order to have a contradiction we choose α^{k+1} such that

$$\alpha^{k+1} \notin < v_{k+1}(x, x_{k+1}) >^{\perp} . \tag{1.64}$$

Let us notice now that, fixed x we can always choose x_{k+1} such that $v_{k+1}(x, x_{k+1})$ is not the null vector (because $(\widehat{e})_{h=1}^{k+1}$ is a basis). Therefore, we can conclude that, there exists $\alpha^{k+1} \neq \overline{0}$ such that (1.64) is satisfied, because $< v_{k+1} >^{\perp}$ is a subspace of dimension $n-1$ in $\mathbb{R}^n$, for each k. □

1.6 An alternative notion of invariance

1.6.1 A definition of consistency that manages the spreads

Theorem 1.10 provides an equivalent characterization of the consistency and it is at the basis of the procedure described at the beginning of Section 1.5.1 for the construction of FDRs. In the multi-curve framework, the presence of the spreads should not be problematic in terms of recalibration procedure, since they are finite-dimensional. Hence, in this section we study conditions that guarantee the finite-dimensional components of (1.6) to be included in the space variable Z, introduced in Definition 1.9. In other words, adopting the notation $\widehat{r}_t = (r_t, Y_t) \in \widehat{\mathcal{H}}$, we investigate under which conditions the existence of a stopping time τ, a process Z_t, taking values on a subset of $\mathbb{R}^n$ and a function $\overline{G} : \mathbb{R}^{m+n} \to \mathcal{H}^{m+1}$ such that:

$$r_t(x) = \overline{G}(x; Y_t, Z_t), \quad x \geq 0, \quad t \in [0, \tau]. \tag{1.65}$$

is guaranteed. (1.65) is equivalent to the definition of $\widehat{r}$-invariance (introduced in Definition 1.9), where the immersion G is defined as:

$$G(y, z) := (\overline{G}(y, z), y). \tag{1.66}$$

To develop the computations, we write the dynamics of a multi-curve model $\mathcal{M}$ as follows:

$$d\widehat{r}_t := \begin{pmatrix} dr_t \\ dY_t \end{pmatrix} = \begin{pmatrix} \mu(r_t, Y_t)dt + \sigma(r_t, Y_t) \circ dW_t \\ \gamma(r_t, Y_t)dt + \beta(r_t, Y_t) \circ dW_t \end{pmatrix} \in \begin{pmatrix} \mathcal{H}^{m+1} \\ \mathbb{R}^m \end{pmatrix}. \tag{1.67}$$

We recall that μ and γ are fully determined by σ and β by non-arbitrage constraints. In this setting, the consistency can be described by the following notion of invariance:

Definition 1.29. A parameterized family $\overline{G}$ is (r, y)-invariant under the action of a process $\widehat{r}$ defined by the dynamics (1.67), if there exists a stopping time $\tau(r_0(x), Y_0)$ and a process Z taking values on $\mathcal{Z} \subseteq \mathbb{R}^n$, such that condition (1.65) holds.

Analogously with the equivalence between Definition 1.8 and Definition 1.9, it is possible to prove that Definition 1.29 is equivalent to Definition 1.8, with $G := (\overline{G}, \mathbb{1}_m)$. As a consequence, the following result can be proved as in Theorem 1.10:

Proposition 1.30. *A parameterized family $\overline{G} : \mathbb{R}^{m+n} \to \mathcal{H}^{m+1}$ such that $G := (\overline{G}, \mathbb{1}_m)$ satisfies Assumption 1.7 defines a manifold $\mathcal{G} := Im[(\overline{G}, \mathbb{1}_m)]$. Then, $(\mathcal{M}, \mathcal{G})$ is invariant, if and only if, for each $(z, y) \in \mathcal{Z} \times \mathbb{R}^m$:*

$$\begin{cases} \mu((\overline{G}(y,z), y)) - \partial_y \overline{G}(y,z)\gamma(\overline{G}(y,z), y) \in Im[\partial_z \overline{G}(y,z)], \\ \sigma_i((\overline{G}(y,z), y)) - \partial_y \overline{G}(y,z)\beta_i(\overline{G}(y,z), y) \in Im[\partial_z \overline{G}(y,z)], \quad i = 1, \dots, d, \end{cases} \tag{1.68}$$

where μ, γ, σ_i and β_i are introduced in equation (1.67). $\partial_z \overline{G}$ and $\partial_y \overline{G}$ stand for the Fréchét differential of $\overline{G}$ with respect the z and y variable respectively.

Proof. First of all, we introduce the notation: $\widetilde{\beta}(y,z) := \beta(\overline{G}(y,z), y)$ and $\widetilde{\gamma}(y,z) := \gamma(\overline{G}(y,z), y)$, for any $(y,z) \in \mathbb{R}^{m+n}$. By Assumption 1.7, $\widetilde{\beta}$ and $\widetilde{\gamma}$ are smooth vector fields from $\mathbb{R}^{m+n}$ to $\mathbb{R}^n$.

($\Rightarrow$) We exploit the equivalence between invariance and $\widehat{r}$-invariance. We suppose that the couple $(\mathcal{M}, \mathcal{G})$ is invariant, that is:

$$d\widehat{r}_t = \begin{pmatrix} r_t \\ Y_t \end{pmatrix} = G(Y_t, Z_t) = \begin{pmatrix} \overline{G}(Y_t, Z_t) \\ Y_t \end{pmatrix}.$$

Thus, we consider the forward-rate components:

$$\begin{aligned} dr_t =& \partial_y \overline{G}(Y_t, Z_t) \circ dY_t + \partial_z \overline{G}_z(Y_t, Z_t) \circ dZ_t \\ =& \partial_y \overline{G}(Y_t, Z_t)(\gamma(r_t, Y_t)dt + \beta(r_t, Y_t) \circ dW_t) + \partial_z \overline{G}_z(Y_t, Z_t)(a(r_t, Y_t)dt + b(r_t, Y_t) \circ dW_t) \\ =& \partial_y \overline{G}(Y_t, Z_t)(\gamma(\overline{G}(Y_t, Z_t), Y_t)dt + \beta(\overline{G}(Y_t, Z_t), Y_t) \circ dW_t) + \partial_z \overline{G}_z(Y_t, Z_t)(a(\overline{G}(Y_t, Z_t), Y_t)dt \\ & \quad + b(\overline{G}(Y_t, Z_t), Y_t) \circ dW_t) \\ =& \Big[\partial_z \overline{G}(Y_t, Z_t)\widetilde{\gamma}(Y_t, Z_t)) + \partial_z \overline{G}(Y_t, Z_t)a(Y_t, Z_t)\Big]dt + \Big[\partial_z \overline{G}(Y_t, Z_t)\widetilde{\beta}(Y_t, Z_t) \\ & \quad + \partial_z \overline{G}(Y_t, Z_t)b(\overline{G}(Y_t, Z_t), Y_t)\Big] \circ dW_t, \end{aligned} \tag{1.69}$$

Since the starting point (y_0, z_0) can be chosen arbitrarily, (1.69) implies that:

$$\begin{cases} \mu((\overline{G}(y,z),y)) = \mathrm{Im}[\partial_z \overline{G}(y,z)] + \partial_y \overline{G}(y,z)\widetilde{\gamma}(y,z) \\ \sigma_i((\overline{G}(y,z),y)) = \mathrm{Im}[\partial_z \overline{G}(y,z)] + \partial_y \overline{G}(y,z)\widetilde{\beta}_i(y,z), \quad i = 1,\dots,d, \end{cases}$$

where d is the dimension of the Brownian motion which drives the model $\mathcal{M}$. In particular, (1.68) holds.

($\Leftarrow$) Let us suppose that (1.68) holds. This implies that there exists $d+1$ vector fields $a, b_i : \mathbb{R}^{m+n} \to \mathbb{R}^n$ such that:

$$\begin{cases} \mu((\overline{G}(y,z),y)) = \partial_y \overline{G}(y,z)\gamma(\overline{G}(y,z),y) + \partial_z \overline{G}(y,z)a(y,z) \\ \sigma_i((\overline{G}(y,z),y)) = \partial_y \overline{G}(y,z)\beta_i(\overline{G}(y,z),y) + \partial_z \overline{G}(y,z)(b_i(y,z), \quad i = 1,\dots,d, \end{cases} \tag{1.70}$$

Let us focus on the drift components (the computations for the volatility term is the same). (1.70) implies that:

$$\begin{aligned} \widehat{\mu}(\overline{G}(y,z),y) &:= \begin{pmatrix} \mu(\overline{G}(y,z),y) \\ \widetilde{\gamma}(y,z) \end{pmatrix} \\ &= \begin{pmatrix} \partial_y \overline{G}(y,z)\gamma(\overline{G}(y,z),y) + \partial_z \overline{G}(y,z)a(y,z) \\ \widetilde{\gamma}(y,z) \end{pmatrix} \\ &= \begin{pmatrix} \partial_y \overline{G}(y,z) & \partial_z \overline{G}(y,z) \\ \mathbb{1}_m & \mathbb{0}_{m\times n} \end{pmatrix} \begin{pmatrix} \widetilde{\gamma}(y,z) \\ a(y,z) \end{pmatrix} \\ &=: K(y,z) \begin{pmatrix} \widetilde{\gamma}(y,z) \\ a(y,z) \end{pmatrix} \end{aligned} \tag{1.71}$$

We notice that $K(y,z)$ is the Fréchét differential of the function G, hence, by Assumption 1.7, it admits left inverse, that we denote by $H(\overline{G}(y,z),y)$. By smoothness of $\overline{G}$, the function $\widetilde{H}(y,z) := H(\overline{G}(y,z),y)$ is smooth too (in the Fréchét sense). Moreover, we notice that:

$$\begin{pmatrix} \widetilde{\gamma}(y,z) \\ a(y,z) \end{pmatrix} = \widetilde{H}(y,z)\widehat{\mu}(\overline{G}(y,z),y)$$

This implies that:

$$\begin{aligned} (\widetilde{\gamma}, a) :& \mathbb{R}^{m+n} \to \mathbb{R}^{m+n} \\ & (y,z) \mapsto (\widetilde{\gamma}(y,z), a(y,z)) \end{aligned}$$

is a smooth function, therefore it is locally Lipschitz continuous. Proceding analogously with the volatility components $(\widetilde{\beta}(y,z), b(y,z))$, we obtain that:

$$\widehat{\sigma}_i(\overline{G}(y,z),y) := \begin{pmatrix} \sigma_i(\widehat{G}(y,z),y) \\ \widetilde{\beta}_i(y,z) \end{pmatrix} = K(y,z) \begin{pmatrix} \widetilde{\beta}_i(y,z) \\ b_i(y,z) \end{pmatrix}. \tag{1.72}$$

Then, there exists a unique local solution of the stochastic differential equation:

$$\begin{pmatrix} dY_t \\ dZ_t \end{pmatrix} := \begin{pmatrix} \widetilde{\gamma}(Y_t,Z_t) \\ a(Y_t,Z_t) \end{pmatrix} dt + \begin{pmatrix} \widetilde{\beta}(Y_t,Z_t) \\ b(Y_t,Z_t) \end{pmatrix} \circ dW_t.$$

Now, we introduce $\widehat{q}_t = (q_t, Y_t) := (\overline{G}(Y_t,Z_t), Y_t)$. By the chain rule, the Stratonovic dynamics of $\widehat{q}$ is:

$$\begin{aligned}
d\widehat{q}_t &= \begin{pmatrix} \partial_y \overline{G}(Y_t,Z_t) \circ dY_t + \partial_z \overline{G}(Y_t,Z_t) \circ dZ_t \\ dY_t \end{pmatrix} \\
&= \begin{pmatrix} \partial_y \overline{G}(Y_t,Z_t)(\widetilde{\gamma}(Y_t,Z_t)dt + \widetilde{\beta}(Y_t,Z_t) \circ dW_t) + \partial_z \overline{G}(Y_t,Z_t)(a(Y_t,Z_t)dt + b(Y_t,Z_t) \circ dW_t) \\ \widetilde{\gamma}(Y_t,Z_t)dt + \widetilde{\beta}(Y_t,Z_t) \circ dW_t \end{pmatrix} \\
&= \begin{pmatrix} (\partial_y \overline{G}(Y_t,Z_t)\widetilde{\gamma}(Y_t,Z_t) + \partial_z \overline{G}(Y_t,Z_t)a(Y_t,Z_t))dt + (\partial_y \overline{G}(Y_t,Z_t)\widetilde{\beta}(Y_t,Z_t) + \partial_z \overline{G}(Y_t,Z_t)b(Y_t,Z_t)) \circ dW_t \\ \widetilde{\gamma}(Y_t,Z_t)dt + \widetilde{\beta}(Y_t,Z_t) \circ dW_t \end{pmatrix} \\
&= \begin{pmatrix} \partial_y \overline{G}(Y_t,Z_t) & \partial_z \overline{G}(Y_t,Z_t) \\ \mathbb{1}_m & \mathbb{0}_{m\times n} \end{pmatrix} \begin{pmatrix} \widetilde{\gamma}(Y_t,Z_t) \\ a(Y_t,Z_t) \end{pmatrix} dt + \begin{pmatrix} \partial_y \overline{G}(Y_t,Z_t) & \partial_z \overline{G}(Y_t,Z_t) \\ \mathbb{1}_m & \mathbb{0}_{m\times n} \end{pmatrix} \begin{pmatrix} \widetilde{\beta}(Y_t,Z_t) \\ b(Y_t,Z_t) \end{pmatrix} \circ dW_t.
\end{aligned}$$

Hence, by (1.71) and (1.72) together with the definition of $\widetilde{\gamma}$ and $\widetilde{\beta}$, the dynamics of $\widehat{q}$ is:

$$d\widehat{q}_t := \widehat{\mu}(\overline{G}(Y_t,Z_t),Y_t)dt + \widehat{\sigma}(\overline{G}(Y_t,Z_t),Y_t) \circ dW_t.$$

Recalling that, by definition, $q_t = \overline{G}(Y_t,Z_t)$, we obtain that:

$$d\widehat{q}_t := \widehat{\mu}(\widehat{q}_t)dt + \widehat{\sigma}(\widehat{q}_t) \circ dW_t.$$

In conclusion, $\widehat{r}$, solution of equation (1.67) and $\widehat{q}$ solve the same stochastic differential equation. By local uniqueness, there exists $\tau > 0$ such that $\widehat{r}_t = \widehat{q}_t$, $\forall t \in [0,\tau)$. As a consequence, $r_t = q_t$ $\forall t \in [0,\tau)$. Therefore, $r_t = \overline{G}(Y_t,Z_t)$ and $\widehat{r}$ satisfies Definition 1.29. □

1.6.2 Conditions for the existence of FDRs under Definition 1.29

In this subsection, we investigate conditions on the coefficients of an interest-rate model $\mathcal{M}$ that guarantee the existence of FDRs according with Definition 1.29, under the assumption that $\mathcal{M}$ admits a FDRs in accordance with Definition 1.13. We notice that, when a model $\mathcal{M}$ possesses FDRs, it is always possible to construct another FDRs in the form of Definition 1.29. Indeed, we consider a FDR of a multi-curve process $\widehat{r}$ given by a function $G = (\overline{G}, \widetilde{G})$ and a finite-dimensional process Z, such that $(r_t, Y_t) = (\overline{G}(Z_t), \widetilde{G}(Z_t))$. Then, we can introduce another FDR for the same model of the form introduced in Definition 1.29, given by $(r_t, Y_t) = (\overline{G}(Z_t), Y_t)$, where the state process is given by:

$$d\begin{pmatrix} Y_t \\ Z_t \end{pmatrix} := \begin{pmatrix} \gamma(\overline{G}(Z_t), Y_t)dt + \beta(\overline{G}(Z_t), Y_t) \circ dW_t \\ a(Z_t)dt + b(Z_t) \circ dW_t \end{pmatrix},$$

and γ and β are introduced in system (1.65).

As discussed in Remark 1.12, we can find conditions to pass from an invariant parameterized family $G : \mathbb{R}^n \to \widehat{\mathcal{H}}$, such that $G(Z_t) = \widehat{r}_t$, to another consistent parameterized function in the sense of Definition 1.29 that is more parsimonious. To find these conditions we notice that, as shown in Section 1.5, the FDRs are obtained by a set of generators for the Lie algebra $\{\widehat{\mu}, \widehat{\sigma}_1, \dots, \widehat{\sigma}_d\}_{LA}$. We denote them by $\xi_1, \dots, \xi_n$. Hence, the embedding G introduced in equation (1.28) is:

$$G(\overline{z}) := (\overline{G}(\overline{z}), \widetilde{G}(\overline{z})), \tag{1.73}$$

where $\overline{z} := (z_1, \dots, z_n)$, $\overline{G}$ takes values on $\mathcal{H}^{m+1}$ and $\widetilde{G}$ takes values on $\mathbb{R}^m$. Then, we consider the following condition:

Condition 1.31. There exists a subspace $\mathbb{R}^p$ of $\mathbb{R}^m$ (where the spreads live), whose elements are denoted by $\widetilde{y} \in \mathbb{R}^p$ with $p \leq m$, that is diffeomorphic to a subspace of the state space through function $\widetilde{G}$. The elements of this diffeormophic sub-space are denoted by w. Therefore, we decompose the state-space vector using the notation $\overline{z} = (z, w) \in \mathbb{R}^{n-p} \times \mathbb{R}^p$.

Without loss of generality, we assume that the vector $\widetilde{y}$ is formed by the first p components of y, then we denote by $\overline{y}$ the last $m-p$ $(y := (\widetilde{y}, \overline{y}))$. Then, if $\widetilde{G} := (\widetilde{G}_p, \widetilde{G}_{m-p}) \in \mathbb{R}^p \times \mathbb{R}^{m-p}$, the invariance between $\mathcal{M}$ and $\mathcal{G}$ implies that $\widetilde{y} = \widetilde{G}_p(z, w) = \widetilde{G}_p[z](w)$ is invertible in w-variable for each choice of $z \in \mathbb{R}^{n-p}$. As a consequence, there exists $\widetilde{G}_p[z]^{-1}$ such that $w = \widetilde{G}_p[z]^{-1}(\widetilde{y})$.

Let us notice that Condition 1.31 is not restrictive, because if there is no such subspace, we can assume $p = 0$. This is the less parsimonious case, introduced at the beginning of this section.

To reconstruct the FDRs including $y \in \mathbb{R}^m$ in the state-space, it is sufficient the last $m-p$ elements of the canonical basis of $\widehat{\mathcal{H}}$ to the set of generators ξ_i. Using the notation introduced in (1.43), we refer to these elements as γ_k, for $k = p+1, \dots, m$. Hence, to handle the vector fields $(\gamma_k)_{k=p+1,\dots,m}$, that we added to the Lie algebra generated by the coefficients $\widehat{\mu}$ and $\widehat{\sigma}$, we must assume that:

Assumption 1.32. The Lie algebra generated by $\xi_1, \dots, \xi_n, \gamma_{p+1}, \dots, \gamma_m$ is finite-dimensional.

A priori, Assumption 1.32 is not guaranteed even if $\{\xi_1, \dots, \xi_n\}_{LA}$ is finite-dimensional. To find conditions for it, we introduce the notion of multi-index ∂_y^α (as in [Bjö04, Definition 7.4]):

Definition 1.33. A multi-index $\alpha \in \mathbb{Z}_+^k$ is any k-vector with nonnegative integer elements. For a multi index $\alpha = (\alpha_1, \dots, \alpha_k)$ the differential operator $\partial_{\overline{y}}^\alpha$ is defined as:

$$\partial_{\overline{y}}^\alpha = \frac{\partial^{\alpha_1}}{\partial y_{p+1}^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial y_{p+2}^{\alpha_2}} \cdots \frac{\partial^{\alpha_k}}{\partial y_k^{\alpha_k}}.$$

Then, the following result holds:

Proposition 1.34. *The following are respectively a necessary and sufficient condition for Assumption 1.32:*

- *If the Lie algebra*

$$\mathcal{N} := \{\xi_1, \dots, \xi_n, \gamma_{p+1}, \dots, \gamma_m\}_{LA} \tag{1.74}$$

is finite dimensional, then:

$$\begin{cases} dim[Span\{\partial_{\hat{y}}^{\alpha} \widehat{\mu}(r, y);\ \alpha \in \mathbb{Z}_+^{m-p}\}] < \infty; \\ dim[Span\{\partial_{\hat{y}}^{\alpha} \widehat{\sigma}_i(r, y);\ \alpha \in \mathbb{Z}_+^{m-p}\}] < \infty, \quad i = 1, \dots, d. \end{cases} \tag{1.75}$$

- *A sufficient condition for Assumption 1.32 is that γ_k commutes with $\widehat{\mu}$ and $\widehat{\sigma}_i$ for each $k = p + 1, \dots, m$.*

Proof. The Lie-algebra $\mathcal{N}$ introduced in (1.74) must contain all the Lie-brackets of the form:

$$[\xi_i, \gamma_k] := \partial_{\hat{r}} \xi_i \gamma_k - \partial_{\hat{r}} \gamma_k \xi_i = \partial_{y_k} \xi_i, \quad k = p + 1, \dots, m.$$

Hence, all the differentials $\partial_y^\alpha \xi_i$ have to be contained in $\{\xi_1, \dots, \xi_n, \gamma_{p+1}, \dots, \gamma_m\}_{LA}$. However, we notice that $\mathcal{N}$ is equal to $\{\widehat{\mu}, \widehat{\sigma}_1, \dots, \widehat{\sigma}_d, \gamma_{p+1}, \dots, \gamma_m\}_{LA}$. Then, we first observe that the vectors γ_k commute each other, thus their Lie bracket is null. Therefore, in order to guarantee that $\dim[\mathcal{N}] < \infty$ it is necessary to guarantee that $\partial_y^\alpha \mu$, and $\partial_y^\alpha \sigma_i$, $i = 1, \dots, d$ for any $\alpha \in \mathbb{Z}_+^{m-p}$ do not generate an infinite dimensional distribution. In particular this implies that (1.75) must be satisfied.

On the other hand, assuming that γ_k commutes with $\widehat{\mu}$ and $\widehat{\sigma}$ is equivalent to require that:

$$\begin{aligned} [\widehat{\mu}, \gamma_k] &= 0, \\ [\widehat{\sigma}_i, \gamma_k] &= 0, \quad i = 1, \dots, d; \end{aligned} \tag{1.76}$$

for each $k = p + 1, \dots, m$. Then, if we consider the successive Lie Brackets, by the Jacobi identity also they commute with γ_k: $[[\widehat{\mu}, \widehat{\sigma}_i], \gamma_k] = -\Big([[\widehat{\sigma}_i, \gamma_k], \widehat{\mu}] + [[\gamma_k, \widehat{\mu}], \widehat{\sigma}_i]\Big) = 0$.

In conclusion, we provided sufficient (equation (1.76)) and necessary (system (1.75)) conditions on the coefficients of the model $\mathcal{M}$, for Assumption 1.32. □

Let us notice that, to prove the existence of FDRs in the case of constant direction volatility model, determined by $\widehat{\sigma}$ of the form (1.39), we studied the Lie algebra $\mathcal{L}^1$, defined in (1.46), that is larger than the Lie algebra generated by the coefficients of the model. We provided conditions that guarantee $\dim[\mathcal{L}^1] < \infty$. Under these conditions, described in the statement of Proposition 1.24, we can construct FDRs in which the spread process is included in the state variable. Indeed, Assumption 1.32 is guaranteed, because $\mathcal{L}^1$ already contains all the vector fields γ_k, $k = 1, \dots, m$.

We apply this property in Example 1.35 below. In particular, we introduce a volatility term which determines a HJM multi-curve model that admits FDRs. Moreover, we establish a diffeomorphism between the spread components of the model and a subset of the components of the state process Z, exploiting the fact that the sigma algebra $\mathcal{L}^1$ already contains the vector fields γ_k.

Example 1.35. We consider a model $\mathcal{M}$, driven by a volatility term:

$$\widehat{\sigma}(y)(x) = \begin{pmatrix} \sigma^0 e^{-a^0 x} \\ y^1 e^{-a^1 x} \\ y^2 e^{-a^2 x} \\ \beta^1 y^1 \\ \beta^2 y^2 \end{pmatrix} \tag{1.77}$$

By simplicity, we consider a model driven by a scalar Brownian motion. However, the extension to a model driven by a multi-dimensional Brownian motion is straightforward.

The study of this class of models is interesting from a financial point of view, because it is natural to suppose that the size of volatility of the forward-rate components and the spread process are proportional (as discussed in Example (1.5.3.1.1)). We notice that, the structure of the volatility is a particular case of the one introduced in (1.39), where $i = 1$ and:

$$\begin{aligned} \varphi^0(r, y) &= \sigma^0, \\ \varphi^j(r, y) &= y^j, \qquad j = 1, 2, \\ \lambda^j(x) &= e^{-a^j x}, \quad j = 0, 1, 2, \\ \beta^j(r, y) &= \beta^j y^j, \qquad j = 1, 2. \end{aligned}$$

The functions λ^j are quasi-exponential for every $j = 0, \ldots, m$, Thus by Proposition 1.24, the FDRs exist. To build them, In particular, for each $j = 0, 1, 2$, there exists natural numbers n^j and p^j and real constants c^j and d^j_k, $k = 0, 1$, such that the conditions introduced in (1.47) are satisfied (we omitted the index i, present in (1.47), because $d = 1$). Repeating the same computations described in Example (1.5.3.1.1), we obtain that:

$$\begin{cases} n^j = 1, & \forall j = 0, 1, 2; \\ c^j = -a^j, & \forall j = 0, 1, 2; \\ p^j = 2, & \forall j = 0, 1, 2; \\ d^j_0 = -2(a^j)^2, & \forall j = 0, 1, 2; \\ d^j_1 = -3a^j, & \forall j = 0, 1, 2. \end{cases}$$

By (1.48), the dimension of the Lie-algebra is:

$$n = 2 + 1 + \sum_{j=0}^{2}(n^j + p^j) = 2 + 1 + (2+1)(1+2) = 12.$$

We adopt the following notation, which is a simplification of the one introduced in Notation 1.26:

$$z = (x^0, x^1, x^2, z^0, z^1, z^2, x_0^0, x_1^0, x_0^1, x_1^1, x_0^2, x_1^2).$$

By (1.50), the FDRs are:

$$\begin{aligned} &G^j(z,x) = r_j^M(x^0 + x) + z_0^j e^{-a^j x} + x_0^j \frac{1}{a^j}\left(e^{-a^j x} - e^{-2a^j x}\right) + x_1^j\left(-e^{-a^j x} + 2e^{-2a^j x}\right), \quad j = 0,1,2;\\ &G^{2+j}(z) = y_j^M + \int_0^{x^0} (r_0^M(s) - r_j^M(s))ds + x^j, \quad j = 1,2. \end{aligned} \tag{1.78}$$

The state process Z is given by the solution of (1.29),where a and b are two vector fields on $\mathbb{R}^n$ respectively, defined by (1.51) and (1.52). Z is a stochastic process taking values on $\mathbb{R}^12$, therefore, we adopt the notation:

$$Z_t = \begin{pmatrix} X_t^0 & X_t^1 & X_t^2 & Z_t^0 & Z_t^1 & Z_t^2 & X_{0,t}^0 & X_{1,t}^0 & X_{0,t}^1 & X_{1,t}^1 & X_{0,t}^2 & X_{1,t}^2 \end{pmatrix}.$$

First of all, let us notice that $X_t^0 = t$ for every t. Moreover, as described in (1.78), we can establish a diffeomorphism between the variable y^j and the variable x^j for every choice of the other components of z. Indeed, by $\widehat{r}$-invariance, the following equivalence holds:

$$Y_t^j = G^{2+j}(Z_t) := y_j^M + \int_0^t (r_0^M(s) - r_j^M(s))ds + X_t^j, \quad j = 1,2.$$

To simplify the notation, we introduce the term $\psi^j(\widehat{r}^M; x^0) = y_j^M + \int_0^{x^0}(r_0^M(s) - r_j^M(s)ds$. In particular, the processes X_t^j are determined by the spread processes Y^j together with the time-variable $X_t^0 = t$, one we fixed the initial value $\widehat{r}^M$. Let us notice that, function G^j for $j = 0,1,2$, introduced in (1.78) does not depend on x^j. As a consequence, there exists a $\widehat{r}$-invariant immersion $G := (\overline{G}, \mathbb{1})$, defined for each $\overline{z} \in \mathbb{R}^{12}$, where

$$\overline{z} = (x^0, y^1, y^2, z^0, z^1, z^2, x_0^0, x_1^0, x_0^1, x_1^1, x_0^2, x_1^2).$$

defined as follows:

$$\begin{aligned} &G^j(z,x) = r_j^M(x^0 + x) + z_0^j e^{-a^j x} + x_0^j \frac{1}{a^j}\left(e^{-a^j x} - e^{-2a^j x}\right) + x_1^j\left(-e^{-a^j x} + 2e^{-2a^j x}\right), \quad j = 0,1,2;\\ &G^{2+j}(\overline{z}) = y^j, \quad j = 1,2. \end{aligned}$$

The dynamics of the state process $\overline{Z}$ is $d\overline{Z}_t = \overline{a}(\overline{Z}_t) + \overline{b}(\overline{Z}_t) \circ dW_t$, where

$$\overline{a}(\overline{z}) = \begin{pmatrix} 1 \\ \gamma^1(\overline{G}(\overline{z}), y) \\ \gamma^2(\overline{G}(\overline{z}), y) \\ -a^0 z^0 \\ -z^1 a^1 - \beta^1 y^1 \\ -z^2 a^2 - \beta^2 y^2 \\ -2(a^0)^2 x_1^0 + 1 \\ x_0^0 - 3a^0 x_1^0 \\ -2(a^1)^2 x_1^1 + 1 \\ x_0^1 - 3a^1 x_1^1 \\ -2(a^2)^2 x_1^2 + 1 \\ x_0^2 - 3a^2 x_1^2. \end{pmatrix}, \quad \overline{b}(\overline{z}) = \begin{pmatrix} 0 \\ \beta^1 y^1 \\ \beta^2 y^2 \\ \sigma^0 \\ y^1 \\ y^2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Remark 1.36. As discussed, the presence of the spreads impacts on the conditions that guarantee the existence of FDRs. However, the spreads are described by a finite-dimensional stochastic process. As a consequence, it should be natural just to focus on the study of existence of FDRs for the forward-rate components, that are infinite-dimensional and then add the spreads to the state space process. This procedure works when the volatility term $\widehat{\sigma}$ of the multi-curve model does not depend on the spread components. Indeed, if $\widehat{\sigma}(\widehat{r}) = \widehat{\sigma}(r)$, the forward-rate components of $\mathcal{M}$ do not depend on the spread process. Thus, we can study the conditions that guarantee the existence of FDRs for a forward-rate model on $\mathcal{H}^{m+1}$ driven by the following dynamics:

$$dr_t = \mu(r_t)dt + \sigma(r_t) \circ dW_t.$$

If we find a finite set of generators to $\{\mu, \sigma_1, \dots, \sigma_d\}_{LA}$ on $\mathcal{H}^{m+1}$, denoted by $\xi_1(r), \dots, \xi_n(r)$, a FDR for the forward-rate components can be built, adopting the strategy described in Section 1.5.1. In conclusion, a function $\overline{G} : \mathbb{R}^n \to \mathcal{H}^{m+1}$ and a finite-dimensional process Z such that $r_t = \overline{G}(Z_t)$ for $t \in [0, \tau)$ exist. On this reasoning, we based the strategy described in Section 1.4.2, under which the enlargement of the domain of the consistent family for the forward-rate components of the Hull-White model is provided.

However, assuming that $\widehat{\sigma}$ does not depend on Y is quite restrictive. Therefore, in general, the forward-rate coefficients μ and σ have to be interpreted as smooth (in the Fréchét sense) functions from $\widehat{\mathcal{H}}$ to $\mathcal{H}^{m+1}$. As a consequence, in general, solving the problem of existence of FDRs for the forward-rate components should be expressed in terms of finite set of generators in $\mathcal{H}^{m+1}$ for $\mathcal{L}^y := \{\mu(\cdot, y), \sigma_1(\cdot, y), \dots, \sigma_d(\cdot, y)\}_{LA}$ for any $y \in \mathbb{R}^m$. Actually, even if this holds, the dependence on $y \in \mathbb{R}^m$ could lead to an infinite-dimensional distribution. In an attempt to solve this problem, we recognized an analogy between the structure of the dynamics of (1.6) and the class of single-curve HJM models driven by a stochastic volatility term. The consistency problem for this class of model is studied in [Bjö04, Chapter 7], in which HJM models r whose volatility is determined by a stochastic process Y are presented. To develop the computations the author supposed that the volatility term Y is independent of r. Since the spread processes of a multi-curve model depend explicitly on the forward-rate components of the model, we cannot make the same assumption.

1.7 Application to market data - a calibration algorithm

In this section, we present a calibration algorithm for the interest-rate market on the basis of the results shown in Section 1.5. We consider a multi-curve forward-rate model, described by system (1.6), that is determined by a set of parameters θ. We consider a model $\mathcal{M}$ that admits FDRs, also depending on θ . Thus, we propose an algorithm that estimates the parameter θ^* such that the l_2 distance between the FDRs of $\mathcal{M}$ and a time series of market data is minimized. Through this procedure, the FDRs determined by θ^* provide the manifold on $\widehat{\mathcal{H}}$ that gives the best representation (in terms of l_2 distance) of the time series of market data taken in analysis. As a consequence, θ^* can be used to produce a realization of $\mathcal{M}$ at the end of the time series of market data.

We consider a three-curve Hull-White model $\mathcal{M}$ driven by a scalar Brownian motion. $\mathcal{M}$ is determined by the volatility term:

$$\widehat{\sigma}(\widehat{r}) = \left(\sigma^0 e^{-a^0 x}, \quad \sigma^1 e^{-a^1 x}, \quad \sigma^2 e^{-a^2 x}, \quad \beta^1, \quad \beta^2\right). \tag{1.79}$$

where β^j, a^j, σ^j are positive constant for $j = 0, 1, 2$. Let us notice that Theorem 1.19 holds for this specification of the model. Indeed, the volatility components are QE functions, since $(\mathbf{F}+a^j)\sigma^j = 0$, $\forall j = 0, 1, 2$. Indeed, the volatility term introduced in (1.79) is a particular case of the volatility structure studied in Section 1.5.2.

1.7.1 Construction of the FDRs

As a first step, we study the vector fields generating $\mathcal{L} := \{\widehat{\mu}, \widehat{\sigma}\}_{LA}$. In view of Remark 1.20, the annihilator of the forward-rate component of volatility (1.79) is:

$$M(\lambda) := (\lambda + a^0)(\lambda + a^1)(\lambda + a^2) = \lambda^3 + \alpha_3\lambda^2 + \alpha_2\lambda^1 + \alpha_1.$$

By equation (1.34), $\dim(\mathcal{L}) := n = 1 + 1 + \deg(M) = 5$ is dimension of the FDRs of $\mathcal{M}$. Then, we consider a state space vector in $\mathbb{R}^5$ denoted by

$$z = (z^0, z_1^0, z_1^1, z_1^2, z_1^3),$$

in accordance with Notation 1.21. We can construct the tangential manifold, determined by the composition of the integral curves of the generators of $\{\widehat{\mu}, \widehat{\sigma}\}_{LA}$. Using the notation of Section

1.5.2, these vector fields are $\widehat{\mu}$, $\widehat{\sigma}$ and

$$\nu^1 = \begin{pmatrix} -a^0\sigma^0 e^{-a^0x} \\ -a^1\sigma^1 e^{-a^1x} \\ -a^2\sigma^2 e^{-a^2x} \\ \sigma^0-\sigma^1 \\ \sigma^0-\sigma^2 \end{pmatrix}, \quad \nu^2 = \begin{pmatrix} (a^0)^2\sigma^0 e^{-a^0x} \\ (a^1)^2\sigma^1 e^{-a^1x} \\ (a^2)^2\sigma^2 e^{-a^2x} \\ -a^0\sigma^0+a^1\sigma^1 \\ -a^0\sigma^0+a^2\sigma^2 \end{pmatrix}, \quad \nu^3 = \begin{pmatrix} -(a^0)^3\sigma^0 e^{-a^0x} \\ -(a^1)^3\sigma^1 e^{-a^1x} \\ -(a^2)^3\sigma^2 e^{-a^2x} \\ (a^0)^2\sigma^0-(a^1)^2\sigma^1 \\ (a^0)^2\sigma^0-(a^2)^2\sigma^2 \end{pmatrix}.$$

The composition of the integral curves of these vector fields, leads to the tangential manifold $G := (G^0, G^1, G^2, G^3, G^4)$. Applying (1.37), G is defined component by component as follows. First of all, we notice that, term S^j, defined in (1.36), has the following form:

$$S^j(x) = \int_0^x \sigma^j e^{-a^js} ds = \frac{\sigma^j}{a^j}(1-e^{-a^jx}), \quad j=0,1,2. \tag{1.80}$$

Then, G satisfies the following equations:

$$\begin{aligned} G^j(x;z) &= r_j^M(x+z^0) + \sum_{k=0}^{3} \mathbf{F}^k\sigma^j(x)z_1^k + \frac{1}{2}\Big(S^j(x+z^0)^2 - S^j(x)^2\Big) - \delta_0^j\Big(S^j(x+z^0)-S^j(x)\Big)\beta^j \\ &= r_j^M(x+z^0) + \sigma^j e^{-a^jx}(z_1^0 - a^jz_1^1 + (a^j)^2z_1^2 - (a^j)^3z_1^3) \\ &\quad + \frac{1}{2}\Big(\frac{\sigma^j}{a^j}\Big)^2 e^{-2a^jx}\Big(e^{-2a^jz^0}-1\Big) - \frac{\sigma^j}{a^j}\Big(\frac{\sigma^j}{a^j} - \delta_0^j\beta^j\Big)e^{-a^jx}\Big(e^{-a^jz^0}-1\Big), \quad j=0,1,2, \\ G^{2+j}(z) &= \sum_{k=1}^{3}\Big(\mathbf{BF}^{k-1}\sigma^0(x) - \mathbf{BF}^{k-1}\sigma^j(x)\Big)z_1^k + \beta^jz_1^0 + y_j^M + \int_0^{z^0}(r_0^M(s)-r_j^M(s))ds \\ &\quad + \frac{1}{2}\int_0^{z^0}\Big(S^0(s)^2 - S^j(s)^2\Big)ds + \beta^j\int_0^{z^0}S^j(s)ds - \frac{1}{2}(\beta^j)^2z^0 \\ &= (\sigma^0-\sigma^j)z_1^1 + (-a^0\sigma^0+a^j\sigma^j)z_1^2 + ((a^j)^2\sigma^0 - (a^j)^2\sigma^j)z_1^3 + \beta^jz_1^0 + y_j^M \\ &\quad + \int_0^{z^0}(r_0^M(s)-r_j^M(s))ds + \frac{1}{2}\Big(\frac{\sigma^0}{a^0}\Big)^2\Big[z^0 - \frac{2}{a^0}\Big(1-e^{-a^0z^0}\Big) + \frac{1}{2a^0}\Big(1-e^{-2a^0z^0}\Big)\Big] \\ &\quad - \frac{1}{2}\Big(\frac{\sigma^j}{a^j}\Big)^2\Big[z^0 - \frac{2}{a^j}\Big(1-e^{-a^jz^0}\Big) + \frac{1}{2a^j}\Big(1-e^{-2a^jz^0}\Big)\Big] + \frac{\sigma^j}{a^j}\beta^j\Big[z^0 - \frac{1}{a^j}\Big(1-e^{-a^jz^0}\Big)\Big] \\ &\quad - \frac{1}{2}(\beta^j)^2z^0, \quad j=1,2. \end{aligned} \tag{1.81}$$

The state process Z_t is determined by $dZ_t = A(Z_t)dt + B(Z_t)\circ dW_t$, where A, B are vector fields on $\mathbb{R}^5$, respectively defined by the coefficients a and b introduced in (1.38). We adopt the notation $Z_t = (Z_t^0, Z_{1t}^0, Z_{1t}^1, Z_{1t}^2, Z_{1t}^3) = (Z_t^0, Z_{1t})$. Still by (1.38), $Z_t^0 = t$. Thus, $Z_t = (t, Z_{1t})$.

1.7.2 Description of the algorithm

The parameters of the Hull-White model will be denoted by the vector

$$\theta := \Big(a^0, \quad \sigma^0, \quad a^1, \quad \sigma^1, \quad a^2, \quad \sigma^2, \quad \beta^1, \quad \beta^2\Big). \tag{1.82}$$

The algorithm we propose is based on [AH02], [AH05] and [Sli10, Section 6] and its goal is to provide best (in terms of l_2 distance) manifold in $\widehat{\mathcal{H}}$ for the time series of market data in analysis that is consistent with the three-curve HW model.

1.7.2.1 Initial families

The FDRs for the HW model depend on the value of the market data at the beginning of the observed time series, by the terms r_0^M, r_1^M, r_2^M, y_1^M and y_2^M of equation (1.81). Following the approach developed in [Sli10, Section 6] in the setting of the multi-currency exchange-rate market, we introduce initial families to describe r_j^M with Nelson Siegel's families (equation (1.11) with parameters $y^0, \dots, y^3$ instead of $z^1, \dots, z^4$) where the term y_3 is the exponential term of the associated component of the Hull-White model ($y_3 = a^j$ if we are considering the j^{th} forward-rate component, $j = 0, 1, 2$). Then:

$$r_j^M(x; y) := y_0 + y_1 e^{-a^j x} + y_2 x e^{-a^j x} = M_j^0(x; a^j) \cdot y, \quad j = 0, 1, 2, \tag{1.83}$$

where $y := (y_0, y_1, y_2)$ and $M_j^0(x; a^j) := (1,\ e^{-a^j x},\ xe^{-a^j})$ is parametric matrix that depends on the terms a^j, x. As a consequence, the initial family of each component of the model depends linearly on the same vector $\overline{y}$ and the only difference is given by the presence of the exponent a^j in the forward-rate components. Moreover, we observe that even if we have assumed that $y_6 = a^j$ for the j^{th} forward-rate curve, the initial family is inconsistent with the model.

1.7.3 The calibration procedure

The market data are defined on the time interval $\{t_0, \dots, t_N\}$. For each day $t \in \{t_0, \dots, t_N\}$ the value of fictitious bonds associated with every tenor δ_j for a set of maturities $\overline{x} := \{x_1, \dots, x_n\}$ and the logarithm of the (spot) spreads are extracted from market data:

$$\text{MK\_data}_t := \Big(B_t^0(x_1), \quad \dots, \quad B_t^0(x_n), \quad B_t^1(x_1), \quad \dots, \quad B_t^2(x_n), \quad Y_t^1, \quad Y_t^2\Big)^\top \in \mathbb{R}^{3n+2}, \tag{1.84}$$

At any $t \in \{t_0, \dots, t_N\}$, the yields associated with the FDRs, defined in (1.81) must be compared with the yields generated by of the market data for any tenor and any maturity x_i. We denote the FDRs by $G^j(x; t, z_1, y; \theta)$ to emphasize the dependence on the parameters that has to be estimated. The coefficient z_1 represents the realization of the components Z_1, of the state variable. Then, it has to be estimated at each t_h. In conclusion, the residual to be minimized, denoted by $\text{Res}_t(\overline{x}; y, z_1; \theta) \in$

$\mathbb{R}^{3n+2}$, is given by the difference between yields generated by the FDRs and market data:

$$\mathrm{Res}_t(\overline{x};y,z_1;\theta)=\begin{pmatrix}\frac{1}{x_1}\Big(-\int_0^{x_1}G^0(u;t,z_1,y;\theta)du-\Big(\log B_t^0(x_1)\Big)\Big)\\ \vdots\\ \frac{1}{x_n}\Big(-\int_0^{x_n}G^0(u;t,z_1,y;\theta)du-\Big(\log B_t^0(x_n)\Big)\Big)\\ \frac{1}{x_1}\Big(-\int_0^{x_1}G^1(u;t,z_1,y;\theta)du-\Big(\log B_t^1(x_1)\Big)\Big)\\ \vdots\\ \frac{1}{x_n}\Big(-\int_0^{x_n}G^2(u;t,z_1,y;\theta)du-\Big(\log B_t^2(x_n)\Big)\Big)\\ G^3(t,z_1,y;\theta)-Y_t^1\\ G^4(t,z_1,y;\theta)-Y_t^2\end{pmatrix} \tag{1.85}$$

By the assumptions made on the initial families (described in Subsection 1.7.2.1) and by the properties of the finite-dimensional realizations, the yields obtained by the parameterized functions G^j are linear functions of (z_1, y). Then, the two-step calibration algorithm is defined as follows:

P.1 For each $t_h \in \{t_0, \dots, t_N\}$ the residual is minimized as function of the time dependent parameters (y, z_1). By linearity, SVD algorithm can be adopted to obtain $y(t_h, \theta)$ and $z_1(t_h, \theta)$, that depend the time-independent parameters θ:

$$(y(t_h,\theta),z_1(t_h,\theta)) := \arg\min_{(y,z_1)} ||\mathrm{Res}_{t_h}(\overline{x};y,z_1;\theta)||.$$

where $\mathrm{Res}_{t_h}(\overline{x};y,z_1;\theta)$ is defined on equation (1.85) and $||\cdot||$ is the Euclidean norm in $\mathbb{R}^{3n+2}$.

P.2 $(y(t_h,\theta),z_1(t_h,\theta))$, estimated at **P.1**, is substituted in $\mathrm{Res}_{t_h}(\overline{x};\overline{y},\overline{z}_1;\theta)$ as the selected (y, z_1). As a consequence, $\mathrm{Res}_{t_h}(\overline{x};y(t_h,\theta),z_1(t_h,\theta);\theta)$ depends only on θ. θ is estimated minimizing the sum of the squared-norm of the residuals along the entire time interval:

$$\theta^* := \arg\min_{\theta>0}\left(\sum_{h=0}^{N}|\mathrm{Res}_{t_h}(\overline{x},y(t_h,\theta),z_1(t_h,\theta);\theta)|^2\right)^{\frac{1}{2}}.$$

To compute θ^*, we adopt the function "optimize.least_squares" of Python (version 3.8.5) library "scipy". This function is based on a reflective trust-region algorithm. Moreover, constraints can be imposed to the domain of the function to be minimized.

1.7.4 Bootstrap of market data

The time series of market data for EURIBOR market is provided by Refinitiv for the time period from 10/08/2016 to 19/11/2021 with business daily basis. Table 1.1 summarizes the data.

We consider the discount curve associated with OIS rate and we consider the risky curves associated with tenor $3M$ (3-months) and $6M$ (6-months). For the bootstrapping technique used to build the data, we based on [GL20, Section 2]. We determined values of the term structures for

discount/risky curve	Market instrument	Quoted maturities
Discount curve	OIS	1W - 2W - 3W - 1M - 2M - 3M - 4M - 5M - 6M - 7M - 8M - 9M - 10M - 11M - 1Y - 15M - 18M - 21M - 2Y - 3Y - 4Y - 5Y - 6Y - 7Y - 8Y - 9Y - 10Y
Three months	FRA	1Mx4M - 2Mx5M - 3Mx6M - 4Mx7M - 5Mx8M - 6Mx9M - 7Mx10M - 8Mx11M - 9Mx12M
	IRS	18M - 2Y - 3Y - 4Y - 5Y - 6Y - 7Y - 8Y - 9Y - 10Y
Six months	FRA	1M+7M - 2Mx8M - 3Mx9M - 4Mx10M - 5Mx11M - 6Mx12M - 9Mx15M - 12Mx18M
	IRS	2Y - 3Y - 4Y - 5Y - 6Y - 7Y - 8Y - 9Y - 10Y

Table 1.1: Summary of market data.

each tenor at maturities $\overline{x} := \{1\text{M}, 2\text{M}, 3\text{M}, 4\text{M}, 5\text{M}, 6\text{M}, 9\text{M}, 1\text{Y}, 2\text{Y}, \dots, 10\text{Y}\}$ from one month to ten years.

1.7.4.1 Bootstrap of discount curve

In order to construct the term structure $T \to B_t^{OIS}(T)$ associated with the discount curve, we consider the formula of an OIS rate with schedule $\mathcal{T} = \{T_0, \dots, T_n\}$:

$$\mathcal{S}_0^{ON}(\mathcal{T}) := \frac{B_0^{OIS}(T_0) - B_0^{OIS}(T_n)}{\sum_{k=1}^n \delta B_0^{OIS}(T_k)}.$$

By market convention, if the contract is signed at $t = t_0$ the spot date is $T_0 = t_0 + 2$. As assumed in [AB13], we impose $t_0 = T_0$. This means that $B_0^{OIS}(T_0) = 1$. Then, for the construction of the term structure, we use the following convention (as described at the beginning of [GL20, Section 2.1]):

- For the maturities up to 21 months we assume that the payment is made only at the maturity. This means that the schedule is $\mathcal{T} = \{T_0, T_n\}$:

$$\mathcal{S}_0^{ON}(\mathcal{T}) = \frac{1 - B_0^{OIS}(T_n)}{\delta B_0^{OIS}(T_n)} \quad \Rightarrow \quad B_0^{OIS}(T_n) = \frac{1}{1 + \delta \mathcal{S}_0^{ON}(\mathcal{T})} = \frac{1}{1 + \tau(T_n - T_0)\mathcal{S}_0^{ON}(\mathcal{T})},$$

  where $\tau(T_n - T_0)$ is equal to the number of days between T_0 and T_n over 360.

- For the maturities beyond 21 months we assume yearly payments: we use the the prices computed with maturities $k = 1, \dots, n-1$ years:

$$\mathcal{S}_0^{ON}(\mathcal{T}) = \frac{1 - B_0^{OIS}(T_n)}{\delta \sum_{k=1}^n B_0^{OIS}(T_k)} \quad \Rightarrow \quad B_0^{OIS}(T_n) = \frac{1 - \delta \mathcal{S}_0^{ON}(\mathcal{T}) \sum_{k=1}^{n-1} B_0^{OIS}(T_k)}{1 + \delta \mathcal{S}_0^{ON}(\mathcal{T})},$$

  and δ is equal to 1 (year).

1.7.4.2 Bootstrap of risky-curves

The risky curve associated with the generic tenor $\delta \in \{3M, 6M\}$ is denoted by $\{T \to B_t^\delta(T)\}$. It is defined by the values of the fictitious tenor bonds, determined by equation (1.2):

$$B_t^\delta(T) = \frac{1+\delta L^\delta(t;T,T+\delta)}{1+\delta L^\delta(t;t,t+\delta)} \frac{B_t^{OIS}(T+\delta)}{B_t^{OIS}(t+\delta)}. \tag{1.86}$$

We cannot use the same strategy proposed in [GL20, Section 2] in order to provide this term structure, since we need the entire risk-free term structure. Thus, as done in [GL20, Appendix A] we interpolate the values of the OIS curve via cubic splines of zero-rates. Moreover, we use the Euribor spot rate time series associated with the tenor δ, denoted by $\{t \to L^\delta(t;t,t+\delta)\}$.

The bootstrapping technique exploits forward-rate agreement quotations to determine the values of fictitious bond for small maturities. For long maturities interest rate swap quotations are used:

- For short maturities, we consider forward-rate agreements, which at a given time t play the role of $L^\delta(t;T,T+\delta)$. Moreover, from the spot rate at time t and the discount curve already provided at time t and with maturity T and $T+\delta$, we get the value $B_t^\delta(T)$.

- For large maturities, we consider the interest rate swap quotations, as done in [GL20, Appendix A]. We recall that an IRS exchange a flow of payments with a fixed rate (whose schedule will be denoted by $\mathcal{T}^x$) with a flow of payments based on a floating rate (in this case we use the Euribor associated with a specific δ and with schedule denoted by $\mathcal{T}^\delta$). The IRS quotation is given by:

$$R^\delta(t;\mathcal{T}^x,\mathcal{T}^\delta) := \sum_{j=1}^{n_\delta} \Big(\frac{\delta B_t^{OIS}(T_j^\delta)}{\sum_{i=1}^{n_x} \delta_i^x B_t^{OIS}(T_i^x)} L^\delta(t;T_{j-1}^\delta,T_j^\delta) \Big), \tag{1.87}$$

  where $L^\delta(t;T_{i-1}^\delta,T_i^\delta) = L^\delta(t;T_{i-1}^\delta,T_{i-1}^\delta+\delta)$. As described in [GR15, Remark 1.5], usually the fixed leg schedule has yearly frequency ($\delta_i^x = 1,\ \forall x,i$) while the floating leg schedule has frequency equal to the tenor δ. Adopting this convention, we use the quotations of the IRS contracts which play the role of $R^\delta(t;\mathcal{T}^x,\mathcal{T}^\delta)$ in equation (1.87) (the values of these instruments which are not quoted in the market are interpolated by cubic splines, as described in [GL20, Section 2]). Doing this, together with the values $B_t^{OIS}(T_i^x)$ we bootstrap the successive values of $L^\delta(t;T_{j-1}^\delta,T_j^\delta)$. Then, we use these bootstrapped values of equation (1.86) to obtain $B_t^\delta(T_{j-1}^\delta)$. Finally, the inverting formula to get the value of the rate with the highest maturity by IRS quotations is the following:

$$L^\delta(t;T_{n_\delta-1}^\delta,T_{n_\delta}^\delta) = \frac{R^\delta(t;\mathcal{T}^d,\mathcal{T}^x)\Big(\sum_{j=1}^{n_x}\delta^x B_t^{OIS}(T_j^x)\Big) - \sum_{i=1}^{n_\delta-1}\delta B_t^{OIS}(T_i^\delta)L^\delta(t;T_{i-1}^\delta,T_i^\delta)}{\delta B_t^{OIS}(T_{n_\delta}^\delta)}.$$

1.7.4.3 Construction of spreads

Finally, we recall the formula to compute the spreads components:

$$S_t^\delta = \frac{1 + \delta L^\delta(t; t, t+\delta)}{1 + \delta L^{OIS}(t; t, t+\delta)},$$

where the rate associated with the OIS term respects:

$$L^{OIS}(t; t, t+\delta) = \frac{1}{\delta}\Big(\frac{1}{B_t^{OIS}(t+\delta)} - 1\Big).$$

1.7.5 Calibration results

To verify the performance of the calibration algorithm, we compare the market data at the end of the time (at $t = t_{N+1}$) with the parameterized family estimated in the time interval $\{t_0, \ldots, t_N\}$. The time-dependent parameters (y, z) at $t = t_{N+1}$ are estimated using the market data at t_{N+1} and the estimation of the time independent parameter θ^* is obtained through the calibration procedure described in Section 1.7.3.

Through a stability analysis (described in Subsection (1.7.5.2) below) we concluded that a time series four months length returns the most stable results. Thus, we analyse a time series of market data with length 4M and which starts at 01/04/2021. The initial value for the time-independent parameter $\theta_0 \in \mathbb{R}^8$ is given by Table 1.2.

	σ	a	β
OIS	$\sigma^0 = 0.00285941$	$a^0 = 0.53041117$	/
Libor - 6M	$\sigma^1 = 0.09546952$	$a^1 = 0.66253001$	$\beta^1 = 0.41734616$
Libor - 6M	$\sigma^2 = 0.09083773$	$a^2 = 0.65812121$	$\beta^2 = 0.82477578$

Table 1.2: Initial values of the model parameters.

Each parameter a^j, β^h in Table 1.2 is chosen randomly in the interval $[0, 1]$, while parameters σ^j are randomly chosen in the interval $[0, 0.1]$. The calibration procedure described in the steps **P.1**, **P.2** provided the estimation, described in Table 1.3.

	σ	a	β
OIS	$\sigma^0 = 0.1643$	$a^0 = 0.3719$	/
Libor - 6M	$\sigma^1 = 0.1590$	$a^1 = 0.3721$	$\beta^1 = 0.4814$
Libor - 6M	$\sigma^2 = 0.1598$	$a^2 = 0.3727$	$\beta^2 = 0.8825$

Table 1.3: Calibrated values of the model parameters.

As discussed at the beginning of this subsection, in terms of the fit of the yields given by

$$G^j(\overline{x}; t_{N+1}, y(t_{N+1}, \theta^*), z_1(t_{N+1}; \theta^*); \theta^*)$$

to the term-structure at t_{N+1}, the results are described in Figure 1.1. Let us notice that, due to

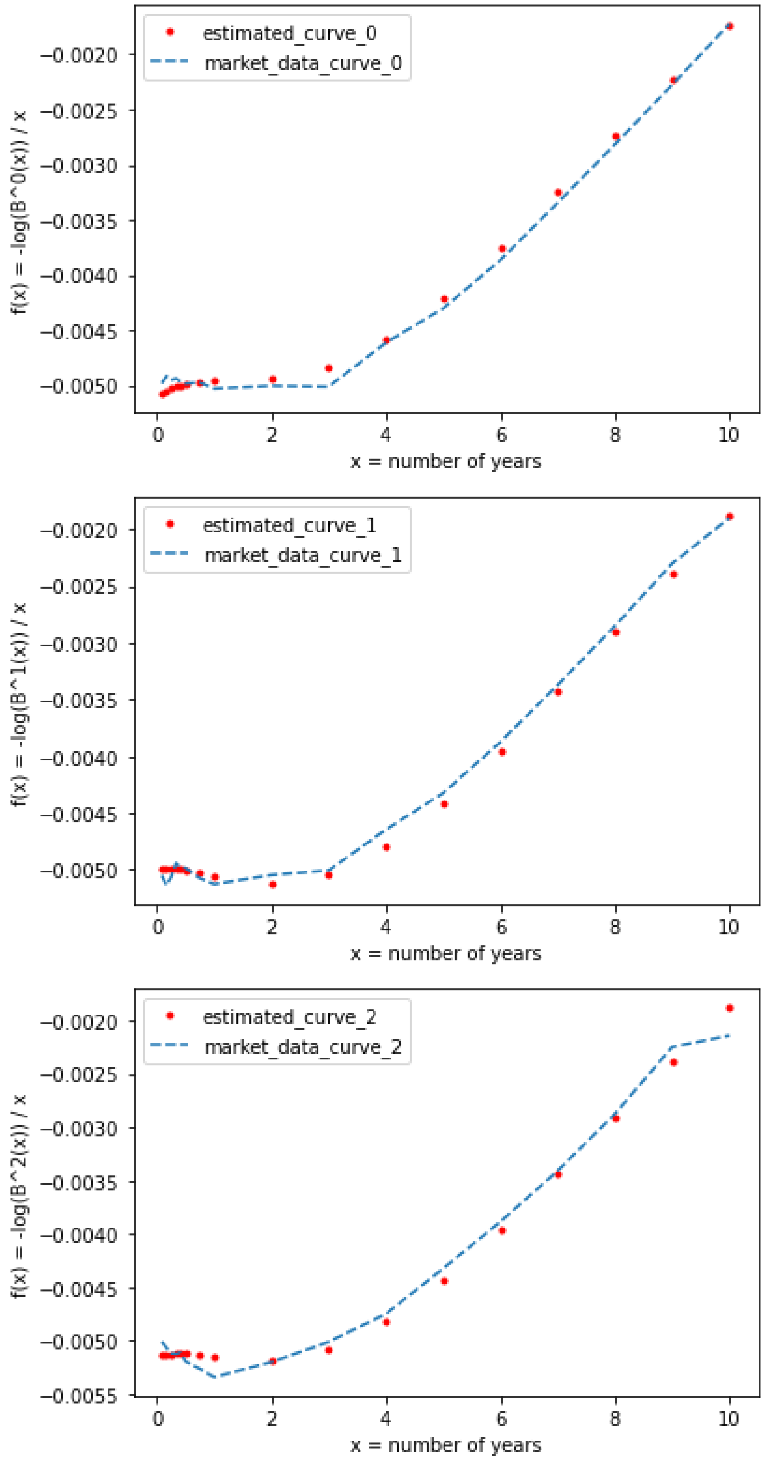

Figure 1.1: Comparison between market data yield and the estimated yield at the end of time series. Top panel: risk-free curve; central panel: 3M curve; bottom panel: 6M curve.

the monetary policies of 2021, the yields are negative for every tenor in all the maturities.

Since the spreads are parameterized as spot processes, we can compare the estimated spread curve with the spread curve obtained by market data over the whole time interval $\{t_0, \ldots, t_N\}$, as shown in Figure 1.2.

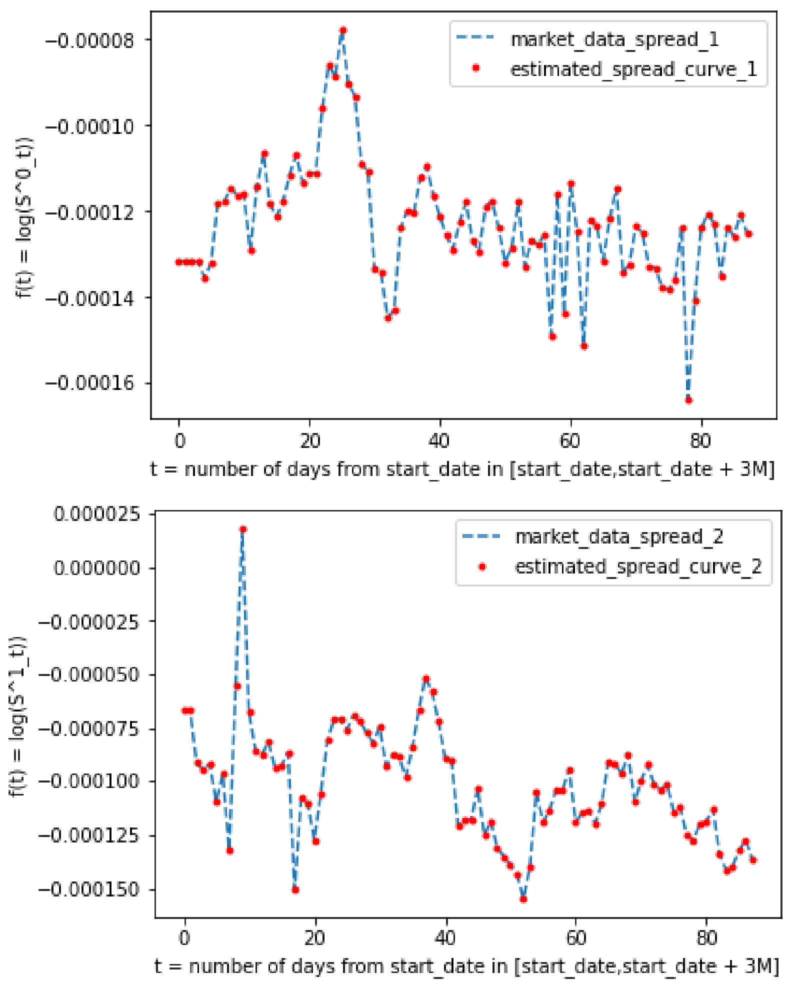

Figure 1.2: Comparison between market data and the estimated spread curve for the whole time series. Top panel: 3M spread curve; bottom panel: 6M spread curve.

In Table 1.4, the relative error between the estimated curves and the market data in each curve is presented. The error value is computed as follows:

1. If $G^j(\overline{x})$ is the estimated j^{th} yield curve at the end of time interval and $M^j(\overline{x})$ is the j^{th} yield curve at the end of time interval obtained by market data, the relative error is obtained by:

$$\text{err}_{\text{yield curve}} = \frac{||G^j(\overline{x}) - M^j(\overline{x})||_n}{||M^j(\overline{x})||_n}, \tag{1.88}$$

2. If $(S_t^j)_{t\in[0,T]}$ is the estimated value for the j^{th} spread in the entire time interval $\{t_0, \ldots, t_N\}$ used to provide the calibration and $(M_t^j)_{t\in\{t_0,\ldots,t_N\}}$ is the value for the j^{th} spread given by

the market data in the considered time interval, the relative error is:

$$\mathrm{err}_{\text{spread curve}} = \frac{||S^j - M^j||}{||M^j||} := \frac{\sqrt{\sum_{i=0}^{N}(S^j_{t_i} - M^j_{t_i})^2}}{\sqrt{\sum_{i=0}^{N}(M^j_{t_i})^2}}. \tag{1.89}$$

	OIS	3M	6M
yields	0.01917	0.01705	0.02385
spread	-	6.92929e-07	8.491172e-07

Table 1.4: Relative errors.

1.7.5.1 Stability with respect to the length of the time series

The length of the time-series to use to provide the estimation is an interesting factor to analyse. Indeed, choosing a time series too short should not give a have sufficient information to get a good estimation. On the other hand, choosing a time series too long should not give a good estimation because the consistency condition is a local property.

We present the error obtained comparing the market data with the manifold estimated at the end of the considered time interval, using different time series of market data. We provide these comparisons separately for each yield-curve and for each spread component. For the error in each k^{th} estimated yield curve, we consider the error function introduced in equation (1.88). To compute the error in the spread, it is convenient to consider the relative error between the estimated spread at t_{N+1}, denoted by S^k and the same value obtained by the market data, M^k for each $k = 1, 2$ (i.e. $\mathrm{err}_{\text{spread}_k} := \frac{|S^k - M^k|}{|M^k|}$).

The estimation procedure is initialized in θ_0 introduced in Table 1.2. We consider a time series of market data ending in 30/07/2021 with length 1M, 2M, 3M, 4M, 5M, 6M (we fix the ending date because we want to compare the market data and the estimated manifold at the end of the considered time interval, i.e. 01/08/2021 for each choice of the length).

The length of the time series that gives the best results in terms of relative error is 4M for an analysis ending at 30/07/2021, as shown in Table 1.5.

length of $\{t_0, \dots, t_N\}$	**Yield curve**			**Spread curve**	
n months	RFRs	Euribor 3M	Euribor 6M	Spread 3M	Spread 6M
1	0.0520	0.0524	0.0368	2.5736e-07	2.4474e-07
2	0.0557	0.0489	0.0371	5.6650e-08	1.6761e-10
3	0.0213	0.0218	0.0387	4.4926e-07	3.8488e-07
4	0.0191	0.0171	0.0239	1.1579e-06	9.6982e-07
5	0.0159	0.0163	0.0334	3.1638e-08	3.4686e-08
6	0.0213	0.0214	0.0393	1.2390e-07	6.3183e-08

Table 1.5: Relative error as a function of the length (in month) of the time window.

1.7.5.2 Stability of time-independent parameters

In Subsection 1.7.5.1 we shown that with a time series of length 4M the calibration procedure for the date 30/07/2021 provides the best result. Another interesting analysis to be conducted is the stability of the estimated parameters in time, using a time series of length 4M. We develop this analysis as follows, starting at $d_0 = 01/04/2021$, with θ_0 given in Table 1.2;

A.1 Apply the calibration algorithm with the time interval of length 4M starting at d_0.

A.2 Compute the calibration procedure with the time interval of length four months starting at data $d_0 + 1$ day, using as initial datum the value θ^* estimated at **A.1**.

A.3 Repeat the previous step, shifting the time series of market data with frequency one day for $\mathcal{N} = 50$ iterations.

In Table 1.6, we present the mean and the standard deviation of the vector made by the estimation of each parameter θ^* at step **A.2**, along the considered time interval:

	a^0	σ^0	a^1	σ^1	a^2	σ^2	β^1	β^2
avg	0.371948	0.164252	0.372120	0.159068	0.372732	0.159813	0.481433	0.882557
std	0.000004	0.000006	0.000003	0.000006	0.000004	0.000004	0.000002	0.000003

Table 1.6: Average and standard deviation of the parameters estimated along the time interval $[d_0, d_0 + 50 \text{ days}]$.

In conclusion, the parameters are stable in time. This stability can be partially motivated by the procedure we used to provide the analysis. Indeed, in line with the commonly adopted recalibration procedure, the initial value θ_0 of the calibration procedure at the iteration i is chosen as the value θ^* estimated at the iteration $i-1$. Under this assumption the stability of the parameters is good, because there is not a significant difference between curves estimated using time series of length four months that differ by a shift of one day.

1.8 Conclusions and further developments

Since the Libor reform there is no agreement on the choice of the benchmark rate to be adopted in the interest-rate market. Adopting RFRs the interest rates can be modelled by a single-curve approach. However, in the Eurozone market, the Euribor rate, that is the benchmark rate, is still computed for a set of different tenors. On the other hand, in the American market, the adoption of the RFR called SOFR is opposed to the desire of the analysts to describe the market via credit sensitive rates (CSRs). As mentioned, financial institutions, such as AFX, have proposed a set of CSRs that are described by a multi-curve setting. However, at the moment, there is not sufficiently high liquidity on the interest-rate derivatives used to construct these rates. Therefore, further analysis on these new rates cannot be provided yet.

An interesting issue to address is the generalization of the results presented in this chapter to general Heath-Jarrow-Morton models defined on suitable Fréchét spaces. Indeed, in analogy with [FT03] and [FT04], the boundedness of the functional $\mathbb{F}$, which is a necessary condition to define the Hilbert space $\mathcal{H}$ introduced in (1.7), does not allow for the geometric analysis of more sophisticated forward-rate models, like CIR. To overcome this issue, we could investigate the solution proposed in [FT03] and [FT04] in the context of single-curve modelling.

A second fascinating application of these results could be the extension of the calibration algorithm, described in Section 1.7, to forward-looking purposes. In particular, it would be interesting to understand if it is possible to apply the parameters determined by the calibration algorithm to estimate future realizations of the forward-rate curve through the finite dimensional process Z that determines the FDRs.

CHAPTER 2

Price formation under asymmetry of information: a mean-field approach

2.1 Introduction

In financial markets, quantifying the information available to an agent is a crucial task, especially when the amount of information accessible to every player is not homogeneous. Our purpose is to understand what is the impact of the equilibrium price ϖ, in a market populated by a large number of agents who can access different sources of randomness.

This chapter is based on a project jointly developed with Alekos Cecchin, Markus Fischer and my PhD supervisor, Claudio Fontana,

We are interested in a market model in which one asset is traded by N small agents, called *standard agents* and one major agent. We suppose that there is a gap between the information that is accessible to the major player and the standard agents. In particular, we assume that the major agent can observe a stochastic factor that is inaccessible to the standard players. We suppose that this stochastic factor impacts on the revenues of the major agent and therefore on her strategies. In this setting, we aim at studying the mechanism of price formation, that is the procedure at the basis of the determination of the price of an asset through the interaction between the rational agents who trade that asset. In particular, we derive an equation for the price under which the demand and supply of the asset are balanced at every time. We call this condition *market clearing*. The price determined in such a way, called *equilibrium price*, depends on the choices of every agent. As a consequence, the structure of the market becomes complex, due to the interdependence between the strategies of the agents. To overcome this problem, we follow the approach described in [FT22a], formulating the problem through a mean-field approach.

Mean-field games (MFGs) are stochastic differential games with infinitely many players and symmetric interactions. The seminal papers [LL07] and [HMC06] present a characterization of the Nash equilibrium for this class of games, through a coupled system of a Hamilton-Jacobi-Bellman

(HJB) and a Fokker-Planck (FP) equation. In [CD18a] and [CD18b] an alternative probabilistic approach, based on the Pontryagin maximum principle is proposed. In particular, the mean-field game solution is proved to be related to the solution to a McKean-Vlasov Forward Backward Stochastic Differential Equation (FBSDE). As we are going to show, we consider an analogy between the equation for the equilibrium price we are going to derive and the consistency condition of a weak mean-field game equilibrium. Hence, related results to the existence of solutions to these kinds of equilibria can be found in [CDL16] and [Lac16].

As discussed in [GNP15] and [Car20], there are many economic and financial applications of mean-field games theory. However, most of these results in mean-field game theory are applied to construct approximate Nash equilibria among the agents, given a response function of the price process exogenously. See, for example, applications to optimal trading ([LM19]) as well as liquidation of portfolio ([FGHP21]), exploitation of exhaustible resources and related problems among many agents' responses to an exogenously given price process ([Fu23; FH20]) and systemic risk [CDL17]. In [ABM20], an application to the optimal management of energy storage and distribution in a smart grid system is presented, while in [ABBC23] a mean-field model for the development of renewable capacities is proposed. In all these works the price is supposed to be defined as an exogenous stochastic process with a prescribed dynamics.

In this chapter, we aim at deriving an equation for the equilibrium price process, determined endogenously by the balance between demand and supply. In this direction a first application in the context of mean-field games is [GLL10], where the balance between demand and supply is analyzed explicitly, but the demand is supposed to be exogenously defined as a function of the asset price. In [GS21], through an analytic approach, the market clearing condition, is imposed. This condition leads to an equation for the price process, that turned out to be deterministic due to the absence of common noise. More general results, involving the presence of a common noise, are obtained in [FT22a; FT22c; Fuj23] through a probabilistic approach. In particular, the authors proved the existence of a solution to a mean-field equation for the equilibrium price process in a homogeneous information setting. Analogous results applied to the electricity market can be found in [FTT20; FTT21; ACP22; SFJ22] and [GGR23]. In [FTT20; FTT21], the intraday market electricity price is determined by a combination between the fundamental electricity price, that is supposed not to be affected by the agents strategies and the average position of the agents trading in the market. In [GGR23], in a commodity market model, the balance between the average quantity of the commodity demanded by the agents and the random supply of the commodity determines the equilibrium price.

We are interested in an asymmetric scenario, where a major agent is counterposed to a family of small symmetric agents. In the theory of mean-field games, models on which major and minor agents interact among each other are studied in [BCY16; CZ16; BLL20] and [CCP20]. An application to finance and economics is given in [Nuñ17]. In the direction of equilibrium price formation, interesting results are described in [FT22b] and [FTT20], with an application to the electricity market in presence of a major agent. Moreover, results related the issue of heterogeneity of information in mean-field game models are discussed in [SC16; FH20; MSZ18; Ber22; BY21; CJ19] and [CJ20]. In particular, in [CJ20] a market model in which the agents have different beliefs about the evolution of the exogenously given price process is presented. More recently, in [BCR23] and [BS24], the authors studied a market model described by a Stackelberg game in which a more informed major agent exploits the additional information to manipulate the market.

In this chapter, we aim at relating the issues linked with an asymmetric information structure to the features of an equilibrium price approach. We adopt the setting [FT22a; FT22b] and [FT22c]. We suppose that the N standard agents observe a common shock, given by a Brownian motion B, an idiosyncratic noise, independent of any other source of randomness and the price process ϖ. On the other hand, we suppose that the major player can observe the common noise B, the price process ϖ, but also an additional stochastic factor c that a priori is inaccessible to the other players.

As a first step, we assume that the price is an exogenous factor for every player. Therefore, the agents have to solve an optimal control problem depending on the random noises they can observe. Assuming that the agents are rational and minimize a cost functional, we obtain a family of optimal control problems depending on the common and idiosyncratic stochastic factors. We aim at determining the equilibrium price determined by the market clearing. By this condition, we are going to derive an equation for the equilibrium price process, depending on the optimal control of all agents. As a consequence, through the observation of the equilibrium price process, the standard agents can deduce the strategy of the major player, that depends on the additional factor c. In particular, the market clearing condition establishes a link between the optimal strategies of all players, leading to a complex structure of the market. Indeed, the equilibrium price ϖ is defined by a combination of the strategies of all agents, that depend on the price ϖ in a recursive way. Therefore, the equation for the equilibrium price ϖ seems intractable from an analytical point of view when N is finite.

To overcome this problem, we follow the strategy described in [FT22b] and study the case with infinitely many standard players. To do so, we assume that the information commonly shared by every standard agent is given by the one generated by the common noise B and the equilibrium price process ϖ. This procedure allows to ignore the effects of the idiosyncratic noises of each individual standard player. However, due to the market structure, the same does not happen to the private information accessible to the major player. As a consequence, we obtain a mean-field equation for the equilibrium price process. We prove the existence of a mean-field solution ϖ^{mf} of this equation. We can follow this approach due to the symmetry of the optimal control problems of every standard agent. However, to construct the solution to this mean-field equation, we cannot use standard tools like fixed point arguments. This is due to the structure of the mean-field equation, that depends on the filtration generated by the solution itself. To circumvent this problem, we recognize an analogy between the structure of the mean-field equation for the price process and the notion of weak equilibrium in the context of mean-field game theory. Therefore, we construct the solution adopting a strategy analogous to the one described in [CDL16] to determine a weak mean-field game equilibrium. We justify the construction of the price process in the mean-field limit, showing that ϖ^{mf} satisfies a weak form of the market clearing condition in the case of finitely many agents.

This chapter is structured as follows. In Section 2.2, we describe the market model and we introduce the concept of equilibrium price process in both the finite-dimensional case and in the mean-field limit. In Section 2.3, we describe our approach to construct the solution to the mean-field equation for the equilibrium price process. It is based on the discretization of the domain in which the solution is defined. Hence, we solve a fixed point problem in a space in which the characterization of compact sets is easier to obtain. We construct a sequence of stochastic processes, defined on the discretized domain, that approximates the equilibrium price. Applying suitable convergence arguments, we prove that the limit in distribution of the sequence of approximate solutions solves the mean-field equation. Finally, in Section 2.4, we prove that, under additional conditions on the structure of the market, we can justify the approximation of the equilibrium price process with its mean-field limit in a market populated by finitely many players.

2.1.1 Notation

We denote a given filtered probability space with the standard notation $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$. When we refer to the canonical space, we use the notation $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}}, \overline{\mathbb{F}})$, denoting with $\mathcal{G} \otimes \mathcal{H}$ the product sigma-algebra generated by two sigma-algebras $\mathcal{G}$ and $\mathcal{H}$.

We denote by L or C a positive real constant. We refer to the time horizon with the constant $T > 0$. Moreover, for every probability space $(\Omega, \mathcal{G}, \mathbb{P})$ endowed with a filtration $\mathbb{G} := (\mathcal{G}_t)_{t\in[0,T]}$ we refer to:

- $\mathbb{L}^2(\mathcal{G}; \mathbb{R})$ as the set of real valued $\mathcal{G}$-measurable square integrable random variables;
- $\mathbb{S}^2(\mathbb{G}; \mathbb{R})$ as the set of real valued $\mathbb{G}$-adapted càdlàg processes X satisfying:

$$||X||_{\mathbb{S}^2(\mathbb{G};\mathbb{R})} := \mathbb{E}\left[\sup_{t\in[0,T]} |X_t|^2\right]^{\frac{1}{2}} < \infty;$$

- $\mathbb{H}^2(\mathbb{G}; \mathbb{R})$ is the set of real valued $\mathbb{G}$-progressively measurable processes Z satisfying:

$$||Z||_{\mathbb{H}^2(\mathbb{G};\mathbb{R})} := \mathbb{E}\left[\left(\int_0^T |Z_t|^2 dt\right)\right]^{\frac{1}{2}} < \infty.$$

We adopt the standard notation $\mathcal{C}([0,T], \mathbb{R})$ for the space of real valued continuous functions on $[0,T]$ and $\mathcal{D}([0,T], \mathbb{R})$, for the space of real valued càdlàg functions on $[0,T]$. Finally, we denote by $\mathcal{M}([0,T], \mathbb{R})$ the set of real valued Lebesgue measurable functions on $[0,T]$.

We adopt the notation $\mathbb{F}^\Theta := (\mathcal{F}_t^\Theta)_{t\in[0,T]}$ to denote the natural filtration generated by the process Θ. The complete and right continuous augmentation of $\mathbb{F}^\Theta$ is denoted by $\overline{\mathbb{F}}^\Theta$.

2.2 The market setup

2.2.1 The probabilistic framework

As mentioned in the introduction, [FT22a; FT22b; FT22c] are at the starting point of this project. As in [FT22a, Section 3.1], we consider a market model populated by N standard agents and one major agent, who plays the role of a large company or investor. The players acting on this market trade a single asset. The goal of every agent is to solve an optimal control problem, that depends on the price of the traded asset, denoted in the following by ϖ. We suppose that every agent is a price taker, thus all agents make their decision considering the price as an exogenous stochastic process. In this setting, we consider the problem of the price formation: we aim at studying the dynamics of the price process ϖ that satisfies the market clearing condition. We consider a market model in which there exists a gap in terms of amount of information that is accessible to the standard agents and the major agent. We suppose that the major agent's revenues depend on an additional stochastic process c, that can be observed by her, but is inaccessible to the standard agents.

We aim at understanding if the less informed agents can deduce some information regarding the strategy of the major agent, through the observation of the price process. Indeed, since the strategy of the major agent depends on the extra factor c, it is a priori inaccessible to the standard agents.

The probabilistic setup is represented by a family of stochastic control problems defined as follows. First, we introduce the following probability spaces:

$$(\Omega^0, \mathcal{F}^0, \mathbb{P}^0), \quad (\Omega^j, \mathcal{F}^j, \mathbb{P}^j)_{j=1,\dots,N}.$$

Each probability space is endowed with a filtration defined as follows:

- For every $j = 1, \dots, N$, on $(\Omega^j, \mathcal{F}^j, \mathbb{P}^j)$ we introduce $\mathbb{F}^j := (\mathcal{F}^j_t)_{t\geq 0}$ as the usual augmentation of the filtration generated by a random variable ξ^j and a Brownian motion $(W^j_t)_{t\geq 0}$ independent of ξ^j. ξ^j is the initial value of the state variable of the j^{th}-standard agent, while the Brownian motions represent the idiosyncratic noises associated with each standard agent;
- On $(\Omega^0, \mathcal{F}^0, \mathbb{P}^0)$, we introduce a one dimensional Brownian motion $(B_t)_{t\geq 0}$ and another real valued stochastic process c, independent of B, which represents the private information of the major player, as discussed in Section 2.2.2.2 below. Hence, we denote by $\mathbb{F}^0 := (\mathcal{F}^0_t)_{t\geq 0}$ as the usual augmentation of the filtration generated by (c, B).

As a consequence, the Brownian motions $(W^j)_{j=1,\dots,N}$ are pairwise independent and are independent of (c, B). We define the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$, $\mathbb{F} := (\mathcal{F}_t)_{t\in[0,T]}$ by:

$$\begin{aligned} \Omega &:= \Omega^0 \times \Omega^1 \times \cdots \times \Omega^N, \\ (\mathcal{F}, \mathbb{P}) &:= (\mathcal{F}^0 \otimes \mathcal{F}^1 \otimes \cdots \mathcal{F}^N, \mathbb{P}^0 \otimes \cdots \otimes \mathbb{P}^N); \\ \mathcal{F}_t &:= \mathcal{F}^0_t \otimes \cdots \otimes \mathcal{F}^N_t, \quad t \in [0,T]. \end{aligned} \tag{2.1}$$

The price process is defined as follows:

Definition 2.1. The price process ϖ is a generic càdlàg process $\varpi := (\varpi_t)_{t\geq 0}$, defined on Ω.

2.2.2 The market

In this section we describe the optimal control problems that must be solved by the agents acting in the market. As discussed, every agent is a priori supposed to be a price taker. An individual agent in the market is called *price taker* if he cannot manipulate the price of the asset. He must accept the prevailing price in the market because he does not have enough market power to influence the price. A price taker chooses his strategies observing the price ϖ as an exogenous stochastic process.

2.2.2.1 The problem of the standard agents

In analogy to the setting proposed in [FT22a], we suppose that every standard agent has a group of customers who trade the securities with the agent and have no direct access to the market. We introduce the number of shares X_t^j possessed by the j^{th} agent at time t. The state process X^j is controlled by the agent through the trading speed, denoted α^j, that belongs to a set of admissible strategies $\mathbb{A}^j$. In particular, α_t^j is the number of shares, traded by the j^{th} agent in the infinitesimal time interval $[t, t+dt]$. Moreover, the position X^j is dependent also on the trades between the j^{th} agent and its individual clients. In particular, the random demand of the private customers of agent i is described by the Brownian motion W^i, multiplied by a factor σ that can also depend on the price process. Moreover, we also allow for the existence of a random demand that affects all the standard agents in the market in the same way. It is described by a factor σ_0, possibly dependent on ϖ, multiplied by the Brownian motion B.

As a consequence, we suppose that each standard agent's state dynamics is described by the following SDE:

$$\begin{cases} dX_t^j = (\alpha_t^j + l(t,\varpi_t))dt + \sigma_0(t,\varpi_t)dB_t + \sigma(t,\varpi_t)dW_t^j, \\ X_0^j = \xi^j, \end{cases} \tag{2.2}$$

for every $j = 1, \dots, N$. We recall that, the sequence of starting points $(\xi^j)_{j=1,\dots,N}$ is i.i.d.

Assumption 2.2. $\mathbb{E}[|\xi^j|^4] < \infty$, for every $j = 1, \dots, N$.

The cost functional to be minimized by the j^{th} agent is:

$$J^j(\alpha^j) := \mathbb{E}\left[\int_0^T f(t, X_t^j, \alpha_t^j, \varpi_t)dt + g(X_T^j, \varpi_T)\right], \quad j = 1, \dots, N, \tag{2.3}$$

where

$$\begin{aligned} f(t,x,a,\varpi) :=& \varpi a + \frac{1}{2}\Lambda a^2 + \overline{f}(t,x,\varpi), \\ g(x,\varpi) :=& \overline{g}_1(\varpi)x + \overline{g}_2(x). \end{aligned}$$

for measurable functions $\overline{f}$ and $\overline{g}$ and a positive constant Λ. For every $j = 1, \dots, N$, the family of admissible controls is given by the set of square-integrable stochastic processes adapted to the filtration generated by the price process ϖ, the Brownian motion B and the idiosyncratic noise W^j, denoted by $\mathbb{F}^{S,j} := \overline{\mathbb{F}}^{\varpi,B,W^j}$:

$$\alpha^j \in \mathbb{A}^j := \mathbb{H}^2(\mathbb{F}^{S,j}; \mathbb{R}). \tag{2.4}$$

In particular, every standard player observes the price process and the Brownian motions B, W^j, but he has no clue about the presence of other sources of information that are affecting the market.

2.2.2.2 The problem of the major agent

The major agent is a large company trading in market. We assume that she does not have a private set of customers for which she provides trading services. In other words, the dynamics of the state variable X^0, describing the number of shares in the portfolio at every time is:

$$dX_t^0 = \beta_t dt + \sigma_0^M(t, \varpi_t) dB_t, \quad X_0^0 = x_0, \tag{2.5}$$

where β represents, in analogy to α^j for the j^{th} standard agent, the strategy determining the trading speed of the major agent. For technical reasons, that are motivated by Remark 2.33 below, we assume that x^0 is a constant. As discussed in the introduction, the gap in the information structure between the major player and the standard agents is described by the presence of an additional stochastic factor c, independent of B that affects the revenues paid to the major player.

We consider a cost functional of the form:

$$J_{(N)}^0(\beta) := \mathbb{E}\left[\int_0^T f_{(N)}^0(t, X_t^0, \beta_t, \varpi_t, c_t) dt + g^0(X_T^0, \varpi_T, c_T)\right], \quad j = 1, \dots, N, \tag{2.6}$$

that has to be minimized, where

$$\begin{aligned} f_{(N)}^0(t, x, b, \varpi, c) :=& \varpi b + \frac{1}{2}\frac{\Lambda^0}{N} b^2 + \overline{f}^0(t, x, \varpi, c), \\ g^0(x, \varpi, c) :=& \overline{g}_1^0(\varpi, c) x. \end{aligned}$$

for some measurable functions $\overline{g}_1^0$, $\overline{g}_2^0$, $\overline{f}^0$, associated with the revenues obtained by trading activities. We assume that:

$$\overline{f}^0(t, x, \varpi, c) := \overline{c}_0^M(t, \varpi, c) x,$$

where $\overline{c}_0^M$ and $\overline{c}_1^M$ are measurable functions defined on $[0, T] \times \mathbb{R} \times \mathbb{R}$. As done for the less informed agent, we consider the set of admissible controls, that is given by, the space of square-integrable stochastic processes adapted to $\mathbb{F}^M := \overline{\mathbb{F}}^{\varpi,B,c}$

$$\beta \in \mathbb{A}^0 := \mathbb{H}^2(\mathbb{F}^M; \mathbb{R}), \tag{2.7}$$

This implies that the major agent can choose the trading speed by relying on the extra factor c.

Remark 2.3. Both the family of standard agents and the major agent have to solve optimal control problem that is dependent on the stochastic factor ϖ, representing the price process. In both cases the objective functional is determined by a running cost and a final cost. The running cost is expressed by the sum of a linear quadratic term in the control variable and a term dependent on the state variable. The linear quadratic term is given by two coefficients. The first component, $\varpi_t \alpha_t$, is the product between the price process and the control. It represents the amount of money exchanged to buy or sell the asset at time t. On the other hand, the quadratic term represents a penalization term due to the fees associated with large trades. The remaining component of the running cost, independent of the control variable, could be interpreted as a combination between costs associated with financial risk and revenues obtained by an appropriate management of the position, described by the state variable. In particular, in the case of the major player, the factor $\overline{c}_0^M$ represents a coupon stream proportional to the number of shares owned by the major player. $\overline{c}_0^M$ and $\overline{g}_1^0$ depend on the stochastic process c that determines the private information of the major agent. This process c can be interpreted as a stochastic factor that impact on the stream of coupon $\overline{c}_0^M$ allowing higher (or lower) revenues. On the other hand, the terminal cost of the major player is determined by the coefficient $\overline{g}_1^0$ that can be interpreted by the liquidity price of the terminal position. In this case, the factor c could represent a stochastic threshold for the liquidity price of the major player. For instance, if $\overline{g}_1^0(\varpi, c) = \max\{\varpi, c\}$, the major player can close its position at a higher price than the equilibrium price.

As described in Section 2.2.5, the presence of the factor N at the denominator of quadratic coefficient of the running cost function $f^0_{(N)}$ guarantees that, when passing to the limit in the number of standard agents, the impact of the major player in the market clearing condition does not disappear.

2.2.3 Technical assumptions

2.2.3.1 Assumptions on the coefficients

We introduce the hypotheses on the coefficients of the model. These hypotheses are an adaptation of [CD18b, Assumption Coefficients MFG with a Common Noise] and [CD18b, Assumption FBSDE MFG with a Common Noise]:

Assumption 2.4 (FBSDE MFG with a Common Noise)**.** There exists a constant $L > 0$ such that:

A1 The coefficients l, σ, σ_0^M and σ_0 introduced in (2.2) and (2.5) are Borel-measurable functions from $[0, T] \times \mathbb{R}$ into $\mathbb{R}$. For every $t \in [0, T]$ the functions $l(t, \cdot)$, $\sigma(t, \cdot)$, $\sigma_0(t, \cdot)$ and $\sigma_0^M(t, \cdot)$ are continuous on $\mathbb{R}$ and for every $\varpi \in \mathbb{R}$:

$$|l(t, \varpi)| + |\sigma_0(t, \varpi)| + |\sigma(t, \varpi)| + |\sigma_0^M(t, \varpi)| \leq L[1 + |\varpi|].$$

A2 For every $t \in [0,T]$ the coefficients $\overline{f}(t,\cdot,\cdot)$, $\overline{g}_1(\cdot)$ and $\overline{g}_2(\cdot)$, introduced in (2.3), are continuous functions from $\mathbb{R}\times\mathbb{R}$ to $\mathbb{R}$ and from $\mathbb{R}$ and $\mathbb{R}$ respectively. For every $t \in [0,T]$ and $\varpi \in \mathbb{R}$, the functions $\overline{f}(t,\cdot,\varpi)$, $\overline{g}_1(\cdot)$ and $\overline{g}_2(\cdot)$ are continuously differentiable. Analogously, the functions c_0^M, c_1^M, are continuous from $[0,T]\times\mathbb{R}\times\mathbb{R}$ to $\mathbb{R}$, while $\overline{g}_1^0$ and $\overline{g}_2^0$ are continuous from $\mathbb{R}\times\mathbb{R}$ to $\mathbb{R}$. Moreover:

$$|\overline{f}(t,x,\varpi)| + |g(x,\varpi)| \le L(1+|x|^2+|\varpi|^2),$$

A3 We assume that the functions $\overline{f}$ and g are convex functions in the x variable.

A4 We require that the volatility functions σ_0, σ_0^M and σ are jointly continuously differentiable in both variables, t and ϖ, they are convex in ϖ variable and their partial derivatives are bounded:

$$|\partial_t\sigma_0(t,\varpi)| + |\partial_\varpi\sigma_0(t,\varpi)| + |\partial_t\sigma(t,\varpi)| + |\partial_\varpi\sigma(t,\varpi)| + |\partial_t\sigma_0^M(t,\varpi)| + |\partial_\varpi\sigma_0^M(t,\varpi)| \le L.$$

Remark 2.5. Let us notice that the L-convexity condition (defined in [CD18a, Assumption (Control of MKV Dynamics]) is already satisfied by $L := \max\{(\Lambda^0)^{-1},(\Lambda)^{-1}\}$, due to the linear quadratic structure of the cost functional:

A5

$$f(t,x,\alpha',\varpi) - f(t,x,\alpha,\varpi) - (\alpha'-\alpha)\partial_\alpha f(t,x,\alpha,\varpi) \ge \frac{1}{2}L^{-1}|\alpha-\alpha'|^2,$$
$$f^0_{(N)}(t,x,\beta',\varpi) - f^0_{(N)}(t,x,\beta,\varpi) - (\beta'-\beta)\partial_\beta f^0_{(N)}(t,x,\beta,\varpi) \ge \frac{1}{2}L^{-1}|\beta-\beta'|^2.$$

Let us notice that L is independent of N.

In line with [CD18b, Section 1.4. - Voll II], we aim at applying the stochastic maximum principle to solve the optimal control problems of every agent. Therefore, we need some additional conditions, that are stated in Assumption 2.6 below. We adopt a forward-backward formulation of the solution to the optimal control problem, because, as we are going to show in Section 2.2.4, the equilibrium price process is going to be determined by a combination of the solutions of the backward components of the FBSDE system associated with the optimal control of every player.

Assumption 2.6 (FBSDE MFG with a common noise)**.**

B1 For a constant $\Lambda > 0$, we denote by $\overline{\Lambda} = -\Lambda^{-1}$. The functions

$$B(t,y,\varpi) := -\overline{\Lambda}(y+\varpi) + l(t,\varpi),$$
$$F(t,x,\varpi) := -\partial_x\overline{f}(t,x,\varpi),$$
$$G(x,\varpi) := \partial_x\overline{g}(x)$$

are continuous in (y,ϖ), (x,ϖ) and x variable respectively for every $t \in [0,T]$.

B2 There exists a constant $L > 0$, such that, such that $B(t, y, \varpi)$ introduced in B1, $\partial_x \overline{f}(t, x, \varpi)$, $\partial_x g(x, \varpi)$, $\overline{c}_0^M(t, \varpi, c)$, $\overline{g}_1^0(\varpi, c)$ satisfy:

$$\begin{aligned} |B(t, \varpi, y)| &\leq L[1 + |\varpi| + |y|], \\ |\partial_x \overline{f}(t, x, \varpi)| + |\partial_x g(x, \varpi)| + |\overline{c}_0^M(t, \varpi, c)| + |\overline{g}_1^0(\varpi, c)| &\leq L. \end{aligned} \tag{2.8}$$

B3 There exists a constant $L > 0$, such that, such that, for every $x, x' \in \mathbb{R}$, for every $t \in [0, T]$, for every $\varpi \in \mathbb{R}$:

$$\begin{aligned} |\partial_x \overline{f}(t, x', \varpi) - \partial_x \overline{f}(t, x, \varpi)| &\leq L|x - x'|; \\ |\partial_x \overline{g}_2(x) - \partial_x \overline{g}_2(x')| &\leq L|x - x'|. \end{aligned} \tag{2.9}$$

2.2.3.2 The compatibility condition

In Subsection 2.2.2.1 and Subsection 2.2.2.2, we defined two optimal control problems, one for the standard agents and one for the major agent. The optimal control problems depend on a stochastic process, respectively ϖ and (ϖ, c), that a priori may not be adapted to the Brownian motions driving the state variable. As discussed in [CD18b, Section 1.1], in this setting, we must handle carefully the interdependence between the stochastic processes determining the randomness of the model. In particular, if we do not specify properly the relations between the filtrations generated by ϖ, c and the Brownian motions B and W^j, $j = 1, \ldots, N$, we may encounter several problems when studying the optimal control problems introduced in Subsections 2.2.2.1 and 2.2.2.2. For instance, if we consider the case of the j^{th}-standard agent, without additional conditions on ϖ, the value of ϖ at a fixed time t may reveal future realizations of the Brownian motion (B, W^j). As a consequence, (B, W^j) would not be anymore a Brownian motion with respect the filtration $\mathbb{F}^{S,j}$. To overcome this problem, we are going to introduce an assumption, that is based on the following definition:

Definition 2.7. In a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, let us consider two filtrations $(\mathcal{G}_t)_{t\in[0,T]}$ and $(\mathcal{F}_t)_{t\in[0,T]}$ We say that $(\mathcal{G}_t)_{t\in[0,T]}$ is *immersed* in $(\mathcal{F}_t)_{t\in[0,T]}$ if

- $\mathcal{G}_t \subseteq \mathcal{F}_t$ for all $t \in [0, T]$.
- martingales with respect to $(\mathcal{G}_t)_{t\in[0,T]}$ remain martingales with respect to $(\mathcal{F}_t)_{t\in[0,T]}$.

The immersion of $(\mathcal{G}_t)_{t\in[0,T]}$ in $(\mathcal{F}_t)_{t\in[0,T]}$ is equivalent to require that:

$$\mathcal{G}_T \text{ is conditionally independent of } \mathcal{F}_t \text{ given } \mathcal{G}_t. \tag{2.10}$$

The immersion condition, introduced in [BY78] with the term $\mathcal{H}$-*hypothesis*, has been widely studied in the literature and has been analysed in many financial applications, notably in credit risk modelling ([EJY00; JR00; CJN12]). For a detailed description of the properties and the characterizations of the immersion condition, we refer to [AJ17, Chapter 3].

We state the compatibility condition, that is defined as follows:

Definition 2.8 (compatibility)**.** A stochastic process θ, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, is *compatible* with a filtration $\mathbb{F} := (\mathcal{F}_t)_{t\in[0,T]}$, if the natural filtration $\mathbb{F}^\theta$ generated by θ is immersed in $\mathbb{F}$.

In analogy to [CD18b, Definition 1.13], we introduce the concept of admissible probabilistic setup:

Definition 2.9. A complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, endowed with a complete and right-continuous filtration $\mathbb{F}$ and on which a process $\theta := (\xi, W, \varpi)$ is defined, is an admissible probabilistic setup if:

1. W is a $\mathbb{F}$-Brownian motion;
2. θ and $\mathbb{F}$ are compatible.

We may refer to the term admissible probabilistic setup by $((\Omega, \mathcal{F}, \mathbb{P}), \mathbb{F}, (\xi, W, \varpi))$.

We notice that, the process ϖ may be not independent of the Brownian motion W. Finally, we introduce the following assumption:

Assumption 2.10.

- For every $j = 1, \ldots, N$, $((\Omega, \mathcal{F}, \mathbb{P}), \mathbb{F}^{S,j}, (\xi^j, (B, W^j), \varpi))$ is admissible;
- $((\Omega, \mathcal{F}, \mathbb{P}), \mathbb{F}^M, (x_0, B, (\varpi, c)))$ is admissible.

Assumption 2.10 guarantees that the relation between the Brownian motions $B, W^1, \ldots, W^N$ and the filtration to which the controls are adapted is not affected by the processes ϖ and c. In particular, B and W^j are respectively an $\mathbb{F}^M$-Brownian motion and an $\mathbb{F}^{S,j}$-Brownian motion, for every j. In general, the admissibility property, introduced in Definition 2.9, is fundamental to guarantee that anticipative controls are excluded. We are going to discuss about this condition in Remark 2.16 below.

2.2.4 The market clearing condition

2.2.4.1 A forward-backward system for the solution to the optimal control problems

In order to apply the stochastic maximum principle, we introduce the reduced Hamiltonians of the stochastic optimal control problem for both the standard agents (denoted by H) and the major agent (denoted by H^0). By the symmetry of the problem of standard agent, the Hamiltonian H is the same for every standard player. For the major player the Hamiltonian H^0 depends also on c:

$$H(t, x, y, \alpha, \varpi) = y(\alpha + l(t, \varpi)) + f(t, x, \alpha, \varpi),$$
$$H^0_{(N)}(t, x^0, y^0, \beta, \varpi, c) = y^0\beta + f^0_{(N)}(t, x, \beta, \varpi, c).$$

Under Assumption 2.4, in both cases, the reduced Hamiltonian is convex in the control variable. Hence, there exists a unique minimizer, that have the following structure:

$$\begin{aligned}\widehat{\alpha}(\varpi, y) &:= -\overline{\Lambda}(y+\varpi),\\ \widehat{\beta}^{(N)}(\varpi, y^0) &:= -N\overline{\Lambda}^0(y^0+\varpi).\end{aligned} \tag{2.11}$$

where $\overline{\Lambda} := -\Lambda^{-1}$ and analogously for Λ^0. Hence, to apply the stochastic maximum principle, in line with [CD18b, Section 1.4], we introduce the following FBSDE systems:

$$\begin{cases} dX_t^j = (-\overline{\Lambda}(Y_t^j + \varpi_t) + l(t, \varpi_t))dt + \sigma_0(t, \varpi_t)dB_t + \sigma(t, \varpi_t)dW_t^j, \\ X_0^j = \xi^j, \\ dY_t^j = -\partial_x \overline{f}(t, X_t^j, \varpi_t)dt + Z_t^{0,j}dB_t + Z_t^j dW_t^j + dM_t^j, \\ Y_T^j := \partial_x g(X_T^j, \varpi_T), \end{cases} \tag{2.12}$$

for each $j = 1, \dots, N$.

$$\begin{cases} dX_t^0 = -N\overline{\Lambda}^0(Y_t^0 + \varpi_t)dt + \sigma_0(t, \varpi_t)dB_t, \\ X_0^0 = x_0, \\ dY_t^0 = -\overline{c}_0^M(t, \varpi_t, c_t)dt + Z_t^0 dB_t + d\widetilde{M}_t^0, \\ Y_T^0 = \overline{g}_1^0(\varpi_T, c_T). \end{cases} \tag{2.13}$$

Let us remark that the martingale M^j appears because the process ϖ may be not adapted to $\mathbb{F}^{B,W^j}$.

By the compatibility condition, ensured by Assumption 2.10, the convexity of the Hamiltonian, guaranteed by condition A5, Assumption 2.4 and Assumption 2.6, we can apply the stochastic maximum principle in the version of [CD18b, Theorem 1.60]. We deduce that:

- Systems (2.12) and (2.13), defined on the filtered space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ introduced in (2.1), admit a unique strong solution respectively denoted by $(X^j, Y^j, Z^{0,j}, Z^j, M^j)$ and (X^0, Y^0, Z^0, M^0).
- The unique optimal controls of the optimal control problems introduced in Subsection 2.2.2.1 and Subsection 2.2.2.2 are respectively defined by:

$$\begin{aligned}\widehat{\alpha}_t^j &:= -\overline{\Lambda}(Y_t^j + \varpi_t),\\ \widehat{\beta}_t^{(N)} &:= -N\overline{\Lambda}^0(Y_t^0 + \varpi_t),\end{aligned}$$

Adopting the terminology introduced in [CD18a, Definition 2.14], we call *adjoint process associated with the optimal control prbolem*, the solution to the backward stochastic differential equation (BSDE) in a FBSDE system associated with the stochastic maximum principle. In particular Y^j and Y^0 in (2.12) and (2.13) are respectively the adjoint processes associated with the optimal controls $\widehat{\alpha}$ and $\widehat{\beta}^{(N)}$.

We briefly recall how the adjoint processes are constructed. As described in [EPQ97, Section 5.1], the adjoint processes are defined by the conditional expectations:

$$
\begin{aligned}
Y_t^j &= \mathbb{E}\left[\partial_x g(X_T^j, \varpi_T) + \int_t^T \partial_x \overline{f}(s, X_s^j, \varpi_t) ds \middle| \mathcal{F}_t^{S,j}\right], \quad t \in [0,T], \\
Y_t^0 &= \mathbb{E}\left[\overline{g}_1^0(\varpi_T, c_T) + \int_t^T \overline{c}_0^M(s, \varpi_s, c_s) ds \middle| \mathcal{F}_t^M\right], \quad t \in [0,T].
\end{aligned}
$$

Indeed, considering the case of Y^j (for Y^0 the computations are analogous), we introduce the martingale:

$$
N_t^j = \mathbb{E}\left[\partial_x g(X_T^j, \varpi_T) + \int_0^T \partial_x \overline{f}(s, X_s^j, \varpi_s) ds \middle| \mathcal{F}_t^{S,j}\right], \quad t \in [0,T]. \tag{2.14}
$$

By the Kunita-Watanabe decomposition theorem ([KW67, Proposition 4.1]), N_t can be decomposed as:

$$
N_t^j = N_0 + \int_0^t \widetilde{Z}_s^{0,j} dB_s + \int_0^t \widetilde{Z}_s^j dW_s^j + \widetilde{M}_t^j, \quad t \in [0,T],
$$

for suitable predictable process $\widetilde{Z}^{0,j}$, $\widetilde{Z}^j$ and a càdlàg local martingale $\widetilde{M}^j$, ortogonal to B and W^j. We define the random process by $\widetilde{Y}_t^j := N_t^j - \int_0^t \partial_x \overline{f}(s, X_s^j, \varpi_s) ds$. We notice that:

$$
\widetilde{Y}_t^j = \mathbb{E}\left[\partial_x g(X_T^j, \varpi_T) + \int_t^T \partial_x \overline{f}(s, X_s^j, \varpi_s) ds \middle| \mathcal{F}_t^{S,j}\right], \quad t \in [0,T].
$$

Moreover:

$$
\begin{aligned}
\widetilde{Y}_t^j &= N_t^j - \int_0^t \partial_x \overline{f}(s, X_s^j, \varpi_s) ds \\
&= N_0^j + \int_0^t \widetilde{Z}_s^{0,j} dB_s + \int_0^t \widetilde{Z}_s^j dW_s^j + \widetilde{M}_t^j - \int_0^t \partial_x \overline{f}(s, X_s^j, \varpi_s) ds \\
&= \widetilde{Y}_0^j - \int_0^t \partial_x \overline{f}(s, X_s^j, \varpi_s) ds + \int_0^t \widetilde{Z}_s^{0,j} dB_s + \int_0^t \widetilde{Z}_s^j dW_s^j + \widetilde{M}_t^j.
\end{aligned}
$$

This implies that:

$$
\widetilde{Y}_t^j = \widetilde{Y}_T^j + \int_t^T \partial_x \overline{f}(s, X_s^j, \varpi_s) ds - \int_t^T \widetilde{Z}_s^{0,j} dB_s - \int_t^T \widetilde{Z}_s^j dW_s^j - \widetilde{M}_T^j + \widetilde{M}_t^j.
$$

Finally, by uniqueness of solutions (see [EPQ97, Theorem 5.1]), it is possible to conclude that $(\widetilde{Y}^j, \widetilde{Z}^{0,j}, \widetilde{Z}^j, \widetilde{M}^j) = (Y^j, Z^{0,j}, Z^j, M^j)$.

Remark 2.11. Under Assumption 2.4 and Assumption 2.6, [CD18b, Theorem 1.60] guarantees also that the stochastic integrals $\int_\cdot^T Z_s^{0,j} dB_s$, $\int_\cdot^T Z_s^j dW_s^j$ and M^j are true martingales, because $Z^{0,j}, Z^j \in \mathbb{H}^2([0,T], \mathbb{R})$ while $M^j \in \mathbb{S}^2([0,T], \mathbb{R})$. The same holds for the solution to the FBSDE associated with the optimal control of the major player.

2.2.4.2 The equilibrium price process

Until now, we have assumed that the price process ϖ is exogenously given, i.e., every standard agent considers ϖ as an external stochastic process that affects the dynamics of the state variable and the objective functional. Under this assumption, the setting consists of a family of stochastic optimal control problems that can be solved separately by each player. As in [FT22a], we want to determine the price process by the equilibrium between demand and supply. This condition, called *market clearing*, is expressed by

$$\beta_t + \sum_{j=1}^{N} \alpha_t^j = 0, \quad \mathrm{dt} \otimes \mathrm{d}\mathbb{P}\text{-a.e.} \tag{2.15}$$

We recall that $\alpha^j \in \mathbb{H}^2(\mathbb{F}^{S,j};\mathbb{R})$, while $\beta \in \mathbb{H}^2(\mathbb{F}^M;\mathbb{R})$. Assuming that agents are rational, the market clearing condition (2.15) implies that $\widehat{\beta}^{(N)}$ must a posteriori be adapted to the filtration generated by the controls $\widehat{\alpha}_t^1, \ldots, \widehat{\alpha}_t^N$, where $\widehat{\alpha}^j$ and $\widehat{\beta}^{(N)}$ denote the optimal controls introduced in (2.11), respectively for the j^{th} standard player and the major player. As a consequence, $\widehat{\beta}^{(N)}$ is adapted to $(\mathcal{F}_t^{\varpi,B,W} \wedge \mathcal{F}_t^{\varpi,B,c})_{t\in[0,T]}$, where $W := (W^1, \ldots, W^N)$ is a Brownian motion defined on $\prod_{i=1}^N \Omega^i$. Hence, it is convenient to show the following result:

Lemma 2.12. *Under the conditions introduced above, the following holds:*

$$\mathcal{F}_t^{\varpi,B,W} \wedge \mathcal{F}_t^{\varpi,B,c} = \mathcal{F}_t^{\varpi,B}, \quad t \in [0,T].$$

Proof. Let us notice that elements of $\mathcal{F}_t^{\varpi,B,W}$ and $\mathcal{F}_t^{\varpi,B,c}$ are generated respectively by sets defined as follows:

$$A := \left\{ (\omega^0,\omega) \in \Omega^0 \times \prod_{i=1}^N \Omega^i : \ \varpi_{\cdot\wedge t}(\omega^0,\omega) \in A^1, \ B_{\cdot\wedge t}(\omega^0) \in A^2, \ W_{\cdot\wedge t}(\omega) \in A^3 \right\},$$
$$D := \left\{ (\omega^0,\omega) \in \Omega^0 \times \prod_{i=1}^N \Omega^i : \ \varpi_{\cdot\wedge t}(\omega^0,\omega) \in D^1, \ B_{\cdot\wedge t}(\omega^0) \in D^2, \ c_{\cdot\wedge t}(\omega^0) \in D^3 \right\},$$

for $A^1, D^1, D^3 \in \mathcal{B}(\mathcal{D}([0,T],\mathbb{R}))$ and $A^2, A^3, B^2 \in \mathcal{B}(\mathcal{C}([0,T],\mathbb{R}))$. Thus, an element of $\mathcal{F}_t^{\varpi,B,W} \wedge \mathcal{F}_t^{\varpi,B,c}$ must be generated by sets that are of the form of both A and D. This implies that $A^1 = D^1$ and $A^2 = D^2$. We now observe that:

$$A = A \cap \left\{ (\omega^0,\omega) \in \Omega^0 \times \prod_{i=1}^N \Omega^i : \quad c_{\cdot\wedge t}(\omega^0) \in \mathcal{D}([0,T],\mathbb{R}) \right\},$$
$$D = D \cap \left\{ (\omega^0,\omega) \in \Omega^0 \times \prod_{i=1}^N \Omega^i : \quad W_{\cdot\wedge t}(\omega) \in \mathcal{C}([0,T],\mathbb{R}) \right\}.$$

Therefore:

$$A = A \cup D = \{(\omega^0,\overline{\omega}) \in \Omega : \ \varpi_{\cdot\wedge t}(\omega^0,\overline{\omega}) \in A^1, \ B_{\cdot\wedge t}(\omega^0) \in A^2\},$$

from which we conclude that elements of $\mathcal{F}_t^{\varpi,B,W} \wedge \mathcal{F}_t^{\varpi,B,c}$ are generated by intersection between pre-images of $\varpi_{\cdot\wedge t}$ and of $B_{\cdot\wedge t}$. As a consequence, we conclude that $\mathcal{F}_t^{\varpi,B,W} \wedge \mathcal{F}_t^{\varpi,B,c} = \mathcal{F}_t^{\varpi,B}$. $\square$

Remark 2.13. In this setting, at the equilibrium, the optimal strategy of the major player is actually revealed by the price process. As a consequence, by the market clearing condition the optimal control of the problem defined by the class of admissible controls $\mathbb{H}^2(\mathbb{F}^{\varpi,B,c};\mathbb{R})$ is adapted to $\mathbb{F}^{\varpi,B}$. Hence, we can restrict family of admissible controls of the major player to $\mathbb{H}^2(\mathbb{F}^{\varpi,B};\mathbb{R})$.

Moreover, the strategy $\widehat{\beta}^{(N)}$ cannot be independent of ϖ. In other words, the market clearing condition imposes a posteriori the major player to be a market maker. This implies that the major player is supposed to be the one who balances the demand and supply of the asset in the market. As a consequence, $\widehat{\beta}^{(N)}$ must be a measurable function of the equilibrium price process ϖ and the Brownian motion B.

By equation (2.11) together with Lemma 2.12 we conclude that the adjoint process Y^0 of the major player is adapted to $\mathbb{F}^{\varpi,B}$. As a consequence, the independence between c and B^1 implies that the coefficients $\overline{g}_1^0$ and $\overline{c}_0^M$ explicitly depend only on ϖ. For instance, since Y_T^0 is $\mathcal{F}_T^{\varpi,B}$-measurable, the coefficient $\overline{g}_1^0(\varpi_T, c_T) = Y_T^0 = g_1^0(\varpi)$, where g_1^0 is a measurable function of the whole trajectory of ϖ. To develop the computations we assume that g_1^0 depends only on ϖ_T. Reasoning analogously for $\overline{c}_0^M$, we can formulate the following assumption:

Assumption 2.14. We assume that $\overline{c}_0^M(t,\varpi,c) = c_0^M(t,\varpi)$ and $\overline{g}_1^0(\varpi,c) = g_1^0(\varpi)$. Moreover, in accordance with Assumption B2, we suppose that:

B4 c_0^M and g_1^0 are bounded (by the same constant L introduced in Assumption B2) and continuous functions of (t,ϖ) and ϖ respectively.

As a consequence, the dynamics of the adjoint process Y^0 is:

$$\begin{cases} dY_t^0 &= -c_0^M(t,\varpi_t)dt + Z_t^0 dB_t + dM_t^0, \\ Y_T^0 &= g_1^0(\varpi_T). \end{cases} \tag{2.16}$$

where c_0^M and g_1^0 are measurable functions of (t,ϖ) and ϖ respectively. M^0 is a martingale adapted to $\mathbb{F}^{\varpi}$. Moreover, we suppose additionally that: The market clearing condition (2.15), applied to the optimal controls introduced in (2.11), defines an equation for the price process. Indeed, the following holds:

$$0 = \widehat{\beta}_t^{(N)} + \sum_{j=1}^N \widehat{\alpha}_t^j = -N\Big(\frac{\overline{\Lambda}}{N}\sum_{j=1}^N (Y_t^j + \varpi_t) + \overline{\Lambda}^0(Y_t^0 + \varpi_t)\Big).$$

As a consequence, the market clearing price process is:

$$\varpi_t = -(\overline{\Lambda} + \overline{\Lambda}^0)^{-1}\Big(\frac{1}{N}\overline{\Lambda}\sum_{j=1}^N Y_t^j + \overline{\Lambda}^0 Y_t^0\Big). \tag{2.17}$$

As in [FT22b], the presence of the normalization factor N in the denominator of the constant Λ^0 guarantees that the strategy of $\widehat{\beta}^{(N)}$ does not become negligible when $N >> 0$.

We are now in a position to give the formal definition of equilibrium price process for the market model populated by $N+1$ agents:

Definition 2.15. We say that ϖ is a *equilibrium price process* if it is a fixed-point of the functional $\Phi^{(N)} : \mathbb{H}^2(\mathbb{F}, \mathbb{R}) \to \mathbb{H}^2(\mathbb{F}, \mathbb{R})$, defined as follows:

$$\Phi_t^{(N)}(\varpi) := -(\overline{\Lambda} + \overline{\Lambda}^0)^{-1}\Big(\frac{1}{N}\overline{\Lambda}\sum_{j=1}^{N} Y_t^{\varpi,j} + \overline{\Lambda}^0 Y_t^{\varpi,0}\Big), \quad t \in [0,T].$$

where $Y^{\varpi,j}$ and $Y^{\varpi,0}$ denote respectively the adjoint process of the j^{th} standard player and the major player, when the exogenous random environment is given by ϖ.

2.2.5 The mean-field limit

2.2.5.1 The formulation of the fixed-point problem

The problem of price formation can be formulated by the existence of a solution to (2.17). However, by market clearing condition, an equilibrium price defined as solution to (2.17) makes the optimal control problems introduced in Section 2.2.2.1 and in Section 2.2.2.2 highly recursive. Indeed, if the stochastic process ϖ, that appears in (2.3) and (2.6), is the solution to (2.17), it is not even clear how to guarantee that the cost functionals, that depend on ϖ, are well-defined and convex with respect to the control variables.

The complexity of the problem is due to the presence of the idiosyncratic noises. To overcome this problem, we recall that the agents are price taker. This implies that the effect of the trading activities of each single agent is negligible, when the number N of standard agents becomes large. To have some insights on a possible strategy to face the problem of price formation and to simplify the structure of the model, it is convenient to study the mean-field limit. To pass to the limit in the number of standard agents, we exploit the fact that the optimal control problem of the standard agents is symmetric and the only difference is given by the idiosyncratic noises that are pairwise independent. By the symmetry of the optimal control problems of the standard players, we can apply the Yamada-Watanabe theorem in the version of [CD18b, Theorem 1.33]. As a consequence, there exists a progressively measurable function $\Phi^Y : \mathbb{R} \times \mathcal{C}([0,T], \mathbb{R}) \times \mathcal{D}([0,T], \mathbb{R}) \times \mathcal{C}([0,T], \mathbb{R}) \to \mathcal{D}([0,T], \mathbb{R})$ such that, the adjoint process Y^j of the j^{th} player, defined by an exogenously given price process ϖ, satisfies

$$Y^j = \Phi^Y(\xi^j, B, \varpi, W^j). \tag{2.18}$$

If $(\xi^j)_{j\in\mathbb{N}}$ is a sequence of i.i.d. random variables, distributed like initial conditions introduced in Section 2.2.2.1 and $(W^j)_{j\in\mathbb{N}}$ is a sequence of pairwise independent Brownian motions, independent also of $(\xi^j)_{j\in\mathbb{N}}$, the sequence of $(\xi^j, B, \varpi, W^j)_{j\in\mathbb{N}}$ is exchangeable in the sense of [Kle13, Definition 12.1]. Therefore, the sequence $(Y_t^j)_{j=1}^N$ defined by (2.18) is exchangeable too. Hence, applying De

Finetti representation theorem ([CD18b, Theorem 2.1], [Kle13, Theorem 12.26]), the limit for N going to infinity of $\frac{1}{N}\sum_{i=1}^{N} Y_t^i$ satisfies:

$$\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} Y_t^i = \mathbb{E}\left[Y_t^1 \middle| \bigcap_{j\geq 1} \sigma\{Y_t^k,\ k\geq j\}\right], \quad a.s. \tag{2.19}$$

We guess that the sigma-algebra $\bigcap_{j\geq 1}\sigma\{Y_t^k,\ k\geq j\}$ is given by the common stochastic factors of all the random variables Y_t^k, that are $\varpi_{\cdot\wedge t}$ and $b_{\cdot\wedge t}$. It would be natural to substitute the empirical mean in equation (2.17) with the conditional expectation $\mathbb{E}[Y_t^1|\mathcal{F}_t^{\varpi,B}]$. Through this substitution, we can consider a market populated by a single typical standard agent and a major agent. Indeed, the mean-field equilibrium price process ϖ^{mf}, defined as the solution to the mean-field limit of (2.17), should be defined in a suitable probabilistic setup $(\Omega,\mathcal{F},\mathbb{P},\mathbb{F})$ by the following equation:

$$\varpi_t^{\mathrm{mf}} = -(\overline{\Lambda}+\overline{\Lambda}^0)^{-1}\Big(\overline{\Lambda}\mathbb{E}[Y_t|\mathcal{F}_t^{\varpi^{\mathrm{mf}},B}] + \overline{\Lambda}^0 Y_t^0\Big), \quad \forall t\in[0,T],\ \mathbb{P}-a.s. \tag{2.20}$$

In (2.20), Y is the adjoint process associated with the optimal control problem of a typical standard agent in the mean-field limit. Analogously, Y^0 is the adjoint process associated with the optimal control problem of the major agent in the mean-field limit. In particular, on $(\Omega,\mathcal{F},\mathbb{P},\mathbb{F})$ we introduce the optimal control problem for the typical standard agent using the same coefficients introduced in Section 2.2.2.1. For the major agent, we proceed analogously, introducing a normalized optimal control problem as the one introduced in Section 2.2.2.2, where the quadratic term of the running cost $f_{(N)}^0$ is not divided by the factor N. Hence, the cost functional for the major player in the mean-field limit is

$$J^0(\beta) := \mathbb{E}\left[\int_0^T f^0(t,X_t^0,\beta_t,\varpi_t^{\mathrm{mf}})dt + g^0(X_T^0,\varpi_T^{\mathrm{mf}})\right], \quad j=1,\dots,N, \tag{2.21}$$

where

$$\begin{aligned} f^0(t,x,b,\varpi) &:= \varpi b + \frac{1}{2}\Lambda^0 b^2 + c_0^M(\varpi)x + c_1^M(\varpi), \\ g^0(x,\varpi) &:= g_1^0(\varpi)x + g_2^0(\varpi). \end{aligned}$$

When we pass to the mean-field limit, it is natural to suppose that the equilibrium price process ϖ^{mf} is independent of the idiosyncratic Brownian motion affecting the typical standard agent. Indeed, passing to the limit the effects of the idiosyncratic noises of every standard agent cancel out.

We want to provide conditions that guarantee the existence of a stochastic process ϖ^{mf} that satisfies the following

P-I On an admissible probabilistic setup $(\Omega,\mathcal{F},\mathbb{P},\mathbb{F})$, we consider:

- the stochastic optimal control problem of the typical standard agent determined by the coefficients introduced in (2.3)

$$\inf_{\alpha\in\mathbb{H}^2(\mathbb{F}^{\varpi,B,W^j};\mathbb{R})} J(\alpha)$$

$$\begin{cases} dX_t = (\alpha_t + l(t,\varpi_t))dt + \sigma_0(t,\varpi_t)dB_t + \sigma(t,\varpi_t)dW_t, \\ X_0 = \xi. \end{cases}$$

- By Remark 2.13, the stochastic optimal control problem determined by the coefficients introduced in (2.21) for the major agent is

$$\inf_{\beta\in\mathbb{H}^2(\mathbb{F}^{\varpi,B};\mathbb{R})} J^0(\beta)$$

$$\begin{cases} dX^0_t = \beta_t dt + \sigma^0(t,\varpi_t)dB_t, \\ X^0_0 = x_0. \end{cases}$$

Under Assumption 2.4 and Assumption 2.6, we introduce the FBSDE systems, associated with the stochastic maximum principle for the standard typical player:

$$\begin{cases} dX_t = (-\overline{\Lambda}(Y_t+\varpi_t) + l(t,\varpi_t))dt + \sigma_0(t,\varpi_t)dB_t + \sigma(t,\varpi_t)dW^1_t, \\ X_0 = \xi, \\ dY_t = -\partial_x\overline{f}(t,X_t,\varpi_t)dt + Z^{0,1}_t dB_t + Z_t dW_t + dM_t, \\ Y_T := \partial_x g(X_T,\varpi_T), \end{cases} \tag{2.22}$$

and the major player:

$$\begin{cases} dX^0_t = -\overline{\Lambda}^0(Y^0_t+\varpi_t)dt + \sigma^M_0(t,\varpi_t)dB_t, \\ X^0_0 = x_0, \\ dY^0_t = -c^M_0(t,\varpi_t)dt + Z^0_t dB_t + dM^0_t, \\ Y^0_T := g^0_1(\varpi_T), \end{cases} \tag{2.23}$$

P-II We consider the stochastic process

$$\Phi_t(\varpi) := -(\overline{\Lambda}+\overline{\Lambda}^0)^{-1}\Big(\mathbb{E}[\overline{\Lambda}Y_t + \overline{\Lambda}^0 Y^0_t|\mathcal{F}^{\varpi,B}_t], \quad t\in[0,T].$$

We aim at proving the existence of a stochastic process ϖ^{mf} such that: $\Phi_t(\varpi^{\mathrm{mf}}) = \varpi^{\mathrm{mf}}_t$ a.s., for all $t\in[0,T]$.

The structure of the solution to (2.20) is complicated due to the presence of the unknown stochastic process, ϖ^{mf}, in both the left and the right hand-side. In the right hand-side it appears both in Y and Y^0 as well as in the filtration on which we are conditioning.

We can remark an analogy between the structure of equation (2.20) and the consistency condition for a weak mean-field game equilibrium in the presence of common noise. This notion, introduced in [CDL16, Definition 3.1] (see also [CD18b, Definition 2.24] and [Lac16, Definition 2.1]), describes a probability distribution μ on the canonical space, that is a version of the condition law of (W, X) (where W is the idiosyncratic noise and X is the optimal state variable) given the realization of the common noise B and the solution itself μ. In other words:

$$\mu = \mathcal{L}(W, X|B, \mu). \tag{2.24}$$

We propose an adaptation of the results described in [CD18b, Chapter III], that are stronger than the one proposed in [CDL16], but allow for the forward-backward formulation of the optimal control problems of each player. In [CD18b, Chapter III], the construction of the solution to (2.24) is performed by discretizing the space $\mathcal{H}$ on which the solution takes values, in order to obtain a sequence of approximated solutions. Adapting this approach to our problem, we consider a random process ϖ taking values on a suitable functional space $\mathcal{H}$. We observe that by the Yamada-Watanabe theorem (see [CD18b, Theorem 1.33]), if ϖ is supposed to be exogenously given, there exists two progressively measurable functionals Ψ^S and Ψ^M such that the processes (X, Y) and (X^0, Y^0), respectively defined in (2.22) and (2.23), satisfy the following functional form

$$\begin{aligned}(X, Y) &= \Psi^S(\xi, (B, W), \varpi),\\ (X^0, Y^0) &= \Psi^M(B, \varpi).\end{aligned}$$

Since the presence of the càdlàg martingales terms in the decomposition of Y and Y^0, we cannot conclude that $\mathcal{H}$ is the space of continuous functions $\mathcal{C}([0,T], \mathbb{R})$, because ν may takes values on $\mathcal{D}([0,T];\mathbb{R})$. As a consequence, we suppose that $\mathcal{H} = \mathcal{D}([0,T];\mathbb{R})$.

Our purpose is to construct a random variable ϖ^{mf}, taking values on a suitable functional space $\mathcal{H}$, as a fixed point of the functional Φ, defined by

$$\begin{aligned}\Phi : \mathcal{D}([0,T];\mathbb{R})^{\Omega^0} &\to \mathcal{D}([0,T];\mathbb{R})^{\Omega^0}\\ \varpi &\mapsto (-(\overline{\Lambda} + \overline{\Lambda}^0)^{-1}\mathbb{E}[\overline{\Lambda} Y_t + \overline{\Lambda}^0 Y_t^0 | \mathcal{F}_t^{\varpi, B}])_{t\in[0,T]}\end{aligned}$$

2.2.5.2 The definition of mean-field equilibrium price process

As discussed in [CD18b, Chapter II], when Ω^0 is not countable and $\mathcal{H}$ is $\mathcal{D}([0,T], \mathbb{R})$, the characterization of the compact sets of $\mathcal{H}^{\Omega^0}$ is complicated. As a consequence, we cannot use standard fixed-point arguments, like Schauder's theorem, to provide the existence of a solution for equation (2.20). To overcome this problem, we adapt the strategies described in [CD18b, Chapter III] and in [CDL16, Section 3] and proceed as follows:

1. We consider an admissible probabilistic setup in the sense of Definition 2.9. In particular, a random process (ξ, B, W, ϖ) satisfying the assumptions of Definition 2.9 is defined. We discretize with n steps in space and l in time the trajectories of B and ϖ. We construct a

fixed point on $\mathcal{D}([0,T];\mathbb{R})^{\widetilde{\mathbb{J}}}$, where $\widetilde{\mathbb{J}}$ is the (finite) discretization of the image of the processes B and ϖ. Since $\mathcal{D}([0,T];\mathbb{R})^{\widetilde{\mathbb{J}}}$ is a finite-product of copies of the functional space $\mathcal{D}([0,T];\mathbb{R})$, we can apply Schauder's fixed point theorem to construct a solution to the fixed point problem.

2. We apply the previous step for each n and l to obtain a sequence of approximated solutions. Proving that this sequence has a weak limit, we can find the conditions that guarantee the weak limit to be a solution to equation (2.20).

Remark 2.16. The procedure described above to construct the solution to equation (2.20) as the weak limit of a sequence of discretized game involves the issue of compatibility. Indeed, as described in [CD18b, Section 2.2.2], it is not sufficient to require the compatibility condition for the optimal control problems defined for the agents in the discretized setting, because the compatibility condition is in general not preserved when passing to a weak limit. As we show in Section 2.3.3.3, we have to lift the sequence of the fixed point obtained in the discretized space. This guarantees that the optimality of the sequence of optimal trajectories computed in the discretized setting is preserved in the weak limit. As a consequence, the stability of the weak equilibria, as solutions to the optimal control problem in the probabilistic setup in which the weak limit is defined, is maintained. Therefore, to ensure that the compatibility condition is conserved in the weak limit, we are going to add to the approximated price process, defined by the fixed point obtained in the discretized space, the adjoint process associated with the optimal control problem of the standard player in the discretized setting. As we are going to prove, this is sufficient to guarantee that the weak limit of the sequence of optimal state variables for the standard player in the discretized setting will result adapted to the filtration generated by the weak limit of the random processes driving the dynamics of the game. For what regards the major player, the affine structure of the cost functional in the x variable allows to avoid these issues, as we are going to show in Section 2.3.3.5.

By Remark 2.16, we need to slightly change the structure of the optimal control problems defined in P-I. We consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ on which a process $(\xi, (B, W^1), \mathcal{W})$ is defined. In particular, $\mathcal{W} := (\varpi, \overline{Y})$ is a random variable taking values on $\mathcal{D}([0,T], \mathbb{R}^2)$. For the moment, $\overline{Y}$ is another stochastic process defined on $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$. We assume that $\mathbb{F}$ is a right continuous and complete filtration, a priori bigger than $\mathbb{F}^{\xi,B,W^1,\mathcal{W}}$, such that $((\Omega, \mathcal{F}, \mathbb{P}), \mathbb{F}, (\xi, (B, W^1), \mathcal{W}))$ is admissible in the sense of Definition 2.9. We notice that for $((\Omega, \mathcal{F}, \mathbb{P}), \mathbb{F}^{\varpi,B}, (x_0, B, \varpi))$ admissibility has already guaranteed, since $\mathbb{F}^{\varpi,B}$ is the generated by (ϖ, B). Let us introduce a relaxed optimal control problem for the typical standard agent as follows

$$J^{\varpi}(\gamma) = \inf_{\gamma \in \mathbb{H}^2(\mathbb{F};\mathbb{R})} \mathbb{E}\left[\int_0^T f(s, X_s, \varpi_s, \gamma_s)ds + g(X_T, \varpi_T)\right]$$
$$\begin{cases} dX_t = (\gamma_t + l(t, \varpi_t))dt + \sigma_0(t, \varpi_t)dB_t + \sigma(t, \varpi_t)dW_t^1, \\ X_0 = \xi. \end{cases} \tag{2.25}$$

The major agent must solve the following optimal control problem

$$J^{0,\varpi}(\beta) = \inf_{\beta \in \mathbb{H}^2(\mathbb{F}^{\varpi,B};\mathbb{R})} \mathbb{E}\left[\int_0^T f^0(s, X_s, \varpi_t, \beta_s)ds + g^0(X_T, \varpi)\right]$$
$$\begin{cases} dX_t = \beta_t dt + \sigma^M(t, \varpi_t)dB_t, \\ X_0 = x_0. \end{cases} \tag{2.26}$$

The FBSDE systems associated with the stochastic maximum principle are denoted by (X, Y, Z^0, Z, M) for (2.25) and $(X^0, Y^0, Z^{0,0}, M^0)$ for (2.26). The optimal control for the typical standard agent is different from the analogous problem introduced in P-I. Indeed, the functional costs are the same, but the filtration $\mathbb{F}$ on which the controls γ are adapted to is supposed not to be a priori the one generated by $(\xi, (B, W^1), \mathcal{W})$, but only compatible with the lifted process $(\xi, (B, W^1), \mathcal{W})$.

In conclusion, a mean-field equilibrium price process as defined as follows:

Definition 2.17 (Mean-field equilibrium price process)**.** We say that $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F}, \xi, (b, w), (\varpi, \overline{Y}))$ is a *mean-field equilibrium price process* if:

- $((\Omega, \mathcal{F}, \mathbb{P}), \mathbb{F}, (\xi, (B, W^1), \mathcal{W}))$ and $((\Omega, \mathcal{F}, \mathbb{P}), \mathbb{F}^{\varpi,B}, (x_0, B, \varpi))$ are admissible in the sense of Definition 2.9;
- ϖ solves equation (2.20), where Y^0 and Y are the adjoint processes associated with the optimal controls of respectively (2.26) and (2.25). We may refer to this property as *consistency condition for the equilibrium price process.*
- $\overline{Y} = Y$.

In Section 2.3, we are going to prove that:

Theorem 2.18. *Under Assumption 2.4 and Assumption 2.6, there exists a mean-field equilibrium price process in the sense of Definition 2.17, with ξ satisfying Assumption 2.2.*

Finally, let us remark that, assuming additionally that the cost functions of the standard agents are affine functions in the x variable, we obtain also stronger results. In this setting, we can introduce a stronger version of Definition 2.17, for which it is not necessary to lift the price process adding $\overline{Y}$. We refer to Section 2.4 below for the analysis of this specific case and in particular, the result concerning the existence of this stronger version of the equilibrium price process (Theorem 2.37).

2.3 Existence of solutions to the mean-field equilibrium price process

Our goal is to prove Theorem 2.18, constructing a sequence of discretized solutions defined on the canonical space that is tight. In this section, we present the conditions and the strategy that guarantee a weak limit of this sequence to be solution to equation (2.20). We proceed as follows:

1. As already mentioned, the main problem is the complexity in the characterization of the compact sets of $\mathcal{D}([0,T],\mathbb{R})^{\Omega^0}$, when Ω^0 is not countable. In Subsection 2.3.1, we present the discretization procedure that enables us to restrict the case $\mathcal{D}([0,T],\mathbb{R})^{\Omega^0}$ to $\mathcal{D}([0,T],\mathbb{R})^{\mathbb{J}}$, where $\mathbb{J}$ is a finite set, depending on two natural numbers n and l. These natural numbers represent the discretization step in space and time respectively. In other words, we are considering càdlàg random processes on finite probability spaces. In this setting, we define an input-output map to reproduce the structure of the equilibrium.

2. In Subsection 2.3.2 we provide, under suitable conditions, the existence of a fixed point for the input-output functional. This fixed point plays the role of a discretized price process. To prove this result, we can apply Schauder's theorem, since the restricted space on which is defined the input-output map is a Polish space.

3. In Subsection 2.3.3 we state the main results: we consider the fixed points $(\varpi^{n,l})_{n,l}$ of the input-output functional as well as the solutions, introduced in the discretized setting, of the state variables $(X^{n,l}, X^{0,n,l})_{n,l}$ associated with the optimal controls. We show that $(\varpi^{n,l})_{n,l}$ and $((X^{n,l}, X^{0,n,l}))_{n,l}$ form tight sequences. Therefore, we show the discretized equilibria are stable, in the sense that there exists a weak limit of $(\varpi^{n,l})_{n,l}$, that is an equilibrium price process for a the optimal control problems, optimally solved by the weak limits of $((X^{n,l}, X^{0,n,l}))_{n,l}$. Finally, we prove that the weak limit we obtain as equilibrium price process satisfies equation (2.20).

2.3.1 The discretization procedure

We aim at constructing a sequence that is tight, in order to extract a weak limit. Thus, similarly to [CD18b, Section 3.3], we must move to the space of trajectories of the stochastic processes involved in the $(N+1)$-player game. To do so, we introduce the canonical spaces:

$$\overline{\Omega}^0 := \mathcal{C}([0,T];\mathbb{R}); \tag{2.27}$$

$$\overline{\Omega}^1 := [0,1) \times \mathcal{C}([0,T];\mathbb{R}). \tag{2.28}$$

On these spaces we define the canonical processes b for $\overline{\Omega}^0$ and (η, w) for $\overline{\Omega}^1$. We denote by $(\overline{\Omega}^0, \overline{\mathcal{F}}^0, \mathcal{W}_1)$, the probability space on which $\overline{\mathcal{F}}^0$ is the Borel sigma-algebra and $\mathcal{W}_1$ is the Wiener measure on $\mathcal{C}([0,T];\mathbb{R})$. Similarly, on $\overline{\Omega}^1$, we consider the Borel σ-algebra $\overline{\mathcal{F}}^1$ and the product measure $\text{Leb}_1 \otimes \mathcal{W}_1$. The complete and right-continuous augmentation of the canonical filtration is denoted by $\overline{\mathbb{F}}^0$ and $\overline{\mathbb{F}}^1$ for $\overline{\Omega}^0$ and $\overline{\Omega}^1$ respectively. Finally, the augmentation of the product space is denoted by $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}}, \overline{\mathbb{F}})$. In the following, we denote the expected value on $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}}, \overline{\mathbb{F}})$ by $\overline{\mathbb{E}}$.

In the next subsections, we are constructing a discretize price process on the filtered probability space introduced above.

2.3.1.1 Discretization of the common noise

We first present the discretization procedure, that is defined on a general finite-dimensional vector space $\mathbb{R}^d$. We consider two integers $l, n \geq 1$:

- l is the step size in the grid space;
- n is the step size in the grid time.

Denoting with $\lfloor x \rfloor$ the floor function applied to x, we introduce the following function:

$$\begin{aligned} \Pi_l^1 &: \mathbb{R} \to \mathbb{R} \\ & x \mapsto \begin{cases} 2^{-l}\lfloor x 2^l \rfloor, & \text{if } |x| \leq 2^l; \\ 2^l \text{ sign } (x), & \text{if } |x| < 2^l; \end{cases} \end{aligned}$$

and its multi-dimensional generalization:

$$\begin{aligned} \Pi_l^d &: \mathbb{R}^d \to \mathbb{R}^d \\ & x \mapsto \Pi_l^d(\overline{x}) := \begin{pmatrix} \Pi_l^1(x_1) & \cdots & \Pi_l^1(x_d) \end{pmatrix}. \end{aligned}$$

Moreover, we consider $\Pi_{l,j}^d : (\mathbb{R}^d)^j \to (\mathbb{R}^d)^j$, defined in the following iterated way:

$$\begin{aligned} \Pi_{l,1}^d &:= \Pi_l^d \\ \Pi_{l,j+1}^d(x^1, \ldots, x^{j+1}) &:= \Big(\overbrace{\Pi_{l,j}^d(x^1, \ldots, x^j)}^{:=(y^1,\ldots,y^j)}, \Pi_l^d\Big(\overbrace{(\Pi_{l,j}^d(x^1, \ldots, x^j))_j}^{=:y^j} + x^{j+1} - x^j\Big)\Big). \end{aligned}$$

We now state the following result, analogous to [CD18b, Lemma 3.17]:

Lemma 2.19. *With the notation introduced above, given $l \in \mathbb{N}$, for every $(x^1, \ldots, x^j) \in (\mathbb{R}^d)^j$ such that for all $i \in \{1, \ldots, j\}$ and $|x^i|_\infty := \max_{k=1,\ldots,d} |(x^i)_k| \leq 2^l - 1$, let:*

$$(y^1, \ldots, y^j) := \Pi_{l,j}^d(x^1, \ldots, x^j).$$

Then, the following holds:

$$\forall j \leq 2^l : \quad |x^i - y^i|_\infty \leq \frac{i}{2^l}, \quad \forall\, i \in \{1, \ldots, j\}.$$

Proof. By induction on i (we are fixing j and we do the induction on $i \leq j$):

(a) $i = 1$. $|x^1|_\infty \leq 2^l - 1$ holds, by hypothesis. Therefore, we consider the following

$$\begin{aligned} |x^1 - y^1|_\infty &:= |x^1 - \Pi_l^d(x^1)|_\infty = \max_{k=1,\ldots,d} |x_k^1 - \Pi_l^1(x_k^1)| \\ &= \max_{k=1,\ldots,d} |x_k^1 - 2^{-l}\lfloor 2^l x_k^1 \rfloor = \Big(\max_{k=1,\ldots,d} |2^l x_k^1 - \lfloor 2^l x_k^1 \rfloor|\Big) 2^{-l} \leq 2^{-l}. \end{aligned}$$

(b) $i \in \{2, \dots, j\}$ with $j \leq 2^l$ and $|x^{i-1} - y^{i-1}|_\infty \leq \frac{i-1}{2^l}$. In this case, we have that:

$$|y^{j-1} + x^i - x^{i-1}|_\infty \leq |x^i|_\infty + |y^{i-1} - x^{i-1}|_\infty \leq 2^l - 1 + \frac{i-1}{2^l} < 2^l - \frac{2^l - j}{2^l} \leq 2^l.$$

Therefore, applying the case $i = 1$, we obtain:

$$|(y^{i-1} + x^i - x^{i-1}) - \Pi_l^d(y^{i-1} + x^i - x^{i-1})|_\infty \leq \frac{1}{2^l}.$$

Finally:

$$\begin{aligned}|x^i - y^i|_\infty &\leq |x^i - (y^{i-1} + x^i - x^{i-1})|_\infty + |(y^{i-1} + x^i - x^{i-1}) - \Pi_l^d(y^{i-1} + x^i - x^{i-1})|_\infty \\ &\leq |y^{i-1} - x^{i-1}|_\infty + \frac{1}{2^l} \leq \frac{i-1}{2^l} + \frac{1}{2^l} = \frac{i}{2^l}.\end{aligned}$$

□

Given an integer n, let $N = 2^n$ and consider the diadic time mesh:

$$t_i = \frac{iT}{N}, \quad i \in \{0, \dots, N\}.$$

We define the discrete random variable, $(V_1, \dots, V_{N-1}) = \Pi_{l,N-1}^d(b_{t_1}, \dots, b_{t_{N-1}})$ and we adopt the notation

$$\overline{V}_j := (V_1, \dots, V_j), \quad j = 1, \dots, N. \tag{2.29}$$

where $(b_t)_{t\in[0,T]}$ is the canonical process on $\overline{\Omega}^0$. In this way, we have obtained a discrete random variable on $(\overline{\Omega}^0, \overline{\mathcal{F}}^0, \overline{\mathbb{P}}^0)$. In particular, we introduce also We recall now [CD18b, Lemma 3.18]:

Lemma 2.20. *Given $i = 1, \dots, N-1$ the random vector $(V_1, \dots, V_i)$ has support equal to $\mathbb{J}^i$, with:*

$$\mathbb{J} := \left\{ -\Lambda, -\Lambda + \frac{1}{\Lambda}, -\Lambda + \frac{2}{\Lambda}, \dots, \Lambda - \frac{1}{\Lambda}, \Lambda \right\}, \quad \Lambda := \frac{1}{2^l}.$$

2.3.1.2 Discretization of input-output map

In this section we want to define a price process ϖ on $\overline{\Omega}^0$ that is adapted to the discretization of the Brownian motion b and that satisfies a discrete version of the equilibrium condition introduced in equation (2.20). To do so, we introduce an input-output map, whose fixed point will be ϖ.

We introduce a discretized input that is a family $\overline{\boldsymbol{\theta}} := (\theta^0, \dots, \theta^{N-1})$ such that, for each $\forall i = 0, \dots, N-1$:

$$\theta^i : \mathbb{J}^i \to \mathcal{C}([t_i, t_{i+1}]; \mathbb{R}),$$

where $\mathbb{J}^0 = \emptyset$. Therefore, θ^0 is supposed to be constant in $\mathcal{C}([t_0, t_1]; \mathbb{R})$. In particular, $\overline{\boldsymbol{\theta}}$ can be thought as an element of $\prod_{i=0}^{N-1} \mathcal{C}([t_i, t_{i+1}]; \mathbb{R})^{\mathbb{J}^i}$

Alternatively, we can introduce a function $(\overline{\theta}_t)_{t\in[0,T]} \in \mathcal{D}([t_0,t_N],\mathbb{R})^{\mathbb{J}^{N-1}}$, defined as:

$$\begin{cases} \overline{\theta}_t(v_1,\dots,v_{N-1}) := (\theta^i(v_1,\dots,v_i))_t, \quad t\in[t_i,t_{i+1}), \quad i\in\{1,\dots,N-1\}, \\ \overline{\theta}_T(v_1,\dots,v_{N-1}) := (\theta^{N-1}(v_1,\dots,v_{N-1}))_T, \end{cases} \tag{2.30}$$

while $\overline{\theta}_t$ is assumed to be not dependent on $(v_1,\dots,v_{N_1})$ for $t\in[t_0,t_1]$.

Remark 2.21. Notice that $\overline{\boldsymbol{\theta}}$ and $(\overline{\theta}_t)_{t\in[0,T]}$ are two different objects for which there exists a one-to-one correspondence, given by the relation introduced in equation (2.30). However, it is convenient to introduce both these objets, because we are going to apply a fixed point argument on $\prod_{i=0}^{N-1}\mathcal{C}([t_i,t_{i+1}];\mathbb{R})^{\mathbb{J}^i}$, to deduce the existence of an element in $\mathcal{D}([t_0,t_N],\mathbb{R})^{\mathbb{J}^{N-1}}$ that satisfies suitable properties.

Through the input map, we can introduce a càdlàg stochastic process ϖ defined on $(\overline{\Omega}^0,\overline{\mathcal{F}}^0,\overline{\mathbb{P}}^0)$ as follows:

$$\varpi_t^\theta := \overline{\theta}_t(V_1,\dots,V_{N-1}), \quad t\in[0,T]. \tag{2.31}$$

We introduce a function ψ defined as

$$\begin{aligned} \psi : [0,1)\times\mathcal{P}_2(\mathbb{R}) &\to \mathbb{R}, \\ (\eta,\mu) &\mapsto \psi(\eta,\mu), \end{aligned}$$

to transport the initial distribution of the standard player's state variable ξ on the canonical space. Indeed, as proved in [CD18a, Lemma 5.29], for every square integrable probability measure μ there exists a measurable mapping ψ such that the image of the Lebesgue measure on $[0,1)$ by $\psi(\cdot,\mu)$ is μ itself. Therefore, if the law of the random variable ξ is denoted by by $\mathcal{L}(\xi)$, then $\psi(\cdot,\mathcal{L}(\xi))$ is a real valued random variable defined on the component $[0,1)$ of $\overline{\Omega}^1$, distributed like ξ. With an abuse of notation, we denote by ξ the random variable $\psi(\cdot,\mathcal{L}(\xi))$.

Let us notice that (ξ,b,w,ϖ^θ) is compatible with the canonical filtration on $(\overline{\Omega},\overline{\mathcal{F}},\overline{\mathbb{P}},\overline{\mathbb{F}})$, because ϖ^θ is adapted to the filtration generated by the Brownian motion b. Thus, we can introduce the optimal control problem for the typical standard player on $(\overline{\Omega},\overline{\mathcal{F}},\overline{\mathbb{P}},\overline{\mathbb{F}})$. The coefficients of this optimal control problem have been already introduced in Section 2.2.2.1 and the class of controls is given by $\mathbb{H}^2(\overline{\mathbb{F}};\mathbb{R})$. Hence, we introduce the FBSDE associated with the stochastic maximum principle, applied to the optimal control problem of the typical standard player:

$$\begin{cases} d\widetilde{X}_t = (-\overline{\Lambda}(\widetilde{Y}_t+\varpi_t^\theta)+l(t,\varpi_t^\theta))dt+\sigma^0(t,\varpi_t^\theta)db_t+\sigma(t,\varpi_t^\theta)dw_t, & \widetilde{X}_0=\xi, \\ d\widetilde{Y}_t = -\partial_x\overline{f}(t,X_t,\varpi_t^\theta)dt+\widetilde{Z}_t^0 db_t+\widetilde{Z}_t dw_t, & \widetilde{Y}_T=\partial_x g(\widetilde{X}_T,\varpi_T^\theta). \end{cases} \tag{2.32}$$

Note that, since the random environment ϖ^θ is adapted to b, there is no càdlàg orthogonal martingale term M in the dynamics of $\widetilde{Y}$. The optimal control is defined by the following function:

$$\widehat{\alpha} := -\overline{\Lambda}(y+\varpi). \tag{2.33}$$

As a second step, we consider the optimal control problem for the major player, when the random environment is the input process ϖ^θ. In analogy to Section 2.2.2.2, we define the optimal control problem on $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}}, \overline{\mathbb{F}})$, using the normalized coefficient f^0 introduced for the running cost in (2.21). In this case, the controls are adapted to the filtration $\mathbb{F}^{\varpi^\theta, b}$. Therefore, the optimal dynamics of the major player is defined by:

$$\begin{cases} d\widetilde{X}_t^0 = -\overline{\Lambda}^0(\widetilde{Y}_t^0 + \varpi_t^\theta)dt + \sigma_0^M(t, \varpi_t^\theta)db_t, & \widetilde{X}_0^0 = x_0, \\ d\widetilde{Y}_t^0 = -c_0^M(t, \varpi_t^\theta)dt + \widetilde{Z}_t^{0,0}db_t, & \widetilde{Y}_T^0 = g_1^0(\varpi_T^\theta). \end{cases} \tag{2.34}$$

The optimal control of the major player is given by:

$$\widehat{\beta} := -\overline{\Lambda}^0(y^0 + \varpi). \tag{2.35}$$

We introduce the discretized output process $\mathbf{\Phi}(\overline{\boldsymbol{\theta}}) := (\varphi^0(\overline{\boldsymbol{\theta}}), \dots, \varphi^{N-1}(\overline{\boldsymbol{\theta}}))$, defined as follows:

$$(\varphi^i(\overline{\boldsymbol{\theta}}))_t := (-(\overline{\Lambda}^0 + \overline{\Lambda})^{-1}\overline{\mathbb{E}}[\overline{\Lambda}\widetilde{Y}_t + \overline{\Lambda}^0\widetilde{Y}_t^0|V_1 = v_1, \dots, V_i = v_i])_{(v_1,\dots,v_i)\in\mathbb{J}^i}, \quad t \in [t_i, t_{i+1}]. \tag{2.36}$$

In particular, $\varphi^i(\overline{\boldsymbol{\theta}}) \in \mathcal{C}([t_i, t_{i+1}]; \mathbb{R}^n)$, for each $i = 0, \dots, N-1$. Analogously to the definition of $\overline{\boldsymbol{\theta}}$ and $(\theta_t)_{t\in[0,T]}$, we can introduce $\mathbf{\Phi}(\overline{\boldsymbol{\theta}})$ as an element of $\mathcal{D}([0,T]; \mathbb{R}^n)^{\mathbb{J}^{N-1}}$:

$$\begin{aligned} \mathbf{\Phi}(\overline{\boldsymbol{\theta}}) : \quad \mathbb{J}^{N-1} &\to \mathcal{D}([0,T]; \mathbb{R}^n) \\ (v_1, \dots, v_{N-1}) &\mapsto (\Phi_t(v_1, \dots, v_{N-1}))_{t\in[0,T]} \end{aligned}$$

where

$$\Phi_t(v_1, \dots, v_{N-1}) := \begin{cases} \varphi^i(\overline{\boldsymbol{\theta}})_t, & t \in [t_i, t_{i+1}), \ i = 0, \dots, N-1 \\ \varphi^{N-1}(\overline{\boldsymbol{\theta}})_t, & t = T, \end{cases}$$

$\mathbf{\Phi}$ is well-defined because the random variable $(V_1, \dots, V_i)$ takes values in $\mathbb{J}^i$.

As done for the input process, we can define the input-output map also as a functional:

$$\begin{aligned} \mathbf{\Phi} : \prod_{i=0}^{N-1} \mathcal{C}([t_i, t_{i+1}]; \mathbb{R}^n)^{\mathbb{J}^i} &\to \prod_{i=0}^{N-1} \mathcal{C}([t_i, t_{i+1}]; \mathbb{R}^n)^{\mathbb{J}^i} \\ \overline{\boldsymbol{\theta}} &\mapsto \mathbf{\Phi}(\overline{\boldsymbol{\theta}}) := (\varphi^0(\overline{\boldsymbol{\theta}}), \dots, \varphi^{N-1}(\overline{\boldsymbol{\theta}})). \end{aligned} \tag{2.37}$$

Let us introduce a more compact notation: we denote $(\overline{\theta}_t)_{t\in[0,T]}$ by $\overline{\theta}$ and the vector $(v_1, \dots, v_i) \in \mathbb{J}^i$, by $\overline{v}_i$, for every $i = 1, \dots, N$. Making use of this notation, the input-output map can be rewritten as a function defined on $\mathcal{D}([0,T]; \mathbb{R}^n)^{\mathbb{J}^{N-1}}$:

$$\begin{aligned} \mathbf{\Phi} : \quad \mathcal{D}([0,T]; \mathbb{R}^n)^{\mathbb{J}^{N-1}} &\to \mathcal{D}([0,T]; \mathbb{R}^n)^{\mathbb{J}^{N-1}} \\ (\overline{\theta}(\overline{v}_{N-1}))_{\overline{v}_{N-1}\in\mathbb{J}^{N-1}} &\mapsto \mathbf{\Phi}(\overline{\theta}) := (\Phi_t(\overline{v}_{N-1})_{t\in[0,T]})_{(\overline{v}_{N-1})\in\mathbb{J}^{N-1}} \end{aligned} \tag{2.38}$$

We aim at proving the existence of a fixed point of the function $\mathbf{\Phi}$, in one of the two equivalent formulations (2.37), (2.38). In order to apply standard fixed point results, we introduce a metric in the two spaces:

$$d(\overline{\boldsymbol{\theta}}^1,\overline{\boldsymbol{\theta}}^2) := \max_{i=0,\dots,N-1}\Big\{\max_{\overline{v}_i\in\mathbb{J}^i}\sup_{t\in[t_i,t_i+1]}\{|\theta^{1,i}(v_i)_t-\theta^{2,i}(v_i)_t|\}\Big\},\ \overline{\boldsymbol{\theta}}^1,\overline{\boldsymbol{\theta}}^2\in\prod_{i=0}^{N-1}\mathcal{C}([t_i,t_{i+1}];\mathbb{R})^{\mathbb{J}^i} \tag{2.39}$$

$$\widetilde{d}(\overline{\theta}^1,\overline{\theta}^2) := \max_{\overline{v}_{N-1}\in\mathbb{J}^{N-1}} d_{MZ}(\overline{\theta}^1_t(\overline{v}_{N-1}),\overline{\theta}^2_t(\overline{v}_{N-1})),\qquad \overline{\theta}^1,\overline{\theta}^2\in\mathcal{D}([0,T];\mathbb{R})^{\mathbb{J}^{N-1}}, \tag{2.40}$$

where d_{MZ} is the distance defined in [Kur91, Section 4], which characterizes the Meyer-Zheng space introduced in [MZ84]. The Meyer-Zheng space is defined as the space of equivalence classes of Lebesgue measurable functions $f:[0,T]\to\mathbb{R}$, where two functions are equivalent if they are equal for almost every $t\in[0,T]$. The Meyer-Zheng space is endowed with the topology of convergence in measure (for the properties of the Meyer-Zheng topology, we refer to [CD18b, Section 3.2.2]).

In both cases the discretized space is a Polish space. Thus, we can apply Schauder's fixed point theorem.

2.3.2 Solution to the discretized game

In this section, we show that the hypotheses of Schauder's fixed point theorem are satisfied. We first recall the statement of Schauder's fixed point theorem (see [Rud91, Theorem 5.28]):

Theorem 2.22. *(Schauder's fixed point) If K is a nonempty compact convex set in a locally convex space X, and $f:K\to K$ is continuous, then there exists $p\in K$, such that $f(p)=p$,*

In order to apply Theorem 2.22 , we have to prove that:

- $\boldsymbol{\Phi}$ is continuous;
- there exists a compact and convex subset $K\subseteq\prod_{i=0}^{N-1}\mathcal{C}([t_i,t_{i+1}];\mathbb{R})^{\mathbb{J}^i}$ such that $\boldsymbol{\Phi}:K\to K$.

As a first step, we prove the continuity in the whole space $\prod_{i=0}^{N-1}\mathcal{C}([t_i,t_{i+1}];\mathbb{R})^{\mathbb{J}^i}$. First, we state the following assumption, analogous to

Assumption 2.23 (Iteration in a Random Environment)**.** There exists a constant $\Gamma_0\geq 0$ such that, if $t\in[0,T]$ and $(\Omega,\mathcal{F},(\mathcal{F}_s)_{s\in[t,T]},\mathbb{P})$ is an admissible probabilistic setup equipped with some process $(b_s,\varpi_s,w_s)_{s\in[t,T]}$ with $x,x'\in\mathbb{R}$, in the sense of Definition 2.9, then for every two solutions $(X_s,Y_s,Z_s,Z^0_s)_{s\in[t,T]}$ and $(X'_s,Y'_s,Z'_s,Z^{0'}_s)_{s\in[t,T]}$ of FBSDE (2.32) with x and x' as respective initial condition at time t, it holds that:

$$\overline{\mathbb{P}}(|Y_t-Y'_t|\leq\Gamma_0|x-x'|)=1. \tag{2.41}$$

We are going to verify that Assumption 2.23 is verified for a couple of adjoint processes Y^θ and $Y^{\theta'}$, respectively associated with two input processes $\overline{\boldsymbol{\theta}}$ and $\overline{\boldsymbol{\theta}}'$. This condition is crucial to verify the continuity of $\boldsymbol{\Phi}$.

Proposition 2.24. *For any couple of elements* $\overline{\boldsymbol{\theta}}, \overline{\boldsymbol{\theta}}' \in \prod_{i=0}^{N-1} \mathcal{C}([t_i, t_{i+1}]; \mathbb{R})^{\mathbb{J}^i}$ *such that* $d(\overline{\boldsymbol{\theta}}, \overline{\boldsymbol{\theta}}') \to 0$, *also*

$$d(\boldsymbol{\Phi}(\overline{\boldsymbol{\theta}}), \boldsymbol{\Phi}(\overline{\boldsymbol{\theta}}')) \to 0, \tag{2.42}$$

where d is defined in equation (2.39).

Proof. Since Φ is defined by (2.38), to prove that the functional is continuous it is sufficient to show that:

$$\sup_{t\in[t_i,t_i+1]} |\varphi_t^i(\overline{\boldsymbol{\theta}})(\overline{v}_i) - \varphi_t^i(\overline{\boldsymbol{\theta}}')(\overline{v}_i)| \overset{\overline{\boldsymbol{\theta}} \xrightarrow{d} \overline{\boldsymbol{\theta}}'}{\longrightarrow} 0, \quad \forall \overline{v}_i \in \mathbb{J}^i, \quad \forall i = 0, \dots, N-1. \tag{2.43}$$

We consider θ and θ' in $\prod_{i=0}^{N-1} \mathcal{C}([t_i, t_{i+1}]; \mathbb{R})^{\mathbb{J}^i}$. We define two input processes ϖ^θ and $\varpi^{\theta'}$ determined by θ and θ' respectively. We denote by Y^θ and $Y^{\theta'}$ the solutions to the backward components in system (2.32) and we adopt the same notation for $Y^{0,\theta}$ and $Y^{0,\theta'}$ for the backward components of (2.34), where the input processes are ϖ^θ and $\varpi^{\theta'}$ respectively. Therefore, equation (2.43) is equivalent to

$$\begin{aligned}
\sup_{t\in[t_i,t_i+1]} |\varphi_t^i(\overline{\boldsymbol{\theta}})(\overline{v}_i) - \varphi_t^i(\overline{\boldsymbol{\theta}}')(\overline{v}_i)| &= \sup_{t\in[t_i,t_i+1]} \Big| - (\overline{\Lambda} + \overline{\Lambda}^0)^{-1} \overline{\mathbb{E}}[\overline{\Lambda} Y_t^\theta + \overline{\Lambda}^0 Y_t^{0,\theta} | V_1 = v_1, \dots, V_i = v_i] \\
&\quad + (\overline{\Lambda} + \overline{\Lambda}^0)^{-1} \overline{\mathbb{E}}[\overline{\Lambda} Y_t^{\theta'} + \overline{\Lambda}^0 Y_t^{0,\theta'} | V_1 = v_1, \dots, V_i = v_i]\Big| \\
&= (\overline{\Lambda} + \overline{\Lambda}^0)^{-1} \Bigg\{ \sup_{t\in[t_i,t_i+1]} \Big| - (\overline{\Lambda} \overline{\mathbb{E}}[Y_t^\theta - Y_t^{\theta'} | V_1 = v_1, \dots, V_i = v_i] \\
&\quad - \overline{\Lambda}^0 \overline{\mathbb{E}}[Y_t^{0,\theta} - Y_t^{0,\theta'} | V_1 = v_1, \dots, V_i = v_i])\Big| \Bigg\} \xrightarrow{d(\overline{\boldsymbol{\theta}}, \overline{\boldsymbol{\theta}}') \to 0} 0
\end{aligned}$$

We consider now the two terms separately. The difference $Y_t^{0,\theta} - Y_t^{0,\theta'}$ converges to zero, by the continuity of the functions c_0^M and g_1^0, that is ensured by Assumption B4. Indeed:

$$\begin{aligned}
(A) &:= \sup_{t\in[t_i,t_{i+1}]} |\overline{\mathbb{E}}[Y_t^{0,\theta} - Y_t^{0,\theta'} | V_1 = v_1, \dots, V_i = v_i]| \\
&\le \sup_{t\in[t_i,t_{i+1}]} \overline{\mathbb{E}}\Big[\Big| g_1^0(\varpi_T^\theta) - g_1^0(\varpi_T^{\theta'}) + \int_t^T (c_0^M(t, \varpi_s^\theta) - c_0^M(t, \varpi_s^{\theta'})) ds \Big| \Big| V_1 = v_1, \dots, V_i = v_i \Big], \\
&\le \mathbb{E}\Big[|g_1^0(\varpi_T^\theta) - g_1^0(\varpi_T^{\theta'})| | V_1 = v_1, \dots, V_i = v_i \Big] \\
&\quad + \sup_{t\in[t_i,t_{i+1}]} \overline{\mathbb{E}}\Big[\int_t^T |c_0^M(t, \varpi_s^\theta) - c_0^M(t, \varpi_s^{\theta'})| ds | V_1 = v_1, \dots, V_i = v_i \Big],
\end{aligned}$$

that converges to zero as $d(\overline{\boldsymbol{\theta}}, \overline{\boldsymbol{\theta}}') \to 0$ by the continuity of g_1^0 and c_0^M (we passed to the limit thanks to the boundedness of c_0^M, which is a sufficient condition to apply the dominated convergence theorem). We consider now the difference $Y_t^\theta - Y_t^{\theta'}$:

$$\begin{aligned}(B) &:= \sup_{t\in[t_i,t_{i+1}]} |\overline{\mathbb{E}}[Y_t^\theta | V_1 = v_1, \dots, V_i = v_i] - \overline{\mathbb{E}}[Y_t^{\theta'} | V_1 = v_1, \dots, V_i = v_i]| \\ &= \sup_{t\in[t_i,t_{i+1}]} |\overline{\mathbb{E}}[Y_t^\theta - Y_t^{\theta'} | V_1 = v_1, \dots, V_i = v_i]|.\end{aligned}$$

The solution to the backward equation determined by θ is:

$$Y_t^\theta = Y_T^\theta + \int_t^T \partial_x \overline{f}(s, X_s^\theta, \varpi_s^\theta) ds - \int_t^T Z_s^{0,\theta} db_s - \int_t^T Z_s^\theta dw_s, \quad t \in [0,T]. \tag{2.44}$$

Then:

$$\begin{aligned}(B) = \sup_{t\in[t_i,t_{i+1}]} &\Bigg|\overline{\mathbb{E}}\Bigg[(\partial_x g(X_T^\theta, \varpi_T^\theta) - \partial_x g(X_T^{\theta'}, \varpi_T^{\theta'})) + \int_t^T (\partial_x \overline{f}(s, X_s^\theta, \varpi_s^\theta) - \partial_x \overline{f}(s, X_s^{\theta'}, \varpi_s^{\theta'}))ds \\ &- \int_t^T (Z_s^{0,\theta} - Z_s^{0,\theta'})db_s - \int_t^T (Z_s^\theta - Z_s^{\theta'})dw_s \Big| V_1 = v_1, \dots, V_i = v_i\Bigg]\Bigg| \\ &= \sup_{t\in[t_i,t_{i+1}]} \Bigg|\overline{\mathbb{E}}\Bigg[(\overline{g}_1(\varpi_T^\theta) - \overline{g}_1(\varpi_T^{\theta'}) + \partial_x \overline{g}_2(X_T^\theta) - \partial_x \overline{g}_2(X_T^{\theta'})) \\ &\quad + \int_t^T (\partial_x \overline{f}(s, X_s^\theta, \varpi_s^\theta) - \partial_x \overline{f}(s, X_s^{\theta'}, \varpi_s^{\theta'}))ds \Big| V_1 = v_1, \dots, V_i = v_i\Bigg]\Bigg| \\ &\le \sup_{t\in[t_i,t_{i+1}]} \overline{\mathbb{E}}\Bigg[|\partial_x \overline{g}_2(X_T^\theta) - \partial_x \overline{g}_2(X_T^{\theta'})| + |\overline{g}_1(\varpi_T^\theta) - \overline{g}_1(\varpi_T^{\theta'})| \\ &\quad + \int_t^T \Big|\partial_x \overline{f}(s, X_s^\theta, \varpi_s^\theta) - \partial_x \overline{f}(s, X_s^{\theta'}, \varpi_s^{\theta'})\Big| ds \Big| V_1 = v_1, \dots, V_i = v_i\Bigg] \\ &\le \sup_{t\in[t_i,t_{i+1}]} \mathbb{E}\Big[|\partial_x \overline{g}_2(X_T^\theta) - \partial_x \overline{g}_2(X_T^{\theta'})| + I\overline{g}_1(\varpi_T^\theta) - \overline{g}_1(\varpi_T^{\theta'})| \\ &\quad + \int_t^T \Big|\partial_x \overline{f}(s, X_s^\theta, \varpi_s^\theta) - \partial_x \overline{f}(s, X_s^{\theta'}, \varpi_s^{\theta'})\Big| ds \Big| V_1 = v_1, \dots, V_i = v_i\Big] \\ &\le \sup_{t\in[t_i,t_{i+1}]} \overline{\mathbb{E}}\Bigg[L(|X_T^\theta - X_T^{\theta'}| + |\varpi_T^\theta - \varpi_T^{\theta'}|) \\ &\quad + \int_t^T \Big|\partial_x \overline{f}(s, X_s^\theta, \varpi_s^\theta) - \partial_x \overline{f}(s, X_s^{\theta'}, \varpi_s^{\theta'})\Big| ds \Big| V_1 = v_1, \dots, V_i = v_i\Bigg].\end{aligned}$$

The martingale terms vanish because the random variable on which we are conditioning is obtained as a function of the trajectory of the canonical process b in $[0, t_i]$. Therefore, the integral terms starting from t_i are both zero. Moreover, the term $|\varpi_T^\theta - \varpi_T^{\theta'}|$ converges to zero in the limit

$d(\overline{\boldsymbol{\theta}}\,\overline{\boldsymbol{\theta}}')\to 0$. Thus, we can consider:

$$\begin{aligned}
(C) := &\sup_{t\in[t_i,t_{i+1}]}\overline{\mathbb{E}}\left[\int_t^T\left|\partial_x\overline{f}(s,X_s^{\theta},\varpi_s^{\theta})-\partial_x\overline{f}(s,X_s^{\theta'},\varpi_s^{\theta'})\right|ds\middle|V_1=v_1,\dots,V_i=v_i\right]\\
&\leq \sup_{t\in[t_i,t_{i+1}]}\overline{\mathbb{E}}\left[\int_t^T\left|\partial_x\overline{f}(s,X_s^{\theta},\varpi_s^{\theta})-\partial_x\overline{f}(s,X_s^{\theta'},\varpi_s^{\theta})\right|ds\right.\\
&\qquad\left.+\int_t^T\left|\partial_x\overline{f}(s,X_s^{\theta'},\varpi_s^{\theta})-\partial_x\overline{f}(s,X_s^{\theta'},\varpi_s^{\theta'})\right|ds\middle|V_1=v_1,\dots,V_i=v_i\right]\\
&\leq \overline{\mathbb{E}}\left[\int_{t_i}^T\left|\partial_x\overline{f}(s,X_s^{\theta},\varpi_s^{\theta})-\partial_x\overline{f}(s,X_s^{\theta'},\varpi_s^{\theta})\right|ds\right.\\
&\qquad\left.+\int_{t_i}^T\left|\partial_x\overline{f}(s,X_s^{\theta'},\varpi_s^{\theta})-\partial_x\overline{f}(s,X_s^{\theta'},\varpi_s^{\theta'})\right|ds\middle|V_1=v_1,\dots,V_i=v_i\right]\\
&\leq \overline{\mathbb{E}}\left[LT\sup_{t\in[t_i,T]}|X_t^{\theta}-X_t^{\theta'}|\right.\\
&\qquad\left.+\int_{t_i}^T\left|\partial_x\overline{f}(s,X_s^{\theta'},\varpi_s^{\theta})-\partial_x\overline{f}(s,X_s^{\theta'},\varpi_s^{\theta'})\right|ds\middle|V_1=v_1,\dots,V_i=v_i\right].
\end{aligned}$$

By Assumption B1, the second term of the equation converges to zero (the function $\partial_x\overline{f}$ is continuous in the ϖ variable), when $\theta\to\theta'$. It is sufficient to prove that

$$\lim_{\theta\to\theta'}\overline{\mathbb{E}}\left[\sup_{t\in[t_i,t_{i+1}]}|X_t^{\theta}-X_t^{\theta'}|\middle|V_1=v_1,\dots,V_i=v_i\right]=0,\quad\forall\overline{v}_i\in\mathbb{J}^i,\quad i=0,\dots,N-1. \tag{2.45}$$

To prove this result, we apply [CD18b, Theorem 1.53]. The assumptions of this result are completely recovered by Assumption B2, except for Assumption 2.23. In Appendix 2.A, we verify that condition

(2.41) is satisfied by Y^{θ} and $Y^{\theta'}$. Hence, we can prove that the following stability condition holds:

$$\begin{aligned}
\overline{\mathbb{E}}\Bigg[\sup_{t\in[t_i,T]}|X_t^{\theta}-X_t^{\theta'}|^2|\overline{\mathcal{F}}_{t_i}\Bigg] &\leq \Gamma\overline{\mathbb{E}}\Bigg[|X_{t_i}^{\theta}-X_{t_i}^{\theta'}|^2+|\partial_x g(X_T^{\theta},\varpi_T^{\theta})-\partial_x g(X_T^{\theta},\varpi_T^{\theta'})|^2+\\
&\quad+\int_{t_i}^{T}\Big[|\partial_x\overline{f}(t,X_t^{\theta},\varpi_t^{\theta})-\partial_x\overline{f}(t,X_t^{\theta},\varpi_t^{\theta'})|^2+\\
&\quad+|-\overline{\Lambda}(Y_t^{\theta}-Y_t^{\theta})+l(t,\varpi_t^{\theta})-l(t,\varpi_t^{\theta'})|^2+\\
&\quad+|\sigma^0(t,\varpi_t^{\theta})-\sigma^0(t,\varpi_t^{\theta'})|^2+|\sigma(t,\varpi_t^{\theta})-\sigma(t,\varpi_t^{\theta'})|^2\Big]dt\Big|\overline{\mathcal{F}}_{t_i}\Bigg]\\
&=\Gamma\overline{\mathbb{E}}\Bigg[|X_{t_i}^{\theta}-X_{t_i}^{\theta'}|^2+L^2|\varpi_T^{\theta}-\varpi_T^{\theta'}|^2+\int_{t_i}^{T}\Big[|\partial_x\overline{f}(t,X_t^{\theta},\varpi_t^{\theta})+\\
&\quad-\partial_x\overline{f}(t,X_t^{\theta},\varpi_t^{\theta'})|^2+l(t,\varpi_t^{\theta})-l(t,\varpi_t^{\theta'})|^2+|\sigma^0(t,\varpi_t^{\theta})-\sigma^0(t,\varpi_t^{\theta'})|^2+\\
&\quad+|\sigma(t,\varpi_t^{\theta})-\sigma(t,\varpi_t^{\theta'})|^2\Big]dt|\overline{\mathcal{F}}_{t_i}\Bigg],
\end{aligned}\tag{2.46}$$

where the constant Γ depends only on Γ_0, L with L in Assumption 2.4. We notice that, under continuity in the ϖ variable (Assumption B1), the terms in equation (2.46) containing the functions $\partial_x\overline{f}, l, \sigma^0$ and σ converge to zero as $d(\boldsymbol{\overline{\theta}},\boldsymbol{\overline{\theta}}')\to 0$. Therefore, the only term we have to deal with is $|X_{t_i}^{\theta}-X_{t_i}^{\theta'}|^2$. We notice that, by the tower property:

$$\overline{\mathbb{E}}\Big[\sup_{t\in[t_i,T]}|X_t^{\theta}-X_t^{\theta'}|^2\Big|V_1=v_1,\dots,V_i=v_i\Big]=\overline{\mathbb{E}}\Big[\overline{\mathbb{E}}\Big[\sup_{t\in[t_i,T]}IX_t^{\theta}-X_t^{\theta'}|^2\Big|\mathcal{F}_{t_i}\Big]\Big|V_1=v_1,\dots,V_i=v_i\Big],$$

because $\{V_1=v_1,\dots,V_i=v_i\}\in\mathcal{F}_{t_i}$. Hence, we can control $\overline{\mathbb{E}}[\sup_{t\in[t_i,T]}|X_t^{\theta}-X_t^{\theta'}|^2|\mathcal{F}_{t_i}]$ under condition (2.46) as follows

$$\begin{aligned}
&\overline{\mathbb{E}}\Bigg[\overline{\mathbb{E}}[\sup_{t\in[t_i,T]}|X_t^{\theta}-X_t^{\theta'}|^2|\mathcal{F}_{t_i}]\Bigg|V_1=v_1,\dots,V_i=v_i\Bigg]\\
&\qquad\leq\Gamma(t_i)\overline{\mathbb{E}}\Bigg[\overline{\mathbb{E}}\Bigg[|X_{t_i}^{\theta}-X_{t_i}^{\theta'}|^2+\int_{t_i}^{T}\Big[|\partial_x\overline{f}(t,X_t^{\theta},\varpi_t^{\theta})-\partial_x\overline{f}(t,X_t^{\theta},\varpi_t^{\theta'})|^2+l(t,\varpi_t^{\theta})-l(t,\varpi_t^{\theta'})|^2\\
&\qquad+|\sigma^0(t,\varpi_t^{\theta})-\sigma^0(t,\varpi_t^{\theta'})|^2+|\sigma(t,\varpi_t^{\theta})-\sigma(t,\varpi_t^{\theta'})|^2\Big]dt\Big|\mathcal{F}_{t_i}\Bigg]\Bigg|V_1=v_1,\dots,V_i=v_i\Bigg]\\
&\quad=\Gamma(t_i)\overline{\mathbb{E}}\Big[|X_{t_i}^{\theta}-X_{t_i}^{\theta'}|^2\Big|V_1=v_1,\dots,V_i=v_i\Big]+(D)
\end{aligned}$$

where

$$(D) := \Gamma(t_i)\overline{\mathbb{E}}\Bigg[\int_{t_i}^{T}\Big[|\partial_x\overline{f}(t,X_t^\theta,\varpi_t^\theta)-\partial_x\overline{f}(t,X_t^\theta,\varpi_t^{\theta'})|^2+l(t,\varpi_t^\theta)-l(t,\varpi_t^{\theta'})|^2+ \\ +|\sigma^0(t,\varpi_t^\theta)-\sigma^0(t,\varpi_t^{\theta'})|^2+|\sigma(t,\varpi_t^\theta)-\sigma(t,\varpi_t^{\theta'})|^2\Big]dt\Bigg|V_1=v_1,\dots,V_i=v_i\Bigg]\xrightarrow{d(\overline{\boldsymbol{\theta}}\,\overline{\boldsymbol{\theta}}')\to 0} 0$$

Let us notice that in $\overline{\mathbb{E}}[|X_{t_i}^\theta - X_{t_i}^{\theta'}|^2|V_1=v_1,\dots,V_i=v_i]$ the price processes ϖ^θ and $\varpi^{\theta'}$ are computed in the dynamics of X^θ and $X^{\theta'}$ respectively until t_i. Thus, they are constant in the event $\{V_1=v_1,\dots,V_i=v_i\}$. However, if we consider condition (2.46) for $t_i=t_0=0$, we can notice that:

$$\overline{\mathbb{E}}\Bigg[\sup_{t\in[0,T]}|X_t^\theta-X_t^{\theta'}|^2|\mathcal{F}_0\Bigg]\le\Gamma_0\overline{\mathbb{E}}\Bigg[\overbrace{|X_0^\theta-X_0^{\theta'}|^2}^{=0}+\int_0^T\Big[|\partial_x\overline{f}(t,X_t^\theta,\varpi_t^\theta)-\partial_x\overline{f}(t,X_t^\theta,\varpi_t^{\theta'})|^2+ \\ +l(t,\varpi_t^\theta)-l(t,\varpi_t^{\theta'})|^2+|\sigma^0(t,\varpi_t^\theta)-\sigma^0(t,\varpi_t^{\theta'})|^2+ \\ +|\sigma(t,\varpi_t^\theta)-\sigma(t,\varpi_t^{\theta'})|^2\Big]dt\Bigg|\mathcal{F}_0\Bigg]\overset{\theta\to\theta'}{\longrightarrow}0.$$

By the tower property, $\overline{\mathbb{E}}\Big[\sup_{t\in[0,T]}|X_t^\theta-X_t^{\theta'}|^2|\Big]\overset{\theta\to\theta'}{\longrightarrow}0$ too. We recall also that $\overline{V}_i=(V_1,\dots,V_i)$ is a discrete random variable, whose support is given by the finite set $\mathbb{J}^i$. Hence, by definition of conditional expectation with respect to an even, the following holds

$$\overline{\mathbb{E}}[|X_{t_i}^\theta-X_{t_i}^{\theta'}|^2]=\sum_{\overline{v}_{\le i}\in\mathbb{J}^i}\overline{\mathbb{E}}[|X_{t_i}^\theta-X_{t_i}^{\theta'}|^2|V_1=v_1,\dots,V_i=v_i]\overline{\mathbb{P}}(V_1=v_1,\dots,V_i=v_i),$$

and $\overline{\mathbb{P}}(V_1=v_1,\dots,V_i=v_i)>0$ for each $\overline{v}_i\in\mathbb{J}^i$. As a consequence, since

$$\overline{\mathbb{E}}[|X_{t_i}^\theta-X_{t_i}^{\theta'}|^2|V_1=v_1,\dots,V_i=v_i]\ge 0,\quad a.s.,$$

we conclude that:

$$\lim_{\theta\to\theta'}\overline{\mathbb{E}}[|X_{t_i}^\theta-X_{t_i}^{\theta'}|^2|V_1=v_1,\dots,V_i=v_i]\le\lim_{\theta\to\theta'}\frac{1}{\overline{\mathbb{P}}(V_1=v_1,\dots,V_i=v_i)}\overline{\mathbb{E}}[|X_{t_i}^\theta-X_{t_i}^{\theta'}|^2] \\ \le\frac{1}{\overline{\mathbb{P}}(V_1=v_1,\dots,V_i=v_i)}\lim_{\theta\to\theta'}\overline{\mathbb{E}}\Bigg[\sup_{t\in[0,T]}|X_t^\theta-X_t^{\theta'}|^2\Bigg]=0.$$

□

2.3.2.1 Compact closure of Im(**Φ**)

Our goal now is to show that the image of the functional **Φ**, introduced in equation (2.37) is contained in a compact set of $\prod_{i=0}^{N-1}\mathcal{C}([t_i,t_{i+1}];\mathbb{R})^{\mathbb{J}^i}$. To this effect, we are going to apply Ascoli's theorem [Rud91, Theorem A.5] to the set of functions:

$$C^i_{\overline{v}_i} := \left\{ \varphi^i_t(\overline{\boldsymbol{\theta}})(\overline{v}_i) : [t_i, t_{i+1}] \longrightarrow \mathbb{R}, \quad \overline{\boldsymbol{\theta}} \in \prod_{j=0}^{N-1} \mathcal{C}([t_j, t_{j+1}]; \mathbb{R})^{\mathbb{J}^j} \right\}, \quad \forall \overline{v}_i \in \mathbb{J}^i, \tag{2.47}$$

defined for every $i = 0, \dots, N-1$, where φ is introduced in equation (2.36). Indeed, if $C^i_{\overline{v}_i}$ has compact closure for each i, also the finite product $\prod_{i=1}^N \prod_{\overline{v}_i} C^i_{\overline{v}_i}$ has compact closure.

To carry out this program, we must prove the following conditions:

1. pointwise-boundedness: there exists a constant C, independent of the input process $\overline{\boldsymbol{\theta}}$ such that:

$$\sup \left\{ |\varphi^i_t(\overline{\boldsymbol{\theta}})(\overline{v}_i)| : \ \overline{\boldsymbol{\theta}} \in \prod_{j=0}^{N-1} \mathcal{C}([t_j, t_{j+1}]; \mathbb{R})^{\mathbb{J}^j} \right\} \leq C, \quad \forall t \in [t_i, t_{i+1}], \quad \forall \overline{v}_i \in \mathbb{J}^i.$$

2. equi-continuity: there exists a constant L, independent of the input process $\overline{\boldsymbol{\theta}}$ such that:

$$|\varphi^i_t(\overline{\boldsymbol{\theta}})(\overline{v}_i) - \varphi^i_s(\overline{\boldsymbol{\theta}})(\overline{v}_i)| \leq L|t-s|, \quad \forall t, s \in [t_i, t_{i+1}]. \tag{2.48}$$

Concerning the pointwise boundedness, let us first notice that, by Assumption B2:

$$\begin{aligned} |Y^\theta_t|^2 &= \left| \overline{\mathbb{E}} \left[\partial_x g(X^\theta_T, \varpi^\theta_T) + \int_t^T \partial_x \overline{f}(s, X^\theta_s, \varpi^\theta_s) ds \middle| \overline{\mathcal{F}}_t \right] \right|^2 \\ &\leq 2\overline{\mathbb{E}} \left[|\partial_x g(X^\theta_T, \varpi^\theta_T)|^2 + (T-t) \int_t^T |\partial_x \overline{f}(s, X^\theta_s, \varpi^\theta_s)|^2 ds \middle| \overline{\mathcal{F}}_t \right] \\ &\leq 2L^2((T-t)^2 + 1) \leq 2L^2(T^2+1) =: C, \quad \text{a.s.} \end{aligned} \tag{2.49}$$

Applying Assumption B4, we conclude that $|Y^{0,\theta}_t|^2 \leq C$ a.s. Let us highlight that the constant C does not depend neither on t nor on $\overline{\boldsymbol{\theta}}$. Therefore, if we consider an arbitrary $\overline{\boldsymbol{\theta}} \in \prod_{j=0}^{N-1} \mathcal{C}([t_j, t_{j+1}]; \mathbb{R})^{\mathbb{J}^j}$:

$$\begin{aligned} |\varphi^i_t(\overline{\boldsymbol{\theta}})(\overline{v}_i)| &= (|\overline{\mathbb{E}}[-(\overline{\Lambda} + \overline{\Lambda}^0)^{-1}(\overline{\Lambda} Y^\theta_t + \overline{\Lambda}^0 Y^{0,\theta}_t)|V_1 = v_1, \dots, V_i = v_i]|^2)^{\frac{1}{2}} \\ &\leq \sqrt{2}(\overline{\Lambda} + \overline{\Lambda}^0)^{-1} \Big(\overline{\Lambda} \left(\overline{\mathbb{E}}[|Y^\theta_t|^2 | V_1 = v_1, \dots, V_i = v_i] \right)^{\frac{1}{2}} + \overline{\Lambda}^0 \left(\overline{\mathbb{E}}[|Y^{0,\theta}_t|^2 | V_1 = v_1, \dots, V_i = v_i] \right)^{\frac{1}{2}} \Big) \\ &\leq 2\sqrt{2}C, \end{aligned}$$

that yields the pointwise boundedness. The second property to prove is equi-continuity. To prove it, it is sufficient to restrict our attention to the case in which $s \leq t$ in $[t_i, t_{i+1}]$, hence the following holds

$$
\begin{aligned}
|\varphi_t^i(\overline{\boldsymbol{\theta}})(\overline{v}_i) - \varphi_s^i(\overline{\boldsymbol{\theta}})(\overline{v}_i)| &= \Big| - (\overline{\Lambda} + \overline{\Lambda}^0)^{-1}\overline{\mathbb{E}}\Big[\overline{\Lambda} Y_t^\theta + \overline{\Lambda}^0 Y_t^{0,\theta} | V_1 = v_1, \dots, V_i = v_i\Big] \\
&\quad + (\overline{\Lambda} + \overline{\Lambda}^0)^{-1}\overline{\mathbb{E}}\Big[\overline{\Lambda} Y_s^\theta + \overline{\Lambda}^0 Y_s^{0,\theta} \Big| V_1 = v_1, \dots, V_i = v_i\Big]\Big| \\
&\leq (\overline{\Lambda} + \overline{\Lambda}^0)^{-1} | (\overline{\Lambda} |\overline{\mathbb{E}}[Y_t^\theta - Y_s^\theta | V_1 = v_1, \dots, V_i = v_i]| \\
&\quad + \overline{\Lambda}^0 |\overline{\mathbb{E}}[Y_t^{0,\theta} - Y_s^{0,\theta} | V_1 = v_1, \dots, V_i = v_i]) | \\
&= \left| \overline{\mathbb{E}}\left[- \int_s^t \partial_x \overline{f}(u, X_u^\theta, \varpi_u^\theta) du + \int_s^t Z_u^{0,\theta} db_u + \int_s^t Z_u^\theta dw_u | V_1 = v_1, \dots, V_i = v_i \right] \right| \\
&\quad + \left| \overline{\mathbb{E}}\left[- \int_s^t c_0^M(t, \varpi_u^\theta) ds + \int_s^t Z_u^{0,0,\theta} db_u | V_1 = v_1, \dots, V_i = v_i \right] \right|.
\end{aligned}
$$

Since both s and t are greater than t_i, the stochastic integrals are independent of $\{V_1 = v_1, \dots, V_i = v_i\}$, thus, recalling that those terms are true martingales (see Remark 2.11), their conditional expectation is zero. As a consequence:

$$
|\varphi_t^i(\overline{\boldsymbol{\theta}})(\overline{v}_i) - \varphi_s^i(\overline{\boldsymbol{\theta}})(\overline{v}_i)| \leq \overline{\mathbb{E}}\Big[\int_s^t (|\partial_x \overline{f}(u, X_u^\theta, \varpi_u^\theta)| + |c_0^M(t, \varpi_u^\theta)|) du | V_1 = v_1, \dots, V_i = v_i\Big]
$$

By Assumptions B2 and B4, we conclude that

$$
|\varphi_t^i(\overline{\boldsymbol{\theta}})(\overline{v}_i) - \varphi_s^i(\overline{\boldsymbol{\theta}})(\overline{v}_i)| \leq 2L(t - s).
$$

The same holds if $t \leq s$. Therefore, equi-continuity is proved and, by Ascoli's theorem, the image of

$$
\varphi^i(\overline{v}_i) : \prod_{j=0}^{N-1} \mathcal{C}([t_j, t_{j+1}]; \mathbb{R})^{\mathbb{J}^j} \to \mathcal{C}([t_i, t_{i+1}]; \mathbb{R}),
$$

has compact closure for the sup norm, for each choice of the vector $\overline{v}_i \in \mathbb{J}^i$. Since $\mathbb{J}^i$ is finite, also the function:

$$
\varphi^i : \prod_{j=0}^{N-1} \mathcal{C}([t_j, t_{j+1}]; \mathbb{R})^{\mathbb{J}^j} \to \mathcal{C}([t_i, t_{i+1}]; \mathbb{R})^{\mathbb{J}^i}
$$

has compact closure of its image. We can conclude that the image of $\boldsymbol{\Phi} = (\varphi^0, \dots \varphi^{N-1})$ has compact closure. Therefore, we can restrict the continuous function $\boldsymbol{\Phi}$ to the compact closure of its image to apply Schauder's fixed point theorem.

2.3.3 Stability of the discretized equilibria

2.3.3.1 Outline of the strategy

In order to adopt the approach of [CD18b, Chapter 3] we take into account the solution to the discretized game contained in Subsection 2.3.1 for each step $\Delta = 2^{-l}$ for $l \in \mathbb{N}$ in the space grid and $N = 2^n$ for $n \in \mathbb{N}$ in the time grid. For each n and l we denote by $(X^{n,l}, Y^{n,l}, Z^{0,n,l}, Z^{n,l})$ the solution to equation (2.32) defined on the canonical space introduced at the beginning of Section 2.3.1. Analogously, we denote by $(X^{0,n,l}, Y^{0,n,l}, Z^{0,0,n,l})$ the solution to equation (2.34) defined on the canonical space. The price process $\varpi^{n,l}$ is defined as an input process of the form of (2.31) and is defined by the fixed point of the functional Φ introduced in equation (2.37). In particular, the process $\varpi^{n,l}$ is a càdlàg process defined as follows:

$$\varpi_t^{n,l} = -(\overline{\Lambda} + \overline{\Lambda}^0)^{-1}\mathbb{E}[\overline{\Lambda} Y_t^{n,l} + \overline{\Lambda}^0 Y_t^{0,n,l} | \overline{V}_i^{n,l}], \quad t \in [t_i, t_{i+1}), \quad i = 0, \dots, N-1, \tag{2.50}$$

where $\overline{V}_i^{n,l} = (V_1^{n,l}, \dots V_i^{n,l})$ is the discretization of the common noise until time $t_i = i\frac{T}{N}$.

The constants n and l determine respectively the number of factors of the product space on which Φ is defined and the number of components of each of these factors (i.e. the cardinality of $\mathbb{J}^i$). As shown in the previous section, for each n and l we consider two FBSDE systems linked by the price process $\varpi^{n,l}$. The one for the typical standard player is:

$$\begin{cases} dX_t^{n,l} &= (-\overline{\Lambda}(Y_t^{n,l} + \varpi_t^{n,l}) + l(t, \varpi_t^{n,l}))dt + \sigma_0(t, \varpi_t^{n,l})db_t + \sigma(t, \varpi_t^{n,l})dw_t, \\ X_0^{n,l} &= \xi, \\ dY_t^{n,l} &= -\partial_x \overline{f}(t, X_t^{n,l}, \varpi_t^{n,l})dt + Z_t^{0,n,l}db_t + Z_t^{n,l}dw_t, \\ Y_T^{n,l} &= \partial_x g(X_T^{n,l}, \varpi_T^{n,l}). \end{cases} \tag{2.51}$$

The one for the major player is:

$$\begin{cases} dX_t^{0,n,l} &= -\overline{\Lambda}^0(Y_t^{0,n,l} + \varpi_t^{n,l})dt + \sigma_0^M(t, \varpi_t^{n,l})db_t, \\ X_0^{0,n,l} &= x_0, \\ dY_t^{0,n,l} &= -c_0^M(t, \varpi_t^{n,l})dt + Z_t^{0,0,n,l}db_t, \\ Y_T^{0,n,l} &= g_1^0(\varpi_T^{n,l}). \end{cases} \tag{2.52}$$

We are going to present the steps **S-I**,. . . ,**S-VI**, that describe the procedure we follow in the next subsections to prove Theorem 2.18.

S-I In Subsection 2.3.3.2, we prove the tightness of the sequences $(X^{n,l})_{n,l}$ and $(X^{0,n,l})_{n,l}$ in $\mathcal{C}([0,T];\mathbb{R})$. Moreover, we show that the sequences $(\varpi^{n,l})_{n,l}$, $(Y^{n,l})_{n,l}$ and $(Y^{0,n,l})_{n,l}$ are tight in $\mathcal{M}([0,T];\mathbb{R})$, where $\mathcal{M}([0,T];\mathbb{R})$ is the Meyer-Zheng space. As a consequence, the sequence $(\xi, b, \varpi^{n,l}, w, X^{n,l}, X^{0,n,l})$, defined on $\overline{\Omega}$, is tight in the space $\Omega_{\text{input}} \times \mathcal{C}([0,T];\mathbb{R}) \times \mathcal{C}([0,T];\mathbb{R})$ where

$$\Omega_{\text{input}} := \mathbb{R} \times \mathcal{C}([0,T];\mathbb{R}) \times \mathcal{D}([0,T];\mathbb{R}) \times \mathcal{C}([0,T];\mathbb{R}).$$

As discussed in Section 2.3.1.2, the compatibility condition between the canonical filtration $\overline{\mathbb{F}}$ and the process $(\xi, b, \varpi^{n,l}, w)$ is guaranteed since $(\xi, b, \varpi^{n,l}, w)$ is adapted to $\overline{\mathbb{F}}$. (as stated in [CD18b, Remark 1.12]).

S-II Once we have proved tightness of the solutions of the two optimal control problems in the discretized setting, we need a stability result related to the optimality of any weak limit. To obtain this result, we consider a weak limit of $(\xi, b, \varpi^{n,l}, w, X^{n,l}, X^{0,n,l})$ defined on a suitable complete probability space $(\Omega^\infty, \mathcal{F}^\infty, \mathbb{P}^\infty)$ and denoted by $(\xi^\infty, b^\infty, \varpi^\infty, w^\infty, X^\infty, X^{0,\infty})$. As we will see in Section 2.3.3.3 we need to lift the environment $(\varpi^{n,l})_{n,l\in\mathbb{N}}$, adding the sequence of adjoint processes of the discretized game $(Y^{n,l}, Y^{0,n,l})_{n,l\in\mathbb{N}}$. In particular, we shall prove that the sequence $(Y^{n,l}, Y^{0,n,l})_{n,l\in\mathbb{N}}$ is tight in $\mathcal{M}([0,T], \mathbb{R}^2)$, so we are allowed to consider a weak limit, that will be denoted by $(Y^\infty, Y^{0,\infty})$. This step will be performed in Section 2.3.3.2. The lift of the random environment is determined by the process $\overline{\mathcal{W}}_t^\infty := (\varpi_t^\infty, Y_t^\infty, Y_t^{0,\infty})$, for every $t \in [0,T]$. Hence, we introduce the process:

$$\Theta^\infty := (\xi^\infty, b^\infty, \overline{\mathcal{W}}^\infty, w^\infty, X^\infty, X^{0,\infty}). \tag{2.53}$$

The filtration generated by this process is denoted by

$$\mathbb{F}^\infty := \overline{\mathbb{F}}^{\xi^\infty, b^\infty, \overline{\mathcal{W}}^\infty, w^\infty, X^\infty, X^{0,\infty}}. \tag{2.54}$$

On $(\Omega^\infty, \mathcal{F}^\infty, \mathbb{P}^\infty)$, we introduce also the following two sub-filtrations:

$$\mathbb{F}^{M,\infty} := \overline{\mathbb{F}}^{b^\infty, \mathcal{W}^{M,\infty}, X^{0,\infty}},, \tag{2.55}$$

$$\mathbb{F}^{S,\infty} := \overline{\mathbb{F}}^{\xi^\infty, b^\infty, \mathcal{W}^{S,\infty}, w^\infty, X^\infty},, \tag{2.56}$$

where $\mathcal{W}^{M,\infty} = (\varpi^\infty, Y^{0,\infty})$ and $\mathcal{W}^\infty = (\varpi^\infty, Y^\infty)$.

S-III In Subsection 2.3.3.3, we show that

- The stochastic process
$$\widetilde{\Theta}^{M,\infty} := (b^\infty, \mathcal{W}^{M,\infty}) \tag{2.57}$$
taking values on
$$\Omega^M_{\text{input}} := \mathcal{C}([0,T]; \mathbb{R}) \times \mathcal{D}([0,T]; \mathbb{R}^2)$$
is compatible with the filtration $\mathbb{F}^{M,\infty}$.
- The stochastic process
$$\widetilde{\Theta}^{S,\infty} := (\xi^\infty, b^\infty, \mathcal{W}^\infty, w^\infty) \tag{2.58}$$
taking values on
$$\Omega^S_{\text{input}} := \mathbb{R} \times \mathcal{C}([0,T]; \mathbb{R}) \times \mathcal{D}([0,T]; \mathbb{R}^2) \times \mathcal{C}([0,T]; \mathbb{R})$$
is compatible with the filtration $\mathbb{F}^{S,\infty}$.

S-IV Once compatibility is verified, we prove that the optimal control problem for the standard player:

$$\inf_{\alpha\in\mathbb{H}^2(\mathbb{F}^{S,\infty};\mathbb{R})} J^{\varpi^\infty}(\alpha), \quad J^{\varpi^\infty}(\alpha) = \mathbb{E}^\infty\left[\int_0^T f(s, X_s, \varpi_s^\infty, \alpha_s)ds + g(X_T, \varpi_T^\infty)\right],$$

$$\begin{cases} dX_t = (\alpha_t + l(t, \varpi_t^\infty))dt + \sigma_0(t, \varpi_t^\infty)db_t^\infty + \sigma(t, \varpi_t^\infty)dw_t^\infty, \\ X_0 = \xi^\infty, \end{cases}$$

is solved by the weak limit X^∞ of the sequence of discretized state variables $(X^{n,l})_{n,l\in\mathbb{N}}$. We prove this result in Subsection 2.3.3.4.

S-V Analogously for the major player, in Subsection 2.3.3.4, we prove that the optimal control problem for the major player:

$$\inf_{\beta\in\mathbb{H}^2(\mathbb{F}^{M,\infty};\mathbb{R})} J^{0,\varpi^\infty}(\beta), \quad J^{0,\varpi^\infty}(\beta) = \mathbb{E}^\infty\left[\int_0^T f^0(s, X_s, \varpi_s^\infty, \beta_s)ds + g^0(X_T, \varpi_T^\infty)\right],$$

$$\begin{cases} dX_t = \beta_t dt + \sigma^M(t, \varpi_t^\infty)db_t^\infty, \\ X_0 = x_0, \end{cases}$$

is solved by the weak limit $X^{0,\infty}$ of the sequence of discretized state variables $(X^{0,n,l})_{n,l\in\mathbb{N}}$.

We denote by $(X^{0,\infty}, Y^{0,\infty}, Z^{0,0,\infty}, M^{0,\infty})$ the solution to the FBSDE associated with the maximum principle applied to this optimal control problem. In order to guarantee that the consistency condition (2.20) is satisfied, we should prove that the backward equation of the FBSDE associated with the maximum principle applied to the optimal control problem of the major player is adapted to $\mathbb{F}^{b^\infty,\varpi^\infty}$. However, we cannot guarantee that this holds, because the controls are adapted to $\mathbb{F}^{M,\infty}$ that is larger than

$$\mathbb{F}^{M,p,\infty} := \overline{\mathbb{F}}^{b^\infty,\varpi^\infty}. \tag{2.59}$$

To circumvent this problem, in Subsection 2.3.3.5, we introduce the optimal control problem with the same coefficients of J^{0,ϖ^∞} and the same state variable, but for which the controls can be chosen in the space $\mathbb{H}^2(\mathbb{F}^{M,p,\infty};\mathbb{R})$. We exploit the affine structure in the x^0 variable of the cost functionals f^0 and g^0 to project the solution to the FBSDE system associated the stochastic maximum principle applied to the optimal control problem defined by controls in $\mathbb{H}^2(\mathbb{F}^{M,\infty};\mathbb{R})$ to the solution to the FBSDE system (denoted by $(\widetilde{X}^{0,\infty}, \widetilde{Y}^{0,\infty}, \widetilde{Z}^{0,0,\infty}, \widetilde{M}^{0,\infty})$) obtained solving the same optimal control problem using the controls chosen in $\mathbb{H}^2(\mathbb{F}^{M,p,\infty};\mathbb{R})$. In particular, we prove that $\widetilde{Y}_t^{0,\infty} = \mathbb{E}^\infty[Y_t^{0,\infty}|\mathcal{F}_t^{\varpi^\infty,b}]$.

S-VI Finally, by **S-IV** and **S-V**, we can introduce another FBSDE, associated with the stochastic maximum principle applied to the optimal control problem defined in **S-IV** for the typical standard agent. We denote its solution by $(X^{\infty}, Y^{\infty}, Z^{0,\infty}, Z^{\infty}, M^{\infty})$. On the other hand, for the major player, instead of considering $(X^{0,\infty}, Y^{0,\infty}, Z^{0,0,\infty}, M^{0,\infty})$, we can consider $(\widetilde{X}^{0,\infty}, \widetilde{Y}^{0,\infty}, \widetilde{Z}^{0,0,\infty}, \widetilde{M}^{0,\infty})$ obtained by restricting the class of controls adapted to $\mathbb{H}^2(\mathbb{F}^{M,p,\infty};\mathbb{R})$. In Subsection 2.3.3.6, we prove that

$$\varpi_t^{\infty} = -(\overline{\Lambda} + \overline{\Lambda}^0)^{-1}\Big(\mathbb{E}^{\infty}\Big[\overline{\Lambda} Y_t^{\infty} | \mathcal{F}_t^{\varpi^{\infty}, b^{\infty}}\Big] + \overline{\Lambda}^0 \widetilde{Y}_t^{0,\infty}\Big), \quad t \in [0,T] \tag{2.60}$$

is verified.

2.3.3.2 Tightness of $(\mathbf{X^{n,l}})_{\mathbf{n,l}\in\mathbb{N}}$, $(\mathbf{Y^{n,l}})_{\mathbf{n,l}\in\mathbb{N}}$, $(\mathbf{X^{0,n,l}})_{\mathbf{n,l}\in\mathbb{N}}$, $(\mathbf{Y^{0,n,l}})_{\mathbf{n,l}\in\mathbb{N}}$ and $(\varpi^{\mathbf{n,l}})_{\mathbf{n,l}\in\mathbb{N}}$

Tightness of $(\mathbf{X^{n,l}})_{\mathbf{n,l}\in\mathbb{N}}$ and $(\mathbf{X^{0,n,l}})_{\mathbf{n,l}\in\mathbb{N}}$ In order to prove the tightness of the sequences $(X^{n,l})_{n,l\in\mathbb{N}}$ and $(X^{0,n,l})_{n,l\in\mathbb{N}}$ in the canonical space we apply Aldous' criterion. We focus on the proof of the tightness of $(X^{n,l})_{n,l\in\mathbb{N}}$. The proof of the tightness of $(X^{0,n,l})_{n,l\in\mathbb{N}}$ is analogous, due to the analogy between Assumption B2 and Assumption B4. We first state the following Lemma, whose proof is given in Appendix 2.C:

Lemma 2.25. *The two following conditions hold:*

1. *There exists a constant $\widetilde{C} > 0$ independent of n and l such that:*

$$\sup_{n,l} \overline{\mathbb{E}}\Bigg[\sup_{t\in[0,T]} |X_t^{n,l}|^2\Bigg] \leq \widetilde{C}; \tag{2.61}$$

2. *for every $\mathbb{F}^{X^{n,l}}$-stopping time τ and positive constant $\delta > 0$, there exists a constant C, independent of n and l such that:*

$$\overline{\mathbb{E}}\Big[|X_{(\tau+\delta)\wedge T}^{n,l} - X_\tau^{n,l}|\Big] \leq C(\delta + \sqrt{\delta}). \tag{2.62}$$

In order to prove the tightness of the sequence $(X^{n,l})_{n,l}$ we apply Aldous criterion [Bil99, Theorem 16.10] that ensures tightness under the following two conditions:

1. $\lim_{a\to\infty} \limsup_{n,l} \overline{\mathbb{P}}(\sup_{t\in[0,m]} |X_t^{n,l}| \geq a) = 0, \quad \forall m \leq T$;

2. $\forall \epsilon > 0,\ \eta > 0, m \geq 0,\ \exists \delta_0 > 0, n_0 \in \mathbb{N},$ such that:

$$\overline{\mathbb{P}}(|X_{(\tau+\delta)\wedge T}^{n,l} - X_\tau^{n,l}| \geq \epsilon) \leq \eta, \quad \forall \delta \leq \delta_0,\ \forall n \geq n_0,\ \forall \tau\ \mathbb{F}^{X^{n,l}} - \text{stopping time}.$$

We first apply Markov inequality to (2.61), thus obtaining

$$\lim_{a\to\infty}\limsup_{n,l}\overline{\mathbb{P}}\left(\sup_{t\in[0,m]}|X_t^{n,l}|\geq a\right)\leq\lim_{a\to\infty}\limsup_{n,l}\frac{1}{a}\overline{\mathbb{E}}\left[\sup_{t\in[0,m]}|X_t^{n,l}|\right]$$
$$\leq\lim_{a\to\infty}\limsup_{n,l}\frac{1}{a}\left(\overline{\mathbb{E}}\left[\sup_{t\in[0,m]}|X_t^{n,l}|^2\right]\right)^{\frac{1}{2}}=0.$$

Therefore, condition 1. of Aldous' criterion is satisfied.

We consider now an $\mathbb{F}^{X^{n,l}}$-stopping time τ and $\delta\in(0,1)$. We observe that, by Markov inequality, $\forall\epsilon>0$ and $\eta>0$:

$$\overline{\mathbb{P}}\left(|X^{n,l}_{(\tau+\delta)\wedge T}-X^{n,l}_\tau|\geq\epsilon\right)\leq\frac{1}{\epsilon}\overline{\mathbb{E}}\left[||X^{n,l}_{(\tau+\delta)\wedge T}-X^{n,l}_\tau|\right]\leq\frac{1}{\epsilon}C\sqrt{\delta}.$$

Hence, if $\delta_0:=(\eta\epsilon C)^2$, condition 2. of Aldous' criterion is satisfied.

In conclusion, the sequence $(X^{n,l})_{n,l\in\mathbb{N}}$ is tight in $\mathcal{D}([0,T];\mathbb{R})$. Moreover, since $(X^{n,l})_{n,l\in\mathbb{N}}$ has continuous paths, the sequence is tight also in $\mathcal{C}[0,T];\mathbb{R})$.

Tightness of $(\mathbf{Y^{n,l}})_{\mathbf{n,l}}$, $(\mathbf{Y^{0,n,l}})_{\mathbf{n,l}}$ and $(\varpi^{\mathbf{n,l}})_{\mathbf{n,l}}$ We prove tightness of the sequence $(\varpi^{n,l})_{n,l\in\mathbb{N}}$ in the space of Meyer-Zheng. The key point is to notice that the process $\varpi^{n,l}$ is a càdlàg process defined on the probability space $(\overline{\Omega},\overline{\mathcal{F}},\mathbb{P})$ endowed with the filtration $(\mathcal{F}^{n,l}_t)_{t\in[0,T]}$ defined by the vector $\overline{V}_j$, introduced in (2.29),

$$\mathcal{F}^{n,l}_t=\sigma\{\overline{V}_j\},\quad\forall t\in[t_j,t_{j+1}),\quad\forall j=0,\dots,2^n.$$

It is therefore sufficient to show that the sequence $(\varpi^{n,l})_{n,l\in\mathbb{N}}$ satisfies the hypotheses of [CD18b, Theorem 3.9]. In particular, we need to verify the following condition:

$$\sup_{n,l}\left\{\overline{\mathbb{E}}\left[|\varpi^{n,l}_T|\right]+V^{n,l}_T(\varpi^{n,l})\right]\right\}<\infty,\tag{2.63}$$

where for each process $(A^{n,l}_t)_{t\in[0,T]}$ adapted to a filtration $(\mathcal{G}^{n,l}_t)_{t\in[0,T]}$, the conditional variance $V^{n,l}_t(A^{n,l})$ is defined by:

$$V^{n,l}_t(A^{n,l}):=\sup_{\Delta\subset[0,t]}\overline{\mathbb{E}}\left[\sum_{i=1}^N\left|\overline{\mathbb{E}}[A^{n,l}_{t_{i+1}}-A^{n,l}_{t_i}|\mathcal{G}^{n,l}_{t_i}]\right|\right],\tag{2.64}$$

where the supremum is taken over all partitions Δ of the time interval $[0,t]$.

The process $\varpi^{n,l}$ is adapted to $(\mathcal{F}^{n,l}_t)_{t\in[0,T]}$ and the conditional variance satisfies:

$$V^{n,l}_T(\varpi^{n,l})=\sup_{N\geq1}\sup_{0\leq\widetilde{t}_0\leq\cdots\leq\widetilde{t}_N}\overline{\mathbb{E}}\left[\sum_{j=0}^{N-1}|\overline{\mathbb{E}}[\varpi^{n,l}_{\widetilde{t}_{j+1}}-\varpi^{n,l}_{\widetilde{t}_j}|\mathcal{F}^{n,l}_{\widetilde{t}_j}]|\right].$$

We notice that the canonical filtration $\overline{\mathbb{F}} = (\overline{\mathcal{F}}_t)_{t\in[0,T]}$ contains $(\mathcal{F}^{n,l}_t)_{t\in[0,T]}$. This implies that:

$$\begin{aligned}
\left|\overline{\mathbb{E}}\Big[\varpi^{n,l}_{\widetilde{t}_{j+1}} - \varpi^{n,l}_{\widetilde{t}_j}|\mathcal{F}^{n,l}_{\widetilde{t}_j}\Big]\right| &\leq \Big|\overline{\mathbb{E}}\Big[-(\overline{\Lambda}+\overline{\Lambda}^0)^{-1}(\overline{\Lambda}\ \overline{\mathbb{E}}[Y^{n,l}_{\widetilde{t}_{j+1}}|\mathcal{F}^{n,l}_{\widetilde{t}_{j+1}}] + \overline{\Lambda}^0\overline{\mathbb{E}}[Y^{0,n,l}_{\widetilde{t}_{j+1}}|\mathcal{F}^{n,l}_{\widetilde{t}_{j+1}}]) \\
&\quad + (\overline{\Lambda}+\overline{\Lambda}^0)^{-1}(\overline{\Lambda}\ \overline{\mathbb{E}}[Y^{n,l}_{\widetilde{t}_j}|\mathcal{F}^{n,l}_{\widetilde{t}_j}] + \overline{\Lambda}^0\overline{\mathbb{E}}[Y^{0,n,l}_{\widetilde{t}_j}|\mathcal{F}^{n,l}_{\widetilde{t}_j}])\Big|\mathcal{F}^{n,l}_{\widetilde{t}_j}\Big]\Big| \\
&\leq (\overline{\Lambda}+\overline{\Lambda}^0)^{-1}\Big| - (\overline{\Lambda}\ \overline{\mathbb{E}}[Y^{n,l}_{\widetilde{t}_{j+1}}|\mathcal{F}^{n,l}_{\widetilde{t}_j}] + \overline{\Lambda}^0\overline{\mathbb{E}}[Y^{0,n,l}_{\widetilde{t}_{j+1}}|\mathcal{F}^{n,l}_{\widetilde{t}_j}]) \\
&\quad + \overline{\Lambda}\ \overline{\mathbb{E}}[Y^{n,l}_{\widetilde{t}_j}|\mathcal{F}^{n,l}_{\widetilde{t}_j}] + \overline{\Lambda}^0\overline{\mathbb{E}}[Y^{0,n,l}_{\widetilde{t}_j}|\mathcal{F}^{n,l}_{\widetilde{t}_j}]\Big| \\
&= (\overline{\Lambda}+\overline{\Lambda}^0)^{-1}\Big| - \overline{\Lambda}\ \overline{\mathbb{E}}[Y^{n,l}_{\widetilde{t}_{j+1}} - Y^{n,l}_{\widetilde{t}_j}|\mathcal{F}^{n,l}_{\widetilde{t}_j}] - \overline{\Lambda}^0\overline{\mathbb{E}}[Y^{0,n,l}_{\widetilde{t}_{j+1}} - Y^{0,n,l}_{\widetilde{t}_j}|\mathcal{F}^{n,l}_{\widetilde{t}_j}]\Big| \\
&= (\overline{\Lambda}+\overline{\Lambda}^0)^{-1}\Bigg| - \overline{\Lambda}\ \overline{\mathbb{E}}\left[\int_{\widetilde{t}_j}^{\widetilde{t}_{j+1}} \partial_x\overline{f}(s, X^{n,l}_s, \varpi^{n,l}_s)ds\Big|\mathcal{F}^{n,l}_{\widetilde{t}_j}\right] \\
&\quad - \overline{\Lambda}^0\overline{\mathbb{E}}\left[\int_{\widetilde{t}_j}^{\widetilde{t}_{j+1}} c^M_0(t, \varpi^{n,l}_s)ds\Big|\mathcal{F}^{n,l}_{\widetilde{t}_j}\right]\Bigg| \\
&\leq (\overline{\Lambda}+\overline{\Lambda}^0)^{-1}\overline{\mathbb{E}}\left[\int_{\widetilde{t}_j}^{\widetilde{t}_{j+1}} (\overline{\Lambda}|\partial_x\overline{f}(s, X^{n,l}_s, \varpi^{n,l}_s)| + \overline{\Lambda}^0|c^M_0(t, \varpi^{n,l}_s)|)ds\Big|\mathcal{F}^{n,l}_{\widetilde{t}_j}\right] \\
&\leq L(\widetilde{t}_{j+1} - \widetilde{t}_j).
\end{aligned}$$

Therefore, it holds that $V^{n,l}_T(\varpi^{n,l}) \leq LT$. Finally, we prove in the same way that the sequence $(Y^{n,l})_{n,l\in\mathbb{N}}$ and $(Y^{0,n,l})_{n,l\in\mathbb{N}}$ are tight in the space of Meyer-Zheng $\mathcal{M}([0,T];\mathbb{R})$. To do so, we need to check the hypotheses of [CD18b, Theorem 3.9]. Let us prove the following result for $(Y^{n,l})_{n,l\in\mathbb{N}}$, the analogous result for $(Y^{0,n,l})_{n,l\in\mathbb{N}}$ can be proved in the same way.

Lemma 2.26. *If $Y^{n,l}$ is the solution to the backward equation of the FBSDE system* (2.32)*, where the price process is given by* (2.50)*, then*

$$\sup_{n,l}[\overline{\mathbb{E}}[|Y^{n,l}_T|] + V^{n,l}_T(Y^{n,l})] < \infty.$$

where $V^{n,l}_T(Y^{n,l})$ is defined in equation (2.64)*.*

Proof. By assumption $\overline{\mathbb{E}}[|Y^{n,l}_T|] = \overline{\mathbb{E}}[|\partial_x g(X^{n,l}_T \varpi^{n,l}_T)|] \leq L$ for every n and l. We can consider:

$$V^{n,l}_T(Y^{n,l}) = \sup_{\widetilde{N}\geq 1}\ \sup_{\widetilde{t}_0\leq\cdots\leq\widetilde{t}_N = T}\ \overline{\mathbb{E}}\left[\sum_{j=0}^{\widetilde{N}-1} |\overline{\mathbb{E}}[Y^{n,l}_{\widetilde{t}_{j+1}} - Y^{n,l}_{\widetilde{t}_j}|\mathcal{F}_{\widetilde{t}_j}]|\right].$$

We compute

$$\begin{aligned}
\overline{\mathbb{E}}\left[\sum_{j=0}^{\widetilde{N}-1} |\overline{\mathbb{E}}[Y^{n,l}_{\widetilde{t}_{j+1}} - Y^{n,l}_{\widetilde{t}_j}|\mathcal{F}_{\widetilde{t}_j}]|\right] &= \overline{\mathbb{E}}\left[\sum_{j=0}^{\widetilde{N}-1}\left|\overline{\mathbb{E}}\left[\int_{\widetilde{t}_j}^{\widetilde{t}_{j+1}} \partial_x\overline{f}(s, X^{n,l}_s, \varpi^{n,l}_s)ds|\mathcal{F}_{\widetilde{t}_j}\right]\right|\right] \\
&\leq \overline{\mathbb{E}}\left[\sum_{j=0}^{\widetilde{N}-1}(\widetilde{t}_{j+1} - \widetilde{t}_j)L\right] = TL,
\end{aligned}$$

for all n and l. Since the estimate does not depend on n or l, we can take the supremum and obtain the result. □

2.3.3.3 Compatibility for the limit optimal control problem for the standard player

In order to prove stability of the equilibria when passing to the limit we need to guarantee the compatibility condition between the process $(\xi^\infty, b^\infty, \varpi^\infty, w^\infty)$, and the filtration generated by the weak limit $(\xi^\infty, b^\infty, \varpi^\infty, w^\infty, X^\infty)$. Since this property does not hold in general, we replace the $\varpi^{n,l}$ with

$$\mathcal{W}^{n,l} := (\varpi^{n,l}, Y^{n,l}), \quad n, l \in \mathbb{N}.$$

In analogy to [CD18b, Chapter 3], we call $\mathcal{W}^{n,l}$ *lifted environment*. As we are going to show in this section, this procedure allows us to guarantee the compatibility condition. By the results proved in Section 2.3.3.2, the sequence $\mathcal{W}^{n,2n}$ is tight, as a consequence, we can consider a weak limit in $\mathcal{M}([0,T];\mathbb{R}^2)$. The weak limit $\mathcal{W}^\infty = (\varpi^\infty, Y^\infty)$ possesses a càdlàg version ([Kur91, Theorem 5.8]). For the moment, we cannot conclude that Y^∞ is the adjoint process of the solution $(X^\infty, \varpi^\infty)$ of the optimal control problem of the typical standard agent defined on $(\Omega^\infty, \mathcal{F}^\infty, \mathbb{P}^\infty)$, because the compatibility between the process driving the SDE of the state variable and the filtration, which the controls are adapted to, is not guaranteed. Therefore, we need to prove the following result:

Lemma 2.27. *The process* $\widetilde{\Theta}^{S,\infty}$ *introduced in* (2.58) *is compatible with the filtration* $\mathbb{F}^{S,\infty}$ *defined in* (2.56).

Proof. Following the proof of [CD18b, Proposition 3.12], we rewrite the filtration generated by Θ^∞ as the filtration generated by a process that does not explicitly depends on X^∞. To do so, we exploit the lifted environment $\mathcal{W}^\infty$.

Step 1 First of all, mimicking the reasoning of *Step 1* of the proof of Proposition 2.28 below, we can prove that the sequence of optimal controls of the optimal control problem for the standard player in the discretized setting forms a tight sequence, thus we are allowed to extract a weak limit, that we denote by $\widehat{\alpha}^\infty$.

We introduce the process $\Psi^\infty := (\xi^\infty, b^\infty, \mathcal{W}^\infty, w^\infty, \widehat{\alpha}^\infty)$ taking values on $\Omega_{\text{input}} \times \mathcal{M}([0,T];\mathbb{R})$. The process Ψ^∞ generates a filtration denoted by $\mathbb{G}^\infty$. Moreover, following the same strategy described in *Step 2* of the proof of Proposition 2.28 below, we can show that $\widehat{\alpha}^\infty$ is adapted to $\mathbb{F}^{S,\infty}$, where $\mathbb{F}^{S,\infty}$ is the filtration introduced in (2.56). As a consequence, we have that $\mathbb{G}^\infty \subseteq \mathbb{F}^{S,\infty}$. Moreover, applying the same reasoning we can prove that X^∞ is adapted to $\mathbb{G}^\infty$. Therefore, we conclude that $\mathbb{G}^\infty = \mathbb{F}^{S,\infty}$. Hence, it is sufficient to show compatibility between $\mathbb{G}^\infty$ and the process $\widetilde{\Theta}^{S,\infty}$ introduced in (2.58). In order to guarantee compatibility between $\widetilde{\Theta}^{S,\infty}$ and $\mathbb{G}^\infty$ it is sufficient to show that $\widehat{\alpha}^\infty$ is compatible with $\widetilde{\Theta}^{S,\infty}$, in the sense of [BL20, Theorem 1.2. (i)], that is

$$\mathcal{F}_t^{\widehat{\alpha}^\infty} \text{ is conditionally independent of } \mathcal{F}_T^{\widetilde{\Theta}^{S,\infty}} \text{ given } \mathcal{F}_t^{\widetilde{\Theta}^{S,\infty}}.$$

This is equivalent to say that, for any arbitrary sets $\widetilde{A}_t \in \mathcal{F}_t^{\widetilde{\Theta}^{S,\infty}}$, $\widetilde{A}_T \in \mathcal{F}_T^{\widetilde{\Theta}^{S\infty}}$, $\widetilde{C}_t \in \mathcal{G}_t^{\infty}$,

$$\overline{\mathbb{P}}(\widetilde{C}_t \cap \widetilde{A}_t | \mathcal{F}_t^{\widetilde{\Theta}^{S,\infty}}) = \overline{\mathbb{P}}(\widetilde{C}_t | \mathcal{F}_t^{\widetilde{\Theta}^{S,\infty}}) \overline{\mathbb{P}}(\widetilde{A}_T | \mathcal{F}_t^{\widetilde{\Theta}^{S,\infty}}).$$

Step 2 We notice that the sets $\widetilde{C}_t, \widetilde{A}_t$ and $\widetilde{A}_T$ can be rewritten as

$$\begin{aligned} \widetilde{C}_t &= \{\Psi^{\infty}_{\cdot \wedge t} \in C_t\}, \quad C_t \in \mathcal{B}(\Omega^S_{\text{input}} \times \mathcal{M}([0,T];\mathbb{R})), \\ \widetilde{A}_t &= \{\widetilde{\Theta}^{S,\infty}_{\cdot \wedge t} \in A_t\}, \quad A_t \in \mathcal{B}(\Omega^S_{\text{input}}) \\ \widetilde{A}_T &= \{\widetilde{\Theta}^{S,\infty} \in A_T\}, \quad A_T \in \mathcal{B}(\Omega^S_{\text{input}}). \end{aligned}$$

Approximating each Borel set for the product sigma-algebra with rectangles in $\Omega^S_{\text{input}} \times \mathcal{D}([0,T];\mathbb{R})$, we obtain that $\mathbb{1}_{\widetilde{C}_t}(\omega) = \mathbb{1}_{C_t^{\text{input}}}(\widetilde{\Theta}^{S,\infty}_{\cdot \wedge t}(\omega)) \cdot \mathbb{1}_{C_t^{\alpha}}(\widehat{\alpha}^{\infty} \cdot \wedge t(\omega))$, where C_t^{input} is the projection on Ω_{input} of C_t and C_t^{α} is the projection on $\mathcal{D}([0,T],\mathbb{R})$ of C_t. Thus, the compatibility condition is equivalent to

$$\begin{aligned} 0 &= \overline{\mathbb{E}}[\mathbb{1}_{\widetilde{C}_t}\mathbb{1}_{\widetilde{A}_t}\mathbb{1}_{\widetilde{A}_T}]\overline{\mathbb{E}}[\mathbb{1}_{\widetilde{A}_t}] - \overline{\mathbb{E}}[\mathbb{1}_{\widetilde{C}_t}\mathbb{1}_{\widetilde{A}_t}]\overline{\mathbb{E}}[\mathbb{1}_{\widetilde{A}_T}\mathbb{1}_{\widetilde{A}_t}] \\ &= \mathbb{E}^{\infty}[\mathbb{1}_{C_t^{\text{input}}}(\widetilde{\Theta}^{S,\infty}_{\cdot\wedge t})\mathbb{1}_{C_t^{\alpha}}(\widehat{\alpha}^{\infty}_{\cdot\wedge t})\mathbb{1}_{A_T}(\widetilde{\Theta}^{S,\infty})\mathbb{1}_{A_t}(\widetilde{\Theta}^{S,\infty}_{\cdot\wedge t})]+ \\ &\quad - \mathbb{E}^{\infty}[\mathbb{1}_{C_t^{\text{input}}}(\widetilde{\Theta}^{S,\infty}_{\cdot\wedge t})\mathbb{1}_{C_t^{\alpha}}(\widehat{\alpha}^{\infty}_{\cdot\wedge t})\mathbb{1}_{A_t}(\widetilde{\Theta}^{S,\infty}_{\cdot\wedge t})]\mathbb{E}^{\infty}[\mathbb{1}_{A_T}(\widetilde{\Theta}^{S,\infty})\mathbb{1}_{A_t}(\widetilde{\Theta}^{S,\infty}_{\cdot\wedge t})] \\ &= \mathbb{E}^{\infty}[\mathbb{1}_{C_t^{\text{input}}\cap A_t}(\widetilde{\Theta}^{S,\infty}_{\cdot\wedge t})\mathbb{1}_{C_t^{\alpha}}(\widehat{\alpha}^{\infty}_{\cdot\wedge t})\mathbb{1}_{A_T}(\widetilde{\Theta}^{S,\infty})]+ \\ &\quad - \mathbb{E}^{\infty}[\mathbb{1}_{C_t^{\text{input}}\cap A_t}(\widetilde{\Theta}^{S,\infty}_{\cdot\wedge t})\mathbb{1}_{C_t^{\alpha}}(\widehat{\alpha}^{\infty}_{\cdot\wedge t})]\mathbb{E}^{\infty}[\mathbb{1}_{A_T}(\widetilde{\Theta}^{S,\infty})\mathbb{1}_{A_t}(\widetilde{\Theta}^{S,\infty}_{\cdot\wedge t})]. \end{aligned}$$

By the last line of this equation, the compatibility condition between $\widehat{\alpha}^{\infty}$ and $\widetilde{\Theta}^{S,\infty}$ is equivalent to require that the sigma-algebra generated by the process $\widehat{\alpha}^{\infty}$ until time $t \in [0,T]$ is conditionally independent of the sigma-algebra generated by the process $\widetilde{\Theta}^{S,\infty}$ given the sigma-algebra generated by $\widetilde{\Theta}^{S,\infty}$ until time t. This condition is guaranteed when:

$$\mathbb{E}^{\infty}\Big[f(\widehat{\alpha}^{\infty}_{\cdot\wedge t})h(\widetilde{\Theta}^{S,\infty}_{\cdot\wedge t})\Big(g(\widetilde{\Theta}^{S,\infty}) - \mathbb{E}^{\infty}[g(\widetilde{\Theta}^{S,\infty})|\mathcal{F}_t^{\widetilde{\Theta}^{S,\infty}}]\Big)\Big] = 0, \tag{2.65}$$

for f, g, h bounded and measurable. We recall that, the sequence $(\widehat{\alpha}^n, \varpi^n, Y^n)$ is converging in $\mathcal{M}([0,T];\mathbb{R}^3)$ to $(\widehat{\alpha}^{\infty}, \varpi^{\infty}, Y^{\infty})$. For every $n \in \mathbb{N}$, condition (2.33) is satisfied. This implies that:

$$\alpha_t^n + \overline{\Lambda}(Y_t^n + \varpi_t^n) = 0, \quad n \in \mathbb{N}.$$

In particular, since $p(a, y, \varpi) := a + \overline{\Lambda}(y + \varpi)$ is continuous, the triplet

$$(\widehat{\alpha}^n, \varpi^n, Y^n, (p(\widehat{\alpha}_t^n, \varpi_t^n, Y_t^n))_{t\in[0,T]})$$

converges in distribution to $(\widehat{\alpha}^{\infty}, \varpi^{\infty}, Y^{\infty}, (p(\widehat{\alpha}_t^{\infty}, \varpi_t^{\infty}, Y_t^{\infty}))_{t\in[0,T]})$, on $\mathcal{M}([0,T];\mathbb{R}^4)$ by [CD18b, Lemma 3.5] and continuous mapping theorem.

Step 3 Moreover, since $p(\widehat{\alpha}^n_t, \varpi^n_t, Y^n_t) = 0$ for each $n \in \mathbb{N}$, also $p(\widehat{\alpha}^\infty_t, \varpi^\infty_t, Y^\infty_t) = 0$. This implies that $\widehat{\alpha}^\infty_t = -\overline{\Lambda}(Y^\infty_t + \varpi^\infty_t)$ for almost every $t \in [0,T]$. Since there exists a unique càdlàg function in an equivalent class defining an element of the Meyer-Zheng space ([DM78, Section IV.44]), the equality holds for every $t \in [0,T]$. As a consequence, $\widehat{\alpha}^\infty$ is adapted to the filtration generated by $\widetilde{\Theta}^{S,\infty}$ and condition (2.65) is satisfied. We can conclude that $\mathbb{F}^{S,\infty} = \overline{\mathbb{F}}^{\widetilde{\Theta}^{S,\infty}}$.

□

Applying the same reasoning, we can prove that $\widetilde{\Theta}^{M,\infty}$, introduced in (2.57), is compatible with the filtration $\mathbb{F}^{M,\infty}$ introduced in (2.55).

2.3.3.4 Optimality of the weak limit

In this section, we prove stability of the discretized equilibria when the number of players goes to infinity. In Lemma 2.27, we proved that the process $\widetilde{\Theta}^{S,\infty} := (\xi^\infty, b^\infty, \mathcal{W}^\infty, w^\infty)$ is compatible with the filtration $\mathbb{F}^{S,\infty}$ introduced in (2.56), we can study the optimal control problem on $(\Omega^\infty, \mathcal{F}^\infty, \mathbb{P}^\infty)$ defined in **S-IV**. We show that the optimal solution to this optimal control problem is the weak limit X^∞ of the sequence $(X^{n,l})_{n,l\in\mathbb{N}}$. To do so, we must apply the following result, whose proof is contained in Appendix 2.D. The following proposition is inspired by [CD18b, Proposition 3.11].

Proposition 2.28. *On the canonical space $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}}, \overline{\mathbb{F}})$, introduced at the beginning of Section 2.3.1, we consider a sequence of càdlàg stochastic processes $(\varpi^n)_{n\in\mathbb{N}}$ adapted to the filtration generated by canonical process b. We introduce a sequence of stochastic control problems defined as follows:*

$$\inf_{\alpha\in\mathbb{H}^2(\overline{F}^n)} J^{\varpi^n}(\alpha),\ J^{\varpi^n}(\alpha) := \overline{\mathbb{E}}\left[\int_0^T f(s, X_s, \varpi^n_s, \alpha_s)ds + g(X_T, \varpi^n_T)\right], \tag{2.66}$$

subject to

$$dX^n_t = (\alpha_t + l(t, \varpi^n_t))dt + \sigma^0(t, \varpi^n_t)db_t + \sigma(t, \varpi^n_t)dw_t, \quad X^n_0 = X^0. \tag{2.67}$$

The controls are supposed to be chosen in the set of $\mathbb{F}^n := \overline{\mathbb{F}}^{b,w}$-progressively measurable processes. Under the assumptions of Assumption 2.4 and Assumption 2.6, there exists a minimizer to the functional (2.66)*. Thus, let us consider the process $X^n := (X^n_t)_{t\in[0,T]}$, solution to equation* (2.67)*, driven by the optimal control $\widehat{\alpha}^n$. Let us introduce moreover the FBSDE* (2.51) *associated with the stochastic maximum principle:*

$$\begin{cases} dX^n_t &= (-\overline{\Lambda}(Y^n_t + \varpi^n_t) + l(t, \varpi^n_t))dt + \sigma^0(t, \varpi^n_t)db_t + \sigma(t, \varpi^n_t)dw_t, \\ X^n_0 &= X^0, \\ dY^n_t &= -\partial_x \overline{f}(t, X^n_t, \varpi^n_t)dt + Z^{0,n}_t db_t + Z^n_t dw_t, \\ Y^n_T &= g(X^n_T, \varpi^n_T). \end{cases}$$

Assume that:

D1 *The sequences* $(X^n)_{n\in\mathbb{N}}$ *and* $(Y^n)_{n\in\mathbb{N}}$ *are tight respectively on* $\mathcal{C}([0,T];\mathbb{R})$ *and* $\mathcal{M}([0,T];\mathbb{R})$.

D2 *The sequences* $(X^n)_{n\in\mathbb{N}}$, $(Y^n)_{n\in\mathbb{N}}$ *and* $(\varpi^n)_{n\in\mathbb{N}}$ *are uniformly square-integrable. Moreover, the fourth-moment of* Y^n *is uniformly bounded in* $n\in\mathbb{N}$.

D3 *The sequence* $(\varpi^n)_{n\in\mathbb{N}}$ *is tight in* $\mathcal{M}([0,T],\mathbb{R})$.

Then, the sequence $(\overline{\mathbb{P}}\circ(X_0,b,\mathcal{W}^n,w,X^n)^{-1})_n$ *is tight on the space* $\widetilde{\Omega}_{input}\times\mathcal{C}([0,T];\mathbb{R})$, *where*

$$\widetilde{\Omega}_{input} := \mathbb{R}\times\mathcal{C}([0,T],\mathbb{R})\times\mathcal{M}([0,T],\mathbb{R}^2)\times\mathcal{C}([0,T],\mathbb{R}) \tag{2.68}$$

and $\mathcal{W}^n_t := (\varpi^n_t, Y^n_t)$, *for* $t\in[0,T]$.

Moreover, if $(X^{0,\infty},b^\infty,\mathcal{W}^\infty,w^\infty,X^\infty)$ *is a* $\widetilde{\Omega}_{input}\times\mathcal{C}([0,T],\mathbb{R})$*-valued process on a complete filtered space* $(\Omega^\infty,\mathbb{F}^\infty,\mathbb{P}^\infty)$ *such that the probability measure* $\mathbb{P}^\infty\circ(X^{0,\infty},b^\infty,\mathcal{W}^\infty,w^\infty,X^\infty)^{-1}$ *is a weak limit of the sequence* $(\overline{\mathbb{P}}\circ(X_0,b,\mathcal{W}^n,w,X^n)^{-1})_n$ *with* $\mathcal{W}^\infty=(\varpi^\infty,Y^\infty)$, *we can associate with the complete and right continuous filtration* $\mathbb{F}^\infty$ *generated by* $(X^{0,\infty},b^\infty,\mathcal{W}^\infty,w^\infty,X^\infty)$ *the stochastic control problem given by the functional:*

$$\inf_{\alpha\in\mathbb{H}^2(\mathbb{F}^\infty;\mathbb{R})} J^{\varpi^\infty}(\alpha), \quad J^{\varpi^\infty}(\alpha) := \mathbb{E}^\infty\left[\int_0^T f(s,X_s,\varpi^\infty_s,\alpha_s)ds + g(X_T,\varpi^\infty_T)\right], \tag{2.69}$$

with state variable defined by:

$$\begin{cases} dX_t = (\alpha_t + l(t,\varpi^\infty_t))dt + \sigma^0(t,\varpi^\infty_t)db^\infty_t + \sigma(t,\varpi^\infty_t)dw^\infty_t, \\ X_0 = X^{0,\infty}. \end{cases} \tag{2.70}$$

Moreover, if the filtration $\mathbb{F}^\infty$ *is compatible with the process* $(X^{0,\infty},b^\infty,\mathcal{W}^\infty,w^\infty)$, *then* X^∞ *is optimal for the stochastic control problem* (2.69) *and* (2.70), *when it is considered in the admissible probabilistic setup* $(\Omega^\infty,\mathcal{F}^\infty,\mathbb{P}^\infty,\mathbb{F}^\infty)$ *on which* $(X^{0,\infty},(b^\infty,w^\infty),\mathcal{W}^\infty)$ *is defined.*

To apply Proposition 2.28 we must check that the hypotheses are respected. In Subsection 2.3.3.2, we proved that D1 and D3 are satisfied. To check D2, we observe that if a sequence $(A^n)_{n\in\mathbb{N}}$ has uniformly bounded fourth moments, then the uniform square integrability is guaranteed. Indeed, applying Fubini's Theorem, together with Cauchy-Schwartz, Markov and Jensen's inequalities, we can observe that:

$$\begin{aligned}\lim_{a\to\infty}\sup_{n\in\mathbb{N}}\overline{\mathbb{E}}\left[\int_0^T |A^n_t|^2\mathbb{1}_{\{|A^n_t|\geq a\}}dt\right] &= \lim_{a\to\infty}\sup_{n\in\mathbb{N}}\int_0^T \overline{\mathbb{E}}\Big[|A^n_t|^2\mathbb{1}_{\{|A^n_t|\geq a\}}\Big]dt,\\ &\leq \lim_{a\to\infty}\sup_{n\in\mathbb{N}}\int_0^T\left(\overline{\mathbb{E}}\Big[|A^n_t|^4\Big]\overline{\mathbb{E}}\Big[\mathbb{1}_{\{|A^n_t|\geq a\}}\Big]\right)^{\frac{1}{2}}dt,\end{aligned}$$

$$
\begin{aligned}
&\leq \lim_{a\to\infty}\sup_{n\in\mathbb{N}} T\left(\overline{\mathbb{E}}\left[\sup_{t\in[0,T]}|A_t^n|^4\right]\overline{\mathbb{E}}\left[\mathbb{1}_{\{\sup_{t\in[0,T]}|A_t^n|\geq a\}}\right)^{\frac{1}{2}}\right] \\
&\leq \lim_{a\to\infty}\sup_{n\in\mathbb{N}} T\frac{1}{\sqrt{a}}\left(\overline{\mathbb{E}}\left[\sup_{t\in[0,T]}|A_t^n|^4\right]\overline{\mathbb{E}}\left[\sup_{t\in[0,T]}|A_t^n|\right]\right)^{\frac{1}{2}} \\
&\leq T\lim_{a\to\infty}\frac{1}{\sqrt{a}}\sup_{n\in\mathbb{N}}\left(\overline{\mathbb{E}}\left[\sup_{t\in[0,T]}|A_t^n|^4\right]\right)^{\frac{5}{8}} = 0.
\end{aligned}
$$

To check D2, it is sufficient to prove the following proposition

Proposition 2.29. *$(X^{n,l})_{n,l\in\mathbb{N}}$, $(Y^{n,l})_{n,l\in\mathbb{N}}$ and $(\varpi^{n,l})_{n,l\in\mathbb{N}}$ have uniformly bounded for moments.*

Proof. As done in (2.49), we focus on $Y^{n,l}$, $Y^{0,n,l}$ and $\varpi^{n,l}$ first. We notice that, by Jensen's inequality:

$$
\begin{aligned}
|Y_t^{n,l}|^{2p} &= \left|\overline{\mathbb{E}}\left[\partial_x\overline{g}(X_T^{n,l},\varpi_T^{n,l}) + \int_t^T \partial_x\overline{f}(s,X_s^{n,l},\varpi_s^{n,l})ds\Big|\overline{\mathcal{F}}_t\right]\right| \\
&\leq L^{2p} + \frac{1}{T-t}\overline{\mathbb{E}}\left[(T-t)^{2p}\int_t^T|\partial_x\overline{f}(s,X_s^{n,l},\varpi_s^{n,l})ds|^{2p}\Big|\overline{\mathcal{F}}_t\right] \\
&\leq L^{2p} + (T-t)^{2p-1}(T-t)L^{2p} = L^{2p}(T^{2p}+1) =: C_1^p
\end{aligned} \tag{2.71}
$$

Hence, $\forall p\in\mathbb{N}$, $|Y_t^{n,l}|^{2p}\leq L^{2p}(T^{2p}+1)$, $\overline{\mathbb{P}}$-a.s.. The same computations applied to $Y^{0,n,l}$ lead to $|Y_t^{0,n,l}|^{2p}\leq C_1^p$, $\overline{\mathbb{P}}$-a.s., for all $p\in\mathbb{N}$. As a consequence, also the fourth moments of $Y^{n,l}$ and $Y^{0,n,l}$ are uniformly bounded on n,l by C_1^p.

We notice that:

$$
\begin{aligned}
|\varpi_t^{n,l}|^{2p} &= |\overline{\mathbb{E}}[-(\overline{\Lambda}+\overline{\Lambda}^0)^{-1}(\overline{\Lambda}Y_t^{n,l}+\overline{\Lambda}^0Y_t^{0,n,l})|V_1=v_1,\dots,V_i=v_i]|^{2p} \\
&\leq (\overline{\Lambda}+\overline{\Lambda})^{-2p}((\overline{\Lambda})^{2p}\overline{\mathbb{E}}[|Y_t^{n,l}|^{2p}|\overline{V}_i] + (\overline{\Lambda}^0)^{2p}\overline{\mathbb{E}}[|Y_t^{0,n,l}|^{2p}|\overline{V}_i]) \\
&\leq (\overline{\Lambda}+\overline{\Lambda})^{-2p}((\overline{\Lambda})^{2p}+(\overline{\Lambda}^0)^{2p})L^{2p}(T^{2p}+1) =: C_2^p.
\end{aligned} \tag{2.72}
$$

Therefore, the fourth moment of $\varpi^{n,l}$ is uniformly bounded in n,l by C_2^p. Finally we consider the forward process $X^{n,l}$. Its dynamics is given by:

$$
X_t^{n,l} = \xi + \int_0^t(-\overline{\Lambda}(Y_s^{n,l}+\varpi_s^{n,l}) + l(s,\varpi_s^{n,l}))ds + \int_0^t\sigma^0(s,\varpi_s^{n,l})db_s + \int_0^t\sigma(s,\varpi_s^{n,l})dw_s.
$$

Therefore:

$$
|X_t^{n,l}|^{2p}\leq 4^{2p-1}\left\{|\xi|^{2p} + \left|\int_0^t -\Lambda(Y_s^{n,l}+\varpi_s^{n,l}) + l(s,\varpi_s^{n,l})ds\right|^{2p} + \left|\int_0^t\sigma^0(s,\varpi_s^{n,l})db_s\right|^{2p} + \left|\int_0^t\sigma(s,\varpi_s^{n,l})dw_s\right|^{2p}\right\}.
$$

We take the supremum over $t\in[0,T]$, and we apply Burkholder-David-Gundy and Jensen inequaliies, together with (2.71) and (2.72), to conclude that:

$$
\begin{aligned}
\overline{\mathbb{E}}\left[\sup_{t\in[0,T]}|X_t^{n,l}|^{2p}\right] &\leq 4^{2p-1}\Bigg\{\overline{\mathbb{E}}[|\xi|^{2p}]+T^{2p-1}\overline{\mathbb{E}}\left[\int_0^T|\overline{\Lambda}(Y_s^{n,l}+\varpi_s^{n,l})+l(s,\varpi_s^{n,l})|^{2p}ds\right]\\
&\quad+\overline{\mathbb{E}}\left[\sup_{t\in[0,T]}\left(\int_0^t\sigma^0(s,\varpi_s^{n,l})db_s\right|^{2p}\right]+\overline{\mathbb{E}}\left[\sup_{t\in[0,T]}\left|\int_0^t\sigma(s,\varpi_s^{n,l})dw_s\right|^{2p}\right]\Bigg\}\\
&\leq 4^{2p-1}\Bigg\{\overline{\mathbb{E}}[|\xi|^{2p}]+T^{2p-1}\overline{\mathbb{E}}\Bigg[\int_0^T3^{2p-1}\Bigg(\overline{\Lambda}|Y_s^{n,l}|^{2p}+\overline{\Lambda}|\varpi_s^{n,l}|^{2p}\\
&\quad+L^{2p}(1+|\varpi_s^{n,l}|)^{2p}\Bigg)ds\Bigg]+C_p\Bigg\{\overline{\mathbb{E}}\left[\left(\int_0^T|\sigma^0(s,\varpi_s^{n,l})|^2ds\right)^p\right]\\
&\quad+\overline{\mathbb{E}}\left[\left(\int_0^T|\sigma(s,\varpi_s^{n,l})|^2ds\right)^p\right]\Bigg\}\Bigg\}\\
&\leq 4^{2p-1}\Bigg\{\overline{\mathbb{E}}[|\xi|^{2p}]+T^{2p-1}\overline{\mathbb{E}}\Bigg[\int_0^T3^{2p-1}\Bigg(\overline{\Lambda}|Y_s^{n,l}|^{2p}+\overline{\Lambda}|\varpi_s^{n,l}|^{2p}\\
&\quad+L^{2p}(1+|\varpi_s^{n,l}|)^{2p}\Bigg)ds\Bigg]+2C_p\left(\overline{\mathbb{E}}\left[\int_0^TL^2(1+|\varpi_t^{n,l}|)^2dt\right]\right)^p\Bigg\}\\
&\leq 4^{2p-1}\left\{\overline{\mathbb{E}}[|\xi|^{2p}]+T^{2p}3^{2p-1}(\overline{\Lambda}(C_1^p+C_2^p)+L^{2p}2^{2p-1}(1+C_2^p))+2C_pL^2C_2^2T\right\}
\end{aligned}
$$

Therefore, considering the case $p=2$ and applying Assumption 2.2, we conclude that:

$$
\overline{\mathbb{E}}\left[\sup_{t\in[0,T]}|X_t^{n,l}|^4\right]\leq 4^3\left\{\overline{\mathbb{E}}[|\xi|^4]+T^43^3(\overline{\Lambda}(C_1^4+C_2^4)+L^42^3(1+C_2^4))+2C_2L^2C_2^2T\right\}=:C_3^2
$$

Letting $C_4:=\max\{C_1^2,C_2^2,C_3^2\}$, we can conclude that

$$
\overline{\mathbb{E}}\left[\sup_{t\in[0,T]}|\varpi_t^{n,l}|^4+\sup_{t\in[0,T]}|Y_t^{n,l}|^4+\sup_{t\in[0,T]}|X_t^{n,l}|^4\right]\leq C_4.
$$

□

In conclusion, we have shown that D2 is satisfied by $(X^{n,l})_{n,l\in\mathbb{N}}$, $(Y^{n,l})_{n,l\in\mathbb{N}}$ and $(\varpi^{n,l})_{n,l\in\mathbb{N}}$. Thus, we can apply Proposition 2.28 and concluding that X^∞ is optimal for the stochastic control problem (2.69) and (2.70), when it is considered in the admissible probabilistic setup $(\Omega^\infty,\mathcal{F}^\infty,\mathbb{P}^\infty,\mathbb{F}^{S,\infty})$ on which the process $(\xi^\infty,(b^\infty,w^\infty),\mathcal{W}^\infty)$ is defined.

Remark 2.30. By the same reasoning, it is possible to prove that conditions D1, D2 and D3 of Proposition 2.28 are satisfied by $(X^{0,n,l})_{n,l\in\mathbb{N}}$, $(Y^{0,n,l})_{n,l\in\mathbb{N}}$ and $(\varpi^{n,l})_{n,l\in\mathbb{N}}$. In particular, the uniform square integrability of $(X^{0,n,l})_{n,l\in\mathbb{N}}$ and $(Y^{0,n,l})_{n,l\in\mathbb{N}}$ is guaranteed by Assumption B4. Hence, the optimal control problem introduced in **S-V** is solved by the weak limit $X^{0,\infty}$ of the sequence $(X^{0,n,l})_{n,l\in\mathbb{N}}$.

2.3.3.5 Measurability of the solution to the optimal control problem S-V

As discussed at the beginning of this Section, we want to project the solution to the optimal control problem introduced in **S-V** to the space of stochastic processes adapted to the filtration $\mathbb{F}^{M,p,\infty}$, introduced in (2.59). This step is crucial to guarantee the consistency condition for the equilibrium price process introduced in (2.60). In particular, we aim at considering the projection to $(\Omega^\infty, \mathcal{F}^\infty, \mathbb{P}^\infty, \overline{F}^{\varpi^\infty, b^\infty})$ of the solution $(X^{0,\infty}, Y^{0,\infty}, Z^{0,0,\infty}, M^{0,\infty})$ to the FBSDE associated with the optimal control problem **S-V**. To do so, we consider the optimal control problem with the same coefficients to the one introduced in **S-V**, but where the class of admissible controls is given by $\mathbb{H}^2(\mathbb{F}^{M,p,\infty};\mathbb{R})$, where $\mathbb{F}^{M,p,\infty}$ is introduced in (2.59). We then introduce the following optimal control problem

S-IV(2)

$$\inf_{\beta\in\mathbb{H}^2(\mathbb{F}^{M,p,\infty};\mathbb{R})} J^{0,\varpi^\infty}(\beta), \quad J^{0,\varpi^\infty}(\beta) := \mathbb{E}^\infty\left[\int_0^T f^0(s, X_s, \varpi_s^\infty, \beta_s)ds + g^0(X_T, \varpi_T^\infty)\right],$$

where the state variable is defined as the solution to

$$\begin{cases} dX_t^0 = \beta_t dt + \sigma^M(t, \varpi_t^\infty)db_t^\infty, \\ X_0^0 = x_0, \end{cases}$$

In analogy to (2.35), the optimal control of the major player is defined by

$$\widehat{\beta}_t^\infty := -\overline{\Lambda}^0(\widetilde{Y}_t^{0,\infty} + \varpi_t^\infty),$$

where $\widetilde{Y}^{0,\infty}$ is the solution to the backward component of the FBSDE:

$$\begin{cases} d\widetilde{X}_t^{0,\infty} = -\overline{\Lambda}^0(\widetilde{Y}_t^{0,\infty} + \varpi_t^\infty)dt + \sigma^M(t, \varpi_t^\infty)db_t^\infty, \\ d\widetilde{Y}_t^{0,\infty} = -c_0^M(t, \varpi_t^\infty)dt + \widetilde{Z}_t^{0,0,\infty}db_t^\infty + d\widetilde{M}_t^{0,\infty}, \\ \widetilde{X}_0^{0,\infty} = x_0, \\ \widetilde{Y}_T^{0,\infty} = g_1^0(\varpi_T^\infty). \end{cases} \tag{2.73}$$

We recall that $\widetilde{Z}^{0,0,\infty}$ and $\widetilde{M}^{0,\infty}$ are the martingales obtained by the Kunita-Watanabe decomposition theorem, computing the conditional expectation of the random variable $g_1^0(\varpi_T^\infty)+\int_t^T c_0^M(s,\varpi_s^\infty)ds$ with respect to $\mathcal{F}_t^{b^\infty,\varpi^\infty}$. In other words:

$$\widetilde{Y}_t^{0,\infty} = \mathbb{E}^\infty\left[g_1^0(\varpi_T^\infty) + \int_t^T c_0^M(s, \varpi_s^\infty)ds \middle| \mathcal{F}_t^{M,p,\infty}\right].$$

As discussed in [CD18b, Remark 1.8], the compatibility condition does not depend on the completeness of filtrations. This allows us to conclude that:

$$\widetilde{Y}_t^{0,\infty} = \mathbb{E}^\infty \left[g_1^0(\varpi_T^\infty) + \int_t^T c_0^M(s, \varpi_s^\infty) ds \middle| \mathcal{F}_t^{b^\infty, \varpi^\infty} \right].$$

Due to the affine structure in the x variable of the cost functional of the major player, we notice that:

$$\widetilde{Y}_t^{0,\infty} = \mathbb{E}^\infty \left[g_1^0(\varpi_T^\infty) + \int_t^T c_0^M(s, \varpi_s^\infty) ds \middle| \mathcal{F}_t^{b^\infty, \varpi^\infty} \right] = \mathbb{E}^\infty[Y_t^{0,\infty} | \mathcal{F}_t^{b^\infty, \varpi^\infty}]. \tag{2.74}$$

We are going to see in the next section that this property is fundamental for the consistency condition (2.60).

2.3.3.6 Consistency condition for the limit game

The main consequence of Proposition 2.28 is that the optimal solution to the stochastic control problem described in **S-IV** is given by the weak limit X^∞ of the state variable, with the controls given by the weak limit $\widehat{\alpha}^\infty$ of the optimal controls for the discretized game, determined by the solution to the FBSDE (2.51). The limit optimal control problem is defined on a suitable probability space $(\Omega^\infty, \mathcal{F}^\infty, \mathbb{P}^\infty)$, endowed with the filtration $\mathbb{F}^{S,\infty} = \overline{\mathbb{F}}^{\xi^\infty, b^\infty, \mathcal{W}^\infty, w^\infty}$. As a consequence, we can introduce the FBSDE associated with the optimal state:

$$\begin{cases} dX_t^\infty &= (\widehat{\alpha}_t^\infty + l(t, \varpi_t^\infty))dt + \sigma_0(t, \varpi_t^\infty) db_t^\infty + \sigma(t, \varpi_t^\infty) dw_t^\infty, \\ X_0^\infty &= \xi^\infty, \\ dY_t^\infty &= -\partial_x \overline{f}(t, X_t^\infty, \varpi_t^\infty) dt + Z_t^{0,\infty} db_t^{1,\infty} + Z_t^\infty dw_t^\infty + dM_t^\infty, \\ Y_T^\infty &= \partial_x g(X_T^\infty, \varpi_T^\infty). \end{cases} \tag{2.75}$$

Since the optimal control is unique, it satisfies $\widehat{\alpha}_t^\infty = -\overline{\Lambda}(Y_t^\infty + \varpi_t^\infty)$. In (2.75), $\varpi^\infty = \pi_\varpi(\mathcal{W}^\infty)$, where π_ϖ denotes the projection on the first component. As a consequence $Y_t^\infty = -(\overline{\Lambda}^{-1} \widehat{\alpha}_t^\infty + \varpi_t^\infty)$, therefore, Y^∞ is a continuous function of $\widehat{\alpha}^\infty$ and ϖ^∞. We recall that $(\widehat{\alpha}^\infty, \varpi^\infty)$ are the weak limit of $(\widehat{\alpha}^n, \varpi^n)$, hence by the continuous mapping theorem Y_t^∞ is distributed as the weak limit of the sequence $-(\overline{\Lambda}^{-1} \widehat{\alpha}_t^n + \varpi_t^n) = Y_t^n$. In conclusion, the adjoint equation Y^∞ is distributed like the second component of the lifted process $\mathcal{W}^\infty$.

The same reasoning can be applied to the optimal control problem introduced in **S-V**. On $(\Omega^\infty, \mathcal{F}^\infty, \mathbb{P}^\infty)$, endowed with $\mathbb{F}^{M,\infty}$, the solution to the optimal control problem is characterized by the following FBSDE:

$$\begin{cases} dX_t^{0,\infty} = & -\overline{\Lambda}^0 (Y_t^{0,\infty} + \varpi_t^\infty) dt + \sigma_0^M(t, \varpi_t^{n,l}) db_t^\infty, \\ X_0^{0,\infty} = & x_0, \\ dY_t^{0,\infty} = & -c_0^M(t, \varpi_t^\infty) dt + Z_t^{0,0,\infty} db_t^\infty, \\ Y_T^{0,\infty} = & g_1^0(\varpi_T^\infty). \end{cases}$$

In the following paragraph, we shall explain how the stability of weak equilibria for the two optimal control problems leads to the solution to the mean-field equation for the price process introduced in (2.20).

Limit in the consistency condition We recall that

$$\varpi_t^{n,l} = -(\overline{\Lambda} + \overline{\Lambda}^0)^{-1}\overline{\mathbb{E}}[\overline{\Lambda} Y_t^{n,l} + \overline{\Lambda}^0 Y_t^{0,n,l} | \overline{V}_i^{n,l}], \quad t \in [t_i, t_{i+1}),$$

for each n, l, where $\overline{V}_i^{n,l} := (V_0^{n,l}, \dots, V_i^{n,l})$. We introduce the function:

$$\begin{aligned} &\mathcal{V}_t^{n,l} := \overline{V}_i^{n,l}, \quad t \in [t_i, t_{i+1}),\ \forall i = 0, 1 \dots, 2^n - 1, \\ &\mathcal{V}_T^{n,l} := \overline{V}_{2^n-1}^{n,l}. \end{aligned} \tag{2.76}$$

As a consequence, $\varpi_t^{n,l} = -(\overline{\Lambda} + \overline{\Lambda}^0)^{-1}\overline{\mathbb{E}}[\overline{\Lambda} Y_t^{n,l} + \overline{\Lambda}^0 Y_t^{0,n,l} | \mathcal{F}_t^{n,l}]$ for all $t \in [0, T]$, where $\mathcal{F}_t^{n,l} := \sigma\{\mathcal{V}_s^{n,l} :\ s \le t\}$. In particular, $\varpi_t^{n,l}$ is $\mathcal{F}_t^{n,l}$-measurable for each $t \in [0, T]$. Thus, we can define $\mathcal{G}_t^{n,l} := \sigma\{\varpi_s^{n,l}, \mathcal{V}_s^{n,l} :\ s \le t\} \equiv \mathcal{F}_t^{n,l}$, for each $t \in [0, T]$. We notice that

$$\overline{\mathbb{E}}[[\varpi_t^{n,l} + (\overline{\Lambda} + \overline{\Lambda}^0)^{-1}(\overline{\Lambda} Y_t^{n,l} + \overline{\Lambda}^0 Y_t^{0,n,l})] | \mathcal{G}_t^{n,l}] = 0, \forall t \in [0, T],\ n, l \in \mathbb{N}.$$

This is equivalent to:

$$\overline{\mathbb{E}}\Big[h(\varpi_{\cdot\wedge t}^{n,l}, \mathcal{V}_{\cdot\wedge t}^{n,l})[\varpi_t^{n,l} + (\overline{\Lambda} + \overline{\Lambda}^0)^{-1}(\overline{\Lambda} Y_t^{n,l} + \overline{\Lambda}^0 Y_t^{0,n,l})]\Big] = 0, \quad \forall t \in [0, T],\ n, l \in \mathbb{N}, \tag{2.77}$$

for every $h \in \mathcal{C}_b(\mathcal{D}([0, T]; \mathbb{R}) \times \mathcal{D}([0, T]; \mathbb{R}))$.

Lemma 2.31. *In the setting developed above, the process $(\mathcal{V}^{n,2n})_n$, defined in (2.76), converges in probability to b.*

Proof. Similarly to [CD18b, Lemma 3.17], we introduce the event

$$A^n := \left\{ \sup_{t \in [0,T]} |b_t| \le 2^{2n} - 1 \right\}.$$

On A^n, for every $i \in \{1, \dots, N-1\}$, it holds that $|V_i^{n,l} - b_{\frac{i}{2n}T}| \le \frac{i}{2^{2n}} \le \frac{1}{2^n}$. Therefore, let us define

$$B^n := \left\{ \sup_{t \in [0,T]} |\mathcal{V}_t^{n,2n} - b_t| \le \frac{1}{2^{n-1}} + \sup_{s,t \in [0,T]:\ |t-s| \le \frac{1}{2^n}} |b_s - b_t| \right\}$$

We observe that, on A^n:

$$\sup_{t \in [0,T]} |\mathcal{V}^{n,2n} - b_t| = \max_{i=0,\dots,2^n-1} \Big(|\mathcal{V}_{\frac{i}{2^n}T}^{n,2n} - b_{\frac{i}{2^n}T}| + \sup_{t \in (t_i, t_i+1)} |\mathcal{V}_t^{n,2n} - b_t| \Big)$$

$$\leq \frac{1}{2^n} + \max_{i=0,\dots,2^n-1}\Big(\sup_{t\in(t_i,t_{i+1})} |\mathcal{V}_t^{n,2n} - b_t + b_{t_i} - b_{t_i}|\Big)$$
$$\leq \frac{1}{2^n} + \max_{i=0,\dots,2^n-1}\Big(|V_i^{n,2n} - b_{t_i}| + \sup_{t\in(t_i,t_{i+1})} |b_{t_i} - b_t|\Big)$$
$$\leq \frac{1}{2^{n-1}} + \sup_{s,t\in[0,T],|s-t|<\frac{1}{2^n}} |b_s - b_t|.$$

This implies that $A^n \subseteq B^n$, for each $n \in \mathbb{N}$: $\overline{\mathbb{P}}(B^n) = \overline{\mathbb{P}}(B^n \cap A^n) + \overline{\mathbb{P}}(B^n \cap A^{nc})$. By the reflection principle:

$$\overline{\mathbb{P}}(A^{nc}) = 2\overline{\mathbb{P}}(b_T \geq 2^n - 1) = 2\Big(1 - \Phi\Big(\frac{2^n-1}{\sqrt{T}}\Big)\Big) \overset{n\to\infty}{\longrightarrow} 0.$$

Thus:

$$\lim_{n\to\infty} \overline{\mathbb{P}}(B^n) = \lim_{n\to\infty} (\overline{\mathbb{P}}(A^n \cap B^n) + \overline{\mathbb{P}}(A^{nc} \cap B^n)) = \lim_{n\to\infty} \overline{\mathbb{P}}(A^n) + \overbrace{\lim_{n\to\infty} \overline{\mathbb{P}}(A^{nc} \cap B^n)}^{\leq \lim_{n\to\infty} \overline{\mathbb{P}}(A^{nc})=0} = 1. \quad (2.78)$$

For any $\epsilon > 0$, we introduce the events:

$$C^{n,\epsilon} := \Big\{ \sup_{t\in[0,T]} |\mathcal{V}_t^{n,2n} - b_t| \geq \epsilon \Big\}$$
$$D^{n,\epsilon} := \Big\{ \sup_{s,t\in[0,T]:\ |t-s|\leq\frac{1}{2^n}} |b_s - b_t| + \frac{1}{2^{n-1}} \geq \epsilon \Big\}.$$

In particular:

$$C^{n,\epsilon} = \Big\{ \sup_{t\in[0,T]} |\mathcal{V}_t^{n,2n} - b_t| \geq \epsilon + \varphi_n - \varphi_n \Big\},$$

where $\varphi_n := \sup_{s,t\in[0,T]:\ |t-s|\leq\frac{1}{2^n}} |b_s - b_t| + \frac{1}{2^{n-1}}$. Therefore, we can rewrite the events $D^{n,\epsilon}, C^{n,\epsilon}, B^n$ as follows:

$$D^{n,\epsilon} = \{\varphi_n - \epsilon \geq 0\}$$
$$C^{n,\epsilon} = \Big\{ \sup_{t\in[0,T]} |\mathcal{V}^{n,2n} - b_t| \geq \epsilon \Big\}$$
$$B^n = \Big\{ \sup_{t\in[0,T]} |\mathcal{V}^{n,2n} - b_t| \leq \varphi_n \Big\}$$

Therefore:

$$C^{n,\epsilon} = (C^{n,\epsilon} \cap D^{n,\epsilon}) \cup (C^{n,\epsilon} \cap D^{n,\epsilon c}) \subseteq D^{n,\epsilon} \cup B^{nc}. \quad (2.79)$$

By Lévy theorem ([KS98, Theorem 2.9.25]),

$$\widetilde{\Omega} := \Big\{ \lim_{n\to\infty} \frac{1}{g(2^{-n})} \sup_{s,t\in[0,T]:\ |t-s|\leq 2^{-n}} |b_s - b_t| = 1 \Big\},$$

then $\overline{\mathbb{P}}(\widetilde{\Omega}) = 1$, with $g(x) := \sqrt{2x \log \frac{1}{x}}$. Hence, for each $\omega \in \widetilde{\Omega}$:

$$\lim_{n\to\infty} \sup_{s,t\in[0,T]:\ |t-s|\leq 2^{-n}} |b_s(\omega) - b_t(\omega)| = 0,$$

meaning that

$$\forall \varepsilon > 0\ \exists\ \overline{n} \in \mathbb{N} : \forall n \geq \overline{n} : \sup_{s,t\in[0,T]:\ |t-s|\leq 2^{-n}} |b_s(\omega) - b_t(\omega)| < \varepsilon.$$

As a consequence, there exists $n_\epsilon \in \mathbb{N}$ such that $\forall n > n_\epsilon$:

$$\sup_{s,t\in[0,T]:\ |t-s|\leq 2^{-n}} |b_s(\omega) - b_t(\omega)| + \frac{1}{2^{n-1}} < \varepsilon, \quad \forall \omega \in \widetilde{\Omega}.$$

In conclusion, it holds that $D^{n,\epsilon} \cap \widetilde{\Omega} = \emptyset$, for all $n \geq \overline{n}_\epsilon$. Thus, by (2.78) and (2.79), the following holds

$$\lim_{n\to\infty} \overline{\mathbb{P}}(C^{n,\epsilon}) \leq \lim_{n\to\infty} (\overline{\mathbb{P}}(D^{n,\epsilon}) + \overline{\mathbb{P}}(B^{n\,c})) = 0.$$

This is equivalent to say that $\{\mathcal{V}^{n,2n}\}$ converges in probability on $\mathcal{C}([0,T];\mathbb{R})$ to b and therefore also in distribution. □

We apply Lemma 2.31, to prove the following result.

Theorem 2.32. *In the framework introduced at the beginning of this section, equation* (2.60) *is satisfied.*

Proof. Let us notice that (2.60) is equivalent to:

$$\mathbb{E}^\infty[h(\varpi^\infty_{\cdot\wedge t}, b^\infty_{\cdot\wedge t})(\varpi^\infty_t + (\overline{\Lambda} + \overline{\Lambda}^0)^{-1}(\overline{\Lambda} Y^\infty_t + \overline{\Lambda}^0 Y^{0,\infty}_t))] = 0, \quad t \in [0,T].$$

If we pass to the limit in equation (2.77), we get:

$$0 = \lim_{n\to\infty} \overline{\mathbb{E}}\Big[h(\varpi^{n,2n}_{\cdot\wedge t}, \mathcal{V}^{n,2n}_{\cdot\wedge t})(\varpi^{n,2n}_t + (\overline{\Lambda} + \overline{\Lambda}^0)^{-1}(\overline{\Lambda} Y^{n,2n}_t + \overline{\Lambda}^0 Y^{0,n,2n}_t))\Big].$$

To bring the limit inside the expectation, we need to approximate $\varpi^{n,2n}_t + (\overline{\Lambda} + \overline{\Lambda}^0)^{-1}(\overline{\Lambda} Y^{n,2n}_t + \overline{\Lambda}^0 Y^{0,n,2n}_t)$ with the bounded function $h_k(\varpi^{n,2n}_t + (\overline{\Lambda} + \overline{\Lambda}^0)^{-1}(\overline{\Lambda} Y^{n,2n}_t + \overline{\Lambda}^0 Y^{0,n,2n}_t))$, where

$$h_k(x) := \begin{cases} x & \text{if } |x| \leq k, \\ k \ \mathrm{sgn}(x) & \text{otherwise.} \end{cases}$$

We proceed as follows:

$$
\begin{aligned}
0 &= \lim_{n\to\infty} \overline{\mathbb{E}}\Big[h(\varpi^{n,2n}_{\cdot\wedge t}, \mathcal{V}^{n,2n}_{\cdot\wedge t})(\varpi^{n,2n}_t + (\overline{\Lambda}+\overline{\Lambda}^0)^{-1}(\overline{\Lambda}Y^{n,2n}_t + \overline{\Lambda}^0 Y^{0,n,2n}_t))\\
&= \lim_{n\to\infty} \Big\{\overline{\mathbb{E}}[h(\varpi^{n,2n}_{\cdot\wedge t}, \mathcal{V}^{n,2n}_{\cdot\wedge t})h_k(\varpi^{n,2n}_t + (\overline{\Lambda}+\overline{\Lambda}^0)^{-1}(\overline{\Lambda}Y^{n,2n}_t + \overline{\Lambda}^0 Y^{0,n,2n}_t))]+\\
&\quad + \overline{\mathbb{E}}\Big[h(\varpi^{n,2n}_{\cdot\wedge t}, \mathcal{V}^{n,2n}_{\cdot\wedge t})(\varpi^{n,2n}_t + (\overline{\Lambda}+\overline{\Lambda}^0)^{-1}(\overline{\Lambda}Y^{n,2n}_t + \overline{\Lambda}^0 Y^{0,n,2n}_t)+\\
&\quad - h_k(\varpi^{n,2n}_t + (\overline{\Lambda}+\overline{\Lambda}^0)^{-1}(\overline{\Lambda}Y^{n,2n}_t + \overline{\Lambda}^0 Y^{0,n,2n}_t)))\Big]\Big\}\\
&= \lim_{n\to\infty} \Big\{\overline{\mathbb{E}}[h(\varpi^{n,2n}_{\cdot\wedge t}, \mathcal{V}^{n,2n}_{\cdot\wedge t})h_k(\varpi^{n,2n}_t + (\overline{\Lambda}+\overline{\Lambda}^0)^{-1}(\overline{\Lambda}Y^{n,2n}_t + \overline{\Lambda}^0 Y^{0,n,2n}_t))] + \mathcal{A}^k_n(t)\Big\},
\end{aligned}
$$

where

$$
\begin{aligned}
\mathcal{A}^k_n(t) := \overline{\mathbb{E}}\Big[&h(\varpi^{n,2n}_{\cdot\wedge t}, \mathcal{V}^{n,2n}_{\cdot\wedge t})(\varpi^{n,2n}_t + (\overline{\Lambda}+\overline{\Lambda}^0)^{-1}(\overline{\Lambda}Y^{n,2n}_t + \overline{\Lambda}^0 Y^{0,n,2n}_t)+\\
&- h_k(\varpi^{n,2n}_t + (\overline{\Lambda}+\overline{\Lambda}^0)^{-1}(\overline{\Lambda}Y^{n,2n}_t + \overline{\Lambda}^0 Y^{0,n,2n}_t)))\Big]\Big\}
\end{aligned}
$$

By the convergence in distribution, the first term converges to:

$$
\begin{aligned}
\lim_{n\to\infty} &\overline{\mathbb{E}}[h(\varpi^{n,2n}_{\cdot\wedge t},\mathcal{V}^{n,2n}_{\cdot\wedge t})h_k(\varpi^{n,2n}_t + (\overline{\Lambda}+\overline{\Lambda}^0)^{-1}(\overline{\Lambda}Y^{n,2n}_t + \overline{\Lambda}^0 Y^{0,n,2n}_t))] =\\
&= \mathbb{E}^\infty[h(\varpi^\infty_{\cdot\wedge t}, b^\infty_{\cdot\wedge t})h_k(\varpi^\infty_t + (\overline{\Lambda}+\overline{\Lambda}^0)^{-1}(\overline{\Lambda}Y^\infty_t + \overline{\Lambda}^0 Y^{0,\infty}_t))],
\end{aligned}
\tag{2.80}
$$

for almost every $t \in [0,T]$. For the second term we introduce the events

$$
\begin{aligned}
A^k_n(t) &:= \{|\varpi^{n,2n}_t + (\overline{\Lambda}+\overline{\Lambda}^0)^{-1}(\overline{\Lambda}Y^{n,2n}_t + \overline{\Lambda}^0 Y^{0,n,2n}_t)| \geq k\},\\
\widetilde{A}^k_n &:= \{\sup_{t\in[0,T]} |\varpi^{n,2n}_t + (\overline{\Lambda}+\overline{\Lambda}^0)^{-1}(\overline{\Lambda}Y^{n,2n}_t + \overline{\Lambda}^0 Y^{0,n,2n}_t)| \geq k\}.
\end{aligned}
$$

Clearly $A^k_n(t) \subseteq \widetilde{A}^k_n$ for every $\in [0,T]$. Denoting by L_h a bound for h, we notice that

$$
\begin{aligned}
\mathcal{A}^k_n(t) &\leq |\overline{\mathbb{E}}[h(\varpi^{n,2n}_{\cdot\wedge t}, \mathcal{V}^{n,2n}_{\cdot\wedge t})((\varpi^{n,2n}_t + (\overline{\Lambda}+\overline{\Lambda}^0)^{-1}(\overline{\Lambda}Y^{n,2n}_t + \overline{\Lambda}^0 Y^{0,n,2n}_t))+\\
&\qquad - h_k(\varpi^{n,2n}_t + (\overline{\Lambda}+\overline{\Lambda}^0)^{-1}(\overline{\Lambda}Y^{n,2n}_t + \overline{\Lambda}^0 Y^{0,n,2n}_t)))]|\\
&\leq \overline{\mathbb{E}}[|h(\varpi^{n,2n}_{\cdot\wedge t}, \mathcal{V}^{n,2n}_{\cdot\wedge t})((\varpi^{n,2n}_t + (\overline{\Lambda}+\overline{\Lambda}^0)^{-1}(\overline{\Lambda}Y^{n,2n}_t + \overline{\Lambda}^0 Y^{0,n,2n}_t))+\\
&\qquad - h_k(\varpi^{n,2n}_t + (\overline{\Lambda}+\overline{\Lambda}^0)^{-1}(\overline{\Lambda}Y^{n,2n}_t + \overline{\Lambda}^0 Y^{0,n,2n}_t)))|\mathbb{1}_{A^k_n(t)}]\\
&\leq L_h\overline{\mathbb{E}}[\overbrace{(|(\varpi^{n,2n}_t + (\overline{\Lambda}+\overline{\Lambda}^0)^{-1}(\overline{\Lambda}Y^{n,2n}_t + \overline{\Lambda}^0 Y^{0,n,2n}_t))| - k)}^{\leq |(\varpi^{n,2n}_t + (\overline{\Lambda}+\overline{\Lambda}^0)^{-1}(\overline{\Lambda}Y^{n,2n}_t + \overline{\Lambda}^0 Y^{0,n,2n}_t))|}\mathbb{1}_{A^k_n(t)}]\\
&\leq L_h\Big[\overline{\mathbb{E}}[|\varpi^{n,2n}_t + (\overline{\Lambda}+\overline{\Lambda}^0)^{-1}(\overline{\Lambda}Y^{n,2n}_t + \overline{\Lambda}^0 Y^{0,n,2n}_t)|^2]\overline{\mathbb{P}}(A^k_n(t))\Big]^{\frac{1}{2}}.
\end{aligned}
$$

Applying (2.49), it is straightforward to show that

$$
\overline{\mathbb{P}}(A^k_n) \leq \frac{1}{k}\overline{\mathbb{E}}\Bigg[\sup_{t\in[0,T]} |\varpi^{n,2n}_t + (\overline{\Lambda}+\overline{\Lambda}^0)^{-1}(\overline{\Lambda}Y^{n,2n}_t + \overline{\Lambda}^0 Y^{0,n,2n}_t)|\Bigg] \leq \frac{1}{k}C
$$

Analogously $\overline{\mathbb{E}}[|\varpi_t^{n,2n} + (\overline{\Lambda} + \overline{\Lambda}^0)^{-1}(\overline{\Lambda} Y_t^{n,2n} + \overline{\Lambda}^0 Y_t^{0,n,2n})|^2] \leq CT$. Let us notice that $\widetilde{C}$ and C do not depend on n. Hence, $\mathcal{A}_n^k(t) \leq \frac{1}{\sqrt{k}}\overline{C}$, where $\overline{C}$ does not depend on n. Moreover, let us observe that

$$
\begin{aligned}
0 &= \lim_{n\to\infty} \overline{\mathbb{E}}[h(\varpi_{\cdot\wedge t}^{n,2n}, \mathcal{V}_{\cdot\wedge t}^{n,2n})(\varpi_t^{n,2n} + (\overline{\Lambda} + \overline{\Lambda}^0)^{-1}(\overline{\Lambda} Y_t^{n,2n} + \overline{\Lambda}^0 Y_t^{0,n,2n}))] \\
&= \lim_{n\to\infty} \left\{\overline{\mathbb{E}}[h(\varpi_{\cdot\wedge t}^{n,2n}, \mathcal{V}_{\cdot\wedge t}^{n,2n}) h_k(\varpi_t^{n,2n} + (\overline{\Lambda} + \overline{\Lambda}^0)^{-1}(\overline{\Lambda} Y_t^{n,2n} + \overline{\Lambda}^0 Y_t^{0,n,2n}))] + \mathcal{A}_n^k(t)\right\} \\
&= \mathbb{E}^\infty[h(\varpi_{\cdot\wedge t}^{\infty}, b_{\cdot\wedge t}^{\infty}) h_k(\varpi_t^{\infty} + (\overline{\Lambda} + \overline{\Lambda}^0)^{-1}(\overline{\Lambda} Y_t^{\infty} + \overline{\Lambda}^0 Y_t^{0,\infty})] + \lim_{n\to\infty} \mathcal{A}_n^k(t).
\end{aligned} \tag{2.80}
$$

This holds for every $k \in \mathbb{N}$. As a consequence, we can take the limit for $k \to \infty$:

$$
\begin{aligned}
0 &= \mathbb{E}^\infty[h(\varpi_{\cdot\wedge t}^{\infty}, b_{\cdot\wedge t}^{\infty}) h_k(\varpi_t^{\infty} + (\overline{\Lambda} + \overline{\Lambda}^0)^{-1}(\overline{\Lambda} Y_t^{\infty} + \overline{\Lambda}^0 Y_t^{0,\infty})] + \lim_{k\to\infty}\lim_{n\to\infty} \mathcal{A}_n^k(t) \\
&= \mathbb{E}^\infty[h(\varpi_{\cdot\wedge t}^{\infty}, b_{\cdot\wedge t}^{\infty})(\varpi_t^{\infty} + (\overline{\Lambda} + \overline{\Lambda}^0)^{-1}(\overline{\Lambda} Y_t^{\infty} + \overline{\Lambda}^0 Y_t^{0,\infty}))]
\end{aligned}
$$

Indeed:

$$
0 \leq \lim_{k\to\infty}\lim_{n\to\infty} \mathcal{A}_n^k(t) \leq \lim_{k\to\infty}\lim_{n\to\infty} \frac{\overline{C}}{\sqrt{k}} = 0.
$$

This implies that $\varpi_t^\infty = -(\overline{\Lambda} + \overline{\Lambda}^0)^{-1}\mathbb{E}^\infty[\overline{\Lambda} Y_t^\infty + \overline{\Lambda}^0 Y_t^{0,\infty} | \mathcal{F}_t^{\varpi^\infty, b^\infty}]$, a.e. $t \in [0,T]$, where $\widetilde{Y}^{0,\infty}$ adjoint process introduced in (2.73). By (2.74), we conclude that:

$$
\varpi_t^\infty = -(\overline{\Lambda} + \overline{\Lambda}^0)^{-1}(\mathbb{E}^\infty[\overline{\Lambda} Y_t^\infty | \mathcal{F}_t^{\varpi^\infty, b^\infty}] + \overline{\Lambda}^0 \widetilde{Y}_t^{0,\infty}). \tag{2.81}
$$

□

Remark 2.33. To guarantee that ϖ^∞ solves (2.81), we replaced $\mathbb{E}^\infty[Y_t^{0,\infty} | \mathcal{F}_t^{b^\infty, \varpi^\infty}]$ with $\widetilde{Y}_t^{0,\infty}$. To apply this substitution, the hypothesis of a deterministic initial value x_0 for the state variable of the major player is crucial. Indeed, if x_0 were defined as a random variable ξ^0 independent of b^∞ and ϖ^∞ and observable only by the major player, then we would have to add ξ^0 to the filtration $\mathbb{F}^{M,\infty}$, introduced in (2.55), to which the controls of the major player are adapted. As a consequence, its projection $\mathbb{F}^{M,\infty,p}$, introduced in (2.59), would be defined as $\mathbb{F}^{\xi^0, \varpi^\infty, b^\infty}$. As a consequence, the random process $\widetilde{Y}^{0,\infty}$ would not be adapted to $\mathbb{F}^{\varpi^\infty, b^\infty}$ and ϖ^∞ would not satisfy (2.81).

Since $\widetilde{Y}^{0,\infty}$ is adapted to $\mathbb{F}^{\varpi^\infty, b^\infty}$, we can define on $(\Omega^\infty, \mathcal{F}^\infty, \mathbb{P}^\infty, \mathbb{F}^{S,\infty})$ the two optimal control problems **S-IV** and **S-IV(2)** and conclude that ϖ^∞ is determined as the solution to (2.81).

We have now all the ingredients to prove Theorem 2.18

Proof of Theorem 2.18. In Subsection 2.3.3.2, we prove that the sequence $(\xi, b, \varpi^{n,l}, w, X^{n,l}, X^{0,n,l})$ is tight on $\Omega_{\text{input}} \times \mathcal{C}([0,T];\mathbb{R}) \times \mathcal{C}([0,T];\mathbb{R})$. Therefore, we introduce a weak limit, defined on a suitable probability space $(\Omega^\infty, \mathcal{F}^\infty, \mathbb{P}^\infty)$ and denoted by $(\xi^\infty, b^\infty, \varpi^\infty, w^\infty, X^\infty, X^{0,\infty})$. Hence, by Lemma 2.27, we show that $\widetilde{\Theta}^{M,\infty}$, introduced in (2.57) and $\widetilde{\Theta}^{S,\infty}$, introduced in (2.58), are respectively compatible with $\mathbb{F}^{M,\infty}$, introduced in (2.55) and $\mathbb{F}^{S,\infty}$, introduced in (2.56). This result, together with Proposition 2.28, guarantees that:

$$(\Omega^\infty, \mathcal{F}^\infty, \mathbb{P}^\infty, \mathbb{F}^{S,\infty}, (\xi^\infty, (b^\infty, w^\infty), \mathcal{W}^\infty)),$$
$$(\Omega^\infty, \mathcal{F}^\infty, \mathbb{P}^\infty, \mathbb{F}^{M,\infty}, (x_0, b^\infty, \mathcal{W}^{M,\infty}))$$

are admissible in the sense of Definition 2.9. Hence, we introduce the optimal control problems defined in **S-IV** and **S-V**. We exploit the linearity of the cost functional of the major player in the x variable to introduce, in **S-IV(2)**, another control problem, where the class of admissible controls is $\mathbb{H}^2(\mathbb{F}^{\varpi^\infty, b^\infty}; \mathbb{R})$. Finally, in Theorem 2.32, we derive the consistency condition (2.60) for the equilibrium price process ϖ^∞. In conclusion, we prove that $(\Omega^\infty, \mathcal{F}^\infty, \mathbb{P}^\infty, \mathbb{F}^{S,\infty}, (\xi^\infty, (b^\infty, w^\infty), \mathcal{W}^\infty))$ satisfies Definition 2.17. □

Remark 2.34 (On the strategy of the major player). The main consequence of Theorem 2.18 is that the major player cannot hide his strategy from the standard player. In analogy to Remark 2.13, we notice that as soon as the major agent adopts a strategy depending on her private information, the typical standard player can deduce it through the observation of the mean-field price process ϖ introduced in Definition 2.17:

$$\widehat{\beta}_t := -\overline{\Lambda}^0 (Y_t^0 + \varpi_t) = (\overline{\Lambda} + \overline{\Lambda}^0)\varpi_t + \mathbb{E}[\overline{\Lambda} Y_t | \mathcal{F}_t^{b,\varpi}] - \overline{\Lambda}^0 \varpi_t = \overline{\Lambda}(\varpi_t + \mathbb{E}[Y_t | \mathcal{F}_t^{b,\varpi}]), \quad t \in [0, T].$$

2.4 A weak version of the market clearing condition

2.4.1 Existence of stronger solutions under suitable conditions

In Section 2.2, we proposed a methodology to construct a stochastic process ϖ^{mf} that solves equation (2.20). As discussed in Section 2.3, we adapted the approach proposed in [CD18b, Chapter 3], based on the discretization of the common source of randomness. The natural question one may ask at this point is: *"given the existence of the solution to* (2.20)*, how is it related to the price process under which the market clearing condition is satisfied in the case of finitely many players?"* In other words, we want to understand how the substitution made passing to the mean-field limit impacts on the market clearing condition. In particular, we shall prove that a weak version of the market clearing condition (2.15) is satisfied, when the N players solve their stochastic optimal control problem taking the process ϖ^{mf} as an exogenous price process, under the additional assumption that the cost functional of the standard player is affine in the x variable.

To prove this result, the first issue we have to face is related to the weak form of Definition 2.17, obtained as the weak limit of a sequence of discretized solutions. In particular, we cannot fix a priori how rich the information structure given by the filtration $\mathbb{F}$ is, because the solution we are going to construct is defined only in distribution. Hence, we have to construct the $(N+1)$-player market (i.e. the objective functional and the dynamics introduced in Sections 2.2.2.1 and 2.2.2.2) in the probabilistic setup in which the weak equilibrium is defined. It is necessary to do so, because otherwise we cannot suppose that the $(N+1)$-player game is defined on a probability space sufficiently rich to guarantee the existence of ϖ^{mf}. We recall that the consistency condition in Definition 2.17 refers to an optimal control problem for the typical standard player that is adapted to a larger filtration (generated also by the Y-component of the lifted environment). The main consequence of the dependence on this larger filtration is that we cannot construct the solution on a canonical space of the form $\Omega^0 \times \Omega^1$, where Ω^0 should be the domain of $(b, \varpi^{\mathrm{mf}})$ and Ω^1 the domain of definition of the idiosyncratic noises (ξ, w). This is due to the fact that we cannot guarantee independence between Y and the idiosyncratic noise w. To overcome this difficulty, we require more structure to the optimal control problem of the standard player. In analogy to Subsection 2.3.3.5, we suppose additionally that the cost functional of the optimal control problem of the standard player is affine in the x variable:

Assumption 2.35. **A6** The functions $\overline{f}$ and $\overline{g}$ introduced in Section 2.2.2.1 satisfy:

$$\overline{f}(t, x, \varpi) = x c^S(t, \varpi),$$
$$\overline{g}(x, \varpi) = x g_1(\varpi).$$

for suitable continuous and bounded functions c^S and g_1.

In other words, the revenues of the standard player are cashflows dependent on the price process ϖ. Under Assumption **A6**, we can exploit the linear structure of the cost functional to define an optimal control problem for the typical standard player, where the admissible controls are adapted to the filtration $\mathbb{F}^{\varpi^\infty, b^\infty, w^\infty}$ (in analogy to the procedure followed in Subsection 2.3.3.5). Therefore, we introduce the following optimal control problem for the typical standard player, defined using the same coefficients introduced in **S-IV**

S-III(2)

$$\inf_{\alpha \in \mathbb{H}^2(\mathbb{F}^{\xi^\infty, b^\infty, \varpi^\infty, w^\infty}; \mathbb{R})} J^{\varpi^\infty}(\alpha), \quad J^{\varpi^\infty}(\alpha) = \mathbb{E}^\infty\left[\int_0^T f(s, X_s, \varpi_s^\infty, \alpha_s) ds + g(X_T, \varpi_T^\infty)\right]$$

and the state process solves the following SDE

$$\begin{cases} dX_t = (\alpha_t + l(t, \varpi_t^\infty))dt + \sigma_0(t, \varpi_t^\infty) db_t^\infty + \sigma(t, \varpi_t^\infty) dw_t^\infty, \\ X_0 = \xi^\infty. \end{cases}$$

The candidate optimal control for this problem is given by the function $\widehat{\alpha}(y, \varpi) := -\overline{\Lambda}(y + \varpi)$ introduced in (2.33). Hence, applying the stochastic maximum principle, the solution to the optimal control problem **S-III(2)** is defined by the following FBSDE:

$$\begin{cases} d\widetilde{X}_t^\infty = (-\overline{\Lambda}(\widetilde{Y}_t^\infty + \varpi_t^\infty) + l(t,\varpi_t^\infty))dt + \sigma(t,\varpi_t^\infty)db_t^\infty + \sigma(t,\varpi_t^\infty)dw_t^\infty, \\ d\widetilde{Y}_t^\infty = -c^S(t,\varpi_t^\infty)dt + \widetilde{Z}_t^{0,\infty}db_t^\infty + d\widetilde{M}_t^\infty, \\ \widetilde{X}_0^\infty = \xi^\infty, \\ \widetilde{Y}_T^\infty = g_1(\varpi_T^\infty). \end{cases} \tag{2.82}$$

In analogy to Subsection 2.3.3.5, it is straightforward to show that:

$$\mathbb{E}^\infty[\widetilde{Y}_t^\infty|\mathcal{F}_t^{b^\infty,\varpi^\infty}] = \mathbb{E}^\infty[Y_t^\infty|\mathcal{F}_t^{b^\infty,\varpi^\infty}],$$

where Y^∞ is the adjoint process associated with the optimal control problem introduced in **S-IV**. We recall that the class of admissible controls, in **S-IV**, was $\mathbb{H}^2(\mathbb{F}^{S,\infty};\mathbb{R})$, with $\mathbb{F}^{S,\infty}$ introduced in (2.56). As a consequence, the price process ϖ^∞, defined in (2.81) is equivalent to:

$$\varpi_t^\infty = -(\overline{\Lambda} + \overline{\Lambda}^0)^{-1}(\mathbb{E}^\infty[\overline{\Lambda}\widetilde{Y}_t^\infty|\mathcal{F}_t^{\varpi^\infty,b^\infty}] + \overline{\Lambda}^0\widetilde{Y}_t^{0,\infty}), \quad t\in[0,T]. \tag{2.83}$$

Inspired by this reasoning, we introduce the following definition

Definition 2.36 (Unlifted mean-field equilibrium price process)**.** We say that $(\Omega,\mathcal{F},\mathbb{P},\mathbb{F},\xi,(b,w),\varpi)$ is an *unlifted mean-field equilibrium price process* if:

- $\mathbb{F} = \mathbb{F}^{\xi,b,w,\varpi}$;
- $((\Omega,\mathcal{F},\mathbb{P}),\mathbb{F},(\xi,(b,w),\varpi))$ and $((\Omega,\mathcal{F},\mathbb{P}),\mathbb{F}^{\varpi,b},(x_0,(b,w),\varpi))$, for a constant $x_0\in\mathbb{R}$, are admissible probabilistic setups in the sense of Definition 2.9.
- ϖ solves equation (2.83), where $\widetilde{Y}^\infty$ and $\widetilde{Y}^{0,\infty}$ are respectively the backward components of the FBSDE (2.82) and (2.73), associated with the stochastic maximum principle.

In the remaining part this section, when the framework is clear from the context, we will refer to an unlifted mean-field equilibrium price process simply by mean-field equilibrium price process.

We can prove now the following result:

Theorem 2.37. *In the setting of Theorem 2.18, under the additional Assumption* ***A6****, an unlifted mean-field equilibrium price process in the sense of Definition 2.36 exists.*

Proof. By Theorem 2.18, a mean-field equilibrium price process

$$(\Omega^\infty,\mathcal{F}^\infty,\mathbb{P}^\infty,\mathbb{F}^\infty,\xi^\infty,(b^\infty,w^\infty),(\varpi^\infty,Y^\infty)),$$

satisfying Definition 2.17 exists. Under Assumption **A6**, we can introduce the optimal control problems **S-III(2)** and **S-IV(2)**. Therefore, by the Yamada-Watanabe theorem, in the version introduced in [CD18b, Theorem 1.33], the solutions of the FBSDE (2.82) and (2.73) are respectively adapted to $\mathbb{F}^{\xi,b,w,\varpi}$ and $\mathbb{F}^{b,\varpi}$. This fact, together with (2.83), yields the existence of an unlifted mean-field equilibrium price process in the sense of Definition 2.36. □

We consider an unlifted mean-field equilibrium price process:

$$(\Omega^\infty, \mathcal{F}^\infty, \mathbb{P}^\infty, \mathbb{F}^{\xi, b^\infty, \varpi^\infty, w^\infty}, \xi^\infty, (b^\infty, w^\infty), \varpi^\infty). \tag{2.84}$$

Following the approach described in the proof of [CD18b, Theorem 3.13], it is convenient to transfer (2.84) on the extended canonical space

$$\begin{cases} \overline{\Omega}^0 := \mathcal{C}([0,T];\mathbb{R}) \times \mathcal{D}([0,T];\mathbb{R}), \\ \overline{\Omega}^1 := [0,1) \times \mathcal{C}([0,T];\mathbb{R}). \end{cases}$$

As we are going to see in the next subsection, this procedure enables us to define the market with finitely many standard agents on the same probability space on which the mean-field equilibrium price process is defined. This is crucial to prove the asymptotic version of the market clearing condition satisfied by the mean-field equilibrium price process.

We equip $\overline{\Omega}^0$ with the law of the process $(\xi^\infty, b^\infty, w^\infty, \varpi^\infty)$ under $\mathbb{P}^\infty$. Moreover, introducing the Borel sigma-algebra $\overline{\mathcal{F}}^0$, we denote the canonical space by $(\overline{\Omega}^0, \overline{\mathcal{F}}^0, \overline{\mathbb{P}}^0)$. On $\overline{\Omega}$, we define the canonical process by $(b, \varpi^{\mathrm{mf}})$, while the canonical filtration is $\overline{\mathbb{F}}^0$. Similarly, we equip $\overline{\Omega}^1$ with the law Leb$\otimes\mathcal{W}$, where $\mathcal{W}$ is the Wiener measure. Introducing, the Borel sigma-algebra $\overline{\mathcal{F}}^1$, we denote the canonical space by $(\overline{\Omega}^1, \overline{\mathcal{F}}^1, \overline{\mathbb{P}}^1)$. Moreover, we define the canonical process by $\overline{\Omega}^1$ by (ξ^1, W^1) and the canonical filtration by $\overline{F}^1$. Hence, we denote by $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{F}}, \overline{\mathbb{P}})$ the completion of the product of the spaces $(\overline{\Omega}^0, \overline{\mathcal{F}}^0, \overline{\mathbb{P}}^0)$ and $(\overline{\Omega}^1, \overline{\mathcal{F}}^1, \overline{\mathbb{P}}^1)$, on which the canonical process $(\xi^1, b, \varpi^{\mathrm{mf}}, W^1)$ is defined. The filtration $\overline{\mathbb{F}}$ is defined as the complete and right continuous augmentation of the canonical filtration. We recall, by [CD18b, Theorem 1.33], there exist two measurable functions Φ^S and Φ^M such that:

$$(\widetilde{X}^\infty, \widetilde{Y}^\infty, \widetilde{Z}^{0,\infty}, \widetilde{Z}^\infty, \widetilde{M}^\infty) = \Phi^S(\xi^\infty, b^\infty, \varpi^\infty, w^\infty),$$
$$(\widetilde{X}^{0,\infty}, \widetilde{Y}^{0,\infty}, \widetilde{Z}^{0,0,\infty}, \widetilde{M}^{0,\infty}) = \Phi^M(b^\infty, \varpi^\infty).$$

We notice that the two distributions $\overline{\mathbb{P}} \circ (\xi^1, b, \varpi^{\mathrm{mf}}, W^1)^{-1}$ and $\mathbb{P}^\infty \circ (\xi^\infty, b^\infty, \varpi^\infty, w^\infty)^{-1}$ are equal. Thus, we can introduce the optimal control problems with the same coefficients of **S-III(2)** and **S-IV(2)** on $\overline{\Omega}$, where the filtration to which the controls are adapted is respectively given by $\overline{\mathbb{F}}$ and $\mathbb{F}^{\varpi^{\mathrm{mf}}, b}$. Again by [CD18b, Theorem 1.33], since

$$\overline{\mathbb{P}} \circ (\xi^1, (b, W^1), \varpi^{\mathrm{mf}})^{-1} = \mathbb{P}^\infty \circ (\xi^\infty, (b^\infty, w^\infty), \varpi^\infty)^{-1},$$

the solutions to the FBSDE systems associated with the stochastic maximum principle applied to the optimal control problems for the standard and the major agent defined on the canonical space are respectively defined by

$$(X^1, Y^1, Z^{0,1}, Z^1, M^1) = \Phi^S(\xi^1, b, \varpi^{\text{mf}}, W^1),$$
$$(X^0, Y^0, Z^{0,0}, M^0) = \Phi^M(b, \varpi^{\text{mf}}).$$

As a consequence, the law of the solutions of the optimal control problems introduced in $\overline{\Omega}$ coincides with the ones defined on Ω^∞. Therefore, we can conclude that:

$$\varpi_t^{\text{mf}} = -(\overline{\Lambda} + \overline{\Lambda}^0)^{-1}(\overline{\Lambda}\ \overline{\mathbb{E}}[Y_t^1|\mathcal{F}_t^{b,\varpi^{\text{mf}}}] + \overline{\Lambda}^0 Y_t^0), \quad \forall t \in [0,T]. \tag{2.85}$$

Indeed for $t \in [0,T]$, we have that

$$\begin{aligned} 0 &= \mathbb{E}^\infty[h(b^\infty_{\cdot\wedge t}, \varpi^\infty_{\cdot\wedge t})[\varpi_t^\infty + (\overline{\Lambda} + \overline{\Lambda}^0)^{-1}(\overline{\Lambda}\widetilde{Y}_t^\infty + \overline{\Lambda}^0 Y_t^{0,\infty})]] \\ &= \overline{\mathbb{E}}[h(b_{\cdot\wedge t}, \varpi^{\text{mf}}_{\cdot\wedge t})[\varpi_t^{\text{mf}} + (\overline{\Lambda} + \overline{\Lambda}^0)^{-1}(\overline{\Lambda} Y_t^1 + \overline{\Lambda}^0 Y_t^0)]]. \end{aligned} \tag{2.86}$$

2.4.2 Asymptotic market clearing condition

We consider $N-1$ copies of the space $(\Omega^1, \mathcal{F}^1, \mathbb{P}^1, \overline{\mathbb{F}}^1)$ (the second component of the canonical space introduced in Section 2.4.1), denoted by $(\Omega^j, \mathcal{F}^j, \mathbb{P}^j, \overline{F}^j)_{j=2}^N$. By construction, for each $j = 2, \ldots, N$, $(\Omega^j, \mathcal{F}^j, \mathbb{P}^j, \overline{F}^j)$ is rich enough to carry a Brownian motion $W^j = (W_t^j)_{t\in[0,T]}$ and a random variable ξ^j distributed as ξ and independent of W^j. Thus, we can define the product space:

$$\begin{cases} \overline{\Omega} & := \Omega^0 \times \Omega^1 \times \Omega^2 \times \cdots \times \Omega^N; \\ (\overline{\mathcal{F}}, \overline{\mathbb{P}}) & := (\mathcal{F}^0 \otimes \mathcal{F}^1 \otimes \cdots \otimes \mathcal{F}^N, \mathbb{P}^0 \otimes \ldots \mathbb{P}^N); \\ \overline{\mathbb{F}} & := (\mathcal{F}_t^0 \otimes \cdots \otimes \mathcal{F}_t^N)_{t\in[0,T]}. \end{cases} \tag{2.87}$$

As a consequence, on $\overline{\Omega}$, N independent Brownian motions $(W^j)_{j=1,\ldots,N}$ are defined. The j^{th} player, whose state variable is given by the process X^j defined on $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}}, \overline{\mathbb{F}})$, must solve its control problem applying controls belonging to the set $\mathbb{A}^j := \mathbb{H}^2(\mathbb{F}^{\varpi^{\text{mf}},b,W^j}; \mathbb{R})$. Analogously, the class of admissible controls for the major player is $\mathbb{A}^0 := \mathbb{H}^2(\mathbb{F}^{\varpi^{\text{mf}},b}; \mathbb{R})$. By Assumptions B3 and B4, the optimal controls for the j^{th}-standard player and the major player are given by:

$$\widehat{\alpha}_t^{j,\text{mf}} := -\overline{\Lambda}(Y_t^j + \varpi_t^{\text{mf}}), \tag{2.88}$$
$$\widehat{\beta}_t^{\text{mf}} := -\overline{\Lambda}^0(Y_t^0 + \varpi_t^{\text{mf}}), \tag{2.89}$$

whatever the price process ϖ^{mf} is. Analogously to equation (2.22), we can introduce the following system of FBSDE in $\overline{\Omega}$ for each $j = 1, \ldots, N$

$$\begin{cases} dX_t^j = (-\overline{\Lambda}(Y_t^j + \varpi_t^{\text{mf}}) + l(t, \varpi_t^{\text{mf}}))dt + \sigma^0(t, \varpi_t^{\text{mf}})db_t + \sigma(t, \varpi_t^{\text{mf}})dW_t^j, \quad X_0^j = \xi^j \\ dY_t^j = -c^S(t, \varpi_t^{\text{mf}})dt + Z_t^{0,j}db_t + Z_t^j dW_t + dM_t^j, \quad Y_T^j = g_1(\varpi_T^{\text{mf}}). \end{cases} \tag{2.90}$$

Moreover, on $\overline{\Omega}$, the FBSDE analogous to (2.23) is defined by

$$\begin{cases} dX_t^0 = -\overline{\Lambda}^0(Y_t^0 + \varpi_t^{\mathrm{mf}})dt + \sigma^0(t, \varpi_t^{\mathrm{mf}})db_t, \quad X_0^0 = x_0, \\ dY_t^0 = -c_0^M(t, \varpi_t^{\mathrm{mf}})dt + Z_t^{0,0}db_t + Z_t^0 dW_t + dM_t^0, \quad Y_T^0 = g_1^0(\varpi_T^{\mathrm{mf}}). \end{cases} \tag{2.91}$$

By construction $(\xi^j)_{j=1}^N$ is a sequence of i.i.d. random variables on $\overline{\Omega}$, distributed like the random variable ξ introduced in Subsection 2.2.2.1. Moreover, also ϖ^{mf}, introduced in (2.85), can be interpreted as a stochastic process defined on $\overline{\Omega}$.

We have now all the ingredients to prove the main result of this section

Theorem 2.38. *Assume that there exists a weak equilibrium ϖ^{mf} solution to* (2.85)*, a unique solution (X, Y, Z^0, Z, M) of system* (2.22) *and a unique solution $(X^0, Y^0, Z^{0,0}, M^0)$ of* (2.23) *in the probabilistic setup $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}}, \overline{\mathbb{F}})$ introduced in* (2.87)*. We consider the family of optimal controls for the $(N+1)$-player game in the form of equations* (2.88)*,* (2.89) *i.e. the optimal control in which the price is the equilibrium price for the limit game introduced in equation* (2.85)*. Then, the following weak version of the market clearing condition, given by the existence of a constant C, holds:*

$$\mathbb{E}\left[\int_0^T \left|\frac{1}{N}\sum_{j=1}^N \widehat{\alpha}_t^{mf,j} + \widehat{\beta}_t^{mf}\right|^2 dt\right] \leq \frac{C}{N}. \tag{2.92}$$

Proof. First of all, we notice that, by (2.85)

$$\begin{aligned} \frac{1}{N}\sum_{j=1}^N \widehat{\alpha}_t^{\mathrm{mf},j} + \widehat{\beta}_t^{\mathrm{mf}} &= \frac{1}{N}\sum_{j=1}^N -\overline{\Lambda}(Y_t^j + \varpi_t^{\mathrm{mf}}) - \overline{\Lambda}^0(Y_t^0 + \varpi_t^{\mathrm{mf}}) \\ &= -\overline{\Lambda}\left(\frac{1}{N}\sum_{j=1}^N Y_t^j\right) - \overline{\Lambda}\varpi_t^{\mathrm{mf}} - \overline{\Lambda}^0 Y_t^0 - \overline{\Lambda}^0\varpi_t^{\mathrm{mf}} \\ &= -\overline{\Lambda}\left(\frac{1}{N}\sum_{j=1}^N Y_t^j\right) + (\overline{\Lambda} + \overline{\Lambda}^0)(\overline{\Lambda} + \overline{\Lambda}^0)^{-1}(\overline{\Lambda}\mathbb{E}[Y_t|\mathcal{F}_t^{\varpi^{\mathrm{mf}},b}] + \overline{\Lambda}^0 Y_t^0) - \overline{\Lambda}^0 Y_t^0 \\ &= -\overline{\Lambda}\left(\frac{1}{N}\sum_{j=1}^N Y_t^j - \overline{\Lambda}\mathbb{E}[Y_t|\mathcal{F}_t^{\varpi^{\mathrm{mf}},b}]\right). \end{aligned} \tag{2.93}$$

Applying the Yamada-Watanabe theorem ([CD18b, Theorem 1.33]) to the FBSDEs defined on $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}}, \overline{\mathbb{F}})$, it is possible to show that there exists a progressively measurable function Ψ such that:

$$(X_t^j, Y_t^j)_{t\in[0,T]} := \Psi(\xi^j, b, (-\mathbb{E}[Y_t|\mathcal{F}_t^{\varpi^{\mathrm{mf}},b}])_{t\in[0,T]} + \Lambda^0 Y^0, W^j), \quad j = 1, \ldots, N, \tag{2.94}$$

where we separated the two stochastic terms defining the process ϖ^{mf}. As a consequence, $(Y^j)_j$ are i.i.d. conditionally to $\mathbb{F}^{\varpi^{\mathrm{mf}},b}$, because the only stochastic factor that is different in j is the idiosyncratic noise and the initial conditions, that form an i.i.d. random sequence in $j = 1, \ldots, N$ (the function Ψ is the same for each j). As we supposed, the idiosyncratic noises are a sequence of pairwise independent Brownian motions, thus the i.i.d. property holds. This implies that:

$$\mathbb{E}[Y_t^j | \mathcal{F}_t^{\varpi^{\mathrm{mf}},b}] = \mathbb{E}[Y_t^1 | \mathcal{F}_t^{\varpi^{\mathrm{mf}},b}], \quad \forall t \in [0,T], \quad \forall j = 1, \ldots, N.$$

Hence, we can substitute $\mathbb{E}[Y_t^j | \mathcal{F}_t^{\varpi^{\mathrm{mf}},b}]$ in the last term of (2.93), thus obtaining

$$\frac{1}{N} \sum_{j=1}^{N} \widehat{\alpha}_{mf,t}^j = -\overline{\Lambda} \bigg(\frac{1}{N} \sum_{j=1}^{N} Y_t^j - \mathbb{E}[Y_t^1 | \mathcal{F}_t^{\varpi^{\mathrm{mf}},b}] \bigg).$$

We notice that the function $F : \Omega^0 \times \Omega^n \to \mathbb{R}$ defined by:

$$F(t, (\omega^0, \ldots, \omega^N)) := \Big| \frac{1}{N} \sum_{j=1}^{N} Y_t^j(\omega^0, \omega^j) - \mathbb{E}[Y_t^1 | \mathcal{F}_t^{\varpi^{\mathrm{mf}},b}](\omega^0) \Big|^2,$$

is nonnegative and measurable (because Y^j is progressive measurable, see [CD18b, Remark 1.34]). This implies that we can apply Fubini's theorem

$$\mathbb{E}\left[\int_0^T \left| \frac{1}{N} \sum_{j=1}^{N} Y_t^j - \mathbb{E}[Y_t^1 | \mathcal{F}_t^{\varpi^{\mathrm{mf}},b}] \right|^2 dt \right] = \int_0^T \mathbb{E}\left[\left| \frac{1}{N} \sum_{j=1}^{N} Y_t^j - \mathbb{E}[Y_t^1 | \mathcal{F}_t^{\varpi^{\mathrm{mf}},b}] \right|^2 \right] dt.$$

We notice that:

$$\begin{aligned}
(A) &:= \mathbb{E}\left[\left| \frac{1}{N} \sum_{j=1}^{N} Y_t^j - \mathbb{E}[Y_t^1 | \mathcal{F}_t^{\varpi^{\mathrm{mf}},b}] \right|^2 \right] = \frac{1}{N^2} \mathbb{E}\left[\left| \sum_{j=1}^{N} \Big(Y_t^j - \mathbb{E}[Y_t^1 | \mathcal{F}_t^{\varpi^{\mathrm{mf}},b}] \Big) \right|^2 \right] \\
&= \frac{1}{N^2} \mathbb{E}\left[\sum_{j=1}^{N} \left| Y_t^j - \mathbb{E}[Y_t^j | \mathcal{F}_t^{\varpi^{\mathrm{mf}},b}] \right|^2 \right] + 2\mathbb{E}\left[\sum_{h,k=1,\ h\neq k}^{N} \Big(Y_t^h - \mathbb{E}[Y_t^h | \mathcal{F}_t^{\varpi^{\mathrm{mf}},b}] \Big) \Big(Y_t^k - \mathbb{E}[Y_t^k | \mathcal{F}_t^{\varpi^{\mathrm{mf}},b}] \Big) \right].
\end{aligned}$$

Moreover

$$\begin{aligned}
(B_{h,k}) &:= \mathbb{E}\Big[\Big(Y_t^h - \mathbb{E}[Y_t^h | \mathcal{F}_t^{\varpi^{\mathrm{mf}},b}] \Big) \Big(Y_t^k - \mathbb{E}[Y_t^k | \mathcal{F}_t^{\varpi,b}] \Big) \Big] \\
&= \mathbb{E}\Big[Y_t^h Y_t^k \Big] - \mathbb{E}\Big[Y_t^h \mathbb{E}\Big[Y_t^k | \mathcal{F}_t^{\varpi^{\mathrm{mf}},b} \Big] \Big] - \mathbb{E}\Big[Y_t^k \mathbb{E}\Big[Y_t^h | \mathcal{F}_t^{\varpi^{\mathrm{mf}},b} \Big] \Big] \\
&\quad + \mathbb{E}\Big[\mathbb{E}\Big[Y_t^k | \mathcal{F}_t^{\varpi^{\mathrm{mf}},b} \Big] \mathbb{E}\Big[Y_t^h | \mathcal{F}_t^{\varpi^{\mathrm{mf}},b} \Big] \Big] \\
&= \mathbb{E}\Big[Y_t^h Y_t^k \Big] - \mathbb{E}\Big[\mathbb{E}\Big[Y_t^k | \mathcal{F}_t^{\varpi^{\mathrm{mf}},b} \Big] \mathbb{E}\Big[Y_t^h | \mathcal{F}_t^{\varpi^{\mathrm{mf}},b} \Big] \Big] = 0,
\end{aligned}$$

where the third equality holds by the tower property and the fourth is due to the fact that $Y^j = \Phi_y(\xi^j, b, \varpi^{\mathrm{mf}}, W^j)$ (Ψ_y is the projection on the Y component of the function Ψ introduced in equation (2.94)). Again by the tower property:

$$\mathbb{E}\Big[Y_t^h Y_t^k\Big] = \mathbb{E}\Big[\mathbb{E}\Big[Y_t^h Y_t^k \Big| \mathcal{F}_t^{\varpi^{\mathrm{mf}},b}\Big]\Big] = \mathbb{E}\Big[\mathbb{E}\Big[Y_t^k | \mathcal{F}_t^{\varpi^{\mathrm{mf}},b}\Big]\mathbb{E}\Big[Y_t^h | \mathcal{F}_t^{\varpi^{\mathrm{mf}},b}\Big]\Big].$$

Hence, we observe that

$$\begin{aligned}\mathbb{E}\Big[\Big|Y_t^j - \mathbb{E}[Y_t^j|\mathcal{F}_t^{\varpi^{\mathrm{mf}},b}]\Big] &= \mathbb{E}\Big[\Big|Y_t^j\Big|^2\Big] - 2\mathbb{E}\Big[Y_t^j\mathbb{E}\Big[Y_t^j\Big|\mathcal{F}_t^{\varpi^{\mathrm{mf}},b}\Big]\Big] + \mathbb{E}\Big[\mathbb{E}^2\Big[Y_t^j\Big|\mathcal{F}_t^{\varpi^{\mathrm{mf}},b}\Big]\Big] \\ &= \mathbb{E}\Big[\Big|Y_t^j\Big|^2\Big] - \mathbb{E}\Big[\mathbb{E}^2\Big[Y_t^j\Big|\mathcal{F}_t^{\varpi^{\mathrm{mf}},b}\Big]\Big]\end{aligned}$$

Since by the Yamada Watanabe theorem $(Y_t^j)_{j=1}^{N_n}$ is a sequence of $\mathcal{F}_t^{\varpi^{\mathrm{mf}},b}$-conditionally i.i.d. random variables, also $\mathbb{E}\Big[Y_t^j\Big|\mathcal{F}_t^{\varpi^{\mathrm{mf}},b}\Big]\Big]$ is i.i.d. and

$$\mathbb{E}\Big[|Y_t^j|^2\Big] = \mathbb{E}\Big[\mathbb{E}\Big[|Y_t^j|^2\Big|\mathcal{F}_t^{\varpi^{\mathrm{mf}},b}\Big]\Big] = \mathbb{E}\Big[\mathbb{E}\Big[|Y_t^1|^2\Big|\mathcal{F}_t^{\varpi^{\mathrm{mf}},b}\Big]\Big] = \mathbb{E}\Big[|Y_t^1|^2\Big].$$

Finally, we conclude that

$$\begin{aligned}\mathbb{E}\left[\left|\frac{1}{N}\sum_{j=1}^N Y_t^j - \mathbb{E}[Y_t^1|\mathcal{F}_t^{\varpi^{\mathrm{mf}},b}]\right|^2\right] &= \frac{1}{N^2}\mathbb{E}\left[\sum_{j=1}^N \Big|Y_t^j - \mathbb{E}[Y_t^j|\mathcal{F}_t^{\varpi^{\mathrm{mf}},b}]\Big|^2\right] \\ &= \frac{1}{N^2}\sum_{j=1}^N\Big(\mathbb{E}\Big[\Big|Y_t^j\Big|^2\Big] - \mathbb{E}\Big[\mathbb{E}^2\Big[Y_t^j\Big|\mathcal{F}_t^{\varpi^{\mathrm{mf}},b}\Big]\Big]\Big) \\ &= \frac{1}{N^2}\sum_{j=1}^N\Big(\mathbb{E}\Big[\Big|Y_t^1\Big|^2\Big] - \mathbb{E}\Big[\mathbb{E}^2\Big[Y_t^1\Big|\mathcal{F}_t^{\varpi^{\mathrm{mf}},b}\Big]\Big]\Big) \\ &= \frac{1}{N}\Big(\mathbb{E}\Big[\Big|Y_t^1\Big|^2\Big] - \mathbb{E}\Big[\Big(\mathbb{E}\Big[Y_t^1\Big|\mathcal{F}_t^{\varpi^{\mathrm{mf}},b}\Big]\Big)^2\Big]\Big) \overset{N\to\infty}{\longrightarrow} 0,\end{aligned}$$

because the two terms in the numerator are finite. Indeed, by Jensen's inequality:

$$\mathbb{E}\Big[\Big(\mathbb{E}\Big[Y_t^1\Big|\mathcal{F}_t^{\varpi^{\mathrm{mf}},b}\Big]\Big)^2\Big] \le \mathbb{E}\Big[\mathbb{E}\Big[|Y_t^1|^2\Big|\mathcal{F}_t^{\varpi^{\mathrm{mf}},b}\Big]\Big] = \mathbb{E}\Big[\Big|Y_t^1\Big|^2\Big] \le \mathbb{E}\Big[\sup_{t\in[0,T]}\Big|Y_t^1\Big|^2\Big] < \infty.$$

In conclusion, since by construction $Y^1 \in \mathbb{S}^2(\mathbb{F};\mathbb{R})$, we have that

$$\begin{aligned}\int_0^T \mathbb{E}\left[\left|\frac{1}{N}\sum_{j=1}^N Y_t^j - \mathbb{E}[Y_t^1|\mathcal{F}_t^{\varpi^{\mathrm{mf}},b}]\right|^2\right]dt &= \int_0^T \frac{1}{N}\Big(\mathbb{E}\Big[\Big|Y_t^1\Big|^2\Big] - \mathbb{E}\Big[\Big(\mathbb{E}\Big[Y_t^1\Big|\mathcal{F}_t^{\varpi^{\mathrm{mf}},b}\Big]\Big)^2\Big]\Big)dt \\ &\le \int_0^T \frac{2}{N}\mathbb{E}\Big[\sup_{s\in[0,T]}\Big|Y_s^1\Big|^2\Big]ds \\ &\le \frac{2T\mathbb{E}\Big[\sup_{s\in[0,T]}\Big|Y_s^1\Big|^2\Big]}{N} =: \frac{C}{N}.\end{aligned}$$

□

2.5 Conclusions and further developments

In this chapter we proved the existence of the mean-field limit of a stochastic process representing the market clearing price ϖ in a market populated by a family of agents accessing different amounts of information. Due to the complexity of the market clearing condition (2.15), we considered the mean-field limit $N \to \infty$ in the number of standard agents. We derived the mean-field equation for the price process (see (2.20)). We proved the existence of a solution to (2.20) in the canonical space. The approximation of the price process ϖ satisfying the market clearing condition for the $(N+1)$-player market determined by the mean-field equation (2.20) can be motivated as in Section 2.4. In particular, we considered a market with finitely many agents (N standard agents and one major agent), for which the cost functional is affine in the x variable and we assumed that every agent takes as given price process the solution to (2.20). Hence, the asymptotic formulation of the market clearing condition expressed by (2.92) is satisfied.

We observed that the market clearing condition (2.15) establishes a relation between the strategies of the major player and the standard agents. The strategy of the major player is not a priori measurable with respect the filtration generated by the stochastic processes that can be observed by the standard players. However, the market clearing condition imposes an intrinsic link between the strategy of the major player, the price process ϖ and the common noise b. Indeed, as discussed in Section 2.2.4.2, $\widehat{\beta}^{(N)}$, defined in (2.11), is adapted to the filtration generated by the processes ϖ and b. This implies that, the way in which the additional source of information is exploited by the major player can be measured by the standard agents through the observation of the market clearing price process. Indeed, when the major player exploits this additional information in the choice of the strategy $\widehat{\beta}^{(N)}$, (i.e. she applies a strategy $\widehat{\beta}^{(N)}$ that does not depend only on b), immediately $\widehat{\beta}^{(N)}$ (and consequently the application of the additional information) can be deduced by the standard players, through the observation of the equilibrium price ϖ. Under this measurability constraint, the target function of the major player is supposed to be adapted to the filtration generated by the common noise b and the market clearing price process ϖ.

To construct the solution to (2.20), we adapted the strategy the developed in [CD18b, Chapter 3] to the setting of continuous flows of probability measure to the context of càdlàg processes. We constructed the solution to (2.20) on the canonical space as the weak limit of approximated solution defined on a discretized space.

Further developments Focusing on the technical challenges of our approach, an aspect worth of further investigation is the issue of compatibility. In particular, in Section 2.3.3.3, we guaranteed the compatibility between the process $(\xi^\infty, b^\infty, \varpi^\infty, w^\infty)$ and the filtration $\mathbb{F}^\infty$ of typical standard player, substituting ϖ^∞ with the lifted process $(\varpi^\infty, Y^\infty)$. The main problem of this approach is that Y^∞ is not independent of w^∞. As a consequence, we cannot transfer the solution on the extended canonical space endowed with a probability measure of the form $\mathbb{P}^0 \otimes \mathbb{P}^1$, where $\mathbb{P}^0$ is associated with the law of the triplet $(b^\infty, \varpi^\infty, Y^\infty)$ and $\mathbb{P}^1$ the law of (ξ^∞, w^∞). This is crucial to define the $(N+1)$-player game in the mean-field setting. It would be interesting to understand if it is possible to either give a financial interpretation to the lifted process $(\varpi^\infty, Y^\infty)$ or lift the price process ϖ^∞, adding a stochastic process that captures the extra randomness given by Y^∞ but that is independent of w^∞.

Another issue to address is related to the proof of the existence of a solution ϖ^N to equation (2.17) for the equilibrium price process for the $(N+1)$-player market. This should be a first step to show that, at least in the case of affine costs in the x variable, the convergence of $(\varpi^N)_{N\in\mathbb{N}}$ to the mean-field price ϖ^{mf} solution to (2.20) is guaranteed. In [FT22c], strong convergence of the $(N+1)$-player market to the mean-field one is proved. However, the approach of [FT22c] cannot be directly applied to our setting.

A possible generalization of the market model can take into account the case in which the major informed player is no longer a price taker, but she can manipulate the price through her strategy, in analogy to [FT22b]. It could be interesting to establish a link between this model and the results of [BCR23].

Generalizing the setting we developed, we could also study the case of a market model populated by two families of agents that access different amount of information. In particular, we could consider the case of a family of less informed agents and a family of more informed agents who share the knowledge of an additional stochastic factor inaccessible to the others.

Finally, it would be interesting to understand if it is possible to interpret the results we provided adopting an approach related to the theory of mean-field games of controls (see [Dje23a; Dje23b]). The equilibrium price process is determined by the market clearing, that is a condition on the optimal controls of every agents. Hence, we could investigate if the mechanism of price formation can be analysed introducing a family of mean-field optimal control problems, whose coefficients depend on a condition on the controls.

Appendix

2.A Stability of the solution to the discretized game

In this section we prove that, under Assumption 2.4 and Assumption 2.6, condition (2.41) is satisfied. First, we prove a preliminary stability result for the solution to system (2.32).

Lemma 2.39. *Let us consider two solutions of system* (2.32)*, denoted by* $(X^1, Y^1, Z^{0,1}, Z^1)$ *and* $(X^2, Y^2, Z^{0,2}, Z^2)$*, for two different compatible processes* $(X^{0,1}, b, w, \varpi^1)$*,* $(X^{0,2}, b, w, \varpi^2)$*. Then, there exists a constant* $C \geq 0$ *(depending on constant* L *introduced in Assumption B3 and on* T*)*

such that:

$$
\begin{aligned}
&\mathbb{E}\left[\int_0^T |Y_t^1 - Y_t^2|^2 dt + \int_0^T (|Z_t^{0,1} - Z_t^{0,2}|^2 + |Z_t^1 - Z_t^2|^2) dt + \sup_{t\in[0,T]} |Y_t^1 - Y_t^2|^2 \middle| \overline{\mathcal{F}}_0\right] \\
&\quad \leq C(T,L)\mathbb{E}\Bigg[\sup_{t\in[0,T]} |X_t^1 - X_t^2|^2 + |\partial_x g(X_T^1, \varpi_T^1) - \partial_x g(X_T^1, \varpi_T^2)|^2 \\
&\qquad + \int_0^T |\partial_x \overline{f}(t, X_t^1, \varpi_t^1) - \partial_x \overline{f}(t, X_t^1, \varpi_t^2)|^2 dt \Bigg| \overline{\mathcal{F}}_0\Bigg].
\end{aligned}
\tag{2.95}
$$

Proof. We recall that:

$$
Y_t^j = \partial_x g(X_T^j, \varpi_T^j) + \int_t^T \partial_x \overline{f}(s, X_s^j, \varpi_s^j) ds - \int_t^T Z_s^{0,1} db_s - \int_t^T Z_s^j dw_s.
$$

Moreover, by Itô's formula we have that

$$
\begin{aligned}
|Y_t^1 - Y_t^2|^2 &= |Y_0^1 - Y_0^2|^2 + \int_0^t (Y_s^1 - Y_s^2) d(Y_s^1 - Y_s^2) + \int_0^t d\langle Y_1 - Y_2\rangle_s \\
&= |Y_0^1 - Y_0^2|^2 + 2\int_0^t (Y_s^1 - Y_s^2)(-\partial_x \overline{f}(s, X_s^1, \varpi_s^1) + \partial_x \overline{f}(s, X_s^2, \varpi_s^2)) ds \\
&\quad + 2\int_0^t (Y_s^1 - Y_s^2)(Z_s^{0,1} - Z_s^{0,2}) db_s + 2\int_0^t (Y_s^1 - Y_s^2)(Z_s^1 - Z_s^2) dw_s \\
&\quad + \int_0^t [|Z_s^{0,1} - Z_s^{0,2}|^2 + |Z_s^1 - Z_s^2|^2] ds.
\end{aligned}
$$

On the other hand:

$$
\begin{aligned}
|Y_t^1 - Y_t^2|^2 &= |Y_T^1 - Y_T^2|^2 - (|Y_T^1 - Y_T^2|^2 - |Y_t^1 - Y_t^2|^2) \\
&= |\partial_x g(X_T^1, \varpi_T^1) - \partial_x g(X_T^2, \varpi_T^2)|^2 - \Big\{ 2\int_t^T (Y_s^1 - Y_s^2)(-\partial_x \overline{f}(s, X_s^1, \varpi_s^1) \\
&\quad + \partial_x \overline{f}(s, X_s^2, \varpi_s^2)) ds + 2\int_t^T (Y_s^1 - Y_s^2)(Z_s^{0,1} - Z_s^{0,2}) db_s \\
&\quad + 2\int_t^T (Y_s^1 - Y_s^2)(Z_s^1 - Z_s^2) dw_s + \int_t^T [|Z_s^{0,1} - Z_s^{0,2}|^2 + |Z_s^1 - Z_s^2|^2] ds \Big\}.
\end{aligned}
$$

We apply the conditional expectation with respect $\overline{\mathcal{F}}_0$ and by B3, obtain

$$
\begin{aligned}
(A_1) :=& \mathbb{E}\bigg[|Y_t^1 - Y_t^2|^2 + \int_t^T [|Z_s^{0,1} - Z_s^{0,2}|^2 + |Z_s^1 - Z_s^2|^2] ds \bigg| \overline{\mathcal{F}}_0 \bigg] = \\
=& \mathbb{E}\bigg[|\partial_x g(X_T^1, \varpi_T^1) - \partial_x g(X_T^2, \varpi_T^2)|^2 + 2\Big[\int_t^T (Y_s^1 - Y_s^2)(\partial_x \overline{f}(s, X_s^1, \varpi_s^1) \\
& - \partial_x \overline{f}(s, X_s^2, \varpi_s^2)) ds\Big] \bigg| \overline{\mathcal{F}}_0 \bigg] \\
\leq& \mathbb{E}\bigg[|\partial_x g(X_T^1, \varpi_T^1) - \partial_x g(X_T^2, \varpi_T^2)|^2 + \Big[\int_t^T \Big(|Y_s^1 - Y_s^2|^2 + |\partial_x \overline{f}(s, X_s^1, \varpi_s^1) \\
& - \partial_x \overline{f}(s, X_s^2, \varpi_s^2)|^2\Big) ds\Big] \bigg| \overline{\mathcal{F}}_0 \bigg] \\
\leq& \mathbb{E}\bigg[|\partial_x g(X_T^1, \varpi_T^1) - \partial_x g(X_T^2, \varpi_T^2)|^2 + \int_t^T |Y_s^1 - Y_s^2|^2 ds + 2\int_t^T [|\partial_x \overline{f}(s, X_s^1, \varpi_s^1) \\
& - \partial_x \overline{f}(s, X_s^1, \varpi_s^2)|^2 + |\partial_x \overline{f}(s, X_s^1, \varpi_s^2) - \partial_x \overline{f}(s, X_s^2, \varpi_s^2)|^2] ds \bigg| \overline{\mathcal{F}}_0 \bigg] \\
\leq& \mathbb{E}\bigg[|\partial_x g(X_T^1, \varpi_T^1) - \partial_x g(X_T^2, \varpi_T^2)|^2 + \int_t^T |Y_s^1 - Y_s^2|^2 ds + 2\int_t^T [|\partial_x \overline{f}(s, X_s^1, \varpi_s^1) \\
& - \partial_x \overline{f}(s, X_s^1, \varpi_s^2)|^2 + L^2 |X_s^1 - X_s^2|^2] ds \bigg| \overline{\mathcal{F}}_0 \bigg] \\
=:& (B_1)
\end{aligned}
\tag{2.96}
$$

To have a more compact notation we define:

$$
\gamma_t := [|Z_t^{0,1} - Z_t^{0,2}|^2 + |Z_t^1 - Z_t^2|^2], \quad t \in [0, T]. \tag{2.97}
$$

We now multiply $\mathbb{E}\Big[|Y_t^1 - Y_t^2|^2 + \int_t^T [|Z_s^{0,1} - Z_s^{0,2}|^2 + |Z_s^1 - Z_s^2|^2] ds \Big| \overline{\mathcal{F}}_0\Big]$ with $\exp(\alpha t)$ and integrate in $[0,T]$. Hence, the left member of equation (2.96) is:

$$
\begin{aligned}
(A_1) &:= \int_0^T e^{\alpha t}\mathbb{E}\Big[|Y_t^1 - Y_t^2|^2 + \int_t^T \gamma_s ds\Big|\overline{\mathcal{F}}_0\Big]dt \\
&= \mathbb{E}\Big[\int_0^T e^{\alpha t}\Big(|Y_t^1 - Y_t^2|^2 + \int_t^T \gamma_s ds\Big)dt\Big|\overline{\mathcal{F}}_0\Big] \\
&= \mathbb{E}\Big[\int_0^T e^{\alpha t}|Y_t^1 - Y_t^2|^2 dt|\overline{\mathcal{F}}_0\Big] + \mathbb{E}\Big[\frac{e^{\alpha t}}{\alpha}\int_t^T \gamma_s ds\Big|_0^T + \int_0^T \frac{e^{\alpha t}}{\alpha}\gamma_t dt\Big|\overline{\mathcal{F}}_0\Big] \\
&= \mathbb{E}\Big[\int_0^T e^{\alpha t}|Y_t^1 - Y_t^2|^2 dt|\overline{\mathcal{F}}_0\Big] + \mathbb{E}\Big[\frac{-1}{\alpha}\int_0^T \gamma_t dt + \int_0^T \frac{e^{\alpha t}}{\alpha}\gamma_t dt\Big|\overline{\mathcal{F}}_0\Big] \\
&= \mathbb{E}\Big[\int_0^T e^{\alpha t}|Y_t^1 - Y_t^2|^2 dt + \int_0^T \Big(\frac{e^{\alpha t}-1}{\alpha}\Big)\gamma_t dt\Big|\overline{\mathcal{F}}_0\Big].
\end{aligned}
\tag{2.98}
$$

The right member can be computed as

$$
\begin{aligned}
(B_1) &:= \int_0^T e^{\alpha t}\mathbb{E}\Big[|\partial_x g(X_T^1, \varpi_T^1) - \partial_x g(X_T^2, \varpi_T^2)|^2 + \int_t^T |Y_s^1 - Y_s^2|^2 ds \\
&\quad + 2\int_t^T [|\partial_x \overline{f}(s, X_s^1, \varpi_s^1) - \partial_x \overline{f}(s, X_s^1, \varpi_s^2)|^2 + L^2|X_s^1 - X_s^2|^2]ds\Big|\overline{\mathcal{F}}_0\Big]dt = \\
&= \mathbb{E}\Big[\Big(\frac{e^{\alpha T}-1}{\alpha}\Big)|\partial_x g(X_T^1, \varpi_T^1) - \partial_x g(X_T^2, \varpi_T^2)|^2 + \int_0^T \Big(\frac{e^{\alpha t}-1}{\alpha}\Big)|Y_t^1 - Y_t^2|^2 dt \\
&\quad + 2\int_0^T \Big(\frac{e^{\alpha t}-1}{\alpha}\Big)[|\partial_x \overline{f}(t, X_t^1, \varpi_t^1) - \partial_x \overline{f}(t, X_t^1, \varpi_t^2)|^2 + L^2|X_t^1 - X_t^2|^2]dt\Big|\overline{\mathcal{F}}_0\Big]
\end{aligned}
\tag{2.99}
$$

Putting together inequalities (2.98) and equation (2.99), we conclude that

$$
\begin{aligned}
(A_1) &= \mathbb{E}\Big[\int_0^T \Big[e^{\alpha t} - \Big(\frac{e^{\alpha t}-1}{\alpha}\Big)\Big]|Y_t^1 - Y_t^2|^2 dt + \int_0^T \Big(\frac{e^{\alpha t}-1}{\alpha}\Big)\gamma_t dt\Big|\overline{\mathcal{F}}_0\Big] \\
&\le (B_1) \\
&= \mathbb{E}\Big[\Big(\frac{e^{\alpha T}-1}{\alpha}\Big)|\partial_x g(X_T^1, \varpi_T^1) - \partial_x g(X_T^2, \varpi_T^2)|^2 + 2\int_0^T \Big(\frac{e^{\alpha t}-1}{\alpha}\Big)[|\partial_x \overline{f}(t, X_t^1, \varpi_t^1) \\
&\quad - \partial_x \overline{f}(t, X_t^1, \varpi_t^2)|^2 + L^2|X_t^1 - X_t^2|^2]dt\Big|\overline{\mathcal{F}}_0\Big]
\end{aligned}
$$

If $\alpha = 1$ the coefficient of $|Y_t^1 - Y_t^2|^2$ is equal to one, thus, applying condition (2.99) and Assumption

B3, we have hat

$$
\begin{aligned}
\mathbb{E}\left[\int_0^T |Y_t^1 - Y_t^2|^2 dt \Big| \mathcal{F}_0\right] &\leq \mathbb{E}\left[\int_0^T |Y_t^1 - Y_t^2|^2 dt + \int_0^T (e^t - 1)\gamma_t dt \Big| \overline{\mathcal{F}}_0\right] \\
&\leq \mathbb{E}\Bigg[2\Big(e^T - 1\Big)L^2|X_T^1 - X_T^2|^2 + 2\Big(e^T - 1\Big)|\partial_x g(X_T^1, \varpi_T^1) \\
&\quad - \partial_x g(X_T^1, \varpi_T^2)|^2 + \int_0^T \Big(e^t - 1\Big)2L^2|X_t^1 - X_t^2|^2 dt \\
&\quad + 2\int_0^T \Big(e^t - 1\Big)|\partial_x \overline{f}(t, X_t^1, \varpi_t^1) - \partial_x \overline{f}(t, X_t^1, \varpi_t^2)|^2 dt \Big| \overline{\mathcal{F}}_0\Bigg] \\
&\leq \mathbb{E}\Bigg[\overbrace{2L^2\Big(e^T - 1\Big)(1+T)}^{:=C_1(T,L)} \sup_{t\in[0,T]} |X_t^1 - X_t^2|^2 \\
&\quad + 2\Big(e^T - 1\Big)\Bigg[\int_0^T |\partial_x \overline{f}(t, X_t^1, \varpi_t^1) - \partial_x \overline{f}(t, X_t^1, \varpi_t^2)|^2 dt \\
&\quad + |\partial_x g(X_T^1, \varpi_T^1) - \partial_x g(X_T^1, \varpi_T^2)|^2\Bigg]\Bigg| \overline{\mathcal{F}}_0\Bigg] \\
&\leq C_A(T,L)\mathbb{E}\Bigg[\sup_{t\in[0,T]} |X_t^1 - X_t^2|^2 + \int_0^T |\partial_x \overline{f}(t, X_t^1, \varpi_t^1) - \partial_x \overline{f}(t, X_t^1, \varpi_t^2)|^2 dt \\
&\quad + |\partial_x g(X_T^1, \varpi_T^1) - \partial_x g(X_T^1, \varpi_T^2)|^2 \Big| \overline{\mathcal{F}}_0\Bigg].
\end{aligned}
\tag{2.100}
$$

Applying inequality (2.96), together with (B3) and (2.100) at $t = 0$, we find

$$
\begin{aligned}
\mathbb{E}\left[\int_0^T \gamma_t dt \Big| \overline{\mathcal{F}}_0\right] &= \mathbb{E}\left[|Y_0^1 - Y_0^2|^2 + \int_0^T \gamma_t dt \Big| \overline{\mathcal{F}}_0\right] \\
&\leq \mathbb{E}\Bigg[|\partial_x g(X_T^1, \varpi_T^1) - \partial_x g(X_T^2, \varpi_T^2)|^2 + \int_0^T |Y_t^1 - Y_t^2|^2 dt \\
&\qquad + 2\int_0^T [|\partial_x \overline{f}(t, X_t^1, \varpi_t^1) - \partial_x \overline{f}(t, X_t^1, \varpi_t^2)|^2 + L^2 |X_t^1 - X_t^2|^2] dt \Big| \overline{\mathcal{F}}_0\Bigg] \\
&\leq \mathbb{E}\Bigg[2|\partial_x g(X_T^1, \varpi_T^1) - \partial_x g(X_T^1, \varpi_T^2)|^2 + 2L^2 |X_T^1 - X_T^2|^2 + \int_0^T |Y_t^1 - Y_t^2|^2 dt \\
&\qquad + 2\int_0^T [|\partial_x \overline{f}(t, X_t^1, \varpi_t^1) - \partial_x \overline{f}(t, X_t^1, \varpi_t^2)|^2 + L^2 |X_t^1 - X_t^2|^2] dt \Big| \overline{\mathcal{F}}_0\Bigg] \\
&\leq \mathbb{E}\Bigg[2|\partial_x g(X_T^1, \varpi_T^1) - \partial_x g(X_T^1, \varpi_T^2)|^2 + 2L^2 |X_T^1 - X_T^2|^2 \\
&\qquad + C_1(T, L) \sup_{t \in [0,T]} |X_t^1 - X_t^2|^2 + 2\Big(e^T - 1\Big)\Bigg[\int_0^T |\partial_x \overline{f}(t, X_t^1, \varpi_t^1) \\
&\qquad - \partial_x \overline{f}(t, X_t^1, \varpi_t^2)|^2 dt + |\partial_x g(X_T^1, \varpi_T^1) - \partial_x g(X_T^1, \varpi_T^2)|^2\Bigg] \\
&\qquad + 2\int_0^T [|\partial_x \overline{f}(t, X_t^1, \varpi_t^1) - \partial_x \overline{f}(t, X_t^1, \varpi_t^2)|^2 + L^2 |X_t^1 - X_t^2|^2] dt \Big| \overline{\mathcal{F}}_0\Bigg] \\
&= \mathbb{E}\Bigg[\Bigg(2L^2 + C_1(T, L) + 2L^2 T\Bigg) \sup_{t \in [0,T} |X_t^1 - X_t^2|^2 \\
&\qquad + 2e^T \Bigg(\int_0^T |\partial_x \overline{f}(t, X_t^1, \varpi_t^1) - \partial_x \overline{f}(t, X_t^1, \varpi_t^2)|^2 dt + |\partial_x g(X_T^1, \varpi_T^1) - \partial_x g(X_T^1, \varpi_T^2)|^2\Bigg) \Bigg| \overline{\mathcal{F}}_0\Bigg] \\
&= C_B(T, L) \mathbb{E}\Bigg[\sup_{t \in [0,T} |X_t^1 - X_t^2|^2 + \int_0^T |\partial_x \overline{f}(t, X_t^1, \varpi_t^1) - \partial_x \overline{f}(t, X_t^1, \varpi_t^2)|^2 dt \\
&\qquad + |\partial_x g(X_T^1, \varpi_T^1) - \partial_x g(X_T^1, \varpi_T^2)|^2 \Bigg| \overline{\mathcal{F}}_0\Bigg].
\end{aligned}
\tag{2.101}
$$

On the other hand, we apply the inequality $\left(\sum_{i=1}^{n} a_i\right)^2 \leq n \sum_{i=1}^{n} a_i^2$ as follows

$$
\begin{aligned}
|Y_t^1 - Y_t^2|^2 = &\bigg|(\partial_x g(X_T^1, \varpi_T^1) - \partial_x g(X_T^2, \varpi_T^2)) + \int_t^T (\partial_x \overline{f}(s, X_s^1, \varpi_s^1) + \\
&- \partial_x \overline{f}(s, X_s^2, \varpi_s^2))ds - \int_t^T (Z_s^{0,1} - Z_s^{0,2})db_s - \int_t^T (Z_s^1 - Z_s^2)dw_s\bigg|^2 \\
\leq 4 &\bigg\{ |\partial_x g(X_T^1, \varpi_T^1) - \partial_x g(X_T^2, \varpi_T^2)|^2 + \bigg|\int_t^T (\partial_x \overline{f}(s, X_s^1, \varpi_s^1) \\
&- \partial_x \overline{f}(s, X_s^2, \varpi_s^2))ds\bigg|^2 + \bigg|\int_t^T (Z_s^{0,1} - Z_s^{0,2})db_s\bigg|^2 + \bigg|\int_t^T (Z_s^1 - Z_s^2)dw_s\bigg|^2 \bigg\} \\
\leq 4 &\bigg\{ |\partial_x g(X_T^1, \varpi_T^1) - \partial_x g(X_T^2, \varpi_T^2)|^2 + \bigg|\int_t^T (\partial_x \overline{f}(s, X_s^1, \varpi_s^1) \\
&- \partial_x \overline{f}(s, X_s^2, \varpi_s^2))ds\bigg|^2 + 2\bigg|\int_0^T (Z_t^{0,1} - Z_t^{0,2})db_s\bigg|^2 + 2\bigg|\int_0^t (Z_s^{0,1} - Z_s^{0,2})db_s\bigg|^2 \\
&+ 2\bigg|\int_0^T (Z_t^1 - Z_t^2)dw_t\bigg|^2 + 2\bigg|\int_0^t (Z_s^1 - Z_s^2)dw_s\bigg|^2 \bigg\}.
\end{aligned}
\tag{2.102}
$$

By Jensen's inequality and Assumption B3, we have that

$$
\begin{aligned}
&\bigg|\int_t^T (\partial_x \overline{f}(s, X_s^1, \varpi_s^1) - \partial_x \overline{f}(s, X_s^2, \varpi_s^2))ds\bigg|^2 \\
&\quad = \bigg|\int_t^T (\partial_x \overline{f}(s, X_s^1, \varpi_s^1) - \partial_x \overline{f}(s, X_s^1, \varpi_s^2) + \partial_x \overline{f}(s, X_s^1, \varpi_s^2) - \partial_x \overline{f}(s, X_s^2, \varpi_s^2))ds\bigg|^2 \\
&\quad \leq 2\bigg\{\bigg|\int_t^T (\partial_x \overline{f}(s, X_s^1, \varpi_s^1) - \partial_x \overline{f}(s, X_s^1, \varpi_s^2))ds\bigg|^2 + \bigg|\int_t^T (\partial_x \overline{f}(s, X_s^1, \varpi_s^2) - \partial_x \overline{f}(s, X_s^2, \varpi_s^2))ds\bigg|^2\bigg\} \\
&\quad \leq 2(T-t)\bigg\{\int_t^T \Big|\partial_x \overline{f}(s, X_s^1, \varpi_s^1) - \partial_x \overline{f}(s, X_s^1, \varpi_s^2)\Big|^2 ds + \int_t^T \Big|\partial_x \overline{f}(s, X_s^1, \varpi_s^2) - \partial_x \overline{f}(s, X_s^2, \varpi_s^2)\Big|^2 ds\bigg\} \\
&\quad \leq 2(T-t)\bigg\{\int_t^T \Big|\partial_x \overline{f}(s, X_s^1, \varpi_s^1) - \partial_x \overline{f}(s, X_s^1, \varpi_s^2)\Big|^2 ds + \int_t^T L^2 \Big|X_s^1 - X_s^2\Big|^2 ds\bigg\}.
\end{aligned}
\tag{2.103}
$$

Let us consider the supremum in $[0,T]$ in equation (2.102). By (2.103) the following holds

$$\begin{aligned}\sup_{t\in[0,T]}|Y_t^1-Y_t^2|^2 \leq 4\Bigg\{&2|\partial_x g(X_T^1,\varpi_T^1)-\partial_x g(X_T^1,\varpi_T^2)|^2+2L^2|X_T^1-X_T^2|^2\\ &+2\sup_{t\in[0,T]}\Bigg\{(T-t)\Big[\int_t^T\Big|\partial_x\overline{f}(s,X_s^1,\varpi_s^1)-\partial_x\overline{f}(s,X_s^1,\varpi_s^2)\Big|^2ds\\ &+\int_t^T L^2\Big|X_s^1-X_s^2\Big|^2ds\Big]\Bigg\}+2\Big|\int_0^T(Z_t^{0,1}-Z_t^{0,2})db_s\Big|^2\\ &+2\Big|\int_0^T(Z_t^1-Z_t^2)dw_t\Big|^2+2\sup_{t\in[0,T]}\Bigg\{\Big|\int_0^t(Z_s^{0,1}-Z_s^{0,2})db_s\Big|^2\\ &+\Big|\int_0^t(Z_s^1-Z_s^2)dw_s\Big|^2\Bigg\}\Bigg\}.\end{aligned}$$

We take now the conditional expectation with respect to $\overline{\mathcal{F}}_0$. By Itô's isometry and Doob's inequality, the following holds:

$$\begin{aligned}&\mathbb{E}\Bigg[\sup_{t\in[0,T]}|Y_t^1-Y_t^2|^2|\overline{\mathcal{F}}_0\Bigg]\\ &\leq\mathbb{E}\Bigg[8|\partial_x g(X_T^1,\varpi_T^1)-\partial_x g(X_T^1,\varpi_T^2)|^2+8L^2(1+T^2)\sup_{t\in[0,T]}|X_t^1-X_t^2|^2\\ &\quad+T\int_0^T\Big|\partial_x\overline{f}(t,X_t^1,\varpi_t^1)-\partial_x\overline{f}(t,X_t^1,\varpi_t^2)\Big|^2dt\Big|\overline{\mathcal{F}}_0\Bigg]+8\mathbb{E}\Bigg[\int_0^T\gamma_t dt\Big|\overline{\mathcal{F}}_0\Bigg]\\ &\quad+8\mathbb{E}\Bigg[\sup_{t\in[0,T]}\Big|\int_0^t Z_s^{0,1}-Z_s^{0,2})db_s\Big|^2\Big|\overline{\mathcal{F}}_0\Bigg]+8\mathbb{E}\Bigg[\sup_{t\in[0,T]}\Big|\int_0^t(Z_s^1-Z_s^2)dw_s\Big|^2\Big|\overline{\mathcal{F}}_0\Bigg]\\ &\leq\mathbb{E}\Bigg[8|\partial_x g(X_T^1,\varpi_T^1)-\partial_x g(X_T^1,\varpi_T^2)|^2+8L^2(1+T^2)\sup_{t\in[0,T]}|X_t^1-X_t^2|^2\\ &\quad+T\int_0^T\Big|\partial_x\overline{f}(t,X_t^1,\varpi_t^1)-\partial_x\overline{f}(t,X_t^1,\varpi_t^2)\Big|^2dt\Big|\overline{\mathcal{F}}_0\Bigg]+8\mathbb{E}\Bigg[\int_0^T\gamma_t dt\Big|\overline{\mathcal{F}}_0\Bigg]+32\mathbb{E}\Bigg[\int_0^T\gamma_t dt\Big|\overline{\mathcal{F}}_0\Bigg].\end{aligned}$$

By inequality (2.101), we conclude

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,T]}|Y_t^1-Y_t^2|^2\middle|\overline{\mathcal{F}}_0\right] &\le \mathbb{E}\Bigg[8L^2(1+T^2)\sup_{t\in[0,T]}|X_t^1-X_t^2|^2+8|\partial_x g(X_T^1,\varpi_T^1)-\partial_x g(X_T^1,\varpi_T^2)|^2\\
&\quad + T\int_0^T\left|\partial_x\overline{f}(t,X_t^1,\varpi_t^1)-\partial_x\overline{f}(t,X_t^1,\varpi_t^2)\right|^2dt\Big|\overline{\mathcal{F}}_0\Bigg]\\
&\quad + 40\Bigg\{\mathbb{E}\Big[\Big(2L^2+C_1(T,L)+2L^2T\Big)\sup_{t\in[0,T}|X_t^1-X_t^2|^2\\
&\quad + 2e^T\Big(\int_0^T|\partial_x\overline{f}(t,X_t^1,\varpi_t^1)-\partial_x\overline{f}(t,X_t^1,\varpi_t^2)|^2dt\\
&\quad + |\partial_x g(X_T^1,\varpi_T^1)-\partial_x g(X_T^1,\varpi_T^2)|^2\Big)\Big|\overline{\mathcal{F}}_0\Big]\Bigg\}\\
&\le C_C(T,L)\mathbb{E}\Bigg[\sup_{t\in[0,T]}|X_t^1-X_t^2|^2+|\partial_x g(X_T^1,\varpi_T^1)-\partial_x g(X_T^1,\varpi_T^2)|^2\\
&\quad + \int_0^T\left|\partial_x\overline{f}(t,X_t^1,\varpi_t^1)-\partial_x\overline{f}(t,X_t^1,\varpi_t^2)\right|^2dt\Big|\overline{\mathcal{F}}_0\Bigg].
\end{aligned} \tag{2.104}$$

In conclusion, considering $C(T,L):=3\max\{C_A(T,L);C_B(T,L);C_C(T,L)\}$, we have the result. □

We now apply the previous result to check for (2.41).

Proposition 2.40. *We consider the t-inizialized filtered probability space* $(\overline{\Omega},\overline{\mathcal{F}},\mathbb{P},\mathbb{F})$ *on which the input process* $(b_s,\varpi_s,w_s)_{t\in[s,T]}$ *is defined. On this setup, Assumption 2.23 is satisfied.*

Proof. We consider the two solutions introduced in Assumption 2.23, that differ only for the initial value of the state process. By Assumption A3:

$$\begin{aligned}
g(x,\varpi)-g(x',\varpi) &= (\overline{g}_1(\varpi)x+\overline{g}_2(x))-(\overline{g}_1(\varpi)x'+\overline{g}_2(x'))\\
&= (\overline{g}_1(\varpi)(x-x')+(\overline{g}_2(x)-\overline{g}_2(x')))\le(\overline{g}_1(\varpi)(x-x')+\partial_x\overline{g}_2(x)(x-x'))\\
&= (x-x')(\overline{g}_1(\varpi)+\partial_x\overline{g}_2(x))=(x-x')\partial_x g(x,\varpi).
\end{aligned}$$

Therefore:

$$g(X_T^{t,x},\varpi_T)-g(X_T^{t,x'},\varpi_T)\le\partial_x g(X_T^{t,x},\varpi_T)(X_T^{t,x}-X_T^{t,x'})=Y_T^{t,x}(X_T^{t,x}-X_T^{t,x'}).$$

Hence, applying integration by parts:

$$\begin{aligned}
Y_T^{t,x}(X_T^{t,x}-X_T^{t,x'}) &= Y_t^{t,x}(X_t^{t,x}-X_t^{t,x'})+\int_t^T Y_s^{t,x}d(X_s^{t,x}-X_s^{t,x'})\\
&\quad + \int_t^T(X_s^{t,x}-X_s^{t,x'})dY_s^{t,x}+\langle X^{t,x}-X^{t,x'},Y^{t,x}\rangle_t^T
\end{aligned}$$

$$
\begin{aligned}
&= Y_t^{t,x}(x-x') + \int_t^T Y_s^{t,x}(\widehat{\alpha}_s^{t,x} + l(s,\varpi_s) - \widehat{\alpha}_s^{t,x'} - l(s,\varpi_s))ds \\
&\quad + \int_t^T Y_s^{t,x}(\sigma^0(s,\varpi_s) - \sigma^0(s,\varpi_s))db_s + \int_t^T Y_s^{t,x}(\sigma(s,\varpi_s) - \sigma(s,\varpi_s))dw_s \\
&\quad + \int_t^T (X_s^{t,x} - X_s^{t,x'})(-\partial_x \overline{f}(s, X_s^{t,x}, \varpi_s))ds + \int_t^T (X_s^{t,x} - X_s^{t,x'}) Z_s^{0,t,x} db_s \\
&\quad + \int_t^T (X_s^{t,x} - X_s^{t,x'}) Z_s^{t,x} dw_s + \int_t^T (\sigma^0(s,\varpi_s) - \sigma^0(s,\varpi_s)) Z_s^{0,t,x} ds \\
&\quad + \int_t^T (\sigma(s,\varpi_s) - \sigma(s,\varpi_s)) Z^{t,x} ds \\
&= Y_t^{t,x}(x-x') + \int_t^T Y_s^{t,x}(\widehat{\alpha}_s^{t,x} - \widehat{\alpha}_s^{t,x'})ds - \int_t^T (X_s^{t,x} - X_s^{t,x'})\partial_x \overline{f}(s, X_s^{t,x}, \varpi_s)ds \\
&\quad \int_t^T (X_s^{t,x} - X_s^{t,x'}) Z_s^{0,t,x} db_s + \int_t^T (X_s^{t,x} - X_s^{t,x'}) Z_s^{t,x} dw_s,
\end{aligned}
$$

where $\widehat{\alpha}^{t,x}$ is the optimal control for the control problem (2.2), (2.3) defined on the t-inizialized probabilistic setup. We consider now the conditional expectation with respect to $\overline{\mathcal{F}}_t$ of $Y_T^{t,x}(X_T^{t,x} - X_T^{t,x'})$. We recall that by Remark 2.11, the stochastic integrals are true martingales. By Assumption A3, the following inequality holds

$$
\begin{aligned}
\mathbb{E}[Y_T^{t,x}(X_T^{t,x} - X_T^{t,x'})|\overline{\mathcal{F}}_t] &= Y_t^{t,x}(x-x') + \mathbb{E}\Big[\int_t^T Y_s^{t,x}(\widehat{\alpha}_s^{t,x} - \widehat{\alpha}_s^{t,x'})ds - \int_t^T (X_s^{t,x} - X_s^{t,x'})\partial_x \overline{f}(s, X_s^{t,x}, \varpi_s)ds|\overline{\mathcal{F}}_t\Big] \\
&\leq Y_t^{t,x}(x-x') + \mathbb{E}\Big[\int_t^T Y_s^{t,x}(\widehat{\alpha}_s^{t,x} - \widehat{\alpha}_s^{t,x'})ds + \int_t^T (\overline{f}(s, X_s^{t,x'}, \varpi_s) - \overline{f}(s, X_s^{t,x}, \varpi_s))ds|\overline{\mathcal{F}}_t\Big].
\end{aligned}
$$

Hence, we have

$$
\begin{aligned}
g(X_T^{t,x}, \varpi_T) - g(X_T^{t,x'}, \varpi_T) \leq Y_t^{t,x}(x-x') + \mathbb{E}\Big[&\int_t^T Y_s^{t,x}(\widehat{\alpha}_s^{t,x} - \widehat{\alpha}_s^{t,x'})ds \\
&+ \int_t^T (\overline{f}(s, X_s^{t,x'}, \varpi_s) - \overline{f}(s, X_s^{t,x}, \varpi_s))ds|\overline{\mathcal{F}}_t\Big],
\end{aligned}
\tag{2.105}
$$

We focus now on the integral part of the cost functional. Let us observe that $\partial_a f(t,x,\varpi,a) = \varpi + \Lambda a$. We recall that, by the optimality condition introduced in Assumption B3:

$$
\widehat{\alpha}_t^{t,x} = -(\Lambda)^{-1}(Y_t^{t,x} + \varpi_t) \Rightarrow Y_t^{t,x} = -\Lambda \widehat{\alpha}_t^{t,x} - \varpi_t \Rightarrow \partial_a f(t, \widehat{X}_t^{t,x}, \varpi_t, \widehat{\alpha}_t^{t,x}) = -Y_t^{t,x},
$$

thanks to the linear-quadratic structure of the cost functional f. Therefore:

$$
\begin{aligned}
g(X_T^{t,x}, \varpi_T) - g(X_T^{t,x'}, \varpi_T) \leq Y_t^{t,x}(x-x') + \mathbb{E}\Big[&\int_t^T (-\partial_a f(t, \widehat{X}_t^{t,x}, \varpi_t, \widehat{\alpha}_t^{t,x}))(\widehat{\alpha}_s^{t,x} - \widehat{\alpha}_s^{t,x'})ds \\
&+ \int_t^T (\overline{f}(s, X_s^{t,x'}, \varpi_s) - \overline{f}(s, X_s^{t,x}, \varpi_s))ds\Big|\overline{\mathcal{F}}_t\Big].
\end{aligned}
\tag{2.106}
$$

We rewrite condition (2.106) as follows:

$$\begin{aligned} Y_t^{t,x}(x'-x) \leq \mathbb{E}\Bigg[g(X_T^{t,x'},\varpi_T) + \int_t^T \overline{f}(s,X_s^{t,x'},\varpi_s)ds\Big|\overline{\mathcal{F}}_t\Bigg] - \mathbb{E}\Bigg[g(X_T^{t,x},\varpi_T) \\ + \int_t^T \overline{f}(s,X_s^{t,x},\varpi_s)ds\Big|\overline{\mathcal{F}}_t\Bigg] + \mathbb{E}\Bigg[\int_t^T (\partial_a f(s,X_s^{t,x},\varpi_s))(\widehat{\alpha}_s^{t,x'} - \widehat{\alpha}_s^{t,x})ds\Big|\overline{\mathcal{F}}_t\Bigg]. \end{aligned} \tag{2.107}$$

By the convexity assumption A3, we observe that:

$$\begin{aligned} (\widehat{\alpha}_s^{t,x'} - \widehat{\alpha}_s^{t,x})(\partial_a f(s,X_s^{t,x},\varpi_s)) &\leq f(s,X_s^{t,x},\varpi_s,\widehat{\alpha}_s^{t,x'}) - f(s,X_s^{t,x},\varpi_s,\widehat{\alpha}_s^{t,x}) - \frac{1}{2}L^{-1}|\widehat{\alpha}_s^{t,x'} - \widehat{\alpha}_s^{t,x}|^2 \\ &\leq (\varpi_s\widehat{\alpha}_s^{t,x'}) + \frac{1}{2}\Lambda(\widehat{\alpha}_s^{t,x'})^2 + \overline{f}(s,X_s^{t,x},\varpi_s) - (\varpi_s\widehat{\alpha}_s^{t,x}) + \frac{1}{2}\Lambda(\widehat{\alpha}_s^{t,x'})^2 \\ &\quad + \overline{f}(s,X_s^{t,x},\varpi_s) - \frac{1}{2}L^{-1}|\widehat{\alpha}_s^{t,x'} - \widehat{\alpha}_s^{t,x}|^2 \\ &= \varpi_s(\widehat{\alpha}_s^{t,x'} - \widehat{\alpha}_s^{t,x}) + f(s,X_s^{t,x'},\varpi_s,\widehat{\alpha}_s^{t,x'}) - \overline{f}(s,X_s^{t,x'},\varpi_s) \\ &\quad - (f(s,X_s^{t,x},\varpi_s,\widehat{\alpha}_s^{t,x}) - \overline{f}(s,X_s^{t,x},\varpi_s)) - \frac{1}{2}L^{-1}|\widehat{\alpha}_s^{t,x'} - \widehat{\alpha}_s^{t,x}|^2. \end{aligned} \tag{2.108}$$

We apply condition (2.108) to inequality (2.107):

$$\begin{aligned} Y_t^{t,x}(x'-x) \leq \mathbb{E}\Bigg[g(X_T^{t,x'},\varpi_T) + \int_t^T f(s,X_s^{t,x'},\varpi_s,\widehat{\alpha}_s^{t,x'})ds\Big|\overline{\mathcal{F}}_t\Bigg] - \mathbb{E}\Bigg[g(X_T^{t,x},\varpi_T) \\ + \int_t^T f(s,X_s^{t,x},\varpi_s,\widehat{\alpha}_s^{t,x})ds\Big|\overline{\mathcal{F}}_t\Bigg] - \frac{1}{2}L^{-1}\mathbb{E}\Bigg[\int_t^T |\widehat{\alpha}_s^{t,x'} - \widehat{\alpha}_s^{t,x}|^2 ds\Big|\overline{\mathcal{F}}_t\Bigg] \end{aligned} \tag{2.109}$$

Exchanging the rule of x and x', we obtain that

$$-(Y_t^{t,x'} - Y_t^{t,x})(x'-x) \leq -L^{-1}\mathbb{E}\Bigg[\int_t^T |\widehat{\alpha}_s^{t,x} - \widehat{\alpha}_s^{t,x'}|^2 ds\Big|\overline{\mathcal{F}}_t\Bigg]. \tag{2.110}$$

We now observe that, by Jensen's inequality and Doob's inequalities, the following holds

$$\begin{aligned} \mathbb{E}\Bigg[\sup_s |X_s^{t,x} - X_s^{t,x'}|^2\overline{\mathcal{F}}_t\Bigg] &= \mathbb{E}\Bigg[\sup_{s\in[t,T]}\Big|x + \int_t^s (\alpha_u^{t,x} + l(u,\varpi_u))du + \int_t^s \sigma^0(u,\varpi_u)db_u \\ &\quad + \int_t^s \sigma(u,\varpi_u)dw_u - \Big[x' + \int_t^s (\alpha_u^{t,x'} + l(u,\varpi_u))du \\ &\quad + \int_t^s \sigma^0(u,\varpi_u)db_u + \int_t^s \sigma(u,\varpi_u)dw_u\Big|^2\Bigg] \\ &\leq 2\mathbb{E}\Bigg[|x-x'|^2 + \sup_{s\in[t,T]}\Big|\int_t^s (\widehat{\alpha}_u^{t,x} - \widehat{\alpha}_u^{t,x'})du\Big|^2\Big|\overline{\mathcal{F}}_t\Bigg] \end{aligned} \tag{2.111}$$

$$\leq 2\mathbb{E}\left[|x-x'|^2 + T\sup_{s\in[t,T]}\int_t^s |\widehat{\alpha}_u^{t,x} - \widehat{\alpha}_u^{t,x'}|^2 du\middle|\overline{\mathcal{F}}_t\right]$$
$$\leq C(T)\left[|x-x'|^2 + \mathbb{E}\left[\int_t^T |\widehat{\alpha}_s^{t,x} - \widehat{\alpha}_s^{t,x'}|^2 ds\middle|\overline{\mathcal{F}}_t\right]\right]$$

We apply now condition (2.104), that can be proved for the couple of input processes $(x,(b_s)_{s\in[t,T]},(w_s)_{s\in[t,T]},(\varpi_s)_{s\in[t,T]})$ and $(x,(b_s)_{s\in[t,T]},(w_s)_{s\in[t,T]},(\varpi_s)_{s\in[t,T]})$, to (2.111). Hence, we have that

$$\mathbb{E}\left[\sup_{s\in[t,T]}|Y_s^{t,x} - Y_s^{t,x'}|^2\middle|\overline{\mathcal{F}}_t\right] \leq C_C(T,L)\mathbb{E}\Big[\sup_s |X_s^{t,x} - X_s^{t,x'}|^2\Big|\overline{\mathcal{F}}_t\Big]$$
$$\leq \overbrace{C(T)C_C(T,L)}^{=:C_D(T,L)}\left[|x-x'|^2 + \mathbb{E}\left[\int_t^T |\widehat{\alpha}_s^{t,x} - \widehat{\alpha}_s^{t,x'}|^2 ds\middle|\overline{\mathcal{F}}_t\right]\right].$$

Therefore, $\overline{\mathbb{P}}$-a.s., by (2.110), the following holds

$$\begin{aligned}|Y_t^{t,x} - Y_t^{t,x'}|^2 &\leq \mathbb{E}\Big[\sup_{s\in[t,T]}|Y_s^{t,x} - Y_s^{t,x'}|^2|\overline{\mathcal{F}}_t\Big]\\ &\leq C_D(T,L)[|x-x'|^2 + L(Y_t^{t,x'} - Y_t^{t,x})(x'-x)]\\ &= \max\Big\{C_D(T,L);\frac{1}{2}L\Big\}\Big[|x-x'|^2 + 2(Y_t^{t,x'} - Y_t^{t,x})(x'-x)\Big].\end{aligned} \tag{2.112}$$

Using the following notation:

$$\begin{cases} a := |Y_t^{t,x} - Y_t^{t,x'}|,\\ b := |x-x'|,\\ c := \max\Big\{C_D(T,L);\frac{1}{2}L\Big\} > 0,\end{cases}$$

equation (2.112) reduces to $|a|^2 \leq c(|b|^2 + 2|a||b|)$. From this we conclude that $|a|^2(1+c) \leq c(|a|+|b|)^2$. Passing to the square root:

$$|a|\sqrt{1+c} \leq \sqrt{c}(|a|+|b|) \Rightarrow |a| \leq \frac{\sqrt{c}}{\sqrt{1+c}-\sqrt{c}}|b|, \quad \overline{\mathbb{P}}\text{-a.s.}$$

that is the statement of Assumption 2.41, with $\Gamma_0 := \frac{\sqrt{c}}{\sqrt{1+c}-\sqrt{c}}$. □

2.B Convergence analysis

Lemma 2.41. *Let us consider a process $\Theta^n := (\Psi^n, b, w)$, defined on a probability space $(\Omega,\mathcal{F},\mathbb{P})$. Assume that Ψ^n takes values on a Polish space $\mathcal{X}$ for each $n\in\mathbb{N}$ and b,w are Brownian motions with respect to the filtration generated by Θ^n for each $n\in\mathbb{N}$. If the sequence $(\Theta^n)_n$ is convergent in distribution on $\mathcal{X}\times\mathcal{C}([0,T];\mathbb{R}^2)$ to a process $\Theta^\infty := (\Psi^\infty, b^\infty, w^\infty)$, defined on a probability space*

$(\Omega^\infty, \mathcal{F}^\infty, \mathbb{P}^\infty)$*, then,* (b^∞, w^∞) *is a two-dimensional Brownian motion with respect to the filtration generated by* Θ^∞*.*

Proof. Applying [KS98, Theorem 3.3.16], it is sufficient to prove that b^∞ and w^∞ are continuous local martingales with respect to $\mathbb{F}^\infty := \mathbb{F}^{\Theta^\infty}$ and the covariations are given by

$$\langle b^\infty_t, b^\infty_t\rangle = t = \langle w^\infty_t, w^\infty_t\rangle, \quad \langle b^\infty_t, w^\infty_t\rangle = 0, \quad \forall t \in [0,T].$$

Step 1 show that b^∞ and w^∞ are local martingales. We focus only on b^∞ (the proof for w^∞ is the fully analogous). It is equivalent to show that:

$$\mathbb{E}^\infty[g(\Theta^\infty_{\cdot\wedge s})(b^\infty_t - b^\infty_s)] = 0, \quad \forall s \le t, \quad \forall g \in \mathcal{C}^0_b(\mathcal{X} \times \mathcal{C}([0,T];\mathbb{R}^2);\mathbb{R}). \tag{2.113}$$

To prove (2.113), we approximate the function g and the difference between the (b, w) at t and s with a bounded and continuous function of Θ^∞. The main issues we need to address are:

- possible discontinuity of the function $x \to x_{\cdot\wedge s}$ in $\mathcal{X}$;
- unboundedness of $f(x) = x_t - x_s$ defined for $x \in \mathcal{C}([0,T],\mathbb{R})$.

First, let us notice that the stopping map $x \to x_{\cdot\wedge s}$ could be not continuous in $\mathcal{X}$, thus a priori we could not have a continuous function of Θ^∞. Thus, denoting $\Theta^{\infty,b} := (\Psi^\infty, w^\infty)$ and $\Theta^{n,b} := (\Psi^n, w)$, condition (2.113) is equivalent to say that:

$$\mathbb{E}^\infty[g(\Theta^{\infty,b}_{\cdot\wedge s}, b^\infty_{\cdot\wedge s})(b^\infty_t - b^\infty_s)] = 0, \quad \forall s \le t,$$

$\forall g : (\mathcal{X} \times \mathcal{C}([0,T];\mathbb{R})) \times \mathcal{C}([0,T];\mathbb{R}) \to \mathbb{R}$ such that:

C1 $g(\cdot, b) : \mathcal{X} \times \mathcal{C}([0,T],\mathbb{R}) \to \mathbb{R}$ is measurable and bounded for each $b \in \mathcal{C}([0,T];\mathbb{R})$;

C2 $g(\Theta^b, \cdot) : \mathcal{C}([0,T],\mathbb{R}) \to \mathbb{R}$ is continuous and bounded for almost every $\Theta^b \in \mathcal{X} \times \mathcal{C}([0,T];\mathbb{R})$.

By [Bil99, Theorem 2.8], since $\mathcal{X} \times \mathcal{C}([0,T],\mathbb{R})$ is separable, the weak convergence of law$(\Theta^{n,b}, b)$ to law$(\Theta^{\infty,b}, b^\infty)$ is equivalent to the convergence of:

$$\begin{cases} \text{law}(\Theta^{n,b}) & \Rightarrow \text{law}(\Theta^{\infty,b}), \\ \text{law}(b) & \Rightarrow \text{law}(b^\infty). \end{cases}$$

In particular, this implies that

$$\text{law}(b^\infty)(B) = \lim_{n\to\infty} \text{law}(b)(B) = \text{law}(b)(B), \quad \forall B \in \mathcal{B}(\mathcal{C}([0,T],\mathbb{R})).$$

And thus, the marginal in $\mathcal{C}([0,T],\mathbb{R})$ is the same. As a consequence:

$$\text{law}(\Theta^{n,b}, b), \text{law}(\Theta^{\infty,b}, b^\infty) \in \{P \in \mathcal{P}(\mathcal{X} \times \mathcal{C}([0,T],\mathbb{R})) : \; P(\cdot \times \mathcal{C}([0,T],\mathbb{R})) = \text{law}(b)\}.$$

We can apply [BL20, Lemma 2.1], to conclude that $\lim_{n\to\infty} \mathbb{E}[h(\Theta^{n,b}, b)] = \mathbb{E}^\infty[h(\Theta^{\infty,b}, b^\infty)]$, for each function h satisfying conditions C1 and C2. To prove condition (2.113), we introduce the function, for each $k \in \mathbb{N}$:

$$\begin{aligned} h_k &: \mathbb{R} \to \mathbb{R} \\ & x \mapsto \begin{cases} x & \text{if } |x| \le k \\ k \cdot \operatorname{sgn}(x) & \text{if } |x| \ge k. \end{cases} \end{aligned} \tag{2.114}$$

We consider now

$$\begin{aligned} \Big|\mathbb{E}^\infty[g(\Theta^\infty_{\cdot\wedge s})(b^\infty_t - b^\infty_s)]\Big| \le & \Big|\mathbb{E}^\infty[g(\Theta^\infty_{\cdot\wedge s})h_k(b^\infty_t - b^\infty_s)]\Big| \\ & + \Big|\mathbb{E}^\infty\Big[g(\Theta^\infty_{\cdot\wedge s})[h_k(b^\infty_t - b^\infty_s) - (b^\infty_t - b^\infty_s)]\Big]\Big| \end{aligned}$$

Since $g(\Theta^\infty_{\cdot\wedge s})h_k(b^\infty_t - b^\infty_s)$ is a bounded function of $\mathcal{X} \times \mathcal{C}([0,T], \mathbb{R})$ satisfying conditions C1 and C2 for each $k \in \mathbb{N}$, we can apply [BL20, Lemma 2.1] to conclude that

$$\mathbb{E}^\infty[g(\Theta^\infty_{\cdot\wedge s})h_k(b^\infty_t - b^\infty_s)] = \lim_{n\to\infty} \mathbb{E}[g(\Theta^n_{\cdot\wedge s})h_k(b_t - b_s)] = 0,$$

because b a Brownian motion with respect to the filtration generated by Θ^n. We can focus on $\mathbb{E}^\infty\Big[g(\Theta^\infty_{\cdot\wedge s})[h_k(b^\infty_t - b^\infty_s) - (b^\infty_t - b^\infty_s)]\Big]$. We introduce the event:

$$A^k_{t,s} := \{|b^\infty_t - b^\infty_s| \ge k\}, \quad k \in \mathbb{N}.$$

If $g(\Theta) \le L_g$ for each Θ, apply the Cauchy-Schwartz inequality and the Markov inequality to

$$\begin{aligned} \Big|\mathbb{E}^\infty\Big[g(\Theta^\infty_{\cdot\wedge s})[h_k(b^\infty_t - b^\infty_s) - (b^\infty_t - b^\infty_s)]\Big]\Big| &\le \mathbb{E}^\infty\Big[\Big|g(\Theta^\infty_{\cdot\wedge s})[h_k(b^\infty_t - b^\infty_s) - (b^\infty_t - b^\infty_s)]\Big|\Big] \\ &\le \mathbb{E}^\infty\Big[L_g(|b^\infty_t - b^\infty_s| - k)\mathbb{1}_{A^k_{t,s}}\Big] \le \mathbb{E}^\infty\Big[L_g|b^\infty_t - b^\infty_s|\mathbb{1}_{A^k_{t,s}}\Big] \\ &\le L_g\mathbb{E}^\infty[|b^\infty_t - b^\infty_s|]\mathbb{P}^\infty(A^k_{t,s}) \le L_g\frac{1}{k}\mathbb{E}^\infty[|b^\infty_t - b^\infty_s|]^2 \end{aligned} \tag{2.115}$$

It suffices to prove that $\mathbb{E}^\infty[|b^\infty_t - b^\infty_s|] < \infty$. To do so, we notice that, $(|b_t - b_s|)_n$ converges in distribution on $\mathbb{R}$ to $|b^\infty_t - b^\infty_s|$, for each t and s. Hence, in order to have the convergence of expectation we must check uniform integrability of $(|b_t - b_s|)_n$, i.e.:

$$\lim_{a\to\infty} \sup_{n\in\mathbb{N}} \mathbb{E}\Big[|b_t - b_s|\mathbb{1}_{A^a_{s,t}}\Big] = 0.$$

We notice that:

$$\sup_{n\in\mathbb{N}} \mathbb{E}\Big[|b_t - b_s|\mathbb{1}_{A^a_{s,t}}\Big] \le \frac{1}{a}(\mathbb{E}[|b_t - b_s|])^2 = \frac{1}{a}\Big(\frac{1}{\sqrt{2\pi}|t-s|}\int_{-\infty}^{\infty} |x|e^{-\frac{x^2}{2(t-s)^2}}\,dx\Big)^2 = \frac{2}{\pi a}(t-s)^2.$$

Then, $\lim_{a\to\infty} \sup_{n\in\mathbb{N}} \mathbb{E}[|b_t - b_s|\mathbb{1}_{A^a_{s,t}}] = 0$. Once we proved the convergence of expectations, we conclude that the process b^∞ is a $\mathbb{F}^{\Theta^\infty}$-martingale, because $\mathbb{E}^\infty[|b^\infty_t - b^\infty_s|] = \lim_{n\to\infty} \mathbb{E}[|b_t - b_s|] =$

$\sqrt{\frac{2}{\pi}}|t-s| < \infty$.
Step 2: In order to prove that $\langle b^\infty, b^\infty\rangle_t = t$, we consider the following result (see [KS98, Exercise 3.38]). Let X be a continuous local martingale and A a continuous, increasing process such $X_0 = A_0$, a.s.. If $Z := \exp\left(X - \frac{1}{2}A\right)$ is a local martingale, then $\langle X\rangle = A$.

Indeed, applying Itô-formula to Z:

$$dZ_t = Z_t(dX_t - \frac{1}{2}dA_t) + \frac{1}{2}Z_t(d\langle X\rangle_t - 0) = Z_t\Big(dX_t - \frac{1}{2}dA_t + \frac{1}{2}d\langle X\rangle_t\Big).$$

As a consequence, the finite variation of Z is given by $\frac{1}{2}Z_t(-dA_t + d\langle X\rangle_t)$, because X is a local martingale. By assumption Z is a local martingale, the finite-variation is zero. Thus, since $X_0 = A_0 = 0$, we conclude that $-dA_t + d\langle X\rangle_t = 0$ implies that $\langle X\rangle_t = A_t, \quad t \in [0,T]$. Since we have already proved that b^∞ and w^∞ are martingales on $(\Omega^\infty, \mathcal{F}^\infty, \mathbb{P}^\infty, \mathbb{F}^\infty)$, we need to prove that $Z_t^\infty := \exp\left(b_t^\infty - \frac{1}{2}t\right)$ is a local martingale on $(\Omega^\infty, \mathcal{F}^\infty, \mathbb{P}^\infty, \mathbb{F}^\infty)$. To do so, it is sufficient to show that

$$\mathbb{E}^\infty[g(\Theta^{\infty,b}_{\cdot\wedge s}, b^\infty_{\cdot\wedge s})(Z_t^\infty - Z_s^\infty)] = 0, \quad \forall s \le t,$$

$\forall g : (\mathcal{X} \times \mathcal{C}([0,T];\mathbb{R})) \times \mathcal{C}([0,T];\mathbb{R}) \to \mathbb{R}$ satisfying E1 and E2. We notice that:

$$\mathbb{E}^\infty[g(\Theta^{\infty,b}_{\cdot\wedge s}, b^\infty_{\cdot\wedge s})(Z_t^\infty - Z_s^\infty)] = \mathbb{E}^\infty\Big[g(\Theta^{\infty,b}_{\cdot\wedge s}, b^\infty_{\cdot\wedge s})\Big(\exp\Big(b_t^\infty - \frac{1}{2}t\Big) - \exp\Big(b_s^\infty - \frac{1}{2}s\Big)\Big)\Big].$$

As in *Step 1*, we consider a continuous, but unbounded function of b_t^∞ and b_s^∞. Thus, we apply the function h_k introduced in equation (2.114):

$$\begin{aligned}\Big|\mathbb{E}^\infty[g(\Theta^\infty_{\cdot\wedge s})(Z_t^\infty - Z_s^\infty)]\Big| \le &\Big|\mathbb{E}^\infty[g(\Theta^\infty_{\cdot\wedge s})h_k(Z_t^\infty - Z_s^\infty)]\Big| \\ &+ \Big|\mathbb{E}^\infty\Big[g(\Theta^\infty_{\cdot\wedge s})[h_k(Z_t^\infty - Z_s^\infty) - (Z_t^\infty - Z_s^\infty)]\Big]\Big|\end{aligned}$$

For the first term, we can pass to the limit:

$$\Big|\mathbb{E}^\infty[g(\Theta^\infty_{\cdot\wedge s})h_k(Z_t^\infty - Z_s^\infty)]\Big| = \lim_{n\to\infty}\Big|\mathbb{E}^\infty[g(\Theta^n_{\cdot\wedge s})h_k(Z_t^1 - Z_s^1)]\Big| = 0,$$

where $Z_t^1 = \exp\left(b_t - \frac{1}{2}t\right)$, and the second equivalence holds because b is a $\mathbb{F}^{\Theta^n}$-Brownian motion. To handle the second term, it is sufficient to prove that $\mathbb{E}^\infty[|Z_t^1 - Z_s^1|] < \infty$ as in (2.115) for $\mathbb{E}^\infty[|b_t^{1,\infty} - b_s^{1,\infty}|] < \infty$. We notice that $\mathbb{E}[|Z_t^1 - Z_s^1|] \le \mathbb{E}[Z_t^1] + \mathbb{E}[Z_s^1] = 2$, then, uniform integrability holds and we can conclude that

$$\lim_{k\to\infty}\Big|\mathbb{E}^\infty\Big[g(\Theta^\infty_{\cdot\wedge s})[h_k(Z_t^\infty - Z_s^\infty) - (Z_t^\infty - Z_s^\infty)]\Big]\Big| = 0.$$

We have proved that $\left(\exp\left(b_t^\infty - \frac{1}{2}t\right)\right)_t$ is a $\mathbb{F}^{\Theta^\infty}$-martingale. In the same way, we can prove that $\left(\exp\left(w_t^\infty - \frac{1}{2}t\right)\right)_t$ is a $\mathbb{F}^{\Theta^\infty}$-martingale.
Step 3: Finally, we need to prove that $\langle b^\infty, w^\infty\rangle_t = 0$ for every $t \in [0,T]$. By [KS98, Theorem

1.5.13], this is equivalent to show that $(b_t^\infty w_t^\infty)_{t\in[0,T]}$ is a martingale with respect to $\sigma\{\Theta^\infty\}$. The proof can be done similarly to the previous steps. Thus, to conclude it is sufficient to prove that $\mathbb{E}[|w_t b_t - w_s b_s|] < \infty$:

$$\mathbb{E}[|w_t b_t - w_s b_s|] \le \mathbb{E}[|w_t|]\mathbb{E}[|b_t|] + \mathbb{E}[|w_s|]\mathbb{E}[|b_s|] = \frac{4}{\pi}|t-s| < \infty.$$

□

2.C Proof of Lemma 2.25

Step 1 We recall that $X^{n,l}$ is defined on the canonical space $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}}, \overline{\mathbb{F}})$ as follows:

$$X^{n,l} = X^0 + \int_0^t (-\overline{\Lambda}(Y_s^{n,l} + \varpi_s^{n,l}) + l(s, \varpi_s^{n,l}))ds + \int_0^t \sigma^0(s, \varpi_s^{n,l})db_s + \int_0^t \sigma(s, \varpi_s^{n,l})dw_s.$$

By Assumption B2, we have that

$$\begin{aligned} |X_t^{n,l}|^2 &\le 4\Big[|X^0|^2 + t\int_0^t |\overline{\Lambda}(Y_s^{n,l} + \varpi_s^{n,l}) + l(s, \varpi_s^{n,l})|^2 ds + \Big|\int_0^t \sigma^0(s, \varpi_s^{n,l})db_s\Big|^2 \\ &\quad + \Big|\int_0^t \sigma(s, \varpi_s^{n,l})dw_s\Big|^2\Big] \\ &\le 4\Big[|X^0|^2 + 3tL^2\int_0^t (1 + |Y_s^{n,l}|^2 + |\varpi_s^{n,l}|^2)ds + \Big|\int_0^t \sigma^0(s, \varpi_s^{n,l})db_s\Big|^2 \\ &\quad + \Big|\int_0^t \sigma(s, \varpi_s^{n,l})dw_s\Big|^2\Big]. \end{aligned}$$

Passing to the supremum:

$$\begin{aligned} \sup_{t\in[0,T]} |X_t^{n,l}|^2 &\le 4\Big[|X^0|^2 + 3TL^2\int_0^T (1 + |Y_s^{n,l}|^2 + |\varpi_s^{n,l}|^2)ds + \sup_{t\in[0,T]}\Big|\int_0^t \sigma^0(s, \varpi_s^{n,l})db_s\Big|^2 \\ &\quad + \sup_{t\in[0,T]}\Big|\int_0^t \sigma(s, \varpi_s^{n,l})dw_s\Big|^2\Big]. \end{aligned}$$

We take the expectation and apply Fubini's Theorem and Doob's inequality. Hence, by Assumption A1, the following holds

$$\begin{aligned} \overline{\mathbb{E}}\Big[\sup_{t\in[0,T]} |X_t^{n,l}|^2\Big] &\le 4\Big[\overline{\mathbb{E}}[|X^0|^2] + 3TL^2\int_0^T (1 + \overline{\mathbb{E}}[|Y_s^{n,l}|^2] + \overline{\mathbb{E}}[|\varpi_s^{n,l}|^2])ds \\ &\quad + \overline{\mathbb{E}}\Big[\sup_{t\in[0,T]}\Big|\int_0^t \sigma^0(s, \varpi_s^{n,l})db_s\Big|^2\Big] + \overline{\mathbb{E}}\Big[\sup_{t\in[0,T]}\Big|\int_0^t \sigma(s, \varpi_s^{n,l})dw_s\Big|^2\Big]\Big] \\ &\le 4\Big[\overline{\mathbb{E}}[|X^0|^2] + 3TL^2\int_0^T (1 + \overline{\mathbb{E}}[|Y_s^{n,l}|^2] + \overline{\mathbb{E}}[|\varpi_s^{n,l}|^2])ds \end{aligned}$$

$$
\begin{aligned}
&+ 4\overline{\mathbb{E}}\left[\int_0^T |\sigma^0(s, \varpi_s^{n,l})|^2 ds\right] + 4\overline{\mathbb{E}}\left[\int_0^T |\sigma(s, \varpi_s^{n,l})|^2 ds\right]\Bigg] \\
\leq\ & 4\Bigg[\overline{\mathbb{E}}[|X^0|^2] + 3TL^2 \int_0^T (1 + \overline{\mathbb{E}}[|Y_s^{n,l}|^2] + \overline{\mathbb{E}}[|\varpi_s^{n,l}|^2]) ds \\
&+ 8L^2\overline{\mathbb{E}}\left[\int_0^t (1 + |\varpi_s^{n,l}|^2) ds\right] + 8L^2\overline{\mathbb{E}}\left[\int_0^T (1 + |\varpi_s^{n,l}|^2) ds\right]\Bigg] \\
\leq\ & 4\Bigg[\overline{\mathbb{E}}[|X^0|^2] + TL^2(3T + 16) + 3TL^2 \int_0^T \overline{\mathbb{E}}[|Y_s^{n,l}|^2] ds \\
&+ L^2(3T + 16)\left[\int_0^T \overline{\mathbb{E}}[|\varpi_s^{n,l}|^2] ds\right]\Bigg] \\
\leq\ & 4\Bigg[\overline{\mathbb{E}}[|X^0|^2] + TL^2(3T + 16) + 3TL^2 \int_0^T \overline{\mathbb{E}}[|Y_s^{n,l}|^2] ds \\
&+ L^2(3T + 16)\left[\sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \overline{\mathbb{E}}[|\varpi_s^{n,l}|^2] ds\right]\Bigg] \\
=\ & 4\Bigg[\overline{\mathbb{E}}[|X^0|^2] + TL^2(3T + 16) + 3TL^2 \int_0^T \overline{\mathbb{E}}[|Y_s^{n,l}|^2] ds \\
&+ L^2(3T + 16)\left[\sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \overline{\mathbb{E}}[| - (\overline{\Lambda} + \overline{\Lambda}^0)^{-1}\overline{\mathbb{E}}[\overline{\Lambda} Y_s^{n,l} + \overline{\Lambda}^0 Y_s^{0,n,l} | \overline{V}_i]|^2] ds\right]\Bigg]
\end{aligned}
$$

We notice that:

$$
\overline{\mathbb{E}}[| - (\overline{\Lambda} + \overline{\Lambda}^0)^{-1}\overline{\mathbb{E}}[\overline{\Lambda} Y_s^{n,l} + \overline{\Lambda}^0 Y_s^{0,n,l} | \overline{V}_i]|^2] \leq 2((\overline{\Lambda} + \overline{\Lambda}^0)^{-1})^2[(\overline{\Lambda})^2\overline{\mathbb{E}}[|Y_s^{n,l}|^2] + (\overline{\Lambda}^0)^2\overline{\mathbb{E}}[|Y_s^{0,n,l}|^2]].
$$

As a consequence, denoting $C_1 := (\overline{\Lambda} + \overline{\Lambda}^0)^{-1}\overline{\Lambda}$, we apply the same reasoning described in (2.49), to conclude that

$$
\begin{aligned}
\overline{\mathbb{E}}\left[\sup_{t\in[0,T]} |X_t^{n,l}|^2\right] \leq\ & 4\Bigg[\overline{\mathbb{E}}[|X^0|^2] + TL^2(3T + 16) + 3TL^2 \int_0^T \overline{\mathbb{E}}[|Y_s^{n,l}|^2] ds \\
&+ L^2(3T + 16)\Big[\sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} (C_1^2\overline{\mathbb{E}}[|Y_s^{n,l}|^2] + (1 - C_1)^2\overline{\mathbb{E}}[|Y_s^{0,n,l}|^2]) ds\Bigg]\Bigg] \\
\leq\ & 4\Bigg[\overline{\mathbb{E}}[|X^0|^2] + TL^2(3T + 16) + [3TL^2 + C_1^2(3TL^2 + 16L^2) \int_0^T \overline{\mathbb{E}}[|Y_s^{n,l}|^2] ds \\
&+ (3T + 16)L^2(C_1 - 1)^2 \int_0^T \overline{\mathbb{E}}[|Y_s^{0,n,l}|^2] ds\Bigg]
\end{aligned}
$$

$$\leq 4\Bigg[\overline{\mathbb{E}}[|X^0|^2] + TL^2(3T+16) + 2L^2(T^2+1)T[3TL^2 + C_1^2(3TL^2+16L^2)] \\ + 2L^2(T^2+1)T(3T+16)L^2(1-C_1)^2\Bigg] =: \widetilde{C}.$$

Step 2: We consider a $\mathbb{F}^{X^{n,l}}$-stopping time τ and a constant $\delta > 0$.

$$\overline{\mathbb{E}}[|X^{n,l}_{(\tau+\delta)\wedge T} - X^{n,l}_{\tau}|] \leq \overline{\mathbb{E}}\Bigg[\Bigg|\int_\tau^{(\tau+\delta)\wedge T} (-\overline{\Lambda}(Y^{n,l}_s + \varpi^{n,l}_s) + l(s,\varpi^{n,l}_s))ds\Bigg| \\ + \Bigg|\int_\tau^{(\tau+\delta)\wedge T} \sigma^0(s,\varpi^{n,l}_s)db_s\Bigg| + \int_\tau^{(\tau+\delta)\wedge T} \sigma(s,\varpi^{n,l}_s)dw_s\Bigg|\Bigg].$$

We focus on the three terms separately:

$$\begin{aligned}(A) &:= \overline{\mathbb{E}}\Bigg[\Bigg|\int_\tau^{(\tau+\delta)\wedge T} (-\overline{\Lambda}(Y^{n,l}_s + \varpi^{n,l}_s) + l(s,\varpi^{n,l}_s))ds\Bigg|\Bigg] \\ &= \overline{\mathbb{E}}\Bigg[\Bigg|\int_0^T \mathbb{1}_{[\tau,\tau+\delta)\wedge T]}(s)(-\overline{\Lambda}(Y^{n,l}_s + \varpi^{n,l}_s) + l(s,\varpi^{n,l}_s))ds\Bigg|\Bigg] \\ &\leq \overline{\mathbb{E}}\Bigg[\Bigg|\int_0^T \mathbb{1}_{[\tau,\tau+\delta)\wedge T]}(s)L(1+|Y^{n,l}_s| + |\varpi^{n,l}_s|)ds\Bigg|\Bigg] \\ &\leq L\overline{\mathbb{E}}\Bigg[\int_0^T \mathbb{1}_{[\tau,\tau+\delta)\wedge T]}(s)ds + \sup_{s\in[0,T]}(|Y^{n,l}_s| + |\varpi^{n,l}_s|)\Bigg[\int_0^T \mathbb{1}_{[\tau,\tau+\delta)\wedge T]}(s)ds\Bigg] \\ &= L\delta\overline{\mathbb{E}}[\sup_{s\in[0,T]}(|Y^{n,l}_s| + |\varpi^{n,l}_s|)].\end{aligned}$$

Applying the same reasoning of (2.49), we observe that

$$\overline{\mathbb{E}}\Bigg[\sup_{s\in[0,T]}|Y^{n,l}_s|\Bigg] \leq \Bigg(\overline{\mathbb{E}}\Bigg[\sup_{s\in[0,T]}|Y^{n,l}_s|^2\Bigg] \leq \sqrt{2L^2(T^2+1)}\Bigg)^{\frac{1}{2}}.$$

For what regards $\varpi^{n,l}$, let us notice that:

$$\overline{\mathbb{E}}\Bigg[\sup_{s\in[0,T]}|\varpi^{n,l}_s|\Bigg] = \overline{\mathbb{E}}\Bigg[\max_{i=0,\dots,N-1}\sup_{s\in[t_i,t_{i+1})}|\overline{\mathbb{E}}[C_1 Y^{n,l}_s + (1-C_1)Y^{0,n,l}_s|\overline{V}_i]|\Bigg] \\ \leq \max_{i=0,\dots,N-1}\Bigg(\overline{\mathbb{E}}\Bigg[\sup_{s\in[t_i,t_{i+1})} C_1|\overline{\mathbb{E}}[Y^{n,l}_s|\overline{V}_i]| + (1-C_1)\sup_{s\in[t_i,t_{i+1})}|\overline{\mathbb{E}}[Y^{0,n,l}_s|\overline{V}_i]|\Bigg]\Bigg).$$

By Assumption B2, for each $s \in [t_i, t_{i+1})$, the following holds

$$\begin{aligned}
|\overline{\mathbb{E}}[Y_s^{n,l}|\overline{V}_i]| &= \bigg|\overline{\mathbb{E}}\bigg[\partial_x g(X_T^{n,l}, \varpi_T^{n,l}) + \int_s^T \partial_x \overline{f}(u, X_u^{n,l}, \varpi_u^{n,l}) du \\
&\quad + \int_s^T Z_u^{0,n,l} db_u + \int_s^T Z_u^{n,l} dw_u |\overline{V}_i\bigg]\bigg| \\
&= \overline{\mathbb{E}}\bigg[|\partial_x g(X_T^{n,l}, \varpi_T^{n,l})| + \int_s^T |\partial_x \overline{f}(u, X_u^{n,l}, \varpi_u^{n,l})| du |\overline{V}_i\bigg] \leq L(T+1).
\end{aligned}$$

Analogously, by Assumption B4:

$$\begin{aligned}
|\overline{\mathbb{E}}[Y_s^{0,n,l}|\overline{V}_i]| &:= \bigg|\overline{\mathbb{E}}\bigg[g_1^0(\varpi_s^{n,l}) + \int_t^T c_0^M(t, \varpi_s^{n,l}) ds - \int_t^T Z_s^{0,0,n,l} db_s |\overline{V}_i\bigg]\bigg| \\
&\leq \overline{\mathbb{E}}\bigg[|g_1^0(\varpi_s^{n,l})| + (T-t)\int_t^T |c_0^M(t, \varpi_s^{n,l})| ds |\overline{V}_i\bigg] \leq L(T+1).
\end{aligned}$$

Therefore,

$$\overline{\mathbb{E}}\bigg[\sup_{s\in[0,T]} |\varpi_s^{n,l}|\bigg] \leq C_1 L(T+1) + (1-C_1)L(T+1) = L(T+1) \tag{2.116}$$

In conclusion:

$$\overline{\mathbb{E}}\bigg[\bigg|\int_\tau^{(\tau+\delta)\wedge T} (-\overline{\Lambda}(Y_s^{n,l} + \varpi_s^{n,l}) + l(s, \varpi_s^{n,l})) ds\bigg|\bigg] \leq L^2\delta(\sqrt{2(T+1)} + (T+1)).$$

For the volatility terms, applying Jensen's inequality together with Itô's isometry, by Assumption A1, the following holds

$$\begin{aligned}
\overline{\mathbb{E}}\bigg[\bigg|\int_\tau^{(\tau+\delta)\wedge T} \sigma^0(s, \varpi_s^{n,l}) db_s\bigg|\bigg] &= \overline{\mathbb{E}}\bigg[\bigg|\int_0^T \mathbb{1}_{[\tau,(\tau+\delta)\wedge T]}(s)\sigma^0(s, \varpi_s^{n,l}) db_s\bigg|\bigg] \\
&\leq \bigg(\overline{\mathbb{E}}\bigg[\bigg|\int_0^T \mathbb{1}_{[\tau,(\tau+\delta)\wedge T]}(s)\sigma^0(s, \varpi_s^{n,l}) db_s\bigg|^2\bigg]\bigg)^{\frac{1}{2}} \\
&= \bigg(\overline{\mathbb{E}}\bigg[\int_0^T \mathbb{1}_{[\tau,(\tau+\delta)\wedge T]}(s)|\sigma^0(s, \varpi_s^{n,l})|^2 ds\bigg]\bigg)^{\frac{1}{2}} \\
&\leq \bigg(\overline{\mathbb{E}}\bigg[\int_0^T \mathbb{1}_{[\tau,(\tau+\delta)\wedge T]}(s) 2L^2(1 + |\varpi_s^{n,l}|^2) ds\bigg]\bigg)^{\frac{1}{2}} \\
&\leq \sqrt{2}L\bigg(\overline{\mathbb{E}}\Big[\int_0^T \mathbb{1}_{[\tau,(\tau+\delta)\wedge T]}(s) ds + \int_0^T \mathbb{1}_{[\tau,(\tau+\delta)\wedge T]}(s) \sup_{t\in[0,T]} |\varpi_t^{n,l}|^2 ds\Big]\bigg)^{\frac{1}{2}} \\
&= \sqrt{2}L\bigg(\delta + \overline{\mathbb{E}}\bigg[\sup_{t\in[0,T]} |\varpi_t^{n,l}|^2\bigg]\delta\bigg)^{\frac{1}{2}} \leq \sqrt{2}L\sqrt{1 + 2L^2(T+1)}\sqrt{\delta}.
\end{aligned}$$

Hence, again by Jensen's inequality together with (2.49), we conclude that

$$\begin{aligned}\overline{\mathbb{E}}\left[\sup_{t\in[0,T]}|\varpi_t^{n,l}|^2\right] &\leq \overline{\mathbb{E}}\left[\max_{i=0,\dots,N-1}\sup_{s\in[t_i,t_{i+1})}|(\overline{\Lambda}+\overline{\Lambda}^0)^{-1}\overline{\mathbb{E}}[\overline{\Lambda}Y_s^{n,l}+\overline{\Lambda}^0Y_s^{0,n,l}|\overline{V}_i]|^2\right]\\ &\leq \overline{\mathbb{E}}\left[\max_{i=0,\dots,N-1}\sup_{s\in[t_i,t_{i+1})}(C_1^2\overline{\mathbb{E}}[|Y_s^{n,l}|^2|\overline{V}_i]+(1-C_1^2)\overline{\mathbb{E}}[|Y_s^{0,n,l}|^2|\overline{V}_i])\right]\\ &\leq (2L^2(T^2+1)).\end{aligned}$$

The computation for $\overline{\mathbb{E}}[|\int_\tau^{(\tau+\delta)\wedge T}\sigma(s,\varpi_s^{n,l})dw_s|]$ is exactly the same. In conclusion:

$$\overline{\mathbb{E}}[|X_{(\tau+\delta)\wedge T}^{n,l}-X_\tau^{n,l}|]\leq L^2\delta(\sqrt{2(T+1)}+(T+1))+2\sqrt{2}L\sqrt{1+2L^2(T+1)}\sqrt{\delta}\leq C\sqrt{\delta},$$

where $C:=L^2(\sqrt{2(T+1)}+(T+1))+2\sqrt{2}L\sqrt{1+2L^2(T+1)}$ and $\delta<1$.

2.D Proof of Proposition 2.28

Proposition 2.28 guarantees the stability of weak equilibria of a sequence of solutions to a family of optimal control problems. In particular, the weak limit of a sequence of solutions to optimal control problems defined in an approximate setting is still solution to an optimal control problem. To prove this result we based on [CD18b, Proposition 3.11]. Before proving Proposition 2.28, we rely on some preliminary results that we are going to apply in the proof.

Theorem 2.42. *([KP91, Theorem 4.7]) For each $n\in\mathbb{N}$, let (ξ^n,η^n) be a $\mathcal{F}^n$-adapted process with sample paths in $\mathcal{D}([0,\infty),\mathbb{R}^2)$. Suppose that:*

E1 *$(\eta^n)_{n\in\mathbb{N}}$ are continuous martingales for every $n\in\mathbb{N}$ such that there exists a sequence of stopping process $\{\tau_n^\alpha\}_{\alpha>0}$ such that $\mathbb{P}^n(\tau_n^\alpha\leq\alpha)\leq\frac{1}{\alpha}$ and*

$$\sup_n\mathbb{E}^n[[\eta^n](t\wedge\tau_n^\alpha)]<\infty. \tag{2.117}$$

E2 *$(\xi^n)_{n\in\mathbb{N}}$ are càdlàg -processes such that for each $\alpha>0$ there exists stopping times $\{\tau_n^\alpha\}$ with $\mathbb{P}^n(\tau_n^\alpha\leq\alpha)\leq\frac{1}{\alpha}$ and such that, for each $t\geq 0$, $\sup_n(\mathbb{E}^n[|\xi^n(t\wedge\tau_n^\alpha)|]+V_t(\xi^n(\cdot\wedge\tau_n^\alpha)))<\infty$, where V_t is defined in equation (2.64).*

If (ξ^n,η^n) converges to (ξ^∞,η^∞) in distribution on $\mathcal{M}([0,T];\mathbb{R})\times\mathcal{D}([0,T];\mathbb{R})$ and η^∞ is continuous, then ξ^∞ admits a version with sample paths in $\mathcal{D}([0,\infty),\mathbb{R})$, η^∞ is a semi-martingale with respect to a filtration to which ξ^∞ and η^∞ are adapted and $(\xi^n,\eta^n,\int\xi^n\eta^n)$ converges in distribution to $(\xi^\infty,\eta^\infty,\int\xi^\infty d\eta^\infty)$ on $\mathcal{M}([0,T];\mathbb{R})\times\mathcal{D}([0,T];\mathbb{R})\times\mathcal{D}([0,T];\mathbb{R})$.

We prove moreover the following result:

Lemma 2.43. *Let us a consider a real valued random variable γ, defined on a probability space*

$(\Omega, \mathcal{F}, \mathbb{P})$ *and measurable with respect to the completion of a sigma-algebra* $\mathcal{G}$*. Then, there another random variable* $\widetilde{\gamma}$*, such that* $\mathbb{P}(\widetilde{\gamma} = \gamma) = 1$ *and* $\widetilde{\gamma}$ *is* $\mathcal{G}$*-measurable.*

Proof. Since γ is measurable with respect to the completion of $\mathcal{G}$, the following holds

$$\forall a \in \mathbb{Q},\ \gamma^{-1}((-\infty, a)) = G_a \cup N_a,$$

where $G_a \in \mathcal{G}$ and $N_a \subset \Omega$ such that $\exists N_a^0 \in \mathcal{G}$, $\mathbb{P}(N_a^0) = 0$ and $N_a \subset N_a^0$. We now define $N := \bigcup_{a \in \mathbb{Q}} N_a^0$. In particular, $\mathbb{P}(N) = 0$. and we define:

$$\widetilde{\gamma}(\varpi) = \begin{cases} \gamma(\omega) & \text{if } \omega \notin N, \\ 0 & \text{if } \omega \in N. \end{cases}$$

Therefore $\mathbb{P}(\gamma \neq \widetilde{\gamma}) = \mathbb{P}(N) = 0$. Moreover:

$$\widetilde{\gamma}^{-1}((-\infty, a)) = \begin{cases} G_a \cup N, & a > 0, \\ G_a \cap (\Omega - N), & a \leq 0 \end{cases}$$

Therefore, $\widetilde{\gamma}$ is a $\mathcal{G}$-measurable version of γ. □

Proof of Proposition 2.28.

Step 1 First, we prove that the family $(\widehat{\alpha}^n)_n$, is tight in the Meyer-Zheng space $\mathcal{M}([0,T];\mathbb{R})$ equipped with its topology. Moreover, any weak limit can be seen as the law of a $\mathbb{R}$-valued process. By condition D1, (X^n, Y^n) forms a tight sequence in the product space $\mathcal{C}([0,T];\mathbb{R}) \times \mathcal{M}([0,T];\mathbb{R}^n)$, where Y^n is the adjoint process of X^n. Moreover, by condition D3, the sequence (Y^n, ϖ^n) is tight. Hence, for each $n \in \mathbb{N}$ the optimal control is defined by:

$$\widehat{\alpha}_t^n := \widehat{\alpha}(Y_t^n, \varpi_t^n) = -\overline{\Lambda}(Y_t^n + \varpi_t^n),$$

where $\widehat{\alpha}(y, \varpi) := -\overline{\Lambda}(y + \varpi)$, is a continuous function on $\mathbb{R}^2$. Let us notice that, by [CD18b, Lemma 3.5], if a function $\theta : [0,T] \times \mathbb{R}^h \to \mathbb{R}^k$ is continuous also the function:

$$\begin{aligned} \Theta : (\mathcal{M}([0,T];\mathbb{R}^h), \delta_{\mathcal{M}}) &\to (\mathcal{M}([0,T];\mathbb{R}^h), \delta_{\mathcal{M}}) \\ x &\mapsto (\theta(t, x_t))_{t \in [0,T]}. \end{aligned}$$

is continuous. As a consequence, applying the continuous mapping theorem, we conclude that if (Y^n, ϖ^n) converges weakly to a process $(Y^\infty, \varpi^\infty)$, then also $\widehat{\alpha}(Y_t^n, \varpi_t^n)$ converges weakly to $\widehat{\alpha}(Y_t^\infty, \varpi_t^\infty)$. In particular, defining $\widehat{\alpha}_t^\infty := \widehat{\alpha}(Y_t^\infty, \varpi_t^\infty)$ the sequence $(\widehat{\alpha}^n)_n$ converges weakly to $\widehat{\alpha}^\infty$ and the weak limit can be seen as an $\mathbb{R}$-valued process. Therefore, up to subsequences, the sequence $(\Theta^n)_{n \in \mathbb{N}}$ defined by:

$$\Theta^n := (X^0, b, \mathcal{W}^n, w, X^n, \widehat{\alpha}^n) \in \widetilde{\Omega}_{\text{input}} \times \mathcal{C}([0,T];\mathbb{R}) \times \mathcal{M}([0,T];\mathbb{R}) \tag{2.118}$$

admits a weak limit $\Theta^\infty := (X^\infty, b^\infty, \mathcal{W}^\infty, w^\infty, X^\infty, \widehat{\alpha}^\infty)$. We denote by $\mathbb{G}^\infty$ the complete and right continuous augmentation of $\mathbb{F}^{\Theta^\infty}$. Let us notice that, a priori, $\mathbb{G}^\infty$ can be strictly larger than $\mathbb{F}^\infty$.

Step 2 We need to prove that the weak limit X^∞ satisfies system (2.70) with $\gamma_t \equiv \widehat{\alpha}^\infty_t$ for each $t \in [0,T]$. Any weak limit of the initial condition X^0 is distributed like $X^{0,\infty}$. Moreover, applying Lemma 2.41, we show that (b^∞, w^∞) is a two-dimensional Brownian motion with respect to the filtration $\mathbb{G}^\infty$.

Since $\overline{\mathbb{E}}[\sup_{t\in[0,T]} |X^n_t|^2]$ and $\overline{\mathbb{E}}[\int_0^T |\widehat{\alpha}^n_t|^2 dt]$ are uniformly bounded in $n \in \mathbb{N}$, also:

$$\mathbb{E}^\infty\left[\sup_{t\in[0,T]} |X^\infty_t|^2 + \int_0^T |\widehat{\alpha}^\infty_t|^2 dt\right] < \infty.$$

Indeed, to prove the square integrability of the optimal control process $\widehat{\alpha}^\infty$ we can apply [CD18b, Lemma 3.6]. To do so, we need to show that:

$$\forall \epsilon > 0 \ \lim_{a\to\infty} \sup_{n\geq 0} \overline{\mathbb{P}}\left(\int_0^T |\widehat{\alpha}(Y^n_t, \varpi^n_t)|^2 \mathbb{1}_{\{|\widehat{\alpha}(Y^n_t,\varpi^n_t)|^2\geq a\}} dt \geq \epsilon\right) = 0. \tag{2.119}$$

Applying Markov's and the Cauchy-Schwartz inequalities, together with Fubini-Tonelli theorem, by condition D2, we observe that

$$\begin{aligned}
&\overline{\mathbb{P}}\left(\int_0^T |\widehat{\alpha}(Y^n_t, \varpi^n_t)|^2 \mathbb{1}_{\{|\widehat{\alpha}(Y^n_t,\varpi^n_t)|^2\geq a\}} dt \geq \epsilon\right)\\
&\qquad \leq \frac{1}{\epsilon}\overline{\mathbb{E}}\left[\int_0^T |\widehat{\alpha}(Y^n_t, \varpi^n_t)|^2 \mathbb{1}_{\{|\widehat{\alpha}(Y^n_t,\varpi^n_t)|^2\geq a\}} dt\right]\\
&\qquad \leq \frac{1}{\epsilon}\overline{\mathbb{E}}\left[2\int_0^T |\overline{\Lambda}|^2(|Y^n_t|^2 + |\varpi^n_t|^2) \mathbb{1}_{\{|\widehat{\alpha}(Y^n_t,\varpi^n_t)|\geq\sqrt{a}\}} dt\right]\\
&\qquad \leq \frac{4|\overline{\Lambda}|^2 T}{\epsilon}\int_0^T \overline{\mathbb{E}}\left[\sup_{t\in[0,T]}(|Y^n_t|^4 + |\varpi^n_t|^4)\right] dt \int_0^T \mathbb{P}(|\widehat{\alpha}(Y^n_t, \varpi^n_t)| \geq \sqrt{a}) dt\\
&\qquad \leq \frac{4|\overline{\Lambda}|^2 T}{\epsilon} CT \frac{1}{\sqrt{a}} \int_0^T \overline{\mathbb{E}}[|\widehat{\alpha}(Y^n_t, \varpi^n_t)|] dt \leq \frac{4|\overline{\Lambda}|^2 T}{\epsilon}(CT)^2 \frac{1}{\sqrt{a}}
\end{aligned}$$

Therefore, condition (2.119) is satisfied and the sequence $\int_0^T |\widehat{\alpha}^n_t|^2 dt$ converges weakly to $\int_0^T |\widehat{\alpha}^\infty_t|^2 dt$. We introduce now the following processes:

$$B^n_t := \int_0^t (\widehat{\alpha}^n_s + l(s, \varpi^n_s)) ds, \ \Sigma^{0,n}_t := \int_0^t \sigma^0(s, \varpi^n_s) db_s, \ \Sigma^n_t := \int_0^t \sigma(s, \varpi^n_s) dw_s.$$

Applying again [CD18b, Lemma 3.6], we prove that B^n converges weakly to $B^\infty := \int_0^\cdot (\widehat{\alpha}^\infty_s + l(s, \varpi^\infty_s)) ds$ on $\mathcal{C}([0,T];\mathbb{R})$. For $\Sigma^{0,n}$ and Σ^n, we proceed differently. Let us consider $\Sigma^{0,n}$. Since ϖ^n converges weakly to ϖ^∞ in $\mathcal{M}([0,T];\mathbb{R})$ and $\sigma(t, \varpi)$ is a continuous map in ϖ

variable, by [CD18b, Lemma 3.5], $(\sigma(t,\varpi_t^n))_{t\in[0,T]}$ converges weakly to $(\sigma(t,\varpi_t^\infty))_{t\in[0,T]}$ in $\mathcal{M}([0,T];\mathbb{R})$.

We want to exploit the continuity of the Brownian motion in order to obtain the convergence in distribution on $\mathcal{C}([0,T],\mathbb{R})$ of the stochastic integrals. In order to do so, we apply a particular case of Theorem 2.42, to the sequence $(\xi^n,\eta^n) := ((\sigma^0(t,\varpi_t^n))_{t\in[0,T]}, b)$. In order to do so, we need to check that conditions E1 and E2 hold for $((\sigma^0(t,\varpi_t^n))_{t\in[0,T]}, b)$. We observe that:

E1 Let us consider $\tau_n^c := \inf_{t\in[0,T]}\{b_t \geq c\}$. Hence, given a constant $c_\alpha > 0$, we introduce $\tau_n^\alpha := \tau_n^{c_\alpha}$:

$$\overline{\mathbb{P}}(\tau_n^\alpha \leq \alpha) = \overline{\mathbb{P}}\bigg(\inf_{t\in[0,T]}\{b_t \geq c_\alpha\} \leq \alpha\}\bigg) = \overline{\mathbb{P}}\bigg(\sup_{t\in[0,\alpha]} b_t \geq c_\alpha\bigg) = 2\overline{\mathbb{P}}(b_\alpha \geq c_\alpha) = 2\bigg(1-\Phi\Big(\frac{c_\alpha}{\sqrt{\alpha}}\Big)\bigg) \leq \frac{1}{\alpha},$$

for a sufficiently large c_α. Moreover, since $[b_t] = t$, (2.117) is satisfied.

E2 We introduce $\tau_n^c := \inf_{t\in[0,T]}\{|\sigma^0(t,\varpi_t^n)| > c\}$. By condition D2, given a constant $c_\alpha > 0$, we define $\tau_n^\alpha := \tau_n^{c_\alpha}$. Hence, by (2.116), the following holds

$$\begin{aligned}\overline{\mathbb{P}}(\tau_n^\alpha \leq \alpha) = \overline{\mathbb{P}}\bigg(\sup_{t\in[0,\alpha]}|\sigma^0(t,\varpi_t^n)| \geq c_\alpha\bigg) &\leq \frac{1}{c_\alpha}\overline{\mathbb{E}}\bigg[\sup_{t\in[0,T]}|\sigma^0(t,\varpi_t^n)|\bigg] \\ &\leq \frac{1}{c_\alpha}\overline{\mathbb{E}}\bigg[\sup_{t\in[0,T]}|\varpi_t^n|\bigg] \leq \frac{1}{c_\alpha}[(3L+\Lambda L_{h^0}T)+\Lambda\sqrt{C^\beta}] \leq \frac{1}{\alpha},\end{aligned}$$

for a sufficiently large c_α. To adopt a lighter notation we denote $\sigma^0(t,\varpi_t^n)$ by $\sigma_t^{0,n}$. We apply [Kur91, Lemma 5.2], which guarantees that

$$\overline{\mathbb{E}}[|\sigma^{0,n}_{\tau_n^\alpha\wedge T}|] \leq V_T(\sigma^{0,n}) + \overline{\mathbb{E}}[|\sigma_T^{0,n}|] \leq \sup_{n\in\mathbb{N}}\Big\{V_T(\sigma^{0,n}) + \overline{\mathbb{E}}[|w_T^n|]\Big\}.$$

In order to use this result, we need to check that $\sup_{n\in\mathbb{N}} V_t(\sigma^{0,n}) < \infty$, for every $t\in[0,T]$. To prove this condition, we need Assumption B4. Indeed, let us notice that, if ϕ is a convex function whose gradient is bounded and $\mathcal{G}$ is a sigma-algebra, the following holds

$$\begin{aligned}|\mathbb{E}[\phi(X)-\phi(Y)|\mathcal{G}]| &= \mathbb{E}[\phi(X)-\phi(Y)|\mathcal{G}] - 2\mathbb{1}_{\{\mathbb{E}[\phi(X)-\phi(Y)|\mathcal{G}]<0\}}\mathbb{E}[\phi(X)-\phi(Y)|\mathcal{G}], \\ &\leq \mathbb{E}[\phi(X)-\phi(Y)|\mathcal{G}] - 2\mathbb{1}_{\{\mathbb{E}[\phi(X)-\phi(Y)|\mathcal{G}]<0\}}(\mathbb{E}[\phi(X)-\phi(Y) \\ &\qquad -\phi'(Y)(X-Y)|\mathcal{G}] - \mathbb{E}[\phi'(Y)(X-Y)|\mathcal{G}]) \\ &\leq \mathbb{E}[\phi(X)-\phi(Y)|\mathcal{G}] - 2\mathbb{1}_{\{\mathbb{E}[\phi(X)-\phi(Y)|\mathcal{G}]<0\}}\mathbb{E}[\phi'(Y)(X-Y)|\mathcal{G}] \\ &\leq \mathbb{E}[\phi(X)-\phi(Y)|\mathcal{G}] + 2\sup_x|\phi'(x)|\mathbb{E}[X-Y|\mathcal{G}].\end{aligned} \tag{2.120}$$

If $\phi(X) = \sigma^0(t, X)$ and $\Delta := \{t_0, \ldots, t_N\}$, then:

$$\begin{aligned}
V_T^\Delta(\sigma^{0,n}) &:= \overline{\mathbb{E}}\left[\sum_i \left|\overline{\mathbb{E}}[\sigma^0(t_{i+1}, \varpi^n_{t_{i+1}}) - \sigma^0(t_i, \varpi^n_{t_i}) | \mathcal{F}^{n,l}_{t_i}]\right|\right] \\
&\leq \overline{\mathbb{E}}\left[\sum_i \overline{\mathbb{E}}[\sigma^0(t_{i+1}, \varpi^n_{t_{i+1}}) - \sigma^0(t_i, \varpi^n_{t_i}) | \mathcal{F}^{n,l}_{t_i}]\right] + 2 \sup_{t,\varpi} |\partial_t \sigma^0(t,\varpi)| \overline{\mathbb{E}}[\sum_i (t_{i+1} - t_i)] \\
&\quad + 2 \sup_{t,\varpi} |\partial_\varpi \sigma^0(t,\varpi)| \overline{\mathbb{E}}\left[\sum_i |\overline{\mathbb{E}}[\varpi^n_{t_{i+1}} - \varpi^n_{t_i} | \mathcal{F}^{n,l}_{t_i}]|\right] \\
&= \overline{\mathbb{E}}[\sigma^0(T, \varpi^n_T) - \sigma^0(0, \varpi^n_0)] + 2 \sup_{t,\varpi} |\partial_t \sigma^0(t,\varpi)| T \\
&\quad + 2 \sup_{t,\varpi} |\partial_\varpi \sigma^0(t,\varpi)| \overline{\mathbb{E}}\left[\sum_i |\overline{\mathbb{E}}[\varpi^n_{t_{i+1}} - \varpi^n_{t_i} | \mathcal{F}^{n,l}_{t_i}]|\right]
\end{aligned}$$

Applying the supremum in $n \in \mathbb{N}$, we conclude that:

$$\begin{aligned}
\sup_{n\in\mathbb{N}} V_t(\sigma^{0,n}) &\leq 2 \sup_{n\in\mathbb{N}} \left\{ \overline{\mathbb{E}}\left[\sup_{t\in[0,T]} |\sigma^0(t, \varpi^n_t)|\right] + \sup_{t,\varpi} |\partial_t \sigma^0(t,\varpi)| T + \sup_{t,\varpi} |\partial_\varpi \sigma^0(t,\varpi)| V_t(\varpi^n) \right\} \\
&\leq 2 \sup_{n\in\mathbb{N}} \left\{ \left(1 + \overline{\mathbb{E}}\left[\sup_{t\in[0,T]} |\varpi^n_t|\right]\right) + LT + LV_t(\varpi^n) \right\} < \infty.
\end{aligned}$$

Tightness of $(\varpi^n)_{n\in\mathbb{N}}$ guarantees that $\sup_{n\in\mathbb{N}} \overline{\mathbb{E}}[|\sigma^{0,n}_{\tau^\alpha_n \wedge T}|] < \infty$. We apply now [Kur91, Lemma 5.4] to the process $\widetilde{\sigma}^{0,n}_t := \mathbb{1}_{[0,\tau)}(t) \sigma^{0,n}_t$ to obtain:

$$V_t(\widetilde{\sigma}^{0,n}) \leq V_t(\sigma^{0,n}) + L\overline{\mathbb{E}}[|\varpi^n_t|].$$

We notice that: $\sigma^{0,n}_{t\wedge\tau} := \mathbb{1}_{[0,\tau)}(t)\sigma^{0,n}_t + \mathbb{1}_{[\tau,\infty)}(t)\sigma^{0,n}_\tau$, for every stopping time τ. Thus:

$$\begin{aligned}
V_t(\sigma^{0,n}_{\cdot\wedge\tau}) &= \sup_{\Delta\subset[0,t]} \overline{\mathbb{E}}\left[\sum_i \left|\overline{\mathbb{E}}[\sigma^{0,n}_{t_{i+1}\wedge\tau} - \sigma^{0,n}_{t_i\wedge\tau} | \mathcal{F}^{n,l}_{t_i}]\right|\right] \\
&\leq \sup_{\Delta\subset[0,t]} \overline{\mathbb{E}}\left[\sum_i \left|\overline{\mathbb{E}}[\widetilde{\sigma}^{0,n}_{t_{i+1}\wedge\tau} - \widetilde{\sigma}^{0,n}_{t_i\wedge\tau} | \mathcal{F}^{n,l}_{t_i}]\right|\right] \\
&\quad + \sup_{\Delta\subset[0,t]} \overline{\mathbb{E}}\left[\sum_i \left|\overline{\mathbb{E}}[\sigma^{0,n}_\tau (\mathbb{1}_{[\tau,\infty)}(t_{i+1}) - \mathbb{1}_{[\tau,\infty)}(t_i)) | \mathcal{F}^{n,l}_{t_i}]\right|\right] \\
&= V_t(\widetilde{\sigma}^{0,n}) + \sup_{\Delta\subset[0,t]} \overline{\mathbb{E}}\left[\sum_i \left|\overline{\mathbb{E}}[\sigma^{0,n}_\tau \mathbb{1}_{(t_i,t_{i+1}]}(\tau) | \mathcal{F}^{n,l}_{t_i}]\right|\right] \\
&\leq V_t(\widetilde{\sigma}^{0,n}) + \sup_{\Delta\subset[0,t]} \overline{\mathbb{E}}\left[\sum_i \overline{\mathbb{E}}\left[\sup_{s\in[t_i,t_{i+1}]} |\sigma^{0,n}_s| \mathbb{1}_{(t_i,t_{i+1}]}(\tau) | \mathcal{F}^{n,l}_{t_i}\right]\right]
\end{aligned}$$

$$
\begin{aligned}
&\leq V_t(\widetilde{\sigma}^{0,n}) + \sup_{\Delta\subset[0,t]} \sum_i \overline{\mathbb{E}}\left[\sup_{s\in[t_i,t_{i+1}]} |\sigma^{0,n}_s| \mathbb{1}_{(t_i,t_{i+1}]}(\tau)\right] \\
&\leq V_t(\widetilde{\sigma}^{0,n}) + \sup_{\Delta\subset[0,t]} \overline{\mathbb{E}}\left[\sup_{s\in[0,T]} |\sigma^{0,n}_s| \sum_i \mathbb{1}_{(t_i,t_{i+1}]}(\tau)\right] \\
&\leq V_t(\widetilde{\sigma}^{0,n}) + \sup_{\Delta\subset[0,t]} \overline{\mathbb{E}}\left[\sup_{s\in[0,T]} |\sigma^{0,n}_s| \mathbb{1}_{(0,T]}(\tau)\right] \\
&\leq V_t(\widetilde{\sigma}^{0,n}) + \overline{\mathbb{E}}\left[\sup_{s\in[0,T]} |\sigma^{0,n}_s|\right] \\
&\leq \overline{\mathbb{E}}\left[\sup_{t\in[0,T]} |\sigma^0(t,\varpi^n_t)|\right] + \sup_{t,\varpi} |\partial_t\sigma^0(t,\varpi)|T + \sup_{t,\varpi} |\partial_\varpi\sigma^0(t,\varpi)| V_t(\varpi^n) \\
&\quad + L\overline{\mathbb{E}}[|\varpi^n_t|] + \overline{\mathbb{E}}\left[\sup_{s\in[0,T]} |\varpi^n_s|\right].
\end{aligned}
$$

Taking the supremum we conclude that $\sup_{n\in\mathbb{N}} V_t(\sigma^{0,n}_{\cdot\wedge\tau^\alpha_n}) < \infty$.

As a consequence, applying Theorem 2.42 to $((\sigma^0(t,\varpi^n_t))_{t\in[0,T]}, b)$, the sequence

$$\left(\sigma^0(\cdot,\varpi^n_\cdot), b, \left(\int \sigma^0(t,\varpi^n_t)db_t\right)\right) \text{ is tight on } \mathcal{M}([0,T],\mathbb{R}) \times \mathcal{D}([0,T];\mathbb{R}^2).$$

Moreover, the weak convergence on $\mathcal{M}([0,T],\mathbb{R}) \times \mathcal{D}([0,T];\mathbb{R}^2)$ holds:

$$\left(\sigma^0(\cdot,\varpi^n_\cdot), b, \int \sigma^0(t,\varpi^n_t)db_t\right)\Bigg) \text{ converges in distribution to } \left(\sigma^0(\cdot,\varpi^\infty_\cdot), b^\infty, \left(\int \sigma^0(t,\varpi^\infty_t)db^\infty_t\right)\right).$$

Since the weak convergence of the stochastic integrals is in $\mathcal{D}([0,T];\mathbb{R})$ and the trajectories are continuous, the tightness can be restricted in $\mathcal{C}([0,T];\mathbb{R})$. In conclusion, doing the same computations for Σ^n we obtain that Σ^n converges in distribution to Σ^∞ on $\mathcal{C}([0,T];\mathbb{R})$. Moreover, up to subsequences, $(\Theta^n, B^n, \Sigma^{0,n}, \Sigma^n)_n$ converges to $(\Theta^\infty, B^\infty, \Sigma^{0,\infty}, \Sigma^\infty)$ on $\widetilde{\Omega}_{\text{input}} \times \mathcal{C}([0,T];\mathbb{R}) \times \mathcal{M}([0,T];\mathbb{R}) \times (\mathcal{C}([0,T];\mathbb{R}))^3$ in distribution. We consider the function:

$$
\begin{aligned}
h : \widetilde{\Omega}_{\text{input}} \times \mathcal{C}([0,T];\mathbb{R}) \times \mathcal{M}([0,T];\mathbb{R}) \times (\mathcal{C}([0,T];\mathbb{R}))^3 &\to \mathcal{C}([0,T];\mathbb{R}) \\
(\Theta, B, \Sigma^0, \Sigma) &\mapsto (X_t - (X^0 + B_t + \Sigma^0_t + \Sigma_t))_{t\in[0,T]}.
\end{aligned}
$$

Since h is continuous, by the continuous mapping theorem, $(h(\Theta^n, B^n, \Sigma^{0,n}, \Sigma^n))_n$ is convergent in distribution on $\mathcal{C}([0,T];\mathbb{R})$. Moreover, $h(\Theta^n, B^n, \Sigma^{0,n}, \Sigma^n) = 0$. The convergent series is constant, so we have the convergence also in probability ([Kal97, Lemma 5.7]). Therefore, extracting a subsequence, we obtain that $\mathbb{P}^\infty$-a.s.:

$$X^\infty - (X^{0,\infty} + B^\infty + \Sigma^{0,\infty} + \Sigma^\infty) = 0,$$

that is:

$$X_t^\infty = X^{0,\infty} + \int_0^t (\widehat{\alpha}_s^\infty + l(s,\varpi_s^\infty))ds + \int_0^t \sigma^0(s,\varpi_s^\infty)db_s^{1,\infty} + \int_0^t \sigma(s,\varpi_s^\infty)dw_s^\infty, \quad t\in[0,T]. \tag{2.121}$$

This implies that $\left(\int_0^t \widehat{\alpha}_s^\infty ds\right)_{t\in[0,T]}$ is $\mathbb{F}^\infty$-adapted. As a consequence: $\overline{\mathbb{P}}^\infty$-a.s.

$$\widehat{\alpha}_t^\infty = \lim_{p\to\infty} p\int_{(t-\frac{1}{p})_+}^t \widehat{\alpha}_s^\infty ds, \text{ a.e. } t\in[0,T].$$

In particular, this holds on the full-Lebesgue measured set in $[0,T]$ in which the càdlàg version of the weak limit $(\widehat{\alpha}^{n,l})_{n,l}$ in Meyer-Zheng topology is defined. We consider the optional projection of the weak limit $(\widehat{\alpha}_t^\infty)_{t\in[0,T]}$ given the filtration $\mathbb{F}^\infty$. We denote it by $(^\circ\widehat{\alpha}_t^\infty)_{t\in[0,T]}$. By definition, $^\circ\widehat{\alpha}_t^\infty = \mathbb{E}[\widehat{\alpha}_t^\infty|\mathcal{F}_t^\infty] =^\circ \widehat{\alpha}_t^\infty$ for almost any $t\in[0,T]$, $\mathbb{P}^\infty$-a.s. (in particular, $(^\circ\widehat{\alpha}^\infty)$ is progressively measurable with respect to $\mathbb{F}^\infty$). As a consequence, for almost every $t\in[0,T]$, equation (2.121) can be rewritten as:

$$X_t^\infty = X^{0,\infty} + \int_0^t (^\circ\widehat{\alpha}_s^\infty + l(s,\varpi_s^\infty))ds + \int_0^t \sigma^0(s,\varpi_s^\infty)db_s^{1,\infty} + \int_0^t \sigma(s,\varpi_s^\infty)dw_s^\infty, \quad t\in[0,T].$$

Step 3 Since the process ϖ^∞ has trajectories in $\mathcal{D}([0,T];\mathbb{R})$, we can find a countable set $\tau\subset[0,T]$ such that:

$$\mathbb{P}^\infty(\varpi^\infty \text{ is continuous on } [0,T]\setminus\tau) = 1.$$

Therefore, for eacn $N\geq 1$, there exists a sequence $0=t_0^N<\cdots<t_N^N=T$ of step length less than $\frac{2T}{N}$ on $[0,T]/\tau$. Therefore, we can introduce a family $(\Phi(t_i^N;\cdot))_{i=0,\ldots,N-1}\subseteq \mathcal{C}_b(\widetilde{\Omega}_{\text{input}}\times\mathcal{C}([0,T];\mathbb{R});A)$. As a consequence, for every $n\in\mathbb{N}\cup\{\infty\}$, we define:

$$\gamma_t^{n,N} := \sum_{i=0}^{N-1}\mathbb{1}_{[t_i^N,t_{i+1}^N)}(t)\Phi(t_i^N;X^0,b_{\cdot\wedge t_i^N},\mathcal{W}^n_{\cdot\wedge t_i^N},w_{\cdot\wedge t_i^N},X^n_{\cdot\wedge t_i^N}), \quad t\in[0,T].$$

In particular, $\gamma^{n,N}$ is $\mathbb{F}^n$-progressively measurable, then:

$$J^{\varpi^n}(\widehat{\alpha}^n)\leq J^{\varpi^n}(\gamma^{n,N}), \quad \forall n\in\mathbb{N}\cup\{\infty\}, N\in\mathbb{N}.$$

In particular, $\Phi(t_i^N;X^0,b_{\cdot\wedge t_i^N},\mathcal{W}^n_{\cdot\wedge t_i^N},w_{\cdot\wedge t_i^N},X^n_{\cdot\wedge t_i^N})$ weakly converges in A by the continuous mapping theorem. As a consequence, the function:

$$\begin{aligned}\Psi : \Omega_{\text{input}}\times\mathcal{C}([0,T];\mathbb{R}) &\to \mathcal{D}([0,T];\mathbb{R})\\ (x^0,b,\mathcal{W},w,x) &\mapsto \sum_{i=0}^{N-1}\mathbb{1}_{[t_i^N,t_{i+1}^N)}(t)\Phi(t_i^N;x^0,b_{\cdot\wedge t_i^N},\mathcal{W}_{\cdot\wedge t_i^N},w_{\cdot\wedge t_i^N},x_{\cdot\wedge t_i^N})\end{aligned}$$

is continuous. This implies that $(\gamma^{n,N})_{n\in\mathbb{N}}$ weakly converges to $\gamma^{\infty,N}$ in $\mathcal{D}([0,T];A)$ for each

N fixed. By the uniform square integrability of $\{Y^n\}_n$ also the optimal controls $\{\widehat{\alpha}^n\}_n$ are uniformly square integrable. By Assumption A2:

$$\lim_{n\to\infty} \mathbb{E}^n\left[\int_0^T f(t, X_t^n, \varpi_t^n, \widehat{\alpha}_t^n)dt + g(X_T^n, \varpi_T^n)\right] = \mathbb{E}^\infty\left[\int_0^T f(t, X_t^\infty, \varpi_t^\infty, \widehat{\alpha}^\infty)dt + g(X_T^\infty, \varpi^\infty)\right]$$

As a consequence, $J^{\varpi^\infty}(\widehat{\alpha}^\infty) \le \lim_{n\to\infty} J^{\varpi^n}(\gamma^{n,N})$, $\forall N \ge 0$. To pass to the limit in the right hand-side we recall that the sequence $(X^0, b, \mathcal{W}^n, w, X^n, \gamma^{n,N})_n$ converges in distribution to $(\xi^\infty, b^\infty, \mathcal{W}^\infty, w^\infty, X^\infty, \gamma^{\infty,N})$ (by the continuous mapping theorem). Let us notice that we can define the weak limit on the same filtered probability space $(\Omega^\infty, \mathcal{F}^\infty, \mathbb{P}^\infty, \mathbb{F}^\infty)$ on which the limit game is defined. Indeed, since the state equation satisfies strong uniqueness when the input $(\xi^\infty, b^\infty, \mathcal{W}^\infty, w^\infty, \gamma^{\infty,N})$ is given, we can define on the filtered space $(\Omega^\infty, \mathcal{F}^\infty, \mathbb{P}^\infty, \mathbb{F}^\infty)$ the control $\gamma^{\infty,N}$ introduced in equation

$$\gamma_t^{\infty,N} := \sum_{i=0}^{N-1} \mathbb{1}_{[t_i^N, t_{i+1}^N)}(t)\Phi(t_i^N; X^{0,\infty}, b^{1,\infty}_{\cdot\wedge t_i^N}, \mathcal{W}^\infty_{\cdot\wedge t_i^N}, w^\infty_{\cdot\wedge t_i^N}, X^\infty_{\cdot\wedge t_i^N}), \quad t \in [0, T],$$

that is $\mathbb{F}^\infty$-progressive measurable by definition.

Denoting by $\widetilde{\mathbb{F}}$ the complete and right-continuous augmentation of

$$(\xi^\infty, b^\infty, \mathcal{W}^\infty, w^\infty, \gamma^{\infty,N}),$$

by the Yamada-Watanabe theorem the solution to the state equation is $\widetilde{\mathbb{F}}$-progressively measurable. As a consequence, on $(\Omega^\infty, \mathcal{F}^\infty, \mathbb{P}^\infty, \widetilde{\mathbb{F}})$, the solution $X^{\infty,N}$ driven by the control $\gamma^{\infty,N}$ is uniquely determined by the law of $(\xi^\infty, b^\infty, (\mathcal{W}^\infty, \gamma^{\infty,N}), w^\infty)$ and therefore it is $\widetilde{\mathbb{F}}$-progressively measurable.

Since $\widetilde{\mathbb{F}} \subseteq \mathbb{F}^\infty$ and b^∞ and w^∞ are $\mathbb{F}^\infty$-Brownian motions, we can define the solution to the state equation with the control $\gamma^{\infty,N}$ on $(\Omega^\infty, \mathcal{F}^\infty, \mathbb{P}^\infty, \mathbb{F}^\infty)$.

We conclude that the functional $(\int_0^T f(t, X_t^n, \varpi_t^n, \gamma_t^{n,N})dt + g(X_T^n, \varpi_T^n))$ converges in law to $\int_0^T f(t, X_t^\infty, \varpi_t^\infty, \gamma_t^{\infty,N})dt + g(X_T^\infty, \varpi_T^\infty)$ and the limit can be seen as the distribution of a process defined on $(\Omega^\infty, \mathcal{F}^\infty, \mathbb{P}^\infty, \mathbb{F}^\infty)$. Finally, to prove that $J^{\varpi^n}(\gamma^{n,N})$ converges to $J^{\varpi^\infty}(\gamma^{\infty,N})$ we need to check square uniform integrability of the controls $(\gamma^{n,N})_n$. This condition is guaranteed by the boundedness of the functionals $\Phi(t_i; \cdot)$ that define $\gamma^{n,N}$.

Therefore, the sequence $J^{\varpi^n}(\gamma^{n,N})$ converges to $J^{\varpi^\infty}(\gamma^{\infty,N})$ and we have:

$$J^{\varpi^\infty}(\widehat{\alpha}_t^\infty) \le J^{\varpi^\infty}(\gamma^{\infty,N}), \quad \forall N \ge 0.$$

We now approximate every $\mathbb{F}^\infty$-progressively measurable process with the limit of a sequence like the one $(\gamma^{\infty,N})_N$, applying the strategy proposed in the third step of the proof of [CD18b, Proposition 3.11]. We proceed as follows:

(a) By convexity, we can approximate γ by $\pi_R(\gamma)$ where, π_R is the projection on $B_0(R)$. As a consequence, we can take γ bounded.

(b) We consider the process $\left(\frac{1}{h}\int_{t-h}^{t}\gamma_s ds\right)_{0\le t\le T}$ where for negative s we impose that $\gamma_s = a$, for an arbitrary $a \in \mathbb{R}$. As a consequence, we can assume that γ has Lipschitz continuous paths.

(c) We can approximate now γ with a sum $\left(\sum_{i=1}^{N}\gamma_{t_{i-1}^N}\mathbb{1}_{[t_i^N,t_{i+1}^N)}(t)\right)_{0\le t\le t^N}$.

(d) Since γ is supposed to be continuous in time, each $\gamma_{t_{i-1}^N}$ is measurable with respect to the completion of $\sigma\{(\xi^\infty, b_s^\infty, \mathcal{W}^\infty, w_s^\infty, X_s^\infty :\ s \le t_{i-1}^N)\}$.

By Lemma 2.43, there exists a version of $\gamma_{t_{i-1}^N}$, which is measurable with respect to $\sigma\{(\xi^\infty, b_s^\infty, \mathcal{W}_s^\infty, w_s^\infty, X_s^\infty :\ s \le t_{i-1}^N)\}$.

This implies that, there exists a bounded and measurable function $\Phi(t_i^N;\cdot)$ such that:

$$\gamma_{t_i^N}^N := \Phi(t_i^N;\xi^\infty, b^\infty_{\cdot\wedge t_i^N}, \mathcal{W}^\infty_{\cdot\wedge t_i^N}, w^\infty_{\cdot\wedge t_i^N}, X^\infty_{\cdot\wedge t_i^N}).$$

(e) Finally, we can approximate the function $\Phi(t_i^N;\cdot)$ in $L^2(\Omega_{\text{input}}\times\mathcal{C}([0,T];\mathbb{R}))$ with continuous functions, applying Lusin theorem.

Step 4 By the previous step the following holds $J^{\varpi^\infty}(\widehat{\alpha}^\infty) \le J^{\varpi^\infty}({}^\circ\widehat{\alpha}^\infty)$. As a consequence:

$$\begin{aligned}
J^{\varpi^\infty}(\widehat{\alpha}^\infty) &= \mathbb{E}^\infty\left[\int_0^T f(s, X_s^\infty, \varpi_s^\infty, \widehat{\alpha}_s^\infty)ds + g(X_T^\infty, \varpi_T^\infty)\right]\\
A3 &\ge \mathbb{E}^\infty\left[\int_0^T f(s, X_s^\infty, \varpi_s^\infty, {}^\circ\widehat{\alpha}_s^\infty)ds + g(X_T^\infty, \varpi_T^\infty)\right]+\\
&\quad + \mathbb{E}^\infty\left[\int_0^T (\widehat{\alpha}_s^\infty - {}^\circ\widehat{\alpha}_s^\infty)\partial_\alpha f(s, X_s^\infty, \varpi_s^\infty, {}^\circ\widehat{\alpha}_s^\infty)ds\right] + \frac{1}{2}L^{-1}\mathbb{E}^\infty\left[\int_0^T |\widehat{\alpha}_s^\infty - {}^\circ\widehat{\alpha}_s^\infty|^2 ds\right]\\
&= J^{\varpi^\infty}({}^\circ\widehat{\alpha}^\infty) + \frac{1}{2}L^{-1}\mathbb{E}^\infty\left[\int_0^T |\widehat{\alpha}_s^\infty - {}^\circ\widehat{\alpha}_s^\infty|^2 ds\right].
\end{aligned}$$

In conclusion

$$\mathbb{E}^\infty\left[\int_0^T |\widehat{\alpha}_s^\infty - {}^\circ\widehat{\alpha}_s^\infty|^2 ds\right] = 0$$

and thus $\widehat{\alpha}_s^\infty$ possesses an $\mathbb{F}^\infty$-progressively measurable modification.

□

CHAPTER 3

On a foundamental theorem for statistical arbitrage opportunities

3.1 Introduction

The term statistical arbitrage is commonly adopted to denote a trading strategy characterized by milder conditions than the ones that define an arbitrage opportunity. We recall that an arbitrage opportunity is defined as a self-financing strategy that guarantees a net profit taking no risks. If a market allows for arbitrage opportunities, it is considered inefficient, because the presence of arbitrage opportunities implies that certain assets are not priced correctly. Indeed, if an arbitrage opportunity is present, a large demand (or supply) of some assets is created, leading to a modification of the prices and consequently the disappearance of the opportunity itself. The characterization of the absence of arbitrage opportunities has been a crucial task in mathematical finance. The study of the conditions that guarantee absence of arbitrage opportunities leaded to the fundamental theorem of asset pricing ([DS94, Theorem 1.1]).

As discussed, it is reasonable to suppose that the market is free of arbitrage opportunity. As a consequence, speculators aim at constructing strategies that lead profit involving some risks that can be either controlled or assessed. These strategies are called *statistical arbitrage opportunities.*

Widely analysed in the literature, the term of statistical arbitrage can refer to different kinds of financial strategies, that extend the definition of arbitrage opportunities. All these different strategies, linked with the term statistical arbitrage, are surveyed in [LBŠ+18]. In this paper, the authors compared the differences between the strategies commonly adopted in the market that are referred as statistical arbitrage opportunities. In particular, a statistical arbitrage can refer to the notion of δ-Arbitrage, defined in [Led95] as an investment strategies satisfying a prescribed condition on its Sharpe ratio. Successively, this definition of δ-Arbitrage was replaced by the notion of Approximate Arbitrage, introduced in [BL00]. Another generalization of the notion of arbitrage opportunity, that is linked by the definition of statistical arbitrage is given by the Good Deal, introduced in [CS00] in the context of incomplete financial markets and investigated successively in [ČH02], [Sta04] and in [BS06]. Finally, there can be found links with the notion of statistical arbitrage in [CGM01] (Acceptable Opportunity) and in [BKL01] (ϵ-Arbitrage). In [HJTW04], in the context of infinite-time horizon models, a statistical arbitrage is defined as a zero-cost, self financing strategy, that asymptotically has positive expected payoff, a probability of loss converging to zero and a time-averaged variance converging to zero (if the probability of a loss does not become zero in finite time). In [Bon03] a statistical arbitrage is defined, in a finite-horizon economy, as a zero-cost trading strategy for which the expected payoff is positive and the conditional expected payoff in each final state of the economy is nonnegative.

In this chapter we focus on the notion of statistical arbitrage analysed in [Bon03]. The author studied the case of a market model determined by a single asset, modelled by a discrete-time stochastic process defined on a finite probability space. The time horizon, denoted by $T > 0$, is finite. The state of the economy at time t is represented by a random variable ξ_t. Therefore, the information structure in the market is determined by the vector $I_t := (\xi_0, \dots, \xi_t)$, for each $t = 0, \dots, T$. A *statistical arbitrage opportunity* (SAO) is defined in [Bon03] as a zero-cost trading strategy for which the expected payoff is positive and the conditional expected payoff in each final state of the economy (i.e. each realization of the random variable ξ_T) is nonnegative. To formalize this setting, the author introduced the vector of augmented information set $I_t^{\xi_T} := (\xi_0, \dots, \xi_t; \xi_T)$. Then, a (SAO) is defined as follows:

Definition 3.1. ([Bon03, Definition 2]) A zero-cost trading strategies, with payoff Z_T (depending on I_T), is called (SAO) if:

1. $\mathbb{E}[Z_T | I_0] > 0$;
2. $\mathbb{E}[Z_T | I_0^{\xi_T}] \geq 0$, for every realization of ξ_T.

In [Bon03, Proposition 1], it is shown shown that the absence of (SAO) is equivalent to the existence of a probability measure $\mathbb{Q}$ such that the price process is a $\mathbb{Q}$-martingale and its kernel with respect the objective measure $\mathbb{P}$ is measurable with respect to ξ_T. To prove this result, the author applied the same techniques described in [HK79] and [HP81] to characterize the absence of classical arbitrage opportunities in a multi-period finite dimensional market. Its generalization to financial markets defined on general probability spaces represents an open problem.

In this chapter we present some preliminary results in the direction of a characterization of the absence of statistical arbitrage opportunities, under a suitable generalization of the definition introduced by Bondarenko, in the case of market models defined on general probability spaces. In the recent work [RRS21], some results in this direction have already been provided. The authors generalized Definition 3.1 to the case of self-financing zero-cost trading strategies that allow for a positive expected payoff and conditional expected payoff that is nonnegative in (almost) every realization of an augmented information set, described by a given sigma-algebra $\mathcal{G}$. In [RRS21, Section 2.2], the authors provided a counterexample to the characterization of absence of (SAO) opportunities proved in [Bon03, Proposition 1]. However, we confirm [Bon03, Proposition 1] in the case of finite dimensional markets.

The chapter is structured as follows. In Section 3.3, we present the mathematical setup. In Section 3.4, we describe the characterization of the absence of statistical arbitrage opportunities in the case of a market defined on a finite probability space. In particular, we prove that the absence of statistical arbitrage opportunities is equivalent to the existence of a martingale measure for the price process, whose Radon-Nykodym derivative with respect to the objective measure $\mathbb{P}$ is $\mathcal{G}$-measurable. Moreover, we study the counterexample proposed in [RRS21], showing that it does not disprove [Bon03, Proposition 1]. In Section 3.5, we show that the existence of the martingale measure introduced in the finite setting is actually equivalent to a stronger condition than the absence of statistical arbitrage opportunities. In Section 3.6, we focus on discrete-time market defined on a general probability space. In particular, we provide conditions under which the characterization that holds in the finite-dimensional case is preserved. Finally, in Section 3.7, we establish a characterization of the absence of statistical arbitrage opportunities in the case of a semimartingale market under a taylor-made condition.

3.2 Notation

In this section we introduce the notation we are adopting along the chapter.

- Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, for any $p, q \in [1, \infty]$ such that $\frac{1}{p} + \frac{1}{q} = 1$, we compactly denote $L^p(\Omega, \mathcal{F}, \mathbb{P})$ by $L^p(\mathcal{F})$. We adopt the convention that $p = 1$ if and only if $q = \infty$ and vice versa. When it is clear from the context, we may adopt the notation $L^p(\mathcal{F})$, to denote also the space of $\mathbb{R}^d$-valued random variables, whose components are in $L^p(\mathcal{F})$.

- For any $p \in [1, \infty)$, the dual of $L^p(\mathcal{F})$ is $L^q(\mathcal{F})$, where $q := \frac{p}{p-1}$ (if $p = 1$, $q = \infty$). The weak topology of $L^p(\mathcal{F})$ induced by $L^q(\mathcal{F})$ is denoted by $\sigma(L^p(\mathcal{F}), L^q(\mathcal{F}))$.

- For $p = \infty$, we denote by $\sigma(L^\infty(\mathcal{F}), L^1(\mathcal{F}))$ the weak-$*$ topology on $L^\infty(\mathcal{F})$, induced by $L^1(\mathcal{F})$.

- Let A be a set in $L^p(\mathcal{F})$, for $p \in [1, \infty]$. We denote:
  - the $\sigma(L^p(\mathcal{F}), L^q(\mathcal{F}))$-closure of A, by $Cl^*(A)$, where $\frac{1}{p} + \frac{1}{q} = 1$.
  - the closure of A in the norm topology of $L^p(\mathcal{F})$ by $\overline{Cl}(A)$.

- In a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we consider a sub-sigma-algebra $\mathcal{G} \subseteq \mathcal{F}$. We recall that the conditional expectation of $X \in L^1(\mathcal{F})$ with respect to $\mathcal{G}$ is a random variable $Y \in L^1(\mathcal{G})$, such that for every $G \in \mathcal{G}$, $\mathbb{E}[Y \mathbb{1}_G] = \mathbb{E}[X \mathbb{1}_G]$. It can be proved that such random variable Y is $\mathbb{P}$-a.s. unique and we denote it by $\mathbb{E}[X|\mathcal{G}]$. In the following, we may denote the conditional expectation $\mathbb{E}[X|\mathcal{G}]$ by $\pi_{\mathcal{G}}(X)$, to emphasize that the conditional expectation can be interpreted as a linear projection:

$$\pi_{\mathcal{G}} : L^1(\Omega, \mathcal{F}, \mathbb{P}) \to L^1(\Omega, \mathcal{G}, \mathbb{P})$$
$$X \mapsto \mathbb{E}[X|\mathcal{G}].$$

In analogy to [JS13, equation (1.1)], we extend the definition of condition expectation to every random variable $X \in L^0(\Omega, \mathcal{F}, \mathbb{P})$, introducing the *generalized conditional expectation*:

$$\mathbb{E}[X|\mathcal{G}] := \begin{cases} \mathbb{E}[X^+|\mathcal{G}] - \mathbb{E}[X^-|\mathcal{G}] & \text{on the set where } \mathbb{E}[|X||\mathcal{G}] < \infty, \\ +\infty & \text{otherwise} \end{cases}$$

where $X^+ := \max\{X, 0\}$ and $X^- := \max\{-X, 0\}$.

- We denote the positive cone on $\mathbb{R}^N$, by $\mathbb{R}^N_+$, i.e. $\mathbb{R}^N_+ := \{x \in \mathbb{R}^N : \; x_i \geq 0, \; \forall i = 1, \dots, N\}$. For every $p \in [1, \infty]$, we denote the positive cone on $L^p(\mathcal{F})$, by $L^p_+(\mathcal{F}) = \{l \in L^p(\mathcal{F}), \; l \geq 0, \; \mathbb{P}\text{-a.s.}\}$.

- We denote the scalar product of two vector $a, b \in \mathbb{R}^d$ by $a \cdot b := \sum_{i=1}^d a_i b_i$. If E is a topological vector space and E^* is its dual, then for every $a \in E$ and $b \in E^*$, we may denote by $\langle b, a \rangle$ the duality introduced in [AB06, Definition 5.90]. In particular, for any $p \in [1, \infty]$, given q such that $\frac{1}{p} + \frac{1}{q} = 1$, for every $f \in L^p(\mathcal{F})$ and $g \in L^q(\mathcal{F})$, $\langle f, g \rangle = \mathbb{E}[f \cdot q]$.

3.3 The formalization of the problem

In this section, we introduce the mathematical framework, the definition of statistical $\mathcal{G}$-arbitrage opportunity and the preliminary properties of the set of statistical $\mathcal{G}$-arbitrage opportunities.

We consider a finite time-horizon T and a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We define the class of admissible trading strategies in two cases, when the price process is a discrete-time stochastic process defined on $\mathbb{T} := \{0, \dots, T\}$ and when the price process is a continuous-time stochastic process defined on the time interval $\mathcal{T} := [0, T]$. In what follows, we denote the price process by S and assume that it is defined by a stochastic process taking values on $\mathbb{R}^d$, expressed already in discounted terms.

Definition 3.2. We consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a discrete time set $\mathbb{T}$. We introduce a filtration $\mathbb{F} := (\mathcal{F}_t)_{t=0}^T$ defined on $\mathbb{T} := \{0, \dots, T\}$. We consider an $\mathbb{R}^d$-valued stochastic process S adapted to $\mathbb{F}$. We introduce the following concepts:

- An admissible trading strategy ϕ is an $\mathbb{R}^d$-valued predictable discrete-time stochastic process $(\phi_i)_{i=1}^T$, i.e. ϕ_i is $\mathcal{F}_{i-1}$-measurable for any $i = 1, \dots, T$. We denote the set of such strategies by $\mathcal{A}$.

- The value process of ϕ is defined by $V_0(\phi) = 0$ and

$$V_t(\phi) := \sum_{i=1}^{t} \phi_i \cdot \Delta S_i, \quad t \in \{1, \dots, T\},$$

  where $\Delta S_i := S_i - S_{i-1}$ for any $i = 1, \dots, T$. We compactly denote $(V_t(\phi))_{t\in[0,T]}$ by $V(\phi)$.

In continuous time, we give the following definition:

Definition 3.3. We consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a time interval $[0, T]$. We introduce a right-continuous and complete filtration $\mathbb{F} := (\mathcal{F}_t)_{t\in[0,T]}$. We consider an $\mathbb{R}^d$-valued locally bounded semimartingale $S := (S^1, \dots, S^d)$. We suppose that S is adapted to $\mathbb{F}$. Then:

- A trading strategy ϕ is every S-integrable and predictable stochastic process on taking values on $\mathbb{R}^d$;

- The value process of ϕ is defined by:

$$V_t(\phi) := \int_0^t \phi_s \cdot dS_s + \phi_0 S_0, \quad t \in [0, T].$$

  We compactly denote $(V_t(\phi))_{t\in[0,T]}$ by $V(\phi)$.

- A trading strategy ϕ is called admissible if $\phi_0 = 0$ and there exists a constant $a \geq 0$ such that $V_t(\phi) \geq -a$, a.s., for all $t \in [0, T]$. In this setting, analogously with Definition 3.2, we denote the set of admissible trading strategies by $\mathcal{A}$.

In analogy to [RRS21, Definition 2.1], we introduce the definition of statistical $\mathcal{G}$-arbitrage opportunity, that can be applied both for discrete-time and continuous-time markets:

Definition 3.4. Let $\mathcal{G} \subseteq \mathcal{F}_T$ be a sigma-algebra. An admissible strategy ϕ is called *statistical $\mathcal{G}$-arbitrage opportunity* if the two following conditions hold:

$$\mathbb{E}[V_T(\phi)] > 0,$$
$$\mathbb{E}[V_T(\phi)|\mathcal{G}] \geq 0, \quad \text{a.s.}$$

Let us notice that, if $\mathcal{G} := \mathcal{F}_T$, we recover the classical notion of arbitrage opportunity. Moreover, in the context of discrete-time markets, if $\mathcal{G} := \sigma\{S_T\}$, we recover the notion of statistical arbitrage opportunities introduced in [Bon03] (Definition 3.1). Finally, By the tower property if $\mathcal{G}_1 \subset \mathcal{G}_2$ are two sigma-algebras on $(\Omega, \mathcal{F}, \mathbb{P})$, the set of statistical $\mathcal{G}_2$-arbitrage opportunities is contained in the set of statistical $\mathcal{G}_1$-arbitrage opportunities. In particular, this implies that absence of statistical $\mathcal{G}$-arbitrage opportunities guarantees that the market is free of standard arbitrage opportunities.

Remark 3.5 (Financial interpretation). By definition, a statistical $\mathcal{G}$-arbitrage opportunity is a trading strategy ϕ whose final portfolio value can be negative at the condition that the aggregated value in each scenario of the sigma-algebra $\mathcal{G}$ is non negative and, moreover, there exists a non negligible set on $\mathcal{F}$ on which $V_T(\phi)$ is strictly positive in average. From an economic point of view, the definition of statistical $\mathcal{G}$-arbitrage can be employed to evaluate if a trading strategy is worth to be applied. For instance, one can introduce a fixed benchmark of scenarios to provide a criterion to establish the quality of a trading strategy. In particular, an investment based on a strategy ϕ can be labelled as a *good* investment if the average final value of the portfolio determined by ϕ is non-negative in all the benchmark scenarios. In this example, denoted by $B_1, \dots, B_m$ the benchmark scenarios, ϕ generates a *good* investment if and only if ϕ is a statistical $\mathcal{G}$-arbitrage opportunity, where $\mathcal{G} := \sigma\{B_1, \dots, B_m\}$.

Let us notice that richer is the sigma-algebra $\mathcal{G}$ stricter become the conditions that guarantee absence of statistical $\mathcal{G}$-arbitrage opportunities. If for example $\mathcal{G}$ is the trivial sigma-algebra the statistical $\mathcal{G}$-arbitrage is a weak definition and it is rather unrealistic to suppose absence of statistical $\mathcal{G}$-arbitrage opportunities. On the other hand, practitioners can be interested in conditions under which the market the free of statistical $\mathcal{G}$-arbitrage opportunities, when $\mathcal{G}$ is defined by a large number of benchmark scenarios.

In what follows, we denote the absence of statistical $\mathcal{G}$-arbitrage opportunities with the symbol NSA($\mathcal{G}$). We introduce the set $K_{\mathcal{G}}$ by:

$$K^0_{\mathcal{G}} := \{\mathbb{E}[V_T(\phi)|\mathcal{G}]|\ \phi \in \mathcal{A}\}. \tag{3.1}$$

Then, NSA($\mathcal{G}$) is equivalent to:

$$K^0_{\mathcal{G}} \cap L^0_+(\Omega, \mathcal{G}, \mathbb{P}) = \{0\}. \tag{NSA($\mathcal{G}$)}$$

Our goal is to find equivalent conditions to NSA($\mathcal{G}$).

Let us recall that, in the framework described in [Bon03] (i.e. $\mathcal{G} = \sigma\{S_T\}$), condition NSA($\mathcal{G}$) is equivalent to the existence of a probability measure $\mathbb{Q}$, equivalent to $\mathbb{P}$ such that $S = (S_t)_{t=0}^T$ is a $\mathbb{Q}$-martingale, i.e. $\mathbb{E}^{\mathbb{Q}}[S_{t+1}|\mathcal{F}_t] = S_t$ and the Radon-Nykodym derivative $\frac{d\mathbb{Q}}{d\mathbb{P}}$ is $\sigma\{S_T\}$-measurable ([Bon03, Proposition 3]).

The natural generalization of this result to the case of statistical $\mathcal{G}$-arbitrage opportunities is the characterization of NSA($\mathcal{G}$) in terms of the following condition:

- On $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t=0}^T)$, there exists a probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ such that $S = (S_t)_{t=0}^T$ is a $\mathbb{Q}$-martingale and the Radon-Nikodym derivative $\frac{d\mathbb{Q}}{d\mathbb{P}}$ is $\mathcal{G}$-measurable. We denote this condition by EMM($\mathcal{G}$).

- On $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,T]})$, there exists a probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ such that $S = (S_t)_{t\in[0,T]}$ is a $\mathbb{Q}$-local martingale and the Radon-Nikodym derivative $\frac{d\mathbb{Q}}{d\mathbb{P}}$ is $\mathcal{G}$-measurable. We denote this condition by ELMM($\mathcal{G}$).

When the market structure is not specified, we denote both these conditions by E(L)MM($\mathcal{G}$).

As discussed in Section 3.1, in [RRS21, Section 2.2] a counterexample to [Bon03, Proposition 3] is proposed in the case in which $\mathcal{G} := \sigma\{S_T\}$. However, as described in Section 3.4.1 below, a gap affects the computations of the counterexample. Then, as we show in Section 3.4 below, NSA($\mathcal{G}$) $\iff$ EMM($\mathcal{G}$) holds for generic sigma-algebra $\mathcal{G}$, when Ω is a finite set.

As expected, in the general case, the difficult implication is NSA($\mathcal{G}$)$\Rightarrow$E(L)MM($\mathcal{G}$). Indeed, in [RRS21, Theorem 2.1], the converse implication is proved. We give the proof of this result, for completeness of presentation:

Theorem 3.6. *Let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,T]})$ be a filtered probability space on which an $\mathbb{R}^d$-valued, locally bounded semimartingale process S is defined. Then, if the set of admissible strategies is the one introduced in Definition 3.3, ELMM($\mathcal{G}$) $\Rightarrow$NSA($\mathcal{G}$).*

Proof. We consider an equivalent local martingale measure $\mathbb{Q}$ for the price S. By assumption the kernel $Z := \frac{d\mathbb{Q}}{d\mathbb{P}}$ is a $\mathcal{G}$-measurable random variable. By contradiction, suppose that there exists a statistical $\mathcal{G}$-arbitrage opportunity ϕ, in the sense of Definition 3.4. We first observe that:

$$\mathbb{E}^{\mathbb{Q}}[V_T(\phi)|\mathcal{G}] = \frac{1}{\mathbb{E}^{\mathbb{P}}[Z|\mathcal{G}]}\mathbb{E}^{\mathbb{P}}[V_T(\phi)Z|\mathcal{G}] = \mathbb{E}^{\mathbb{P}}[V_T(\phi)|G], \quad \text{a.s.}$$

Therefore, by definition of statistical $\mathcal{G}$-arbitrage and the equivalenece between $\mathbb{Q}$ and $\mathbb{P}$:

$$\mathbb{E}^{\mathbb{Q}}[V_T(\phi)|\mathcal{G}] \geq 0, \quad a.s.$$

Moreover, since ϕ is admissible, $(V_t(\phi))_{t\in[0,T]}$ is a $\mathbb{Q}$-local martingale ([AS94, Proposition 3.3]). Applying then Fatou's lemma it is possible to show that $(V_t(\phi))_{t\in[0,T]}$ is a $\mathbb{Q}$-supermartingale. This implies that $\mathbb{E}^{\mathbb{Q}}[V_T(\phi)] \leq V_0(\phi) = 0$. Therefore, $0 = \mathbb{E}^{\mathbb{Q}}[V_T(\phi)|\mathcal{G}] = \mathbb{E}^{\mathbb{P}}[V_T(\phi)|\mathcal{G}]$ *a.s.*, that is a contradiction. ☐

The same reasoning can be applied in the case of discrete-time markets, under the notation introduced in Definition 3.2. Indeed, for every $\phi = (\phi_t)_{t=0}^T$ predictable, if S is a $\mathbb{Q}$-martingale, also $(V_t(\phi))_{t=0}^T$ is $\mathbb{Q}$-martingale:

$$\mathbb{E}^{\mathbb{Q}}[V_{t+1}(\phi)|\mathcal{F}_t] = \mathbb{E}^{\mathbb{Q}}\Big[\sum_{i=1}^{t+1}\phi_i \cdot \Delta S_i|\mathcal{F}_t\Big] = \sum_{i=1}^{t}\phi_i \cdot \Delta S_i + \phi_{t+1} \cdot \mathbb{E}^{\mathbb{Q}}[\Delta S_{t+1}|\mathcal{F}_t] = \sum_{i=1}^{t}\phi_i \cdot \Delta S_i = V_t(\phi).$$

As a consequence, EMM($\mathcal{G}$) implies that $\mathbb{E}^{\mathbb{Q}}[V_{t+1}(\phi)|\mathcal{F}_t] = 0$, for every $\phi \in \mathcal{A}$. Thus, no statistical $\mathcal{G}$-arbitrage opportunities are allowed.

Remark 3.7. The converse implication, NSA($\mathcal{G}$) $\Rightarrow$ ELMM($\mathcal{G}$), has to be handled more carefully, as we are going to see in the next sections. However, assuming additionally that there exists a unique element in

$$\mathcal{M}^e(S) := \{\mathbb{Q} \sim \mathbb{P} \text{ probability measure on } (\Omega, \mathcal{F}) \text{ s.t. } S \text{ is a } \mathbb{Q}\text{-local martingale}\}, \tag{3.2}$$

it is possible to show that NSA($\mathcal{G}$) is equivalent to ELMM($\mathcal{G}$). This result, shown in [RRS21, Theorem 3.3], is based on the application of [DS95, Theorem 16]) and [KL17, Proposition 6].

3.4 Absence of statistical $\mathcal{G}$-arbitrage opportunities in finite dimensional markets

As discussed in Section 3.1, when $\mathcal{M}^e(S) \neq \{\mathbb{Q}\}$, condition NSA($\mathcal{G}$)$\Rightarrow$ E(L)MM($\mathcal{G}$) is not guaranteed. In this section, we show that if the market is defined on a finite probability space, the equivalence is verified, meaning that

($\bigstar$) If Ω is a finite set, then NSA($\mathcal{G}$) $\Longleftrightarrow$ EMM($\mathcal{G}$).

This result, stated in Theorem 3.10 below, extends and confirms the characterization first provided in [Bon03, Proposition 3]. However, it is in contrast to the counterexample proposed in [RRS21, Section 2.2]. In the next section, we describe the gap that affected this counterexample, proving that it does not disprove condition ($\bigstar$).

3.4.1 Analysis of the counterexample proposed in [RRS21, Section 2.2]

In the example studied in [RRS21, Section 2.2], the authors analysed the trinomial model described by Figure 3.4.1. In particular, they considered a probability space $\Omega = \{\omega_1, \dots, \omega_6\}$ and $T = 2$.

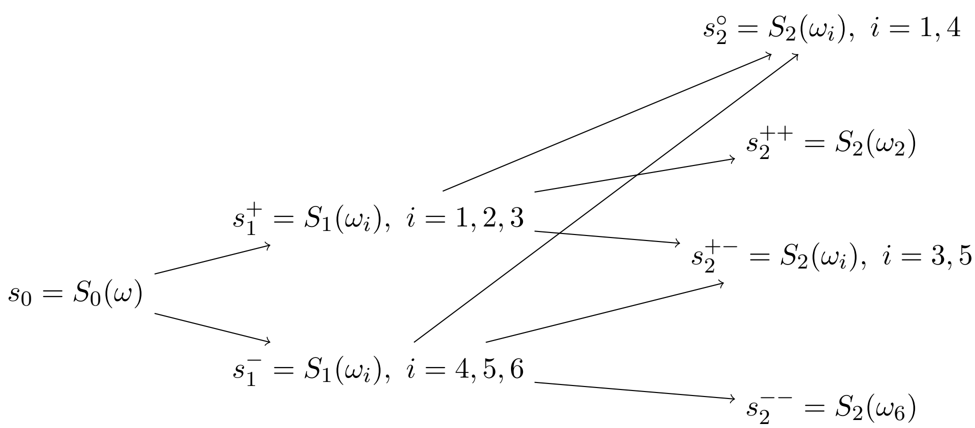

Figure 3.4.1: The tree of the price process.

In particular, the price process S is defined as:

$$\begin{cases} S_1(\omega_1) = S_1(\omega_2) = S_1(\omega_3) = s_1^+, \\ S_1(\omega_4) = S_1(\omega_5) = S_1(\omega_6) = s_1^-, \\ S_2(\omega_1) = S_2(\omega_4) = s_2^o, \\ S_2(\omega_3) = S_2(\omega_5) = s_2^{+-}, \\ S_2(\omega_2) = s_2^{++}, \\ S_2(\omega_6) = s_2^{--}, \end{cases}$$

Let us notice that a trading strategy $\phi = (\phi_1, \phi_2^+, \phi_2^-)$ in the sense of Definition 3.2 is any predictable process. The portfolio value associated with ϕ at $t = 1, 2$ is $V_t(\phi) = \sum_{i=1}^t \phi_i \Delta S_i$, where $\phi_2 \in \{\phi_2^+, \phi_2^-\}$.

The authors studied the absence of statistical $\mathcal{G}$-arbitrage opportunities for the following sigma-algebra:

$$\mathcal{G} := \sigma(S_2) = \sigma(\{\omega_1, \omega_4\}, \{\omega_3, \omega_5\}, \{\omega_2\}, \{\omega_6\}).$$

By Definition 3.4, a stastistical $\mathcal{G}$-arbitrage is an admissible trading strategy $\phi := (\phi_1, \phi_2^+, \phi_2^-)$ such that the following inequalities holds:

$$\begin{cases} \mathbb{E}[V_2(\phi)|\{\omega_2\}] \geq 0, & \phi_1 \Delta S_1(\omega_2) + \phi_2^+ \Delta S_2(\omega_2) \geq 0, \\ \mathbb{E}[V_2(\phi)|\{\omega_6\}] \geq 0, & \phi_1 \Delta S_1(\omega_6) + \phi_2^- \Delta S_2(\omega_6) \geq 0, \\ \mathbb{E}[V_2(\phi)|\{\omega_1, \omega_4\}] \geq 0, & \phi_1 \Delta S_1(\omega_1) p_1 + \phi_2^+ \Delta S(\omega_1) p_1 + \phi_1 \Delta S_1(\omega_4) p_4 + \phi_2^- \Delta S(\omega_4) p_4 \geq 0, \\ \mathbb{E}[V_2(\phi)|\{\omega_3, \omega_5\}] \geq 0, & \phi_1 \Delta S_1(\omega_3) p_3 + \phi_2^+ \Delta S(\omega_3) p_3 + \phi_1 \Delta S_1(\omega_5) p_5 + \phi_2^- \Delta S(\omega_5) p_5 \geq 0, \end{cases} \tag{3.3}$$

and at least one of them is strictly positive. As discussed in Section 3.1, one aims at characterizing the absence of statistical $\mathcal{G}$-arbitrage in terms of the existence of a martingale measure $\mathbb{Q}$ whose kernel Z is $\mathcal{G}$-measurable, i.e. Z must satisfy

$$Z(\omega_1) = Z(\omega_4) \text{ and } Z(\omega_3) = Z(\omega_5).$$

In [RRS21, Lemma 2.6] the authors provided a characterization of the absence of statistical $\mathcal{G}$-arbitrage opportunities for the class of trinomial models described in Figure 3.4.1. Then, they constructed an example of a market free of statistical $\mathcal{G}$-arbitrage opportunities and, applying [RRS21, Lemma 2.6], they shown that there is no martingale measure $\mathbb{Q}$ for the price process S with a $\mathcal{G}$-measurable kernel.

However, we show that there is a gap in the computations in the proof of [RRS21, Lemma 2.6] and we propose, an alternative result, Lemma 3.8 below, under which this gap is corrected. In Subsection 3.4.1.1, we study the trinomial model studied in [RRS21, Section 2.2] and we apply Lemma 3.8 to show that EMM($\mathcal{G}$) is not satisfied. Finally, in Subsection 3.4.1.2, we construct explicitly the statistical arbitrage opportunity, proving that equivalence (★) cannot be excluded by this example.

Lemma 3.8. *Let* $\nu_1 := \frac{p_1}{p_4}$ *and* $\nu_2 := \frac{p_3}{p_5}$. *In the trinomial model above, suppose that the following two conditions hold:*

1. $\nu_1 = -\frac{\Delta S_2(\omega_3)}{\Delta S_2(\omega_1)} \nu_2$;
2. $\nu_2 \in [\Gamma_1, \Gamma_2], \quad \Gamma_1 < \Gamma_2$,

where

$$\Gamma_1 := \frac{-\Delta S_1(\omega_5) + \Delta S_2(\omega_5) \frac{\Delta S_1(\omega_6)}{\Delta S_2(\omega_6)}}{\Delta S_1(\omega_3) - \Delta S_2(\omega_3) \frac{\Delta S_1(\omega_2)}{\Delta S_2(\omega_2)}}, \tag{3.4}$$

$$\Gamma_2 := \frac{\frac{\Delta S_1(\omega_6)}{\Delta S_2(\omega_6)}(\Delta S_2(\omega_4) + \Delta S_2(\omega_5)) - (\Delta S_1(\omega_4) + \Delta S_1(\omega_5))}{\Delta S_1(\omega_3) - \Delta S_1(\omega_1)\frac{\Delta S_2(\omega_3)}{\Delta S_2(\omega_1)}}, \tag{3.5}$$

Then, the market is free from statistical $\mathcal{G}$-arbitrage if and only if $\nu_2 \in (\Gamma_1, \Gamma_2)$.

Proof. First, we make some remarks on system (3.3), which determines the existence of statistical $\mathcal{G}$-arbitrage strategies. System (3.3) is equivalent to require that every row of $A\xi \geq 0$ is nonnegative, where

$$A := \begin{pmatrix} \Delta S_1(\omega_2) & \Delta S_2(\omega_2) & 0 \\ \Delta S_1(\omega_6) & 0 & \Delta S_2(\omega_6) \\ \Delta S_1(\omega_1)\nu_1 + \Delta S_1(\omega_4) & \Delta S_2(\omega_1)\nu_1 & \Delta S_2(\omega_4) \\ \Delta S_1(\omega_3)\nu_2 + \Delta S_1(\omega_5) & \Delta S_2(\omega_3)\nu_2 & \Delta S_2(\omega_5) \end{pmatrix} \tag{3.6}$$

and $\xi := (\phi_1, \phi_2^+, \phi_2^-)$. Moreover, at least one row must be strictly positive.

We change the basis of A doing operations on columns, obtaining

$$\begin{aligned} A &\to \begin{pmatrix} 0 & \Delta S_2(\omega_2) & 0 \\ \Delta S_1(\omega_6) & 0 & \Delta S_2(\omega_6) \\ \Big(\Delta S_1(\omega_1) - \frac{\Delta S_2(\omega_1)}{\Delta S_2(\omega_2)}\Delta S_1(\omega_2)\Big)\nu_1 + \Delta S_1(\omega_4) & \Delta S_2(\omega_1)\nu_1 & \Delta S_2(\omega_4) \\ \Big(\Delta S_1(\omega_3) - \frac{\Delta S_2(\omega_3)}{\Delta S_2(\omega_2)}\Delta S_1(\omega_2)\Big)\nu_2 + \Delta S_1(\omega_5) & \Delta S_2(\omega_3)\nu_2 & \Delta S_2(\omega_5) \end{pmatrix}, \\ &\to \begin{pmatrix} 0 & \Delta S_2(\omega_2) & 0 \\ 0 & 0 & \Delta S_2(\omega_6) \\ B_1 & \Delta S_2(\omega_1)\nu_1 & \Delta S_2(\omega_4) \\ B_2 & \Delta S_2(\omega_3)\nu_2 & \Delta S_2(\omega_5) \end{pmatrix} =: \widetilde{A}, \end{aligned} \tag{3.7}$$

where

$$B_1 := \nu_1\Big(\Delta S_1(\omega_1) - \frac{\Delta S_2(\omega_1)}{\Delta S_2(\omega_2)}\Delta S_1(\omega_2)\Big) + \Delta S_1(\omega_4) - \frac{\Delta S_2(\omega_4)}{\Delta S_2(\omega_6)}\Delta S_1(\omega_6). \tag{3.8}$$

$$B_2 := \nu_2\Big(\Delta S_1(\omega_3) - \Delta S_2(\omega_3)\frac{\Delta S_1(\omega_2)}{\Delta S_2(\omega_2)}\Big) + \Delta S_1(\omega_5) - \Delta S_2(\omega_5)\frac{\Delta S_1(\omega_6)}{\Delta S_2(\omega_6)}. \tag{3.9}$$

In particular, to pass from A to $\widetilde{A}$, we use the following change of basis:

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{\Delta S_1(\omega_2)}{\Delta S_2(\omega_2)} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{\Delta S_1(\omega_6)}{\Delta S_2(\omega_6)} & 0 & 1 \end{pmatrix}, \tag{3.10}$$

i.e. $AE_1E_2 = \widetilde{A}$. Since the matrices E_1, E_2 are invertible, the following holds:

$$A\xi \geq 0 \Rightarrow \widetilde{A}(E_1E_2)^{-1}\xi =: \widetilde{A}\widetilde{\xi} \geq 0.$$

We prove now the two implications of the Lemma.

($\Leftarrow$) Suppose that $\nu_2 \in (\Gamma_1, \Gamma_2)$. In order to prove that the model is free of statistical $\mathcal{G}$-arbitrage, we need to show that $\mathrm{Im}(\widetilde{A}) \cap \mathbb{R}^4_+ = \{0\}$, where:

$$\mathrm{Im}\widetilde{A} := \left\{ \alpha \begin{pmatrix} 0 \\ 0 \\ B_1 \\ B_2 \end{pmatrix} + \beta \begin{pmatrix} \Delta S_2(\omega_2) \\ 0 \\ \Delta S_2(\omega_1)\nu_1 \\ \Delta S_2(\omega_3)\nu_2 \end{pmatrix} + \gamma \begin{pmatrix} 0 \\ \Delta S_2(\omega_6) \\ \Delta S_2(\omega_4) \\ \Delta S_2(\omega_5) \end{pmatrix} : \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \in \mathbb{R}^3 \right\}.$$

We assume by contradiction that there exists $v \in \mathrm{Im}\widetilde{A}$ such that $v \geq 0$ and at least one coordinate is strictly positive. We denote with latin numbers the components of the vector $\widetilde{A}\widetilde{\xi}$, where $\widetilde{\xi} = (\alpha, \beta, \gamma)^\top$.

$$\begin{aligned} &(\widetilde{A}\widetilde{\xi})_{\mathrm{I}} \geq 0 \qquad \beta\Delta S_2(\omega_2) = \beta(s^{++} - s_1^+) \geq 0 \Rightarrow \quad \beta \geq 0 \quad (> 0 \Leftrightarrow \beta > 0), \\ &(\widetilde{A}\widetilde{\xi})_{\mathrm{II}} \geq 0 \qquad \gamma\Delta S_2(\omega_6) = \gamma(s_2^{--} - s_1^-) \geq 0 \Rightarrow \gamma \leq 0 \quad (< 0 \Leftrightarrow \gamma < 0); \end{aligned} \tag{3.11}$$

Moreover, since rows $(\widetilde{A}\widetilde{\xi})_{III}$ and $(\widetilde{A}\widetilde{\xi})_{IV}$ are nonnegative, then their sum is nonnegative too:

$$\alpha C_1 + \beta(\Delta S_2(\omega_1)\nu_1 + \Delta S_2(\omega_3)\nu_2) + \gamma(\Delta S_2(\omega_5) + \Delta S_2(\omega_4)) \geq 0, \tag{3.12}$$

where

$$\begin{aligned} C_1 :=& \nu_1\Big(\Delta S_1(\omega_1) - \Delta S_2(\omega_1)\frac{\Delta S_1(\omega_2)}{\Delta S_2(\omega_2)}\Big) + \nu_2\Big(\Delta S_1(\omega_3) - \Delta S_2(\omega_3)\frac{\Delta S_1(\omega_2)}{\Delta S_2(\omega_2)}\Big) \\ &+ \frac{\Delta S_1(\omega_6)}{\Delta S_2(\omega_6)}(-\Delta S_2(\omega_4) - \Delta S_2(\omega_5)) + \Delta S_1(\omega_4) + \Delta S_1(\omega_5). \end{aligned}$$

Then, since by hypothesis $\nu_1 = -\frac{\Delta S_2(\omega_3)}{\Delta S_2(\omega_1)}\nu_2$, the coefficient multiplied by β in (3.12) is zero. Moreover, by hypothesis:

$$\nu_2 < \Gamma_2 =: \frac{N_2}{D_2}, \tag{3.13}$$

where N_2 and D_2 are respectively the numerator and denominator of Γ_2 introduced in (3.5).

Thus, we observe that

$$D_2 := \Delta S_1(\omega_3) - \Delta S_1(\omega_1)\frac{\Delta S_2(\omega_3)}{\Delta S_2(\omega_1)} = (s_1^+ - s_0) - (s_1^+ - s_0)\frac{s_2^{+-} - s_1^+}{s_2^o - s_1^+} > 0.$$

Then,

$$\begin{aligned}C_1 :=& -\frac{\Delta S_2(\omega_3)}{\Delta S_2(\omega_1)}\nu_2\Big(\Delta S_1(\omega_1) - \Delta S_2(\omega_1)\frac{\Delta S_1(\omega_2)}{\Delta S_2(\omega_2)}\Big) + \nu_2\Big(\Delta S_1(\omega_3) - \Delta S_2(\omega_3)\frac{\Delta S_1(\omega_2)}{\Delta S_2(\omega_2)}\Big)\\ &+ \frac{\Delta S_1(\omega_6)}{\Delta S_2(\omega_6)}(-\Delta S_2(\omega_4) - \Delta S_2(\omega_5)) + \Delta S_1(\omega_4) + \Delta S_1(\omega_5)\\ =&\nu_2\Big(-\frac{\Delta S_2(\omega_3)}{\Delta S_2(\omega_1)}\Delta S_1(\omega_1) + \Delta S_1(\omega_3)\Big) + \frac{\Delta S_1(\omega_6)}{\Delta S_2(\omega_6)}(-\Delta S_2(\omega_4) - \Delta S_2(\omega_5))\\ &+ \Delta S_1(\omega_4) + \Delta S_1(\omega_5).\end{aligned} \tag{3.14}$$

Equation (3.14), together with condition (3.13), leads to $C_1 = \nu_2 D_2 - N_2$, so that $C_1 < 0$. Then, condition (3.12) becomes:

$$\alpha \overbrace{C_1}^{<0} + \overbrace{\gamma}^{\leq 0} \overbrace{(\Delta S_2(\omega_5) + \Delta S_2(\omega_4))}^{>0} \geq 0 \Leftrightarrow \alpha \leq 0. \tag{3.15}$$

Finally, by hypothesis:

$$\nu_2 > \frac{-\Delta S_1(\omega_5) + \Delta S_2(\omega_5)\frac{\Delta S_1(\omega_6)}{\Delta S_2(\omega_6)}}{\Delta S_1(\omega_3) - \Delta S_2(\omega_3)\frac{\Delta S_1(\omega_2)}{\Delta S_2(\omega_2)}} = \Gamma_1 =: \frac{N_1}{D_1}.$$

Then, since $D_1 = (s_1^+ - s_0) - (s_2^{+-} - s_1^+)\frac{s_1^+ - s_0}{s_2^{++} - s_1^+} > 0$:

$$B_2 = \nu_2 D_1 - N_1 > 0. \tag{3.16}$$

Studying condition $(\widetilde{A}\widetilde{\xi})_{IV} \geq 0$, we have that:

$$(\widetilde{A}\widetilde{\xi})_{IV} \geq 0 \Leftrightarrow B_2\alpha + \beta\Delta S_2(\omega_3)\nu_2 + \gamma\Delta S_2(\omega_5) \geq 0.$$

Thus:

$$\overbrace{\alpha}^{\leq 0} \overbrace{B_2}^{>0} + \overbrace{\beta}^{\geq 0} \overbrace{\Delta S_2(\omega_3)\nu_2}^{<0} + \overbrace{\gamma}^{\leq 0} \overbrace{\Delta S_2(\omega_5)}^{>0} \leq 0,$$

and equality holds if and only if $\alpha = \beta = \gamma = 0$. In conclusion, absence of statistical $\mathcal{G}$-arbitrage opportunities is guaranteed if $\nu_2 \in (\Gamma_1, \Gamma_2)$.

($\Rightarrow$) It is sufficient to prove that if $\nu_2 \in \{\Gamma_1, \Gamma_2\}$, then it is always possible to build a statistical $\mathcal{G}$-arbitrage. As done in the first part of the proof, we consider a generic vector $\widetilde{\xi} := (\alpha, \beta, \gamma)^\top$ and we denote the components of $\widetilde{A}\widetilde{\xi}$ by latin numbers, where $\widetilde{A}$ is defined in equation (3.7). We are going to determine the values of the components of $\widetilde{\xi}$ such that every row of $\widetilde{A}\widetilde{\xi}$ is nonnegative and at least one row is strictly positive. The first two rows of system $\widetilde{A}\widetilde{\xi} \geq 0$ are independent from the choice of ν_2 and lead to conditions (3.11). There are two possibilities:

(a) If $\nu_2 = \Gamma_2$. By equation (3.14), $C_1 = 0$. Then, a necessary condition for the existence of a statistical $\mathcal{G}$-arbitrage is $(\widetilde{A}\widetilde{\xi})_{III} + (\widetilde{A}\widetilde{\xi})_{IV} \geq 0$, that is (3.12). Thus, under condition $\nu_1 = -\frac{\Delta S_2(\omega_3)}{\Delta S_2(\omega_1)}\nu_2$, inequality (3.12) becomes:

$$\overbrace{C_1}^{=0}\alpha + \overbrace{\gamma}^{\leq 0}\overbrace{(\Delta S_2(\omega_5) + \Delta S_2(\omega_4))}^{>0} \geq 0 \Leftrightarrow \gamma = 0.$$

Therefore, in this case we have no constraints on α but it must hold $\gamma = 0$. Recalling that B_2 is defined in (3.9), we consider condition $(\widetilde{A}\widetilde{\xi})_{IV} \geq 0$:

$$B_2\alpha + \beta\Delta S_2(\omega_3)\nu_2 + \overbrace{\gamma}^{=0}\Delta S_2(\omega_5) \geq 0.$$

Since there are no conditions on α, we can choose:

$$\alpha \geq -\frac{\Delta S_2(\omega_3)\nu_2}{B_2}\beta, \tag{3.17}$$

we can divide by B_2 because since $\nu_2 = \Gamma_2 > \Gamma_1$ condition (3.16) holds.

Finally, we have to find conditions on $\widetilde{\xi}$ such that $(\widetilde{A}\widetilde{\xi})_{III} \geq 0$. First, let us notice that since $\nu_2 = \Gamma_2$, then, since B^1 is defined in (3.8), $0 = C_1 = B_1 + B_2$. By condition $\nu_1 = -\frac{\Delta S_2(\omega_3)}{\Delta S_2(\omega_1)}\nu_2$, the following holds:

$$(\widetilde{A}\widetilde{\xi})_{III} := \alpha B_1 + \Delta S_2(\omega_1)\nu_1\beta + \Delta S_2(\omega_4)\overbrace{\gamma}^{=0} = -\alpha B_2 - \Delta S_2(\omega_3)\nu_2\beta = -(\widetilde{A}\widetilde{\xi})_{IV}.$$

Thus, if we impose the equivalence in condition (3.17) we have that $(\widetilde{A}\widetilde{\xi})_{III} = 0 = (\widetilde{A}\widetilde{\xi})_{IV}$. Finally, let us consider the strategy:

$$\begin{pmatrix}\alpha & \beta & \gamma\end{pmatrix} = \begin{pmatrix}-\frac{\Delta S_2(\omega_3)\nu_2}{B_2}\beta & \beta & 0\end{pmatrix}. \tag{3.18}$$

By (3.11), with $\beta > 0$, the following holds:

$$\begin{aligned}(\widetilde{A}\widetilde{\xi})_{I} &> 0,\\ (\widetilde{A}\widetilde{\xi})_{II} &= 0,\\ (\widetilde{A}\widetilde{\xi})_{III} &= 0,\\ (\widetilde{A}\widetilde{\xi})_{IV} &= 0,\end{aligned}$$

Thus, if $\boldsymbol{\nu_2 = \Gamma_2}$ there exists a statistical $\mathcal{G}$-arbitrage.

(b) $\nu_2 = \Gamma_1 < \Gamma_2$. In this case, condition (3.13), together with (3.14) and (3.15), leads to $C_1 < 0$ and then $\alpha \leq 0$. Moreover, $B_2 = 0$ because condition (3.16) is an equality. This implies that $B_1 = B_1 + B_2 = C_1 < 0$. In that case, we can consider the strategy:

$$\begin{pmatrix}\alpha & \beta & \gamma\end{pmatrix} = \begin{pmatrix}\alpha & 0 & 0\end{pmatrix},$$

with $\alpha < 0$, we have that:

$$\begin{aligned}(\widetilde{A}\widetilde{\xi})_{I} &= \beta\Delta S_2(\omega_2) = 0,\\ (\widetilde{A}\widetilde{\xi})_{II} &= \gamma\Delta S_2(\omega_6) = 0,\\ (\widetilde{A}\widetilde{\xi})_{III} &= \alpha B_1 > 0,\\ (\widetilde{A}\widetilde{\xi})_{IV} &= \alpha B_2 = 0.\end{aligned}$$

Thus, we have a statistical $\mathcal{G}$-arbitrage also in this case.

□

Remark 3.9. Note that in the statement of [RRS21, Lemma 2.6], the absence of statistical $\mathcal{G}$-arbitrage opportunities is equivalent to $\nu_1 = -\frac{\Delta S_2(\omega_3)}{\Delta S_2(\omega_1)}\nu_2$ and $\nu_2 \in (\Gamma_1, \Gamma_2], \quad \Gamma_1 < \Gamma_2$. However, as we have seen in the proof of Lemma 3.8, for $\nu_2 = \Gamma_2$ a statistical $\mathcal{G}$-arbitrage exists.

3.4.1.1 The counterexample

In this subsection, we consider the trinomial model described in [RRS21, Section 2.2]. We show that we can guarantee the existence of a martingale measure for the prices process S, determined by Figure 3.4.1, whose kernel with respect the objective measure $\mathbb{P}$ is $\mathcal{G}$-measurable. As in [RRS21, Section 2.2], we consider a tree as the one described in Figure 3.4.1, where:

$$(s_0, s_1^{+}, s_1^{-}, s_2^{++}, s_2^{+-}, s_2^{--}, s_2^{o}) = (10, 12, 8, 13, 10, 6, 14).$$

The model is arbitrage-free, but it is incomplete. We consider the following filtration:

$$\begin{cases}\mathcal{F}_0 = \{\emptyset, \Omega\},\\ \mathcal{F}_1 = \{\emptyset, \Omega, \{\omega_1, \omega_2, \omega_3\}, \{\omega_4, \omega_5, \omega_6\}\},\\ \mathcal{F}_2 = \mathcal{P}(\Omega).\end{cases}$$

The space of the martingale measures is $\mathcal{M}^e(S) := \{\mathbb{Q} : \ E^{\mathbb{Q}}[S_i|\mathcal{F}_{i-1}] = S_{i-1}, \quad i = 1, 2\}$. Since the first two elements of the filtration are $\mathcal{F}_t = \sigma(S_t)$ for $t = 0, 1$, the condition that a vector

$\mathbb{Q} = (q_1, \dots, q_6) = (\mathbb{Q}(\omega_1), \dots, \mathbb{Q}(\omega_6)) \in \mathbb{R}^6$ has to satisfy to be in $\mathcal{M}^e(S)$ is:

$$\begin{cases} \mathbb{E}^{\mathbb{Q}}[S_2|S_1 = s_1^+] = s_1^+, \\ \mathbb{E}^{\mathbb{Q}}[S_2|S_1 = s_1^-] = s_1^-, \\ \mathbb{E}^{\mathbb{Q}}[S_1|S_0 = s_0] = s_0, \\ \sum_{i=1}^6 q_i = 1, \\ q_i \geq 0, \quad \forall i. \end{cases} \tag{3.19}$$

These conditions are explicitly given by:

$$\begin{aligned} 12 = \mathbb{E}^{\mathbb{Q}}[S_2|S_1 = s_1^+] &= \mathbb{E}^{\mathbb{Q}}[S_2|S_1 = 12] \\ &= 14\mathbb{Q}(S_2 = 14|S_1 = 12) + 13\mathbb{Q}(S_2 = 13|S_1 = 12) + 10\mathbb{Q}(S_2 = 10|S_1 = 12) \\ &= \frac{1}{\mathbb{Q}(\omega_1) + \mathbb{Q}(\omega_2) + \mathbb{Q}(\omega_3)}(14\mathbb{Q}(\omega_1) + 13\mathbb{Q}(\omega_2) + 10\mathbb{Q}(\omega_3)) \\ &= \frac{1}{q_1 + q_2 + q_3}(14q_1 + 13q_2 + 10q_3), \\ 8 = \mathbb{E}^{\mathbb{Q}}[S_2|S_1 = s_1^-] &= \mathbb{E}^{\mathbb{Q}}[S_2|S_1 = 8] \\ &= 14\mathbb{Q}(S_2 = 14|S_1 = 8) + 10\mathbb{Q}(S_2 = 10|S_1 = 8) + 6\mathbb{Q}(S_2 = 6|S_1 = 8), \\ &= \frac{1}{\mathbb{Q}(\omega_4) + \mathbb{Q}(\omega_5) + \mathbb{Q}(\omega_6)}(14\mathbb{Q}(\omega_4) + 10\mathbb{Q}(\omega_5) + 6\mathbb{Q}(\omega_6)) \\ &= \frac{1}{q_4 + q_5 + q_6}(14q_4 + 10q_5 + 6q_6), \\ 10 = \mathbb{E}^{\mathbb{Q}}[S_1|S_0 = s_0] &= \mathbb{E}^{\mathbb{Q}}[S_1|S_0 = 10] \\ &= 12\mathbb{Q}(S_1 = 12|S_0 = 10) + 8\mathbb{Q}(S_1 = 8|S_0 = 10) \\ &= 12\mathbb{Q}(\{\omega_1, \omega_2, \omega_3\}) + 8\mathbb{Q}(\{\omega_4, \omega_5, \omega_6\}) \\ &= 12(q_1 + q_2 + q_3) + 8(q_4 + q_5 + q_6)m \\ 1 &= q_1 + q_2 + q_3 + q_4 + q_5 + q_6. \end{aligned}$$

Solving this system, we find the same results of [RRS21, Section 2.2]:

$$\begin{cases} q_1 = -\frac{3}{4}q_2 + \frac{1}{4}, \\ q_3 = -\frac{1}{4}q_2 + \frac{1}{4}, \\ q_4 = q_6 - \frac{1}{4}, \\ q_5 = -2q_6 + \frac{3}{4}, \\ q_2 \in (0, \frac{1}{3}), \\ q_6 \in (\frac{1}{4}, \frac{3}{8}). \end{cases} \tag{3.20}$$

Under the hypotheses of Lemma 3.8, the conditions under which absence of statistical $\mathcal{G}$-arbitrage holds (i.e. if $\nu_2 \in (\Gamma_1, \Gamma_2)$) are:

$$\begin{aligned} &\nu_1 = -\frac{\Delta S_2(\omega_3)}{\Delta S_2(\omega_1)}\nu_2 = -\frac{10-12}{14-12}\nu_2 = \nu_2, \\ &\nu_2 \in (\Gamma_1, \Gamma_2) = \left(\frac{2}{3}, 3\right). \end{aligned} \tag{3.21}$$

We prove that the model described in the counterexample, on which $\nu_2 = \Gamma_2$ is imposed, allows for a statistical $\mathcal{G}$-arbitrage.

We consider the kernel $Z = (Z_i)_i := \left(\frac{q_i}{p_i}\right)_i$ of $\mathbb{Q}$ with respect to $\mathbb{P} = (p_1, \dots, p_6)$. Z is $\mathcal{G}$-measurable if and only if $Z_1 = Z_4$ and $Z_3 = Z_5$. This condition implies that $\frac{q_1}{q_4} = \nu_1$ and $\frac{q_3}{q_5} = \nu_2$. Adding these conditions to system (3.20) we have:

$$\begin{cases} \nu_1 q_4 = -\frac{3}{4}q_2 + \frac{1}{4}, \\ \nu_2 q_5 = -\frac{1}{4}q_2 + \frac{1}{4}, \\ q_4 = q_6 - \frac{1}{4}, \\ q_5 = -2q_6 + \frac{3}{4}, \end{cases} \rightarrow \begin{cases} -\frac{3}{4\nu_1}q_2 + \frac{1}{4\nu_1} = q_6 - \frac{1}{4}, \\ -\frac{1}{4\nu_2}q_2 + \frac{1}{4\nu_2} = -2q_6 + \frac{3}{4}, \end{cases}$$

By conditions (3.21):

$$\begin{cases} q_2\left(\frac{3}{2\nu_1} + \frac{1}{4\nu_1}\right) = -\frac{1}{4} + \frac{3}{4\nu_1} \rightarrow q_2 = \frac{3-\nu_1}{7}, \\ q_6 = \frac{5\nu_1 - 1}{14\nu_1}. \end{cases}$$

Then, $\nu_1 \in \left(\frac{2}{3}, 3\right)$ implies that $q_2 \in \left(0, \frac{1}{3}\right)$ and $q_6 \in \left(\frac{1}{4}, \frac{1}{3}\right)$. Thus, differently from [RRS21, equation (8)], the existence of a martingale measure $\mathbb{Q}$ whose kernel is $\mathcal{G}$-measurable is allowed.

3.4.1.2 Construction of a statistical arbitrage

We show that it is possible to provide a statistical $\mathcal{G}$-arbitrage for the model studied in [RRS21, Section 2.2]. In particular, $\nu_2 = \Gamma_2$ and $p = (0.15, 0.2, 0.3, 0.05, 0.1, 0.2)$. Using the same notation of the proof of Lemma 3.8, we consider C_1 introduced in (3.14), so that, $C_1 = 0$. Moreover, adopting notation (3.10), the following holds:

$$\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = (E_1 E_2)^{-1} \begin{pmatrix} \phi_1 \\ \phi_2^+ \\ \phi_2^- \end{pmatrix}.$$

We consider the vector (α, β, γ) defined in (3.18). Then:

$$\begin{cases} \gamma = 0, \\ \beta \geq 0, \\ \alpha \geq -\frac{\beta \Delta S_2(\omega_3)\nu_2}{B_2} = \frac{3}{7}\beta. \end{cases}$$

Recalling that the matrix $\widetilde{A}$ is defined on (3.7), hence we notice that

$$\widetilde{A} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -2 \\ -14 & 6 & 6 \\ 14 & -6 & 2 \end{pmatrix} \to \widetilde{A} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} \beta \\ 0 \\ -14\alpha + 6\beta \\ 14\alpha - 6\beta \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{3.22}$$

Choosing $\alpha = 6$ and $\beta = 14$, all the (3.22) is satisfied. Now we compute the two matrices E_1, E_2 and their product:

$$E_1 \cdot E_2 = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{\Delta S_1(\omega_2)}{\Delta S_2(\omega_2)} & 1 & 0 \\ -\frac{\Delta S_1(\omega_6)}{\Delta S_2(\omega_6)} & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

We conclude that:

$$\begin{pmatrix} \phi_1 \\ \phi_2^+ \\ \phi_2^- \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 6 \\ 14 \\ 0 \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \\ -6 \end{pmatrix}. \tag{3.23}$$

At this point, it is easy to check that system (3.3) is explictly given by:

$$\begin{cases} \phi_1 \Delta S_1(\omega_2) + \phi_2^+ \Delta S_2(\omega_2) = 14 > 0, \\ \phi_1 \Delta S_1(\omega_6) + \phi_2^- \Delta S_2(\omega_6) = -12 + 12 = 0, \\ \phi_1 \Delta S_1(\omega_1) p_1 + \phi_2^+ \Delta S(\omega_1) p_1 + \phi_1 \Delta S_1(\omega_4) p_4 + \phi_2^- \Delta S(\omega_4) p_4 = \frac{180}{100} + \frac{60}{100} - \frac{60}{100} - \frac{180}{100} = 0, \\ \phi_1 \Delta S_1(\omega_3) p_3 + \phi_2^+ \Delta S(\omega_3) p_3 + \phi_1 \Delta S_1(\omega_5) p_5 + \phi_2^- \Delta S(\omega_5) p_5 = \frac{360}{100} - \frac{120}{100} - \frac{120}{100} - \frac{120}{100} = 0. \end{cases}$$

Thus, the admissible trading strategy in equation (3.23) is a statistical arbitrage. In conclusion, this example does not disprove the statement of [Bon03]. Indeed, the analysed model, on which a martingale measure $\mathbb{Q}$ whose kernel is $\mathcal{G}$-measurable does not exist, admits statistical $\mathcal{G}$-arbitrage opportunities.

3.4.2 The characterization when Ω is a finite set

In Section 3.4.1 we shown that the counterexample proposed in [RRS21, Section 2.2] is affected by a gap and it does not disprove condition ($\bigstar$) in finite dimensional markets. In this section, we show that condition ($\bigstar$) holds, generalizing [Bon03, Proposition 3] to a rigorous mathematical setting.

We consider a finite probability space $(\Omega, \mathcal{F}, \mathbb{P})$. In particular, there exists $N \in \mathbb{N}$ such that $\Omega := \{\omega_1, \dots, \omega_N\}$. We can assume without loss of generality that $\mathcal{F} = \mathcal{P}(\Omega)$ is the power sigma-algebra. We introduce a discounted price process $S = (S_t)_{t=0}^T$ adapted to a filtration $(\mathcal{F}_t)_{t=0}^T$.

We follow the proof of [HP81, Theorem 2.7] to obtain a characterization of the existence of a martingale measure $\mathbb{Q}$ whose kernel $Z := \frac{d\mathbb{Q}}{d\mathbb{P}}$ is $\mathcal{G}$-measurable:

Theorem 3.10. *Let Ω be a finite set and S a $\mathbb{R}^d$-valued stochastic process on $(\Omega, \mathcal{F}, \mathbb{P})$. Moreover, we consider a sigma-algebra $\mathcal{G} \subseteq \mathcal{F}$. Then,*

$$\text{EMM}(\mathcal{G}) \iff \text{NSA}(\mathcal{G})$$

Proof. The only implication to prove is NSA($\mathcal{G}$) $\Rightarrow$ELMM($\mathcal{G}$) , because EMM($\mathcal{G}$)$\Rightarrow$ NSA($\mathcal{G}$) is a particular case of Theorem 3.6.

We aim at applying the set separation theorem [Bre11, Theorem 1.7] to $K^0_{\mathcal{G}}$, introduced in (3.1) and $L^0_+(\mathcal{F})$. Clearly, $K^0_{\mathcal{G}}$ is convex and closed under the Euclidean topology in $\mathbb{R}^N$. $L^0_+(\mathcal{F})$ is not compact, but we can consider the convex hull P of the unit vectors $(\mathbb{1}_{\{\omega_n\}})_{n=1,\dots,N}$ of $L^\infty_+(\mathcal{F})$:

$$P := \Big\{ \sum_{n=1}^N \mu_n \mathbb{1}_{\{\omega_n\}} \Big| \mu_n \geq 0 : \sum_{n=1}^N \mu_n = 1 \Big\}.$$

By condition $NSA(\mathcal{G})$, it holds that $K^0_{\mathcal{G}} \cap P = \emptyset$. Since P is convex and compact we can apply hyperplane separation theorem, [Bre11, Theorem 1.7], there exists $q \in (L^\infty(\Omega, \mathcal{F}, \mathbb{P}))^* = L^1(\Omega, \mathcal{F}, \mathbb{P}) = \mathbb{R}^N$ and two constants $\alpha < \beta$ such that:

$$\langle q, f \rangle \leq \alpha, \quad \forall f \in K^0_{\mathcal{G}},$$
$$\langle q, f \rangle \geq \beta, \quad \forall h \in P,$$

where $\langle q, f \rangle := \sum_{i=1}^N q_i f_i$.

Since $K_{\mathcal{G}}$ is a linear space we have that $\alpha \geq 0$. Moreover, we can assume $\alpha = 0$. Indeed, if $f \in K^0_{\mathcal{G}}$ and $\lambda > 0$ is an arbitrary real constant, $\lambda f \in K^0_{\mathcal{G}}$ too. Then:

$$\langle q, \lambda f \rangle \leq \alpha \Rightarrow \langle q, f \rangle \leq \frac{\alpha}{\lambda}, \quad \forall \lambda > 0.$$

As a consequence, $\beta > 0$ and, since $\mathbb{1}_{\{\omega_n\}} \in P$ for all $n = 1, \dots, N$:

$$\langle q, \mathbb{1}_{\{\omega_n\}} \rangle = q_n > 0, \quad \forall n = 1, \dots, N,$$

In conclusion, replacing q with $\frac{q}{\sum_{n=1}^N q_i}$, we can construct a measure $\mathbb{Q} := \frac{1}{\sum_{n=1}^N q_i}(q_1, \dots, q_N)$ equivalent to $\mathbb{P}$ such that:

$$\mathbb{E}^{\mathbb{Q}}[f] = \langle q, f \rangle \leq 0, \quad \forall f \in K^0_{\mathcal{G}}. \tag{3.24}$$

Condition (3.24) implies that $\mathbb{E}^{\mathbb{Q}}[E^{\mathbb{P}}[V_T(\phi)|\mathcal{G}]] \leq 0$, for all ϕ. From $\mathbb{Q}$ we can construct a probability measure $\widetilde{\mathbb{Q}}$ that is an EMM($\mathcal{G}$). We observe that:

$$\begin{aligned} 0 &\geq \mathbb{E}^{\mathbb{Q}}[\mathbb{E}^{\mathbb{P}}[V_T(\phi)|\mathcal{G}]] \\ &= \mathbb{E}^{\mathbb{P}}[Z\mathbb{E}^{\mathbb{P}}[V_T(\phi)|\mathcal{G}]] \\ &= \mathbb{E}^{\mathbb{P}}[\mathbb{E}^{\mathbb{P}}[Z|\mathcal{G}]\mathbb{E}^{\mathbb{P}}[V_T(\phi)|\mathcal{G}]] \\ &= \mathbb{E}^{\mathbb{P}}[\mathbb{E}^{\mathbb{P}}[Z|\mathcal{G}]V_T(\phi)], \end{aligned}$$

where $Z := \left(\frac{q_1}{p_1}, \dots, \frac{q_n}{p_n}\right)$. Since $Z(\omega) > 0$ for each $\omega \in \Omega$, we have that:

$$\mathbb{E}^{\mathbb{P}}[Z|\mathcal{G}](\omega_n) := \widetilde{q}_n > 0, \quad n = 1, \dots, N.$$

Then, we can define a new probability measure as follows:

$$\widetilde{\mathbb{Q}} := \frac{1}{\sum_{n=1}^{N}(p_n\widetilde{q}_n)} \begin{pmatrix} p_1\widetilde{q}_1 & \cdots & p_N\widetilde{q}_N \end{pmatrix}.$$

It holds that $\widetilde{\mathbb{Q}} \sim \mathbb{P}$ by construction and

$$\mathbb{E}^{\widetilde{\mathbb{Q}}}[V_T(\phi)] = \mathbb{E}^{\mathbb{P}}[\mathbb{E}^{\mathbb{P}}[Z|\mathcal{G}]V_T(\phi)] \leq 0.$$

Repeating the same computations with $-\phi$ we conclude that $\mathbb{E}^{\widetilde{\mathbb{Q}}}[V_T(\phi)] = 0$ for every ϕ. By [DS06, Lemma 2.2.6], we conclude that $\widetilde{\mathbb{Q}}$ is an EMM($\mathcal{G}$). □

3.5 Characterization of ELMM($\mathcal{G}$) in general probability spaces

The strategy we applied to prove Theorem 3.10 cannot be generalized to a general probability space. Indeed, the hyperplane separation theorem needs topological constraints, such as the closure of the convex cone $K^0_{\mathcal{G}}$ introduced in (3.1). In the infinite-dimensional setting there are delicate topological issues to handle in order to guarantee the conditions necessary to apply the hyperplane separation theorem. Therefore, even if E(L)MM($\mathcal{G}$)⇒NSA($\mathcal{G}$) still holds, it is unclear if the converse implication is true. Our goal is to provide a characterization of E(L)MM($\mathcal{G}$) in the case of a general market model.

In this section, we consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a finite horizon T, a complete and right-continuous filtration $\mathbb{F} := (\mathcal{F}_t)_{t\in[0,T]}$ and an $\mathbb{R}^d$-valued locally bounded semimartingale S representing the price process. We assume that S is adapted to $\mathbb{F}$.

To characterize condition E(L)MM($\mathcal{G}$), we adapt the proof of the Kreps-Yan theorem in the version of [DS06, Theorem 5.2.2] to our setting. First of all, we restrict condition NSA($\mathcal{G}$) to the space $L^\infty(\mathcal{G})$. To do so, we introduce:

$$K_{\mathcal{G}} := K^0_{\mathcal{G}} \cap L^\infty(\mathcal{G}). \tag{3.25}$$

Then, NSA($\mathcal{G}$) implies that $K_{\mathcal{G}} \cap L^\infty_+(\mathcal{G}) = \{0\}$.

Following [DS06, Chapter 5], we restrict the set of admissible strategies to the family of simple strategies. We recall this concept, introduced in [DS06, Definition 5.1.1]:

Definition 3.11. We say that an $\mathbb{R}^d$-valued process $\phi = (\phi_t)_{t=0}^\infty$ is a *simple strategy* if ϕ is of the form:

$$\phi = \sum_{i=1}^n \phi_i \mathbb{1}_{]\tau_{i-1},\tau_i]},$$

where $0 \leq \tau_1 \leq \cdots \leq \tau_n < \infty$ a.s. are finite stopping times and ϕ_i are $\mathcal{F}_{\tau_{i-1}}$-measurable $\mathbb{R}^d$-valued random variables, for some $n \in \mathbb{N}$. Moreover, we say that a simple strategy ϕ is *admissible* if, in addition, the stopped process S^{τ_n} and the random variables ϕ_i are uniformly bounded. We denote the set of simple admissible trading strategies by $\mathcal{A}^{\text{simple}}$.

By Definition 3.11, the value of the portfolio obtained by a simple strategy ϕ at $t \in [0, T]$ is:

$$V_t(\phi) = \sum_{i=1}^n \phi_i \cdot (S_{\tau_i \wedge t} - S_{\tau_{i-1} \wedge t}).$$

In the following, we provide a first characterization of ELMM($\mathcal{G}$) in terms of the closure with respect the weak-$*$ topology of a convex cone containing the restriction of $K_{\mathcal{G}}$ to the conditional expectation of the portfolio values at T obtained using only admissible simple strategies. To do so, we introduce the cone $K^{\text{simple}}_{\mathcal{G}}$ in $L^\infty(\mathcal{F})$:

$$K^{\text{simple}}_{\mathcal{G}} := \{\mathbb{E}[V_T(\phi)|\mathcal{G}] : \ \phi \in \mathcal{A}^{\text{simple}}\}. \tag{3.26}$$

We observe that $K^{\text{simple}}_{\mathcal{G}}$ is defined as the image with respect to the operator $\pi_{\mathcal{G}}$ of the family of portfolio values, obtained from admissible simple strategies. It is convenient to introduce the following convex cone:

$$C^{\text{simple}}_{\mathcal{G}} := K^{\text{simple}}_{\mathcal{G}} - L^\infty_+(\mathcal{G}) = \{\mathbb{E}[V_T(\phi)|\mathcal{G}] - l : \ \phi \in \mathcal{A}^{\text{simple}}, \ l \in L^\infty_+(\mathcal{G})\}. \tag{3.27}$$

$C^{\text{simple}}_{\mathcal{G}}$ is defined as the image with respect $\pi_{\mathcal{G}}$ of the family of the random variables $X \in L^\infty(\mathcal{F})$ for which there exists $\phi \in \mathcal{A}^{\text{simple}}$ such that $X \leq V_T(\phi)$ a.s. In other words, $C^{\text{simple}}_{\mathcal{G}}$ is composed by the conditional expectation with respect to $\mathcal{G}$ of random variables that can be superreplicated by an admissible strategy $\phi \in \mathcal{A}^{\text{simple}}$. Note that

$$C^{\text{simple}}_{\mathcal{G}} \cap L^\infty_+(\mathcal{G}) = \{0\} \tag{3.28}$$

is equivalent to

$$K^{\text{simple}}_{\mathcal{G}} \cap L^\infty_+(\mathcal{G}) = \{0\}, \tag{3.29}$$

The following remark justifies why it is convenient to consider (3.28) instead of (3.29).

Remark 3.12. We aim at applying the Hahn-Banach separation theorem to separate $L^\infty_+(\mathcal{G})$ and the closure with respect to a convenient topology in $L^\infty(\mathcal{G})$ of either $K^{\text{simple}}_{\mathcal{G}}$ or a closed convex cone containing $K^{\text{simple}}_{\mathcal{G}}$. However, it is possible to construct examples which lead to counter intuitive situations. For instance, in [Sch94, Example 3.1], the author described the situation of a convex cone A such that $Cl^*(A) \cap L^\infty_+(\Omega, \mathcal{F}, \mathbb{P}) = \{0\}$, while $B := A - L^\infty_+(\Omega, \mathcal{F}, \mathbb{P})$ satisfies $Cl^*(B) = L^\infty(\mathcal{F})$. Then, replacing $C^{\text{simple}}_{\mathcal{G}}$ with its weak-$*$ closure in (3.28) and $K^{\text{simple}}_{\mathcal{G}}$ with its weak-$*$ closure in (3.29), we obtain that $Cl^*(C^{\text{simple}}_{\mathcal{G}}) \cap L^\infty_+(\mathcal{G}) = \{0\}$ is a stronger condition.

In conclusion, admitting only simple strategies, it is convenient to define the absence of statistical $\mathcal{G}$-arbitrage strategies as:

$$C^{\text{simple}}_{\mathcal{G}} \cap L^\infty_+(\Omega, \mathcal{G}, \mathbb{P}) = \{0\}. \qquad (NSA^s(\mathcal{G}))$$

3.5.1 The characterization of E(L)MM($\mathcal{G}$)

Adapting the proof of the Kreps-Yan theorem [DS06, Theorem 5.2.2], it is possible to prove the equivalence between ELMM($\mathcal{G}$) and the following condition:

$$Cl^*(C^{\text{simple}}_{\mathcal{G}}) \cap L^\infty_+(\mathcal{G}) = \{0\},. \qquad (\text{NFL}(\mathcal{G}))$$

Analogously to “no free lunch” condition introduced in [Kre81], this strengthening of the condition of no-arbitrage is tailor-made to state the following result:

Theorem 3.13. *On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we consider a sigmal-algebra $\mathcal{G} \subseteq \mathcal{F}$ and an $\mathbb{R}^d$-valued locally bounded semimartingale $S := (S_t)_{t\in[0,T]}$. Then:*

$$ELMM(\mathcal{G}) \iff NFL(\mathcal{G}).$$

Proof.

- ELMM($\mathcal{G}$)⇒NFL($\mathcal{G}$) Analogously to Theorem 3.6 this implication is the easiest one. We consider a local martingale measure $\mathbb{Q}$ satisfying ELMM($\mathcal{G}$). By definition, for every couple of stopping times $\sigma_1 < \sigma_2$ a.s., which are localizing for S and every $\phi_1 \in L^\infty(\mathcal{F}_{\sigma_1})$,

$$\mathbb{E}^{\mathbb{Q}}[\phi_1 \cdot (S_{\sigma_2} - S_{\sigma_1})] = 0$$

  Therefore, by linearity $\mathbb{E}^{\mathbb{Q}}[k] = 0$ for every $k \in K^{\text{simple}}_{\mathcal{G}}$. s a consequence, $\mathbb{E}^{\mathbb{Q}}[c] \leq 0$ for any $c \in C^{\text{simple}}_{\mathcal{G}}$. Consider now an element $c \in Cl^*_{\mathcal{G}}(C^{\text{simple}}_{\mathcal{G}})$. By construction, there exists a net $(c_\delta)_{\delta\in I} \subseteq C^{\text{simple}}_{\mathcal{G}}$ such that $\lim_{\delta\in I} c_\delta = c$. We recall that the expectation $f \mapsto \mathbb{E}^{\mathbb{Q}}[f]$ is a continuous operator with respect the weak-$*$ topology. Therefore, we conclude that $\mathbb{E}^{\mathbb{Q}}[c] \leq 0$. By contradiction, assume that $c \geq 0$ a.s. and there exists a set $A \in \mathcal{G}$, with $\mathbb{P}(A) > 0$, such that $c > 0$ on A. This would imply that $\mathbb{E}^{\mathbb{Q}}[c] > 0$, thus, yielding a contradiction.

- NFL($\mathcal{G}$) ⇒ELMM($\mathcal{G}$)
  first step. Let us notice that by NFL($\mathcal{G}$) we can apply the Hahn-Banach theorem ([Rud91,

Theorem 3.5]) on $L^\infty(\Omega, \mathcal{G}, \mathbb{P})$ to separate $Cl^*(C_{\mathcal{G}}^{\text{simple}})$ and every fixed $f \in L_+^\infty(\mathcal{G})$, $f \neq 0$. In particular, there exists an element $Z \in L_+^1(\Omega, \mathcal{G}, \mathbb{P})$ such that

$$\begin{aligned}&\mathbb{E}^{\mathbb{P}}[Zf] > 0,\\ &\mathbb{E}^{\mathbb{P}}[Zh] \leq 0, \quad \forall h \in Cl^*(C_{\mathcal{G}}^{\text{simple}}).\end{aligned}$$

second step. We apply the exhaustion argument to show that the Z we found is strictly positive a.e. in Ω. Let us introduce

$$D := \{d \in L_+^1(\mathcal{G}) \setminus \{0\} : \ \mathbb{E}^{\mathbb{P}}[dh] \leq 0, \ \forall h \in Cl^*(C_{\mathcal{G}}^{\text{simple}})\}.$$

Clearly, $D \neq \emptyset$ since $0 \in D$. We consider the family $\mathcal{S}$ of subsets of Ω defined as:

$$\mathcal{S} := \{\{g > 0\} : \ g \in D\}.$$

$\mathcal{S}$ is closed under countable unions. Indeed, for a sequence $(g_n) \subset D$ we can find strictly positive scalars (α_n) such that $\sum_{n=1}^\infty \alpha_n g_n \in D$ (let us consider for instance $\alpha_n := \frac{1}{2^n \max_{j=1,\dots,n} \mathbb{E}[|g_j|]}$). Therefore, there exists $g_0 \in D$ such that:

$$\mathbb{P}(g_0 > 0) = \sup_{g \in D} \mathbb{P}(g > 0).$$

In particular, $\mathbb{P}(g_0 > 0) = 1$. Indeed, if $\mathbb{P}(g_0 > 0) < 1$, we could apply the first step to $f := \mathbb{1}_{\{g_0 = 0\}}$. This would yield the existence of $g_1 \in D$ such that $\mathbb{E}^{\mathbb{Q}}[fg_1] > 0$. As a consequence, we could have $g_0 + g_1 \in D$, but its support is strictly greater than the one of g_0, thus yielding a contradiction.

Now, normalising g_0 so that $\mathbb{E}^{\mathbb{P}}[g_0] = 1$, we can define a measure $\mathbb{Q}$ through $\frac{d\mathbb{Q}}{d\mathbb{P}} = g_0$, such that:

$$\mathbb{E}^{\mathbb{P}}[g_0 h] \leq 0 \quad \forall h \in K_{\mathcal{G}}^{\text{simple}} \Rightarrow \mathbb{E}^{\mathbb{Q}}[h] = \mathbb{E}^{\mathbb{P}}[g_0 h] = \mathbb{E}^{\mathbb{P}}[g_0 \mathbb{E}^{\mathbb{P}}[V_T(\phi)|\mathcal{G}]] \leq 0, \quad \forall \phi \in \mathcal{A}^{\text{simple}}. \tag{3.30}$$

third step. Since $g_0 \in D \subset L_+^1(\mathcal{G})$, the measure $\mathbb{Q}$ we constructed in the previous step has $\mathcal{G}$-measurable kernel. We prove now that $\mathbb{E}^{\mathbb{Q}}[V_T(\phi)] \leq 0$ for each $\phi \in \mathcal{A}^{\text{simple}}$. From (3.30), we conclude that:

$$0 \geq \mathbb{E}^{\mathbb{Q}}[\mathbb{E}^{\mathbb{P}}[V_T(\phi)|\mathcal{G}]] = \mathbb{E}^{\mathbb{Q}}[\mathbb{E}^{\mathbb{Q}}[V_T(\phi)|\mathcal{G}]] = \mathbb{E}^{\mathbb{Q}}[V_T(\phi)], \quad \forall \phi \in \mathcal{A}^{\text{simple}}. \tag{3.31}$$

Substituting ϕ with $-\phi$ in equation (3.31), we obtain that

$$\mathbb{E}^{\mathbb{Q}}[V_T(\phi)] = 0, \quad \forall \phi \in \mathcal{A}^{\text{simple}}. \tag{3.32}$$

fourth step. Following the proof of [DS06, Lemma 5.1.3] we show now that $E^{\mathbb{Q}}[V_T(\phi)] = 0$

implies that S is a local $\mathbb{Q}$-martingale. Since S is locally bounded, there exists a sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$ increasing to infinity a.s. such that S^{τ_n} is bounded for each $n \in \mathbb{N}$. We prove that S^{τ_n} is a $\mathbb{Q}$-martingale. We introduce two arbitrary stopping times σ_1, σ_2 such that $0 \leq \sigma_1 \leq \sigma_2 \leq \tau^n$ a.s. for some $n \in \mathbb{N}$ and an arbitrary bounded $\mathcal{F}_{\sigma_1}$-measurable random variable h. The strategy $\phi := h\mathbb{1}_{]\sigma_1,\sigma_2]}$ is admissible and simple. For this strategy, (3.32) reduces to

$$0 = \mathbb{E}^{\mathbb{Q}}[h \cdot (S_{\sigma_2} - S_{\sigma_1})], \quad \forall h \in L^\infty(\mathcal{F}_{\sigma_1}),$$

thus, proving the martingale property.

□

Theorem 3.13 states the equivalence between condition ELMM($\mathcal{G}$) and NFL($\mathcal{G}$). Condition NFL($\mathcal{G}$), that is stronger than NSA($\mathcal{G}$), does not have a clear interpretation from an economical point of view, because the characterization of the weak-$*$ closure has to be made in terms of converging nets. We have to deal with converging nets instead of converging sequences (for which an economical interpretation would be much clearer, as discussed in [DS06, Section 5.2]), because $L^\infty(\mathcal{G})$ is not first-countable with respect the weak-$*$ topology induced by $L^1(\mathcal{G})$.

To find a strategy to overcome this problem, we recall the procedure proposed in [DS94] to characterize the absence of classical arbitrage opportunities. Our goal is to find the conditions that the sigma-algebra $\mathcal{G}$ must satisfy in order to exploit the results developed in [DS94] and characterize ELMM($\mathcal{G}$) by a condition that is interpretable from a financial point of view.

We consider an $\mathbb{R}^d$-valued, locally bounded semimartingale S. We introduce the set of admissible strategies $\mathcal{A}$ and define:

$$\begin{aligned} K_0^{\mathcal{A}} &:= \{V_T(\phi) \mid \phi \in \mathcal{A}\}, \\ C_0^{\mathcal{A}} &:= \{k - l \mid k \in K_0^{\mathcal{A}} \text{ and } l \in L^0_+(\mathcal{F})\}. \end{aligned} \tag{3.33}$$

$K_0^{\mathcal{A}}$ is a convex cone in $L^0(\mathcal{F})$. Analogously with Remark 3.12, the absence of arbitrage opportunities (NA) is defined as

$$C^{\mathcal{A}} := C_0^{\mathcal{A}} \cap L^\infty(\mathcal{F}). \tag{3.34}$$

Indeed, (NA) is defined in [DS94, Definition 2.8] by:

$$C^{\mathcal{A}} \cap L^\infty_+(\mathcal{F}) = \{0\}. \tag{NA}$$

As proved in [DS06, Proposition 5.1.7], a sufficient condition for the absence of standard arbitrage opportunities is the existence of a probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$, such that S is a $\mathbb{Q}$-local martingale. This property, that we denote by ELMM, is however not necessary, as shown in the counter example proposed in [DS06, Proposition 5.1.7]. As a consequence, if $\mathcal{A} = \mathcal{A}^{\text{simple}}$, a characterization of ELMM is obtained applying the Kreps-Yan theorem ([Kre81]). In particular, ELMM is equivalent to the so called *no free lunch condition*:

$$Cl^*(C^{\mathcal{A}}) \cap L^\infty_+(\Omega, \mathcal{F}, \mathbb{P}) = \{0\}. \tag{NFL}$$

The authors observed that the difficulties related to the weak-$*$ topology are due to the restriction of the set of admissible strategies to simple strategies. Therefore, they extended the class of admissible strategies considering the set $\mathcal{A}$ introduced in Definition 3.3. In conclusion, the authors proved that ELMM is equivalent to:

$$\overline{Cl}(C^{\mathcal{A}}) \cap L^{\infty}_{+}(\Omega, \mathcal{F}, \mathbb{P}) = \{0\}. \tag{NFLVR}$$

In particular, condition NFLVR deals with the closure of $C^{\mathcal{A}}$ with respect the norm closure of $L^{\infty}(\mathcal{F})$, replacing the more complicated weak-$*$ convergence. The crucial point is that NFLVR implies that $C^{\mathcal{A}}$ is already weak-$*$ closed ([DS94, Theorem 4.2]).

In the context of absence of statistical $\mathcal{G}$-arbitrage opportunities, we notice that the cone $C^{\text{simple}}_{\mathcal{G}}$, introduced in (3.27) is the image of $C^{\mathcal{A}^{\text{simple}}}$ with respect the conditional expectation $\pi_{\mathcal{G}}$. In particular, $NSA^s(\mathcal{G})$ is equivalent to

$$Cl^*(\pi_{\mathcal{G}}(C^{\mathcal{A}^{\text{simple}}})) \cap L^{\infty}_{+}(\mathcal{G}) = \{0\}.$$

By relying on the same reasoning, condition NSA($\mathcal{G}$), restricted to $L^{\infty}(\mathcal{G})$, is equivalent to:

$$\pi_{\mathcal{G}}(C^{\mathcal{A}}) \cap L^{\infty}_{+}(\mathcal{G}) = \{0\}.$$

Then, we introduce the following assumption:

Assumption 3.14. For every $p \in [1, \infty]$, the sigma-algebra $\mathcal{G}$ is such that $\pi_{\mathcal{G}} : L^p(\mathcal{F}) \to L^p(\mathcal{G})$ is a closed map. In other words, for every $C \in L^p(\mathcal{F})$ weak closed, also $\pi_{\mathcal{G}}(C)$ is weak closed on $L^p(\mathcal{G})$. If $p = \infty$, the weak topology has to be replaced by the weak-$*$ topology.

Under Assumption 3.14, we can transfer to $\pi_{\mathcal{G}}(C^{\mathcal{A}})$ the topological properties of $C^{\mathcal{A}}$. In particular, in Section 3.6 and Section 3.7 we are going to exploit Assumption 3.14 to transfer the closure of the cone $C^{\mathcal{A}}$ guaranteed by the results of [DMW90] and [DS94], to its image with respect to $\pi_{\mathcal{G}}$. Then, applying Theorem 3.13, we establish the equivalence between NSA($\mathcal{G}$) and ELMM($\mathcal{G}$). We point out, Assumption 3.14 is a tailor-made condition to apply this procedure. In Appendix 3.A, we prove that $\pi_{\mathcal{G}} : L^p(\mathcal{F}) \to L^p(\mathcal{G})$ is a continuous map with respect the weak topology (if $p = \infty$, the continuity is proved for the weak-$*$ topology). However, it is at present unclear how to characterize the closedness in Assumption 3.14.

Remark 3.15. Let us notice that NSA($\mathcal{G}$) and ELMM($\mathcal{G}$) are equivalent when $\mathcal{G}$ is the sigma-algebra generated by a benchmark set of events $B_1, \dots, B_m \in \mathcal{F}$. Indeed, in this case $L^p(\mathcal{G}) = \mathbb{R}^m$ and every linear subspace of $L^{\infty}(\mathcal{G})$ is closed. Since $\pi_{\mathcal{G}} : L^p(\mathcal{F}) \to \mathbb{R}^m$ is a continuous linear map, $\pi_{\mathcal{G}}(K_0 \cap L^{\infty}(\mathcal{F}))$ is closed in $L^{\infty}(\mathcal{G})$. Hence, we conclude through the same strategy applied in the proof of Theorem 3.10. This reasoning ensures that the equivalence between NSA($\mathcal{G}$) and ELMM($\mathcal{G}$) is guaranteed in the motivating example introduced in Remark 3.5.

3.6 Absence of statistical $\mathcal{G}$-arbitrage opportunities in discrete-time markets

In this Section, we focus on discrete-time models defined on a general probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We adopt Definition 3.2 and we propose two results. Theorem 3.16 states that, under Assumption 3.14, NSA($\mathcal{G}$) is equivalent to EMM($\mathcal{G}$). On the other hand, Theorem 3.19 states the equivalence between NSA($\mathcal{G}$) and EMM($\mathcal{G}$) under an additional assumption.

Theorem 3.16. *In the setting of Definition 3.2, consider a sigma-algebra $\mathcal{G} \subseteq \mathcal{F}$. Then, under Assumption 3.14, the following holds:*

$$\text{EMM}(\mathcal{G}) \iff \text{NSA}(\mathcal{G})$$

Proof. EMM($\mathcal{G}$)$\Rightarrow$ NSA($\mathcal{G}$) is granted by Theorem 3.6.
NSA($\mathcal{G}$)$\Rightarrow$EMM($\mathcal{G}$)
NSA($\mathcal{G}$) guarantees absence of classical arbitrage opportunities that, in the case of a discrete-time market, corresponds to:

$$C_0^{\mathcal{A}} \cap L_+^0(\mathcal{F}) = \{0\}, \tag{NA}$$

where $C_0^{\mathcal{A}}$ is introduced in (3.33) and $\mathcal{A}$ is the class of admissible strategies introduced in Definition 3.2. By [Sch92, Lemma 2.1], the cone $C_0^{\mathcal{A}}$ is closed in $L^0(\Omega, \mathcal{F}, \mathbb{P})$, under the topology of the convergence in probability. Let us notice that NSA($\mathcal{G}$) guarantees that

$$\pi_{\mathcal{G}}(C_0^{\mathcal{A}}) \cap L_+^0(\mathcal{G}) = \{0\}.$$

We then consider the cone

$$C_1^{\mathcal{A}} := C_0^{\mathcal{A}} \cap L^1(\mathcal{F}). \tag{3.35}$$

Since $C_0^{\mathcal{A}}$ is closed in $L^0(\mathcal{F})$, the set $C_1^{\mathcal{A}}$ is closed in the norm topology of $L^1(\mathcal{F})$. We notice the $C_1^{\mathcal{A}}$ is a closed convex cone in $L^1(\mathcal{F})$, then by [Bre11, Theorem 3.7], $C_1^{\mathcal{A}}$ is closed with respect the weak topology, that is $\sigma(L^1(\mathcal{F}), L^\infty(\mathcal{F}))$. By Assumption 3.14, $\pi_{\mathcal{G}}(C_1^{\mathcal{A}})$ is a $\sigma(L^1(\mathcal{G}), L^\infty(\mathcal{G}))$-closed convex cone.

Since $\pi_{\mathcal{G}}(C_1^{\mathcal{A}})$ contains $L_-^1(\mathcal{G})$ and $\pi_{\mathcal{G}}(C_1^{\mathcal{A}}) \cap L_+^1(\mathcal{G}) \subset \pi_{\mathcal{G}}(C_0^{\mathcal{A}}) \cap L_+^0(\mathcal{G}) = \{0\}$, we can apply [DS06, Theorem 5.2.3] to prove the existence of a probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{G})$, such that

$$\begin{cases} \mathbb{Q} \sim \mathbb{P}; \\ \frac{d\mathbb{Q}}{d\mathbb{P}} \in L^\infty(\mathcal{G}); \\ \mathbb{E}^{\mathbb{Q}}[c] \leq 0, \quad \forall c \in \pi_{\mathcal{G}}(C_1^{\mathcal{A}}). \end{cases} \tag{3.36}$$

Finally, let $t = 0, 1 \ldots, T-1$ and $A \in \mathcal{F}_t$. Let us notice that $\phi := \mathbb{1}_A \mathbb{1}_{]t+1,t]} \in \mathcal{A}$, thus $V_T(\phi) = \mathbb{1}_A \cdot (S_{t+1} - S_t) \in C_1^{\mathcal{A}}$. For the same reason, also $-\phi \in \mathcal{A}$, then $V_T(-\phi) = -\mathbb{1}_A \cdot (S_{t+1} - S_t) \in C_1^{\mathcal{A}}$.

By (3.36), we conclude that

$$0 = \mathbb{E}^{\mathbb{Q}}[V_T(\phi)] = \mathbb{E}^{\mathbb{Q}}[\mathbb{1}_A \cdot (S_{t+1} - S_t)], \quad \forall A \in \mathcal{F}_t,$$

thus proving the martingale property. □

We prove now the equivalence NSA($\mathcal{G}$) and EMM($\mathcal{G}$) under the additional condition:

Assumption 3.17. In the framework introduced in Definition 3.2, We suppose that $T = 1$ and $\mathcal{F}_0 \subseteq \mathcal{G}$

This result is based on the adaptation of the [Rog94] To prove it, we apply the following result:

Proposition 3.18. *([KS09, Proposition 2.1.5]) Let $\xi \in L^0(\mathcal{G})$ and let $\mathcal{F}_0 \subseteq \mathcal{G}$. Then, the following are equivalent:*

1. *for every $\phi \in L^0(\mathcal{F}_0)$, inequality $\phi\xi \geq 0$, holds as the equality;*
2. *there exists a bounded random variable $\rho > 0$ such that $\mathbb{E}[\rho|\xi|] < \infty$ and $\mathbb{E}[\rho\xi|\mathcal{F}_0] = 0$.*

Theorem 3.19. *In the setting of Definition 3.2, let $T = 1$. Moreover, let $\mathcal{G}$ be a sigma-algebra satisfying Assumption 3.17. Then,*

$$NSA(\mathcal{G}) \iff EMM(\mathcal{G}).$$

Proof. EMM($\mathcal{G}$)$\Rightarrow$ NSA($\mathcal{G}$) is follows by Theorem 3.6.
NSA($\mathcal{G}$)$\Rightarrow$ EMM($\mathcal{G}$) In this setting, the admissible strategies $\phi \in \mathcal{A}$ are $\mathcal{F}_0$-measurable random variables. A statistical $\mathcal{G}$-arbitrage opportunity, is defined by $\mathbb{E}[\phi\Delta S] > 0$ and $\mathbb{E}[\phi\Delta S|\mathcal{G}] \geq 0$ a.s., where $\Delta S = S_1 - S_0$. First, we notice that NSA($\mathcal{G}$) implies that:

$$\mathbb{E}[\phi\Delta S|\mathcal{G}] \geq 0 \text{ a.s.} \Rightarrow \mathbb{E}[\phi\Delta S|\mathcal{G}] = 0 \text{ a.s.} \tag{3.37}$$

By Assumption 3.17, every $\phi \in L^0(\mathcal{F}_0)$ is $\mathcal{G}$-measurable. As a consequence, (3.37) is equivalent to:

$$\phi\mathbb{E}[\Delta S|\mathcal{G}] \geq 0 \text{ a.s.} \Rightarrow \phi\mathbb{E}[\Delta S|\mathcal{G}] = 0 \text{ a.s.} \tag{3.38}$$

Then, we apply [KS09, Proposition 2.1.5] to $\xi := \mathbb{E}[\Delta S|\mathcal{G}]$. By (3.38), there exists a bounded random variable $\rho > 0$ in $L^0(\mathcal{G})$ such that $\mathbb{E}[\rho|\xi|] < \infty$ and $\mathbb{E}[\rho\xi|\mathcal{F}_0] = 0$. Thus, for any $h \in L^\infty(\mathcal{F}_0)$, $\mathbb{E}[\rho h\xi|\mathcal{F}_0] = 0$. As a consequence, $\mathbb{E}[\rho h\xi] = 0$. Since, ρ is bounded, we can define $\widetilde{\rho} := \frac{\rho}{\mathbb{E}[|\rho|]}$. Hence, $\widetilde{\rho}$ defines an equivalent probability measure $\mathbb{Q}$ such that $\frac{d\mathbb{Q}}{d\mathbb{P}} = \widetilde{\rho}$ and $\mathbb{E}^{\mathbb{Q}}[h\xi] = 0$, for all $h \in L^\infty(\mathcal{F}_0)$. We notice that:

$$0 = \mathbb{E}^{\mathbb{Q}}[h\xi] = \mathbb{E}^{\mathbb{P}}[\widetilde{\rho}h\mathbb{E}^{\mathbb{P}}[\Delta S|\mathcal{G}]] = \mathbb{E}^{\mathbb{P}}[\mathbb{E}^{\mathbb{P}}[\widetilde{\rho}h\Delta S|\mathcal{G}]] = \mathbb{E}^{\mathbb{Q}}[h\Delta S], \quad \forall h \in L^\infty(\mathcal{F}_0),$$

that is sufficient to conclude that S is a $\mathbb{Q}$-martingale. □

In [KS09, Section 2.1.4], the result is generalized inductively to the case of a discrete-time market defined for a time horizon $T > 1$. To apply the same reasoning and generalize Theorem 3.19, we have to assume that $\mathcal{F}_{t-1} \subseteq \mathcal{G}$. However, this condition seems restrictive from a financial point of view.

3.7 Absence of statistical $\mathcal{G}$-arbitrage opportunities in a general semimartingale model

In this section, we consider the general case of a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, on which an $\mathbb{R}^d$-valued, locally bounded, semimartingale S is defined. The set of admissible strategies is $\mathcal{A}$, introduced in Definition 3.3. As in Theorem 3.16, our goal is to apply Assumption 3.14 to guarantee that the topological properties of the cone $C^{\mathcal{A}}$ provided by NFLVR are preserved when we introduce $\pi_{\mathcal{G}}(C^{\mathcal{A}})$.

Theorem 3.20. *In the setting of Definition 3.2, we consider a sigma-algebra $\mathcal{G} \subseteq \mathcal{F}$. Under Assumption 3.14, if NFLVR holds, then:*

$$NSA(\mathcal{G}) \iff ELMM(\mathcal{G}).$$

Proof. We recall that, by NFLVR, the convex cone $C^{\mathcal{A}}$ introduced in (3.34) is $\sigma(L^\infty(\mathcal{F}), L^1(\mathcal{F}))$-closed ([DS94, Theorem 4.2]). By Assumption 3.14, $\pi_{\mathcal{G}}(C^{\mathcal{A}})$ is $\sigma(L^\infty(\mathcal{G}), L^1(\mathcal{G}))$-closed in. Hence, applying [DS06, Theorem 5.2.3], we conclude that there exists a probability measure $\mathbb{Q}$ on $(\Omega, \mathcal{G}, \mathbb{P})$, equivalent to $\mathbb{P}$ such that $\frac{d\mathbb{Q}}{d\mathbb{P}} \in L^1(\Omega, \mathcal{G}, \mathbb{P})$ and $\mathbb{E}^{\mathbb{Q}}[c] \leq 0$, $\forall c \in C^{\mathcal{A}}$. This implies that

$$\mathbb{E}^{\mathbb{Q}}[k] \leq 0, \quad \forall k \in K_0^{\mathcal{A}} \cap L^\infty(\Omega, \mathcal{F}, \mathbb{P}).$$

Mimicking the *fourth step* of Theorem 3.13, we conclude that S is a $\mathbb{Q}$-local martingale. □

Observe that the proof is particularly short, thanks to NFLVR-based theory.

Let us notice that, Theorem 3.16 characterizes the absence of statistical $\mathcal{G}$-arbitrage opportunities NSA($\mathcal{G}$) in terms of ELMM($\mathcal{G}$) under Assumption 3.14. On the other hand, Theorem (3.20) requires the additional condition of NFLVR. The key difference is that in a discrete time market, the closure of the cone $C^{\mathcal{A}}$ is guaranteed by the absence of standard arbitrage opportunities NA, that follows from NSA($\mathcal{G}$). On the other hand, there is nothing that suggests a link between NSA($\mathcal{G}$) and NFLVR.

3.8 Conclusions and further developments

In Theorem 3.10 we confirmed the result described in [Bon03, Proposition 3], showing that, when the price process is defined on a finite probability space, the absence of statistical $\mathcal{G}$-arbitrage opportunities NSA($\mathcal{G}$) is equivalent to EMM($\mathcal{G}$). To generalize this result to more general cases, we proposed a condition, introduced in Assumption 3.14. Assumption 3.14 implies that the image with respect to the conditional expectation $\pi_{\mathcal{G}}$ of a weak closed set in $L^p(\mathcal{F})$ remains weak closed in $L^p(\mathcal{G})$. This condition is tailor-made to exploit the properties provided in [DS94] and prove that NSA($\mathcal{G}$) is equivalent to EMM($\mathcal{G}$).

The first task we should investigate is the existence of conditions on $\mathcal{G}$ under which Assumption 3.14 holds, or at least, when the image with respect $\pi_{\mathcal{G}}$ of the cone $C^{\mathcal{A}}$, introduced in (3.34), is closed. We recall that, in the case of discrete-time markets, the cone $C^{\mathcal{A}}$ defined in (3.35) is already weak closed in $L^1(\mathcal{F})$. This condition is ensured by the absence of standard arbitrage opportunities NA. Moreover, we recall that NA follows from NSA($\mathcal{G}$). Thus, assuming NSA($\mathcal{G}$), the weak closure of $C^{\mathcal{A}}$ in $L^1(\mathcal{F})$ is guaranteed. This reasoning is crucial to prove the equivalence between NSA($\mathcal{G}$) and ELMM($\mathcal{G}$) in the case of a discrete time market model under Assumption 3.14.

In the general semimartingale case, we shown that equivalence between NSA($\mathcal{G}$) and ELMM($\mathcal{G}$) is guaranteed assuming NFLVR, in addition to Assumption 3.14. An interesting extension of this result to delve into is a characterization of NSA($\mathcal{G}$) beyond NFLVR. If we do not assume NFLVR, we cannot exploit the results already provided by [DS94]. Therefore, we should find alternative approaches to guarantee that the image of superreplicable claims with respect the conditional expectation operator is closed in the weak-$*$ topology of $L^\infty(\mathcal{G})$. Another interesting development in this direction could be a NFLVR condition in a statistical sense. In other words, we could search for a condition on the set of admissible strategies $\mathcal{A}$ or on $\mathcal{G}$, under which the image of the cone of super-replicable claims with respect to $\pi_{\mathcal{G}}$ is already weak-$*$ closed in $L^\infty(\mathcal{G})$.

A second interesting problem to investigate is a link between the definition of statistical arbitrage introduced in [Bon03] and the one proposed in [HJTW04]. As discussed in [RRS21, Remark 2.3], the authors noticed that, iterating a statistical $\mathcal{G}$-arbitrage over time, under some stationary condition, one can obtain a statistical arbitrage strategy in the asymptotic sense described in [HJTW04]. It could be interesting to understand if the results proved in [Bon03] can be extended, not only in the direction of market models defined on more general probability spaces, but to also the case of financial markets defined on an infinite time horizon.

Appendix

3.A Properties of conditional expectation

This Appendix is devoted to the description of some properties of the conditional expectation with respect a sigma-algebra $\mathcal{G}$, interpreted as a linear operator on $L^p(\mathcal{F})$.

We recall that $L^\infty(\mathcal{G})$ endowed with the weak-$*$ topology is not first countable. Indeed, the weak topology on a normed space X is first countable if and only if X if finite-dimensional (see [AB06, Theorem 6.26]). Thus, we cannot characterize the weak-$*$ closure of a set A through the limits of sequences in A, but we have to consider the more general concept of converging net. Moreover, in the context of absence of statistical $\mathcal{G}$-arbitrage opportunities, we have to deal also with the conditional expectation operator, that is in general discontinuous with respect the norm topology, as we are going to see in the following example.

Example 3.21. Let $(\Omega, \mathcal{F}, \mathbb{P}) := ([0,1], \mathcal{B}([0,1]), \text{Leb}([0,1]))$. We consider the sigma-algebra $\mathcal{G}$ generated by the set of the form $(0, \frac{1}{2} + q)$ where $q \in \mathbb{Q} \cap [0, \frac{1}{2}]$. We now consider the sequence of functions:

$$f_n(\omega) := 2n\mathbb{1}_{(0,\frac{1}{2n})}(\omega).$$

As a consequence: $\lim_{n\to\infty} f_n(\omega) = 0$ for almost every $\omega \in [0,1]$. On the other hand:

$$\lim_{n\to\infty} \pi_{\mathcal{G}}(f_n) = \lim_{n\to\infty} \mathbb{E}[f_n|\mathcal{G}] = \lim_{n\to\infty} \int_0^1 2n\mathbb{1}_{(0,\frac{1}{2n})}(\omega)d\omega = 1.$$

On the other hand, the conditional expectation $\pi_{\mathcal{G}}$ is a continuous map, when $L^\infty(\mathcal{F})$ and $L^\infty(\Omega, \mathcal{G}, \mathbb{P})$ are both endowed with the weak-$*$ topology. To prove this continuity, we need to introduce some concepts:

Definition 3.22. We consider two topological space (X, τ_X) and (Y, τ_Y). A function $f : X \to Y$ is continuous, if $f^{-1}(U) \in \tau_X$ for every $U \in \tau_Y$. Moreover, a function $f : X \to Y$ is continuous at $x \in X$, if for every $U \in \tau_Y$ such that $f(x) \in U$, then $f^{-1}(U) \in \tau_X$ and $x \in f^{-1}(U)$.

Definition 3.23.

- A direction $\succeq$ on a set D is a reflexive transitive binary relation with the property that each pair has an upper bound, that is:

$$\forall \alpha, \beta \in D \text{ there exists } \gamma \in D \text{ such that } \gamma \succeq \alpha \text{ and } \gamma \succeq \beta.$$

- A *net* in a set X is a function $x : D \to X$, where D is a directed set, that is called *index set* of x.

- We say that a net $(x_\delta)_{\delta \in I}$ in a topological space (X, τ_X) converges to x if for every $V \in \tau_X$, such that $x \in V$, there exists some index δ_V such that $x_\delta \in V$, for all $\delta \succeq \delta_V$. We denote the convergence with the symbol $x_\delta \to x$.

Finally, we recall the following results

Theorem 3.24. *([AB06, Theorem 2.14]) Let us consider a topological space (X, τ_X). Then, the following are equivalent:*

1. *x belongs to the closure of a set C in X;*

2. *there exists a net $(x_\delta)_{\delta \in D}$ in C such that $x_\delta \to x$ in X.*

Theorem 3.25. *([AB06, Theorem 2.28]) Let us consider two topological spaces (X, τ_X) and (Y, τ_Y). For a function $f : X \to Y$ and a point $x \in X$ the following are equivalent:*

1. *The function f is continuous at x;*

2. If a net $(x_\delta)_{\delta\in D}$ converges to x in X, then $(f(x_\delta))_{\delta\in D}$ converges to $f(x)$ in Y.

We are now ready to prove the following proposition.

Proposition 3.26. *Let us consider a probability space $(\Omega,\mathcal{F},\mathbb{P})$ and a sub-sigma-algebra $\mathcal{G}\subseteq\mathcal{F}$. Then, for any $p,q\in[1,\infty]$ such that $\frac{1}{p}+\frac{1}{q}=1$:*

$$\begin{aligned}\pi_{\mathcal{G}}:(L^p(\Omega,\mathcal{F},\mathbb{P}),\sigma(L^p(\mathcal{F}),L^q(\mathcal{F})))&\to(L^p(\Omega,\mathcal{G},\mathbb{P}),\sigma(L^p(\mathcal{G}),L^q(\mathcal{G})))\\ X&\mapsto\pi_{\mathcal{G}}(X):=\mathbb{E}[X|\mathcal{G}].\end{aligned}\tag{3.39}$$

is a well-defined and continuous. Moreover, if $(x_\delta)_{\delta\in I}$ is a net converging to x in $L^p(\mathcal{F})$ such that $(\pi_{\mathcal{G}}(x_\delta))_{\delta\in I}$ converges to $y\in L^p(\mathcal{G})$, then $y=\pi_{\mathcal{G}}(x)$.

Proof.

- The fact that $\pi_{\mathcal{G}}$ is well-defined is a straightforward consequence of Jensen's inequality.
- To prove continuity, we apply Theorem 3.25. In particular, we consider a net $(x_\delta)_{\delta\in I}$ in $L^\infty(\mathcal{F})$ such that $x_\delta\to x\in L^\infty(\mathcal{F})$ with respect the $\sigma(L^p(\mathcal{F}),L^q(\mathcal{F}))$-topology. By definition, $\sigma(L^p(\mathcal{F}),L^q(\mathcal{F}))$-convergence amounts to

$$\lim_{\delta\in I}\mathbb{E}[x_\delta\phi]=\mathbb{E}[x\phi],\quad\forall\phi\in L^q(\mathcal{F}).$$

  By the tower property, for any $\widetilde{\phi}\in L^q(\mathcal{G})$, we have that:

$$\lim_{\delta\in I}\mathbb{E}[\pi_{\mathcal{G}}(x_\delta)\phi]=\lim_{\delta\in I}\mathbb{E}[\mathbb{E}[x_\delta|\mathcal{G}]\widetilde{\phi}]=\lim_{\delta\in I}\mathbb{E}[x_\delta\widetilde{\phi}]=\mathbb{E}[x\widetilde{\phi}]=\mathbb{E}[\mathbb{E}[x|\mathcal{G}]\widetilde{\phi}]=\mathbb{E}[\pi_{\mathcal{G}}(x)\widetilde{\phi}].$$

  This holds for arbitrary x and therefore continuity follows.

- Finally, we show that if $(x_\delta)_{\delta\in I}$ in $L^p(\mathcal{F})$ is a net converging to x under $\sigma(L^p(\mathcal{F}),L^q(\mathcal{F}))$-topology, such that $\pi_{\mathcal{G}}(x_\delta)\to y$ in $L^p(\mathcal{G})$, under $\sigma(L^p(\mathcal{G}),L^q(\mathcal{G}))$-topology, then $y=\pi_{\mathcal{G}}(x)$. By definition $\pi_{\mathcal{G}}(x_\delta)\to y$ is equivalent to:

$$\lim_{\delta\in I}\mathbb{E}[\pi_{\mathcal{G}}(x_\delta)\widetilde{\phi}]=\mathbb{E}[y\widetilde{\phi}],\quad\forall\widetilde{\phi}\in L^q(\Omega,\mathcal{G},\mathbb{P}).$$

  On the other hand:

$$\lim_{\delta\in I}\mathbb{E}[\pi_{\mathcal{G}}(x_\delta)\widetilde{\phi}]=\lim_{\delta\in I}\mathbb{E}[\mathbb{E}[x_\delta|\mathcal{G}]\widetilde{\phi}]=\mathbb{E}[x\widetilde{\phi}],\quad\forall\widetilde{\phi}\in L^q(\mathcal{G}).$$

  For every $G\in\mathcal{G}$, $\mathbb{1}_G\in L^\infty(\Omega,\mathcal{G},\mathbb{P})\subseteq L^q(\mathcal{G})$, for every $q\in[1,+\infty]$. Thus, $\mathbb{E}[y\mathbb{1}_G]=\mathbb{E}[x\mathbb{1}_G]$, for any $G\in\mathcal{G}$. Since $y\in L^p(\mathcal{G})$ and the conditional expectation is unique, it follows that $y=\mathbb{E}[x|\mathcal{G}]=\pi_{\mathcal{G}}(x)$.

□

References

[AB06] C. D. Aliprantis and K. C. Border (2006): *Infinite Dimensional Analysis - A Hitchhiker's Guide*. 3rd ed. Springer Berlin, Heidelberg.

[AB13] F. M. Ametrano and M. Bianchetti (2013): 'Everything you always wanted to know about multiple interest rate curve bootstrapping but were afraid to ask'. *Available at SSRN 2219548*.

[ABBC23] C. Alasseur, M. Basei, C. Bertucci and A. Cecchin (2023): 'A mean field model for the development of renewable capacities'. *Mathematics and Financial Economics* 17(4), 695–719. DOI: `10.1007/s11579-023-00348-6`. URL: `https://doi.org/10.1007/s11579-023-00348-6`.

[ABM20] C. Alasseur, I. Ben Taher and A. Matoussi (2020): 'An extended mean field game for storage in smart grids'. *Journal of Optimization Theory and Applications* 184, 644–670.

[ACP22] R. Aïd, A. Cosso and H. Pham (2022): 'Equilibrium price in intraday electricity markets'. *Mathematical Finance* 32(2), 517–554.

[AH02] F. Angelini and S. Herzel (2002): 'Consistent initial curves for interest rate models'. *Journal of Derivatives* 9(4), 8–17.

[AH05] F. Angelini and S. Herzel (2005): 'Consistent calibration of HJM models to cap implied volatilities'. *Journal of Futures Markets* 25(11), 1093–1120.

[AJ17] A. Aksamit and M. Jeanblanc (2017): *Enlargement of Filtration with Finance in View*. 1st ed. Springer.

[AS94] J.-P. Ansel and C. Stricker (1994): 'Couverture des actifs contingents et prix maximum'. In: *Annales de l'IHP Probabilités et statistiques*. Vol. 30. 2, pp. 303–315.

[BC99] T. Björk and B. J. Christensen (1999): 'Interest rate dynamics and consistent forward rate curves'. *Mathematical Finance* 9(4), 323–348.

[BCR23] P. Bergault, P. Cardaliaguet and C. Rainer (2023): 'Mean Field Games in a Stackelberg problem with an informed major player'. arXiv preprint - arXiv:2311.05229.

[BCY16] A. Bensoussan, M. H. Chau and S. C. Yam (2016): 'Mean field games with a dominating player'. *Applied Mathematics & Optimization* 74, 91–128.

[Ber22] C. Bertucci (2022): 'Mean field games with incomplete information'. arXiv preprint - arXiv:2205.07703.

[BG99] T. Björk and A. Gombani (1999): 'Minimal realizations of interest rate models'. *Finance and Stochastics* 3, 413–432.

[Bil99] P. Billingsley (1999): *Convergence of Probability Measures*. 2nd ed. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc.

[Bjö04] T. Björk (2004): *On the geometry of interest rate models*. Ed. by R. A. Carmona, E. Çinlar, I. Ekeland, E. Jouini, J. A. Scheinkman and N. Touzi. Paris-Princeton Lectures on Mathematical Finance 2003. Lecture Notes in Mathematics. Vol. 1847. Berlin, Heidelberg: Springer, Berlin, Heidelberg.

[BKL01] D. Bertsimas, L. Kogan and A. W. Lo (2001): 'Hedging derivative securities and incomplete markets: An epsilon-arbitrage approach'. *Operations research* 49(3), 372–397.

[BL00] A. E. Bernardo and O. Ledoit (2000): 'Gain, loss, and asset pricing'. *Journal of political economy* 108(1), 144–172.

[BL02] T. Björk and C. Landén (2002): 'On the construction of finite dimensional realizations for nonlinear forward rate models'. *Finance and Stochastics* 6, 303–331.

[BL20] M. Beiglbock and D. Lacker (2020): 'Denseness of adapted processes among causal couplings'. arXiv preprint arXiv:1805.03185.

[BLL20] C. Bertucci, J.-M. Lasry and P.-L. Lions (2020): 'Strategic advantages in mean field games with a major player'. *Comptes Rendus. Mathématique* 358(2), 113–118.

[Bon03] O. Bondarenko (2003): 'Statistical arbitrage and securities prices'. *The Review of Financial Studies* 16(3), 875–919.

[Bre11] H. Brezis (2011): *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Vol. 2. 3. Springer.

[BS01] T. Björk and L. Svensson (2001): 'On the existence of finite-dimensional realizations for nonlinear forward rate models'. *Mathematical Finance* 11(2), 205–243.

[BS06] T. Björk and I. Slinko (2006): 'Towards a general theory of good-deal bounds'. *Review of Finance* 10(2), 221–260.

[BS24] P. Bergault and L. Sánchez-Betancourt (2024): 'A Mean Field Game between Informed Traders and a Broker'. arXiv preprint - arXiv:2401.05257.

[BY21] A. Bensoussan and S. C. P. Yam (2021): 'Mean field approach to stochastic control with partial information'. *ESAIM: Control, Optimisation and Calculus of Variations* 27, 89.

[BY78] P. Brémaud and M. Yor (1978): 'Changes of filtrations and of probability measures'. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 45(4), 269–295.

[Car20] R. Carmona (2020): 'Applications of mean field games in financial engineering and economic theory'. In: *Mean Field Games - Agent Based Models to Nash Equilibria*. Ed. by F. Delarue. Vol. 78. Proceedings of Symposia in Applied Mathematics. American Mathematical Society, pp. 165-218 165–21.

[CCP20] P. Cardaliaguet, M. Cirant and A. Porretta (2020): 'Remarks on Nash equilibria in mean field game models with a major player'. *Proceedings of the American Mathematical Society* 148(10), 4241–4255.

[CD18a] R. Carmona and F. Delarue (2018a): *Probabilistic Theory of Mean Field Games with Application*. 1st ed. Vol. I. Springer Cham.

[CD18b] R. Carmona and F. Delarue (2018b): *Probabilistic Theory of Mean Field Games with Application*. 1st ed. Vol. II. Springer Cham.

[CDL16] R. Carmona, F. Delarue and D. Lacker (2016): 'Mean field games with common noise'. *The Annals of Probability* 44(6), 3740–3803.

[CDL17] R. Carmona, F. Delarue and D. Lacker (2017): 'Mean field games of timing and models for bank runs'. *Applied Mathematics & Optimization* 76, 217–260.

[CFG16] C. Cuchiero, C. Fontana and A. Gnoatto (2016): 'A general HJM framework for multiple yield curve modelling'. *Finance and Stochastics* 20(2), 267–320.

[CGM01] P. Carr, H. Geman and D. B. Madan (2001): 'Pricing and hedging in incomplete markets'. *Journal of financial economics* 62(1), 131–167.

[ČH02] A. Černỳ and S. Hodges (2002): 'The theory of good-deal pricing in financial markets'. In: *Mathematical Finance—Bachelier Congress 2000: Selected Papers from the First World Congress of the Bachelier Finance Society, Paris, June 29–July 1, 2000*. Springer, pp. 175–202.

[CJ19] P. Casgrain and S. Jaimungal (2019): 'Trading algorithms with learning in latent alpha models'. *Mathematical Finance* 29(3), 735–772.

[CJ20] P. Casgrain and S. Jaimungal (2020): 'Mean-field games with differing beliefs for algorithmic trading'. *Mathematical Finance* 30(3), 995–1034.

[CJN12] D. Coculescu, M. Jeanblanc and A. Nikeghbali (2012): 'Default times, no-arbitrage conditions and changes of probability measures'. *Finance and Stochastics* 16, 513–535.

[CS00] J. H. Cochrane and J. Saa-Requejo (2000): 'Beyond arbitrage: Good-deal asset price bounds in incomplete markets'. *Journal of political economy* 108(1), 79–119.

[CZ16] R. Carmona and X. Zhu (2016): 'A probabilistic approach to mean field games with major and minor players'. *The Annals of Applied Probability* 26(3), 1535–1580.

[Dje23a] M. F. Djete (2023a): 'Large population games with interactions through controls and common noise: convergence results and equivalence between open-loop and closed-loop controls'. *ESAIM: Control, Optimisation and Calculus of Variations* 29, 39.

[Dje23b] M. F. Djete (2023b): 'Mean field games of controls: on the convergence of Nash equilibria'. *The Annals of Applied Probability* 33(4), 2824–2862.

[DM78] C. Dellacherie and P. Meyer (1978): *Probabilities and Potential*. Vol. 1. North-Holland mathematics studies. Hermann.

[DMW90] R. C. Dalang, A. Morton and W. Willinger (1990): 'Equivalent martingale measures and no-arbitrage in stochastic securities market models'. *Stochastics.* 29(2), 185–201.

[DS06] F. Delbaen and W. Schachermayer (2006): *The Mathematics of Arbitrage*. Springer.

[DS94] F. Delbaen and W. Schachermayer (1994): 'A general version of the fundamental theorem of asset pricing'. *Mathematische Annalen* 300(1), 463–520.

[DS95] F. Delbaen and W. Schachermayer (1995): 'The no-arbitrage property under a change of numéraire'. *Stochastics and Stochastic Reports* 53(3-4), 213–226.

[EJY00] R. J. Elliott, M. Jeanblanc and M. Yor (2000): 'On models of default risk'. *Mathematical Finance* 10(2), 179–195.

[EPQ97] N. El Karoui, S. Peng and M. C. Quenez (1997): 'Backward stochastic differential equations in finance'. *Mathematical finance* 7(1), 1–71.

[FGGS20] C. Fontana, Z. Grbac, S. Gümbel and T. Schmidt (2020): 'Term structure modelling for multiple curves with stochastic discontinuities'. *Finance and Stochastics* 24(2), 465–511.

[FGHP21] G. Fu, P. Graewe, U. Horst and A. Popier (2021): 'A mean field game of optimal portfolio liquidation'. *Mathematics of Operations Research* 46(4), 1250–1281.

[FH20] G. Fu and U. Horst (2020): 'Mean-field leader-follower games with terminal state constraint'. *SIAM Journal on Control and Optimization* 58(4), 2078–2113.

[Fil99] D. Filipović (1999): 'A note on the Nelson-Siegel family'. *Mathematical Finance* 9(4), 349–359.

[FLM24] C. Fontana, G. Lanaro and A. Murgoci (2024): 'The geometry of multi-curve interest rate models'. arXiv preprint - arXiv:2401.11619.

[FT03] D. Filipović and J. Teichmann (2003): 'Existence of invariant manifolds for stochastic equations in infinite dimension'. *Journal of Functional Analysis* 197(2), 398–432.

[FT04] D. Filipović and J. Teichmann (2004): 'On the geometry of the term structure of interest rates'. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 460(2041), 129–167.

[FT08] D. Filipović and S. Tappe (2008): 'Existence of Lévy term structure models'. *Finance and Stochastics* 12(1), 83–115.

[FT22a] M. Fujii and A. Takahashi (2022a): 'A mean field game approach to equilibrium pricing with market clearing condition'. *SIAM Journal on Control and Optimization* 60(1), 259–279.

[FT22b] M. Fujii and A. Takahashi (2022b): 'Equilibrium price formation with a major player and its mean field limit'. *ESAIM: Control, Optimisation and Calculus of Variations* 28, 21.

[FT22c] M. Fujii and A. Takahashi (2022c): 'Strong Convergence to the Mean Field Limit of a Finite Agent Equilibrium'. *SIAM Journal on Financial Mathematics* 13(2), 459–490.

[FTT10a] D. Filipović, S. Tappe and J. Teichmann (2010a): 'Term structure models driven by Wiener processes and Poisson measures: existence and positivity'. *SIAM Journal on Financial Mathematics* 1(1), 523–554.

[FTT10b] D. Filipović, S. Tappe and J. Teichmann (2010b): 'Jump-diffusions in Hilbert spaces: existence, stability and numerics'. *Stochastics: An International Journal of Probability and Stochastics Processes* 82(5), 475–520.

[FTT20] O. Féron, P. Tankov and L. Tinsi (2020): 'Price formation and optimal trading in intraday electricity markets with a major player'. *Risks* 8(4), 133.

[FTT21] O. Féron, P. Tankov and L. Tinsi (2021): 'Price formation and optimal trading in intraday electricity markets'. In: *Network Games, Control and Optimization: 10th International Conference, NetGCooP 2020, France, September 22–24, 2021, Proceedings 10*. Springer, pp. 294–305.

[Fu23] G. Fu (2023): 'Extended mean field games with singular controls'. *SIAM Journal on Control and Optimization* 61(1), 285–314.

[Fuj23] M. Fujii (2023): 'Equilibrium pricing of securities in the co-presence of cooperative and non-cooperative populations'. *ESAIM: Control, Optimisation and Calculus of Variations* 29, 56.

[GGR23] D. A. Gomes, J. Gutierrez and R. Ribeiro (2023): 'A random-supply mean field game price model'. *SIAM Journal on Financial Mathematics* 14(1), 188–222.

[GL20] C. Gerhart and E. Lütkebohmert (2020): 'Empirical analysis and forecasting of multiple yield curves'. *Insurance: Mathematics and Economics* 95, 59–78.

[GLL10] O. Guéant, J.-M. Lasry and P. L. Lions (2010): 'Mean Field Games and Oil Production'. In: *The Economics of Sustainable Development*. Ed. by Economica. URL: https://hal.science/hal-01393104.

[GNP15] D. A. Gomes, L. Nurbekyan and E. Pimentel (July 2015): *Economic models and mean-field games theory*. ISBN: 978-85-244-0404-7.

[GR15] Z. Grbac and W. J. Runggaldier (2015): *Interest Rate Modeling: Post-Crisis Challenges and Approaches*. 1st ed. Springer Cham.

[GS21] D. A. Gomes and J. Saúde (2021): 'A mean-field game approach to price formation'. *Dynamic Games and Applications* 11(1), 29–53.

[HJTW04] S. Hogan, R. Jarrow, M. Teo and M. Warachka (2004): 'Testing market efficiency using statistical arbitrage with applications to momentum and value strategies'. *Journal of Financial economics* 73(3), 525–565.

[HK79] J. M. Harrison and D. M. Kreps (1979): 'Martingales and arbitrage in multiperiod securities markets'. *Journal of Economic theory* 20(3), 381–408.

[HMC06] M. Huang, R. P. Malhamé and P. E. Caines (2006): 'Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle'.

[HP81] J. M. Harrison and S. R. Pliska (1981): 'Martingales and stochastic integrals in the theory of continuous trading'. *Stochastic processes and their applications* 11(3), 215–260.

[JR00] M. Jeanblanc and M. Rutkowski (2000): 'Modelling of default risk: an overview'. *Mathematical finance: theory and practice*, 171–269.

[JS13] J. Jacod and A. Shiryaev (2013): *Limit Theorems for Stochastic Processes*. Vol. 288. Springer.

[Kal97] O. Kallenberg (1997): *Foundations of Modern Probability*. Vol. 2. Springer.

[KL17] S. Kassberger and T. Liebmann (2017): 'Additive portfolio improvement and utility-efficient payoffs'. *Mathematics and Financial Economics* 11, 241–262.

[Kle13] A. Klenke (2013): *Probability Theory: A Comprehensive Course*. 2nd ed. Springer London.

[KP91] T. G. Kurtz and P. Protter (1991): 'Weak limit theorems for stochastic integrals and stochastic differential equations'. *The Annals of Probability*, 1035–1070.

[Kre81] D. M. Kreps (1981): 'Arbitrage and equilibrium in economies with infinitely many commodities'. *Journal of Mathematical Economics* 8(1), 15–35.

[KS09] Y. Kabanov and M. Safarian (2009): *Markets with Transaction Costs: Mathematical Theory*. Springer.

[KS12] I. Karatzas and S. Shreve (2012): *Brownian Motion and Stochastic Calculus*. Vol. 113. Springer Science & Business Media.

[KS98] I. Karatzas and S. E. Shreve (1998): *Brownian Motion and Stochastic Calculus*. 2nd ed. Springer New York, NY, pp. 47–127.

[Kur91] T. G. Kurtz (1991): 'Random time changes and convergence in distribution under the Meyer-Zheng conditions'. *The Annals of probability*, 1010–1034.

[KW67] H. Kunita and S. Watanabe (1967): 'On square integrable martingales'. *Nagoya Mathematical Journal* 30, 209–245.

[Lac16] D. Lacker (2016): 'A general characterization of the mean field limit for stochastic differential games'. *Probability Theory and Related Fields* 165, 581–648.

[Lan12] S. Lang (2012): *Fundamentals of Differential Geometry*. Vol. 191. Springer.

[Lan19] G. Lanaro (2019): 'The geometry of interest rates in a post-crisis framework'. MA thesis. Università degli Studi di Padova.

[LBŠ+18] M. Lazzarino, J. Berrill, A. Šević et al. (2018): 'What is statistical arbitrage?' *Theoretical Economics Letters* 8(05), 888.

[Led95] O. R. H. Ledoit (1995): 'Essays on risk and return in the stock market'. PhD thesis. Massachusetts Institute of Technology.

[LL07] J.-M. Lasry and P.-L. Lions (2007): 'Mean field games'. *Japanese journal of mathematics* 2(1), 229–260.

[LM19] C.-A. Lehalle and C. Mouzouni (2019): 'A mean field game of portfolio trading and its consequences on perceived correlations'. arXiv preprint - arXiv:1902.09606.

[MSZ18] J. Ma, R. Sun and Y. Zhou (2018): 'Kyle-Back equilibrium models and linear conditional mean-field SDEs'. *SIAM Journal on Control and Optimization* 56(2), 1154–1180.

[MZ84] P. A. Meyer and W. Zheng (1984): 'Tightness criteria for laws of semimartingales'. In: *Annales de l'IHP Probabilités et Etatistiques*. Vol. 20. 4, pp. 353–372.

[NS87] C. R. Nelson and A. F. Siegel (1987): 'Parsimonious modeling of yield curves'. *Journal of Business* 60(4), 473–489.

[Nuñ17] G. Nuño (2017): 'Optimal social policies in mean field games'. *Applied Mathematics & Optimization* 76, 29–57.

[Rog94] L. C. G. Rogers (1994): 'Equivalent martingale measures and no-arbitrage'. *Stochastics: An International Journal of Probability and Stochastic Processes* 51(1-2), 41–49.

[RRS21] C. Rein, L. Rüschendorf and T. Schmidt (2021): 'Generalized statistical arbitrage concepts and related gain strategies'. *Mathematical Finance* 31(2), 563–594.

[Rud91] W. Rudin (1991): *Functional Analysis*. Second Edition. International series in pure and applied mathematics. Mcgraw-Hill inc.

[SC16] N. Sen and P. E. Caines (2016): 'Mean field game theory with a partially observed major agent'. *SIAM Journal on Control and Optimization* 54(6), 3174–3224.

[Sch92] W. Schachermayer (1992): 'A Hilbert space proof of the fundamental theorem of asset pricing in finite discrete time'. *Insurance: Mathematics and Economics* 11(4), 249–257.

[Sch94] W. Schachermayer (1994): 'Martingale measures for discrete-time processes with infinite horizon'. *Mathematical Finance* 4(1), 25–55.

[SFJ22] A. V. Shrivats, D. Firoozi and S. Jaimungal (2022): 'A mean-field game approach to equilibrium pricing in solar renewable energy certificate markets'. *Mathematical Finance* 32(3), 779–824.

[Sli10] I. Slinko (2010): 'On finite dimensional realizations of two-country interest rate models'. *Mathematical Finance* 20(1), 117–143.

[Sta04] J. Staum (2004): 'Fundamental theorems of asset pricing for good deal bounds'. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics* 14(2), 141–161.

[Tap10] S. Tappe (2010): 'An alternative approach on the existence of affine realizations for HJM term structure models'. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 466(2122), 3033–3060.

[Tap12] S. Tappe (2012): 'Existence of affine realizations for Lévy term structure models'. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 468(2147), 3685–3704.