University of Trento

Department of Mathematics

Doctor of Philosophy in Mathematics
XXXVI Cycle

SDEs and MFGs towards Machine Learning applications

Supervisor
Prof. Luca Di Persio

PhD Student
Matteo Garbelli

Academic Year 2022–2023

Committee members:

Luca Di Persio, Associate Professor, Università di Verona
Markus Fischer, Associate Professor, Università di Padova
Marco Patacca, Researcher, Università di Perugia

Trento, 04 December 2023

"You get pseudo-order
when you seek order;
you only get a measure of order and control
when you embrace randomness."
N.N. Taleb

Contents

Introduction

The term *stochastic* from the Greek word $\sigma\tau\acute{o}\chi o\varsigma$ (guess, aim, target) finds its linguistic roots in the expression *aiming at guess* or *being able at conjecturing*. Its etymology captures the essence of unpredictability inherent in those natural and human phenomena characterized by a probabilistic nature. Over time, this concept has evolved into a cornerstone of mathematical analysis, an indispensable tool for comprehending complex systems that bear the imprint of randomness.

The principal objective of this dissertation is to explore the applications of stochastic analysis across diverse domains, all of which are intrinsically imbued with elements of uncertainty and stochasticity. This work endeavours to address critical questions that have surfaced in recent years within various subfields of stochastic analysis. The dissertation is organized into three major sections, each concentrating on a distinct yet interrelated area of research: the initial segment focuses on *Time-Delayed Backward Stochastic Differential Equations*; the subsequent portion is devoted to the study of *Mean Field Games* and the associated *Master Equation*; the final segment examines the utilization of *Machine Learning techniques in financial and economic decision-making processes.*

In recent years, the study of models incorporating temporal dependence on antecedent states of the solution has gained prominence. Consequently, Stochastic Delay Differential Equations (SDDEs) are formulated as a natural extension of conventional Stochastic Differential Equations (SDEs), allowing the coefficients to be functions of both the current and past states. This work concentrates on a specialized subclass of Backward Stochastic Differential Equations (BSDEs) featuring time-delayed coefficients. In such equations, the generator function at a given time t can be influenced by the historical values of the solution.

The main references for the backward time delay framework rely on the work by Lukas Delong and Peter Imkeller [76, 74, 75, 77, 78]. BSDEs represent a cornerstone of stochastic modelling with applications ranging from finance to physics. Since many financial problems can be related to stochastic optimization problems with target terminal conditions, BSDEs have become an active field of research stimulated by new financial and actuarial applications.

For example, BSDEs might be used to define nonlinear expectations or to solve stochastic optimal control problems by means of a probabilistic representation of solutions to Partial Differential Equations (PDEs).

The connection between probability theory and PDEs is a widely analysed subject, synthetized by the well-known Feynman-Kac formula ([135]) which states that solutions for a large class of second order PDEs of both elliptic and parabolic type, can be expressed in terms of expectation of a diffusion process. One of the first fundamental results has been formalized by Pardoux and Peng in [177, 178], which have shown the connection between BSDEs and a system of semi-linear PDEs, proving the nonlinear Feynman-Kac formula within the Markovian setting. Concerning the non-Markovian scenario, we know from [184, 185] that a nonlinear Feynman-Kac formula can be still established associating a path-dependent PDE to a non-Markovian BSDE. More recently, the introduction of horizontal and vertical derivatives of non-anticipative functionals on path spaces by Dupire [92] and by Cont and Fournié [61, 62, 63] facilitated the formulation of a new class of path dependent PDEs and the introduction of the so-called *viscosity solution* concept, see [94, 95, 185], for more details. For a complete overview of stochastic calculus with delay we refer to [170], [171] and [219].

We continue our analysis by focusing on Mean Field Games (MFGs). MFG theory was introduced in 2006 by Lasry and Lions ([148, 149, 150, 151]) to describe the asymptotic behaviour of differential games with a large number of players, also called agents. A similar definition was given in the same years by Caines, Huang and Malhamé [129]. The term *mean field* refers to the highly symmetric form of interaction and the indistinguishability of the players whose number is usually denoted by N. The solution of MFG corresponds to a Nash equilibrium, i.e. a configuration where each player's strategy is optimal if she/he has no benefit to deviate whereas all the other players follow a fixed strategies profile. A formal description of this limit problem is obtained by analyzing the so-called *MFG system*, namely a forward-backward PDEs system: a forward Fokker-Planck equation that describes the evolution of the distribution of players in the state space and a backward Hamilton-Jacobi-Bellman equation that solves the optimal control problem. An equivalent approach deals with a *Forward-Backward SDEs* (FBSDEs) system used to represent the coupled dynamics of the state, control, and cost processes for each player in the system. The forward equation describes the dynamics of the state variable for a representative player while the BSDE is derived from the player's cost function and the related value function. A third approach is based on a non-local transport PDE defined on the space of probability measures, the so-called *Master Equation*. It takes into account both consistency and optimal strategy in one single equation. The MFG system has been widely studied, and most of the usual topics as the existence, uniqueness and stability of solutions have been proved in many different frameworks. Conversely, results about the convergence problem, namely when the N-players differential game approaches the MFG system, are considerably fewer. In particular, most of the literature mainly considers the case where the players' state space is the torus $\mathbb{T}^d$, hence a periodic framework, or the whole space $\mathbb{R}^d$. Nevertheless when dealing with *real world problems* as those related to biological or financial models, see, e.g., [1], [36], [35], [34] [119], boundary conditions become unavoidable characteristics we have to deal with, being obliged to consider non periodic settings in bounded domains. Thus, we consider Dirichlet boundary conditions in order to model absorption, i.e. players leave the game when touching boundaries of the state space (assumed to be a smooth domain of the Euclidean space). We analyze the related Master Equation in a bounded domain with homogeneous Dirichlet conditions, addressing the convergence problem of the Nash equilibria towards the solution of the Master Equation.

Thereafter, we focus on a mean-field approach that extends the game theory setting beyond competition. Given that MFGs involve competitive interactions among multiple agents, we delve into a complementary perspective known as Mean Field Control (MFC). MFC [47] confronts problems from a cooperative setting, where a single agent optimally adapts its actions based on the expected behaviour of a large population. The objective of MFC is to optimize the overall system's performance by adjusting the agent's controls. Following this intuition, we present a parallelism between MFC problem and the supervised learning paradigm for Deep Neural Networks (DNNs) in order to provide a mathematical framework for supervised learning in a high-dimensional scenario. In the supervised learning paradigm, an algorithm learns from labelled training data to make predictions or decisions. The algorithm is trained by input-output pairs, where the input is the data or features, and the output is the corresponding target or label. The goal of supervised learning is to learn a mapping from inputs to outputs, allowing the algorithm to make predictions on new, unseen data by training a parametric Machine Learning (ML) model, such as a Neural Network (NN), whose parameters are optimized to minimize a chosen loss function. Thus, we analyze a Mean Field Optimal Control Problem [93] where the expected loss is computed over the joint distributions of input-target pairs. Finally, we formalize a learning process in terms of the so-called Mean Field Optimal Transport, introduced in [13], where the mass/information is transported by agents that cooperate to minimize a loss function by selecting optimal controls/parameters.

At last, we develop a couple of ML algorithms for decision-making problems by working on two distinct learning paradigms. First, we construct different Recurrent Neural Networks (RNNs), a class of NNs specific for working on time series data that we use to make predictions: we apply the supervised scheme in the context of time series forecasting, facilitated by a dynamic rolling window mechanism. We deal with a classical problem in modern portfolio theory: following a historical perspective, we predict realized volatility to implement an allocation strategy known as the Risk Parity strategy where all the assets are constrained to contribute in the same proportion to the overall risk of the portfolio. Second, we focus on Reinforcement Learning (RL), i.e. a type of ML where an agent learns to make a sequence of decisions in an environment to maximize a cumulative reward. Without explicitly labelled data, the agent explores the environment through a trial and error procedure while receiving feedback for any taken action. We refer to the monograph [204] for a complete technical overview and for mathematical details. We deploy a single RL algorithm with continuous control [155] in the electricity market: we develop a single-agent RL algorithm to model the optimal behaviour in terms of offering curve of an operator acting on the day-ahead auction market.

The thesis consists of six chapters, corresponding to six different papers. Each chapter represents a self-contained and autonomous paper. Hence, the notation changes chapter by chapter.

Part I: BSDEs with time delayed generators

The first part represents an independent part of the thesis and concerns the study of BSDEs with time–delayed coefficients. Along this research line, our results are applied to two different scenarios.

Chapter 1 is based on [87]. Within the latter, we formulate a Feynman-Kac representation theorem for a class of BSDEs endowed with time-delayed generators and subject to jumps, which are intrinsically tied to path-dependent, nonlinear Kolmogorov equations. Precisely, our focus lies on a system of FBSDEs that are influenced by a time-delayed generator and driven by Lévy-type noise. We derive a nonlinear Feynman-Kac representation formula that correlates the solution emanating from the FBSDE framework with the solution to a path-dependent, nonlinear Kolmogorov equation, characterized by both temporal delays and jump discontinuities. The theoretical constructs are subsequently applied to extend the paradigmatic Large Investor Problem, incorporating stock price dynamics governed by a jump-diffusion model.

Chapter 2 deals with [83]. Here, we establish the existence and uniqueness of solutions for BSDEs featuring time-delayed generators, specifically within the regime of infinitesimal delays or, equivalently, small Lipschitz constants. Additionally, we demonstrate the continuity of these solutions with respect to the associated increasing process, requiring merely uniform convergence without necessitating convergence in variation. Furthermore, we extend the existence theorem to encompass scenarios with arbitrary delay intervals by imposing conditions of monotonicity and linearity upon the generators. As a concrete instantiation of our theoretical framework, we furnish an application pertinent to actuarial science.

Part II: Mean Field Game theory

In the second part, we focus on the context of strategic interactions and collective decision-making, dealing with MFGs theory.

Chapter 3 is devoted to [85]. We generalize the theory developed in [41]: the focal point is the study of a non-local transport equation, called the *Master Equation* whose trajectories represent the solutions of the MFG system. We analyze the Master Equation and the related convergence problem considering a bounded domain with homogeneous Dirichlet conditions. Concerning the N-players differential game, the player's dynamic ends when touching the boundary. We analyze the well-posedness of the Master Equation and the regularity of its solutions for a suitable class of parabolic equations. Such results are then exploited to consider the convergence problem of the Nash equilibria towards the Master Equation solution, also proving the convergence of optimal trajectories. Eventually, we apply our findings to solve a toy model related to an optimal liquidation scenario.

Chapter 4 moves from [81]. We sketch an unified perspective for Optimal Transport (OT) and Mean Field Control (MFC) theories to analyse the learning process for Neural Networks algorithms in a high-dimensional framework. We consider a Mean Field Neural Network in the context of MFC theory referring to the mean field formulation of OT theory that may allow the development of efficient algorithms in a high dimensional framework while providing a powerful tool in the context of explainable Artificial Intelligence.

Part III: Machine Learning Applications in Finance and Economics

Our exploration expands into the application of Machine Learning (ML) algorithms in finance and economics, focusing on two data-driven algorithms for decision making problems.

Chapter 5 focuses on [84]. We present a hybrid method for computing volatility forecasts that can be used to implement a risk-controlled strategy for a multi-asset portfolio consisting of both US and international equities. Recent years have been characterized by extremely low yields, with 2022 marked by rising interest rates and an increasing inflation rate. These factors produced new challenges for both private and institutional investors, including the need for robust forecast methods for financial assets' volatilities. Addressing such a task, we focus on a hybrid solution that combines classical statistical models with specific classes of

Recurrent Neural Networks (RNNs). In particular, we first use the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) approach within the preprocessing phase to capture volatility clustering, striking an efficient balance between computational effort and accuracy, to then apply RNN architectures to maximise performances of volatility forecasts later used as input factors for risk-controlled investment strategies. In terms of portfolio allocation, we focus on a simplified version of the Risk Parity method that was first proposed by the Research Division of S&P Global in [20]. This version ignores the contribution of cross-correlations among assets, nevertheless providing encouraging results. Indeed, we show the effectiveness of the chosen approach by providing forward-looking risk parity portfolio strategies that outperform standard risk/return portfolio structures.

Chapter 6 is based on [82]. In the context of electricity market auctions, we address the stochastic optimal control problem utilizing a single-agent Reinforcement Learning paradigm. Predominantly, energy transactions occur within the framework of day-ahead markets. In these markets, both purchasers and vendors of energy submit their bids for specified future intervals, detailing both the quantity of energy they seek to exchange and the associated financial terms. Notably, existing literature on energy market dynamics often omits the temporal evolution of such markets, concentrating instead on predictive models that emulate the terminal pricing structure emergent from supply-demand equilibria. Our investigation is devoted to the formulation of an optimized bidding strategy for energy sellers, leveraging a single-agent Reinforcement Learning algorithm, specifically the Deep Deterministic Policy Gradient (DDPG) method, operational over a continuous action space. Within this computational model, the objective function maximizes the expected future returns by modulating the bid curve as a control action. The state space is constituted by historical data, encapsulating variables such as past pricing, production overhead, and generation capabilities across both renewable and non-renewable energy sources. Through iterative learning, the agent refines its bidding strategy to attain increasingly efficient market participation, thereby optimizing its long-term payoff.

Part I

BSDEs with time delayed generators

1 Feynman-Kac formula for BSDEs with jumps and time delayed generators associated to path-dependent nonlinear Kolmogorov equations

1.1 Introduction

In this chapter, we analyze a stochastic process described by a system of Forward-Backward Stochastic Differential Equations (FBSDEs). The forward path-dependent equation is driven by a Lévy process, while the backward one presents a path-dependent behaviour with dependence on a small delay δ. We establish a non-linear Feynman–Kac representation formula associating the solution of the latter FBSDE system to the one of a path-dependent nonlinear Kolmogorov equation with delay and jumps. In particular, we prove that the stochastic system allows to uniquely construct a solution to the parabolic Partial Differential Equation (PDE), in the spirit of the Pardoux-Peng approach, see [177, 178], i.e.:

$$\begin{cases} -\partial_t u(t,\phi) - \mathcal{L}u(t,\phi) - f\big(t,\phi,u(t,\phi),\partial_x u(t,\phi)\sigma(t,\phi),\mathcal{J}u(t,\phi),(u(\cdot,\phi))_t\big) = 0\,, \\ u(T,\phi) = h(\phi)\,, \qquad\qquad (t,\phi)\in[0,T]\times\Lambda\,, \end{cases} \tag{1.1}$$

$T<\infty$ being a fixed time horizon and Λ being $D([0,T];\mathbb{R}^d)$, i.e. the space of càdlàg $\mathbb{R}^d$-valued functions defined on the interval $[0,T]$.

The integro–differential operator $\mathcal{J}$ is associated with the jump behaviour and is defined as

$$\mathcal{J}u(t,\phi) := \int_{\mathbb{R}\setminus\{0\}} \Big[u\big(t,\phi^{t,\gamma(t,\phi,z)}\big) - u(t,\phi)\Big]\,\lambda(z)\nu(dz)\,, \tag{1.2}$$

where $\phi^{t,\gamma}$ models the vertical perturbation of the path ϕ defined by

$$\phi^{t,\gamma}(\theta) := \phi(\theta)\mathbb{1}_{\theta\in[0,t)} + [\phi(t)+\gamma]\,\mathbb{1}_{\theta\in[t,T]}\,, \qquad \theta,t\in[0,T]\,,\gamma\in\mathbb{R}^d\,, \tag{1.3}$$

with $\gamma : [0,T] \times \Lambda \times \{\mathbb{R} \setminus \{0\}\} \to \mathbb{R}^d$ being a continuous, non-anticipative functional encoding the shifting at the (right) endpoint while λ models the intensity of the jumps and ν represents the associated Lévy measure.

The second order differential operator $\mathcal{L}$, associated to the diffusion, is defined by

$$\begin{aligned}\mathcal{L}u(t,\phi) :=& \frac{1}{2} Tr\left[\sigma(t,\phi)\sigma^*(t,\phi)\partial^2_{xx}u(t,\phi)\right] + \langle b(t,\phi), \partial_x u(t,\phi)\rangle \\ &+ \int_{\mathbb{R}\setminus\{0\}} \Big(u(t,\phi^{\gamma(t,\phi,z)}) - u(t,\phi) - \partial_x u(t,\phi)\gamma(t,\phi,z)\Big)\,\nu(dz)\,,\end{aligned} \tag{1.4}$$

with $b : [0,T] \times \Lambda \to \mathbb{R}^d$ and $\sigma : [0,T] \times \Lambda \to \mathbb{R}^{d\times l}$ being two non-anticipative functionals.

Also, for a fixed delay $\delta > 0$ we set

$$(u(\cdot,\phi))_t := (u((t+\theta)^+,\phi))_{\theta\in[-\delta,0]}\,. \tag{1.5}$$

Let us underline that the study of a path-dependent Kolmogorov equation whose generator f depends on both a delayed term $(u(\cdot,\phi))_t$ and on a jump operator $\mathcal{J}u(t,\phi)$, represents the main novelty we provide in this chapter.

Under appropriate assumptions on the coefficients, the deterministic non-anticipative functional $u : [0,T] \times \Lambda \to \mathbb{R}$ given by the representation formula

$$u(t,\phi) := Y^{t,\phi}(t)\,, \tag{1.6}$$

is a mild solution of the Kolmorogov Equation (1.1).

Concerning the stochastic process $Y^{t,\phi}(t)$ in Eq. (1.6), provided conditions in Assumptions 1.2 - 1.4 are fulfilled, we prove that the quadruple

$$\left(X^{t,\phi}, Y^{t,\phi}, Z^{t,\phi}, U^{t,\phi}\right)_{s\in[t,T]}\,,$$

is the unique solution of the system of FBSDEs on $[t,T]$ given by

$$\begin{cases} X^{t,\phi}(s) = \phi(t) + \int_t^s b(r, X^{t,\phi})dr + \int_t^s \sigma(r, X^{t,\phi})dW(r) \\ \qquad\qquad + \int_t^s \int_{\mathbb{R}\setminus\{0\}} \gamma(r, X^{t,\phi}, z)\tilde{N}(dr,dz) \\ Y^{t,\phi}(s) = h(X^{T,\phi}) + \int_s^T f(r, X^{t,\phi}, Y^{t,\phi}(r), Z^{t,\phi}(r), \tilde{U}^{t,\phi}(r), Y^{t,\phi}_r)dr \\ \qquad\qquad - \int_s^T Z^{t,\phi}(r)dW(r) - \int_s^T \int_{\mathbb{R}\setminus\{0\}} U^{t,\phi}(r,z)\tilde{N}(dr,dz)\,, \end{cases} \tag{1.7}$$

where W stands for a l-dimensional standard Brownian motion. Assuming a delay $\delta \in \mathbb{R}^+$, the notation $Y_r^{t,\phi}$, appearing in the generator f of the backward dynamic in the system (1.7), stands for the delayed path of the process $Y^{t,\phi}$ restricted to $[r-\delta, r]$, namely

$$Y_r^{t,\phi} := \left(Y^{t,\phi}((r+\theta)^+)\right)_{\theta\in[-\delta,0]}. \tag{1.8}$$

In Eq. (1.7), $\tilde{N}$ models a compensated Poisson random measure, independent from W, with associated Lévy measure ν. This stochastic term appears also in the definition of the integral term $\tilde{U} : \Omega \times [0,T] \to \mathbb{R}$ related to the jump process by

$$\tilde{U}^{t,\phi}(r) = \int_{\mathbb{R}\setminus\{0\}} U^{t,\phi}(r,z)\lambda(z)\nu(dz); \tag{1.9}$$

we need it in order to introduce to express the solution of (1.1) via the FBSDE system (1.7).

A connection between time-delayed BSDEs and path-dependent PDEs has been proved by an *infinite dimensional lifting* approach in [107] and [163]. In [11], the authors consider a BSDE driven by a Brownian motion and a Poisson random measure that provides a viscosity solution of a system of parabolic integral-partial differential equations. In [65], the existence of a viscosity solution to a path-dependent nonlinear Kolmogorov equation (without jumps) and the corresponding nonlinear Feynman–Kac representation has been proved.

In this chapter, we deal with the notion of *mild solution* which can be seen as an intermediate notion for solutions of a PDE lying in between the notions of classical and viscosity solutions. In [109], the authors provide the definition of mild solution for nonlinear Kolmogorov equations along with its link with a specific stochastic process. The latter has been also proved for Semilinear Parabolic Equations in [110], where the definition of the *generalized directional gradient* is firstly introduced. The concept of mild solution together with the generalized directional gradient to handle path-dependent Kolmogorov equation with jumps and delay has been widely analyzed in the functional formulation, see, e.g., [64]. Moreover, a discrete-time approximation for solutions of a system of decoupled FBSDEs with jumps has been proved in [30] by means of Malliavin calculus tools.

Concerning the theory of BSDE with a dependence on a delay, in [77], the authors proved the existence of a solution for a BSDE with a time-delayed generator that depends on the past values of the solution. In particular, both existence and uniqueness are proved assuming a sufficiently small time horizon T or a sufficiently small Lipschitz constant for the generator. Let us underline that the latter has an equivalent within our setting, as we state in Remark 1.5. Moreover, in [76, 78] the authors defined a path-dependent BSDE with time delayed generators driven by Brownian motions and Poisson random measures, with coefficients depending on the whole solution's path.

In [9], following a different approach, namely considering systems with memory and jumps, the authors provide a characterization of a strong solution for a delayed SDE with jumps, considering both L^p-type space and càdlàg processes to derive a non-linear Feynman–Kac representation theorem.

The present chapter is structured as follows: we start stating notations and problem setting in Section 1.1.1, according to the theoretical framework developed by [76] and [77]; in Section 1.3 we study the well-posedness of the path-dependent BSDE mentioned appearing in the Markovian FBSDEs system (1.7) following the approach in [65] by additionally considering jumps; in Section 1.4 we provide a Feynman-Kac formula relating the BSDE to the Kolmogorov Equation defined in (1.1) to then generalise results in [64] by considering a dependence in the generator f of the backward dynamic on a delayed L^2 term, namely $Y_r^{t,\phi}$, for a small delay δ; in Section 1.5 we derive the existence of a mild solution for the Kolmogorov Equation within the setting developed in [110]; finally, in Section 1.7, we provide an application based on the analyzed theoretical setting, namely a version of the *Large Investor Problem* characterised by a jump-diffusion dynamic.

1.1.1 Notation and Problem formulation

On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we consider a standard l-dimensional Brownian motion W and a homogeneous Poisson random measure N on $\mathbb{R}^+ \times (\mathbb{R} \setminus \{0\})$, independent from W, with intensity ν. With the notation $\mathbb{R}_0 := \mathbb{R} \setminus \{0\}$, we also define the compensated Poisson random measure $\tilde{N}$ defined on $\mathbb{R}^+ \times \mathbb{R}_0$ by

$$\tilde{N}(dt, dz) := N(dt, dz) - \nu(dz)dt\,. \tag{1.10}$$

For the sake of completeness, let us recall that the term $\nu(dz)dt$ represents the *compensator* of the random measure N and we assume that

$$\int_{\mathbb{R}_0} |z|^2 \nu(dz) < \infty\,, \tag{1.11}$$

We refer to, e.g., [12] for further details about the stochastic integration in the presence of jumps. We remark that the assumption in (1.11) is a standard condition when dealing with financial applications.

1.2 The forward-backward delayed system

In this section we introduce the delayed forward-backward system, assuming path-dependent coefficients for the forward and the backward components, a dependence on a small delay into the generator f and the presence of jumps modelled via a compound Poisson measure. Furthermore, the equation is formulated on a general initial time t and initial values. Thus, we need to equip the backward equation with a suitable condition in $[0,t)$, as we introduced in Equation (1.15).

On previously defined probability space, we consider a filtration $\mathbb{F}^t = \{\mathcal{F}_s^t\}_{s\in[0,T]}$, which is nothing but the one jointly generated by $W(s\wedge t)-W(t)$ and $N(s\wedge t,\cdot)-N(t,\cdot)$, augmented by all $\mathbb{P}$-null sets. We emphasize that $\mathbb{F}^t$ depends explicitly on t, namely the arbitrary initial time in $[0,T]$ for the dynamic in Eq. (1.7).

Furthermore, the components of the solution of the backward dynamic are defined in the following Banach spaces:

- $\mathbb{S}_t^2(\mathbb{R})$ denotes the space of (equivalence class of) $\mathbb{F}^t$-adapted, product measurable càdlàg processes $Y:\Omega\times[0,T]\to\mathbb{R}$ satisfying
$$\mathbb{E}\left[\sup_{t\in[0,T]}|Y(t)|^2\right]<\infty;$$
- $\mathbb{H}_t^2(\mathbb{R}^l)$ denotes the space of (equivalence class of) $\mathbb{F}^t$-predictable processes $Z:\Omega\times[0,T]\to\mathbb{R}^l$ satisfying
$$\mathbb{E}\left[\int_0^T|Z(t)|^2dt\right]<\infty;$$
- $\mathbb{H}_{t,\nu}^2(\mathbb{R})$ denotes the space of (equivalence class of) $\mathbb{F}^t$-predictable processes $U:\Omega\times[0,T]\times\mathbb{R}_0\to\mathbb{R}$ satisfying
$$\mathbb{E}\left[\int_0^T\int_{\mathbb{R}_0}|U(t,z)|^2\nu(dz)dt\right]<\infty.$$

The spaces $\mathbb{S}_t^2(\mathbb{R})$, $\mathbb{H}_t^2(\mathbb{R}^l)$ and $\mathbb{H}_{t,\nu}^2(\mathbb{R})$ are endowed with the following norms:

$$||Y||^2_{\mathbb{S}_t^2(\mathbb{R})}=\mathbb{E}\left[\sup_{t\in[0,T]}e^{\beta t}|Y(t)|^2\right],$$

$$||Z||^2_{\mathbb{H}_t^2(\mathbb{R}^l)}=\mathbb{E}\left[\int_0^T e^{\beta t}|Z(t)|^2dt\right],$$

and

$$||U||^2_{\mathbb{H}_{t,\nu}^2(\mathbb{R})}=\mathbb{E}\left[\int_0^T\int_{\mathbb{R}_0}e^{\beta t}|U(t,z)|^2\nu(dz)dt\right],$$

for some $\beta > 0$, to be precised later.
The main goal is to find a family of stochastic processes $\left(X^{t,\phi}, Y^{t,\phi}, Z^{t,\phi}, U^{t,\phi}\right)$ for $t, \phi \in [0,T] \times \Lambda$ adapted to $\mathbb{F}^t$ such that the following decoupled forward-backward system holds a.s.

$$\begin{cases} X^{t,\phi}(s) = \phi(t) + \displaystyle\int_t^s b(r, X^{t,\phi}) dr + \int_t^s \sigma(r, X^{t,\phi}) dW(r) + \\ \qquad + \displaystyle\int_t^s \int_{\mathbb{R}_0} \gamma(r, X^{t,\phi}, z) \tilde{N}(dr, dz) \,, & s \in [t,T] \\ X^{t,\phi}(s) = \phi(s) \,, & s \in [0,t] \\ Y^{t,\phi}(s) = h(X^{t,\phi}) + \displaystyle\int_s^T f\left(r, X^{t,\phi}, Y^{t,\phi}(r), Z^{t,\phi}(r), \tilde{U}^{t,\phi}(r), Y_r^{t,\phi}\right) dr \\ \qquad - \displaystyle\int_s^T Z^{t,\phi}(r) dW(r) - \int_s^T \int_{\mathbb{R}_0} U^{t,\phi}(r,z) \tilde{N}(dr, dz) \,, & s \in [t,T] \\ Y^{t,\phi}(s) = Y^{s,\phi}(s) \,, \quad Z^{t,\phi}(s) = U^{t,\phi}(s,z) = 0 \,, & s \in [0,t] \,, \end{cases} \tag{1.12}$$

recalling that the term $\tilde{U}^{t,\phi}$ was introduced in (1.9).

It is worth mentioning that, differently from [76], we work in a non-Markovian setting, enforcing an initial condition over all the interval $[0,t]$. More precisely, for both forward and backward equations, the values of the solution $X^{t,\phi}$ need to be known in the time interval $[0,t]$. Analogously, regarding the backward component, the values of $Y^{t,\phi}$, $Z^{t,\phi}$ and $U^{t,\phi}$ need also to be prescribed for $s \in [0,t]$.

Remark 1.1. *The δ-delayed feature concerns only Y, but we emphasize that it is possible to generalize this result to treat the case where both Z and U depend on their past values for a fixed delay δ. For the sake of simplicity, we consider the case with Y_r, hence limiting ourselves to just one, the process Y, L^2 delayed term. As a consequence, the latter implies that we will have to consider a larger functional space to properly define the contraction which is an essential step to prove the fix point argument in Th. 1.6.*

The forward path-dependent SDE with jumps

We first study the forward component of X appearing in the system (1.12). It is defined according to the following equation:

$$\begin{cases} X^{t,\phi}(s) = \phi(t) + \int_t^s b(r, X^{t,\phi})dr + \int_t^s \sigma(r, X^{t,\phi})dW(r) + \\ \qquad + \int_t^s \int_{\mathbb{R}_0} \gamma(r, X^{t,\phi}, z)\tilde{N}(dr, dz)\,, \qquad s \in [t,T] \\ X^{t,\phi}(s) = \phi(s)\,, \qquad\qquad s \in [0,t]\,. \end{cases} \tag{1.13}$$

More precisely, we say that $X^{t,\phi}$ is a solution to equation (1.13) if the process $s \mapsto X^{t,\phi}(s)$ is $\mathbb{F}^t$-adapted, $\mathbb{P}$-a.s. continuous and (1.13) is satisfied for any $s \in [0,T]$, $\mathbb{P}$-a.s.

Recalling that a Borel-measurable function $\varphi : [0,T] \times \Lambda \times S$ is *non-anticipative* if $\varphi(t,\phi,e) = \varphi(t,\phi(\cdot \wedge t),e)$, for all $(t,\phi,e) \in [0,T] \times \Lambda \times S$, where S is an arbitrary topological space, we assume the following assumptions to hold.

Assumption 1.2. *Let us consider three non-anticipative functions $b : [0,T] \times \Lambda \to \mathbb{R}^d$, $\sigma : [0,T] \times \Lambda \to \mathbb{R}^{d\times l}$ and $\gamma : [0,T] \times \Lambda \times \mathbb{R} \to \mathbb{R}^d$ such that*

(A_1) b, σ and γ are continuous;

(A_2) there exists $\ell > 0$ such that

$$|b(t,\phi) - b(t,\phi')| + |\sigma(t,\phi) - \sigma(t,\phi')| + ||\gamma(t,\phi,\cdot) - \gamma(t,\phi',\cdot)||_{L^2} \le \ell ||\phi - \phi'||_{L^\infty}$$

for any $t \in [0,T]$, $\phi, \phi' \in \Lambda$;

(A_3) the following bound

$$\int_{\mathbb{R}_0} \sup_{\phi \in \Lambda} |\gamma(r,\phi,z)|^2 \nu(dz) < \infty\,,$$

holds.

The existence and the pathwise uniqueness for a solution of forward SDE with jumps under Lipschitz coefficients is a known result, already classical for the case without path-dependent coefficients, see, e.g. [201]. For the sake of completeness, we report the following proposition:

Proposition 1.3. *If (A_1), (A_2), (A_3) hold, then there exists a solution to* (1.13) *and this solution is pathwise unique.*

The proof for an equivalent path-dependent setting is stated in Theorem 2.12 in [9]. Both existence and uniqueness are derived via a Picard iteration approach in the so-called Delfour-Mitter space that we introduce in subsection 1.4.1. It turns out that our equation (1.13) can be reformulated with Delfour-Mitter coefficients such as Eq. (1.20) by transformations of the coefficients we define in Eq. (1.19). Hypotheses ($\mathbf{D_1}$) and ($\mathbf{D_2}$) required for the existence result can be deduced from our conditions (A_1) – (A_3). A similar approach, but for a different class of integrators is used in Theorem 5.2.15 in [24].

The backward delayed path-dependent SDE with jumps

We now focus on the BSDE appearing in the system (1.12), namely

$$\begin{cases} Y^{t,\phi}(s) = h(X^{t,\phi}) + \displaystyle\int_s^T f\big(r, X^{t,\phi}, Y^{t,\phi}(r), Z^{t,\phi}(r), \tilde{U}^{t,\phi}(r), Y_r^{t,\phi}\big)dr \\ \qquad - \displaystyle\int_s^T Z^{t,\phi}(r)dW(r) - \int_s^T \int_{\mathbb{R}_0} U^{t,\phi}(r,z)\tilde{N}(dr,dz), \qquad s \in [t,T] \\ Y^{t,\phi}(s) = Y^{s,\phi}(s), \qquad Z^{t,\phi}(s) = U^{t,\phi}(s,z) = 0, \qquad s \in [0,t]. \end{cases} \tag{1.14}$$

for a finite time horizon $T < \infty$ and $\phi \in \Lambda := D\big([0,T];\mathbb{R}^d\big)$. The path-dependent process $X^{t,\phi}$ represents the solution of the forward SDE with jumps of Eq. (1.13), while $\tilde{N}$ models the compensated Poisson random measure described in Eq. (1.10) and W is a l-dimensional Brownian motion.

We recall that, when we fix the delay term δ, the notation $Y_r^{t,\phi}$ stands for the path of the process restricted to $[r-\delta, r]$, according to Eq.(1.8). Notice that the terminal condition enforced by h depends on the solution of the forward SDE (1.13) as well as the solution (Y,Z,U) of the backward component considered in the time interval $[t,T]$.

Differently from the framework studied by Delong in [76], we consider a general initial time $s \in [0,t)$. As highlighted in [65], the Feynman-Kac formula would fail with standard prolongation.

Thus, an additional initial condition has to be satisfied over the interval $[0,t]$, given by

$$Y^{t,\phi}(s) = Y^{s,\phi}(s), \qquad s \in [0,t). \tag{1.15}$$

We remark that the supplementary initial condition stated in Eq. (1.15) represents one of the main differences between Theorem 1.6 and Theorem 14.1.1 in [76].

1.3 The Well Posedness of the BSDE

Concerning the delayed backward SDE (1.14), we will assume the following to hold.

Assumption 1.4. *Let* $f:[0,T]\times\Lambda\times\mathbb{R}\times\mathbb{R}^l\times\mathbb{R}\times L^2([-\delta,0];\mathbb{R})\to\mathbb{R}$*,* $h:\Lambda\to\mathbb{R}$ *and* $\lambda:\mathbb{R}_0\to\mathbb{R}_+$ *(introduced in* (1.9)*) such that the following holds:*

(A_4) There exist $L,K,M>0$*,* $p\geq 1$ *and a probability measure* α *on* $\mathcal{B}([-\delta,0])$ *such that, for any* $t\in[0,T]$*,* $\phi\in\Lambda$*,* $(y,z,u),(y',z',u')\in\mathbb{R}\times\mathbb{R}^l\times\mathbb{R}$ *and* $\hat{y},\hat{y}'\in L^2([-\delta,0];\mathbb{R})$*, we have*

$$\begin{aligned}
&(i) && \phi\mapsto f(t,\phi,y,z,u,\hat{y}) \text{ is continuous},\\
&(ii) && |f(t,\phi,y,z,u,\hat{y})-f(t,\phi,y',z',u',\hat{y})|\leq L(|y-y'|+|z-z'|+|u-u'|),\\
&(iii) && |f(t,\phi,y,z,u,\hat{y})-f(t,\phi,y,z,u,\hat{y}')|^2\leq K\int_{-\delta}^0\left(|\hat{y}(\theta)-\hat{y}'(\theta)|^2\right)\alpha(d\theta),\\
&(iv) && |f(t,\phi,0,0,0,0)|<M(1+\|\phi\|_T^p).
\end{aligned}$$

(A_5) The function f *is non-anticipative.*

(A_6) The function h *is continuous and* $|h(\phi)|\leq M(1+\|\phi\|_T^p)$*, for all* $\phi\in\Lambda$*.*

(A_7) The function λ *is measurable and* $\lambda(z)\leq M(1\wedge|z|)$*, for all* $z\in\mathbb{R}_0$*.*

The following remark generalizes a classical result, see, e.g., Theorem 14.1.1 in [76], Theorem 2.1 in [77] or Theorem 2.1 in [78].

Remark 1.5. *In order to show both existence and uniqueness of a solution to the backward part of the system* (1.7) *and to obtain the continuity of* $Y^{t,\phi}$ *with respect to* ϕ*, we need to impose* K *or* δ *to be small enough. More precisely, we will assume that there exists a constant* $\chi\in(0,1)$*, such that:*

$$K\,\frac{\chi e^{\left(\chi+\frac{6L^2}{\chi}\right)\delta}}{(1-\chi)L^2}\max\{1,T\}<\frac{1}{578}\,. \tag{1.16}$$

The main difference between our result and Theorem 3.4 in [65] relies on the presence of a jump component in the dynamics of the unknown process $Y^{t,\phi}$: this further term implies a stronger bound in the condition enforced in Eq. (1.16).

Hence, if K or δ are small enough to satisfy the condition stated in Eq. (1.16), then there exists a unique solution of (1.14) and the following theorem holds

Theorem 1.6. *Let assumptions (A_1) – (A_7) hold. If condition* (1.16) *is satisfied, then there exists a unique solution* $(Y^{t,\phi},Z^{t,\phi},U^{t,\phi})_{(t,\phi)\in[0,T]\times\Lambda}$ *of the BSDE* (1.14) *such that* $(Y^{t,\phi},Z^{t,\phi},U^{t,\phi})\in\mathbb{S}_t^2(\mathbb{R})\times\mathbb{H}_t^2(\mathbb{R}^l)\times\mathbb{H}_{t,\nu}^2(\mathbb{R})$ *for all* $t\in[0,T]$ *and the application* $t\mapsto(Y^{t,\phi},Z^{t,\phi},U^{t,\phi})$ *is continuous from* $[0,T]$ *into* $\mathbb{S}_0^2(\mathbb{R})\times\mathbb{H}_0^2(\mathbb{R}^l)\times\mathbb{H}_{0,\nu}^2(\mathbb{R})$*.*

The proof of Th. 1.6 is provided in Appendix 6.6 and it is mainly based on the Banach fixed point theorem.

We emphasize that similar results hold also for multi-valued processes, namely $Y : \Omega \times [0,T] \to \mathbb{R}^m$, $Z : \Omega \times [0,T] \to \mathbb{R}^{m\times l}$ and $U : \Omega \times [0,T] \times (\mathbb{R}^n \setminus \{0\}) \to \mathbb{R}^m$. Further difficulties may arise, due to the presence of correlation between the different components of $Y \in \mathbb{S}^2_t(\mathbb{R}^m)$ or the necessity of introducing the *n-fold iterated stochastic integral*, see [16], [30] or [64, Sec. 2.1] for further details.

1.4 The Feynman-Kac formula

In what follows we prove that the solution of of Eq. (1.7), namely the path-dependent forward-backward system with delayed generator f and driven by a Lévy process, can be connected to the solution of path-dependent PIDE represented by the non-linear Kolmogorov equation (1.1).

1.4.1 The Delfour-Mitter space

According to [9], we need the solution of the forward SDE (1.7) to be a Markov process to derive the Feynman-Kac formula. The Markov property of the solution is fully known for the SDE without jumps, i.e. when $\gamma = 0$, see [170] (Th. III. 1.1). Moreover, the Markov property also holds, by enlarging the state space, for the solution in a setting analogous to that of Eq. (1.13), see [64] (Prop. 2.6) where driving noises with independent increments are considered. Since $X^{t,\phi} : \Omega \times [0,T] \times D([0,T];\mathbb{R}^d) \to \mathbb{R}^d$ in not Markovian, we enlarge the state space by considering the process X as a *process of the path*, by introducing a suitable Hilbert space, as described in [107] and [163], where they present a product-space reformulation of (1.13) splitting the present state $X(t)$ from the past trajectory X_t by a particular choice of the state space. Accordingly, we enlarge the state space of our interest, starting from paths defined on the Skorohod space $D\Big([0,T];\mathbb{R}^d\Big)$ to then consider a *new* functional space, the so-called *Delfour-Mitter space* $M^2 := L^2([-T,0];\mathbb{R}^d) \times \mathbb{R}^d$, by exploiting the continuous embedding of $D\Big([-T,0];\mathbb{R}^d\Big)$ into $L^2([-T,0];\mathbb{R}^d)$, as in-depth analyzed in, e.g., [9] and [64].

It is worth mentioning that M^2 has a Hilbert space structure, endowed with the following scalar product

$$\langle \phi, \psi \rangle_{M^2} = \langle \phi, \psi \rangle_{L^2} + \phi(0) \cdot \psi(0) \,,$$

with associated norm

$$||\phi||^2_{M^2} = ||\phi||_{L^2} + |\phi(0)|^2 \,,$$

where $\cdot$ and $|\cdot|$ stand for the scalar product in $\mathbb{R}^d$, resp. for the Euclidean norm in $\mathbb{R}^d$, while

$\langle\cdot,\cdot\rangle_{L^2}$, resp. $||\cdot||_{L^2}$, indicates the scalar product, resp. the norm in $L^2 := L^2([-T,0];\mathbb{R}^d)$.

For $t \in [0,T]$, $\phi \in \Lambda$ and $(\varphi,x) \in M^2$, let us set:

- compatible initial conditions $x^{t,\phi} := \phi(t)$ and $\eta^{t,\phi} \in D([-T,0];\mathbb{R}^d)$ defined by

$$\eta^{t,\phi}(\theta) := \begin{cases} \phi(t+\theta), & \theta \in [-t,0]; \\ \phi(0), & \theta \in [-T,t); \end{cases} \tag{1.17}$$

- $(\varphi,x)^t \in \Lambda$ defined by

$$(\varphi,x)^t(\theta) := \begin{cases} \varphi(\theta - t), & \varphi \in D([-T,0];\mathbb{R}^d),\ \theta \in [0,t); \\ x, & \varphi \in D([-T,0];\mathbb{R}^d),\ \theta \in [t,T]; \\ 0, & \varphi \notin D([-T,0];\mathbb{R}^d); \end{cases} \tag{1.18}$$

- $\tilde{b} : [0,T] \times M^2 \to \mathbb{R}^d$, $\tilde{\sigma} : [0,T] \times M^2 \to \mathbb{R}^{d\times l}$, $\tilde{\gamma} : [0,T] \times M^2 \times \mathbb{R} \to \mathbb{R}^d$ defined by

$$\begin{aligned} \tilde{b}(t,\varphi,x) &:= b(t,(\varphi,x)^t), \\ \tilde{\sigma}(t,\varphi,x) &:= \sigma(t,(\varphi,x)^t), \\ \tilde{\gamma}(t,\varphi,x,z) &:= \gamma(t,(\varphi,x)^t,z), \end{aligned} \tag{1.19}$$

where $b : [0,T] \times \Lambda \to \mathbb{R}^d$, $\sigma : [0,T] \times \Lambda \to \mathbb{R}^{d\times l}$ and $\gamma : [0,T] \times \Lambda \times \mathbb{R} \to \mathbb{R}^d$ are the given coefficients of Eq. (1.13).

We emphasize that $(\varphi,x)^t$ is well defined since it does not depend on the choice of the representative in the class of $\varphi \in L^2$, and the continuous embedding $D\Big([-T,0];\mathbb{R}^d\Big) \subset M^2$ is also injective.

We can now rewrite the forward equation (1.13) in the M^2-setting:

$$\begin{cases} X^{t,\eta,x}(s) = x + \displaystyle\int_t^s \tilde{b}(r, X_r^{t,\eta,x}, X^{t,\eta,x}(r))dr + \int_t^s \tilde{\sigma}(r, X_r^{t,\eta,x}, X^{t,\eta,x}(r))dW(r) \\ \qquad + \displaystyle\int_t^s \int_{\mathbb{R}_0} \tilde{\gamma}(r, X_r^{t,\eta,x}, X^{t,\eta,x}(r), z)\tilde{N}(dr,dz), \qquad s \in [t,T] \\ \Big(X_t^{t,\eta,x}, X^{t,\eta,x}(t)\Big) = (\eta, x). \end{cases} \tag{1.20}$$

Here, for a process X, the term X_r means $(X((r+\theta)^+))_{\theta\in[-T,0]}$, which is slightly different from the delayed term introduced in (1.8). Since this notation occurs only in the context of the M^2-setting for the forward equation, the reader can clearly distinguish between the two uses.

The link between the solution of equation Eq. (1.13) and that of equation Eq. (1.20) is then provided by

$$X^{t,\phi} = X^{t,\eta^{t,\phi},x^{t,\phi}}, \forall (t,\phi) \in [0,T] \times \Lambda. \tag{1.21}$$

In order to take advantage of the Feynman-Kac formula already derived in [64] in the case of non-delayed but still path-dependent BSDE, we have to additionally impose that b, σ, γ, f and h are locally Lipschitz-continuous with respect to $\phi \in \Lambda$ in the L^2-norm. Thus, in order to have the same regularity for the solution of the BSDE system with forward Eq. (1.20), we require that the coefficients b, σ and γ to be locally Lipschitz.

Assumption 1.7. *There exists $K \geq 0$ and $m \geq 0$ such that:*

(A_8) for all $t \in [0,T]$ and for all $\phi_1, \phi_2 \in \Lambda$,

$$|b(t,\phi_1) - b(t,\phi_2)|^2 + |\sigma(t,\phi_1) - \sigma(t,\phi_2)|^2 + \int_{\mathbb{R}_0} |\gamma(t,\phi_1,z) - \gamma(t,\phi_2,z)|^2 \nu(dz)$$
$$\leq K||\phi_1 - \phi_2||^2_{L^2}(1 + ||\phi_1||^2_{L^2} + ||\phi_2||^2_{L^2});$$

(A_9) for all $t \in [0,T]$, $y \in \mathbb{R}$, $z \in \mathbb{R}^l$, $u \in \mathbb{R}$, $\hat{y} \in L^2([-\delta,0];\mathbb{R})$ and for all $\phi_1, \phi_2 \in \Lambda$,

$$|f(t,\phi_1,y,z,u,\hat{y}) - f(t,\phi_2,y,z,u,\hat{y})| \leq$$
$$K(1 + ||\phi_1||_{L^2} + ||\phi_2||_{L^2} + |y|)^m \cdot (1 + |z| + |u|)||\phi_1 - \phi_2||_{L^2}.$$

(A_{10}) for all $\phi_1, \phi_2 \in \Lambda$,

$$|h(\phi_1) - h(\phi_2)| \leq K(1 + ||\phi_1||_{L^2} + ||\phi_2||_{L^2})^m ||\phi_1 - \phi_2||_{L^2}.$$

We remark that $\tilde{b}$, $\tilde{\sigma}$ and $\tilde{\gamma}$, as defined by (1.19) are not necessarily locally Lipschitz, since $D\left([0,T];\mathbb{R}^d\right)$ is dense in $L^2\left([0,T];\mathbb{R}^d\right)$, but they are set to constants outside $D\left([0,T];\mathbb{R}^d\right)$. However, one can define these coefficients first on $[0,T] \times D\left([0,T];\mathbb{R}^d\right) \times \mathbb{R}^d$ and then extend to $[0,T] \times M^2$ by density, using the uniform continuity provided by assumption (A_8). It is this version of the above functions that we will use in the sequel.

Within this setting, lifting the state space turns out to be particularly convenient in order to investigate differentiability properties of the solution and to relate the solution of Eq. (1.20) (combined with the backward equation) to the solutions of the non-linear Kolmogorov equation defined by Eq. (1.1) on $[0,T] \times \Lambda$.

Remark 1.8. *It is also possible to work directly on the Skorohod space D. However, since D is not a separable Banach space, one has to consider weaker topologies on D, following a semi-group approach like the one developed by Peszat and Zabczyk in [186].*

1.4.2 Main theorem

In what follows we provide the main result, namely a nonlinear version of the Feynman–Kac formula in the case where the process $X^{t,\phi}$ has jumps and the generator of the backward dynamic f depends on the past values of Y.

Theorem 1.9 (Feynman-Kac formula)**.** *Under hypotheses (A_1) – (A_{10}) with condition* (1.16) *being verified, let $(X^{t,\phi}, Y^{t,\phi}, Z^{t,\phi}, U^{t,\phi})_{(t,\phi)\in[0,T]\times\Lambda}$ be the solution of the forward-backward system* (1.12)*.*

Let $u:[0,T]\times\Lambda\to\mathbb{R}$ be the deterministic function defined by

$$u(t,\phi) = Y^{t,\phi}(t), \quad (t,\phi)\in[0,T]\times\Lambda. \tag{1.22}$$

Then u is a non-anticipative function and there exist constants $C>0$ and $m\geq 0$ such that, for all $t\in[0,T]$ and $\phi_1,\phi_2\in\Lambda$,

$$|u(t,\phi_1)-u(t,\phi_2)| \leq C(1+||\phi_1||_{L^2}+||\phi_2||_{L^2})^m||\phi_1-\phi_2||_{L^2}.$$

Moreover, the following formula holds:

$$Y^{t,\phi}(s) = u(s, X^{t,\phi}), \ \forall s\in[0,T], \tag{1.23}$$

for any $(t,\phi)\in[0,T]\times\Lambda$.

To prove the representation formula (1.22), we adapt the proof of Theorem 4.10 of [65] by adding the contribution of U and $\tilde{U}$, respectively modelling the process and the integral term connected to the jump component.

Proof. We follow the Picard iteration scheme, hence considering the iterative process of the BSDE with delayed generator driven by Lévy process described by

$$\begin{aligned} Y^{n+1,t,\phi}(s) = h(X^{t,\phi}) + \int_s^T f\big(r, X^{t,\phi}, Y^{n+1,t,\phi}(r), Z^{n+1,t,\phi}(r), \tilde{U}^{n+1,t,\phi}(r), Y_r^{n,t,\phi}\big)dr \\ - \int_s^T Z^{n+1,t,\phi}(r)dW(r) - \int_t^T\int_{\mathbb{R}} U^{n+1,t,\phi}(r,z)\tilde{N}(dr,dz), \end{aligned}$$

with $Y^{0,t,\phi} \equiv 0$, $Z^{0,t,\phi} \equiv 0$ and $U^{0,t,\phi} \equiv 0$ and initial condition

$$Y^{n+1,t,\phi}(s) = Y^{n+1,s,\phi}(s), \qquad Z^{n+1,t,\phi}(s) = U^{n+1,t,\phi}(s,z) = 0, \quad s \in [0,t].$$

Let us suppose that there exists a non-anticipative functional $u_n : [0,T] \times \Lambda \to \mathbb{R}$ such that u_n is locally Lipschitz and $Y^{n,t,\phi}(s) = u_n(s, X^{t,\phi})$ for every $t,s \in [0,T]$ and $\phi \in \Lambda$.

Since $Y^{n,t,\phi}(r+\theta) = u_n(r+\theta, X^{t,\phi})$ if $r+\theta \geq 0$ and $Y^{n,t,\phi}(r+\theta) = Y^{n,t,\phi}(0) = u_n(0, X^{t,\phi})$ if $r+\theta < 0$, by defining

$$\tilde{u}_n(t,\phi) := \Big(u_n((t+\theta)^+, \phi) \Big)_{\theta \in [-\delta,0]},$$

the delayed term reads

$$Y_r^{n,t,\phi} = \tilde{u}_n(r, X^{t,\phi})$$

and the above equation becomes

$$\begin{cases} Y^{n+1,t,\phi}(s) = h(X^{t,\phi}) \\ + \displaystyle\int_s^T f\big(r, X^{t,\phi}, Y^{n+1,t,\phi}(r), Z^{n+1,t,\phi}(r), \tilde{U}^{n+1,t,\phi}(r), \tilde{u}_n(r, X^{t,\phi})\big) dr \\ \qquad - \displaystyle\int_s^T Z^{n+1,t,\phi}(r) dW(r) - \int_s^T \int_{\mathbb{R}\setminus\{0\}} U^{n+1,t,\phi}(r,z) \tilde{N}(dr,dz), \quad s \in [t,T] \\ Y^{n+1,t,\phi}(s) = Y^{n+1,s,\phi}(s), \quad Z^{n+1,t,\phi}(s) = U^{n+1,t,\phi}(s,z) = 0, \qquad s \in [0,t]. \end{cases}$$

For fixed n, we define $\psi_n : [0,T] \times D([-T,0];\mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^l \times \mathbb{R} \to \mathbb{R}$ as

$$\psi_n(t,\varphi,x,y,z,u) := f(t, (\varphi,x)^t, y, z, u, \tilde{u}_n(t, (\varphi,x)^t)), \tag{1.24}$$

and $\tilde{h} : D([-T,0];\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}$ as

$$\tilde{h}(\varphi,x) := h((\varphi,x)^T). \tag{1.25}$$

Since u_n is locally Lipschitz-continuous, one can show that ψ_n and $\tilde{h}$ are also locally Lipschitz in φ. Therefore, one can extend them by density to $[0,T] \times M^2 \times \mathbb{R} \times \mathbb{R}^l \times \mathbb{R}$, respectively to M^2, keeping the locally Lipschitz property.

After setting $\eta = \eta^{t,\phi}$, $x = x^{t,\phi}$ (see (1.17)), we enlarge the state space and write the forward–backward system in an equivalent way as (for simplicity, we omit the dependence on

$n)$

$$\begin{cases} \mathrm{d}X^{\tau,\eta,x}(t) & = \tilde{b}(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t))\mathrm{d}t + \tilde{\sigma}(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t))\mathrm{d}W(t) \\ & \quad + \int_{\mathbb{R}_0} \tilde{\gamma}(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t), z)\tilde{N}(\mathrm{d}t, \mathrm{d}z) \\ (X_\tau^{\tau,\eta,x}, X^{\tau,\eta,x}(\tau)) & = (\eta, x) \\ \mathrm{d}Y^{\tau,\eta,x}(t) & = \psi_n\left(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t), Y^{\tau,\eta,x}(t), Z^{\tau,\eta,x}(t), \tilde{U}^{\tau,\eta,x}(t)\right)\mathrm{d}t \\ & \quad + Z^{\tau,\eta,x}(t)\mathrm{d}W(t) + \int_{\mathbb{R}_0} U^{\tau,\eta,x}(t,z)\tilde{N}(\mathrm{d}t, \mathrm{d}z) \\ Y^{\tau,\eta,x}(T) & = \tilde{h}(X_T^{\tau,\eta,x}, X^{\tau,\eta,x}(T)). \end{cases}$$

It is easy to check that the two sets of solutions, $\left(X^{t,\eta,x}, Y^{t,\eta,x}, Z^{t,\eta,x}, U^{t,\eta,x}\right)$ and $\left(X^{t,\phi}, Y^{n+1,t,\phi}, Z^{n+1,t,\phi}, U^{n+1,t,\phi}\right)$ are equivalent.

We have already observed that the extended coefficients $\tilde{b}$, $\tilde{\sigma}$, $\tilde{\gamma}$ satisfy conditions (A1)-(A2) in [64, pp. 8-9]; in a similar manner, ψ_n and $\tilde{h}$ satisfy (B1) and (B2) in [64, p. 24]. Then we can exploit Theorem 4.5 in [64] and infer that there exists a locally Lipschitz function $\bar{u}_n : [0,T] \times M^2 \to \mathbb{R}$ such that the following representation formula holds:

$$Y^{\tau,\eta,x}(t) = \bar{u}_n\left(t, X_t^{\tau,\eta,x}, X^{\tau,\eta,x}(t)\right).$$

Let us define the following locally Lipschitz, non-anticipative functional $u_{n+1} : [0,T] \times \Lambda \to \mathbb{R}$ by

$$u_{n+1}(t,\phi) := \bar{u}_n(t, \eta^{t,\phi}, \phi(t)),$$

where the time shifting $\eta^{t,\phi}$ is defined according to (1.17). Then

$$Y^{n+1,t,\phi}(s) = u_{n+1}(s, X^{t,\phi}),\ \forall s \in [0,T].$$

Notice that $\left(Y^{n,t,\phi}, Z^{n,t,\phi}, U^{n,t,\phi}\right)$ is the Picard iterative sequence needed to construct the solution $\left(Y^{t,\phi}, Z^{t,\phi}, U^{t,\phi}\right)$:

$$\left(Y^{n+1,\cdot,\phi}, Z^{n+1,\cdot,\phi}, U^{n+1,\cdot,\phi}\right) = \Gamma(Y^{n,\cdot,\phi}, Z^{n,\cdot,\phi}, U^{n,\cdot,\phi}),$$

where Γ is the contraction defined in the proof of Theorem 1.6. By applying Theorem 1.6, we then have:

$$\lim_{n\to\infty} \mathbb{E}\big(\sup_{s\in[0,T]} \left|Y^{n,t,\phi}(s) - Y^{t,\phi}(s)\right|^2\big) = 0.$$

Of course, $u_n(t,\phi)$ converges to $u(t,\phi) := \mathbb{E}[Y^{t,\phi}(t)]$, for every $t \in [0,T]$ and $\phi \in \Lambda$, hence implying that the nonlinear Feynman–Kac formula $Y^{t,\phi}(s) = u(s, X^{t,\phi})$ holds. From its definition, it is clear that u is non-anticipative. Regarding the locally Lipschitz property, it can be proven by applying Itô's formula to $|Y^{t,\phi_1} - Y^{t,\phi_2}|^2$ and resorting to standard calculus. □

1.5 Mild solution of the Kolmogorov Equation

In this section, we prove the existence of a mild solution of the path-dependent partial integro-differential equation (PPIDE) Kolmorogov equation (1.1) showing a dependence both on a delayed term and on integral term modelling jumps.

Let us start by recalling from [64] the notion of the *Markov transition semigroup* corresponding to the operator $\mathcal{L}$ introduced by (1.4). From [64, Prop. 2.6], we know that the strong solution $X^{0,\eta,x} \in M^2$ of Eq. (1.20) is a *Markov process* in the sense that $\mathbb{P}$-a.s.,

$$\mathbb{P}((X_t^{0,\eta,x}, X^{0,\eta,x}(t)) \in B|\mathcal{F}_s) = \mathbb{P}((X_t^{0,\eta,x}, X^{0,\eta,x}(t)) \in B|(X_s^{0,\eta,x}, X^{0,\eta,x}(s))),$$

for all $(\eta, x) \in M^2$ and all Borel sets B of M^2, the proof of this fundamental result being given in several works, as, e.g. [9, Th. 3.9], [192, Prop. 3.3] or [186, Sec. 9.6]; for more details, see, e.g., [110].

Denoting by $B_p(S)$ the space of Borel functions with at most polynomial growth on a metric space S, the transition semigroup $\tilde{P}_{t,s}$, acting on $B_p(M^2)$ is then defined by

$$\tilde{P}_{t,s}[\varphi](\eta, x) := \mathbb{E}[\varphi(X_s^{t,\eta,x}, X^{t,\eta,x}(s))], \qquad \varphi \in B_p(M^2),\ (\eta, x) \in M^2.$$

Coming back to our setting, we define $P_{t,s} : B_p(\Lambda) \to B_p(\Lambda)$ by

$$P_{t,s}[\varphi](\phi) := \mathbb{E}[\varphi(X^{t,\phi}(\cdot \wedge s)], \qquad \varphi \in B_p(\Lambda),\ \phi \in \Lambda.$$

Obviously (see (1.17) for the notations), for $\varphi \in B_p(\Lambda)$ and $\phi \in \Lambda$,

$$P_{t,s}[\varphi](\phi) = \tilde{P}_{t,s}[\varphi](\eta^{t,\phi}, x^{t,\phi}). \tag{1.26}$$

In order to introduce the notion of *mild solution*, we will need to define the *generalized directional gradient* of a function $u : [0,T] \times \Lambda \to \mathbb{R}$, following the approach in [110], also described in [64]. Suppose that the function u satisfies the following locally Lipschitz-continuity condition:

$$\begin{aligned} &|u(t,\phi_1) - u(t,\phi_2)| \le C||\phi_1 - \phi_2||_{L^2}(1 + ||\phi_1||_{L^2} + ||\phi_2||_{L^2})^m\,; \\ &|u(t,0)| \le C\,; \end{aligned} \tag{1.27}$$

Since the procedure of defining the generalized directional gradient takes place in Hilbert spaces, we will have to appeal again to the M^2-lifting, as we have already done for the coefficients of the forward equation. We define first

$$v(t,\varphi,x) := u(t,(\varphi,x)^t)$$

for $(t,\varphi,x) \in [0,T] \times D([-T,0];\mathbb{R}^d) \times \mathbb{R}^d$ and then extend it to $[0,T] \times M^2$ by density. Again, this is possible due to condition (1.27). Then, according to [110, Th. 3.1], there exists a Borel function $\zeta : [0,T] \times M^2 \to \mathbb{R}^d$ such that

$$\langle v(\cdot, X^{t,\eta,x}_{\cdot}, X^{t,\eta,x}(\cdot)), W(\cdot)\rangle_{[t,\tau]} = \int_t^\tau \zeta(s, X^{t,\eta,x}_s, X^{t,\eta,x}(\cdot))ds, \tag{1.28}$$

for any $0 \le t \le \tau \le T$ and $(\eta,x) \in M^2$, where $\langle X,Y\rangle_{[t,\tau]}$ denotes the *joint quadratic variation* of a pair of real stochastic processes $(X(t),Y(t))_{t\in[0,T]}$ on the interval $[t,\tau]$,

$$\langle X,Y\rangle_{[t,\tau]} := \lim_{\epsilon \searrow 0} \frac{1}{\epsilon} \int_t^\tau (X(s+\epsilon) - X(s)) \cdot (Y(s+\epsilon) - Y(s))ds,$$

with the limit taken in probability and $X(s),Y(s)$ defined as $X(s)$, respectively $Y(s)$ for $s > T$.

The set of all functions ζ with the above property is called the *generalized directional gradient* of v and is denoted $\nabla^\sigma v$. Its name and notation come from the observation that if v (and, consequently u) and the coefficients of the forward equation are sufficiently regular, then $\nabla_{(\eta,x)} v \cdot \tilde{\sigma} \in \nabla^\sigma v$ (see [110, Remark 3.3]).

We come back to the Λ-setting and define $\nabla^\sigma u$, the generalized directional gradient of u as the set of all non-anticipative functions $\xi : [0,T] \times \Lambda \to \mathbb{R}$ such that there exists $\zeta \in \nabla^\sigma v$ satisfying

$$\zeta(t,\varphi,x) = \xi(t,(\varphi,x)^t)$$

for all $(t,\varphi,x) \in [0,T] \times D([-T,0];\mathbb{R}^d) \times \mathbb{R}^d$. Such a function can be defined by setting $\xi(t,\phi) := \zeta(t,\eta^{t,\phi},x^{t,\phi})$.

It is clear that relation (1.28) translates into

$$\langle u(\cdot, X^{t,\phi}), W(\cdot)\rangle_{[t,\tau]} = \int_t^\tau \xi(s, X^{t,\phi})ds,$$

for all $(t,\phi) \in [0,T] \times \Lambda$ and all $\tau \in [t,T]$, if $\xi \in \nabla^\sigma u$. From this relation, it is clear (see also Remark 3.2 in [110]) that if $\xi, \hat{\xi} \in \nabla^\sigma u$, then

$$\xi(s, X^{t,\phi}) = \hat{\xi}(s, X^{t,\phi}) \ ds\text{-a.e., } \mathbb{P}\text{-a.s.},$$

so there is no ambiguity if we write $\nabla^\sigma u(s, X^{t,\phi})$.

We have now all the ingredients for introducing the notion of mild solution.

Definition 1.6. *A non-anticipative function $u : [0,T] \times \Lambda \to \mathbb{R}$ is a mild solution to Eq.* (1.1) *if u satisfies* (1.27) *and the following equality holds true for all $(t,\phi) \in [0,T] \times \Lambda$ and $\xi \in \nabla^\sigma u$:*

$$u(t,\phi) = P_{t,T}[h](\phi) + \int_t^T P_{t,s}\left[f\left(\cdot, u(s,\cdot), \xi(s,\cdot), \mathcal{J}u(s,\cdot)\right), (u(\cdot,\cdot))_s\right](\phi)ds\,, \tag{1.29}$$

where $(u(\cdot,\cdot))_s$ is the delayed term defined in (1.5)*.*

Given the definition of $P_{t,T}$ and the above mentioned remark, we can reformulate relation (1.29) as

$$\begin{aligned} u(t,\phi) = \mathbb{E}[h(X_T^{t,\phi})] \\ + \mathbb{E}\bigg[\int_t^T f\left(X^{t,\phi}, u(s,X^{t,\phi}), \nabla^\sigma u(s,X^{t,\phi}), \mathcal{J}u(s,X^{t,\phi}), (u(\cdot,X^{t,\phi}))_s\right) ds\bigg]. \end{aligned} \tag{1.30}$$

The following theorem represents the core result of this section.

Theorem 1.10 (Existence and Uniqueness)**.** *Let assumptions (A_1) – (A_{10}) and* (1.16) *hold true. Then the function u defined by* (1.22) *is the unique mild solution to the path-dependent partial integro-differential Eq.* (1.1)*.*

Proof.
Existence. Let us consider the backward component of the FBSDE described in Eq. (1.12) for $s \in [t,T]$

$$\begin{aligned} Y^{t,\phi}(s) = h(X^{t,\phi}) + \int_s^T f\big(r, X^{t,\phi}, Y^{t,\phi}(r), Z^{t,\phi}(r), \tilde{U}^{t,\phi}(r), Y_r^{t,\phi}\big)dr \\ - \int_s^T Z^{t,\phi}(r)dW(r) - \int_s^T \int_{\mathbb{R}\setminus\{0\}} U^{t,\phi}(r,z)\tilde{N}(dr,dz)\,, \qquad s \in [t,T]\,. \end{aligned} \tag{1.31}$$

By Theorem 1.9, the non-anticipative function u defined by (1.22) is satisfying (1.27) (the second part comes the continuity of $t \mapsto Y^{t,\phi}$, asserted in Theorem 1.6) and the representation formula (1.23) holds. Moreover, by means of Eq. (1.22), we can write the delayed term Y_r as a function of the path of solution of the forward dynamic $X^{t,\phi}$ and, thus, by defining

$$\tilde{u}(t,\phi) := \Big(u((t+\theta)^+, \phi)\Big)_{\theta\in[-\delta,0]}\,,$$

we can rewrite Eq. (1.31), leading to

$$\begin{cases} Y^{t,\phi}(s) = h(X^{t,\phi}) + \int_s^T f(r, X^{t,\phi}, Y^{t,\phi}(r), Z^{t,\phi}(r), \tilde{U}^{t,\phi}(r), \tilde{u}(r, X^{t,\phi}))dr \\ \qquad - \int_s^T Z^{t,\phi}(r)dW(r) - \int_s^T \int_{\mathbb{R}\setminus\{0\}} U^{t,\phi}(r,z)\tilde{N}(dr,dz), \qquad s \in [t,T] \\ Y^{t,\phi}(s) = Y^{s,\phi}(s), \qquad Z^{t,\phi}(s) = U^{t,\phi}(s,z) = 0, \qquad s \in [0,t], \end{cases} . \tag{1.32}$$

At this point, we enlarge the state space going through M^2 coefficients analogously to the proof of Theorem 1.6, obtaining

$$\begin{aligned} Y^{t,\eta,x}(s) &= \tilde{h}(X_T^{s,\eta,x}, X^{s,\eta,x}(T)) \\ &+ \int_s^T \psi(r, X_r^{t,\eta,x}, X^{t,\eta,x}(r), Y^{t,\eta,x}(r), Z^{t,\eta,x}(r), \tilde{U}^{t,\eta,x}(r))dr \\ &- \int_s^T Z^{t,\eta,x}(r)dW(r) - \int_t^T \int_{\mathbb{R}_0} U^{t,\eta,x}(r,z)\tilde{N}(dr,dz), \end{aligned} \tag{1.33}$$

where the map $\tilde{h}$ was defined in Eq. (1.25) and ψ is defined similarly to ψ_n introduced in Eq. (1.24):

$$\psi(t,\varphi,x,y,z,u) := f(t,(\varphi,x)^t, y, z, u, \tilde{u}(t,(\varphi,x)^t))$$

on $[0,T] \times D([-T,0];\mathbb{R}^d) \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^l \times \mathbb{R}$ and extending it by density. Since $\tilde{h}$ and ψ are locally Lipschitz, hence satisfying the conditions required by Theorem 4.8 in [64], we can apply this result in order to conclude that the function $v : [0,T] \times M^2 \to \mathbb{R}$ defined by

$$v(t,\eta,x) := Y^{t,\eta,x}(t)$$

is a mild solution (in the sense of [64], but similar to ours) of the following PPIDE

$$\begin{cases} -\partial_t v(t,\eta,x) - \tilde{\mathcal{L}} v(t,\eta,x) \\ \qquad -\psi\big(t,\eta,x,v(t,\eta,x), \partial_x v(t,\eta,x)\tilde{\sigma}(t,\eta,x), \tilde{\mathcal{J}} v(t,\eta,x)\big) = 0, \\ v(T,\eta,x) = \tilde{h}(\eta,x), \qquad (t,\eta,x) \in [0,T] \times M^2, \end{cases}$$

where $\tilde{\mathcal{L}}$ and $\tilde{\mathcal{J}}$ are the straightforward modifications of $\mathcal{L}$, respectively $\mathcal{J}$ in M^2. Since $v(t,\phi,x) = u(t,(\varphi,x)^t)$ for any $(t,\phi,x) \in [0,T] \times D([-T,0];\mathbb{R}^d)$, by playing on the relation (1.26) and the connection between the formulations of the generalized directional gradient in the càdlàg, respectively M^2 cases, it is straightforward to show that u is a mild solution of (1.1).

Uniqueness. Let us take two mild solutions u^1 and u^2 of the path-dependent PDE (1.1). We define

$$f^i(t,\phi,y,z,w) := f(t,\phi,y,z,w,\left(u^i(\cdot,\phi)\right)_t), \quad i=\overline{1,2}.$$

Using these drivers we can consider the following BSDEs:

$$\begin{aligned} Y^{t,\phi}(s) &= h(X^{t,\phi}) + \int_s^T f^i(r,X^{t,\phi},Y^{t,\phi}(r),Z^{t,\phi}(r),\tilde{U}^{t,\phi}(r))dr \\ &\quad - \int_s^T Z^{t,\phi}(r)\,dW(r) - \int_s^T \int_{\mathbb{R}_0} U^{t,\phi}(r,z)\tilde{N}(dr,dz), \quad i=\overline{1,2}, \end{aligned} \tag{1.34}$$

for which there exist unique solutions $\left(Y^{i,t,\phi}, Z^{i,t,\phi}, U^{i,t,\phi}\right) \in \mathbb{S}_t^2(\mathbb{R}) \times \mathbb{H}_t^2(\mathbb{R}^l) \times \mathbb{H}_{t,\nu}^2(\mathbb{R})$ for $i=\overline{1,2}$.

By Theorem 1.9 we see that

$$Y^{i,t,\phi}(s) = v^i(s,X^{t,\phi}), \quad \text{for all } s\in[0,T], \quad \text{a.s.},$$

for any $(t,\phi)\in[0,T]\times\Lambda$, where $v^i:[0,T]\times\Lambda\to\mathbb{R}$, $i=\overline{1,2}$ are defined by

$$v^i(t,\phi) := Y^{i,t,\phi}(t),\ (t,\phi)\in[0,T]\times\Lambda.$$

Hence, by the existence part, we obtain that the functions v^i are solutions of the PDE of type (1.1), but without the delayed terms $\left(v^i(\cdot,\phi)\right)_t$:

$$\begin{cases} -\partial_t v^i(t,\phi) - \mathcal{L}v^i(t,\phi) - f^i(t,\phi,v^i(t,\phi),\partial_x v^i(t,\phi)\sigma(t,\phi),\mathcal{J}v^i(t,\phi)) = 0, \\ v^i(T,\phi) = h(\phi), \quad i=\overline{1,2}. \end{cases} \tag{1.35}$$

Since u^i is also solution to equation (1.35), by using the uniqueness part of Theorem 4.5 from [64] we get that (after embedding these equations in M^2, as we did in the previous part)

$$u^i(t,\phi) = v^i(t,\phi), \quad (t,\phi)\in[0,T]\times\Lambda, \quad i=\overline{1,2}.$$

Hence

$$Y^{i,t,\phi}(s) = v^i(s,X^{t,\phi}) = u^i(s,X^{t,\phi}),\ i=\overline{1,2},$$

so BSDEs (1.34) become a single equation,

$$\begin{aligned} Y^{i,t,\phi}(s) = h(X^{t,\phi}) + \int_s^T f(r, X^{t,\phi}, Y^{i,t,\phi}(r), Z^{i,t,\phi}(r), \tilde{U}^{i,t,\phi}(r), Y_r^{i,t,\phi}) dr \\ - \int_s^T Z^{i,t,\phi}(r) dW(r) - \int_s^T \int_{\mathbb{R}_0} U^{t,\phi}(r,z) \tilde{N}(dr,dz), \end{aligned} \tag{1.36}$$

with $i = \overline{1,2}$, for which we have uniqueness from Theorem 1.6.

Therefore $Y^{1,t,\phi} = Y^{2,t,\phi}$ and, consequently

$$u^1(t,\phi) = Y^{1,t,\phi}(t) = Y^{2,t,\phi}(t) = u^2(t,\phi).$$

□

Remark 1.11. *From Theorem 4.5 in [64], besides relation* (1.23)*, it can also be inferred that for every* $(t,\phi) \in [0,T] \times \Lambda$ *the following representation formulas hold,* $\mathbb{P}$*-a.s. and for a.e.* $s \in [t,T]$*:*

$$\begin{aligned} Z^{t,\phi}(s) &= \nabla^\sigma u(s, X^{t,\phi}); \\ U^{t,\phi}(s,z) &= u\big(s, (X^{t,\phi})^{t,\gamma(s,X^{t,\phi},z)}\big) - u(s, X^{t,\phi}), \end{aligned}$$

where we use the notation introduced in (1.3)*. These formulas could have been used to prove directly the existence part of the above result, by taking the expectation in equation* (1.32) *and using relation* (1.30)*.*

1.7 Financial Application

In this section, we provide a financial application moving from the model studied in, e.g. [67], or [96]. We consider a generalization of the so-called Large Investor Problem, where a *large investor* wishes to invest in a given market, buying or selling a stock. The investor has the peculiarity that his actions on the market can affect the stock price. We refer to Example 14.1 in [76] for a detailed example of the time-delayed setting. Other results on replication problems for large investors in different settings are also presented in [65] and [96].

1.7.1 A perfect replication problem for a large investor

Concerning the problem of the perfect replication strategy for a large investor in a setting with memory, we generalize Example 14.1 in [76] by asking, in addition to path-dependent coefficients, the dynamic of the risky asset driven by a Poisson random measure.

We denote the investor's strategy by π and the investment portfolio by X^π and we assume that its past X_r^π may affect directly the stock coefficients μ, σ and γ and the bond rate. Consequently, we consider the following dynamic

$$\begin{cases} \dfrac{dS_0(t)}{S_0(t)} = r(t, X^\pi(t), X_t^\pi)dt\,, \\ S_0(0) = 1\,, \\ \dfrac{dS_i(t)}{S_i(t)} = \mu_i\,(t, X^\pi(t), X_t^\pi)\,dt + \sigma_i(t, X^\pi(t), X_t^\pi)dW(t) \\ \qquad\qquad + \displaystyle\int_{\mathbb{R}_0} \gamma_i(t, X^\pi(t), X_t^\pi, z)\tilde{N}(dt, dz), \\ S_i(0) = s_i > 0\,, \end{cases} \tag{1.37}$$

where r_i, μ_i, σ_i and γ_i, $i = \overline{1,2}$ are $\mathbb{F}^{W,\tilde{N}}$-predictable processes, $\mathbb{F}^{W,\tilde{N}}$ being the natural filtration associated to the Brownian motion W and to the Poisson random measure $\tilde{N}$, with compensator defined according to Eq. (1.10).

The total amount of the portfolio of the large investor is described by

$$dX^\pi(t) = \pi_1(t)\frac{dS_1(t)}{S_1(t)} + \pi_2(t)\frac{dS_2(t)}{S_2(t)} + (X^\pi(t) - \pi_1(t) - \pi_2(t))\frac{dS_0(t)}{S_0(t)}dt\,.$$

where, at any time $t \in [0,T]$, $\pi_i(t)$ represents the amount invested in the risky asset S_i, while $X^\pi(t) - \pi_1(t) - \pi_2(t)$ is the amount invested in the riskless one S_0.

Let us denote, for simplicity, $\phi := (\phi_1, \phi_2)$ for ϕ among the symbols S, π, μ, σ or γ and by $\cdot$ the scalar product in $\mathbb{R}^2$. The goal is to find an admissible replicating strategy $\pi \in \mathcal{A}$ for a claim $h(S(T))$.

We have that the portfolio X evolves according to

$$\begin{aligned} dX^\pi(t) &= \frac{\pi_1(t)}{S_1(t)}dS_1(t) + \frac{\pi_2(t)}{S_2(t)}dS_2(t) + \frac{X^\pi\,(t) - \pi(t)}{S_0\,(t)}dS_0\,(t) \\ &= \pi(t) \cdot \left[\mu\,(t, X^\pi, X_t^\pi)\,dt + \sigma(t, X^\pi, X_t^\pi)dW(t) + \int_{\mathbb{R}_0} \gamma(t, X^\pi, X_t^\pi, z)\tilde{N}(dt, dz)\right] \\ &\quad + [X^\pi\,(t) - \pi_1(t) - \pi_2(t)]\,r(t, X^\pi(t), X_r^\pi)dt\,, \end{aligned}$$

Hence, for $t \in [0,T]$, we have

$$X^\pi(t) = h\,(S(T)) + \int_t^T F\,(s, X^\pi(s), X_s^\pi, \pi(s))\,ds - \int_t^T Z(s)dW(s) - \int_t^T \int_{\mathbb{R}_0} U(s,z)\tilde{N}(ds, dz), \tag{1.38}$$

with final condition $X^\pi(T) = h(S)$ by denoting

$$\begin{aligned}
&F(s, X^\pi(s), X^\pi_s, \pi(s)) := -r(s, X^\pi(s), X^\pi_s)(X^\pi(s) - \pi_1(s) - \pi_2(s)) - \pi(s) \cdot \mu(s, X^\pi(s), X^\pi_s);\\
&Z^\pi(s) := \pi(s) \cdot \sigma(s, X^\pi(s), X^\pi_s);\\
&U^\pi(s,z) := \pi(s) \cdot \gamma(s, X^\pi(s), X^\pi_s, z).
\end{aligned} \tag{1.39}$$

The generator F can be rewritten to accommodate the dependence of Z^π and U^π by introducing the following transformation

$$\bar{F}\big(s, X^\pi(s), X^\pi_s, Z^\pi(s), \tilde{U}^\pi(s)\big) = F(s, X^\pi(s), X^\pi_s, \pi(s)) \tag{1.40}$$

since from Eq. (1.39) we have

$$\begin{bmatrix}\pi_1(s) & \pi_2(s)\end{bmatrix} = \begin{bmatrix}Z^\pi(s) & U^\pi(s,z)\end{bmatrix} \begin{bmatrix}\sigma_1(s, X^\pi(s), X^\pi_s) & \gamma_1(s, X^\pi(s), X^\pi_s, z)\\ \sigma_2(s, X^\pi(s), X^\pi_s) & \gamma_2(s, X^\pi(s), X^\pi_s, z)\end{bmatrix}^{-1}$$

by imposing that the matrix $\begin{bmatrix}\sigma^{\mathrm{T}} & \gamma^{\mathrm{T}}\end{bmatrix}$ is invertible.

We then ask the coefficients r, μ, σ and γ to be such that the function $\bar{F} : [0,T] \times \mathbb{R} \times L^2([-\delta, 0]; \mathbb{R}) \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ satisfies assumptions (A_4), (A_5) and (A_9) .

Furthermore, we need a couple of simplifying assumptions in order to fit the theoretical framework of the chapter.

First, we introduce a functional aiming at encoding the forward process in order to decouple the terminal condition in the BSDE (1.38) from the stock forward dynamic (1.37). The risky assets vector S can be explicitly written, by a variation of constants formula, as a functional of the random coefficients of the stock (W and $\tilde{N}$), on the wealth X and on the processes Z and U depending on the allocation strategy π. Hence $h(S(T))$ can be written as

$$h(S(T)) = \bar{h}\big(W, \tilde{N}, X^\pi, Z^\pi, U^\pi\big) \tag{1.41}$$

where $\bar{h}$ is asked to satisfy, besides conditions (A_6) and (A_{10}) in the first two coefficients, a Lipschitz condition in the last three with Lipschitz constant K_1.

Thus, we can rewrite (1.38) as

$$\begin{aligned}
X(t) = \bar{h}\big(W, \tilde{N}, X, Z, U\big) + \int_t^T \bar{F}\big(s, X(s), X_s, Z(s), \tilde{U}(s)\big) ds - \int_t^T Z(s)\, dW(s)\\
- \int_t^T \int_{\mathbb{R}_0} U(s,z) \tilde{N}(ds, dz) \quad t \in [0,T].
\end{aligned} \tag{1.42}$$

If we assume that $\delta = T$ and the Lipschitz constants K and K_1 are sufficiently small, then the time-delayed BSDE (1.42) has a unique solution. The setting is a little more general than ours, in the sense that the final condition also depends on the path of the solution of the backward equation (1.42), but one can use Theorem 14.1.1 in [76].

The further modelling concerns introducing the Markovian setting, *i.e.* allowing the initial time and value to vary. Hence, we introduce the forward processes $\bar{W}$ and $\bar{N}$ by

$$\begin{cases} (\bar{W}(s), \bar{N}(s)) = \left(\phi_1(t) + \int_t^s dW(r), \phi_2(t) + \int_t^s \int_{\mathbb{R}_0} \tilde{N}(dr,dz)\right), \quad s \in [t,T], \\ (\bar{W}(s), \bar{N}(s)) = (\phi_1(s), \phi_2(s)), \quad s \in [0,t), \end{cases} \tag{1.43}$$

with Brownian motion W, compensated Random measure $\tilde{N}$ and an initial càdlàg datum $\phi = (\phi_1, \phi_2) \in \Lambda$.

Finally, by recalling the coefficients $\bar{F}$, Z and U defined in Eq. (1.39) we may write the following decoupled forward–backward stochastic system:

$$\begin{cases} \bar{W}^{t,\phi}(s) = \phi_1(t) + \int_t^s dW(r), \quad s \in [t,T], \\ \bar{N}^{t,\phi}(s) = \phi_2(t) + \int_{\mathbb{R}_0} \int_t^s d\tilde{N}(r,z), \quad s \in [t,T], \\ \left(\bar{W}^{t,\phi}(s), \bar{N}^{t,\phi}\right) = \phi(s), \quad s \in [0,t), \\ X^{t,\phi}(s) = \bar{h}(\bar{W}^{t,\phi}, \bar{N}^{t,\phi}, X^{t,\phi}, Z^{t,\phi}, U^{t,\phi}) + \int_s^T \bar{F}(r, X^{\pi}(r), X^{\pi}_r, Z^{t,\phi}(r), \tilde{U}^{t,\phi}(r)) dr \\ \qquad\qquad - \int_s^T Z^{t,\phi}(r) dW(r) - \int_s^T \int_{\mathbb{R}_0} U^{t,\phi}(r) \tilde{N}(dr,dz), \quad s \in [t,T], \\ X^{t,\phi}(s) = X^{s,\phi}(s), \quad Z^{t,\phi}(s) = U^{t,\phi}(s,z) = 0, \quad (s,z) \in [0,t) \times \mathbb{R}_0. \end{cases} \tag{1.44}$$

where the BSDE coefficients are defined according to (1.39).

Then, imposing the same assumptions as before (where $t = 0$) there exists a unique solution $\left(X^{t,\phi}, Z^{t,\phi}, U^{t,\phi}\right)_{(t,\phi)\in[0,T]\times\Lambda}$ for the BSDE in (1.44). Again, the final condition is a little more general than the one in the theoretical framework, but this can be easily modified by adapting the proof of Theorem 1.6, as it is done, for example in Theorem 14.1.1 in [76].
Moreover, by Theorem 1.9, the solution of the backward equation in the above system can be expressed as

$$X^{t,\phi}(s) = u(s, \bar{W}^{t,\phi}, \bar{N}^{t,\phi}), \quad \text{for all } s \in [t,T],$$

for every $(t,\phi) \in [0,T] \times \Lambda$, where $u(t,\phi) := X^{t,\phi}(t)$. Furthermore, u is the mild solution, according to Definition 1.6, of the following path-dependent PDE:

$$\begin{cases} \partial_t u(t,\phi)+\frac{1}{2}\partial^2_{xx}u(t,\phi)+\int_{\mathbb{R}_0}\Big(u(t,\phi^{t,(0,1)})-u(t,\phi)-\partial_y u(t,\phi)\Big)\nu(dz) \\ \qquad\qquad +\bar{F}\left(t,u(t,\phi),\left(u\left(\cdot,\phi\right)\right)_t,\partial_x u\left(t,\phi\right),\mathcal{J}u(t,\phi)\right)=0, \\ u(T,\phi)=\bar{h}(\phi,u\left(\cdot,\phi\right),\partial_x u\left(\cdot,\phi\right),u(\cdot,\phi^{\cdot,(0,1)})-u(\cdot,\phi)), \end{cases}$$

with $(t,\phi)\in[0,T)\times\Lambda$, with $u(t,\phi^{t,(0,1)})$ corresponding to the vertical perturbation as defined in Eq. (1.3).

We also present a concrete example of a jump-diffusion model for option pricing that can help the reader link the forward dynamic to a tractable application. This example has a clear limitation if applied to our setting, such as no path-dependence in the coefficients of the stock dynamic (1.45) but only in the terminal condition h of the BSDE. Moreover, if we need to assume suitable conditions on μ, σ and γ, e.g. asking μ bounded and σ, γ constant (w.r.t. X), so that the terminal condition also satisfies the required Lipschitz condition.

Example 1.12. (Forward SDE with a discrete number of jumps). *The stock price may present a jump–diffusion dynamic with a discrete number of jumps triggered by a Poisson process, namely we consider the following equation*

$$\frac{dS(t)}{S(t)}=\mu\left(t\right)dt+\sigma(t)dW(t)+d\sum_{i=1}^{N(t)}\left(V_i-1\right), \tag{1.45}$$

in place of Eq. (1.37)*, where $N(t)$ is a standard Poisson process of fixed rate and jumps size $\{V_i\}$ modelled as a sequence of independent, identically, distributed non-negative random variables. We consider independence among all the sources of randomness, namely W and N. We refer to, e.g., [141] for a detailed treatment of this kind of jump-diffusion. Firstly, we notice that the forward Eq.* (1.45) *can be explicitly solved by the following*

$$S\left(t\right)=s_0\exp\left[\int_0^t\left(\mu\left(s\right)-\frac{1}{2}\sigma^2\left(s\right)\right)ds+\int_0^t\sigma\left(s\right)dW\left(s\right)\right]\sum_{i=1}^{N(t)}V_i\,. \tag{1.46}$$

If we assume no path-dependence in the coefficients, such as in Eq. (1.45)*, then we may encode in the terminal condition of the BSDE* (1.38)*, a dependence only on W, N and $\{V_i\}$. Hence, by introducing the following functional*

$$h\left(S\right)=\tilde{h}\left(W,N,\{V_i\}\right), \tag{1.47}$$

we can decouple the FBSDEs system to fit the setting of Theorem 1.6.

1.8 Conclusions and Future Development

The core result of this chapter relies on deriving a stochastic representation for the solutions of a non-linear PDE and associating the PDE solution to a FBSDE with jumps and a time-delayed generator. The presence of jumps both in the forward and backward dynamic and, moreover, the dependence of the generator on a (small) time-delayed coefficient represents the central aspect of novelty arising in the analysis of this kind of FBSDE system. Furthermore, we present an application for a large investor problem admitting a jump–diffusion dynamic. Throughout the chapter, we mention some possibilities to generalize the setting of our equations such as considering the dependence of f also on a delayed term for the processes Z and U, see Remark 1.1 for more details, or in-depth analyzing the choice of a weaker topology, see Remark 1.8. A different modelling choice deals with considering a further delay term affecting the forward process, see [160] for more details. Furthermore, it might deserve attention to investigate a discretization scheme for this equation, e.g. in line with [30], for the considered equations to obtain a numerical approach based on Neural Networks methods to efficiently compute an approximated solution for the considered FBSDE.

2 Time-Delayed Generalized BSDEs

2.1 Introduction

BSDE were introduced in the linear case by Bismut [26], as adjoint equations involved in the control of SDEs. The nonlinear case was considered by Pardoux and Peng first in [179] and then in [180, 183], where they established a connection between BSDEs and semilinear parabolic partial differential equations (PDEs), by the so-called nonlinear Feynman–Kac formula, in a similar framework to the one analyzed in Chapter 1,. It was this kind of applications which triggered an impressive amount of research on the subject. Concerning parabolic PDEs with Neumann boundary conditions, Pardoux and Zhang discovered that their solutions can be linked to BSDEs involving the integral with respect to continuous increasing processes (Stieltjes integral).

This work represents the first step in establishing a probabilistic representation formula of the solutions of delayed path-dependent parabolic PDEs with Neumann boundary conditions. It consists of studying the well posedness of the associated BSDEs, *i.e.* existence and uniqueness of solutions, as well as stability with respect to terminal data and coefficients. As already shown in [66] for the case of such PDEs considered on the whole space, the generator of the associated BSDE has to take into account the delayed-path of its solution. As a result, our present work is concerned with the following BSDE:

$$\begin{cases} dY(t) = -F(t,Y(t),Z(t),Y_t,Z_t)dt - G(t,Y(t),Y_t)dA(t) \\ \qquad\qquad +Z(t)dW(t), \quad t\in[0,T]; \\ Y(T)=\xi, \end{cases} \tag{2.1}$$

where the generators F and G depend also on the past of the solution $(Y(t),Z(t))$, denoted by (Y_t,Z_t). Here, if $\boldsymbol{x}:[-\delta,T]\to\mathbb{R}^n$ is a function and $t\in[0,T]$, $\boldsymbol{x}_t:[-\delta,0]\to\mathbb{R}^n$ denotes the delayed-path of $\boldsymbol{x}$, defined as

$$\boldsymbol{x}_t(\theta) := \boldsymbol{x}(t+\theta),\ \theta\in[-\delta,0],$$

where $\delta > 0$ is a fixed delay. The coefficient A is a continuous real valued increasing process.

We recall that time-delayed BSDEs were first introduced in [72] and [73] where the authors obtained the existence and uniqueness of the solution of the time–delayed BSDE

$$Y(t) = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z(s)dW(s), \quad 0 \le t \le T, \tag{2.2}$$

where

$$Y_s := (Y(r))_{r\in[0,s]} \quad \text{and} \quad Z_s := (Z(r))_{r\in[0,s]}.$$

In particular, the aforementioned existence and uniqueness result holds true if the time horizon T or the Lipschitz constant for the generator f are sufficiently small.

The motivation behind the introduction of a driving force dA and the corresponding integral goes beyond the link with PDE and can be traced in actuarial applications since the classical works [106] and [203]. In the context of insurance, a BSDE such as the one described in Equation (2.1) can be used to model the evolution of a hedging strategy for an insurance portfolio over time. In this framework, the Riemann-Stieltjes integral is linked to the sum of claims with respect to an increasing continuous process that might model the cumulative distribution of some events.

This chapter is organized as follows. In the remaining part of the first section, we introduce the notations and set the framework of our problem. In section 2.3 we derive a result of existence and uniqueness for small delay (or small Lipschitz constant) for BSDE (2.1), based on Banach's fixed point theorem, expressed in Theorem 2.3. Moreover, we provide in Proposition 2.5, the well-posedness result for an arbitrary delay for a specific case assuming monotone (in the delayed term) and linear coefficients. Section 2.4 is devoted to the problem of stability of solutions with respect to terminal data ξ and coefficients F, G and A. Lastly, in Section 2.5, we present an insurance application dealing with a variable annuity investment that suits the theoretical setting. The main difficulty encountered is to prove the convergence of the solutions of the approximating BSDEs when the increasing process A is approximated uniformly, but not in variation. In order to tackle this problem, we use a stochastic variant of Helly-Bray theorem, proved in the Appendix section, as it may be an interesting result for use in other applications.

2.1.1 Problem setting and notations

On the Euclidean space $\mathbb{R}^n$ we consider the Euclidean norm and scalar product, denoted by $|\cdot|$ and $\langle\cdot,\cdot\rangle$, respectively. If $n, k \in \mathbb{N}^*$, $\mathbb{R}^{n\times k}$ denotes the space of real $n \times k$-matrices, equipped with the Frobenius norm (the Euclidean norm when this space is identified with $\mathbb{R}^{nk}$), denoted as well by $|\cdot|$.

For $s<t$, $C([s,t];\mathbb{R}^n)$ represents the set of continuous functions $\boldsymbol{x}:[s,t]\to\mathbb{R}^d$, endowed with the sup-norm: $\|\boldsymbol{x}\|_{C([s,t];\mathbb{R}^n)}:=\sup_{r\in[s,t]}|\boldsymbol{x}(r)|$; $BV([s,t];\mathbb{R}^n)$ denotes the set of right-continuous functions with bounded variation $\boldsymbol{\eta}$: $[s,t]\to\mathbb{R}^n$, *i.e.* with a finite total variation. Recall that the total variation of $\boldsymbol{\eta}$ on $[s,t]$ is defined as

$$\mathrm{V}_s^t(\boldsymbol{\eta}):=\sup\textstyle\sum_{i=1}^n|\boldsymbol{\eta}(t_i)-\boldsymbol{\eta}(t_{i-1})|,$$

where the sup is taken on all the partitions $s=t_0<t_1<\cdots<t_n=t$. The standard norm on $BV([s,t];\mathbb{R}^n)$ is given by

$$\|\boldsymbol{\eta}\|_{BV([s,t];\mathbb{R}^n)}:=|\boldsymbol{\eta}(s)|+\mathrm{V}_s^t(\boldsymbol{\eta}).$$

We will simply denote $C[s,t]$, $BV[s,t]$ instead of $C([s,t];\mathbb{R})$, $BV([s,t];\mathbb{R})$, respectively.

If $\boldsymbol{x}:[s,t]\to\mathbb{R}^n$ is a Borel-measurable function and $\boldsymbol{\eta}\in BV([s,t];\mathbb{R}^n)$, by $\int_s^t\langle\boldsymbol{x}(r)d\boldsymbol{\eta}(r)\rangle$ we denote the sum

$$\sum_{i=1}^n\int_s^t\langle\boldsymbol{x}_i(r)d\boldsymbol{\eta}_i(r)\rangle,$$

where $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n$ and $\boldsymbol{\eta}_1,\ldots,\boldsymbol{\eta}_n$ are the components of $\boldsymbol{x}$, respectively $\boldsymbol{\eta}$, in the case where the Lebesgue-Stieltjes integrals are well-defined and the sum makes sense.

We fix now the framework of our problem, to be utilized throughout the chapter.

Let $T>0$ be a finite horizon of time, $d,m\in\mathbb{N}^*$ and $\delta\in(0,T]$ a fixed time-delay. Let $(\Omega,\mathcal{F},\mathbb{P})$ be a complete probability space, W a d-dimensional Brownian motion and $\mathbb{F}=\{\mathcal{F}_t\}_{t\in[0,T]}$ the filtration generated by W, augmented by the null-probability subsets of Ω. The stochastic process $A:\Omega\times[0,T]\to\mathbb{R}$ is an increasing $\mathbb{F}$-adapted process with $A(0)=0$, $\mathbb{P}$-almost surely. In the following definition, we present the reference Banach spaces. They represent the continuous analogous of the càdlàg processes introduced in Section 1.2.

Definition 2.2. *Let $p\geq 2$ and $\beta\geq 0$.*

(i) *$\mathcal{S}^{p,m}$ denotes the space of continuous $\mathbb{F}$–progressively measurable processes $Y:\Omega\times[0,T]\to\mathbb{R}^m$ such that*

$$\mathbb{E}\left[\sup\nolimits_{0\leq s\leq T}|Y(s)|^p\right]<+\infty.$$

(ii) *$\mathcal{S}_\beta^{p,m}$ denotes the space of continuous $\mathbb{F}$–progressively measurable processes $Y:\Omega\times[0,T]\to\mathbb{R}^m$ such that*

$$\mathbb{E}\left[\sup\nolimits_{0\leq s\leq T}e^{\beta A(s)}|Y(s)|^p\right]+\mathbb{E}\left[\int_0^T e^{\beta A(s)}|Y(s)|^2dA(s)\right]^{p/2}<+\infty.$$

(iii) *$\mathcal{H}_\beta^{p,m\times d}$ denotes the space of $\mathbb{F}$–progressively measurable processes $Z:\Omega\times[0,T]\to\mathbb{R}^{m\times d}$*

such that

$$\mathbb{E}\left[\int_0^T e^{\beta A(s)}|Z(s)|^2 ds\right]^{p/2} < +\infty.$$

Instead of $\mathcal{H}_0^{p,m\times d}$ we will write $\mathcal{H}^{p,m\times d}$. The space $\mathcal{S}_\beta^{p,m} \times \mathcal{H}_\beta^{p,m\times d}$ (in fact, its quotient with respect to $\mathbb{P}\times\mathbb{P}dt$-a.e. equality) is naturally equipped with the following norm

$$\|(Y,Z)\|_{p,\beta}^p = \mathbb{E}\left[\sup_{0\le s\le T} e^{\beta A(s)}|Y(s)|^p\right] + \mathbb{E}\left[\int_0^T e^{\beta A(s)}|Y(s)|^2 dA(s)\right]^{p/2} + \mathbb{E}\left[\int_0^T e^{\beta A(s)}|Z(s)|^2 ds\right]^{p/2}.$$

2.3 Existence and uniqueness

We consider the following BSDE

$$Y(t) = \xi + \int_t^T F(s,Y(s),Z(s),Y_s,Z_s)ds + \int_t^T G(s,Y(s),Y_s)dA(s) - \int_t^T Z(s)dW(s), \quad t\in[0,T], \tag{2.3}$$

with $\xi \in L^2(\Omega,\mathcal{F}_T,\mathbb{P};\mathbb{R}^m)$ and the generators $F:\Omega\times[0,T]\times\mathbb{R}^m\times\mathbb{R}^{m\times d}\times L^2([-\delta,0];\mathbb{R}^m)\times L^2([-\delta,0];\mathbb{R}^{m\times d})\to\mathbb{R}^m$, $G:\Omega\times[0,T]\times\mathbb{R}^m\times L^2([-\delta,0];\mathbb{R}^m)\to\mathbb{R}^m$ such that the functions $F(\cdot,y,z,\hat{y},\hat{z})$ and $G(\cdot,y,\hat{y})$ are $\mathbb{F}$–progressively measurable, for any $(y,z,\hat{y},\hat{z})\in\mathbb{R}^m\times\mathbb{R}^{m\times d}\times L^2([-\delta,0];\mathbb{R}^m)\times L^2([-\delta,0];\mathbb{R}^{m\times d})$, respectively for any $(y,\hat{y})\in\mathbb{R}^m\times L^2([-\delta,0];\mathbb{R}^m)$.

Recall that, for a function $\boldsymbol{x}:[-\delta,T]\to\mathbb{R}^n$ and some $t\in[0,T]$, $\boldsymbol{x}_t:[-\delta,0]\to\mathbb{R}^n$ denotes the delayed-path of $\boldsymbol{x}$, defined as

$$\boldsymbol{x}_t(\theta) := \boldsymbol{x}(t+\theta),\ \theta\in[-\delta,0].$$

In order to define Y_s and Z_s even for $s<\delta$, we prolong by convention, Y by $Y(0)$ and Z by 0 on the negative real axis.

In what follows we present the assumptions required in this section. We suppose that there exist constants β, L, $\tilde{L}>0$, bounded progressively measurable stochastic processes $K,\tilde{K}:\Omega\times[0,T]\to\mathbb{R}_+$ and ρ, $\tilde{\rho}$ probability measures on $([-\delta,0],\mathcal{B}([-\delta,0]))$ such that:

(B_0) $\mathbb{E}\left[e^{\beta A(T)}\left(1+|\xi|^2\right)\right] < +\infty$;

(B_1) $\mathbb{E}\left[\int_0^T e^{\beta A(t)}|F(t,0,0,0,0)|^2 dt + \int_0^T e^{\beta A(t)}|G(t,0,0)|^2 dA(t)\right] < +\infty.$

(B$_2$) for any $t \in [0,T]$, $(y,z), (y',z') \in \mathbb{R}^m \times \mathbb{R}^{m\times d}$, $\hat{y}, \hat{y}' \in L^2([-\delta,0];\mathbb{R}^m)$ and $\hat{z}, \hat{z}' \in L^2([-\delta,0];\mathbb{R}^{m\times d})$, we have

$$(i) \quad |F(t,y,z,\hat{y},\hat{z}) - F(t,y',z',\hat{y},\hat{z})| \leq L(|y-y'| + |z-z'|), \ \mathbb{P}\text{-a.s.};$$

$$\begin{aligned}(ii) \quad & |F(t,y,z,\hat{y},\hat{z}) - F(t,y,z,\hat{y}',\hat{z}')|^2 \\ & \leq K(t) \int_{-\delta}^{0} \left(|\hat{y}(\theta) - \hat{y}'(\theta)|^2 + |\hat{z}(\theta) - \hat{z}'(\theta)|^2 \right) \rho(d\theta), \ \mathbb{P}\text{-a.s.};\end{aligned}$$

(B$_3$) for any $t \in [0,T]$, $y, y' \in \mathbb{R}^m$ and $\hat{y}, \hat{y}' \in L^2([-\delta,0];\mathbb{R}^m)$, we have

$$(i) \quad |G(t,y,\hat{y}) - G(t,y',\hat{y})| \leq \tilde{L}|y-y'|, \ \mathbb{P}\text{-a.s.};$$

$$(ii) \quad |G(t,y,\hat{y}) - G(t,y,\hat{y}')|^2 \leq \tilde{K}(t) \int_{-\delta}^{0} |\hat{y}(\theta) - \hat{y}'(\theta)|^2 \tilde{\rho}(d\theta), \ \mathbb{P}\text{-a.s.};$$

Remark 2.1. *Let us underline that latter conditions differ from those used in [72], since we allow T to be arbitrary, but different from the delay $\delta \in [0,T]$. This allows to separate the Lipschitz constant L w.r.t. (y,z) from the Lipschitz constant K w.r.t. $(\hat{y},\hat{z})$; therefore the restriction on the coefficients can avoid the constant L.*

Remark 2.2. *Existence and uniqueness of a solution to the backward system (2.3) will be proved exploiting a standard Banach's fixed point argument which requires K or δ to be small enough.*

More precisely, by denoting $K_1 := \sup_{s\in[0,T]} K(s)$, $\tilde{K}_1 := \sup_{s\in[0,T]} \tilde{K}(s)$ and

$$\omega_\delta := \sup_{t\in[0,T-\delta]} \left(A(t+\delta) - A(t) \right),$$

we will assume the existence of a positive constant $c < c_{\beta,\tilde{L}} := \min\left\{ \frac{\beta^2 - 8\tilde{L}^2}{4\beta^2}, \frac{1}{584} \right\}$ such that

(H$_1$) $K_1 \cdot \max\{1,T\} \cdot \frac{e^{(8L^2+\frac{1}{2})\delta + \beta\omega_\delta}}{4L^2} \leq c, \quad \mathbb{P}$*-a.s.;*

(H$_2$) $4\tilde{K}_1 \cdot A(T) \cdot \frac{e^{(8L^2+\frac{1}{2})\delta + \beta\omega_\delta}}{\beta} \leq c, \quad \mathbb{P}$*-a.s.*

Our first result states existence and uniqueness of equation (2.3).

Theorem 2.3. *Let us assume that* (B$_0$)–(B$_3$) *hold true and $\beta > 2\sqrt{2}\tilde{L}$. If conditions* (H$_1$) *and* (H$_2$) *are satisfied then there exists a unique solution $(Y,Z) \in \mathcal{S}_\beta^{2,m} \times \mathcal{H}_\beta^{2,m\times d}$ for (2.3).*

Proof. The existence and uniqueness will be obtained by the Banach fixed point theorem.

Let us consider the map $\Gamma : \mathcal{S}^{2,m}_\beta \times \mathcal{H}^{2,m\times d}_\beta \to \mathcal{S}^{2,m}_\beta \times \mathcal{H}^{2,m\times d}_\beta$, defined in the following way: for $(U,V) \in \mathcal{S}^{2,m}_\beta \times \mathcal{H}^{2,m\times d}_\beta$, $\Gamma(U,V) = (Y,Z)$, where the couple of adapted processes (Y,Z) is the solution to the equation

$$Y(t) = \xi + \int_t^T F(s,Y(s),Z(s),U_s,V_s)ds + \int_t^T G(s,U(s),U_s)dA(s) \\ - \int_t^T Z(s)dW(s), \quad t \in [0,T]. \quad (2.4)$$

The existence of a unique solution $(Y,Z) \in \mathcal{S}^{2,m} \times \mathcal{H}^{2,m\times d}$ is guaranteed by [179]. Indeed, if we denote

$$B(t) := \int_0^t G(s,U(s),U_s)dA(s), \quad t \in [0,T];$$
$$\hat{F}(t,y,z) := F(t,y - B(t),z,U_t,V_t), \quad t \in [0,T],\ (y,z) \in \mathbb{R}^m \times \mathbb{R}^{m\times d},$$

then (Y,Z) is a solution to equation (2.3) if and only if $(Y + B,Z)$ solves the equation

$$\hat{Y}(t) = \xi + B(T) + \int_t^T \hat{F}(s,\hat{Y}(s),Z(s),U_s,V_s)ds - \int_t^T Z(s)dW(s), \quad t \in [0,T].$$

Since $\hat{F}$ is Lipschitz with respect to (y,z), it remains to prove that $\mathbb{E}\int_0^T |\hat{F}(t,0,0)|^2 dt < +\infty$ and $\xi + B(T) \in L^2(\Omega,\mathcal{F}_T,\mathbb{P};\mathbb{R}^m)$. We have (remember that $K_1 := \sup_{s\in[0,T]} K(s)$ and $\tilde{K}_1 := \sup_{s\in[0,T]} \tilde{K}(s)$):

$$\begin{aligned}
&\mathbb{E}\int_0^T |\hat{F}(t,0,0)|^2 dt = \mathbb{E}\int_0^T |F(t,-B(t),0,U_t,V_t)|^2 dt \le 3\mathbb{E}\int_0^T |F(t,0,0,0,0)|^2 dt \\
&\quad + 3L^2\mathbb{E}\int_0^T |B(t)|^2 dt + 3\mathbb{E}\int_0^T K(t)\int_{-\delta}^0 \left(|U(t+\theta)|^2 + |V(t+\theta)|^2\right)\rho(d\theta)dt \\
&\le 3\mathbb{E}\int_0^T |F(t,0,0,0,0)|^2 dt + 3L^2\mathbb{E}\int_0^T |B(t)|^2 dt \\
&\quad + 3T\mathbb{E}\left[K_1 \sup_{t\in[0,T]} |U(t)|^2\right] + 3\mathbb{E}K_1\int_0^T |V(t)|^2 dt.
\end{aligned}$$

Since (B_1) holds and K_1 is bounded, we only have to show that $\mathbb{E}\int_0^T |B(t)|^2 dt < +\infty$ and

$\mathbb{E}|B(T)|^2 < +\infty$. We have

$$\begin{aligned}
&\mathbb{E}\int_0^T \left|\int_0^t G(s,U(s),U_s)dA(s)\right|^2 dt \\
&\leq \mathbb{E}\int_0^T \left[\int_0^t e^{\beta A(s)} |G(s,U(s),U_s)|^2 dA(s) \cdot \int_0^t e^{-\beta A(s)} dA(s)\right] dt \\
&\leq \frac{T}{\beta}\mathbb{E}\int_0^T e^{\beta A(t)} |G(t,U(t),U_t)|^2 dA(t) \leq \frac{2T}{\beta}\mathbb{E}\int_0^T e^{\beta A(t)} |G(t,0,0)|^2 dA(t) \\
&\quad + \frac{2T}{\beta}\mathbb{E}\int_0^T e^{\beta A(t)} L^2 |U(t)|^2 dA(t) + \frac{2T}{\beta}\mathbb{E}\int_0^T e^{\beta A(t)} \tilde{K}(t) \int_{-\delta}^0 |U(t+\theta)|^2 \tilde{\rho}(d\theta) dA(t) \\
&\leq \frac{2T}{\beta}\mathbb{E}\int_0^T e^{\beta A(t)} |G(t,0,0)|^2 dA(t) + \frac{2TL^2}{\beta}\mathbb{E}\int_0^T e^{\beta A(t)} |U(t)|^2 dA(t) \\
&+ \frac{2T}{\beta}\mathbb{E}\tilde{K}_1 A(T) e^{\beta\omega_\delta} \sup_{t\in[0,T]} e^{\beta A(t)} |U(t)|^2 < +\infty,
\end{aligned}$$

by (B_1) and (H_2), which proves the claim (along the way we have also proven that $\mathbb{E}|B(T)|^2 < +\infty$).

The proof that $(Y,Z) \in \mathcal{S}_\beta^{2,m} \times \mathcal{H}_\beta^{2,m\times d}$ is very similar to that of Proposition 1.1 from [181], so it is left to the reader.

Let us prove that Γ is a contraction with respect to the equivalent norm

$$\|(Y,Z)\|_{2,\alpha,\beta,a,b}^2 := \mathbb{E}\big(\sup\nolimits_{t\in[0,T]} e^{\alpha t+\beta A(t)} |Y(t)|^2\big) + a\mathbb{E}\int_0^T e^{\alpha s+\beta A(s)} |Y(s)|^2 dA(s) + b\mathbb{E}\int_0^T e^{\alpha s+\beta A(s)} |Z(s)|^2 ds.$$

where $\alpha := 8L^2 + \frac{1}{2}$ and the constants $a,b > 0$ are yet to be chosen.

Let us consider $(U^1,V^1),(U^2,V^2) \in \mathcal{S}_\beta^{2,m} \times \mathcal{H}_\beta^{2,m\times d}$ and $(Y^1,Z^1) := \Gamma(U^1,V^1)$, $(Y^2,Z^2) := \Gamma(U^2,V^2)$. For the sake of brevity, we will denote in what follows

$$\begin{aligned}
&\Delta F(s) := F(s,Y^1(s),Z^1(s),U_s^1,V_s^1) - F(s,Y^2(s),Z^2(s),U_s^2,V_s^2), \\
&\Delta G(s) := G(s,U^1(s),U_s^1) - G(s,U^2(s),U_s^2), \\
&\Delta U(s) := U^1(s) - U^2(s), \quad \Delta V(s) := V^1(s) - V^2(s), \\
&\Delta Y(s) := Y^1(s) - Y^2(s), \quad \Delta Z(s) := Z^1(s) - Z^2(s).
\end{aligned}$$

Exploiting Itô's formula we have, for any $t \in [0,T]$

$$
\begin{aligned}
&e^{\alpha t+\beta A(t)}|\Delta Y(t)|^2+\int_t^T e^{\alpha s+\beta A(s)}|\Delta Y(s)|^2(\alpha ds+\beta dA(s))+\int_t^T e^{\alpha s+\beta A(s)}|\Delta Z(s)|^2 ds\\
&=e^{\alpha T+\beta A(T)}|\Delta Y(T)|^2-2\int_t^T e^{\alpha s+\beta A(s)}\langle \Delta Y(s),\Delta Z(s)dW(s)\rangle\\
&\quad+2\int_t^T e^{\alpha s+\beta A(s)}\langle \Delta Y(s),\Delta F(s)\rangle ds+2\int_t^T e^{\alpha s+\beta A(s)}\langle \Delta Y(s),\Delta G(s)\rangle dA(s)\,.
\end{aligned}
$$

From assumptions (B_2)–(B_3) we obtain,

$$
\begin{aligned}
&2\Big|\int_t^T e^{\alpha s+\beta A(s)}\langle \Delta Y(s),\Delta F(s)\rangle ds\Big|\le 2\int_t^T e^{\alpha s+\beta A(s)}|\langle \Delta Y(s),\Delta F(s)\rangle| ds\\
&\le 8L^2\int_t^T e^{\alpha s+\beta A(s)}|\Delta Y(s)|^2 ds+\frac{1}{8L^2}\int_t^T e^{\alpha s+\beta A(s)}|\Delta F(s)|^2 ds\\
&\le 8L^2\int_t^T e^{\alpha s+\beta A(s)}|\Delta Y(s)|^2 ds+\frac{1}{2}\int_t^T e^{\alpha s+\beta A(s)}\left(|\Delta Y(s)|^2+|\Delta Z(s)|^2\right)ds\\
&\quad+\frac{K_1 T}{4L^2}e^{\alpha\delta+\beta\omega_\delta}\cdot\sup\nolimits_{s\in[0,T]}\left(e^{\alpha s+\beta A(s)}|\Delta U(s)|^2\right)\\
&\quad+\frac{K_1}{4L^2}e^{\alpha\delta+\beta\omega_\delta}\cdot\int_0^T e^{\alpha s+\beta A(s)}|\Delta V(s)|^2 ds
\end{aligned}
$$

and

$$
\begin{aligned}
&2\Big|\int_t^T e^{\alpha s+\beta A(s)}\langle \Delta Y(s),\Delta G(s)\rangle dA(s)\Big|\le 2\int_t^T e^{\alpha s+\beta A(s)}|\langle \Delta Y(s),\Delta G(s)\rangle| dA(s)\\
&\le \frac{\beta}{2}\int_t^T e^{\alpha s+\beta A(s)}|\Delta Y(s)|^2 dA(s)+\frac{2}{\beta}\int_t^T e^{\alpha s+\beta A(s)}|\Delta G(s)|^2 dA(s)\\
&\le \frac{\beta}{2}\int_t^T e^{\alpha s+\beta A(s)}|\Delta Y(s)|^2 dA(s)+\frac{4\tilde{L}^2}{\beta}\int_t^T e^{\alpha s+\beta A(s)}|\Delta U(s)|^2 dA(s)\\
&\quad+\frac{4\tilde{K}_1\,A(T)}{\beta}e^{\alpha\delta+\beta\omega_\delta}\cdot\sup\nolimits_{s\in[0,T]}\left(e^{\alpha s+\beta A(s)}|\Delta U(s)|^2\right).
\end{aligned}
$$

By (H_1) and (H_2), we have

$$
\begin{aligned}
\Big(\tfrac{K_1 T}{4L^2}+\tfrac{4\tilde{K}_1\,A(T)}{\beta}\Big)e^{\alpha\delta+\beta\omega_\delta}&\le 2c,\quad \mathbb{P}\text{-a.s.};\\
\tfrac{K_1}{4L^2}e^{\alpha\delta+\beta\omega_\delta}&\le c,\quad \mathbb{P}\text{-a.s},
\end{aligned}
$$

(recall that $\alpha := 8L^2 + \frac{1}{2}$). Therefore,

$$\begin{aligned}
&e^{\alpha t+\beta A(t)}|\Delta Y(t)|^2 + \frac{\beta}{2}\int_t^T e^{\alpha s+\beta A(s)}|\Delta Y(s)|^2 dA(s) \\
&\quad + \frac{1}{2}\int_t^T e^{\alpha s+\beta A(s)}|\Delta Z(s)|^2 ds \\
&\leq -2\int_t^T e^{\alpha s+\beta A(s)}\langle \Delta Y(s), \Delta Z(s)\, dW(s)\rangle + \frac{4\tilde{L}^2}{\beta}\int_t^T e^{\alpha s+\beta A(s)}|\Delta U(s)|^2 dA(s) \\
&\quad + 2c\sup_{s\in[0,T]}\left(e^{\alpha s+\beta A(s)}|\Delta U(s)|^2\right) + c\int_0^T e^{\alpha s+\beta A(s)}|\Delta V(s)|^2 ds.
\end{aligned} \tag{2.5}$$

Since $e^{\alpha s+\beta A(s)}\Delta Y \in \mathcal{S}^{2,m}$ and $\Delta Z \in \mathcal{H}^{2,m\times d}$, one can show that

$$\mathbb{E}\left[\int_0^T e^{\alpha s+\beta A(s)}\langle \Delta Y(s), \Delta Z(s)\, dW(s)\rangle\right] = 0,$$

hence

$$\begin{aligned}
&\frac{\beta}{2}\mathbb{E}\int_0^T e^{\alpha s+\beta A(s)}|\Delta Y(s)|^2 dA(s) + \frac{1}{2}\mathbb{E}\int_0^T e^{\alpha s+\beta A(s)}|\Delta Z(s)|^2 ds \\
&\leq \frac{4\tilde{L}^2}{\beta}\mathbb{E}\int_0^T e^{\alpha s+\beta A(s)}|\Delta U(s)|^2 dA(s) + 2c\mathbb{E}\left[\sup_{s\in[0,T]}\left(e^{\alpha s+\beta A(s)}|\Delta U(s)|^2\right)\right] \\
&\quad + c\mathbb{E}\left[\int_0^T e^{\alpha s+\beta A(s)}|\Delta V(s)|^2 ds\right].
\end{aligned} \tag{2.6}$$

On the other hand, by Burkholder–Davis–Gundy's inequality, we have

$$\begin{aligned}
&2\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_t^T e^{\alpha s+\beta A(s)}\langle \Delta Y(s), \Delta Z(s)\rangle dW(s)\right|\right] \\
&\qquad\qquad \leq \frac{1}{2}\mathbb{E}\left(\sup_{t\in[0,T]} e^{\alpha t+\beta A(t)}|\Delta Y(t)|^2\right) + 72\,\mathbb{E}\int_0^T e^{\alpha s+\beta A(s)}|\Delta Z(s)|^2 ds.
\end{aligned}$$

Hence, by (2.5),

$$\begin{aligned}
&\frac{1}{2}\mathbb{E}\left(\sup_{t\in[0,T]} e^{\alpha t+\beta A(t)}|\Delta Y(t)|^2\right) \\
&\leq 72\,\mathbb{E}\int_0^T e^{\alpha s+\beta A(s)}|\Delta Z(s)|^2 ds + \frac{4\tilde{L}^2}{\beta}\mathbb{E}\int_0^T e^{\alpha s+\beta A(s)}|\Delta U(s)|^2 dA(s) \\
&\quad + 2c\mathbb{E}\left[\sup_{s\in[0,T]}\left(e^{\alpha s+\beta A(s)}|\Delta U(s)|^2\right)\right] + c\mathbb{E}\left[\int_0^T e^{\alpha s+\beta A(s)}|\Delta V(s)|^2 ds\right].
\end{aligned}$$

Thus, with $a := \frac{\lambda\beta}{2}$, $b := \frac{\lambda}{2} - 144$ and some $\lambda > 288$, by taking into account (2.6), we obtain

$$\begin{aligned}
&\mathbb{E}\big(\sup_{t\in[0,T]} e^{\alpha t+\beta A(t)}|\Delta Y(t)|^2\big) + a\int_0^T e^{\alpha s+\beta A(s)}|\Delta Y(s)|^2 dA(s) \\
&\quad + b\mathbb{E}\int_0^T e^{\alpha s+\beta A(s)}|\Delta Z(s)|^2 ds \\
&\leq 2c(2+\lambda)\,\mathbb{E}\left[\sup_{s\in[0,T]}\left(e^{\alpha s+\beta A(s)}|\Delta U(s)|^2\right)\right] + \frac{4\tilde{L}^2}{\beta}(2+\lambda)\mathbb{E}\int_0^T e^{\alpha s+\beta A(s)}|\Delta U(s)|^2 dA(s) \\
&\quad + c(2+\lambda)\,\mathbb{E}\int_0^T e^{\alpha r+\beta A(r)}|\Delta V(r)|^2 dr,
\end{aligned}$$

so

$$\|(\Delta Y, \Delta Z)\|^2_{2,\alpha,\beta,a,b} \leq \mu_\lambda \|(\Delta U, \Delta V)\|^2_{2,\alpha,\beta,a,b},$$

where

$$\mu_\lambda := \max\left\{c(2+\lambda), \tfrac{8\tilde{L}^2(2+\lambda)}{\lambda\beta^2}, \tfrac{2c(2+\lambda)}{\lambda-288}\right\}.$$

Since $c < c_{\beta,\tilde{L}}$, we can take λ slightly bigger than $\frac{1}{2c_{\beta,\tilde{L}}} - 2$, such that $2c(2+\lambda) < 1$ and so $\mu_\lambda < 1$ (by the definition of $c_{\beta,\tilde{L}}$).

It follows that the application Γ is a contraction on the Banach space $\mathcal{S}^{2,m}_\beta \times \mathcal{H}^{2,m\times d}_\beta$. Therefore, by Banach fixed point theorem, there exists a unique fixed point $(Y,Z) = \Gamma(Y,Z)$ in the space $\mathcal{S}^{2,m}_\beta \times \mathcal{H}^{2,m\times d}_\beta$, which completes our proof. □

Remark 2.4. *Let us underlined that the condition on A to be increasing can be relaxed assuming it to be a continuous bounded variation $\mathbb{F}$-adapted process with $A_0 = 0$, $\mathbb{P}$-a.s. Indeed, by considering the increasing process $\tilde{A}(t) := \|A\|_{BV([0,t])}$, $t \in [0,T]$ and the Radon–Nikodym derivative $\gamma(t) := \frac{dA(t)}{d\tilde{A}(t)}$, $t \in [0,T]$, we have that $|\gamma(t)| \leq 1$, $\forall t \in [0,T]$, $\mathbb{P}$-a.s. and the BSDE (2.3) can be rewritten as*

$$\begin{aligned}
Y(t) = \xi + \int_t^T F(s, Y(s), Z(s), Y_s, Z_s)ds + \int_t^T \tilde{G}(s, Y(s), Y_s)d\tilde{A}(s) \\
- \int_t^T Z(s)dW(s), \quad t \in [0,T],
\end{aligned}$$

where the new coefficient $\tilde{G} : \Omega \times [0,T] \times \mathbb{R}^m \times L^2([-\delta,0];\mathbb{R}^m) \to \mathbb{R}^m$ is defined as $\tilde{G}(t,y,\hat{y}) := \gamma(t)G(t,y,\hat{y})$, still satisfying the condition (B$_3$), *by replacing A with $\tilde{A}$ and with the same $\tilde{L}$ and $\tilde{K}$.*

As shown in [72], conditions as (H$_1$) and (H$_2$) restricting the magnitude of the delay are necessary. However, in the same paper the authors provide some examples ($F \equiv KY(t-T)$

and $F \equiv K\int_0^t Y(s)ds$, with $K \le 0$) in which the delay can be considered of arbitrary length. The next result is a first attempt to get rid of the restrictive assumptions concerning the delay, by imposing monotonicity and linearity on generators F and G.

More precisely, we assume that $m=1$, $\xi \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$ and we require F and G not depending on Z_s), namely $F : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^d \times L^2([-\delta,0]) \to \mathbb{R}$ and $G : \Omega \times [0,T] \times \mathbb{R} \times L^2([-\delta,0]) \to \mathbb{R}$. Moreover, we require that:

(D$_1$) $\hat{y} \mapsto F(t,y,z,\hat{y})$ and $\hat{y} \mapsto G(t,y,\hat{y})$ are non-increasing with respect to the positive cone of $L^2([-\delta,0])$ for all $(t,y,z) \in [0,T] \times \mathbb{R} \times \mathbb{R}^d$, $\mathbb{P}$-a.s.;

(D$_2$) $F(t,y,z,\hat{y}) = F_0(t) + F_1(y,z,\hat{y})$, $G(t,y,\hat{y}) = G_0(t) + G_1(y,\hat{y})$, with F_1 and G_1 linear.

Thus, the BSDE (2.3) reduces to the following one:

$$Y(t) = \xi + \int_t^T [F_0(s) + F_1(Y(s), Z(s), Y_s)ds] + \int_t^T [G_0(s) + G_1(Y(s), Y_s)]\, dA(s) \\ - \int_t^T Z(s)dW(s), \quad t \in [0,T]. \quad (2.7)$$

Proposition 2.5. *Assume conditions* (D$_1$), (D$_2$) *and* (B$_0$)–(B$_3$) *hold. If $\beta > 2\sqrt{2}\tilde{L}$, then there exists a solution $(Y,Z) \in \mathcal{S}^{2,1}_\beta \times \mathcal{H}^{2,1\times d}_\beta$ for (2.7).*

Proof. As in the proof of theorem 2.3, we consider the map $\Gamma : \mathcal{S}^{2,1}_\beta \to \mathcal{S}^{2,1}_\beta$, defined in the following way: for $U \in \mathcal{S}^{2,1}_\beta$, $\Gamma(U) = Y$, where the couple of adapted processes (Y,Z) is the solution to the equation

$$Y(t) = \xi + \int_t^T [F_0(s) + F_1(Y(s), Z(s), U_s)ds] + \int_t^T [G_0(s) + G_1(Y(s), U_s)]\, dA(s) \\ - \int_t^T Z(s)dW(s), \quad t \in [0,T].$$

Using the same type of computations as in the above proof, it is easy to see that even without conditions (H$_1$) and (H$_2$), Γ is still a Lipschitz-continuous function. By a classical comparison theorem for BSDEs, if $U^1(t) \le U^2(t)$ $\mathbb{P}dt$-a.e., then $Y^1(t) \le Y^2(t)$, $\forall t \in [0,T]$, $\mathbb{P}$-a.s., with $Y^i(t) := \Gamma(U^i)$, $i = \overline{1,2}$. This shows that Γ is non-increasing with respect to the positive cone of $\mathcal{S}^{2,1}_\beta$.

One can use now an argument from [154, Theorem 2.2] to show that there exist $\underline{U}, \overline{U} \in \mathcal{S}^{2,1}_\beta$ such that $\Gamma([\underline{U}, \overline{U}]) \subseteq [\underline{U}, \overline{U}]$, where $[\underline{U}, \overline{U}] := \left\{U \in \mathcal{S}^{2,1}_\beta \mid \underline{U}(t) \le U(t) \le \overline{U}(t),\ \mathbb{P}dt\text{-a.e.}\right\}$. Obviously, $[\underline{U}, \overline{U}]$ is a closed, convex set of the Banach space $\mathcal{S}^{2,1}_\beta$.

Let $Y^0 := \underline{U}$ and, by recursion, $Y^{n+1} := \Gamma(Y^n)$. By the monotonicity property of Γ, it is easy to show that $\forall t \in [0,T]$, $\mathbb{P}$-a.s.,

$$\underline{U}(t) = Y^0(t) \le Y^2(t) \le \cdots \le Y^{2n}(t) \le \cdots \le Y^{2n+1}(t) \le \cdots \le Y^3(t) \le Y^1(t) \le \overline{U}(t).$$

Let $\underline{Y}(t) := \lim_{n\to\infty} Y^{2n}(t)$ and $\overline{Y}(t) := \lim_{n\to\infty} Y^{2n+1}(t)$. Since $\underline{U}, \overline{U} \in \mathcal{S}_\beta^{2,1}$, for any $H \in L^2(\Omega; BV[0,T])$ or $H \in L^2(\Omega \times [0,T], \mathbb{P}dA(\cdot))$ we have, by the dominated convergence theorem,

$$\lim_{n\to\infty} \mathbb{E}\int_0^T e^{\beta A(t)} Y^{2n}(t) H(t)dt = \mathbb{E}\int_0^T e^{\beta A(t)} \underline{Y}(t)H(t)dt \text{ and}$$
$$\lim_{n\to\infty} \mathbb{E}\int_0^T e^{\beta A(t)} Y^{2n+1}(t) H(t)dt = \mathbb{E}\int_0^T e^{\beta A(t)} \overline{Y}(t)H(t)dt.$$

Hence $(e^{\beta A(\cdot)/2}Y^{2n})$ and $(e^{\beta A(\cdot)/2}Y^{2n+1})$ converge weakly to $e^{\beta A(\cdot)/2}\underline{Y}$, respectively $e^{\beta A(\cdot)/2}\overline{Y}$, in both $L^2(\Omega; C[0,T])$ and $L^2(\Omega \times [0,T], \mathbb{P}dA(\cdot))$. By Mazur's lemma (applied two times), for any $n \in \mathbb{N}$ there are convex combinations, let us call them $\underline{Y}^n$ and $\overline{Y}^n$, of the elements of $(Y^{2k})_{k\ge n}$, respectively $(Y^{2k+1})_{k\ge n}$, such that $(e^{\beta A(\cdot)/2}\underline{Y}^n)$ and $(e^{\beta A(\cdot)/2}\overline{Y}^n)$ converge strongly in both $L^2(\Omega; C[0,T])$ and $L^2(\Omega\times[0,T], \mathbb{P}dA(\cdot))$ to $e^{\beta A(\cdot)/2}\underline{Y}$, respectively $e^{\beta A(\cdot)/2}\overline{Y}$. Therefore, $(\underline{Y}^n)$ and $(\overline{Y}^n)$ converge strongly in $\mathcal{S}_\beta^{2,1}$ to $\underline{Y}$, respectively $e^{\beta A(\cdot)}\overline{Y}$; thus, $\lim_{n\to\infty}\Gamma(\underline{Y}^n) = \Gamma(\underline{Y})$ and $\lim_{n\to\infty}\Gamma(\overline{Y}^n) = \Gamma(\overline{Y})$.

On the other hand, by the linearity of F_1 and G_1, $\Gamma(\underline{Y}^n)$ and $\Gamma(\overline{Y}^n)$ are convex combinations of the elements of $(Y^{2k+1})_{k\ge n}$, respectively $(Y^{2k})_{k\ge n}$, so $e^{\beta A(\cdot)/2}\Gamma(\underline{Y}^n)$ and $e^{\beta A(\cdot)/2}\Gamma(\overline{Y}^n)$ converge pointwisely to $e^{\beta A(\cdot)/2}\overline{Y}$, respectively $e^{\beta A(\cdot)/2}\underline{Y}$. Consequently, $\Gamma(\underline{Y}) = \overline{Y}$ and $\Gamma(\overline{Y}) = \underline{Y}$. Then, setting $Y = \frac{1}{2}\underline{Y} + \frac{1}{2}\overline{Y}$, we have $\Gamma(Y) = Y$, which proves our claim. $\square$

2.4 Dependence on parameters

Let us consider, for all $n \in \mathbb{N}^*$, the following BSDEs which approximate (2.3):

$$Y^n(t) = \xi^n + \int_t^T F^n(s, Y^n(s), Z^n(s), Y_s^n, Z_s^n)ds + \int_t^T G^n(s, Y^n(s), Y_s^n)dA^n(s) - \int_t^T Z^n(s)dW(s), \quad t \in [0,T], \quad (2.8)$$

In order to unify the notations, we will sometimes denote ς^0 instead of ς, if ς is ξ, A, F, G, Y or Z. We suppose that the coefficients ξ^n, A^n, F^n, G^n, $n \ge 0$, satisfy conditions (B_2)–(B_3), (H_1), (H_2) with processes K^n, $\tilde{K}^n$, but the same constants β, c, L, $\tilde{L}$. Moreover, we have to impose that $\beta > 2\sqrt{2}\tilde{L}$.

We suppose that there exists $p > 1$ such that

(B$_0'$) $\sup_{n\in\mathbb{N}}\mathbb{E}\left[e^{p\beta A^n(T)}|\xi^n|^{2p}\right]<+\infty$.

(B$_0''$) $\sup_{n\in\mathbb{N}}\mathbb{E}\left[e^{qA^n(T)}\right]<+\infty$, for any $q>0$.

(B$_1'$) $\sup_{n\in\mathbb{N}}\mathbb{E}\left[\left(\int_0^T e^{\beta A^n(t)}|F^n(t,0,0,0,0)|^2\,dt\right)^p+\left(\int_0^T e^{\beta A^n(t)}|G^n(t,0,0)|^2\,dA^n(t)\right)^p\right]<+\infty$.

Under these assumptions, there exists a unique solution $(Y^n,Z^n)\in\mathcal{S}_\beta^{2,m}\times\mathcal{H}_\beta^{2,m\times d}$ to equation (2.8). In fact, one can now prove by standard computations that $(Y^n,Z^n)\in\mathcal{S}_\beta^{p,m}\times\mathcal{H}_\beta^{p,m\times d}$, $\forall n\in\mathbb{N}$ and

$$\sup_{n\in\mathbb{N}}\|(Y^n,Z^n)\|_{p,\beta}<+\infty. \tag{2.9}$$

Our aim is to show that if the coefficients (ξ^n,A^n,F^n,G^n) of equation (2.8) converge to (ξ,A,F,G), then (Y^n,Z^n) converge to (Y,Z) in $\mathcal{S}^{2,m}\times\mathcal{H}^{2,m\times d}$. Let now specify in what sense the convergence of the coefficients takes place. We define

$$\Delta_nF:=\sup_{t\in[0,T],\ (y,z)\in\mathbb{R}^m\times\mathbb{R}^{m\times d},\ (\hat{y},\hat{z})\in L^2\left([-\delta,0];\mathbb{R}^m\times\mathbb{R}^{m\times d}\right)}|F^n(t,y,z,\hat{y},\hat{z})-F(t,y,z,\hat{y},\hat{z})|;$$
$$\Delta_nG:=\sup_{t\in[0,T],\ y\in\mathbb{R}^m\times\mathbb{R}^{m\times d},\ \hat{y}\in L^2([-\delta,0];\mathbb{R}^m)}|G^n(t,y,\hat{y})-G(t,y,\hat{y})|$$

and impose

(C$_1$) $\mathbb{E}\left[|\xi^n-\xi|^{2p}\right]\to 0$ as $n\to\infty$;

(C$_2$) $\mathbb{E}\sup_{t\in[0,T]}|A^n(t)-A(t)|\to 0$ as $n\to\infty$;

(C$_3$) $\mathbb{E}[(\Delta_nF)^{2p}+(\Delta_nG)^{2p}]\to 0$ as $n\to\infty$.

The uniform convergence from assumption (C$_3$) can be relaxed to a weaker type of convergence; however, we will work with this hypothesis for the sake of keeping computations as simple as possible.

We now state a technical result that needed for the proof of Theorem 2.8: it is a variant of the Helly-Bray theorem for the stochastic case and is also stronger than Proposition 3.4 from [222]:

Proposition 2.6. *Let $(X_n,H_n):(\Omega_n,\mathcal{F}_n,\mathbb{P}_n)\to C([0,T];\mathbb{R}^d)^2$, $n\geq 1$, be a sequence of random variables, converging in distribution to a random variable $(X,H):(\Omega,\mathcal{F},\mathbb{P})\to C([0,T];\mathbb{R}^d)^2$. If for all $n\geq 1$, H_n is $\mathbb{P}_n$-a.s. with bounded variation and*

$$\lim_{\nu\to+\infty}\sup_{n\geq 1}\mathbb{P}_n\left(\|H_n\|_{BV([0,T];\mathbb{R}^d)}>\nu\right)=0, \tag{2.10}$$

then H is $\mathbb{P}$-a.s. with bounded variation and the sequence of $C[0,T]$-valued random variables $\left(\int_0^\cdot\langle X_n(s),dH_n(s)\rangle\right)_{n\geq 1}$ converges in distribution to $\int_0^\cdot\langle X(s),dH_n(s)\rangle$.

As expected, the proof of this result uses a deterministic Helly-Bray type theorem aiming uniform convergence. We state and prove this result:

Lemma 2.7. *Let* $(\boldsymbol{x}_n)_{n\geq 1} \subseteq C([0,T];\mathbb{R}^d)$ *and* $(\boldsymbol{\eta}_n)_{n\geq 1} \subseteq BV([0,T];\mathbb{R}^d)$ *be two sequences of functions such that:*

i) $\boldsymbol{x}_n$ *converges uniformly to a function* $\boldsymbol{x} \in C([0,T];\mathbb{R}^d)$*;*

ii) $\boldsymbol{\eta}_n$ *converges uniformly to a function* $\boldsymbol{\eta}$*;*

iii) $\sup_{n\geq 1} \|\boldsymbol{\eta}_n\|_{BV([0,T];\mathbb{R}^d)} < +\infty$.

Then $\boldsymbol{\eta} \in BV([0,T];\mathbb{R}^d)$*,* $\|\boldsymbol{\eta}\|_{BV([0,T];\mathbb{R}^d)} \leq \liminf_{n\to\infty} \|\boldsymbol{\eta}_n\|_{BV([0,T];\mathbb{R}^d)}$ *and the sequence of continuous functions* $\left(\int_0^{\cdot} \langle \boldsymbol{x}_n(s), d\boldsymbol{\eta}_n(s)\rangle\right)_{n\geq 1}$ *converges uniformly to* $\int_0^{\cdot} \langle \boldsymbol{x}(s), d\boldsymbol{\eta}(s)\rangle$*.*

Proof. The first two assertions are well-known, so we skip their proof.

Let us prove the last one. We say that a tuple $\pi = (t_0, \dots, t_k)$ is a *partition* of $[0,T]$ if $0 = t_0 < t_1 < \cdots < t_{k_N} = T$.

We consider $\pi^N = (t_0^N, \dots, t_{k_N}^N)$, $N \in \mathbb{N}^*$ partitions of the interval $[0,T]$ such that

$$\lim_{N\to\infty} \sup_{0\leq i < t_{k_N}^N} \left| t_{i+1}^N - t_i^N \right| = 0.$$

Let $\boldsymbol{x}^N : [0,T] \to \mathbb{R}^d$ be a step-function approximating $\boldsymbol{x}$, defined by

$$\boldsymbol{x}^N := \mathbf{1}_{\{0\}} \boldsymbol{x}(0) + \sum\nolimits_{i=1}^{k_N} \mathbf{1}_{(t_{i-1}, t_i]} \boldsymbol{x}(t_i).$$

Let $M := \sup_{n\geq 1} \|\boldsymbol{\eta}_n\|_{BV([0,T];\mathbb{R}^d)}$. Then

$$\begin{aligned}
&\left| \int_0^t \langle \boldsymbol{x}_n(s), d\boldsymbol{\eta}_n(s)\rangle - \int_0^t \langle \boldsymbol{x}(s), d\boldsymbol{\eta}(s)\rangle \right| \leq \left| \int_0^t \langle \boldsymbol{x}_n(s) - \boldsymbol{x}(s), d\boldsymbol{\eta}_n(s)\rangle \right| \\
&\quad + \left| \int_0^t \left\langle \boldsymbol{x}(s) - \boldsymbol{x}^N(s), d(\boldsymbol{\eta}_n - \boldsymbol{\eta})(s) \right\rangle \right| + \left| \int_0^t \left\langle \boldsymbol{x}^N(s), d(\boldsymbol{\eta}_n - \boldsymbol{\eta})(s) \right\rangle \right| \\
&\leq \|\boldsymbol{x}_n - \boldsymbol{x}\|_{C([0,T];\mathbb{R}^d)} \mathrm{V}_0^T(\boldsymbol{\eta}_n) + \left\|\boldsymbol{x}^N - \boldsymbol{x}\right\|_{C([0,T];\mathbb{R}^d)} \left(\mathrm{V}_0^T(\boldsymbol{\eta}_n) + \mathrm{V}_0^T(\boldsymbol{\eta}) \right) \\
&\quad + \sum\nolimits_{i=1}^{k_N} |\boldsymbol{x}(t_i)| \cdot |(\boldsymbol{\eta}_n - \boldsymbol{\eta})(t_i \wedge t) - (\boldsymbol{\eta}_n - \boldsymbol{\eta})(t_{i-1} \wedge t)|\,.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\sup_{t\in[0,T]} \left| \int_0^t \langle \boldsymbol{x}_n(s), d\boldsymbol{\eta}_n(s)\rangle - \int_0^t \langle \boldsymbol{x}(s), d\boldsymbol{\eta}(s)\rangle \right| \leq M \|\boldsymbol{x}_n - \boldsymbol{x}\|_{C([0,T];\mathbb{R}^d)} \\
&\quad + 2M \left\|\boldsymbol{x}^N - \boldsymbol{x}\right\|_{C([0,T];\mathbb{R}^d)} + 2 \left(\sum\nolimits_{i=1}^{k_N} |\boldsymbol{x}(t_i)| \right) \|\boldsymbol{\eta}^n - \boldsymbol{\eta}\|_{C([0,T];\mathbb{R}^d)}\,.
\end{aligned}$$

It follows that

$$\limsup_{n\to\infty} \sup_{t\in[0,T]} \left| \int_0^t \langle \boldsymbol{x}_n(s), d\boldsymbol{\eta}_n(s)\rangle - \int_0^t \langle \boldsymbol{x}(s), d\boldsymbol{\eta}(s)\rangle \right| \leq 2M \left\| \boldsymbol{x}^N - \boldsymbol{x} \right\|_{C([0,T];\mathbb{R}^d)}.$$

Since $\lim_{N\to\infty} \left\| \boldsymbol{x}^N - \boldsymbol{x} \right\|_{C([0,T];\mathbb{R}^d)} = 0$, we finally get

$$\lim_{n\to\infty} \sup_{t\in[0,T]} \left| \int_0^t \langle \boldsymbol{x}_n(s), d\boldsymbol{\eta}_n(s)\rangle - \int_0^t \langle \boldsymbol{x}(s), d\boldsymbol{\eta}(s)\rangle \right| = 0.$$

□

Let us now proceed with the proof of the main result of this section, which follows the same steps as that of Proposition 3.4 from [222].

Proof of the proposition 2.6. Let $\mathbf{W} := C([0,T];\mathbb{R}^d)$, $\mathbf{V} := C([0,T];\mathbb{R}^d) \cap BV([0,T];\mathbb{R}^d)$ and, for $\nu > 0$,

$$\mathbf{V}_\nu := \left\{ \boldsymbol{\eta} \in \mathbf{V} \mid \|\boldsymbol{\eta}\|_{BV([0,T];\mathbb{R}^d)} \leq \nu \right\}.$$

By the first part of Lemma 2.7, $\mathbf{V}_\nu$ is a closed subset of the Banach space $\mathbf{W}$.

Let us consider the function $\Lambda : \mathbf{W} \times \mathbf{W} \to \mathbf{W}$ defined by

$$\Lambda(\boldsymbol{x}, \boldsymbol{\eta})(t) := \begin{cases} \int_0^t \langle \boldsymbol{x}(s), d\boldsymbol{\eta}(s)\rangle, & (x, \boldsymbol{\eta}) \in \mathbf{W} \times \mathbf{V}; \\ 0, & (x, \boldsymbol{\eta}) \in \mathbf{W} \times (\mathbf{W} \setminus \mathbf{V}). \end{cases}$$

By the last conclusion of Lemma 2.7, the restriction $\Lambda|_{\mathbf{W}\times\mathbf{V}_\nu}$ is continuous.

Let now $R_n := \mathbb{P}^n \circ (X^n, H^n)^{-1}$ and $R_0 := \mathbb{P} \circ (X, H)^{-1}$, the distribution probabilities of (X^n, H^n), respectively (X, H). By the assumptions of the theorem, $(R_n)_{n\geq 1}$ converges weakly to R_0, *i.e.*

$$\lim_{n\to\infty} \int_{\mathbf{W}\times\mathbf{W}} \Phi(\boldsymbol{x}, \boldsymbol{\eta})\, R_n(d\boldsymbol{x}, d\boldsymbol{\eta}) = \int_{\mathbf{W}\times\mathbf{W}} \Phi(\boldsymbol{x}, \boldsymbol{\eta})\, R_0(d\boldsymbol{x}, d\boldsymbol{\eta}), \tag{2.11}$$

for every bounded continuous functional $\Phi : \mathbf{W} \times \mathbf{W} \to \mathbb{R}$.

First of all, by Portmanteau lemma,

$$\limsup_{n\to\infty} R_n(\mathbf{W} \times \mathbf{V}_\nu) \leq R_0(\mathbf{W} \times \mathbf{V}_\nu),\ \forall \nu > 0.$$

Since, by condition (2.10),

$$\lim_{\nu\to+\infty} \inf_{n\geq 1} R_n(\mathbf{W} \times \mathbf{V}_\nu) = 1, \tag{2.12}$$

we get $\lim_{\nu\to+\infty} R_0(\mathbf{W} \times \mathbf{V}_\nu) = 1$, *i.e.* $R_0(\mathbf{W} \times \mathbf{V}) = 1$, meaning that H is $\mathbb{P}$-a.s. of bounded variation.

Let now $\phi : C[0,T] \to \mathbb{R}$ be an arbitrary bounded continuous functional. It remains to prove that $\lim_{n\to\infty} \mathbb{E}\phi(\Lambda(X^n, H^n)) = \mathbb{E}\phi(\Lambda(X, H))$, which can be written as

$$\lim_{n\to\infty} \int_{\mathbf{W}\times\mathbf{W}} (\phi\circ\Lambda) dR_n = \int_{\mathbf{W}\times\mathbf{W}} (\phi\circ\Lambda) dR_0.$$

Since $\phi\circ\Lambda|_{\mathbf{W}\times\mathbf{V}_\nu}$ is bounded and continous, it can be extended to a continuous functional $\Phi_\nu : \mathbf{W}\times\mathbf{W} \to \mathbb{R}$, bounded by $M := \sup_{\mathbf{z}\in C[0,T]} \phi(\mathbf{z})$; hence, by (2.11),

$$\lim_{n\to\infty} \int_{\mathbf{W}\times\mathbf{W}} \Phi_\nu(\boldsymbol{x},\boldsymbol{\eta})\, R_n(d\boldsymbol{x}, d\boldsymbol{\eta}) = \int_{\mathbf{W}\times\mathbf{W}} \Phi_\nu(\boldsymbol{x},\boldsymbol{\eta})\, R_0(d\boldsymbol{x}, d\boldsymbol{\eta}).$$

Let us estimate the term

$$T_{n,\nu} := \left| \int_{\mathbf{W}\times\mathbf{W}} (\Phi_\nu(\boldsymbol{x},\boldsymbol{\eta}) - \phi\circ\Lambda)\, R_n(d\boldsymbol{x}, d\boldsymbol{\eta}) \right|,$$

for $n \in \mathbb{N}$ (including then the case $n = 0$). We have

$$\begin{aligned} T_{n,\nu} &\le \int_{\mathbf{W}\times\mathbf{W}} |\Phi_\nu(\boldsymbol{x},\boldsymbol{\eta}) - \phi\circ\Lambda|\, R_n(d\boldsymbol{x}, d\boldsymbol{\eta}) = \int_{\mathbf{W}\times(\mathbf{W}\setminus\mathbf{V}_\nu)} |\Phi_\nu(\boldsymbol{x},\boldsymbol{\eta}) - \phi\circ\Lambda|\, R_n(d\boldsymbol{x}, d\boldsymbol{\eta}) \\ &\le 2M R_n(\mathbf{W}\times(\mathbf{W}\setminus\mathbf{V}_\nu)) = 2M(1 - R_n(\mathbf{W}\times\mathbf{V}_\nu)). \end{aligned}$$

Hence, by (2.12) and its consequence

$$\lim_{\nu\to+\infty} \sup_{n\ge 0} T_{n,\nu} = 0.$$

Finally, for all $n \ge 1$ and $\nu > 0$,

$$\begin{aligned} &\left| \int_{\mathbf{W}\times\mathbf{W}} (\phi\circ\Lambda) dR_n - \int_{\mathbf{W}\times\mathbf{W}} (\phi\circ\Lambda) dR_0 \right| \\ &\qquad \le \left| \int_{\mathbf{W}\times\mathbf{W}} \Phi_\nu(\boldsymbol{x},\boldsymbol{\eta})\, R_n(d\boldsymbol{x}, d\boldsymbol{\eta}) - \int_{\mathbf{W}\times\mathbf{W}} \Phi_\nu(\boldsymbol{x},\boldsymbol{\eta})\, R_0(d\boldsymbol{x}, d\boldsymbol{\eta}) \right| + T_{n,\nu} + T_{0,\nu} \end{aligned}$$

and therefore

$$\limsup_{n\to\infty} \left| \int_{\mathbf{W}\times\mathbf{W}} (\phi\circ\Lambda) dR_n - \int_{\mathbf{W}\times\mathbf{W}} (\phi\circ\Lambda) dR_0 \right| \le 2 \sup_{n\ge 0} T_{n,\nu},\ \forall \nu > 0$$

which, by passing to the limit as $\nu \to 0$, yields the desired conclusion. □

We are now in the position for presenting the main result of this section, namely the convergence in law of the solution of Eq. (2.8) to the BSDE (2.3):

Theorem 2.8. *Assume that the above assumptions are fulfilled. Then*

$$\lim_{n\to\infty}\mathbb{E}\left[\sup_{t\in[0,T]}|Y^n(t)-Y(t)|^2+\int_0^T|Z^n(t)-Z(t)|^2\,dt\right]=0.$$

Proof. Let us denote for short

$$\begin{aligned}\Delta_nY(t):=Y^n(t)-Y(t)\,,\quad \Delta_nZ(t):=Z^n(t)-Z(t)\,;\quad \Delta_n\xi:=\xi^n(t)-\xi(t)\\ \omega_\delta^n:=\sup_{t\in[0,T-\delta]}\left(A^n(t+\delta)-A^n(t)\right).\end{aligned}$$

Exactly as in the proof of Theorem 2.3, by (H_1) and (H_2), we have

$$\begin{aligned}\Big(\tfrac{K_1T}{4L^2}+\tfrac{4\tilde{K}_1\,A(T)}{\beta}\Big)e^{\alpha\delta+\beta\omega_\delta}\le 2c\,,\quad \mathbb{P}\text{-a.s.};\\ \tfrac{K_1}{4L^2}e^{\alpha\delta+\beta\omega_\delta}\le c\,,\quad \mathbb{P}\text{-a.s},\end{aligned}$$

for all $n\in\mathbb{N}$, where $\alpha=8L^2+\frac{1}{2}$. Let us apply Itô's formula to $e^{\alpha t+\beta A(t)}\,|Y^n(t)-Y(t)|^2$:

$$\begin{aligned}&e^{\alpha t+\beta A^n(t)}|\Delta_nY(t)|^2+\int_t^T e^{\alpha s+\beta A^n(s)}|\Delta_nY(s)|^2\left(\alpha\,ds+\beta\,dA^n(s)\right)+\int_t^T e^{\alpha s+\beta A^n(s)}|\Delta_nZ(s)|^2ds\\ &=e^{\alpha T+\beta A^n(T)}|\Delta_n\xi|^2-2\int_t^T e^{\alpha s+\beta A^n(s)}\langle\Delta_nY(s),\Delta_nZ(s)dW(s)\rangle\\ &\quad+2\int_t^T e^{\alpha s+\beta A^n(s)}\langle\Delta_nY(s),F^n(s,Y^n(s),Z^n(s),Y^n_s,Z^n_s)-F(s,Y(s),Z(s),Y_s,Z_s)\rangle ds\\ &\quad+2\int_t^T e^{\alpha s+\beta A^n(s)}\langle\Delta_nY\left(s\right),G^n(s,Y^n(s),Y^n_s)dA^n(s)-G(s,Y(s),Y_s)dA(s)\rangle\,.\end{aligned}$$

From assumptions (B_2)–(B_3) and (B_1'), we have, with $K_1^n:=\sup_{t\in[0,T]}K^n$ and $\tilde{K}_1^n:=$

$\sup_{t\in[0,T]} \tilde{K}^n$,

$$2\int_t^T e^{\alpha s+\beta A^n(s)}\langle \Delta_n Y(s), F^n(s,Y^n(s),Z^n(s),Y^n_s,Z^n_s) - F(s,Y(s),Z(s),Y_s,Z_s)\rangle ds$$
$$\le 8L^2\int_t^T e^{\alpha s+\beta A^n(s)}|\Delta_n Y(s)|^2 ds + \frac{1}{2}|\Delta_n F|^2\int_t^T e^{\alpha s+\beta A^n(s)}ds$$
$$+\frac{1}{2}\int_t^T e^{\alpha s+\beta A^n(s)}\left(|\Delta_n Y(s)|^2+|\Delta_n Z(s)|^2\right)dr$$
$$+\frac{K_1^n T e^{\alpha\delta+\beta\omega_\delta^n}}{4L^2}\sup_{s\in[0,T]}\left(e^{\alpha s+\beta A^n(s)}|\Delta_n Y(s)|^2\right)$$
$$+\frac{K_1^n e^{\alpha\delta+\beta\omega_\delta^n}}{4L^2}\int_0^T e^{\alpha r+\beta A^n(r)}|\Delta_n Z(r)|^2 dr$$

and, for all $b>0$,

$$2\int_t^T e^{\alpha s+\beta A^n(s)}\langle \Delta_n Y(s), G^n(s,Y^n(s),Y^n_s)dA^n(s) - G(s,Y(s),Y_s)dA(s)\rangle$$
$$= 2\int_t^T e^{\alpha s+\beta A^n(s)}\langle \Delta_n Y(s), G^n(s,Y^n(s),Y^n_s) - G^n(s,Y(s),Y_s)\rangle dA^n(s)$$
$$+2\int_t^T e^{\alpha s+\beta A^n(s)}\langle \Delta_n Y(s), G^n(s,Y^n(s),Y^n_s) - G(s,Y(s),Y_s)\rangle dA^n(s)$$
$$+2\int_t^T e^{\alpha s+\beta A^n(s)}\langle \Delta_n Y(s), G(s,Y(s),Y_s)\rangle\left(dA^n(s)-dA(s)\right)$$
$$\le 2\int_t^T e^{\alpha s+\beta A^n(s)}\langle \Delta_n Y(s), G(s,Y(s),Y_s)\rangle\left(dA^n(s)-dA(s)\right)$$
$$+b\int_t^T e^{\alpha s+\beta A^n(s)}|\Delta_n Y(s)|^2 dA^n(s) + \frac{4\tilde{L}^2}{b}|\Delta_n G|^2\int_t^T e^{\alpha s+\beta A^n(s)}dA^n(s)$$
$$+\frac{\beta}{2}\int_t^T e^{\alpha s+\beta A^n(s)}|\Delta_n Y(s)|^2 dA^n(s) + \frac{4\tilde{L}^2}{\beta}\int_t^T e^{\alpha s+\beta A^n(s)}|\Delta_n Y(s)|^2 dA^n(s)$$
$$+\frac{4\tilde{K}_1^n A^n(T)e^{\alpha\delta+\beta\omega_\delta^n}}{\beta}\sup_{s\in[0,T]}\left(e^{\alpha s+\beta A^n(s)}|\Delta_n Y(s)|^2\right).$$

Since $\alpha = 4L^2 + 1$ and $\beta > 2\sqrt{2}\tilde{L}$, one can choose $b := \frac{\beta}{2} - \frac{4\tilde{L}^2}{\beta}$ and so we obtain

$$\begin{aligned}
&e^{\alpha t+\beta A^n(t)}|\Delta_n Y(t)|^2 + \frac{1}{2}\int_t^T e^{\alpha s+\beta A^n(s)}|\Delta_n Z(s)|^2 ds\\
&\le e^{\alpha T+\beta A^n(T)}|\Delta_n \xi|^2 - 2\int_t^T e^{\alpha s+\beta A^n(s)}\langle \Delta_n Y(s), \Delta_n Z(s) dW(s)\rangle\\
&\quad +2\int_t^T e^{\alpha s+\beta A^n(s)}\langle \Delta_n Y(s), G(s, Y(s), Y_s)\rangle\,(dA^n(s) - dA(s))\\
&\quad +\frac{1}{2}|\Delta_n F|^2 \int_t^T e^{\alpha s+\beta A^n(s)} ds + \frac{4\tilde{L}^2}{b}|\Delta_n G|^2 \int_t^T e^{\alpha s+\beta A^n(s)} dA^n(s)\\
&\quad +\frac{K_1 T e^{\alpha\delta+\beta\omega_\delta^n}}{4L^2}\sup_{s\in[0,T]}\left(e^{\alpha s+\beta A^n(s)}|\Delta_n Y(s)|^2\right)\\
&\quad +\frac{K_1 e^{\alpha\delta+\beta\omega_\delta^n}}{4L^2}\int_0^T e^{\alpha s+\beta A^n(s)}|\Delta_n Z(s)|^2 ds\\
&\quad +\frac{4\tilde{K}_1 A^n(T) e^{\alpha\delta+\beta\omega_\delta^n}}{\beta}\sup_{s\in[0,T]}\left(e^{\alpha s+\beta A^n(s)}|\Delta_n Y(s)|^2\right).
\end{aligned}$$

Therefore, by conditions (H_1) and (H_2),

$$\begin{aligned}
&\frac{1}{2}\int_t^T e^{\alpha s+\beta A^n(s)}|\Delta_n Z(s)|^2 ds\\
&\le 2\int_t^T e^{\alpha s+\beta A^n(s)}\langle \Delta_n Y(s), G(s, Y(s), Y_s)\rangle\,(dA^n(s) - dA(s))\\
&\quad +\frac{1}{2}|\Delta_n F|^2 \int_t^T e^{\alpha s+\beta A^n(s)} ds + \frac{4\tilde{L}^2}{b}|\Delta_n G|^2 \int_t^T e^{\alpha s+\beta A^n(s)} dA^n(s)\\
&\quad +2c\sup_{s\in[0,T]}\left(e^{\alpha s+\beta A^n(s)}|\Delta_n Y(s)|^2\right) + c\int_0^T e^{\alpha s+\beta A^n(s)}|\Delta_n Z(s)|^2 ds.
\end{aligned}$$

Exploiting Burkholder–Davis–Gundy's inequality, we have that

$$\begin{aligned}
&2\mathbb{E}\Big[\sup_{t\in[0,T]}\Big|\int_t^T e^{\alpha s+\beta A^n(s)}\langle \Delta_n Y(s), \Delta_n Z(s) dW(s)\rangle\Big|\Big]\\
&\le \frac{1}{4}\mathbb{E}\big(e^{\alpha s+\beta A^n(s)}|\Delta_n Y(s)|^2\big) + 144\,\mathbb{E}\int_0^T e^{\alpha s+\beta A^n(s)}|\Delta_n Z(s)|^2 ds\,.
\end{aligned}$$

As in the proof of Theorem 2.3, we obtain

$$\begin{aligned}
&\mathbb{E}\big(\sup\nolimits_{s\in[t,T]} e^{\alpha s+\beta A^n(s)}|\Delta_n Y(s)|^2\big)+\mathbb{E}\int_0^T e^{\alpha s+\beta A^n(s)}|\Delta_n Z(s)|^2 ds\\
&\le C\mathbb{E}\left[|\Delta_n\xi|^{2p}+|\Delta_n F|^{2p}+|\Delta_n G|^{2p}\right]\cdot\mathbb{E}e^{\beta q A^n(T)}\\
&\quad +C\mathbb{E}\sup_{t\in[0,T]}\left|\int_t^T e^{\alpha s+\beta A^n(s)}\langle\Delta_n Y(s),G(s,Y(s),Y_s)\rangle\,(dA^n(s)-dA(s))\right|.
\end{aligned}$$

where C is a positive constant and $q := \frac{p}{p-1}$.

By conditions (C_1) and (B_0''),

$$\lim_{n\to\infty}\mathbb{E}\left[|\Delta_n\xi|^{2p}+|\Delta_n F|^{2p}+|\Delta_n G|^{2p}\right]\cdot\mathbb{E}e^{\beta q A^n(T)}=0.$$

It remains to prove that

$$\lim_{n\to\infty}\mathbb{E}\sup_{t\in[0,T]}\left|\int_t^T X^n(s)dH^n(s)\right|=0,$$

where, for $s\in[0,T]$,

$$\begin{aligned}
X^n(s) &:= e^{\alpha s+\beta A^n(s)}\langle\Delta_n Y(s),G(s,Y(s),Y_s)\rangle;\\
H^n(s) &:= A^n(s)-A(s).
\end{aligned}$$

One can prove that

$$\mathbb{E}\sup_{t\in[0,T]}|X^n(t)|^p$$

is uniformly bounded (with respect to n), by (2.9). Obviously, by (B_0''),

$$\sup_{n\in\mathbb{N}^*}\mathbb{E}\sup_{t\in[0,T]}|H^n(t)|^2<+\infty.$$

Hence, the sequence $(X^n,H^n)_{n\in\mathbb{N}^*}$ is tight in $C[0,T]^2$. By Prokhorov's theorem, we can extract a sequence, say $(X^{n_k},H^{n_k})_{k\in\mathbb{N}^*}$, convergent in distribution to some stochastic process (X,H) with continuous paths. Since, by (C_2), $\lim_{n\to\infty}\mathbb{E}\sup_{t\in[0,T]}|H^n(t)|=0$, H must be $\mathbb{P}$-a.s. equal to 0. The condition (B_0'') also implies that $\sup_{n\in\mathbb{N}}\mathbb{E}\|H^n\|^a_{BV[0,T]}<+\infty$, for every $a>1$, so $\|H^n\|_{BV[0,T]}$ is bounded in probability (*i.e.*, it satisfies condition (2.10)). We can now apply Proposition 2.6, proved as an auxiliary result in the Appendix section, in order to derive the convergence in distribution to 0 of the process

$$\left(\int_0^t X^n(s)dH^n(s)\right)_{t\in[0,T]}.$$

Since, for some $\nu > 0$, the functional $\phi_\nu : C[0,T] \to \mathbb{R}$, defined by

$$\phi_\nu(\boldsymbol{x}) := \sup_{t\in[0,T]} |\boldsymbol{x}(T) - \boldsymbol{x}(t)| \wedge \nu,$$

is bounded and continuous, it follows that

$$\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_t^T X^n(s)dH^n(s)\right| \wedge \nu\right] = 0,$$

for every $\nu > 0$. Since, by Markov's inequality, for some $a \in (1,p)$

$$\begin{aligned}\mathbb{E}\sup_{t\in[0,T]}\left|\int_t^T X^n(s)dH^n(s)\right| \le \mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_t^T X^n(s)dH^n(s)\right| \wedge \nu\right]&\\ + \frac{1}{\nu^a}\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_t^T X^n(s)dH^n(s)\right|^a\right]&\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_t^T X^n(s)dH^n(s)\right|^a\right] &\le \mathbb{E}\left[\left(\sup_{t\in[0,T]}|X^n(t)|^a\right)\|H^n\|^a_{BV[0,T]}\right]\\ &\le \left(\mathbb{E}\left[\sup_{t\in[0,T]}|X^n(t)|^p\right]\right)^{\frac{a}{p}}\left(\mathbb{E}\,\|H^n\|^{\frac{p}{a(p-a)}}_{BV[0,T]}\right)^{1-\frac{a}{p}},\end{aligned}$$

it follows that

$$\lim_{n\to\infty}\mathbb{E}\sup_{t\in[0,T]}\left|\int_t^T X^n(s)dH^n(s)\right| = 0,$$

which concludes our proof. □

2.5 Hedging a stream of payments with time-delayed GBSDE

In this last section we present a risk management application for an insurance product, the so called variable annuity instrument, whose composition can be controlled by the insurer selecting an appropriate strategy. The composition for the underlying investment portfolio can be controlled internally by the insurer to reduce the overall risk of the policyholder investment. Specifically, inspired by [71], we consider an insurance product where the policyholder withdraws some guaranteed amounts as a fraction of the maximum value of the investment and, additionally, is subjected to a continuous payment triggered by an increasing continuous process A modelling the cumulative function of claims (or, e.g. of fees for the management of the wealth). At maturity the remaining value is converted into a life-time annuity with a guaranteed consumption rate C.

We consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with associated natural filtration $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq R}$ generated by a Brownian motion $W := (W(t), 0 \leq t \leq T)$ and a finite time horizon $T \leq \infty$.

The goal of the investor is to replicate the insurance by investing into the assets and to quantify the risk of the investing activities. In the terminology of [136], we focus on an investment composed by a risk free asset S_0 and a risky asset D.

The price of the risk free bond $S_0 := (S_0(t), 0 \leq t \leq T)$ is given by the equation

$$\frac{dS_0(t)}{S_0(t)} = r(t)dt, \qquad S_0(0) = 1, \tag{2.13}$$

where r describes the risk free interest rate being a non-negative $\mathbb{F}$-progressively measurable stochastic process.

The price of the risky asset $D := (D(t), 0 \leq t \leq T)$ with maturity T is given by

$$\frac{dD(t)}{D(t)} = (r(t) + \sigma(t)\theta(t))\,dt + \sigma(t)dW(t), \qquad S(0) = x, \tag{2.14}$$

where the volatility $\sigma := (\sigma(t), 0 \leq t \leq T)$ and the risk premium $\theta := (\theta(t), 0 \leq t \leq T)$ are $\mathbb{F}$-progressively measurable processes.

On the other hand, the stream of liabilities $P(t) := (P(t), 0 \leq t \leq T)$ depends on the past value of the portfolio by the following:

$$P(t) = \gamma \sup_{s \in [0,t]} \{X(s)\}\,dt + \int_0^t X(s-\delta)dA(s). \tag{2.15}$$

The first term models a guaranteed withdrawal amount as a fraction $\gamma \in (0,1)$ of the running maximum value of the investment value. Instead, the second term models a Stieltjes integral representing the total amount of continuous claims that depends on a past value of the investment and that are triggered by the increasing continuous function A. We emphasize that if we consider no dependence on the value of the investment X, i.e. only $\int_0^t dA(s)$, we obtain the well-known case with A representing a cumulative consumption process. See, e.g., [136] for a detailed description or [97] for the problem of utility maximization under a drawdown constraint setting.

We consider a self financing investment portfolio $X := (X(t), 0 \leq t \leq T)$, while the admissible strategy $\pi := (\pi(t), 0 \leq t \leq T)$ denotes the amount invested in the risky asset D.

We denote $\mu(t) = r(t) + \theta(t)\sigma(t)$ and we write the dynamic of X by the following SDE

$$\begin{aligned}dX(t) =&\pi(t)\frac{dD(t)}{D(t)} + (X(t)-\pi(t))\frac{dS_0(t)}{S_0(t)} - dP(t) \\ =&\pi(t)\left(\mu(t)dt + \sigma(t)dW(t)\right) + (X(t)-\pi(t))\,r(t)dt \\ &-\gamma \sup_{s\in[0,t]}\{X(s)\}\,dt - \int_0^t X(s-\delta)dA(s) \\ X(T) =&Ca(T)\,,\end{aligned} \tag{2.16}$$

a being the annuity factor $a(T) = \mathbb{E}^{\mathbb{Q}}\left[\int_T^\infty e^{-\int_T^s r(u)du} ds | \mathcal{F}_T\right]$.

Equation (2.16) models a variable annuity contract where the policyholder's contributions are invested into two assets (D and S_0). Positive returns are distributed to policyholder account based on on the maximum value of the investment and on a prescribed process A (hedging fee) while the remaining value at maturity is received as a life-time annuity.

From [71], we know that there exists a unique equivalent martingale measure $\mathbb{Q} \sim \mathbb{P}$ under which the discounted price process S is a $(\mathbb{Q},\mathbb{F})$-martingale. Thus, we perform the following change of variables

$$Y(t) = X(t)e^{-\int_0^t r(s)ds}\,, \qquad Z(t) = \pi(t)\sigma(t)e^{-\int_0^t r(s)ds}\,, \qquad 0 \le t \le T \tag{2.17}$$

giving the following dynamic for the discounted portfolio process $Y := (Y(t))_{0\le t\le T}$ under the measure $\mathbb{Q}$

$$\begin{aligned}Y(t) = C\tilde{a}(T) + \int_t^T \gamma \sup_{u\in[0,s]}\left\{Y(u)e^{-\int_u^s r(v)dv}\right\} ds + \\ + \int_t^T Y(s-\delta)e^{-\int_0^{s-\delta} r(v)dv} dA(s) - \int_t^T Z(s)dW^{\mathbb{Q}}(s)\,,\end{aligned} \tag{2.18}$$

$W^{\mathbb{Q}}$ being a $\mathbb{Q}$-Brownian motion.

By assuming that conditions (B_0) - (B_3) and (H_1) - (H_2) hold true and by applying Theorem 2.3, we obtain existence and uniqueness of the solution of equation (2.18). Moreover, the stability of the investment under a perturbation (in uniform norm) of the distribution of the prescribed cumulative distribution is obtained by Theorem 2.8, letting to model robust hedging for the investment with respect to a modification of the prescribed cumulative distribution of future claims.

2.6 Conclusions and further developments

In this chapter, we develop a theoretical framework to study a BSDE with a time-delayed generator whose dynamic depends also on the Stieltjes integral term. Under regular assumptions of the coefficients and small delay, we prove the well-posedness of the problem in terms of existence, uniqueness and stability under a perturbation in uniform norm. We also provide an application of our results for a BSDE in insurance setting. Moreover, we obtain the global (in time) well–posedness of the BSDE for an arbitrary delay that represents a novel result in the literature, representing a first attempt to handle (globally) time delayed BSDE. Providing a solid theoretical background for this setting could open up new directions for applications.

Concerning further direction of research, other extensions would consider the forward reflected SDE linked to the Stieltjes integral in (2.1) to investigate the corresponding FBSDE with delayed generator and possible connections with the nonlinear PDE with Neumann boundary conditions in the spirit of [181]. Another possibility concerns considering Stieltjes integration with respect to increasing functions that are not necessarily continuous, dealing with dynamics driven by Poisson random measure.

Part II

Mean Field Games theory

3 The Master Equation with Absorption

3.1 Introduction

In this chapter we study the convergence problem in a bounded domain $\Omega \subset \mathbb{R}^d$, via *Master Equation*, whose precise definition is given in Eq. (3.7), requiring an *absorbing* condition at the boundary $\partial\Omega$, namely we ask for each player to remain in the game as long as his dynamic lies in the interior $int(\Omega)$ of the domain Ω.

In this framework, the dynamic of the player i, $1 \le i \le N$, is described by the following Stochastic Differential Equation (SDE), with initial time $t_0 > 0$ and final time $T > t_0$:

$$\begin{cases} dX_t^{i,t_0} = b(t, X_t^{i,t_0}, \alpha_t^i)dt + \sqrt{2}\sigma(X_t^{i,t_0})dB_t^i \\ X_{t_0}^{i,t_0} = x_0^i. \end{cases} \tag{3.1}$$

α_t^i being the control of the player i, i.e. a progressively measurable process taking values in a certain set A, while $b : [0,T] \times \Omega \times A \to \mathbb{R}^d$ is the *drift* function and $\sigma : \Omega \to \mathbb{R}^{d\times d}$ represents the *diffusion matrix*. Henceforth, for simplicity, we will shortly write X_t^i.

The process $X_\cdot^i$ is Markovian providing standard hypotheses on the drift and the coefficient terms. Therefore, we are allowed to define the hitting time τ at $\partial\Omega$ in $[0,T]$ as

$$\tau(X_\cdot^i) = \inf\Big\{ t \in [t_0, T] \,|\, X_t^i \in \partial\Omega \Big\} \wedge T\,, \tag{3.2}$$

where $a \wedge b := \min\{a, b\}$ and we abbreviate $\tau(X_t^i)$ by τ^i, whenever no confusion is possible.

From now on, we indicate a vector of $\mathbb{R}^{Nd}$ with $\boldsymbol{v} := (v_1, \dots, v_N)$, with $v_i \in \mathbb{R}^d$, $i \in \{1, ..., N\}$. A *cost functional* is associated with each player i, which chooses his own strategy in order to minimize it. To obtain a convergence result, we have to require the cost functional to have a symmetric structure. Namely, we assume that the cost for the player i depends on both its position x^i and on the other players' empirical distribution, defined as follows:

$$m_{\mathbf{x}}^{N,i} = \frac{1}{N-1} \sum_{j \ne i} \delta_{x_j} \mathbb{1}_{\{x_j \in int(\Omega)\}}\,, \tag{3.3}$$

where δ_x is the Dirac mass at x. The indicator function specifies that we are interested only in those players in the interior of Ω. Let us underline that the presence of such an indicator function slightly modifies the classic theoretical setting of the problem: for the rest of the chapter, $m_{\mathbf{x}}^{N,i}$ belongs to $\mathcal{P}^{sub}$, the space of subprobability measure, as in-depth analyzed in Section 3.2.2.

Consequently, the cost functional is defined in the following way

$$J_i^N(t_0,\mathbf{x}_0,\boldsymbol{\alpha}) = \mathbb{E}\left[\int_{t_0}^{\tau^i}\left(L\left(s,X_s^i,\alpha_s^i\right)+F\left(s,X_s^i,m_{\mathbf{X}_s}^{N,i}\right)\right)ds + G\left(X_{\tau^i}^i,m_{\mathbf{x}}^{N,i}\right)\right], \tag{3.4}$$

where L, F and G are, respectively, the Lagrangian term, the running cost and the final cost to pay at the hitting time, or at the final time when no boundary hitting happens.

Each player wishes to minimize its cost functional, which also depends on the strategies of the other players. The set of strategies $\boldsymbol{\alpha^*}$ which allows the agents to play their optimal strategy, in relation to the other agents, is called *Nash equilibrium*. Namely,

$$J_i^N(t_0,x_0,\boldsymbol{\alpha^*}) \le J_i^N(t_0,x_0,\alpha_i,(\alpha_j^*)_{j\neq i}).$$

We define $v_i^N(t,\boldsymbol{x}) := J_i^N(t,\boldsymbol{x},\boldsymbol{\alpha}^*)$. Exploiting both the Itô-Doeblin formula and the dynamic programming principle, we have that $\boldsymbol{\alpha^*}$ is a Nash equilibrium if v_i^N solves a coupled system of parabolic equations, traditionally called *Nash system*. Namely, for $(t,\boldsymbol{x})\in(0,T)\times\Omega^N$ and for $i=1,...,N$, we have

$$\begin{cases} -\partial_t v_i^N - \displaystyle\sum_{j=1}^N \operatorname{tr}\big(a(x_j)D^2_{x_jx_j}v_i^N\big) + H(t,x_i,D_{x_i}v_i^N) \\ \qquad\qquad + \displaystyle\sum_{j\neq i}^N H_p(t,x_j,D_{x_j}v_j^N)\cdot D_{x_j}v_i^N = F(t,x_i,m_{\mathbf{x}}^{N,i}), \\ v_i^N(T,\mathbf{x}) = G(x_i,m_{\mathbf{x}}^{N,i}), \\ v_i^N(t,\mathbf{x})_{|x_i\in\partial\Omega} = 0, \forall t\in[0,T] \end{cases} \tag{3.5}$$

$H(t,x,p) = \sup\limits_{\alpha\in A}(-b(t,x,\alpha)\cdot p - L(t,x,\alpha))$ being the Hamiltonian of the system, and with H_p its derivative in p and $a=\sigma\sigma^*$.

Knowing the initial distribution, independently from the individual starting position on $\partial\Omega$, implies that these players suddenly stop their dynamic at the initial time, hence making no contribution to the evolution of other players. Therefore, for all i,j, the following equation holds:

$$v_i^N(t,x_1,\dots,x_{j-1},x,x_{j+1},\dots,x_N) = v_i^N(t,x_1,\dots,x_{j-1},y,x_{j+1},\dots,x_N), \tag{3.6}$$

for all $x, y \in \partial\Omega$. Hence, the value function v_i^N does not depend on x_j when $x_j \in \partial\Omega$. Roughly speaking, the degrees of freedom of the solution v_i^N decreases from N to $N -$ (*number of players on* $\partial\Omega$). This fact has a direct correspondence also in the dynamic of the game in terms of a leak of mass in the Fokker-Planck equation, as we will see in (3.10) and in the $\mathcal{P}^{sub}$ formulation of the Master Equation w.r.t. to a *new* boundary condition as proved in Corollary 3.13.

It is possible to provide a suitable approximation for the asymptotic configuration when $N \to +\infty$ by introducing the MFG system. Roughly speaking, a generic player behaves according to the following dynamic:

$$\begin{cases} dX_t = b(X_t, \alpha_t)\, dt + \sqrt{2}\sigma(X_t)\, dB_t\,, \\ X_{t_0} = x_0\,, \end{cases}$$

and the cost functional is given by $\big($here $\tau = \tau(X.)\big)$

$$J(t_0, x_0, \alpha.) = \mathbb{E}\left[\int_{t_0}^{\tau} \big(L(s, X_s, \alpha_s) + F(X_s, m(s))\big)\, ds + G(X_\tau, m(\tau))\right].$$

The solution of a MFG represents the analogous to a Nash equilibrium for a non-cooperative game, see [39, 41, 45, 148, 149] for more details. Moreover, the optimal strategies in the MFG system provide approximated ε-Nash equilibria in the Nash system, as proved, e.g., in [44, 130, 139]. However, the lack of compactness properties of the problem creates many difficulties in the "very" convergence proof, i.e. in proving the convergence of the Nash equilibria in the N-players game towards the optimal strategies of the MFG system.

In order to handle this problem, Lasry and Lions in [150] proved that the solutions of the Mean Field Games are just the trajectories of a non-local transport equation, called the *Master Equation*, in the space of measures. In particular, if $(t, x, m) \in (0, T) \times \Omega \times \mathcal{P}^{sub}(\Omega)$, then the Master Equation is defined as follows:

$$
\begin{cases}
-\partial_t U(t,x,m) - \operatorname{tr}\left(a(x) D_x^2 U(t,x,m)\right) + H\left(t,x,D_x U(t,x,m)\right) \\
\quad - \displaystyle\int_\Omega \operatorname{tr}\left(a(y) D_y D_m U(t,x,m,y)\right) dm(y) \\
\quad + \displaystyle\int_\Omega D_m U(t,x,m,y) \cdot H_p(t,y,D_x U(t,y,m)) dm(y) = F(x,m) \\
\qquad\qquad\qquad\qquad \text{in } [0,T] \times \Omega \times \mathcal{P}^{sub}(\Omega)\,, \\
U(T,x,m) = G(x,m) \qquad \text{in } \Omega \times \mathcal{P}^{sub}(\Omega)\,, \\
U(t,x,m) = 0 \qquad \text{for } (t,x,m) \in [0,T] \times \partial\Omega \times \mathcal{P}^{sub}(\Omega)\,, \\
\dfrac{\delta U}{\delta m}(t,x,m,y) = 0 \qquad \text{for } (t,x,m,y) \in [0,T] \times \Omega \times \mathcal{P}^{sub}(\Omega) \times \partial\Omega\,,
\end{cases}
\tag{3.7}
$$

where $D_m U$ and $\frac{\delta U}{\delta m}$ are two derivatives, considered with respect to the measure variable m, whose precise definition will be given later.

Furthermore, when we consider the limiting procedure by taking the number of agents N to $+\infty$, the solution of the Nash system will converge to the solution of the Master equation, as proved by Cardaliguet, Delarue, Lasry and Lions in [41] for the periodic setting ($\Omega = \mathbb{T}^d$).

For the sake of completeness, let us underline that our formulation of (3.7), is typically referred to as *First order Master Equation*, or *Master Equation without common noise*, appearing when the dynamic of the generic player has the form (3.1). Besides, in the *Second order Master Equation*, or *Master Equation with common noise*, the dynamic of the generic player presents an additional Brownian term dW_t, common to all the players. In the latter case, the Master Equation presents other terms depending on $D_{mm}U$, a second order derivative of U with respect to the measure, and the overall system difficulty significantly grows.

After being introduced in [150], the Master Equation has been studied in many papers, almost always in the periodic case $\Omega = \mathbb{T}^d$, or in the whole space $\Omega = \mathbb{R}^d$. In [33], Buckdahn, Li, Peng proved the well-posedness of the first order Master equation without coupling terms, by means of probabilistic arguments, while Chassagneux, Crisan and Delarue, in [51], provided a first exhaustive existence and uniqueness result of solution, still without common noise. Moreover, Gangbo and Swiech, in [115], gave a short time existence for the Master Equation in the presence of common noise.

The most relevant result was certainly given in [41], where Cardaliaguet, Delarue, Lasry and Lions proved the well posedness of the Master Equation, with and without common noise, in a periodic setting. As regards other boundary conditions, in [194], Ricciardi proved existence and uniqueness results for the first order Master Equation with Neumann boundary conditions.

Moreover, Carmona and Delarue in [45] derived convergence results in the whole space, while in [69, 68], Delarue, Lacker and Ramanan used the Master Equation to analyse the large deviation problem as well as the central limit theorem. Concerning the major-minor problem, Cardaliaguet, Cirant and Porretta studied a convergence result in [40]. For finite state problems, we refer to the works of Bayraktar and Cohen in [15] and by Cecchin and Pelino in [49]. Finally, Bertucci, Lasry and Lions in [23] studied the Master Equation for the Finite State Space Planning Problem. More recently, in the setting of MFG with exhaustible resources, i.e. the continuum limit of a dynamic game of exhaustible resources modelling Cournot competition between producers, see e.g. [117, 116] for more details, Graber and Sircar study the Master Equation for a MFG of controls with absorption in [118]. Differently from the classical setting studied in this chapter where the interaction is through the mean of the state variable, in [118] the authors studied a model of Cournot competition between producers whose states depend on the empirical measure over controls. Concerning a theoretical overview of MFG of control, we refer to [43].

Other important papers about the Master Equation and the convergence problem are given by [14, 46, 48, 50, 89, 90, 175, 143, 144, 113, 132, 114, 165].

Concerning the MFG model with absorption, the first-order system has been widely studied, see for example [36], where the authors studied MFG with absorption through the introduction of a renormalized empirical measure, or [104, 103], where a monotone regularized version of the problem is considered. In [35] a model of bank run is considered within the framework of a MFG model with absorption and common noise. In [18], the authors consider a jump-diffusion dynamic with controlled jumps for the players, providing an example of an illiquid inter-bank market model, where the banks can change their reserves only at the jump times of some exogenous Poisson processes with a common constant intensity. Similar problems, such as minimal-time MFGs where agents want to leave a given bounded domain through a part of its boundary in minimal time with application to crowd motion, have been studied in [166]. Similarly, in [21, 22, 31], optimal stopping MFGs were studied for a model where a representative agent chooses both the optimal control and the optimal time to exit the game. A very general result about the Mean Field Games system, which includes the Dirichlet case as well as the periodic one and the Neumann case, was given by Porretta in [190].

Our aim is to both treat MFG with absorption and to study the related Master Equation (without common noise) for a bounded domain, assuming homogeneous Dirichlet boundary conditions. We refer to [194] and [195] for an equivalent result in the Neumann case.

The main contributions of this work can be summarized as follows:

- *Existence of the Master Equation solution for a bounded domain with Dirichlet setting*: the solution corresponds to an equilibrium of a (random) flow of sub-probability measures. Here, the main novelty relies upon the study of the Master Equation (3.7) in $\mathcal{P}^{sub}(\Omega)$ instead of $\mathcal{P}(\Omega)$, the space of Borel probability measures in Ω. This particular measure space is essential to handle the dissipation of mass related to the framework of Dirichlet's boundary conditions. Furthermore, we use specific tools such as the generalized Wasserstein distance introduced in [188] to deal with subprobability measures.

- *A new boundary condition in the Master Equation*: we stress the fact that the last boundary condition $\frac{\delta U}{\delta m} = 0$ in (3.7) is completely new in the literature. It relies on the fact that the Dirichlet boundary condition in the Nash system (3.5) provides a smoothing effect near the boundary while approaching the asymptotic configuration.

The chapter is basically divided into three parts: first, we study the Master Equation (3.7), proving the existence, uniqueness and regularity of solutions; then we use previously obtained results on the Master Equation to prove the convergence problem, in a suitable sense; while in the third part, we provide an economical application, inspired by the models proposed by [43, 127, 145]. More precisely, in Section 3.2, we introduce the basic notations and hypotheses we will need throughout the chapter, and we state the two main results we want to prove. In Section 3.6 we give a formal proof for the well-posedness of the Master Equation, which becomes rigorous provided some strong regularity results of the function U. In Sections 3.7 and 3.9 we state the latter regularity results for U: in particular, Section 3.7 is devoted to the study of the Fokker-Planck equation and the Mean Field Games, whereas in Section 3.9 we analyze a MFG linearized system which provides a suitable regularity for the function U. In Section 3.11 we derive results about the convergence of the solution and we conclude with 3.12, presenting an optimal liquidation toy model.

The function U is defined as in (3.9), and various estimates, such as global bounds and global Lipschitz regularity, are established. Notably, one of the key challenges in demonstrating that U satisfies (3.7) is establishing its $\mathcal{C}^1$ continuity with respect to m. This step necessitates a meticulous analysis of the linearized mean field game system (refer to [41, 194]) to establish strong regularity of U in both the spatial and measure variables.

It is worth highlighting that these regularity estimates demand strong regularity not only in the spatial domain but also in the measure variable.

In the spatial domain, the regularity is derived in [41] through differentiation of the equation with respect to x. However, in the case of Dirichlet boundary conditions, as well as in general for any boundary conditions, such methods are not directly applicable. Instead, we obtain these bounds by employing a distinct set of space-time estimates that require careful handling.

Moreover, it's essential to note that regularity estimates for the Dirichlet parabolic equation necessitate compatibility conditions between initial and boundary data. Unfortunately, these compatibility conditions cannot always be guaranteed within this context. Consequently, we extend the estimates obtained in [41] through an in-depth investigation of the regularity of solutions to the Fokker-Planck equation.

A noteworthy novelty in this work is the meticulous study of the Fokker-Planck equation and the linearized Mean Field Games system. This study is conducted within negative-order Hölder spaces, and the well-posedness is established via approximation using smooth functions. However, it's important to emphasize that smooth functions are not dense in the dual of Hölder spaces, as pointed out in [132]. As a result, we must confine our analysis to appropriate subsets of these dual spaces. Notably, these ideas can be applied also to adjust the Neumann case [194].

3.2 Problem Formulation and Main Results

In the first part of the chapter, we analyze the well-posedness of the Master Equation defined on $[0,T]\times\Omega\times\mathcal{P}^{sub}(\Omega)$ with a homogeneous Dirichlet condition on $\partial\Omega$.

Our aim is to find a solution $U:[0,T]\times\Omega\times\mathcal{P}^{sub}(\Omega)\to\mathbb{R}$ which solves strongly the equation (3.7). Following the classical framework, described in [41], for each initial data $(t_0,m_0)\in[0,T]\times\mathcal{P}^{sub}(\Omega)$, we consider the MFG system in $[t_0,T]\times\Omega$ with a homogeneous Dirichlet conditions:

$$\begin{cases} -u_t-\operatorname{tr}(a(x)D^2u)+H(t,x,Du)=F(t,x,m(t))\,, \\ m_t-\sum\limits_{ij}\partial^2_{ij}(a_{ij}(x)m)-\operatorname{div}(mH_p(t,x,Du))=0\,, \\ m(t_0)=m_0, \qquad\qquad u(x,T)=G(x,m(T))\,, \\ u_{|\partial\Omega}=0\,, \qquad m_{|\partial\Omega}=0\,, \end{cases} \tag{3.8}$$

where, as always, a backward Hamilton-Jacobi-Bellman equation for the value function u of the generic player is coupled with a forward Fokker-Planck equation for the density m of the population. Then we define

$$U(t_0,x,m_0)=u(t_0,x) \tag{3.9}$$

and we want to prove that, under certain assumptions, U is the solution of the Master Equation (3.7).

First and foremost, we point out that assuming a Dirichlet boundary condition implies the following preliminary estimate on m:

$$\|m(t)\|_{L^1} \leq 1 \qquad \forall\, m_0 \in \mathcal{P}^{sub}(\Omega)\,. \tag{3.10}$$

This is a classical result, under certain assumptions on H (which will be included in our hypotheses), and it was proved, for example, in *Proposition 3.10* of [190] for $m_0 \in L^1(\Omega)$; by a density argument, it can be easily extended for $m_0 \in \mathcal{P}^{sub}(\Omega)$.

Hence, m can not be represented as a probability measure as it occurs in the periodic setting [41] or in the Neumann formulation [194]. This is why we have to work with $m \in \mathcal{P}^{sub}(\Omega)$.

Actually, this is not a strange result: following the stochastic interpretation and the N-players game, the agents exit the bounded boundary when they hit Ω and, thus, the corresponding distribution $m_{\boldsymbol{x}}^{N,i}$ should decrease in terms of mass.

3.2.1 Hölder spaces

In what follows we briefly recall basic notions about the Banach functions spaces then used throughout the chapter, see, e.g., [146, 194], for more details.

Let $T > 0$ and $\Omega \subset \mathbb{R}^d$ be the closure of an open, bounded and connected set, with the boundary of class $\mathcal{C}^{2+\alpha}$, for some $\alpha > 0$; we denote with Q_T the set $Q_T := [0,T] \times \Omega$. We call $d(\cdot)$ the oriented distance function from the boundary of Ω, defined as

$$d(x) = \begin{cases} \ \ \operatorname{dist}(x, \partial\Omega) & x \in \Omega\,, \\ -\operatorname{dist}(x, \partial\Omega) & x \notin \Omega\,. \end{cases}$$

Thanks to [70], we have $d(\cdot) \in \mathcal{C}^{2+\alpha}$ in a neighbourhood of the boundary. Since we are only interested in the local character of d near $\partial\Omega$, when we write $d(\cdot)$, we mean a $\mathcal{C}^{2+\alpha}$ function coinciding with d in a neighbourhood of the boundary.

For $n \geq 0$ and $\alpha \in (0,1)$, $\mathcal{C}^{n+\alpha}(\Omega)$ is defined as the space of functions n-times differentiable, with derivatives α-Hölder continuous. Being $\phi \in \mathcal{C}^{n+\alpha}(\Omega)$, its norm is defined in the following way:

$$\|\phi\|_{n+\alpha} := \sum_{|\ell| \leq n} \left\| D^l \phi \right\|_\infty + \sum_{|\ell| = n} \sup_{x \neq y} \frac{|D^\ell \phi(x) - D^\ell \phi(y)|}{|x-y|^\alpha}\,.$$

Similarly, the parabolic space $\mathcal{C}^{\frac{n+\alpha}{2}, n+\alpha}(Q_T)$ consists of functions ϕ admitting derivatives

$D_t^r D_x^s \phi$, with $2r+s \le n$, and with norm

$$\|\phi\|_{\frac{n+\alpha}{2},n+\alpha} := \sum_{2r+s\le n} \|D_t^r D_x^s \phi\|_\infty + \sum_{2r+s=n} \sup_t \|D_t^r D_x^s \phi(t,\cdot)\|_\alpha$$
$$+ \sum_{0<n+\alpha-2r-s<2} \sup_x \|D_t^r D_x^s \phi(\cdot,x)\|_{\frac{n+\alpha-2r-s}{2}} .$$

In order to work with Dirichlet boundary conditions, we define the spaces $\mathcal{C}^{n+\alpha,D}(\Omega)$ and $\mathcal{C}^{\frac{n+\alpha}{2},n+\alpha,D}(Q_T)$ as the subspaces of functions belonging to $\mathcal{C}^{n+\alpha}(\Omega)$, resp. to $\mathcal{C}^{\frac{n+\alpha}{2},n+\alpha}(Q_T)$, vanishing at $\partial\Omega$. When there is no risk of confusion, we will omit the set Ω or Q_T.

Analogously, we define the spaces $\mathcal{C}^{0,\alpha}$, $\mathcal{C}^{\alpha,0}$, $\mathcal{C}^{1,2+\alpha}$. For the sake of completeness, also because of its relevance throughout the work, let us specify the norm equipping the latter functional space:

$$\|\phi\|_{1,2+\alpha} := \|\phi\|_\infty + \|\phi_t\|_{0,\alpha} + \|D_x\phi\|_\infty + \left\|D_x^2\phi\right\|_{0,\alpha} .$$

Eventually, we have to work with suitable subsets of the dual spaces of $\mathcal{C}^{n+\alpha}$ and $\mathcal{C}^{n+\alpha,D}$. In particular, we denote as $\mathcal{C}^{-(n+\alpha)}$ the set

$$\mathcal{C}^{-(n+\alpha)} := \left\{ \rho \in \left(\mathcal{C}^{(n+\alpha)}\right)' \,\middle|\, \langle \rho,\phi\rangle = \sum_{|\gamma|\le n} \int_\Omega \partial^\gamma \phi(x)\,\rho_\gamma(dx) \quad \forall \phi \in \mathcal{C}^{n+\alpha}, \quad (\rho_\gamma)_\gamma \text{ measures on } \Omega \right\}$$

In the same way we define the set $\mathcal{C}^{-(n+\alpha),D}$ and $\mathcal{C}^{-\frac{\alpha}{2},-\alpha}$. The norms on these spaces are inherited by the classical dual spaces norm:

$$\|\rho\|_{-(n+\alpha)} = \sup_{\|\phi\|_{n+\alpha}\le 1} \langle \rho,\phi\rangle , \qquad \|\rho\|_{-(n+\alpha),D} = \sup_{\|\phi\|_{n+\alpha,D}\le 1} \langle \rho,\phi\rangle ,$$
$$\|\rho\|_{-\left(\frac{\alpha}{2},\alpha\right)} = \sup_{\|\phi\|_{\frac{\alpha}{2},\alpha}\le 1} \langle \rho,\phi\rangle$$

The importance of these subspaces is given by the following result.

Lemma 3.1. *The space $\mathcal{C}^{-(n+\alpha)}$ is a norm-closed subset of $\left(\mathcal{C}^{n+\alpha}\right)'$. Moreover, if $\rho \in \mathcal{C}^{-n}$, then there exist a sequence $\{\rho_k\}_k \subset \mathcal{C}^{2+\alpha}$ such that $\rho_k \to \rho$ in $\mathcal{C}^{-(n+\alpha)}$ and $\rho_{k|\partial\Omega} = 0$.*

Proof. Thanks to [132, Lemma 2.6], we know that $\mathcal{C}^{-(n+\alpha)}$ is a norm-closed subset of $\left(\mathcal{C}^{n+\alpha}\right)'$. Let $\rho \in \mathcal{C}^{-n}$. Then there exist measures $\{\rho_\gamma\}_{|\gamma|\le n}$ such that $\rho = \sum\limits_{|\gamma|\le n} \partial^\gamma \rho_\gamma(dx)$. Defining $\rho_\gamma(\mathbb{R}^d \setminus \Omega) = 0$, we can consider $\{\rho_\gamma\}_\gamma$ as measures on $\mathbb{R}^d$.

Let $\{n_\varepsilon\}_\varepsilon$ be smooth mollifiers. We define $\rho_\gamma^k := \rho_\gamma * \eta_{\frac{1}{k}}$ and $\tilde\rho_k := \sum\limits_{|\gamma|\le n} \partial^\gamma \rho_\gamma^k(dx)$. Thanks to [132, Lemma 2.3, Lemma 2.7], we have $\tilde\rho_k \in \mathcal{C}^\infty$ and $\tilde\rho_k \to \rho$ in $\mathcal{C}^{-(n+\beta)}$ for all $\beta > 0$.

Then, we consider a nonnegative smooth function $\xi_k(s) \in \mathcal{C}^\infty(\mathbb{R})$ such that $\xi_k(s) = 1$ for $|s| \geq \frac{1}{k}$ and $\xi_k(0) = 0$. We define $\rho_k(x) = \tilde{\rho}_k(x)\xi_k(d(x))$. Then we have $\rho_k \in \mathcal{C}^{2+\alpha}$ and the convergence to ρ in $\mathcal{C}^{-(n+\alpha)}$ is preserved. □

3.2.2 Subprobability Measures and Generalized Wasserstein Measure

Let us start the present subsection with a *proper* notion of distance for elements belonging to $\mathcal{P}^{sub}(\Omega)$. Since the usual Wasserstein distance $W_p(\mu,\nu)$ is defined only when the two measures μ,ν have the same mass, in what follows we rely on the so-called *generalized Wasserstein distance* which allows computing distance between measures with different masses, see e.g. [188, 189] for a complete description and [105, 197] for the corresponding probabilistic interpretation based on optimal transport.

According to [189], we introduce also the following definition which is the one we will use throughout the chapter.

Definition 3.3. *Let $m_1, m_2 \in \mathcal{P}^{sub}(\Omega)$ be two Borel sub-probability measures on Ω. We call the generalized Wasserstein distance between m_1 and m_2, and we write $\mathbf{d}_1(m_1,m_2)$ the quantity*

$$\mathbf{d}_1(m_1,m_2) := \sup_{Lip(\phi)\leq 1,\, \|\phi\|_{C^0}\leq 1} \int_\Omega \phi(x) d(m_1 - m_2)(x)\,, \tag{3.11}$$

We remark that $\mathbf{d}_1(m_1,m_2)$ is equivalent to the well-known *flat metric* over the space of Radon measures with finite mass on Ω.

Furthermore, we define a suitable derivation of U with respect to the (sub)measure m as it appears in the Master Equation (3.7). Namely, we have:

Definition 3.4. *Let $U : \mathcal{P}^{sub}(\Omega) \to \mathbb{R}$. We say that U is of class $\mathcal{C}^1$ if there exists a continuous map $K : \mathcal{P}^{sub}(\Omega) \times \Omega \to \mathbb{R}$ such that, for all m_1, $m_2 \in \mathcal{P}^{sub}(\Omega)$ we have*

$$\lim_{s\to 0} \frac{U(m_1 + s(m_2 - m_1)) - U(m_1)}{s} = \int_\Omega K(m_1,x)(m_2(dx) - m_1(dx))\,. \tag{3.12}$$

We define $\dfrac{\delta U}{\delta m}(m,x) := K(m,x)$.

Remark 3.2. *Condition* (3.12) *defines uniquely the derivative K. Conversely, in the classical setting, where $U : \mathcal{P}(\Omega) \to \mathbb{R}$, see e.g. [41, 194], the function K is defined up to an additive constant, since, for $c \in \mathbb{R}$ and $m_1, m_2 \in \mathcal{P}(\Omega)$, $\int_\Omega c\, d(m_2 - m_1) = 0$. Hence, in those cases a normalization condition is needed to uniquely identify K.*

Let us note that, exploiting Eq. (3.12), we have the following equality for $m_1, m_2 \in \mathcal{P}^{sub}(\Omega)$,

$$U(m_2) - U(m_1) = \int_0^1 \int_\Omega \frac{\delta U}{\delta m}(m_1 + s(m_2 - m_1), x)(m_2(dx) - m_1(dx))ds\,.$$

which turns out to be very useful in our computations. Moreover, we note that, if $\frac{\delta U}{\delta m}$ is $\mathcal{C}^1$ in the space variable, we have

$$|U(m_2) - U(m_1)| \le \sup_m \left\| D_x \frac{\delta U}{\delta m}(m, \cdot) \right\|_\infty \mathbf{d}_1(m_1, m_2)\,,$$

suggesting us define the *intrinsic derivative* of U with respect to m, which directly appears in the Master Equation (3.7).

Definition 3.5. *Let $U : \mathcal{P}^{sub}(\Omega) \to \mathbb{R}$. If U is of class $\mathcal{C}^1$ and $\frac{\delta U}{\delta m}$ is of class $\mathcal{C}^1$ with respect to the last variable, we define the intrinsic derivative $D_m U : \mathcal{P}^{sub}(\Omega) \times \Omega \to \mathbb{R}^d$ as*

$$D_m U(m, x) := D_x \frac{\delta U}{\delta m}(m, x)\,.$$

3.5.1 Assumptions

We require the following hypotheses throughout the report:

Hypotheses 1. *Let $0 < \alpha < 1$. Assume that*

(i) $\|a(\cdot)\|_{1+\alpha} < \infty$ *and* $\exists \mu > \lambda > 0$ *s.t.* $\forall \xi \in \mathbb{R}^d$

$$\lambda |\xi|^2 \le \langle a(x)\xi, \xi \rangle \le \mu |\xi|^2\,;$$

(ii) $H : [0,T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}$, $G : \Omega \times \mathcal{P}^{sub}(\Omega) \to \mathbb{R}$ *and* $F : [0,T] \times \Omega \times \mathcal{P}^{sub}(\Omega) \to \mathbb{R}$ *are smooth functions with H locally Lipschitz with respect to the last variable;*

(iii) $\exists C > 0$ *s.t.*

$$0 < H_{pp}(t, x, p) \le C I_{d\times d}\,;$$

(iv) F is increasing in the last variable, i.e.

$$\int_\Omega (F(t, x, m) - F(t, x, m'))\, d(m - m')(x) \ge 0\,;$$

moreover

$$\sup_{m \in \mathcal{P}^{sub}(\Omega)} \left(\|F(\cdot,\cdot,m)\|_{\frac{\alpha}{2},\alpha} + \left\| \frac{\delta F}{\delta m}(\cdot,\cdot,m,\cdot) \right\|_{\frac{\alpha}{2},\alpha,2+\alpha} \right) + \mathrm{Lip}\left(\frac{\delta F}{\delta m}\right) \le C_F\,,$$

with

$$\operatorname{Lip}\left(\frac{\delta F}{\delta m}\right) := \sup_{m_1 \neq m_2}\left(\|m_1 - m_2\|^{-1}_{-(2+\alpha),D}\left\|\frac{\delta F}{\delta m}(\cdot,\cdot,m_1,\cdot) - \frac{\delta F}{\delta m}(\cdot,\cdot,m_2,\cdot)\right\|_{\frac{\alpha}{2},\alpha,2+\alpha}\right);$$

(v) G satisfies the same estimates as F with α and $1+\alpha$ replaced by $2+\alpha$, i.e.

$$\sup_{m\in\mathcal{P}^{sub}(\Omega)}\left(\|G(\cdot,m)\|_{2+\alpha} + \left\|\frac{\delta G}{\delta m}(\cdot,m,\cdot)\right\|_{2+\alpha,2+\alpha}\right) + \operatorname{Lip}\left(\frac{\delta G}{\delta m}\right) \le C_G\,,$$

with

$$\operatorname{Lip}\left(\frac{\delta G}{\delta m}\right) := \sup_{m_1 \neq m_2}\left(\|m_1 - m_2\|^{-1}_{-(2+\alpha),D}\left\|\frac{\delta G}{\delta m}(\cdot,m_1,\cdot) - \frac{\delta G}{\delta m}(\cdot,m_2,\cdot)\right\|_{2+\alpha,2+\alpha}\right);$$

(vi) Moreover, we require the following Dirichlet boundary conditions:

$$\frac{\delta F}{\delta m}(t,x,m,y)_{|y\in\partial\Omega} = 0\,, \qquad \frac{\delta G}{\delta m}(x,m,y)_{|y\in\partial\Omega} = 0\,, \qquad G(x,m)_{|x\in\partial\Omega} = 0\,,$$

for all $m \in \mathcal{P}^{sub}(\Omega)$.

We note that in hypotheses (vi) we have two standard compatibility assumptions for G and $\frac{\delta G}{\delta m}$, naturally linked to the boundary conditions framework. In particular, the condition on G is essential to have a classical solution for both the Master Equation (3.7) and the MFG system (3.8), whereas the condition on $\frac{\delta G}{\delta m}$ is a compatibility condition for the linearized MFG system, see e.g. *Corollary* 3.13.

We underline that condition on F is a novelty, and it will play a crucial role when considering the Dirichlet boundary condition of $\frac{\delta U}{\delta m}$.

3.5.2 Convergence in the Nash system

The second part of the chapter is focused on the convergence problem. As already said, the *value function* v_i^N is defined as the cost function related to the optimal control, and functions $\{v_i^N\}_i$ solve the Nash system (3.5).

The idea is to define proper finite dimensional projections U along trajectories $m_{\boldsymbol{x}}^{N,i}$. To this end, we define the following functions u_i^N:

$$u_i^N(t,\mathbf{x}) = U(t,x_i,m_{\mathbf{x}}^{N,i})$$

then proving that they *almost* solve the Nash system (3.5), with an error of order $\frac{1}{N}$, and such that, within suitable spaces, $|u_i^N - v_i^N| \to 0$, for $N \to +\infty$.

The main Theorem of this part is the following:

Theorem 3.3. *Assume Hypotheses 1 hold true and let $(v^{N,i})_{i\in\{1,\dots,N\}}$ be the solution of* (3.5) *and U the solution of the master equation* (3.7)*. Fix $N \geq 1$ and $(t_0,m_0) \in [0,T] \times \mathcal{P}^{sub}(\Omega)$.*

(i) For any $\boldsymbol{x} \in (\Omega)^N$, let $m_{\boldsymbol{x}}^N := \dfrac{1}{N}\sum_i^N \delta_{x_i}\mathbb{1}_{\{x_j\in int(\Omega)\}}$. Then,

$$\sup_{i=1,\dots,N}\left|v^{N,i}(t_0,\boldsymbol{x}) - U(t_0,x_i,m_{\boldsymbol{x}}^N)\right| \leq \frac{C}{N}$$

(ii) For any $i \in \{1,\dots,N\}$ and $x_i \in \Omega$, we define

$$w^{N,i}(t_0,x_i,m_0) := \int_{\Omega^{N-1}} v^{N,i}(t_0,\boldsymbol{x})\prod_{j\neq i} m_0(dx_j)\,,$$

where $\boldsymbol{x} = (x_1,\dots,x_N)$. Then,

$$||w^{N,i}(t_0,\cdot,m_0) - U(t_0,\cdot,m_0)||_{L^1(m_0)} \leq \begin{cases} CN^{-1/d} & \text{if } d\geq 3 \\ CN^{-1/2}\log(N) & \text{if } d=2 \\ CN^{-1/2} & \text{if } d=1 \end{cases} \tag{3.13}$$

3.6 Well-posedness of the Master Equation

A formal proof of existence and uniqueness of solutions for (3.7) can be easily given directly using the very definition of U. The difficulty here relies on the proof of the $\mathcal{C}^1$ character of U with respect to m. Once proved this and the boundary conditions of U, the existence and uniqueness theorem are rigorously proved, as we will see in the following theorem.

Theorem 3.4. *Suppose hypotheses 1 are satisfied. Take U defined as in* (3.9)*, suppose that $\frac{\delta U}{\delta m}$ exists, being also bounded in $\mathcal{C}_x^{2+\alpha} \times C_y^{2+\alpha}$, uniformly in t and continuously in m, and that boundary conditions of* (3.7) *hold for U. Then U is the unique classical solution U of the Master Equation* (3.7)*.*

Proof. Existence. By a density argument, we only need to prove that U solves (3.7) at (t_0,x,m_0), where m_0 is a smooth and positive function satisfying $m_{0|\partial\Omega}=0$ and $\int_\Omega m_0(x)dx \le 1$. Taking (u,m) as the solution of the MFG system starting from m_0 at time t_0, by the definition of U, we have that for all $s>t_0$

$$U(s,x,m(s)) = u(s,x)\,,$$

therefore, we can write

$$\partial_t U(t_0,x,m_0) = \lim_{h\to 0}\frac{U(t_0+h,x,m_0)-U(t_0+h,x,m(t_0+h))}{h} + u_t(t_0,x)\,.$$

Again by the very definition of U:

$$u_t(t_0,x) = -\mathrm{tr}(a(x)D_x^2 U(t_0,x,m_0)) + H(t_0,x,D_xU(t_0,x,m_0)) - F(t_0,x,m_0)\,.$$

As regards the limit part, we define $m_s := (1-s)m(t_0)+sm(t_0+h)$. Since U is $\mathcal{C}^1$ with respect to m, then:

$$\begin{aligned}
&-\lim_{h\to 0}\int_0^1\int_\Omega \frac{\delta U}{\delta m}(t_0+h,x,m_s,y)\frac{(m(t_0+h,y)-m(t_0,y))}{h}\,dyds = \\
&\quad -\int_\Omega \frac{\delta U}{\delta m}(t_0,x,m_0,y)\Bigg(\sum_{ij}\partial^2_{ij}(a_{ij}(x)m(t_0,y)) \\
&\qquad\qquad + \mathrm{div}\big(m(t_0,y)H_p(t_0,y,Du(t_0,y))\big)\Bigg)dy\,.
\end{aligned}$$

Integrating by parts, and using the boundary conditions of $\frac{\delta U}{\delta m}$ and m, we can rewrite the right-hand side as

$$\begin{aligned}
&\int_\Omega \Bigg[H_p(t_0,y,D_xU(t_0,y,m_0))D_mU(t_0,x,m_0,y) \\
&\qquad\qquad - \mathrm{tr}\Big(a(y)D_yD_mU(t_0,x,m_0,y)\Big)\Bigg]dm_0(y)\,,
\end{aligned}$$

therefore

$$\begin{aligned}\partial_t U(t,x,m) = &-\mathrm{tr}\left(a(x)D_x^2 U(t,x,m)\right) + H\left(t,x,D_x U(t,x,m)\right)\\ &-\int_\Omega \mathrm{tr}\left(a(y)D_y D_m U(t,x,m,y)\right)dm(y)\\ &+\int_\Omega D_m U(t,x,m,y)\cdot H_p(t,y,D_x U(t,y,m))dm(y) - F(t,x,m)\,,\end{aligned}$$

hence concluding the existence part.

Uniqueness. Suppose that there exists a second solution V of (3.7). We fix (t_0,m_0), with m_0 smooth s.t. $m_{0|\partial\Omega}=0$, and we take $\tilde m$ as the solution of

$$\begin{cases}\tilde m_t - \sum_{ij}\partial^2_{ij}(a_{ij}(x)\tilde m) - \mathrm{div}\big(\tilde m H_p(t,x,D_x V(t,x,\tilde m))\big) = 0\,,\\ \tilde m(t_0) = m_0\,,\\ \tilde m_{|\partial\Omega} = 0\end{cases}$$

which is well-posed since $D_x V$ is Lipschitz continuous with respect to the measure. Then we define $\tilde u(t,x) = V(t,x,\tilde m(t))$ and we want to prove that $\tilde u$ solves a Hamilton-Jacobi equation. Using the equations of V and $\tilde m$, we get

$$\begin{aligned}\tilde u_t(t,x) = &V_t(t,x,\tilde m(t)) + \int_\Omega \frac{\delta V}{\delta m}(t,x,\tilde m(t),y)\,\tilde m_t(t,y)\,dy\\ = &V_t(t,x,\tilde m(t)) + \int_\Omega \frac{\delta V}{\delta m}(t,x,\tilde m(t),y)\sum_{ij}\partial^2_{ij}(a_{ij}(x)\tilde m(t,y))\,dy\\ &+\int_\Omega \frac{\delta V}{\delta m}(t,x,\tilde m(t),y)\,\mathrm{div}\big(\tilde m H_p(t,x,D_x V(t,x,\tilde m))\big)\,dy\,.\end{aligned}$$

Previous integrals can be easily estimated with an integration by parts, and taking into account the boundary condition for V. As regards the first, we have

$$\begin{aligned}&\int_\Omega \frac{\delta V}{\delta m}(t,x,\tilde m(t),y)\sum_{ij}\partial^2_{ij}(a_{ij}(x)\tilde m(t,y))\,dy =\\ &\qquad\int_\Omega \mathrm{tr}(a(y)D_y D_m V(t,x,\tilde m(t),y))\,\tilde m(t,y)\,dy\,,\end{aligned}$$

while for the second

$$\begin{aligned}&\int_\Omega \frac{\delta V}{\delta m}(t,x,\tilde m(t),y)\,\mathrm{div}\big(\tilde m H_p(t,x,D_x V(t,x,\tilde m))\big)\,dy\\ &\qquad= -\int_\Omega H_p(t,x,D_x V(t,x,\tilde m))D_m V(t,x,\tilde m,y)\tilde m(t,y)dy\,.\end{aligned}$$

Previous estimates allow us to write

$$\begin{aligned}\tilde{u}_t(t,x) = & V_t(t,x,\tilde{m}(t)) + \int_\Omega \operatorname{tr}(a(y)D_yD_mV(t,x,\tilde{m},y))\,d\tilde{m}(y)\\ & - \int_\Omega H_p(t,y,D_xV(t,y,\tilde{m}))D_mV(t,x,\tilde{m},y)\,d\tilde{m}(y)\\ = & -\operatorname{tr}(a(x)D^2\tilde{u}(t,x)) + H(t,x,D\tilde{u}(t,x)) - F(t,x,\tilde{m}(t))\,,\end{aligned}$$

hence $(\tilde{u},\tilde{m})$ is a solution of the MFG system (3.8). Then, by uniqueness of solutions for the MFG system, we get $(\tilde{u},\tilde{m}) = (u,m)$, implying that $V(t_0,x,m_0) = U(t_0,x,m_0)$, whenever m_0 is smooth. Therefore, by density, the uniqueness is proved. □

In the next two sections, we will prove the $\mathcal{C}^1$ character of U with respect to m, the $\mathcal{C}^2$ regularity in x and y of $\frac{\delta U}{\delta m}$ and the boundary conditions for U, hence justify the hypotheses previously given, also allowing us to then apply Theorem 3.4 and finally prove the well-posedness of (3.7).

3.7 The Fokker-Planck equation and the MFG system

Let us first give the following technical lemma, providing a regularity result for a linear PDE with non-homogeneous drift.

Lemma 3.5. *Suppose a satisfies hypothesis (i) of 1 , let $q = \frac{d+2}{1-\alpha}$, $b, f \in L^q(Q_T)$, $\psi \in \mathcal{C}^{1+\alpha,D}(\Omega)$, and z be the solution of*

$$\begin{cases} -z_t - \operatorname{tr}(a(x)D^2z) + b(t,x)\cdot Dz = f(t,x)\,,\\ z(T) = \psi\,,\\ z_{|\partial\Omega} = 0\,. \end{cases}$$

Then z satisfies

$$\|z\|_{\frac{1+\alpha}{2},1+\alpha} \le C\left(\|f\|_{L^q} + \|\psi\|_{1+\alpha}\right). \tag{3.14}$$

Moreover, if $b, f \in \mathcal{C}^{0,\alpha}(Q_T)$ and $\psi \in \mathcal{C}^{2+\alpha,D}(\Omega)$, it holds

$$\|z\|_{1,2+\alpha} \le C\left(\|f\|_{0,\alpha} + \|\psi\|_{2+\alpha}\right). \tag{3.15}$$

Proof. We exploit ideas stated in [194]. In particular, if $f \in \mathcal{C}(Q_T)$, $b \in \mathcal{C}(\Omega)$, then we conclude with *Theorem 5.1.1* of [158]. While, in the general case, we write $z = z_1 + z_2$, where z_1 and z_2 respectively satisfy:

$$\begin{cases} -(z_1)_t - \operatorname{tr}(a(x)D^2 z_1) = 0\,, \\ z_1(T) = \psi\,, \\ z_{1|\partial\Omega} = 0\,, \end{cases}$$

and

$$\begin{cases} -(z_2)_t - \operatorname{tr}(a(x)D^2 z_2) + b \cdot Dz_2 = f - b \cdot Dz_1\,, \\ z_2(T) = 0\,, \\ z_{2|\partial\Omega} = 0\,. \end{cases}$$

For z_1, we apply *Theorem 5.1.11* of [158] to obtain $\|z_1\|_{\frac{1+\alpha}{2},1+\alpha} \le C\|\psi\|_{1+\alpha}$. For z_2, by Corollary of *Theorem IV.9.1* in [146], we have

$$\|z_2\|_{\frac{1+\alpha}{2},1+\alpha} \le C\|f - bDz_1\|_{L^q} \le C\left(\|f\|_{L^q} + \|\psi\|_{1+\alpha}\right)\,,$$

concluding the first part. Then, if $b, f \in \mathcal{C}^{0,\alpha}(Q_T)$ and $\psi \in \mathcal{C}^{2+\alpha,D}$, by (3.14), we have $f - b \cdot Dz \in \mathcal{C}^{0,\alpha}$, allowing to apply *Theorem 5.1.13* in [158] to conclude. □

3.7.1 The Fokker-Planck equation

From now on, q will denote the quantity $q := \frac{d+2}{1-\alpha}$, and p will be the conjugate exponent of q, i.e. $p := \frac{d+2}{d+1+\alpha}$.

The present section is focused on the study of the following Fokker-Planck equation

$$\begin{cases} \mu_t - \operatorname{div}(a(x)D\mu) - \operatorname{div}(\mu b) = \operatorname{div}(c)\,, \\ \mu(t_0) = \mu_0\,, \\ \mu_{|\partial\Omega} = 0\,, \end{cases} \tag{3.16}$$

where $c \in L^1(Q_T)$, $\mu_0 \in \mathcal{C}^{-(1+\alpha)}$, and $b \in L^q(Q_T)$. Let us underline that the latter equation plays a fundamental role in studying both the Mean Field Games system and the linearized one.

The main difficulty here relies on low regularity of both data and coefficients. Hence, the main idea is to start with the regular case, where we have the existence and uniqueness of solutions, to then obtain estimates that we will then exploit to pass to the limit in the general case.

Let us start by giving a suitable definition of a solution for the system (3.16):

Definition 3.8. *Let $c \in L^1$, $\mu_0 \in \mathcal{C}^{-1}$, $b \in L^q$. We say that a function $\mu \in \mathcal{C}([0,T];\mathcal{C}^{-(1+\alpha),D}) \cap L^1(Q_T)$ is a weak solution of* (3.16) *if, for all $t \in (t_0,T]$ and ϕ satisfying in $[t_0,t]\times\Omega$ the following linear equation*

$$\begin{cases} -\phi_t - \operatorname{div}(aD\phi) + bD\phi = \psi\,, \\ \phi(t) = \xi\,, \\ \phi_{|\partial\Omega} = 0\,, \end{cases} \tag{3.17}$$

with $\psi \in L^\infty(\Omega)$ and $\xi \in \mathcal{C}^{1+\alpha,D}$, it holds true:

$$\langle \mu(t),\xi\rangle + \int_{t_0}^t \int_\Omega \mu(s,x)\psi(s,x)\,dxds = \langle \mu_0,\phi(t_0,\cdot)\rangle - \int_{t_0}^t\int_\Omega c(s,x)\cdot D\phi(s,x)\,dxds\,, \tag{3.18}$$

where $\langle\cdot,\cdot\rangle$ denotes, respectively, the duality between $\mathcal{C}^{-(1+\alpha),D}$ and $\mathcal{C}^{1+\alpha,D}$, $\mathcal{C}^{-(1+\alpha)}$ and $\mathcal{C}^{1+\alpha}$.

Note that Lemma 3.5 implies $\phi \in \mathcal{C}^{\frac{1+\alpha}{2},1+\alpha}$, therefore the terms on the right-hand side are well-defined and the definition is well-posed.

Proposition 3.6. *Let $c \in L^1$, $\mu_0 \in \mathcal{C}^{-1}$, $b \in L^q$. Then there exists a unique solution for* (3.16)*, which satisfies, for a certain C depending on a and b,*

$$\sup_{t\in[t_0,T]} \|\mu(t)\|_{-(1+\alpha),D} + \|\mu\|_{L^p} \le C\left(\|\mu_0\|_{-(1+\alpha)} + \|c\|_{L^1}\right)\,, \tag{3.19}$$

Moreover, if $\mu_0^n \to \mu_0$ in $\mathcal{C}^{-(1+\alpha)}$, $b^n \to b$ in L^q, $c^n \to c$ in L^1, we have $\mu^n \to \mu$ in $\mathcal{C}([0,T];\mathcal{C}^{-(1+\alpha),D}) \cap L^p(Q_T)$, where μ^n and μ are the solutions related to (μ_0^n, b^n, f^n) and (μ_0, b, f).

Proof. Without loss of generality, we can consider $t_0 = 0$.

Existence: smooth case. Assume that f, b and μ_0 are smooth functions, with μ_0 satisfying $\mu_0(x)_{|x\in\partial\Omega} = 0$. Then, splitting the divergence terms in (3.16), we obtain a linear equation and the existence of solutions is guaranteed by [146, 158].

Let ϕ be a solution of (3.17), with $\psi = 0$ and $\xi \in \mathcal{C}^{1+\alpha,D}$. Using ϕ as a test function for μ, we get by Lemma 3.5

$$\begin{aligned} \langle \mu(t),\xi\rangle &= \langle \mu_0,\phi(0,\cdot)\rangle - \int_0^t\int_\Omega c(s,x)\cdot D\phi(s,x)\,dxds \\ &\le C\|\xi\|_{1+\alpha}\left(\|\mu_0\|_{-(1+\alpha)} + \|c\|_{L^1}\right)\,. \end{aligned}$$

Passing to the *sup* for $\|\xi\|_{1+\alpha} \le 1$, we obtain

$$\sup_{t\in[0,T]} \|\mu(t)\|_{-(1+\alpha),D} \le C\left(\|\mu_0\|_{-(1+\alpha)} + \|c\|_{L^1}\right)\,, \tag{3.20}$$

and we are left to prove the L^p bound for μ. To this end, we take ϕ as the solution of (3.17), with $t=T$, $\xi=0$ and $\psi \in L^q$. Then, again by Lemma 3.5, we have

$$\begin{aligned}\int_0^T\int_\Omega \mu\psi\,dxds &= \langle \mu_0, \phi(0,\cdot)\rangle - \int_0^T\int_\Omega c(s,x)\cdot D\phi(s,x)\,dxds \\ &\le C\|\psi\|_{L^q}\left(\|\mu_0\|_{-(1+\alpha)} + \|c\|_{L^1}\right).\end{aligned}$$

Passing to the sup for $\|\mu\|_{L^q}\le 1$, we get

$$\|\mu\|_{L^p} \le C\left(\|\mu_0\|_{-(1+\alpha)} + \|c\|_{L^1}\right),$$

which concludes the proof in the regular case.

Existence: general case. We proceed adapting and adjusting [194, Proposition 5.3]. By Proposition 3.1, we take smooth data μ_0^k, c_k, b_k, with $\mu_0^k(x)=0$ for $x\in\partial\Omega$, and converging to μ_0, c, b, respectively in $\mathcal{C}^{-(1+\alpha),D}$, L^1 and L^q.

We consider μ^k as the solution of (3.16). The previous convergences imply, for a certain $C>0$,

$$\begin{aligned}||\mu_0^k||_{-(1+\alpha)} &\le C\|\mu_0\|_{-(1+\alpha)},\\ \|b_k\|_{L^q} &\le C\|b\|_{L^q},\\ ||c_k||_{L^1} &\le C\|c\|_{L^1}.\end{aligned}$$

Considering the function $\mu^{k,h} := \mu^k - \mu^h$, which satisfies (3.16) with data $b=b_k$, $c = c_k - c_h + \mu_h(b_k - b_h)$, $\mu^0 = \mu_0^k - \mu_0^h$, then, by estimates (3.19), we have

$$\begin{aligned}&\sup_t ||\mu^{k,h}(t)||_{-(1+\alpha),D} + ||\mu^{k,h}||_{L^p}\\ &\le C\left(||\mu_0^k - \mu_0^h||_{-(1+\alpha)} + ||c_k - c_h||_{L^1} + ||\mu_h(b_k-b_h)||_{L^1}\right),\end{aligned}$$

By the uniform bound of $\{\mu_k\}_k$ in L^p, the last term can be estimated as follows

$$\left\|\mu^h(b_k - b_h)\right\|_{L^1} \le C||b_k - b_h||_{L^q}, \tag{3.21}$$

and since $\{\mu_0^k\}_k$, $\{c_k\}_k$ and $\{b_k\}_k$ are Cauchy sequences, the right-hand side goes to 0 and $\{\mu^k\}_k$ is a Cauchy sequence, too. Then $\mu^k \to \mu$ in $\mathcal{C}([0,T];\mathcal{C}^{-(1+\alpha),N})\cap L^p(Q_T)$, for a certain μ also satisfying (3.19). We are then left to prove that μ satisfies (3.18), for ϕ satisfying (3.17). To this end, we consider the test function ϕ_k related to μ^k, so that

$$\langle \mu^k(t),\xi\rangle + \int_0^t\int_\Omega \mu^k(s,x)\psi(s,x)\,dxds = \langle \mu_0^k, \phi_k(0,\cdot)\rangle - \int_0^t\int_\Omega c_k(s,x)\cdot D\phi_k(s,x)\,dxds, \tag{3.22}$$

Concerning convergence of ϕ_k towards ϕ, since $\phi_k - \phi$ satisfies

$$\begin{cases} -(\phi_k-\phi)_t - \operatorname{div}(aD(\phi_k-\phi)) + b_k D(\phi_k-\phi) = (b_k-b)D\phi\,, \\ (\phi_k-\phi)(t) = 0\,, \\ (\phi_k-\phi)_{|\partial\Omega} = 0\,, \end{cases}$$

by Lemma 3.5, we have

$$||\phi_k-\phi||_{\frac{1+\alpha}{2},1+\alpha} \le C||(b_k-b)D\phi||_{L^q} \to 0\,,$$

allowing us to pass to the limit in (3.22), hence ending the existence part.

Uniqueness and stability. Let μ_1 and μ_2 be two solutions. Then $\mu := \mu_1 - \mu_2$ solves the same problem with $c = \mu_0 = 0$, and by (3.19) we have $\|\mu\|_{L^1} = \sup_{t\in[0,T]} \|\mu(t)\|_{-(1+\alpha),D} = 0$, concluding the uniqueness part.

For the stability part, we consider $c_n \to c$, $\mu_0^n \to \mu_0$ and $b_n \to b$. Then the function $\mu^n - \mu$ satisfies (3.16), with $b\,,\mu_0$ and f respectively replaced by b^n, $\mu_0^n - \mu_0$, and $c^n - c + \mu(b^n - b)$. Hence, (3.19) implies

$$\begin{aligned} &\sup_{t\in[0,T]} \|\mu^n-\mu\|_{-(1+\alpha),D} + \|\mu^n-\mu\|_{L^p} \\ &\le C\left(||\mu_0^n-\mu_0||_{-(1+\alpha)} + ||c^n-c||_{L^1} + ||\mu(b^n-b)||_{L^1}\right)\,. \end{aligned}$$

Proceeding as done for the existence in the general case, we derive the right-hand side convergence to 0, which implies $\mu^n \to \mu$ in $\mathcal{C}([0,T];\mathcal{C}^{-(1+\alpha),D}) \cap L^p(Q_T)$, hence concluding the proof. □

As a direct consequence of Prop. (3.6), we can derive a further estimate in the space $\mathcal{C}^{-(2+\alpha),D}$.

Corollary 3.7. *Suppose hypotheses of Proposition 3.6 are satisfied together with $b \in \mathcal{C}^{\frac{\alpha}{2},\alpha}$. Then the function μ satisfies*

$$\sup_{t\in[0,T]} \|\mu(t)\|_{-(2+\alpha),D} + \|\mu\|_{-\left(\frac{\alpha}{2},\alpha\right)} \le C\left(\|\mu_0\|_{-(2+\alpha)} + \|c\|_{L^1}\right)\,, \tag{3.23}$$

$$\tag{3.24}$$

Proof. Since $b \in \mathcal{C}^{\frac{\alpha}{2},\alpha}$, we can take ϕ as the solution of (3.17), with $\xi \in C^{2+\alpha,D}(\Omega)$, $\psi = 0$, so that, applying the results of [146], [158], we have $\|\phi\|_{1+\frac{\alpha}{2},2+\alpha} \leq C\|\xi\|_{2+\alpha}$, and (3.18) implies

$$\begin{aligned}\langle\mu(t),\xi\rangle &= \langle\mu_0,\phi(0,\cdot)\rangle - \int_0^T\int_\Omega c(s,x)\cdot D\phi(s,x)\,dxds\\ &\leq C\left(\|\mu_0\|_{-(2+\alpha)} + \|c\|_{L^1}\right)\|\xi\|_{2+\alpha}\,.\end{aligned}$$

Now, as in Proposition 3.6, choose ϕ as the solution of (3.17), with $t=T$, $\xi=0$ and $\psi\in\mathcal{C}^{\frac{\alpha}{2},\alpha}$. Then, again by Lemma 3.5, we have

$$\begin{aligned}\int_0^T\int_\Omega \mu\psi\,dxds &= \langle\mu_0,\phi(0,\cdot)\rangle - \int_0^T\int_\Omega c(s,x)\cdot D\phi(s,x)\,dxds\\ &\leq C\|\psi\|_{\frac{\alpha}{2},\alpha}\left(\|\mu_0\|_{-(2+\alpha)} + \|c\|_{L^1}\right)\,.\end{aligned}$$

Passing to the sup in the two inequalities for $\|\xi\|_{1+\frac{\alpha}{2},2+\alpha}\leq 1$ and $\|\mu\|_{L^q}\leq 1$, we get (3.23). $\square$

3.8.1 Mean Field Games system and Lipschitz regularity of U

Now, we turn out to the study of the Mean Field Games system and its properties to derive essential results we will then exploit to obtain some first regularity estimates of the solution U of the Master Equation.

In particular, the MFG system has the following form:

$$\begin{cases}-\partial_t u - \mathrm{tr}(a(x)D^2u) + H(t,x,Du) = F(t,x,m(t))\,,\\ \partial_t m - \sum\limits_{i,j}\partial^2_{ij}(a_{ij}(x)m) - \mathrm{div}(mH_p(t,x,Du)) = 0\,,\\ u(T) = G(x,m(T))\,, \qquad m(0)=m_0\,,\\ u_{|\partial\Omega=0}\,, \qquad m_{|\partial\Omega=0}\,.\end{cases} \tag{3.25}$$

The first result is obtained by studying some regularity properties of (3.25) uniformly in m_0.

Proposition 3.8. *The system* (3.25) *has a unique classical solution* (u,m), *with* $u\in\mathcal{C}^{1+\frac{\alpha}{2},2+\alpha}$ *and* $m\in\mathcal{C}([t_0,T];\mathcal{P}^{sub}(\Omega))$. *Moreover,* $m(t)$ *has a positive density for* $t>t_0$, *and the following estimates hold:*

$$\|m\|_{\frac{\alpha}{2},-(1+\alpha),D} + \|m\|_{L^p} + \|u\|_{1+\frac{\alpha}{2},2+\alpha} \leq C\,. \tag{3.26}$$

Furthermore, if (u_1,m_1) *and* (u_2,m_2) *are two solutions of* (3.25), *with* $m_1(t_0)=m_{01}$, $m_2(t_0)=$

m_{02}, *then it holds*

$$\begin{aligned} &\|u_1-u_2\|_{1,2+\alpha}+\|m_1-m_2\|_{L^p(Q_T)} \le C\mathbf{d}_1(m_{01},m_{02})\,, \\ &\sup_{t\in[0,T]}\|m_1(t)-m_2(t)\|_{-(1+\alpha),D} \le C\mathbf{d}_1(m_{01},m_{02})\,. \end{aligned} \tag{3.27}$$

Proof. Step 1: Existence and uniqueness of solutions. We want to apply Schauder's fixed point Theorem. Let X be the convex set, closed for the uniform distance, defined as follows:

$$X := \left\{\gamma \in \mathcal{C}([t_0,T];\mathcal{P}^{sub}(\Omega)) \text{ s.t. } \|\gamma\|_{\frac{\alpha}{2},-(1+\alpha),D} \le M \ \forall s,t \in [t_0,T]\right\},$$

where M will be specified later. We define a map $\Phi : X \to X$ as follows.
For $\gamma \in X$, let $u = u_\gamma$ be the solution of the HJB equation:

$$\begin{cases} -u_t - \operatorname{tr}(a(x)D^2u) + H(t,x,Du) = F(t,x,\gamma(t))\,, \\ u(T) = G(x,\gamma(T))\,, \\ u_{|\partial\Omega} = 0\,. \end{cases} \tag{3.28}$$

Using hypotheses on F and G and *Theorem V.6.1* of [146], we know that there exists a unique classical solution $u \in \mathcal{C}^{1+\frac{\alpha}{2},\alpha}$. Moreover, by Taylor formula, we can write $H(t,x,Du) = H(t,x,0)+V(t,x)\cdot Du$ for a certain $V \in L^\infty$, thanks to the global boundedness of Du, obtaining a linear equation satisfied by u. Then, exploiting both the Corollary of *Theorem IV.9.1* and *Theorem IV.5.2* of [146] we get

$$\|u\|_{1+\frac{\alpha}{2},2+\alpha} \le C\left(\|F(\cdot,\cdot,\gamma(\cdot))\|_{\frac{\alpha}{2},\alpha}+\|G(\cdot,\gamma(T))\|_{2+\alpha}\right),$$

Define $\Phi(\beta) = m$, where $m \in \mathcal{C}([t_0,T];\mathcal{P}^{sub}(\Omega))$ is the solution of the FP equation

$$\begin{cases} m_t - \sum\limits_{i,j}\partial^2_{ij}(a_{ij}(x)m) - \operatorname{div}(mH_p(t,x,Du)) = 0\,, \\ m(t_0) = m_0\,, \\ m_{|\partial\Omega} = 0\,. \end{cases} \tag{3.29}$$

The existence of a unique solution for (3.29) is guaranteed by Proposition 3.6. Moreover, we know from (3.19) that

$$\sup_{t\in[t_0,T]}\|m(t)\|_{-(1+\alpha),D}+\|m\|_{L^p} \le C\|m_0\|_{-(1+\alpha)}\,,$$

for a certain $p>1$. Therefore, to check that $m \in X$, we are left with proving that, for some $C>0$, it holds

$$\|m(t)-m(s)\|_{-(1+\alpha),D} \leq C|t-s|^{\frac{\alpha}{2}}\,, \qquad \text{for all } t \neq s\,.$$

Subtracting the distributional formulation (3.18) in t and s, we have

$$\int_\Omega \xi(x)m(t,dx) - \int_\Omega \phi(s,x)m(s,dx) + \int_s^t \int_\Omega \psi(r,x)m(r,dx)dr = 0\,, \tag{3.30}$$

for each $\xi \in \mathcal{C}^{1+\alpha,D}$, $\psi \in L^\infty$ and ϕ satisfying

$$\begin{cases} -\phi_t - \mathrm{tr}(a(x)D^2\phi) + H_p(t,x,Du)\cdot D\phi = \psi\,, \\ \phi(t) = \xi\,, \\ \phi_{|\partial\Omega} = 0\,. \end{cases} \tag{3.31}$$

We choose $\psi = 0$. Thanks to Lemma 3.5 we know that $\phi \in \mathcal{C}^{1+\frac{\alpha}{2},1+\alpha}$ and its norm is bounded according to (3.14). Coming back to (3.30), we obtain

$$\int_\Omega \xi(x)(m(t,dx)-m(s,dx)) = \int_\Omega (\phi(t,x)-\phi(s,x))m(s,dx) \leq C\|\phi(t)-\phi(s)\|_{1+\alpha}$$
$$\|m(s)\|_{-(1+\alpha),D} \leq C|t-s|^{\frac{\alpha}{2}}\|m_0\|_{-(1+\alpha)}\|\xi\|_{1+\alpha}\,,$$

and taking the sup over the $\xi \in \mathcal{C}^{1+\alpha,D}$ with $\|\xi\|_{1+\alpha} \leq 1$,

$$\|m(t)-m(s)\|_{-(1+\alpha),D} \leq C|t-s|^{\frac{\alpha}{2}}\|m_0\|_{-(1+\alpha)}\,.$$

Choosing $M = C\|m_0\|_{-(1+\alpha)}$, we have proved that $m \in X$.
Since X is convex and closed and $\Phi(X) \subseteq X$, to apply Schauder's theorem we need to show that:

- $\Phi(X)$ is relatively compact;
- Φ is continuous.

In this way, the closure of the convex envelope of $\Phi(X)$, say $\hat{X} := \overline{\mathrm{inv}(\Phi(X))}$, is compact and convex, and for the closure and the convexity of X we have $\Phi(X) \subseteq \hat{X} \subseteq X$. So we can consider the restriction $\Phi_| : \hat{X} \to \hat{X}$, which satisfies the classical hypotheses of the Schauder's Theorem. Hence, the existence of a fixed point remains guaranteed.

We start proving the relatively compactness of $\Phi(X)$. Let $\{\gamma_n\}_n \subset X$ and let u_n and m_n be the related solutions. Applying Ascoli-Arzelà's Theorem we have $u_{n_k} \to u$ in $\mathcal{C}^{1,2}$, for a

certain subsequence $\{u_{n_k}\}_k$ and $u \in X$.
To prove the convergence of $\{m_{n_k}\}_n$, we take ϕ_{n_k} as the solution of (3.31) with Du replaced by Du_{n_k} and $\psi = 0$. The difference $\phi_{k,h} := \phi_{n_k} - \phi_{n_h}$ satisfies

$$\begin{cases} -(\phi_{k,h})_t - \mathrm{tr}(a(x)D^2\phi_{k,h}) + H_p(t,x,Du_{n_k})\cdot D\phi_{k,h} \\ \qquad\qquad\qquad\qquad = (H_p(t,x,Du_{n_h}) - H_p(t,x,Du_{n_k}))\cdot D\phi_{n_h}\,, \\ \phi_{k,h}(t) = 0\,, \\ \phi_{k,h|\partial\Omega} = 0\,, \end{cases}$$

then Lemma 3.5 implies

$$\begin{aligned} \|\phi_{k,h}\|_{\frac{1+\alpha}{2},1+\alpha} &\le C\,\|(H_p(t,x,Du_{n_h}) - H_p(t,x,Du_{n_k}))\cdot D\phi_{n_h}\|_\infty \\ &\le C\,\|Du_{n_h} - Du_{n_k}\|_\infty \le \omega(k,h)\,, \end{aligned}$$

where $\omega(k,h) \to 0$ when $k,h \to \infty$.

Using (3.30) with (m_{n_k}, ϕ_{n_k}) and (m_{n_h}, ϕ_{n_h}), for $k,h \in \mathbb{N}$, $s = t_0$, subtracting the two equalities, we get

$$\sup_{t\in[t_0,T]} \|m_{n_k}(t) - m_{n_h}(t)\|_{-(1+\alpha),D} \le \omega(k,h) \implies \exists m \in X \text{ s.t. } m_{n_k} \to m \text{ in } X\,,$$

hence concluding the compactness part. The continuity is an easy consequence of the previous arguments.

In particular, we apply Schauder's theorem and obtain a classical solution of the problem (3.25). The estimate (3.26) follows from the above estimates for (3.28) and (3.29).

We skip the uniqueness part, which is a standard argument, see, e.g., *Proposition 3.3* of [194].

Step 2. Let (u_1, m_1) and (u_2, m_2) be two classical solutions of (3.25), with $m_1(t_0) = m_{01}$, $m_2(t_0) = m_{02}$. We take ϕ as the solution of (3.31) related to u_1, with $\psi = 0$, and we note that ϕ is also a good test function for the equation of m_2, since it satisfies, for $\psi = (H_p(t,x,Du_2) - H_p(t,x,Du_1))\cdot D\phi \in L^\infty$,

$$\begin{cases} -\phi_t - \mathrm{tr}(a(x)D^2\phi) + H_p(t,x,Du_2)\cdot D\phi = \psi\,, \\ \phi(t) = \xi\,, \\ \phi_{|\partial\Omega} = 0\,, \end{cases}$$

Then, subtracting the weak formulations of m_1 and m_2 related to ψ, we find

$$\int_\Omega \xi(x)(m_1(t)-m_2(t))\,dx = \tag{3.32}$$

$$\int_{t_0}^t\int_\Omega (H_p(t,x,Du_1)-H_p(t,x,Du_2))D\phi\, m_2(s,x)dx\,ds + \langle \phi(0,\cdot), m_{01}-m_{02}\rangle\,. \tag{3.33}$$

By Lipschitz continuity of both ϕ and H_p, with respect to p, we get

$$\int_\Omega \xi(x)(m_1(t)-m_2(t))\,dx \le C\int_{t_0}^t\int_\Omega |Du_1-Du_2|m_2(s,x)dx\,ds + C\mathbf{d}_1(m_{01},m_{02})\,,$$

We want to estimate the first term in the right-hand side with Young's inequality. To this end, we consider the quantities

$$\tilde c = \int_{t_0}^t\int_\Omega m_2(s,x)\,dxds\,, \qquad \tilde m_2(s,x) = \frac{m_2(s,x)}{\tilde c}\,.$$

Then $\tilde m_2$ is a probability measure in $[0,t]\times\Omega$, and Young's inequality implies that

$$\int_{t_0}^t\int_\Omega |Du_1-Du_2|m_2(s,x)dx\,ds = \tilde c\int_{t_0}^t\int_\Omega |Du_1-Du_2|\tilde m_2(s,x)dx\,ds$$
$$\le \tilde c\left(\int_{t_0}^t\int_\Omega |Du_1-Du_2|^2\tilde m_2(s,x)dx\right)^{\frac12} \le C\left(\int_{t_0}^t\int_\Omega |Du_1-Du_2|^2 m_2(s,x)dx\right)^{\frac12},$$

since $\tilde c$ is bounded thanks to the L^p bound of m_2.

Using the Lasry-Lions monotonicity argument (see *Lemma 3.1.2* of [41]), we have

$$\int_{t_0}^T\int_\Omega |Du_1-Du_2|^2(m_1(t,dx)+m_2(t,dx))dt \le$$
$$\le C\int_\Omega (u_1(0,x)-u_2(0,x))(m_{01}(dx)-m_{02}(dx))$$
$$\le C\|u_1-u_2\|_{\frac{1+\alpha}{2},1+\alpha}\mathbf{d}_1(m_{01},m_{02})\,.$$

Hence, we obtain

$$\int_\Omega \xi(x)(m_1(t)-m_2(t))\,dx$$
$$\le C\left(\|u_1-u_2\|^{\frac12}_{\frac{1+\alpha}{2},1+\alpha}\mathbf{d}_1(m_{01},m_{02})^{\frac12} + \mathbf{d}_1(m_{01},m_{02})\right),$$

and finally, taking the sup over the ξ with $\|\xi\|_{-(1+\alpha),D}\le 1$ and over $t\in[0,T]$,

$$\sup_{t\in[0,T]} \|m_1(t)-m_2(t)\|_{-(1+\alpha),D} \tag{3.34}$$

$$\leq C\left(\|u_1-u_2\|^{\frac{1}{2}}_{\frac{1+\alpha}{2},1+\alpha}\mathbf{d}_1(m_{01},m_{02})^{\frac{1}{2}}+\mathbf{d}_1(m_{01},m_{02})\right). \tag{3.35}$$

Let us now call $u:=u_1-u_2$, then u solves:

$$\begin{cases} -u_t-\mathrm{tr}(a(x)D^2u)+V(t,x)Du=f(t,x)\,,\\ u(T)=g(x)\,, \qquad u_{|\partial\Omega}=0\,, \end{cases}$$

where

$$\begin{aligned} V(t,x)&=\int_0^1 H_p(t,x,\lambda Du_1(t,x)+(1-\lambda)Du_2(t,x)d\lambda;\\ f(t,x)&=\int_0^1\int_\Omega \frac{\delta F}{\delta m}(t,x,m_\lambda(t),y)(m_1(t,dy)-m_2(t,dy))d\lambda;\\ g(x)&=\int_0^1\int_\Omega \frac{\delta G}{\delta m}(x,m_\lambda(T),y)(m_1(T,dy)-m_2(T,dy))d\lambda\,, \end{aligned}$$

where m_λ is defined as follows

$$m_\lambda(\cdot):=\lambda m_1(\cdot)+(1-\lambda)m_2(\cdot)\,.$$

As to apply Lemma 3.5, we estimate the regularity of the data:

$$\begin{aligned} &\sup_{t\in[0,T]}\|f(t,\cdot)\|_\alpha\\ &\qquad\leq \sup_{t\in[0,T]}\int_0^1\left\|\frac{\delta F}{\delta m}(t,\cdot,m_\lambda(t),\cdot)\right\|_{\alpha,(1+\alpha,D)} d\lambda\,\|m_1(t)-m_2(t)\|_{-(1+\alpha),D}\\ &\qquad\leq C\sup_{t\in[0,T]}\|m_1(t)-m_2(t)\|_{-(1+\alpha),D}\,, \end{aligned}$$

analogously

$$\|g(\cdot)\|_{2+\alpha}\leq C\sup_{t\in[0,T]}\|m_1(t)-m_2(t)\|_{-(1+\alpha),D}\,. \tag{3.36}$$

So, Eq. (3.15) implies

$$\|u_1-u_2\|_{1,2+\alpha}\leq C\sup_{t\in[0,T]}\|m_1(t)-m_2(t)\|_{-(1+\alpha),D}\,. \tag{3.37}$$

Coming back to (3.34), we have

$$
\begin{aligned}
&\sup_{t\in[0,T]} \|m_1(t)-m_2(t)\|_{-(1+\alpha),D} \\
&\qquad \le C\left(\left(\sup_{t\in[0,T]} \|m_1(t)-m_2(t)\|_{-(1+\alpha),D}\right)^{\frac{1}{2}} \mathbf{d}_1(m_{01},m_{02})^{\frac{1}{2}} + \mathbf{d}_1(m_{01},m_{02})\right),
\end{aligned}
$$

hence, by a generalized Young's inequality:

$$
\sup_{t\in[0,T]} \|m_1(t)-m_2(t)\|_{-(1+\alpha),D} \le C\mathbf{d}_1(m_{01},m_{02})\,. \tag{3.38}
$$

Plugging this estimate in (3.37), we finally obtain

$$
\|u_1-u_2\|_{1,2+\alpha} \le C\mathbf{d}_1(m_{01},m_{02})\,.
$$

For the L^p inequality, we consider $m := m_1 - m_2$. Then m solves the equation

$$
\begin{cases}
m_t - \sum_{i,j} \partial^2_{ij}(a_{ij}(x)m) - \operatorname{div}(m(H_p(t,x,Du_1)) = \\
\qquad\qquad \operatorname{div}(m_2(H_p(t,x,Du_2)-H_p(t,x,Du_1)))\,, \\
m(t_0) = m_{01} - m_{02}\,, \\
[m_1-m_2]_{|\partial\Omega} = 0\,,
\end{cases}
$$

i.e. m is a solution of (3.16) with $f = \operatorname{div}(m_2(H_p(t,x,Du_2)-H_p(t,x,Du_1)))$, $\mu_0 = m_{01}-m_{02}$, $b = H_p(t,x,Du_1)$. Then estimates (3.19) imply

$$
\|m_1-m_2\|_{L^p(Q_T)} \le C\left(\|\mu_0\|_{-(1+\alpha)} + \|f\|_{L^1}\right).
$$

We estimate the right-hand side term. As regards μ_0 we have

$$
\|\mu_0\|_{-(1+\alpha)} = \sup_{\|\phi\|_{1+\alpha}\le 1} \int_\Omega \phi(x)(m_{01}-m_{02})(dx) \le C\mathbf{d}_1(m_{01},m_{02})\,.
$$

For the f term we argue in the following way:

$$
\begin{aligned}
\|f\|_{L^1} &= \int_0^T \sup_{\|\phi\|_{W^{1,\infty}}\le 1} \left(\int_\Omega H_p(x,Du_2)-H_p(x,Du_1)D\phi\, m_2(t,dx)\right) dt \\
&\le C\|u_1-u_2\|_{\frac{1+\alpha}{2},1+\alpha} \le C\mathbf{d}_1(m_{01},m_{02})\,,
\end{aligned}
$$

which allows us to conclude. □

Previous proposition (3.8) allows to state that

$$\begin{aligned} &\sup_{t\in[0,T]}\sup_{m\in\mathcal{P}^{sub}(\Omega)} \|U(t,\cdot,m)\|_{2+\alpha}\le C\,,\\ &\sup_{t\in[0,T]}\sup_{m_1\neq m_2}\left[(\mathbf{d}_1(m_1,m_2))^{-1}\|U(t,\cdot,m_1)-U(t,\cdot,m_2)\|_{2+\alpha}\right]\le C\,, \end{aligned}\tag{3.39}$$

which are two initial regularity results for the function U.

3.9 Linearized system and differentiability of U with respect to the measure

This section is devoted to the study of the following *linearized MFG system*:

$$\begin{cases} -z_t-\mathrm{tr}(a(x)D^2z)+H_p(t,x,Du)Dz=\dfrac{\delta F}{\delta m}(t,x,m(t))(\rho(t))+h(t,x)\,,\\ \rho_t-\displaystyle\sum_{ij}\partial^2_{ij}(a_{ij}(x)\rho)-\mathrm{div}(\rho(H_p(t,x,Du)))+\\ \qquad\qquad\qquad\qquad -\mathrm{div}(mH_{pp}(t,x,Du)Dz+c)=0,\\ z(T,x)=\dfrac{\delta G}{\delta m}(x,m(T))(\rho(T))+z_T(x)\,,\qquad \rho(t_0)=\rho_0\,,\\ z_{|\partial\Omega}=0\,,\qquad \rho_{|\partial\Omega}=0\,, \end{cases}\tag{3.40}$$

where $z_T\in\mathcal{C}^{2+\alpha,D}$, $\rho_0\in\mathcal{C}^{-(1+\alpha)}$, $h\in\mathcal{C}^{0,\alpha}([t_0,T]\times\Omega)$, $c\in L^1([t_0,T]\times\Omega)$, and where we define for F (and for G)

$$\frac{\delta F}{\delta m}(t,x,m(t))(\rho(t)):=\left\langle\frac{\delta F}{\delta m}(t,x,m(t),\cdot),\rho(t)\right\rangle,$$

where the duality is between $\mathcal{C}^{-(1+\alpha),D}$ and $\mathcal{C}^{(1+\alpha),D}$.

The study of this system plays a crucial role in proving the $\mathcal{C}^1$ character of U in terms of m. In particular, if we define the couple (v,μ) as the solution of (3.40) with $h=c=z_T=0$ and $\mu(t_0)=\mu_0$, we obtain

$$v(t_0,x)=\left\langle\frac{\delta U}{\delta m}(t_0,x,m_0,\cdot),\mu_0\right\rangle.\tag{3.41}$$

Let us start giving a suitable definition of solution for the previous system.

Definition 3.10. *A couple (z,ρ) is a solution of* (3.40) *if $z \in \mathcal{C}^{1,2+\alpha}$, is a classical solution of the linear HJB equation and $\rho \in \mathcal{C}([0,T];\mathcal{C}^{-(1+\alpha),D}(\Omega)) \cap L^1(Q_T)$ solves the Fokker-Planck equation, accordingly with Definition 3.8.*

Let us underline that $c \in L^1([t_0,T]\times\Omega) \implies \operatorname{div}(c) \in L^1([t_0,T];W^{-1,\infty}(\Omega))$. Therefore the well-posedness of the Fokker-Planck equation is included in Proposition 3.6.

The following Proposition states existence, uniqueness and regularity for the problem defined by eq. (3.40). The proof relies on results previously obtained for the Fokker-Planck equation, and it proceeds as the Neumann case, see *Propositions 5.8 and 5.11* of [194]), therefore we omit it.

Proposition 3.9. *Let hypotheses 1 hold for a certain $0<\alpha<1$, and let $\rho_0 \in \mathcal{C}^{-1}$. Then there exists a unique solution $(z,\rho) \in \mathcal{C}^{1,2+\alpha} \times \Big(\mathcal{C}([0,T];\mathcal{C}^{-(1+\alpha),D}(\Omega)) \cap L^1(Q_T)\Big)$ for the system* (3.40)*, satisfying*

$$\|z\|_{1,2+\alpha} + \sup_t \|\rho(t)\|_{-(1+\alpha),D} + \|\rho\|_{L^p} \le CM\,, \tag{3.42}$$

where C depends on H, p is defined as in section 3.7.1, and M is given by

$$M := \|z_T\|_{2+\alpha} + \|\rho_0\|_{-(1+\alpha)} + \|h\|_{0,\alpha} + \|c\|_{L^1}\,. \tag{3.43}$$

Moreover, the solution (v,μ) related to $h=c=z_T=0$ satisfies

$$\|v\|_{1,2+\alpha} + \sup_{t\in[0,T]} \|\mu(t)\|_{-(2+\alpha),D} + \|\mu\|_{-\left(\frac{\alpha}{2},\alpha\right)} \le C\|\mu_0\|_{-(2+\alpha)}\,. \tag{3.44}$$

Throughout the rest of the chapter, we will denote with (v,μ) a solution to the system (3.40), with $h=c=z_T=0$ and $\mu_0 := \rho_0$. We will refer to this system as the *pure linearized system.* Instead, the general system (3.40), with solution (z,ρ), will be called the *general linearized system.*

To prove $U \in \mathcal{C}^1$ as well as the related representation formula (3.41), we have to prove that the pure linearized system has a fundamental solution, which is the content of the next proposition.

Proposition 3.10. *Let hypotheses 1 hold. Then there exists a function $K:[0,T]\times\Omega\times\mathcal{P}^{sub}(\Omega)\times\Omega \to \mathbb{R}$ such that, for any solution (v,μ), with initial data (t_0,m_0,μ_0), we have*

$$v(t_0,x) = {}_{-(1+\alpha)}\langle \mu_0, K(t_0,x,m_0,\cdot)\rangle_{1+\alpha} \tag{3.45}$$

Moreover, K is twice differentiable with respect to both x and y, and it holds

$$\sup_{(t,m)\in[0,T]\times\mathcal{P}^{sub}(\Omega)} \|K(t,\cdot,m,\cdot)\|_{2+\alpha,2+\alpha} \leq C\,, \tag{3.46}$$

Proof. Let $\mu_0 = \delta_y$ be the Dirac function at $y \in \Omega$. We define $K(t_0,x,m_0,y) = v(t_0,x;\delta_y)$, where $v(\cdot,\cdot;\mu_0)$ indicates the function v related to the pure linearized system with initial data μ_0. Exploiting (3.42), we have that K is twice differentiable w.r.t. x, and it holds

$$\|K(t_0,\cdot,m_0,y)\|_{2+\alpha} \leq C\|\delta_y\|_{-(1+\alpha)} = C$$

Furthermore, the linearity character of (3.40) implies

$$\frac{K(t_0,x,m_0,y+he_i)-K(t_0,x,m_0,y)}{h} = v(t_0,x;\Delta_h^i\delta_y)\,,$$

where we use the notation $\Delta_h^i\delta_y = \frac{1}{h}(\delta_{y+he_i}-\delta_y)$. The linear character of the pure linearized system directly implies that the solution (v,μ) is stable with respect to the initial condition μ_0, allowing to pass to the limit and obtain

$$\frac{\partial K}{\partial y_i}(t_0,x,m_0,y) = v(t_0,x;-\partial_{y_i}\delta_y)\,,$$

with the derivative that has to be intended in a distributional sense. As to prove the existence and second derivatives' bounds, we consider the incremental ratio

$$R_{i,j}^h(x,y) := \frac{\partial_{y_i}K(t_0,x,m_0,y+he_j)-\partial_{y_i}K(t_0,x,m_0,y)}{h}\,. \tag{3.47}$$

Hence, estimate (3.44) together with Lagrange Theorem imply that, for $|l|\leq 2$, we have

$$\begin{aligned}
&\left|D_x^l R_{i,j}^h(x,y) - D_x^l R_{i,j}^k(x,y)\right| = \left|D_x^l v\left(t_0,x;\Delta_h^j(-\partial_{y_i}\delta_y)-\Delta_k^j(-\partial_{y_i}\delta_y)\right)\right| \\
&\leq C\left\|\Delta_h^j(-\partial_{y_i}\delta_y)-\Delta_k^j(-\partial_{y_i}\delta_y)\right\|_{-(2+\alpha)} \\
&= \sup_{\|\phi\|_{2+\alpha}\leq 1}\left(\frac{\partial_{y_i}\phi(y+he_j)-\partial_{y_i}\phi(y)}{h} - \frac{\partial_{y_i}\phi(y+ke_j)-\partial_{y_i}\phi(y)}{k}\right) \\
&= \sup_{\|\phi\|_{2+\alpha}\leq 1}\left(\partial^2_{y_iy_j}\phi(y_{\phi,h})-\partial^2_{y_iy_j}\phi(y_{\phi,k})\right) \leq \sup_{\|\phi\|_{2+\alpha}\leq 1}|y_{\phi,h}-y_{\phi,k}|^\alpha \leq |h|^\alpha+|k|^\alpha,
\end{aligned}$$

for a certain $y_{\phi,h}$, resp. $y_{\phi,k}$ in the line segment between y and $y+he_j$ (resp. $y+ke_j$), and where we have used the same notation as seen above for $\Delta_h^j(-\partial_{y_i}\delta_y)$.

The latter proves that (3.47), together with its first and second derivative w.r.t x, are Cauchy sequences in h, implying that $D_x^l\frac{\delta U}{\delta m}$ is twice differentiable w.r.t. y and for all $|l|\leq 2$.

As to conclude estimate (3.46), we prove the estimate for the second order derivatives w.r.t. y, the first derivative being simpler. By estimate (3.44) on v, if $y, y' \in \Omega$, then

$$\begin{aligned}\left\| R^h_{i,j}(\cdot,y) - R^h_{i,j}(\cdot,y')\right\|_{2+\alpha} &= \left\| v\big(t_0,\cdot;\Delta^j_h(-\partial_{y_i}\delta_y + \partial_{y_i}\delta_{y'})\big)\right\|_{2+\alpha} \\ &\le C\left\|\Delta^j_h(-\partial_{y_i}\delta_y + \partial_{y_i}\delta_{y'})\right\|_{-(2+\alpha)}.\end{aligned}$$

Passing to the limit for $h \to 0$, we have

$$\Delta^j_h(-\partial_{y_i}\delta_y + \partial_{y_i}\delta_{y'}) \longrightarrow \partial_{y_j}\partial_{y_i}\delta_y - \partial_{y_j}\partial_{y_i}\delta_{y'}\,, \text{ in } \mathcal{C}^{-(2+\alpha)}\,.$$

Then, by Ascoli-Arzelà and previously obtained convergence result for $R^h_{i,j}$, we have

$$\begin{aligned}&\left\|\partial^2_{y_i y_j}\left(K(t_0,\cdot,m_0,y) - K(t,\cdot,m_0,y')\right)\right\|_{2+\alpha} \\ &\qquad \le C\left\|\partial^2_{y_i y_j}(\delta_y - \delta_{y'})\right\|_{-(2+\alpha)} \le C|y-y'|^\alpha\,,\end{aligned}$$

which proves (3.46). The representation formula (3.45) is an immediate consequence of the linear character of the equation and of the density of the Dirac functions generated set, hence concluding the proof. □

We are now ready to consider the main topic of this section: we want to prove that the function K is actually the derivative of U with respect to the measure. Let us underline that the differentiability with respect to the measure m will be the key for proving U is indeed a classical solution of the Master Equation (3.7).

Theorem 3.11. *Suppose hypotheses 1 hold. Then the function U defined by* (3.9) *is differentiable with respect to m, with the derivative given by*

$$\frac{\delta U}{\delta m}(t,x,m,y) = K(t,x,m,y)\,.$$

Proof. Within the present proof, given two functions a_1, a_2, we define $a_{1+\tau} := \tau a_2 + (1-\tau)a_1$, for $\tau \in [0,1]$.

We will prove a more general fact, the representation formula for U being then a direct consequence of it.

In particular, if (u_1, m_1) and (u_2, m_2) are two solutions of (3.25), with initial conditions (t_0, m_{01}) and (t_0, m_{02}), and (v, μ) is the solution of the pure linearized system related to (u_2, m_2), with initial condition $(t_0, m_{01} - m_{02})$, then a sort of first-order Taylor expansion of

U with respect to m holds, namely:

$$\|u_1-u_2-v\|_{1,2+\alpha}+\sup_{t\in[0,T]}\|m_1(t)-m_2(t)-\mu(t)\|_{-(1+\alpha),D}\le C\mathbf{d}_1(m_{01},m_{02})^2\,. \tag{3.48}$$

As to prove above *expansion*, let us start defining $(z,\rho)=(u_1-u_2-v,m_1-m_2-\mu)$. Then (z,ρ) satisfies the general linearized system (3.40) related to (u_2,m_2), with data $h=h_1+h_2$, $c=c_1+c_2$ and z_T given by:

$$\begin{aligned}
h_1 &= -\int_0^1 (H_p(t,x,Du_{1+s})-H_p(t,x,Du_2))\cdot D(u_1-u_2)\,ds\,,\\
h_2 &= \int_0^1\int_\Omega\left(\frac{\delta F}{\delta m}(t,x,m_{1+s}(t),y)-\frac{\delta F}{\delta m}(t,x,m_2(t),y)\right)(m_1(t)-m_2(t))(dy)ds,\\
c_1(t) &= (m_1(t)-m_2(t))H_{pp}(t,x,Du_2)(Du_1-Du_2)\,,\\
c_2(t) &= m_1\int_0^1 (H_{pp}(t,x,Du_{1+s})-H_{pp}(t,x,Du_2))\,(Du_1-Du_2)\,ds\,,\\
z_T &= \int_0^1\int_\Omega\left(\frac{\delta G}{\delta m}(x,m_{1+s}(T),y)-\frac{\delta G}{\delta m}(x,m_2(T),y)\right)(m_1(T)-m_2(T))(dy)ds\,.
\end{aligned}$$

Applying (3.42) one has

$$\|z\|_{1,2+\alpha}+\sup_{t\in[0,T]}\|\rho(t)\|_{-(1+\alpha),D}\le C\left(\|h\|_{0,\alpha}+\|c\|_{L^1}+\|z_T\|_{2+\alpha}\right)\,. \tag{3.49}$$

We want to estimate the right-hand side term to obtain the desired Taylor expansion. As regards h, exploiting eq. (3.27) and Hölder norm properties, we have:

$$\begin{aligned}
\|h_1\|_{0,\alpha} &= \left\|\int_0^1\int_0^1 s\,\langle H_{pp}(t,x,Du_{1+rs})\,(Du_1-Du_2)\,,\,(Du_1-Du_2)\rangle\,drds\right\|_{0,\alpha}\\
&\le C\mathbf{d}_1(m_{01},m_{02})^2\,,
\end{aligned}$$

Analogously, again by eq. (3.27) and exploiting regularity of both F and G, the same estimate also holds for h_2 and z_T. For the function c, we can write

$$\begin{aligned}
\|c_1\|_{L^1} = \int_0^T\int_\Omega & H_{pp}(t,x,Du_2)(Du_1-Du_2)(m_1(t,dx)-m_2(t,dx))\,dt\\
&\le C\|u_1-u_2\|_{1,2+\alpha}\|m_1(t)-m_2(t)\|\le C\mathbf{d}_1(m_{01},m_{02})^2\,,
\end{aligned}$$

then, proceeding as above, we have

$$
\begin{aligned}
&\|c_2\|_{L^1} = \\
&\quad \int_0^1 \int_0^T \int_\Omega (H_{pp}(t,x,Du_{1+s}) - H_{pp}(t,x,Du_2))\,(Du_1 - Du_2) m_1(t,dx)\,dtds \\
&\quad \le C\,\|Du_1 - Du_2\|_\infty^2 \le C\mathbf{d}_1(m_0^1, m_0^2)^2 .
\end{aligned}
$$

Substituting these estimates in (3.49), we obtain (3.48). Using the representation (3.45) for v, we get

$$
\begin{aligned}
\left\| U(t_0,\cdot,m_{01}) - U(t_0,\cdot,m_{02}) - \int_\Omega K(t_0,\cdot,m_{02},y)(m_{01}-m_{02})(dy) \right\|_\infty \\
\le C\mathbf{d}_1(m_{01},m_{02})^2 .
\end{aligned}
$$

As a straightforward consequence, we have that U is differentiable with respect to m and

$$
\frac{\delta U}{\delta m}(t,x,m,y) = K(t,x,m,y)\,,
$$

hence concluding the proof. □

Using (3.46) we also obtain the following strong regularity estimate for $\frac{\delta U}{\delta m}$:

$$
\sup_t \left\| \frac{\delta U}{\delta m}(t,\cdot,m,\cdot) \right\|_{2+\alpha,2+\alpha} \le C\,. \tag{3.50}
$$

This gives sense to the Master Equation (3.7), since the quantity D_yD_mU is now well-defined. However, to apply Theorem 3.4 we still need to prove the continuity of D_yD_mU in the measure variable.

In the next result we prove a Lipschitz bound for $\frac{\delta U}{\delta m}$, with respect to the measure m. This bound plays a key role, since not only it implies the continuity of D_yD_mU with respect to m, but it is also used in the next section to apply Lemma 3.14.

Theorem 3.12. *Let Assumptions 1 hold. Then the derivative of the solution of the Master Equation $\frac{\delta U}{\delta m}$ is Lipschitz continuous with respect to the measure m:*

$$
\sup_{t\in[0,T]} \sup_{m_1\neq m_2} (\mathbf{d}_1(m_1,m_2))^{-1} \left\| \frac{\delta U}{\delta m}(t,\cdot,m_1,\cdot) - \frac{\delta U}{\delta m}(t,\cdot,m_2,\cdot) \right\|_{2+\alpha,2+\alpha} \le C \tag{3.51}
$$

where C depends on n, F, G, H and T.

Proof. We consider, for $i=1,2$, the solution (v_i,μ_i) of the linearized system (3.40) related to (u_i,m_i).

To avoid too heavy notations, we take $t_0=0$ and we define

$$\begin{aligned} &H_i'(t,x):=H_p(t,x,Du_i(t,x))\,, && H_i''(t,x)=H_{pp}(t,x,Du_i(t,x))\,,\\ &F'(t,x,m,\mu)=\int_\Omega \frac{\delta F}{\delta m}(t,x,m,y)\,\mu(dy)\,, && G'(x,m,\mu)=\int_\Omega \frac{\delta G}{\delta m}(x,m,y)\,\mu(dy)\,. \end{aligned}$$

Then the couple $(z,\rho):=(v_1-v_2,\mu_1-\mu_2)$ satisfies the following linear system:

$$\begin{cases} -z_t-\mathrm{tr}(a(x)D^2z)+H_1'\cdot Dz=F'(t,x,m_1(t),\rho(t))+h\,,\\ \rho_t-\sum\limits_{i,j}\partial^2_{ij}(a_{ij}(x)\rho)-\mathrm{div}(\rho H_1')-\mathrm{div}(m_1H_1''Dz+c)=0\,,\\ z(T,x)=G'(x,m_1(T),\rho(T))+z_T\,, \qquad \rho(t_0)=0\,,\\ z_{|\partial\Omega}=0\,, \qquad \rho_{|\partial\Omega}=0\,, \end{cases}$$

where

$$\begin{aligned} &h(t,x)=h_1(t,x)+h_2(t,x)\,,\\ &h_1(t,x)=F'(t,x,m_1(t),\mu_2(t))-F'(t,x,m_2(t),\mu_2(t))\,,\\ &h_2(t,x)=\big(H_1'(t,x)-H_2'(t,x)\big)\cdot Dv_2(t,x)\,,\\ &c(t,x)=\mu_2(t)\big(H_1'-H_2'\big)(t,x)+\big[(m_1H_1''-m_2H_2'')\big](t,x)\,,\\ &z_T(x)=G'(x,m_1(T),\mu_2(T))-G'(x,m_2(T),\mu_2(T))\,. \end{aligned}$$

Applying (3.42) we obtain this estimate on z:

$$\|z\|_{1,2+\alpha}\le C\left(\|z_T\|_{2+\alpha}+\|h\|_{0,\alpha}+\|c\|_{L^1}\right).$$

Now we estimate the terms in the right-hand side.

The term with z_T, thanks to (3.44) and the hypothesis *(v)* of 1, is immediately estimated:

$$\|z_T\|_{2+\alpha}\le\left\|\frac{\delta G}{\delta m}(\cdot,m_1(T),\cdot)-\frac{\delta G}{\delta m}(\cdot,m_2(T),\cdot)\right\|_{2+\alpha,2+\alpha}\|\mu_2(T)\|_{-(2+\alpha),D}\le C\mathbf{d}_1(m_{01},m_{02})\|\mu_0\|_{-(2+\alpha)}\,.$$

As regards the space estimate for h, we have

$$\|h(t,\cdot)\|_\alpha\le\|F'(\cdot,m_1(t),\mu_2(t))-F'(\cdot,m_2(t),\mu_2(t))\|_\alpha+\big\|(H_1'-H_2')(t,\cdot)Dv_2(t,\cdot)\big\|_\alpha\,.$$

The first term is bounded as z_T :

$$\|F'(\cdot,m_1(t),\mu_2(t))-F'(\cdot,m_2(t),\mu_2(t))\|_\alpha \le C\mathbf{d}_1(m_{01},m_{02})\|\mu_0\|_{-(2+\alpha)}\,.$$

The second term, using (3.27) and (3.44), can be estimated in this way:

$$\|(H_1'-H_2')(t,\cdot)Dv_2(t,\cdot)\|_\alpha \le C\|(u_1-u_2)(t)\|_{1+\alpha}\|v_2(t)\|_{1+\alpha} \le C\mathbf{d}_1(m_{01},m_{02})\|\mu_0\|_{-(2+\alpha)}\,.$$

In summary,

$$\|h\|_{0,\alpha} = \sup_{t\in[0,T]}\|h(t,\cdot)\|_\alpha \le C\mathbf{d}_1(m_{01},m_{02})\|\mu_0\|_{-(2+\alpha)}\,.$$

Finally, we estimate $\|c\|_{L^1}$. We have

$$\begin{aligned}\|c\|_{L^1} &= \int_0^T\int_\Omega (H_1'-H_2')(t,x)\,\mu_2(t,dx)\,dt + \int_0^T\int_\Omega H_1''(t,x)Dv_2(t,x)(m_1(t)-m_2(t))(dx)\,dt\\ &+\int_0^T\int_\Omega (H_1'-H_2')(t,x)\,Dv_2(t,x)\,m_2(t,dx)\,dt \le C\|u_1-u_2\|_{\frac{1+\alpha}{2},1+\alpha}\|\mu_2\|_{-(\frac{\alpha}{2},\alpha)}\\ &+C\|u_1\|_{\frac{1+\alpha}{2},1+\alpha}\|v_2\|_{\frac{1+\alpha}{2},1+\alpha}\|m_1-m_2\|_{L^1}+C\|u_1-u_2\|_{\frac{1+\alpha}{2},1+\alpha}\|v_2\|_{\frac{1+\alpha}{2},1+\alpha}\,.\end{aligned}$$

The first term in the right-hand side, thanks to (3.27) and (3.44), is bounded by

$$C\|u_1-u_2\|_{\frac{1+\alpha}{2},1+\alpha}\|\mu_2\|_{-(\frac{\alpha}{2},\alpha)} \le C\mathbf{d}_1(m_{01},m_{02})\|\mu_0\|_{-(2+\alpha)}\,.$$

The second and the third term are estimated in the same way, using (3.26), (3.27) and (3.44). Then

$$\|c\|_{L^1} \le C\mathbf{d}_1(m_{01},m_{02})\|\mu_0\|_{-(2+\alpha)}\,.$$

Putting together all these estimates, we finally obtain:

$$\|z\|_{1,2+\alpha} \le C\mathbf{d}_1(m_{01},m_{02})\|\mu_0\|_{-(2+\alpha)}\,.$$

Since

$$z(t_0,x) = \int_\Omega\left(\frac{\delta U}{\delta m}(t_0,x,m_1,y)-\frac{\delta U}{\delta m}(t_0,x,m_2,y)\right)\mu_0(dy)\,,$$

we have proved (3.51). □

We stress the fact that in the bound (3.51) a $2+\alpha$ norm in the last variable is strongly required, to make this estimate valid also for the quantity D_yD_mU.

To conclude, we are still left proving the boundary condition for $\frac{\delta U}{\delta m}$ in (3.7) to make true all the hypotheses needed to apply Theorem 3.4. The latter will be the last result of this section.

Corollary 3.13. *If hypotheses 1 hold true, then we have the following boundary conditions for U:*

$$\frac{\delta U}{\delta m}(t,x,m,y)=0\,, \qquad \forall x\in\partial\Omega, y\in\Omega, t\in[0,T], m\in\mathcal{P}^{sub}(\Omega)\,,$$
$$\frac{\delta U}{\delta m}(t,x,m,y)=0\,, \qquad \forall x\in\Omega, y\in\partial\Omega, t\in[0,T], m\in\mathcal{P}^{sub}(\Omega)\,.$$

Proof. The proof of the first boundary condition is trivial, since $\frac{\delta U}{\delta m}(t_0,x,m_0,y)=v(t_0,x;\delta_y)$ and v satisfies a Dirichlet boundary condition.

For the second condition, let us consider $y\in\partial\Omega$. To prove that $v(t_0,x;\delta_y)=0$, it is enough to check that $v=0$ solves

$$\begin{cases} -v_t-\mathrm{tr}(a(x)D^2v)+H_p(t,x,Du)\cdot Dv=\dfrac{\delta F}{\delta m}(t,x,m(t))(\mu(t))\,,\\ v(T,x)=\dfrac{\delta G}{\delta m}(x,m(T))(\mu(T))\,,\\ v_{|\partial\Omega}=0\,,\end{cases} \tag{3.52}$$

where μ is the unique solution, in the sense of Definition 3.8, of

$$\begin{cases} \mu_t-\sum\limits_{ij}\partial^2_{ij}(a_{ij}(x)\mu)-\mathrm{div}(\mu H_p(t,x,Du))=0\,,\\ \mu(t_0)=\mu_0\,,\\ \mu_{|\partial\Omega}=0\,.\end{cases}$$

In this way we have $mH_{pp}(t,x,Du)Dv=0$, and so the couple (v,μ) turns out to be a solution of

$$\begin{cases} -v_t-\mathrm{tr}(a(x)D^2v)+H_p(t,x,Du)Dv=\dfrac{\delta F}{\delta m}(t,x,m(t))(\rho(t))\,,\\ \mu_t-\sum\limits_{ij}\partial^2_{ij}(a_{ij}(x)\mu)-\mathrm{div}\big(\mu H_p(t,x,Du)+mH_{pp}(t,x,Du)Dv\big)=0,\\ v(T,x)=\dfrac{\delta G}{\delta m}(x,m(T))(\mu(T))\,, \qquad \mu(t_0)=\mu_0\,,\\ v_{|\partial\Omega}=0\,, \qquad \mu_{|\partial\Omega}=0\,,\end{cases}$$

which is exactly the pure linearized system.

Due to the linearity character of (3.52), we only need to prove that

$$\frac{\delta F}{\delta m}(t,x,m(t))(\mu(t))=\frac{\delta G}{\delta m}(x,m(T))(\mu(T))=0\,.$$

Thanks to the hypotheses 1, both $\frac{\delta F}{\delta m}(x,m(t),\cdot)$ and $\frac{\delta G}{\delta m}(x,m(T),\cdot)$ satisfy a Dirichlet boundary condition, being also elements of $\mathcal{C}^{2+\alpha}$. Then, choosing $\phi(t,y)$ satisfying (3.17) with $\psi(t,y)=0$ and $\xi(y)=\frac{\delta F}{\delta m}(t,x,m(t),y)$, we have

$$\begin{aligned}\frac{\delta F}{\delta m}(t,x,m(t))(\mu(t)) &= \left\langle \mu(t), \frac{\delta F}{\delta m}(t,x,m(t),\cdot)\right\rangle \\ &= \langle \mu_0, \phi(0,\cdot)\rangle = \langle \delta_y, \phi(0,\cdot)\rangle = \phi(0,y) = 0\,,\end{aligned}$$

since ψ satisfies a Dirichlet boundary condition. The same computations hold for $\frac{\delta G}{\delta m}$, therefore we have:

$$\frac{\delta U}{\delta m}(t_0,x,m_0,y) = \left\langle \frac{\delta U}{\delta m}(t_0,x,m_0,\cdot),\delta_y\right\rangle = v(t_0,x) = 0\,.$$

□

3.11 Convergence of the Nash System

The aim of this section is twofold: proving that, for an integer $N \geq 2$, a classical solution $(v^{N,i})_{i\in\{1,\ldots,N\}}$ of the Nash system (3.5) converges in a suitable sense to the solution of the Master Equation (3.7), and to prove that also optimal trajectories converge.

3.11.1 Finite Dimensional Projections of U

The symmetrical structure of the problem suggests considering *suitable* finite dimensional projections of U, along the empirical distributions $m_{\mathbf{x}}^{N,i}$. Therefore, for $i \in \{1,\ldots,N\}$, we define:

$$u^{N,i}(t,\mathbf{x}) := U(t,x_i,m_{\mathbf{x}}^{N,i})\,, \tag{3.53}$$

where $\mathbf{x} = (x_1,\ldots,x_N) \in \Omega^N$ and $m_{\boldsymbol{x}}^{N,i}$ is defined as in (3.3). Exploiting the regularity of U, we already know that

$$u_i^N \in \mathcal{C}^{1,2+\alpha}$$

with respect to the couple (t,x_i).

In order to prove a regularity result for $(x_j)_{j\neq i}$, we need two technical results about the derivative of the solution of (3.7) with respect to m.

Theorem (3.12) implies the following lemma, which turns out to be crucial to prove the representation formula for the derivatives of u_i^N. Since the related proof is analogous to *Proposition 7.1, 7.4* of [41], we omit it.

Lemma 3.14. *Suppose hypotheses 1 are satisfied, then, if $m \in \mathcal{P}^{sub}(\Omega)$ and $\phi \in L^2(m,\mathbb{R}^d)$ is a bounded vector field, we have*

$$\left|\left|U(t,\cdot,(id+\phi)\sharp m) - U(t,\cdot,m) - \int_\Omega D_m U(t,\cdot,m,y)\cdot\phi(y)dm(y)\right|\right|_{1+\alpha} \le C||\phi||^2_{L^2(m)} \quad (3.54)$$

Now we can prove the following representation formula for the derivatives of u_i^N:

Proposition 3.15. *The following equations for the derivative of u_i^N hold for all $j \neq i$:*

$$D_{x_j} u_i^N(t,\boldsymbol{x}) = \frac{1}{N-1} D_m U(t,x_i,m_{\boldsymbol{x}}^{N,i},x_j)\,, \quad (3.55)$$

$$D^2_{x_i,x_j} u_i^N(t,\boldsymbol{x}) = \frac{1}{N-1} D_x D_m U(t,x_i,m_{\boldsymbol{x}}^{N,i},x_j)\,, \quad (3.56)$$

$$\left|D^2_{x_j,x_j} u_i^N(t,\boldsymbol{x}) - \frac{1}{N-1} D_y D_m U(t,x_i,m_{\boldsymbol{x}}^{N,i},x_j)\right| \le \frac{C}{N^2}\,, \quad (3.57)$$

where the last inequality holds a.e. $\boldsymbol{x} \in \Omega^N$.

Proof. We can limit ourselves to prove (3.55) and (3.57), since the second formula is a direct consequence of the first one. We consider $\mathbf{x} = (x_j)_{j\in\{1,\dots,N\}}$, $\mathbf{v} = (v_j)$ s.t. $\boldsymbol{x},\boldsymbol{v} \in \Omega^N$, with $x_j \neq x_k$ for any $j \neq k$, $v_i = 0$, i being fixed. By density of $\boldsymbol{x}$ with $x_j \neq x_k$ in Ω^N, it is sufficient to prove the first formula just in this case.

For $\varepsilon = \min_{j\neq k}|x_j - x_k|$, we consider a smooth vector field $\varphi : \Omega \to \mathbb{R}^d$ such that $\varphi(x) = v_j$ for $x \in B(x_j,\varepsilon/3)$ and $x + \varphi(x) \in \Omega$ for all $x \in \Omega$. Therefore, since $u_i^N(t,\mathbf{x}+\mathbf{v}) = U(t,x_i,(id+\varphi)\sharp m_{\mathbf{x}}^{N,i})$, Lemma 3.54 implies:

$$u_i^N(t,\mathbf{x}+\mathbf{v}) = u_i^N(t,\mathbf{x}) + \int_\Omega D_m U(t,x_i,m_{\mathbf{x}}^{N,i},y)\cdot\varphi(t)dm_{\mathbf{x}}^{N,i}(y) + o\left(||v||_{L^2(m_{\mathbf{x}}^{N,i}))}\right)\,,$$

exploiting the definition of $m_{\boldsymbol{x}}^{N,i}$, we compute the integral obtaining

$$u_i^N(t,\mathbf{x}+\mathbf{v}) = u_i^N(t,\mathbf{x}) + \frac{1}{N-1}\sum_{j\neq i} D_m U(t,x_i,m_{\boldsymbol{x}}^{N,i},x_j)\cdot v_j + o(|v|)\,,$$

concluding the proof of (3.55). Concerning the last inequality, we first show that $D^2_{x_j,x_j}u_i^N$ exists almost everywhere. Indeed

$$\begin{aligned}|D_{x_j}u_i^N(t,\mathbf{x}) - D_{x_j}u_i^N(t,\mathbf{y}))| &\le \frac{C}{N}|D_m U(t,x_i,m_{\mathbf{x}}^{N,i},x_j) - D_m U(t,x_i,m_{\mathbf{x}}^{N,i},y_j)|\\ &+\frac{C}{N}\left(|D_m U(t,x_i,m_{\mathbf{x}}^{N,i},y_j) - D_m U(t,x_i,m_{\mathbf{y}}^{N,i},y_j)|\right) \le \frac{C}{N}|\boldsymbol{x}-\boldsymbol{y}|\,,\end{aligned}$$

where the last inequality comes from (3.50) and (3.51). This implies that $D^2_{x_j,x_j}u_i^N$ exists a.e., and that it is also globally bounded by $\frac{C}{N}$. Considering $\mathbf{e}_{\mathbf{j}}^{\mathbf{k}} = (e_{jk}^1, ..., e_{jk}^N)$ with $e_{jk}^l = \delta_{lj} e_k$ as a basis over $(\mathbb{R}^d)^N$, we can write

$$\left| \frac{D_{x_j} u^{N,i}(t, \mathbf{x} + h\mathbf{e}_{\mathbf{j}}^{\mathbf{k}}) - D_{x_j} u^{N,i}(t, \mathbf{x})}{h} - \frac{1}{N-1} \partial_{y_k} D_m U(t, x_i, m_{\mathbf{x}}^{N,i}, x_j) \right|$$
$$\leq \frac{C}{N} \left| \frac{D_m U(t, x_i, m_{\mathbf{x}+h\mathbf{e}_{\mathbf{j}}^{\mathbf{k}}}^{N,i}, x_j + he_k) - D_m U(t, x_i, m_{\mathbf{x}}^{N,i}, x_j + he_k)}{h} \right| + \frac{C}{N}$$
$$\left| \frac{D_m U(t, x_i, m_{\mathbf{x}}^{N,i}, x_j + he_k) - D_m U(t, x_i, m_{\mathbf{x}}^{N,i}, x_j)}{h} - \partial_{y_k} D_m U(t, x_i, m_{\mathbf{x}}^{N,i}, x_j) \right|$$
$$\leq \frac{C}{N} \left| \partial_{y_k} D_m U(t, x_i, m_{\mathbf{x}}^{N,i}, x_h) - \partial_{y_k} D_m U(t, x_i, m_{\mathbf{x}}^{N,i}, x_j) \right|$$
$$+ \frac{C}{Nh} \mathbf{d}_1 \left(m_{\mathbf{x}+h\mathbf{e}_{\mathbf{j}}^{\mathbf{k}}}^{N,i}, m_{\mathbf{x}}^{N,i} \right) \leq \frac{C}{N^2}$$

for h sufficiently small, x_h being in the line segment between x_j and $x_j + he_k$, thanks to (3.51). Passing to the limit, as $h \to 0$, we conclude. □

We can prove now that each finite projection $(u_i^N)_{1 \leq i \leq N}$ of U is an approximated solution of the Nash system (3.5).

Theorem 3.16. *Assume Assumptions 1 hold, we have that* $u_i^N \in \mathcal{C}^1([0,T] \times \Omega^N)$, $u_i^N(t, \cdot) \in W^{2,\infty}(\Omega^N)$ *and* u_i^N *solves almost everywhere the following equation*

$$\begin{cases} -\partial_t u_i^N - \sum_j \operatorname{tr}\big(a(x_j) D_{x_j x_j}^2 u_i^N\big) + H(t, x_i, D_{x_i} u_i^N) \\ \qquad + \sum_{j \neq i} H_p(t, x_j, D_{x_j} u_i^N) \cdot D_{x_j} u_i^N = F(t, x_i, m_{\boldsymbol{x}}^{N,i}) + r_i^N(t, \boldsymbol{x}), \\ u_i^N(T, \boldsymbol{x}) = G(x_i, m_{\boldsymbol{x}}^{N,i}), \\ u_i^N(t, \boldsymbol{x})_{|x_i \in \partial\Omega} = 0, \qquad i = 1, ..., N. \end{cases} \tag{3.58}$$

where $||r_i^N||_\infty \leq \dfrac{C}{N}$.

Equation (3.58) models an approximated version of the Nash system (3.5). The homogeneous boundary condition for U implies that $u_i^N(t, \mathbf{x})\big|_{x_i \in \partial\Omega} = 0$ for each $i = 1, ..., N$. The rest of the proof follows analogously to the one of Prop. *Proposition 6.3* of [41], therefore we omit it.

3.11.2 Convergence

The main convergence result is realized by comparing (u_i^N), the projections of the solution of the Master Equation defined in Equation (3.53), with (v_i^N), namely with the solution of the Nash system defined in (3.5).

Since both of these solutions are symmetrical, there exist two functions V^N and U^N : $\Omega \times \Omega^{N-1} \to \mathbb{R}$, such that, for all $x \in \Omega$, the functions $(y_1, \dots, y_{N-1}) \to V^N(x, (y_1, \dots, y_{N-1})$ are invariant under permutations. Therefore, we can write

$$v_i^N(t, \mathbf{x}) = V^N(t, x_i, (x_1, \dots x_{i-1}, x_{i+1}, \dots, x_N)),$$

$$u_i^N(t, \mathbf{x}) = U^N(t, x_i, (x_1, \dots x_{i-1}, x_{i+1}, \dots, x_N)).$$

We fix the initial condition and the corresponding initial spatial distribution $(t_0, m_0) \in [0, T) \times \mathcal{P}^{sub}(\Omega)$, also defining $\mathbf{M} = (M^i)_{i \in \{1, \dots, N\}}$ as a family of independent and identically distributed (i.i.d) random variables of law $m_0 \in P^{sub}$. Let also $((B_t^i)_{t \in [0,T]})_{i \in \{1, \dots, N\}}$ be a family of N independent d-dimensional Brownian motion associated with each player i, being also independent of $(Z_i)_{i \in \{1, \dots, N\}}$.

We consider the processes $(\mathbf{Y}_t)_t$ and their associated hitting times $\boldsymbol{\tau}_Y = \left[\tau_Y^1, \dots, \tau_Y^N\right]$, solution of the following coupled system of SDEs

$$\begin{cases} dY^i_{t \wedge \tau_Y^i} = -H_p(t, Y^i_{t \wedge \tau_Y^i}, D_{x_i} v_i^N(t, \mathbf{Y}_{t \wedge \boldsymbol{\tau}_Y}))dt + \sqrt{2}\sigma(Y^i_{t \wedge \tau_Y^i}) dB_t^i, \\ Y^i_{t_0} = M^i, \end{cases} \tag{3.59}$$

with $\mathbf{Y}_{t \wedge \boldsymbol{\tau}_Y} := \mathbf{Y}_{t \wedge \boldsymbol{\tau}_Y} = \left[Y^1_{t \wedge \tau_Y^1}, \dots Y^i_{t \wedge \tau_Y^i} \dots, Y^N_{t \wedge \tau_Y^N}\right]$ while the process $dY^i_{t \wedge \tau_Y^i}$ evolves according to Eq. (3.59), but only before the corresponding hitting time for $t \in [t_0, \tau_Y^i]$. Accordingly, the correspondent processes $(\boldsymbol{X}_t)_t$ and $\boldsymbol{\tau}_X = \left[\tau_X^1, \dots \tau_X^N\right]$, being related to the projections u_i^N solve the following system

$$\begin{cases} dX^i_{t \wedge \tau_X^i} = -H_p(t, X^i_{t \wedge \tau_X^i}, D_{x_i} u_i^N(t, \mathbf{X}_{t \wedge \boldsymbol{\tau}_X}))dt + \sqrt{2}\sigma(X^i_{t \wedge \tau_X^i}) dB_t^i, \\ X^i_{t_0} = M^i. \end{cases} \tag{3.60}$$

with $\mathbf{X}_{t \wedge \boldsymbol{\tau}_X} := \mathbf{X}_{t \wedge \boldsymbol{\tau}_X} = \left[X^1_{t \wedge \tau_X^1}, \dots X^i_{t \wedge \tau_X^i} \dots, X^N_{t \wedge \tau_X^N}\right]$.

Moreover, we define the minimum between τ_X^i and τ_Y^i as follows $\tau^i := \tau_X^i \wedge \tau_Y^i$.

Remark 3.17. *Let us underline that, evaluating projections over processes $\boldsymbol{Y}$, u_i^N may show some frozen components corresponding to those players that have reached the boundary. For* *example, if we assume an arbitrate time* $t = \tau_Y^i \geq \tau_Y^1$, u_i^N *for which a player has already reached* $\partial\Omega$*, then* u_i^N *reads*

$$u_i^N \left(t, Y^1_{\tau^1_Y}, \dots \{Y^i_t\}_{t \wedge \tau^i_Y} \dots, \{Y^N_t\}_{t \leq \tau^N_Y} \right)$$

with a constant, fixed component corresponding to $Y_{\tau^1_Y}$*.*

The following result holds:

Proposition 3.18. *Assume Hypotheses 1 hold, then, for any* $1 \leq i \leq N$*, we have*

$$\mathbb{E}\left[\sup_{t \in [t_0, \tau^i]} |Y^i_{t \wedge \tau^i} - X^i_{t \wedge \tau^i}| \right] \leq \frac{C}{N}, \tag{3.61}$$

$$\mathbb{E}\left[\int_{t_0}^{t \wedge \tau^i_Y} |D_{x_i} v_i^N(t, \boldsymbol{Y}_{t \wedge \tau^i_Y}) - D_{x_i} u_i^N(t, \boldsymbol{Y}_{t \wedge \tau^i_Y})|^2 dt \right] \leq \frac{C}{N^2}. \tag{3.62}$$

Moreover, $\mathbb{P}$*-almost surely and for all* $i = 1, \dots, N$*, it holds*

$$|u_i^N(t_0, \boldsymbol{M}) - v_i(t_0, \boldsymbol{M})| \leq \frac{C}{N}, \tag{3.63}$$

where C *is a deterministic constant not depending on* t_0*,* m_0 *and* N*.*

Proof. First step. The proof proceeds as in *Theorem 6.2.1* in [41], but we decided to include it both for the reader's convenience and since it is not restricted to the case $a(x) = Id$.

For the sake of simplicity, but without loss of generality, we can assume $t_0 = 0$. In what follows we use the following notations:

$$\begin{aligned}
U_t^{N,i} &= u_i^N \left(t, Y^1_{t \wedge \tau^1_Y}, \dots, Y^N_{t \wedge \tau^N_Y} \right), \\
DU_t^{N,i,j} &= D_{x_j} u_i^N \left(t, Y^1_{t \wedge \tau^1_Y}, \dots, Y^N_{t \wedge \tau^N_Y} \right), \\
V_t^{N,i} &= v_i^N \left(t, Y^1_{t \wedge \tau^1_Y}, \dots, Y^N_{t \wedge \tau^N_Y} \right), \\
DV_t^{N,i,j} &= D_{x_j} v_i^N \left(t, Y^1_{t \wedge \tau^1_Y}, \dots, Y^N_{t \wedge \tau^N_Y} \right).
\end{aligned}$$

Since v_i^N solves the Nash system (3.5), applying the Itô formula, see [5], to $\left(U_t^{N,i} - V_t^{N,i}\right)^2$, we can deduce the following expansion for $d\left(U_t^{N,i} - V_t^{N,i}\right)^2$:

$$(A_t + B_t)dt + 2\sqrt{2}\left(U_t^{N,i} - V_t^{N,i}\right) \sum_{j=1}^N \left[\sigma(Y^i_{t \wedge \tau^i_Y})(DV_t^{N,i,j} - DU_t^{N,i,j}) \right] dB_t^i,$$

where

$$
\begin{aligned}
A_t &= 2\left(U_t^{N,i} - V_t^{N,i}\right)\left(H(t, Y^i_{t\wedge\tau^i_Y}, DU_t^{N,i,i}) - H(t, Y^i_{t\wedge\tau^i_Y}, DV_t^{N,i,i})\right)\\
&-2\left(U_t^{N,i} - V_t^{N,i}\right)\left(DU_t^{N,i,i}(H_p(t, Y^i_{t\wedge\tau^i_Y}, DU_t^{N,i,i}) - H_p(t, Y_t, DV_t^{N,i,i}))\right)\\
&-2\left(U_t^{N,i} - V_t^{N,i}\right)\left((DU_t^{N,i,i} - DV_t^{N,i,i})H_p(t, Y_{t\wedge\tau^i_Y}, DV_t^{N,i,i}) - r_i^N(t, \mathbf{Y}_{t\wedge\tau_Y}))\right).
\end{aligned}
$$

and

$$
\begin{aligned}
B_t &= 2\sum_{j=1}^N \langle a(Y^i_{t\wedge\tau^i_Y})(DU_t^{N,i,j} - DV_t^{N,i,j}), DU_t^{N,i,j} - DV_t^{N,i,j}\rangle\\
&-2(U_t^{N,i} - V_t^{N,i})\sum_{j=1}^N DU_t^{N,i,j}\left(H_p(t, Y^j_{t\wedge\tau^i_Y}, DV_t^{N,j,j}) - H_p(t, Y^j_{t\wedge\tau^i_Y}, DU_t^{N,j,j})\right).
\end{aligned}
$$

Let us underline that $|DU_t^{N,i,j}| \leq M$, $|DV_t^{N,i,j}| \leq M$ for a certain $M > 0$, and that H and D_pH are both locally Lipschitz continuous. Moreover, $DU_t^{N,i,i} = D_xU(t, Y^i_{t\wedge\tau^i_Y}, m^{N,i}_{\mathbf{Y}_{t\wedge\tau_Y}})$ is bounded, independently of i, N, $DU_t^{N,i,j} = (1/(N-1))D_mU(t, Y^i_{t\wedge\tau^i_Y}, m^{N,i}_{\mathbf{Y}_{t\wedge\tau_Y}})$ is bounded from C/N and, by Theorem 3.16, we also have that $r^{N,i} \leq C/N$

Integrating from t to τ^i_T the above-squared difference and taking the conditional expectation given $\mathbf{Z}$, we deduce

$$
\begin{aligned}
\mathbb{E}^{\mathbf{Z}}\left[|U_t^{N,i} - V_t^{N,i}|^2\right] &+ 2\sum_{j=1}^N \mathbb{E}^{\mathbf{Z}}\left[\int_t^T |DU_s^{N,i,j} - DV_s^{N,i,j}|^2 ds\right]\\
&\leq \frac{C}{N}\int_t^T \mathbb{E}^{\mathbf{Z}}\left[|U_s^{N,i} - V_s^{N,i}|\right]\\
&+ C\int_t^T \mathbb{E}^{\mathbf{Z}}\left[|U_s^{N,i} - V_s^{N,i}| \cdot |DU_s^{N,i,i} - DV_s^{N,i,i}|\right] ds\\
&+ \frac{C}{N}\sum_{j=1, j\neq i}^N \int_t^T \mathbb{E}^{\mathbf{Z}}\left[|U_s^{N,i} - V_s^{N,i}| \cdot |DU_s^{N,j,j} - DV_s^{N,j,j}|\right] ds.
\end{aligned}
$$

Following a standard convexity argument, as, e.g., in [41], and by the generalized Young's inequality we get

$$
\begin{aligned}
&\mathbb{E}^{\mathbf{M}}[|U_t^{N,i}-V_t^{N,i}|^2]+\mathbb{E}^{\mathbf{M}}\left[\int_t^T|DU_s^{N,i,i}-DV_s^{N,i,i}|^2ds\right]\\
&\quad\le\frac{C}{N^2}+C\int_t^T\mathbb{E}^{\mathbf{Z}}[|U_s^{N,i}-V_s^{N,i}|]^2ds\\
&\quad+\frac{1}{2N}\sum_{j=1}^N\mathbb{E}^{\mathbf{M}}\left[\int_t^T|DU_s^{N,j,j}-DV_s^{N,j,j}|^2ds\right].
\end{aligned}
\tag{3.64}
$$

By taking the mean over $i\in\{1,...,N\}$, we can get rid of the second term in the left-hand side.

By the Gronwall's Lemma (eventually increasing the constant C), we deduce

$$
\sup_{0\le t<\tau_Y^i}\left[\frac{1}{N}\sum_{i=1}^N\mathbb{E}^{\mathbf{M}}[U_t^{N,i}-V_t^{N,i}]^2\right]\le\frac{C}{N^2}.
\tag{3.65}
$$

Plugging Eq. (3.65) into Eq. (3.64), we deduce that

$$
\frac{1}{N}\sum_{j=1}^N\mathbb{E}^{\mathbf{M}}\left[\int_t^T|DU_s^{N,j,j}-DV_s^{N,j,j}|^2ds\right]\le\frac{C}{N^2}.
$$

Inserting this inequality into Eq. (3.64) and applying again Gronwall's Lemma, we finally obtain

$$
\sup_{t\in[0,\tau_Y^i]}\mathbb{E}^{\mathbf{M}}[|U_t^{N,i}-V_t^{N,i}|^2]+\mathbb{E}^{\mathbf{M}}\left[\int_0^T|DU_s^{N,i,i}-DV_s^{N,i,i}|^2ds\right]\le\frac{C}{N^2},
\tag{3.66}
$$

proving the bound in Eq. (3.62).

Second step. We now derive (3.63).

Evaluating Eq. (3.66) at $t=0$, we obtain $\mathbb{P}$-almost surely for $i\in\{1,...,N\}$

$$
|U_0^{N,i}-V_0^{N,i}|=|u^{N,i}(0,\mathbf{M})-v^{N,i}(0,\mathbf{M})|\le\frac{C}{N},
$$

hence proving Eq. (3.63).

Third step. We are now in the position to prove bound for optimal trajectories, namely Eq. (3.61).

In view of the dynamics, encoded in (3.59) and (3.60) and respectively satisfied by processes $(Y_{i,t})_{t\in[0,\tau^i]}$ and $(X_{i,t})_{t\in[0,\tau^i]}$, we know that

$$
\begin{aligned}
\left(X^i_{t\wedge\tau^i_X} - Y^i_{t\wedge\tau^i_Y}\right)^2 \le \Bigg(\int_0^{\tau^i} & \left(D_pH\left(s, X^i_{s\wedge\tau^i_X}, D_{x_i}u^{N,i}(s, \mathbf{X}_{s\wedge\tau_X})\right)\right. \\
& \left.\left. - D_pH\left(s, Y^i_{s\wedge\tau^i_Y}, D_{x_i}v^{N,i}(s, \mathbf{Y}_{s\wedge\tau_Y})\right)\right) ds \right)^2 \\
& + \left(\int_0^{\tau^i} \sqrt{2}\left(\sigma(X^i_{s\wedge\tau^i_X}) - \sigma(Y^i_{s\wedge\tau^i_Y})\right) dB^i_s \right)^2 .
\end{aligned}
$$

Since D_pH is Lipschitz continuous and U is regular, bounds in Prop. 3.15 implies

$$
\begin{aligned}
\left(X^i_{t\wedge\tau^i_X} - Y^i_{t\wedge\tau^i_Y}\right)^2 \le C \Bigg[& \int_0^{\tau^i} \left(|X^i_{s\wedge\tau^i_X} - Y^i_{s\wedge\tau^i_Y}| + \frac{1}{N}\sum_{j\neq i} |X^j_{s\wedge\tau^i_X} - Y^j_{s\wedge\tau^i_Y}| \right) ds \\
+ \int_0^{\tau^i} & D_pH\left(s, Y^i_{s\wedge\tau^i_Y}, D_{x_i}u^{N,i}(s, \mathbf{Y}_{s\wedge\boldsymbol{\tau}_Y})\right) \\
& - D_pH\left(s, Y^i_{s\wedge\tau^i_Y}, D_{x_i}v^{N,i}(s, \mathbf{Y}_{s\wedge\boldsymbol{\tau}_Y})\right) ds \Bigg]^2 \\
+ \left(\int_0^{\tau^i} \right. & \left. \sqrt{2}\left(\sigma(X^i_s) - \sigma(Y^i_s)\right) dB^i_s \right)^2 .
\end{aligned}
$$

Taking the sup over $t \in [0,\chi]$ with $\chi \in [0,\tau^i]$ and the conditional expectation over $\mathbf{M}$, we obtain

$$
\begin{aligned}
\mathbb{E}^{\mathbf{M}}\left[\sup_{t\in[0,\chi]} |X^i_{t\wedge\tau^i_X} - Y^i_{t\wedge\tau^i_Y}|^2\right] \le \mathbb{E}^{\mathbf{M}} & \left[\int_0^T |DU^{N,i,i}_s - DV^{N,i,i}_s|^2 ds\right] \\
+ C\int_0^{\chi} & \left(\mathbb{E}^{\mathbf{M}}\left[\sup_{t\in[0,s]} |X^i_{t\wedge\tau^i_X} - Y^i_{t\wedge\tau^i_Y}|^2\right] \right. \\
& \left. + \frac{1}{N}\sum_{j\neq i} \mathbb{E}^{\mathbf{M}}\left[\sup_{t\in[0,s]} |X^j_{t\wedge\tau^j_X} - Y^j_{t\wedge\tau^j_Y}|^2\right] \right) ds \\
+ \sup_{t\in[0,s]} & \mathbb{E}^{\mathbf{M}}\left[\left(\int_0^{\chi} \sqrt{2}\left(\sigma(X^i_{t\wedge\tau^i_X}) - \sigma(Y^i_{t\wedge\tau^i_Y})\right) dB^i_t\right)^2\right] .
\end{aligned}
\tag{3.67}
$$

By Itô isometry and exploiting the Lipschitz assumption on σ, we know that

$$
\begin{aligned}
& \sup_{t\in[0,s]} \mathbb{E}^{\mathbf{M}}\left[\left(\int_0^{\chi} \sqrt{2}\left(\sigma(X^i_{t\wedge\tau^i_X}) - \sigma(Y^i_{t\wedge\tau^i_Y})\right) dB^i_t\right)^2\right] \\
& = \sup_{t\in[0,s]} \mathbb{E}^{\mathbf{M}}\left[\int_0^{\chi} 2\left(\sigma(X^i_{t\wedge\tau^i_X}) - \sigma(Y^i_{t\wedge\tau^i_Y})\right)^2 dt\right] \\
& \le \int_0^{\chi} \left(\mathbb{E}^{\mathbf{M}} \sup_{t\in[0,s]} |X^i_{t\wedge\tau^i_X} - Y^i_{t\wedge\tau^i_Y}|^2 ds \right) .
\end{aligned}
$$

We now take the sum over players $i \in \{1,...,N\}$ and enforce Eq. (3.66). Then, by Gronwall's Lemma, we know that

$$\sum_{i=1}^{N} \mathbb{E}^{\mathbf{M}} \left[\sup_{t\in[0,\tau^i]} |X^i_{t\wedge\tau^i_X} - Y^i_{t\wedge\tau^i_Y}| \right] \le C\,. \tag{3.68}$$

Inserting inequality (3.68) in (3.67), and applying again Gronwall's inequality, we obtain (3.61).

□

Coming back to Theorem 3.3, it states the fundamental result about the convergence of $(v^{N,i})_{i\in\{1,...,N\}}$, the solution of the Nash system (3.58) towards U, namely the classical solution to the Master Equation (3.7) stated with Dirichlet boundary conditions.

Proof of Theorem 3.3. For the sake of simplicity, we can choose $m_0 = 1$, then apply the estimate provided by Equation (3.63). Thus,

$$\left| U(t_0, M_i, m^{N,i}_{\mathbf{M}}) - v^{N,i}(t_0,\mathbf{M}) \right| \le \frac{C}{N} \qquad a.e. \quad i \in \{1,...,N\}\,,$$

where $\mathbf{M} = (M_1,...,M_N)$ are i.i.d. random variables, uniformly distributed in Ω. The support of $\mathbf{M}$ is Ω^N, thus from the continuity of U we have

$$\left| U(t_0, x_i, m^{N,i}_{\mathbf{x}}) - v^{N,i}(t_0,\mathbf{x}) \right| \le \frac{C}{N} \qquad \forall \mathbf{x} \in \Omega^N\,, \quad i \in \{1,...,N\}\,.$$

Let us underline that latter estimates hold for any $\mathbf{x} \in \Omega^N$. Moreover, since U is Lipschitz continuous with respect to m, we can replace $U(t_0,x_i,m^{N,i}_{\mathbf{x}})$ with $U(t_0,x_i,m^N_{\mathbf{x}})$, since

$$\left| U(t_0, x_i, m^{N,i}_{\mathbf{x}}) - U(t_0,x_i,m^N_{\mathbf{x}}) \right| \le C\mathbf{d}_1(m^{N,i}_{\mathbf{x}}, m^N_{\mathbf{x}}) \le \frac{C}{N}\,.$$

To prove *(ii)* we need the convergence results of empirical measure stated in [2, 79, 108].

First, we move from the sub-probability measure m_0, defined over Ω to m_0^*, which corresponds to a possible probability measure defined over $\mathcal{P}(\mathbb{R}^n)$, and such that its restriction on Ω coincides with $m_0 \in \mathcal{P}^{sub}(\Omega)$, namely $m_0^*(x)|_\Omega = m_0(x), \forall x \in \Omega$. We remark that m_0^* is one of the possible probability measures since the extension is not uniquely determined. Then, the following inequality holds

$$\begin{aligned} &\int_{\Omega^{N-1}} |u^N_i(t_0,\mathbf{x}) - U(t_0,x_i,m_0)| \prod_{j\neq i} m_0(dx_j) \\ &\quad \le \int_{\mathbb{R}^{N-1}} |u^N_i(t_0,\mathbf{x}) - U(t_0,x_i,m_0^*)| \prod_{j\neq i} m_0^*(dx_j)\,. \end{aligned}$$

Since U is Lipschitz continuous, and we are now working with probability measures, we can apply results from [2, 79, 108], obtaining

$$\begin{aligned}&\textstyle\int_{\mathbb{R}^{N-1}} |u_i^N(t_0,\mathbf{x}) - U(t_0,x_i,m_0^*)| \prod_{j\neq i} m_0^*(dx_j)\\ &\qquad \leq C \textstyle\int_{\mathbb{R}^{N-1}} \mathbf{d}_1(m_x^{N,i},m_0^*) \prod_{j\neq i} m_0^*(dx_j) \leq C\omega_N\,,\end{aligned}$$

where ω_N is defined as the right side in Eq. (3.13).

The latter implies that

$$\begin{aligned}&||w_i^N(t_0,\cdot,m_0) - U(t_0,\cdot,m_0)||_{L^1(m_0)} =\\ &= \int_\Omega \left| \int_{\Omega^{N-1}} \Big(v_i^N(t_0,\mathbf{x}) - U(t_0,x_i,m_0)\Big) \prod_{j\neq i} m_0(dx_j) \right| m_0(dx_i)\\ &\leq \mathbb{E}\Big[|v_i^N(t,\mathbf{M}) - u_i^N(t,\mathbf{M})|\Big] + \int_{\Omega^N} |u_i^N(t_0,\mathbf{x}) - U(t_0,x_i,m_0^*)| \prod_j m_0^*(dx_j)\\ &\leq \frac{C}{N} + C\omega_N \leq C\omega_N\end{aligned}$$

holds, hence proving both Eq. (3.13) and the Theorem.

□

3.12 Application to a Toy Model for Optimal Liquidations

We will study a toy model of mean field type which can be seen as the limiting problem of an N player game. In particular, we adapt the setting described in [43, 127, 145], with multiple investors, letting players stop their investments when a given, fixed, condition is met. We consider a variable interest rate - constant for certain periods of time and then variable at the discretion of the issuer of the loan. The investment ends when the price reaches a certain level, e.g. one prescribed by the contractor, which could also be (from the *low level* perspective) the default one. Thus, we introduce a set in which the players want to operate within a fixed lower bound $l \in \mathbb{R}$ and a fixed upper one $u \in \mathbb{R}$.

We focus on an N-player game with an Itô-diffusion setting for a model with a controlled drift and state-independent controls. The N players, indexed by $i \in \mathcal{I} = \{1,2,...,N\}$, have a common investment horizon $[0,T]$. Each player i uses a self-financing strategy $(\pi_t^i)_{t\in[0,T]}$ and can control its own drift. The corresponding wealth, denoted by $(X_t^i)_{t\in[0,T]}$, represents the amount of the investment, and it evolves accordingly to the following SDE:

$$dX_t^i = \pi_t^i dt + \sqrt{2}\sigma dB_t^i\,, \tag{3.69}$$

assuming a control $\pi_t^i \in \mathcal{A}$ and a constant volatility σ. The self-financing strategies π^i belong to the set of admissible policies $\mathcal{A}$ defined as

$$\mathcal{A} = \left\{ \pi_\cdot^i : [0,T] \to L^2(\tilde{\Omega}) \text{ } \mathcal{F}\textit{-progr. measurable with } \mathbb{E}\left[\int_0^{\tau^i} \sigma^2 \pi_s^2 ds \right] < \infty \right\}$$

with τ^i corresponding to the hitting time of the i^{th} wealth, which corresponds to the time when X_t^i reaches one of the critical values defined by the bounds, l or u:

$$\tau^i = \inf \left\{ t \in [0,T] \, : \, X_t^i = l \vee X_t^i = u \right\}$$

Following the setting stated in [145], players optimize their expected terminal utility being also concerned with the performance of their competitors. Accordingly, we fix an arbitrary policy $(\pi_1, ..., \pi_{i-1}, \pi_{i+1}, ..., \pi_N)$ for the other players, whereas the i^{th} player aims to maximize

$$J^i(t, \boldsymbol{x}, \pi_s^i) = \mathbb{E}\left[\int_t^{\tau^i} \frac{1}{2} \left(\pi_s^i \right)^2 ds + G\left(X_{\tau^i}^i, m_{\boldsymbol{X}_{\tau^i}^i}^{N,i} \right) | \boldsymbol{X_t} = \boldsymbol{x} \right],$$

with the terminal payoff G defined as follows:

$$G(x,m) = x(1-x) \int_0^1 z(1-z) dm(z) \, .$$

We underline that functions $H(p) = \frac{1}{2}|p|^2$ and G both satisfy hypotheses 1.

For our 1-dimensional trading example, we enforce a simplified choice of the domain by prescribing both lower and upper limits, i.e. $l = 0$ and $u = 1$. Moreover, we fix the volatility $\sigma = 1$.

Accordingly, the Nash system takes the following form.

$$\begin{cases} -\partial_t v_i^N - \displaystyle\sum_{j=1}^N \partial_{x_j x_j}^2 v_i^N(t, \boldsymbol{x}) + \frac{1}{2} \left(\partial_{x_i} v_i^N(t, \boldsymbol{x}) \right)^2 \\ \qquad + \displaystyle\sum_{j \neq i}^N \partial_{x_j} v_j^N(t, \boldsymbol{x}) \partial_{x_j} v_i^N(t, \boldsymbol{x}) = 0 \, , \\ v_i^N(T, \mathbf{x}) = \dfrac{1}{N-1} x_i(1-x_i) \displaystyle\sum_{j \neq i}^N x_j(1-x_j) \, , \\ v_i^N(t, \mathbf{x})_{|x_i \in \{0,1\}} = 0 \, , \quad i = 1, ..., N \, . \end{cases}$$

We will study models of mean field type which can be seen as the limiting problem of the above-described one. When $N \to \infty$, the dynamic for a representative player is described by

$$dX_t = \pi_t dt + \sqrt{2} dB_t .$$

The corresponding functional reads

$$J(x, \pi_t) = \mathbb{E}\left[\int_0^\tau \frac{1}{2} {\pi_s}^2 ds + X_\tau (1 - X_\tau) \int_0^1 z(1-z) dm_\tau(z) \,\middle|\, X_t = x \right],$$

where m_τ is the density of the process (X_τ).

The Master Equation for this simplified setting reads

$$\begin{cases} -\partial_t U(t,x,m) - \partial_{xx} U(t,x,m) + \frac{1}{2} \left(\partial_x U(t,x,m)\right)^2 + \\ \qquad - \int_0^1 \partial_y \left(D_m U(t,x,m,y)\right) dm(y) + \\ \qquad + \int_0^1 D_m U(t,x,m,y) \cdot \partial_x U(t,y,m) dm(y) = 0 \\ U(T,x,m) = x(1-x) \int_0^1 z(1-z) dm(z) \\ U(t,0,m) = U(t,1,m) = 0 \\ \frac{\delta U}{\delta m}(t,x,m,0) = \frac{\delta U}{\delta m}(t,x,m,1) = 0 . \end{cases}$$

Although the solution of the Nash system is not easy to compute, we can directly approach the Master Equation, using the definition of U accordingly to (3.9). In particular, the Hamilton-Jacobi equation for our formulation reads as follows:

$$\begin{cases} -u_t - u_{xx} + \frac{1}{2} (u_x)^2 = 0 \\ u(T,x) = x(1-x) \int_0^1 z(1-z) dm_T(z) , \qquad \forall x \in [0,1], , \\ u(t,0) = u(t,1) = 0 , \qquad \forall t \in [t_0, T] . \end{cases}$$

By the change of variable $w = e^{-\frac{1}{2}u}$, see also [42], the equation for the function w becomes

$$\begin{cases} -w_t - w_{xx} = 0 , \\ w(T,x) = \exp\left(-\frac{1}{2} x(1-x) \int_0^1 z(1-z) dm_T(z) \right) , \\ w(t,0) = w(t,1) = 1 . \end{cases}$$

hence a classical heat equation whose solution can be computed with standard methods. Coming back to the change of variable $u = -2\log(w)$ and (3.9), the function $U(t_0, \cdot, m_0) = u(t_0, \cdot)$ admits the following representation:

$$U(t_0, x, m_0) = -2\ln\left(1 + \sum_{k=1}^{+\infty} b_k e^{-k^2\pi^2(T-t_0)} \sin(k\pi x)\right),$$

being $b_k = 2\int_0^1 e^{-\frac{1}{2}y(1-y)\int_0^1 z(1-z)dm_T(z)} \sin(k\pi y)dy$ and $m_T = m(T, \cdot)$, where m is the solution of the Fokker-Planck equation

$$\begin{cases} m_t - m_{xx} - (mu_x)_x = 0\,, \\ m(t_0, x) = m_0\,, & \forall x \in [0,1]\,, \\ m(t,0) = m(t,1) = 0\,, & \forall t \in [t_0, T]\,. \end{cases}$$

4 From Optimal Control to Mean Field Optimal Transport via Stochastic Neural Networks

4.1 Introduction

In recent years, parametric Machine Learning (ML) applications have shown brilliant performances in capturing relevant symmetries and hidden patterns characterizing the specific knowledge base. Specifically, Neural Networks (NNs), i.e. systems of interconnected artificial neurons, constitute a fundamental tool to capture complex patterns and to make accurate predictions for various applications, ranging from computer vision and natural language processing to robotics and reinforcement learning. Their growing popularity steers an increasing demand for a deep mathematical description of the underlying training procedures, specifically in high dimensions to tackle the curse of dimensionality.

Along this latter research challenge, we consider a novel class of NNs, termed *Mean Field Neural Networks* (MFNNs), which are defined as the limiting object of a *population of NNs* when its number of components tends to infinity. Our aim concerns deriving a unified perspective for this class of models based on existing symmetries between Mean Field Control (MFC) theory and the Optimal Transport (OT) method. Our approach is based on an *infinite dimensional lifting* which allows getting new insights into relationships between data in the corresponding finite-dimensional scenario.

We start the analysis by looking at the continuous idealization of a specific class of NNs, namely Residual NNs also named ResNets, whose training process in a Supervised Learning scenario is stated as a Mean Field Optimal Control Problem (MFOCP). We consider a deterministic dynamic that evolves in terms of an ordinary differential equation (ODE). Moreover, the training problem of a ResNet is shown to be equivalent to a MFOCP of Bolza type, see [93] and [28] for further details.

The next passage in our analysis concerns introducing a noisy component into the dynamics of the ODE moving to a Stochastic Differential Equation (SDE) that allows us to consider the inherent uncertainty connected to the variations in the real-world data, symmetrically allowing for integrating stochastic aspects into the learning process. Although this second model does not include any mean field terms, it allows the development of a class of algorithms known as Stochastic NNs (SNNs). In [7] the authors develop a sample-wise backpropagation method for SNNs based on Backward SDE that models the gradient (w.r.t. the parameters) process of the loss function representing a feasible tool for quantifying the uncertainty of the learning process. Another possible approach for probabilistic learning is studied in [25] where the authors develop the so-called Stochastic Deep Network (SDN), namely a NN architecture that can use as input data not only single vectors but also *random vectors* modelling probability distribution of given inputs. Following their analysis, the SDN is considered as an architecture based on the composition of maps between probability measures performing inference tasks and solving ML problems over the space of probability measures.

In the last passage, we merge the stochastic aspect with the mean field one by considering the so-called Mean Field Optimal Transport (MFOT) formulation, recently introduced in [13]. We describe the MFC tools relevant to formalize the training process, hence we formulate the training problem as MFOT in an infinite-dimensional setting. Considering the collective interactions and distributions of the network's parameters may facilitate the analysis of the network behaviour on a macroscopic level, hence improving the interpretability, scalability, and robustness of NNs models, while adding knowledge by highlighting the hidden symmetries and relations between data.

We highlight that the symmetry between mean field models and ML algorithms is also studied in [128], where the authors establish a mathematical relationship between the MFG framework and normalizing flows, a popular method for generative models composed of a sequence of invertible mappings. Similarly, in [38], the authors analyze Generative Adversarial Networks (GANs) from the perspectives of MFGs providing a theoretical connection between GANs, OT and MFG and numerical experiments.

The chapter is organized as follows: in Section 4.2, we introduce the mathematical formalism of the supervised learning paradigm, while providing the description of the continuous idealization of a Residual NN stated as a MFOCP; in Section 4.3, we introduce a noisy component into the network dynamic, thus focusing on Stochastic NNs formalized as Stochastic Optimal Control problems; in Section 4.4 we review the MFG setting in a cooperative scenario defined in terms of MFC theory. Then, we consider recently developed Mean Field Optimal Transport methods that allow rephrasing MFC problems into OT ones. We also illustrate related approximation schemes and possible connection to an abstract class of NNs that respect the MFOT structure. We conclude by reviewing some methods to *learn*, i.e.

approximate, mean field functions that depend on probability distribution obtained as the limiting object of empirical measures.

4.2 Residual Neural Networks as a Mean Field Optimal Control Problem

In this section, we present the workflow to treat a feed-forward NN, specifically a Residual NN, as a dynamical system based on the work in [152]. The main reference for this part is the well-known paper [93], where the authors introduce a continuous idealization of Deep Learning (DL) to study the Supervised Learning (SL) procedure which is stated as an optimal control problem by considering the associated population risk minimization problem.

4.2.1 The Supervised Learning paradigm

Following [88, 153], the SL problem aims at estimating the function $\mathcal{F}:\mathcal{X}\to\mathcal{Y}$, commonly known as the Oracle. The space $\mathcal{X}$ can be identified with a subset of $\mathbb{R}^d$ related to input arrays (such as images, string texts or time series) while $\mathcal{Y}$ is the corresponding target set. Here for simplicity, we consider $\mathcal{X}$ and $\mathcal{Y}$ Euclidean spaces with different dimensions. Thus, training begins with a set of N input-target pairs $\{x_0^i, y_T^i\}_{i=1}^N$ where

- $x_0^i \in \mathbb{R}^d$ denotes the inputs of the NN;
- $x_T^i = \mathcal{F}(x_0^i) \in \mathbb{R}^d$ denotes the outputs of the NN;
- $y_T^i \in \mathbb{R}^l$ the corresponding targets.

We assume the same dimension of the Euclidean space for NN inputs and outputs allowing to explicitly write a dynamic in terms of a difference equation. Hence, for a ResNet, see [123] for more details, with T layers, the Feed-Forward propagation is given by

$$x_{t+1} = x_t + f(x_t,\theta_t) \qquad t = 0,\dots,T-1 \tag{4.1}$$

with $f:\mathbb{R}^d\to\mathbb{R}^d$ being a parameterized function and θ_t being the trainable parameters, e.g. bias, weights of layer t–th that belong to a measurable set $\mathcal{U}$ with values in a subset of the Euclidean space $\mathbb{R}^m$.

Remark 4.1. *Following [214], we report an example of a domain for parameters of NN with ReLU activation functions. We define the following parameter domain*

$$\Theta = \left\{(a,w,b)\in\mathbb{R}\times\mathbb{R}^d\times\mathbb{R} : a^2 < ||w|| + b^2\right\}$$

with activation functions $\phi : \Theta \to \mathbb{R}$ *defined as*

$$\phi(\theta;x) = a\sigma(w^T x + b)\,, \quad \theta = (a,w,b)\,, \quad \sigma(z) = z^+ = max\{z,0\},$$

4.2.2 Empirical Risk Minimization

We aim at minimizing, over the set of measurable parameters Θ, a *terminal loss function* $\Phi : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ plus a regularization term L, to derive a Supervised Learning problem as an Empirical Risk Minimization (ERP) problem, namely

$$\min_{\theta \in \mathcal{U}} \left[\frac{1}{K} \sum_{i=1}^{N} \Phi(x_T^i, y^i) + \sum_{t=0}^{T-1} L(\theta_t) \right] \tag{4.2}$$

over N training data samples indexed by i. We write $\theta = [\theta_0, \ldots \theta_{T-1}]$ to identify the set of all parameters of the network.

If we consider no regularization of the parameters, i.e. $L = 0$, and a quadratic loss function in terms of Φ, then Eq. (4.2) reads

$$J^{ERM}(\theta) = \min_{\theta \in \mathcal{U}} \left[\frac{1}{K} \sum_{i=1}^{N} ||x_T^i - y^i||^2 \right] = \min_{\theta \in \mathcal{U}} \left[\frac{1}{K} \sum_{i=1}^{N} ||f(x^i, \theta) - y^i||^2 \right] \tag{4.3}$$

being $x^i = [x_0, \ldots, x_{T_1}]$ the discrete state process defined in Eq. (4.1).

Optimizing J^{ERM} by computing its gradient is computationally expensive, especially if the number of data K is very large.

To handle the curse of dimensionality, it is usually common to initialize parameters from a θ^0 from a probability distribution, to then optimize their choice inductively according to a Stochastic Gradient Descent scheme

$$\theta^{k+1} = \theta^k - \eta_t \frac{1}{K} \sum_{i=1}^{K} ||f(x^i, \theta) - y^i|| \nabla_\theta f(x^i, \theta) \tag{4.4}$$

with learning rate η_t over K optimization steps.

For the sake of completeness, before going to the limit (we pass from a discrete set of training data to the corresponding distribution), we point out in the following remark that it is also possible to associate a measure corresponding to the empirical distribution of the parameters when the number of neurons goes to infinity.

Remark 4.2. *A different approach, as illustrated, e.g., by Sirignano and Spilipupouls in [200], consists in associating to each layer the corresponding empirical measure and building a measure to describe the whole network, hence working with the empirical measure of controls,* *rather than states as done in Section 4.4. Following the perspective of mean-field term in controls, the SGD equation* (4.4) *can be formalized as a minimization method over the set of probability distributions. Moreover, the training of NN is based on the correspondence between the empirical measure of neurons μ_N and the function f_N that is approximated by NN. Specifically, it has been proved that the training via gradient descent of an over-parameterized 1-hidden layer NN with infinite width is equivalent to gradient flow in Wasserstein space [57, 88, 93, 102]. Conversely, in the small learning rate regime the training is equivalent to a SDE, see, e.g., [56].*

From here on, we deal with empirical distribution and measures associated to the training data.

4.2.3 Population Risk Minimization as Mean Field Optimal Control Problem

In what follows we move from the discrete setting to the corresponding continuous idealization by:

- going from layer index T to continuous parameter t;
- passing from a discrete set of inputs/output to distribution μ that represents the joint distribution in $\mathcal{W}_2(\mathbb{R}^d \times \mathbb{R}^l)$ modelling the input-label distribution;
- passing from empirical risk minimization to population risk (i.e. minimization over expectation $\mathbb{E}$).

In particular, we pass to the limit in the number of data samples (number of input-target pairs) assuming also a continuous dynamic in place of layers discretization. The latter limit allows us to describe the dynamic of the state process x by the following Ordinary Differential Equation (ODE)

$$\dot{x}_t = f(x_t, \theta_t), \qquad t \in [0, T], \tag{4.5}$$

in place of the finite difference equation (4.1). We identify the input-target pairs as sampled from a given distribution μ allowing us to write the SL problem as a Population Risk Minimization (PRM) problem.

In summary, we aim at approximating the Oracle function $\mathcal{F}$ using a provided set of training data sampled by a (known) distribution μ_0 by optimizing weights θ_t to achieve maximal proximity between x_T (output) and y_T (target). Thus, we consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and we assume inputs x_0 in $\mathbb{R}^d$ to be sampled from a distribution $\mu_0 \in \mathcal{P}(\mathbb{R}^d)$, with corresponding target y_T in $\mathbb{R}^l$ sampled from a distribution $\nu \in \mathcal{P}(\mathbb{R}^l)$, while the joint probability distribution μ, that models the distribution of the input-target pairs, is defined by $\mu := \mathcal{P}(x_0, y_T)$, belongs to the Wasserstein space $\mathcal{W}_2(\mathbb{R}^{(d+l)})$ and has μ_0 and ν as its marginals. We recall that given a metric space (X, d), the p-Wasserstein space $\mathcal{W}_p(X)$ is defined as the set of all Borel probability measures on X with finite p-moments.

The marginal distributions are obtained by projecting the joint probability distribution μ over the subspaces of inputs and output, respectively. We identify the first marginal, i.e. the projection over $\mathbb{R}^l$, with the distribution of inputs

$$\mu_0 = \int_{R^l} \mu(x, y) dy \ ,$$

while the distribution of targets reads

$$\nu = \int_{R^d} \mu(x, y) dx \, .$$

Moreover, we assume the controls θ_t depend on the whole distribution of input-target pairs capturing the mean-field aspect of the training data. We consider a measurable set of admissible controls, i.e. training weights, $\Theta \subseteq \mathbb{R}^m$ and we state a Mean Field Optimal Control Problem (MFOCP) to solve the following PRM problem:

$$\begin{aligned}
&\inf_{\theta \in L^\infty([0,T],\Theta)} J^{PRM}(\theta) := \mathbb{E}_\mu \Big[\Phi(x_T, y_T) + \int_0^T L(x_t, \theta_t) dt \Big] \\
&\dot{x}_t = f(x_t, \theta_t) \quad 0 \le t \le T \quad x_0 \sim \mu_0, , \quad (x_0, y_T) \sim \mu
\end{aligned} \tag{4.6}$$

We briefly report basic assumptions allowing us to have a solution for (4.6):

- $f : \mathbb{R}^d \times \Theta \to \mathbb{R}^d$, $L : \mathbb{R}^d \times \Theta \to \mathbb{R}$, $\Phi : \mathbb{R}^d \times \mathbb{R}^l \to \mathbb{R}$ are bounded;
- f, L, Φ are Lipschitz continuous with respect to x, with the Lipschitz constants of f and L being independent of parameters θ, and
- μ has finite support in $\mathcal{W}_2(\mathbb{R}^{(d+l)})$

Problem (4.6) can be approached through two different methods: the first one is based on the Hamilton-Jacobi-Bellman (HJB) equation in the Wasserstein space, while the second one is based on a Mean Field Pontryagin Principle. We refer to [112] and [133] for viscosity solutions to the HJB equation in the Wasserstein space of probability measures, and to [19] for solving the constrained optimal control problems via Pontryagin Maximum Principle.

For the sake of completeness, let us also cite [10] where the authors introduce a BSDE technique to solve the related Stochastic Maximum Principle allowing us to consider the uncertainty associated with NN. The authors employ a Stochastic Differential Equation (SDE) in place of the ODE appearing in (4.6) to continuously approximate a Stochastic Neural Network (SNN). We deepen this approach in the next paragraph.

4.3 Stochastic Neural Network as a Stochastic Optimal Control Problem

In this paragraph, we generalize the previous setting considering a noisy dynamic, namely adding a stochastic integral to the deterministic setting described by the ODE in Problem (4.6). The reference model corresponds to Stochastic NN whose discrete state process is described by the following equation

$$X_{n+1} = X_n + hF(X_n, \theta_n) + \sqrt{h}\sigma_n \omega_n \,, \qquad n = 0, 1, \dots, N-1 \tag{4.7}$$

$\{\omega_n\}$ being a sequence of i.i.d. standard Gaussian random variables. We refer to [25] for a theoretical and computational analysis of SNN.

Eq. (4.7) can be generalized in a continuous setting. To this end, we consider a complete filtered probability space $\left(\Omega, \mathcal{F}, \mathbb{F}^W, \mathbb{P}\right)$, and we introduce the following SDE

$$X_t = X_0 + \int_0^T F\left(X_s, \theta_s\right) + \int_0^T \sigma_s dW_s \,, \tag{4.8}$$

with standard Brownian motion $W := (W_t)_{0 \leq t \leq T}$ and diffusion term σ. Analogously as for ResNets, the index $T > 0$ represents a continuous parameter modelling the width of the layer, being X_T the output of the network.

Here, we report the theory developed in [7] to study Eq. (4.8) in the framework of SOC problem by introducing the control process $u = [\theta, \sigma]$. Thus we consider also the diffusion σ as a trainable parameter of the model. We start by translating the SDE (4.8) into the following controlled process, written in differential form

$$dX_t = f(X_t, u_t)dt + g(u_t)dW_t, \qquad 0 \leq t \leq T, \tag{4.9}$$

where $f(X_t, u_t) = F(X_t, \theta_t)$ and $g(u_t) = \sigma_t$. As in classical control theory applied to ML, the aim is to select the control u that minimizes the discrepancy between the SNN output and the data. Accordingly, we define the cost function for our stochastic optimal control problem as

$$J(u) := \mathbb{E}\left[\Phi\left(X_T, \Lambda\right)\right] \,, \tag{4.10}$$

Λ being a random variable that corresponds to the target of a given input, i.e. X_0. Then the optimal control $u^\star$ is the one that solves

$$J(u^\star) = \inf_{u \in \mathcal{U}[O,T]} J(u)$$

above the class of measurable control $\mathcal{U}$.

At this point, we are able to write the optimization problem that represents the analogous of Eq. (4.6) with stochastic evolution (where also the diffusion is considered as a model parameter) but without reference to the mean field aspect of the learning procedure.

$$\begin{aligned} &\inf_{u \in L^\infty([0,T],\mathcal{U})} J(u) := \mathbb{E}\left[\Phi\left(X_T, \Lambda\right)\right] \\ &dX_t = f(X_t, u_t)dt + g(u_t)dW_t, \quad 0 \le t \le T \end{aligned} \tag{4.11}$$

Following in [7], we address the Stochastic Maximum Principle approach to solve the stochastic optimal control problem stated in (4.11). Firstly, the functional J is differentiated with respect to the control with a derivative in Gateaux sense over $[0,T]$

$$J'_u(t,u_t) = \mathbb{E}\left[f_u^i(X_t,u_t)^T Y_t + g'_u(u_t)^T Z_t\right]. \tag{4.12}$$

Then, by the martingale representation of Y_t, the following Backward SDE is introduced

$$dY_t = f_x^i(X_t,u_t^\star)^T Y_t + Z_t dW_t\,, \quad Y_T = \Phi'_x(X_T,\Lambda) \tag{4.13}$$

to model the back-propagation of the forward state process Eq. defined in (4.9) associated with the optimal control $u^\star$.

Finally, the problem is solved by the gradient descent method with step-size η_k

$$u_t^{k+1} = u_t^k - \eta_k J'_u(t,u_t^k)\,, \quad k = 0,1,2,\ldots, \quad 0 \le t \le T,. \tag{4.14}$$

The authors also provide a numerical scheme whose main benefit is to derive an estimate of the uncertainty connected to the output of this stochastic class of NNs.

We remark that here it is not possible to write the chain rule for Eq. (4.14) as previously done for Eq. (4.4) since the presence of the stochastic integral term that, differently from classical ML theory, makes the back-propagation itself a stochastic process, see Eq. (4.13). However, modern programming libraries (e.g., TensorFlow or PyTorch) perform the computation (4.14) automatically reducing the computational cost, hence allowing to go towards a mean field formulation (in terms of multiple interacting agents) of previous problems.

4.4 Mean Field Neural Network as a Mean Field Optimal Transport

In this section, we focus on the connection between SOC and OT highlighting potential symmetries specifically for a class of infinite dimensional stochastic games.

4.4.1 Optimal Transport

As seen in Section 4.3, SOC deals with finding the optimal control policy for a dynamic system in the presence of uncertainty. Conversely, OT theory focuses on finding the optimal map to transport from one distribution to another. More precisely, given two marginal distribution $\mu \in \mathcal{P}(\mathbb{R}^d)$ and $\nu \in \mathcal{P}(\mathbb{R}^d)$, the classical OT problem in the Kantorovich formulation reads

$$\inf_{\pi \in \Pi(\mu,\nu)} \int c(x.y)\pi(dx,dy) \tag{4.15}$$

where c is a cost function and $\Pi(\mu,\nu)$ corresponds to the set of couplings between μ and ν.

We focus on the setting where μ and ν are distributions computed on $\mathbb{R}^d$, i.e. $\mu \sim (X_1,\dots,X_d)$ and $\nu \sim (Y_1,\dots,Y_d)$. The Monge formulation reads

$$\inf_{T:T\#\mu=\nu} \int c(x,T(x))\mu(dx) \tag{4.16}$$

where the infimum is computed over all measurable maps $T:\mathbb{R}^d \to \mathbb{R}^d$ with the pushforward constraint $T\#\mu = \nu$.

The possibility to link a SOC problem, hence the related mathematical formulation of a specific learning procedure, to the corresponding OT formulation relies on lifting the SOC problem in a proper Wasserstein space. For example, considering the SOC problem introduced in (4.11), the stochastic process X_t described by Eq. (4.9) can be viewed as a vehicle of mass transportation under an initial measure μ_0.

We mention that there are also specific scenarios where the dynamics of the stochastic control problem can be interpreted as a mass transportation problem, provided that certain assumptions on functionals and cost are guaranteed. For example in [167, 168] and similarly in [205], the authors focus on extending an OT problem into the corresponding SOC formulation for a cost which depends on the drift and the diffusion coefficients of a continuous semimartingale and the minimization is run among all continuous semimartingales with given initial and terminal distributions.

For example, in [168] the authors consider a special form for the cost function, namely $c(x,y) = L(y-x)$ with $L(u):\mathbb{R}^d \to [0,+\infty]$ convex in u proving its equivalence to a proper

SOC problem based on the so-called graph property. Indeed, we can define an image measure as $\pi_g : \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ mapping x into $(x, g(x))$. Thus, for any measurable map $g : \mathbb{R}^d \to \mathbb{R}^d$, the following equality between the two formulations holds

$$\int_{\mathbb{R}^d} L(g(x) - x)\mu(dx) = \int_{\mathbb{R}^d \times \mathbb{R}^d} L(y - x)\pi_g(dxdy) \tag{4.17}$$

Thus, μ_g models a probability measure on $\mathbb{R}^d \times \mathbb{R}^d$ with marginals μ and ν.

For the problem stated in (4.17), we know from [209] that an optimal measure $\pi^\star$ always exists. Moreover, if the optimal measure $\pi^\star$ is supported by the graph of a measurable map, we say that the graph property holds; that is, if for any $\pi^\star$ optimal for (4.15), there exists a set Γ satisfying $\pi^\star(\Gamma) = 1$ with $\Gamma = (x, \gamma(x))$ for some measurable mapping γ that resembles the NN parameters introduced in Section 4.2 and analogously $\gamma(x)$ represents the corresponding output according to Eq. (4.1).

4.4.2 Mean Field Games

In the context of Mean Field Games (MFG), i.e. stochastic games where a large number of agents interact and influence each other, it is particularly explicitable the link between SOC and OT, specifically according to the variational formulation of MFG which is directly linked to the dynamic formulation of OT by Benamou and Brenier, see, e.g., [17] for an in–depth analysis.

In Section 4.2 we focus on deterministic evolution by means of an Eq. (4.5) with the mean field interactions captured by the loss function as an expectation given a known joint measure μ between input-target in the corresponding Mean Field Optimal Problem (4.22). On the other hand, in Section 4.3, we introduce the stochastic process in Eq. (4.8) and state the learning problem as a SOC as shown in Eq. (4.10) without focusing on the interaction during the evolution but looking at just a single trajectory. As the final step, the further natural step relies on extending the previous equation to a McKean-Vlasov setting where the dynamic of a random variable X depends on the other N random variables by the mean of the distribution in order to merge the two scenarios presented in Sec. 4.2 and Sec. 4.3 while extending the problem stated in (4.10) by allowing the presence of a mean-field term.

Indeed, instead of considering a single evolution as in Eq. (4.9), we introduce the following McKean-Vlasov SDE for N particles/agents

$$X_t^i = X_0^i + \int_0^T b\left(X_s^i, m_{X_s}^N, \theta_s\right) + \int_0^T \sigma dW_s\,, i = 1, \dots, N \tag{4.18}$$

with X_0^i being the initial states. We assume a measurable drift $b : [0,T] \times \mathcal{W}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}$, a constant diffusion σ and we define the empirical distribution $m^N_{X_s}$ as

$$m^N_{X_s} = \frac{1}{N}\sum_{j=1}^{N} \delta_{X^i_s}\,. \tag{4.19}$$

The main idea would be to model multiple SNNs generalizing the dynamic defined in (4.9), including the dependence on a mean-field term in the drift allows us to model the shared connections between the neurons of different SNNs.

At the limit $N \to \infty$, the *population of SNNs* corresponds to the evolution of a representative SNN while the empirical measure m^N tends to the probability measure m belonging to the Wasserstein space $W_2(\mathbb{R}^d)$, i.e. the space of probability measures on $\mathbb{R}^d$ with a finite second–order moment, that captures a measure of interaction among SSNs.

More precisely, we introduce the following setting, that we need to define the solution of a MFG.

- a finite time horizon $T > 0$;
- $\mathcal{Q} \subseteq \mathbb{R}^d$ is the state space;
- $\mathcal{W}_2(\mathcal{Q})$ the space of probability measure over $\mathcal{Q}$;
- $(x,m,\alpha) \in \mathcal{Q} \times \mathcal{W}_2(\mathcal{Q}) \times \mathbb{R}^k$ describes the agent state, the mean-field term and the agent control;
- $f : \mathcal{Q} \times \mathcal{W}_2(\mathcal{Q}) \times \mathbb{R}^k \to \mathbb{R}$, $(x,m,\alpha) \mapsto f(x,m,\alpha)$ and $g : \mathcal{Q} \times \mathcal{W}_2(\mathcal{Q}) \to \mathbb{R}$, $(x,m) \mapsto g(x,m)$ provides the running and, resp., the terminal cost;
- $b : \mathcal{Q} \times \mathcal{W}_2(\mathcal{Q}) \times \mathbb{R}^k \to \mathbb{R}^d$ represents the drift function ;
- $\sigma > 0$ is the volatility of the state.

Definition 4.5 (MFG equilibrium)**.** *We consider a MFG problem with a given initial distribution $m_0 \in \mathcal{W}_2(\mathcal{Q})$. A Nash equilibrium is a flow of probability measures $\hat{m} = (\hat{m}(t,\cdot))_{0\le t\le T}$ in $\mathcal{W}_2(\mathcal{Q})$ plus a feedback control $\hat{\alpha} : [0,T] \times \mathcal{Q} \to \mathbb{R}^k$ satisfying the following two conditions:*

1. *$\hat{\alpha}$ minimizes J^{MFG}_m over α:*

$$\mathbb{E}\left[\int_0^T f(X^{m,\alpha}_t, m(t,\cdot), \alpha(t, X^{m,\alpha}_t))dt + g(X^{m,\alpha}_T, m(T,\cdot))\right]$$

where $(X_t^{m,\alpha})$ *solves the SDE*

$$dX_t^{m,\alpha} = b(X_t^{m,\alpha}, m(t,\cdot), \alpha(t, X_t^{m,\alpha}))dt + \sigma dW_t$$

W *being a d-dimensional Brownian motion and* $X_0^{m,\alpha}$ *has distribution* m_0*;*

2. *for all* $t \in [0,T]$*,* $\hat{m}$ *is the probability distribution of* $X_t^{\hat{m},\hat{\alpha}}$*.*

4.5.1 Mean Field Control

Differently from MFG where players are modelled as competitors, Mean Field Control (MFC) models a framework that considers a large population of agents aiming to cooperate and optimize individual objectives. In the MFC setting, each agent cost depends on a mean field term representing the average behaviour of all agents. Accordingly, the solution of a MFC is defined in the following way:

Definition 4.6 (MFC optimum)**.** *Given* $m_0 \in \mathcal{W}_2(\mathcal{Q})$*, a feedback control* $\alpha^* : [0,T] \times \mathcal{Q} \to \mathbb{R}^k$ *is an optimal control for the MFC problem if it minimizes over* α J^{MFC} *defined by*

$$\mathbb{E}\left[\int_0^T f(X_t^\alpha, m^\alpha(t,\cdot), \alpha(t, X_t^\alpha))dt + g(X_T^\alpha, m^\alpha(T,\cdot))\right] \tag{4.20}$$

where $m^\alpha(t,\cdot)$ *is the probability distribution of the law of* X_t^α*, under the constraint that the process* $(X_t^\alpha)_{t\in[0,T]}$ *solves the following SDE of McKean-Vlasov type*

$$dX_t^\alpha = b(X_t^\alpha, m^\alpha(t,\cdot), \alpha(t, X_t^\alpha))dt + \sigma dW_t \tag{4.21}$$

with X_0^α *having distribution* m_0*.*

We refer to [47] for an extensive treatment of McKean-Vlasov control problems (4.20).

By considering the joint optimization problem of the entire population, MFC enables the analysis of large-scale systems with cooperative agents and provides insights into the emergence of collective behaviour. One possibility relies on stating the dynamic in Eq. (4.6) in terms of probability measures. For example, we can consider a continuity equation such as the Fokker-Planck equation to consider the evolution of the density function. Along this setting, we cite the measure-theoretical approach for NeurODE developed in [28] where the authors introduce a forward continuity equation in the space of measures with a constrained dynamic in the form of an ODE. Conversely, within the cooperative setting, we can also rely on a novel approach, named Mean Field Optimal Transport, introduced in [13], that we explore in the next paragraph.

4.6.1 Mean Field Optimal Transport

The Mean Field Optimal Transport deals with a framework where all the agents cooperate (such as in MFC) in order to minimize a total cost without terminal cost but with an additional constraint since also the final distribution is prescribed. We notice that the setting with fixed initial and terminal distributions resembles the one introduced in the Population Risk minimization problem described in Section 4.2. We follow the numerical scheme introduced in section 3.1 in [13] to approximate feedback controls, namely we introduce the following model.

Definition 4.7 (Mean Field Optimal Transport)**.** *Let $\mathbb{R}^d$, describe the state space and denote by $\mathcal{W}_2(\mathbb{R}^d)$ the set of square-integrable probability measures on $\mathbb{R}^d$. Let $f : \mathbb{R}^d \times \mathcal{W}_2(\mathbb{R}^d) \times \mathbb{R}^k \to \mathbb{R}$ be the running cost, $g : \mathbb{R}^d \times \mathcal{W}_2(\mathbb{R}^d) \to \mathbb{R}$ be the terminal cost, $b : \mathbb{R}^d \times \mathcal{W}_2(\mathbb{R}^d) \times \mathbb{R}^k \to \mathbb{R}^d$ the drift function and $\sigma \in \mathbb{R}$ the non-negative diffusion. Given two distributions ρ_0 and $\rho_T \in \mathcal{W}_2(\mathbb{R}^d)$ the aim of MFOT is to compute the optimal feedback control $v : [0,T] \times \mathbb{R}^d \to \mathbb{R}^m$ minimizing*

$$J^{MFOT} : v \mapsto \mathbb{E}\left[\int_0^T f(X_t^v, \mu^v(t), v(t, X_t^v))dt\right] \tag{4.22}$$

being $\mu^v(t)$ the distribution of process X_t^v whose dynamics is given by

$$\begin{cases} X_0^v \sim \rho_0 \qquad X_T^v \sim \rho_T \\ dX_t^v = b(X_t^v, \mu^v(t), v(t, X_t^v))dt + \sigma dW_t, \qquad t \in [0,T] \end{cases} \tag{4.23}$$

with ρ_0 and ρ_T are the prescribed initial and terminal distributions.

This type of problem incorporates mean field interactions in the drift and the running cost. Furthermore, it encompasses classical OT as a special case by considering $b(x,\mu,a) = a$, $f(x,\mu,a) = \frac{1}{2}a^T a$ and $\sigma = 0$.

The integration of MFC and OT allows to both tackle the weight optimization problem in NN, and to model the flow of information or mass between layers of neurons, while the optimal weights may be computed as the minimizers of the functional with respect to controls v

$$v^\star = \min_{v \in \mathcal{U}} J^{MFOT}(v) \tag{4.24}$$

along all the trajectories X^v,being $\mathcal{U}$ the set of admissible controls.

Thus, we look at the MFNN as a collection of identical, interchangeable, indistinguishable NNs where the dynamic of the representative agents is a generalization of a SNN (4.7) allowing a dependence on the term $\mu^v(t)$ modelling the mean field interactions. By considering the MFNN dynamic as a population of interconnected NNs, we can employ mean-field control to analyze the collective behaviour and interactions of networks, accounting for their impact on the overall network performance.

To summarize, we are looking at this novel class of NN, i.e. MFNN, as the asymptotically configuration of NNs in a cooperative setting.

We remark that the representative agent does not know the mean field interaction terms since it depends on the whole population but an approximated version can be recursively learned. For example in [13] the authors present different numerical schemes to solve MFOT:

1. Optimal control via direct approximation of controls v;
2. Deep Galerkin Method for solving a forward-backward systems of PDEs;
3. Augmented Lagrangian Method with Deep Learning exploiting the variational formulation of MFOT and the primal/dual approach.

We briefly review the direct method (1) to approximate controls of feedback type by an optimal control formulation. The controls are assumed of feedback form and can be approximated by

$$g(x,\mu) = G(\mathcal{W}_2(\mu,\rho_T)), \qquad \mu \in \mathcal{W}_2(\mathbb{R}^d) \tag{4.25}$$

where $G : \mathbb{R}_+ \to \mathbb{R}_+$ is an increasing function. The idea is to use the function in Eq. (4.25) as a penalty for being far from the target distribution ρ_T as the terminal cost to embed the problem into classical MFG/MFC literature. Intuitively, Eq. (4.25) corresponds to the infinite dimensional analogous of the loss function of the leveraged NN algorithm, being μ the final distribution that has to be close as possible in the sense of Wasserstein metric to the target distribution ρ_T.

In view of obtaining a numerically tractable version of the SDE (4.23), one may consider a classical discretization Euler-Maruyama scheme, also requiring the set of controls v to be restricted to the ones approximated by a NNs v_θ with parameters θ. Moreover, approximating the mean field term m by its finite dimensional counterpart, see Eq. (4.19), allows to develop a stable numerical algorithm, see Section 3.1 in [13] for further details, particularly w.r.t. the linked numerical implementation.

4.7.1 Other approaches for learning Mean Field function

For the sake of completeness, we also mention two different methods to deal with the approximation of mean field function that can be used in parallel with MFOT:

- the first data-driven approach, presented in [8], has been considered to solve a stochastic optimal control problem, where the unknown model parameters were estimated in real-time using a *direct filter method*. This method involves transitioning from the stochastic maximum principle to approximate the conditional probability density functions of the parameters given an observation, which is a set of random samples;

- in [187], the authors report a map that by operating over an appropriate classes of neural networks, specifically the *Bin density-based approximation* and *Cylindrical approximation*, is able to reconstruct a mapping between the Wasserstein space of probability measures and an infinite–dimensional function space on a similar setting to MFG.

4.8 Conclusions and further directions

In the present chapter, we provided an all-around overlook of methods at the intersection of parametric ML, MFC and OT. By assuming a dynamical system viewpoint we considered the deterministic, ODE–based, setting of the supervised learning problem, to then incorporate noisy components, allowing for stochastic NNs definition, hence introducing the MFOT approach. The latter, derived as the limit in the number of training data, recasts the classical learning process in a mean field optimal transport one. As a result, we gained a unified perspective for the parameters optimization process characterizing ML models with a specified learning dynamic, within the framework of OT and MFC, which may concede to efficiently handle high dimensional data sets.

We empathise that the major limitation of MFOT (4.22) concerns the fact that many of its convergence results, such as those related to corresponding forward–backward systems, still need to be verified. Nevertheless, it represents an indubitably fertile and stimulating research ground that should be enhanced since it permits the derivation of techniques that may significantly improve the robustness of algorithms, particularly when dealing with huge sets of training data potentially perturbed by random noise components, while also allowing to highlight hidden symmetries within data.

Part III

Machine Learning Applications in Finance and Economics

5 Volatility Forecasting with Hybrid Neural Networks Methods for Risk Parity Investment Strategies

5.1 Introduction

The capital market environment of the last ten years has been characterized by falling interest rate levels and rising equity markets. Furthermore, during the last 12 months, volatility levels have increased significantly as a result of rising inflation risks and growing interest rate levels in developed markets. As a consequence, efficient risk management techniques are getting increasing attention since both institutional and private investors are looking for new ways to assess the risks produced by market change.

Accordingly, we focus our attention on the so-called *risk parity portfolio* looking for new approaches to lower risk levels linked to equity markets by means of better forecasts computed by Machine Learning (ML) approaches. We remark that risk-parity techniques form the foundation of many portfolio allocation approaches for specific asset classes. These methods consider the relative risk contribution of each asset, which is typically measured by its realized volatility, to the overall risk of the portfolio.

As a consequence, an asset class with a low risk level, i.e. lower realized volatility, has a higher weight in the overall portfolio if compared with an alternative characterized by higher risk levels, respectively higher volatility. Moving from this landmark, in our analysis we consider hybrid solutions to estimate risk levels by employing both classical statistical models, namely Generalized Autoregressive Conditional Heteroscedasticity (GARCH) type models, and Neural Networks (NNs) methods, specifically different classes of Recurrent Neural Networks (RNNs) architectures, to forecast volatility levels as input factors for a portfolio allocation.

As a result, we obtain an improved performance of the forward-looking risk estimators showing better risk-return profiles for investment strategies.

Our main contributions can be summarized as follows:

- We use different RNN architectures (LSTM, GRU and a model with mixed hidden units) combined with a *faster* forecast computed by a GARCH model to provide accurate and robust volatility predictions. In particular, we use the GARCH model at the level of pre-processing stage to get information about the position of volatility clusters, while the RNNs architectures are later used to improve the precision of the forecasts themselves;
- We perform a comparison analysis between different RNNs solutions in terms of their performances as well as considering a convex ensemble of them;
- Later, we move from volatility forecast to feed a data-driven version of a Risk Parity portfolio firstly presented in [20], in a historical perspective for the reference period of October 2020 - December 2022 where we consider both weekly and monthly re-balancing procedures.

Organization of the chapter

The chapter is divided into five sections. In Section 5.2, we provide details about the general Risk Parity framework and recall basic facts about both the GARCH process and NN tools. We also present an in-depth discussion of GRU and LSTM hidden units, as well as the specific models we use in our simulation. In Section 5.3, we present the numerical results obtained for volatility forecasting of the S&P 500 index, which serves as our reference, and conduct a comparison analysis among different ML architectures. In Section 5.4, we present our findings regarding the risk-adjusted strategy in light of the obtained ML-based volatility estimates. We then perform a historical simulation of the Risk Parity portfolio and re-balancing on both a weekly and monthly basis. Finally, in Section 5.5, we provide our concluding remarks, highlighting the most interesting aspects of our research and suggesting possible future research directions.

5.1.1 Related Literature

Over the last years, ML, and in particular Artificial Neural Networks (ANNs) solutions have received increasing attention within a plethora of industrial sectors. Indeed, the fast development of efficient, reliable and relatively not expensive, hardware solutions, such as ensemble, distributed architectures or the use of combined GPUs, has offered the opportunity to implement a series of well known ML-based solutions for the solution of *practical industrial tasks*. This is the case, in particular, of the applications of forecasting solutions in the sectors of manufacturing, transportation, pharmacy as well as within the financial one.

Regarding the implications of machine learning (ML) in finance, it is nearly impossible to provide even a partial summary of the vast amount of studies that have been conducted by practitioners, mathematicians, and computer scientists from diverse backgrounds. Notably, the monographs by Cohen and Dixon serve as important landmarks in this field. Additionally, it is worth noting that there is a considerable body of literature that explores the integration of recurrent neural networks (RNNs), particularly Long Short-Term Memory (LSTM) networks, with specific technologies across a wide range of fields. Examples include the use of LSTM networks for malware threat detection in Internet of Things (IoT) [216], energy consumption forecasting [182], automated diagnosis support in healthcare [217], and the prediction of air pollutant concentrations [134].

We will continue by recalling some of the most significant contributions in forecasting realized volatility. In [173], the authors proposed a novel approach to model stochastic volatility dynamics by combining Stochastic Volatility (SV) models with Long Short-Term Memory (LSTM). This resulted in the development of the LSTM-SV model, which overcomes the short-term memory problem encountered when dealing with numerical approximations of financial time series where the volatility term is modelled through a stochastic process. The proposed model has the ability to capture non-linear dependence in the latent volatility process, providing better forecasting performance than standard SV solutions, such as approximating Heston-type models using Monte Carlo approaches.

In [156], the authors analyzed the performance of different hybrid models that combine Artificial Neural Networks (ANN) with GARCH-type models for volatility forecasting. They demonstrated that the GARCH-ANN model performs better than standard solutions for volatility forecasting in a Chinese energy market index. In [142], a similar approach was applied with a pre-processing stage to forecast price volatility levels of cryptocurrencies like Bitcoin.

In [98], a combination of a Recurrent Neural Network (RNN) with a dimension-reducing symbolic representation was applied to overcome two fundamental limitations of Machine Learning (ML) methods trained on raw numerical time series data: high sensitivity to hyperparameters and the initialization of random weights. The proposed solution was used for time series forecasting.

Finally, in [138], the authors proposed a new hybrid model that combines the LSTM model with various generalized GARCH-type models to forecast stock price volatility. The provided solution showed promising results.

Concerning the integration of LSTM with existing well-known financial method, we mention [162] where the authors integrate the Merton determinants model to forecast the credit default swap. Another example of a modification of LSTM is the modified adaptive

LSTM model (consisting of two LSTM layers followed by a pair of batch normalization (BN) layers, a dropout layer and a binary classifier) presented in [101]. We also cite [202] and [193] to have an overview of possible ensemble techniques following a comparison analysis.

More recently, in [223], the authors provide a comparison between different ML methods finding that NNs dominate linear regressions and tree models for predicting realized volatility, due to their ability to model complex uncovered interactions in time series. Further, in [86], a VolTarget approach has been proposed as a risk-adjusted procedure alternative to Risk Parity, moving from historical realized volatility values then exploiting LSTM steered solutions, see also [3, 4, 37] for further details. Previous examples are worth to be blended with efforts conducted within the so-called *Risk Parity and Budgeting* approach, see [196], as a general (set of) method(s) to treat the Risk Parity setting, namely the scenario we're dealing with.

5.2 Methodology and Neural Networks Models

Financial applications such as risk assessment and forecasting have all benefited from the widespread use of ML, particularly because of the ability of NNs-based algorithms to find connections among data points in view of reconstructing time-series patterns of interest. Throughout this section, we provide a theoretical description of models and methods that we implemented to obtain an effective and robust Risk Parity strategy.

5.2.1 The Risk Parity strategy

A Risk Parity portfolio aims at equalizing the risk allocation across asset classes by over-weighting safer assets and under-weighting riskier ones in terms of their weights in the market portfolio. There exist different ways and methods to construct a risk parity portfolio, basically depending on how we measure and define risk. It is worth mentioning that risk parity weighting does not require investors to make assumptions about expected returns representing a significant advantage over mean-variance optimization.

In this chapter, we consider a multi-asset portfolio formed by N indexes indexed by $i = \{1, ..., N\}$ whose risk we assume to be proportional to their historical volatility σ_i.

The log returns of the i^{th} asset are calculated as a time series r_t^i by

$$r_t^i = \log p_{t+1}^i - \log p_t^i. \tag{5.1}$$

with p_t^i being the value of daily closing price evaluated of index i^{th} at time t. The log returns series are used as opposed to the daily closing prices as they often fit a Gaussian distribution [6].

We measure the corresponding historical volatility σ_t^i as the standard deviation of the log return process conditioned to the σ-algebra $\mathcal{F}_{t-1}^i$ generated by log returns $r_0^i, ..., r_{t-1}^i$. Specifically, the volatility time series are computed as the standard deviation over n days, i.e. the number of samples within that time window, of the log return series, over the time window given by the interval $[\tau, \tau + n]$

$$\sigma_t^i = \sqrt{\frac{1}{n} \sum_{t=\tau}^{\tau+n} \left(r_t^i - \mu_{\tau,\tau+n}^i \right)^2}, \tag{5.2}$$

where $\mu_{\tau,\tau+n}^i = \dfrac{1}{n} \sum_{t=\tau}^{\tau+n} r_t^i$ is the mean of the log returns within a time window of n days. Throughout this project, we set n equal to 21, hence dealing with 21 days standard deviation.

Generally speaking, the marginal contribution MC_i of the i^{th} asset to the total risk of the portfolio is computed by

$$MC_t^i = w_t^i \sigma_t^i \beta_t^i \tag{5.3}$$

where the coefficients β_i depend on all the assets in the portfolio by means of the covariance matrix

$$\beta_t^i = \frac{\mathrm{Cov}(\sigma_t^i, \sigma_t^p)}{(\sigma_t^p)^2}. \tag{5.4}$$

being σ_t^p the overall volatility of the portfolio. The goal is to optimize all asset classes' weights ω^i to have equal marginal contributions MC^i for all t in the time horizon.

However, the computational complexity of the covariance matrix estimation is proportional to the square of the assets number, resulting in a computational, time consuming process. Thus, in this project, we use a simpler version of risk parity, introduced in [20], to approximate Eq. (5.3). We construct the portfolio by neglecting the contribution of coefficient β_t^i, namely by dropping the assumption of correlation between indexes. We assume each asset's weight to be proportional to the inverse of its standard deviation, i.e. the i^{th} asset weight for each time step t is computed according to

$$w_t^i = \frac{\dfrac{1}{\sigma_t^i}}{\displaystyle\sum_{l=1}^{N} \frac{1}{\sigma_t^l}} \tag{5.5}$$

for a N indexes portfolio. The result is that on a relative basis, lower volatile assets will have a higher weight than higher volatile assets, see, e.g., [20, 52].

5.2.2 The GARCH(1,1) model

Modelling asset volatility is a crucial challenge since accurate forecasts of asset returns volatility contribute to a precise risk assessment. Some financial models, such as the Black-Scholes one, have been extensively used to estimate the fair price of European-style options, often overlooking the presence of heteroskedasticity related effects by assuming a homogeneous volatility. Latter assumption produces elegant closed-form formula, despite assuming the stationarity of variance is far from being able to realistically depict financial application, especially within the *investment management arena*. To forecast prices and rates of financial instruments, in 1982 Robert F. Engle introduced the Auto-Regressive Conditional Heteroskedasticity (ARCH) model to provide a more realistic model to estimate financial market volatility, see, e.g., [99] for more details.

In this model, the log-returns r_t defined by Eq. (5.1) are computed in approximated form $\hat{r}_t$ as white noise multiplied by the historical volatility σ_t, while the conditional variance process is given by means of an autoregressive structure

$$\begin{aligned} \hat{r}_t &= \delta_t \hat{\sigma}_t, && (5.6) \\ \hat{\sigma}_t^2 &= \omega + \sum_{i=1}^{p} \alpha_i \hat{r}_{t-i} && (5.7) \end{aligned}$$

where δ_t are independent and identically distributed (i.i.d.) random variables with $\mathbb{E}(\delta_t) = 0$ and $Var(\delta_t) = 1$, independent of σ_k, for all $k \leq t$. The lag length $p \geq 0$ is itself a parameter for the model, and the case $p = 0$ represents the *basic* scenario characterized by white noise.

In 1986, Tim Bollerslev improved the ARCH model by allowing ${\sigma_t}^2$ to have its own autoregressive structure, see [27, 207] for further details. The equation for the GARCH(p,q) (generalised ARCH) model is:

$$\begin{aligned} \hat{r}_t &= \delta_t \hat{\sigma}_t, && (5.8) \\ \hat{\sigma}_t^2 &= \omega + \sum_{i=1}^{p} \alpha_i \hat{r}_{t-i} + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2. && (5.9) \end{aligned}$$

Practically, in applications, the GARCH$(1,1)$ model has become widely used in financial time series modelling because of its relatively simple implementation. Indeed, the likelihood function is easier to manage if compared with continuous time models, the latter needing numerical approximations for the solution of associated stochastic difference equations, see, e.g., [122, 137, 212], for further details. Taking advantage of the GARCH $(1,1)$ implementation agility, we use it to forecast the volatility of 4 indexes index over the period 1999–2020. More precisely, we apply the GARCH-approach within the preprocessing phase aiming to capture the so-called volatility clustering with a good compromise in terms of computational efforts (computational time needed for the task) and accuracy. Indeed, it is known, see, e.g., [60, 121], that financial time series often exhibit a behaviour where large price changes tend to cluster together, resulting in the persistence of the amplitudes of related financial movements.

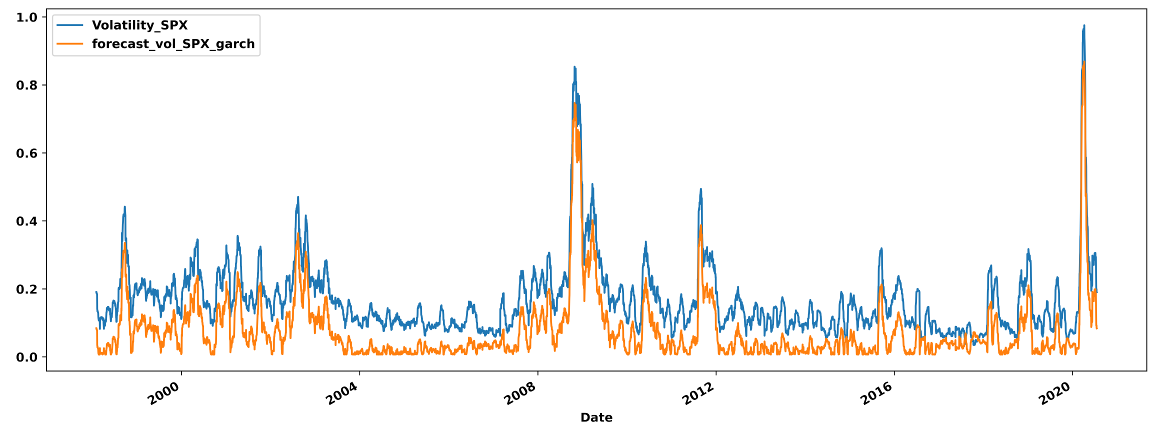


Figure 5.1: GARCH(1,1) volatility output w.r.t. S&P500 data from February 1999 to June 2020.

Using historical data, the weighted sum of the observed volatility is computed by Eq. (5.12). As shown in Figure 5.1, forecasts obtained by GARCH(1,1) are poor in terms of accuracy, while catching high variance peaks as well as corresponding volatility clusters.

5.2.3 The forecast task

We aim at deriving an algorithm able to exploit GARCH-based predictions to get information about volatility clusters, then entrusting a NN-steered solution to increase the accuracy of the forecasts.

The task of the forecast algorithm concerns approximating the non-linear mapping f represented as

$$\mathcal{Y}^i = f\left(\mathcal{X}^i\right) \tag{5.10}$$

overall the K batches indexed by i^{th} in a iterative way. At this stage, we remark that a possible limitation of the model is the fact that it assumes that there exists a function linking volatility inputs with targets over distinct batches.

We perform a rolling window mechanism to model each observation of the series on its past recent values, that called lags. Specifically, we can rewrite Eq. (5.10) at the level of each minibatch as

$$y_{[t,t+5]} = f\left(\mathbf{X}_{[t-25,t-1]}\right) \tag{5.11}$$

The choice of the *look-back window* has been empirically provided. Indeed, the best lag parameter, which turned out to be equal to 25 days, corresponds to such days providing the lowest MSE value among different tested configurations.

We tried different lag parameter values, namely 25, 50 and 100 days, obtaining rather similar results in terms of errors, hence proving the stability of the proposed algorithm w.r.t the look-back window. Indeed, the predictability of a historical series depends largely on its most recent values, which are considered to be the most important for the network during the training phase. This fact is supported by empirical evidence and is widely recognized in the field.

We report in Figure 5.2 the rolling mechanism used to apply the Supervised Learning procedure for time series.

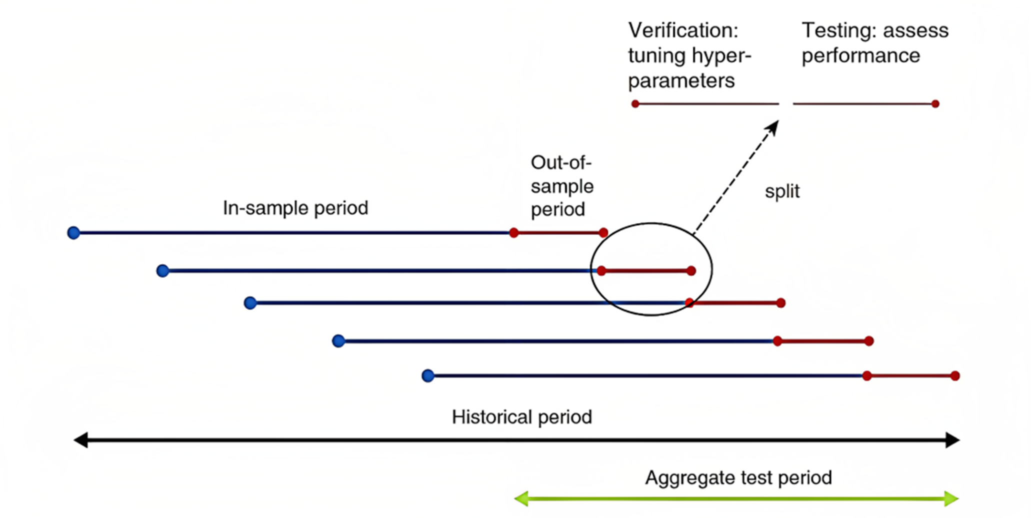


Figure 5.2: Rolling Mechanism with the in-sample period equals to 25 days and the out-of-sample period equals to 5 days.

In the input $\mathbf{X}$, then subdivided into a training and a testing set, we consider supplementary features, hence going beyond the time series of historical volatility. Specifically, the final dataset consists of the following time series:

- historical volatility σ computed by Eq. (5.2);
- the volatility forecast $\hat{\sigma}$ computed via the GARCH(1,1) model. Equation (5.8) with

$p = q = 1$ reads

$$\hat{\sigma}_t^2 \quad = \quad \alpha\sigma_{t-1}^2 + \beta\delta_{t-1}^2 + \omega \tag{5.12}$$

that is adapted since it depends on volatility values at time $(t-1)$);

- weekly averaged volatility, i.e. the mean of the volatility of the last 5 business days;
- monthly averaged volatility, i.e. the mean of the volatility of the last 20 business days;

In Fig. 5.3, we provide a graphical representation for a subset of the training set for the considered horizon (January 2000 - December 2020). Then, we evaluate the model over a test set corresponding from October 2020 to December 2022.

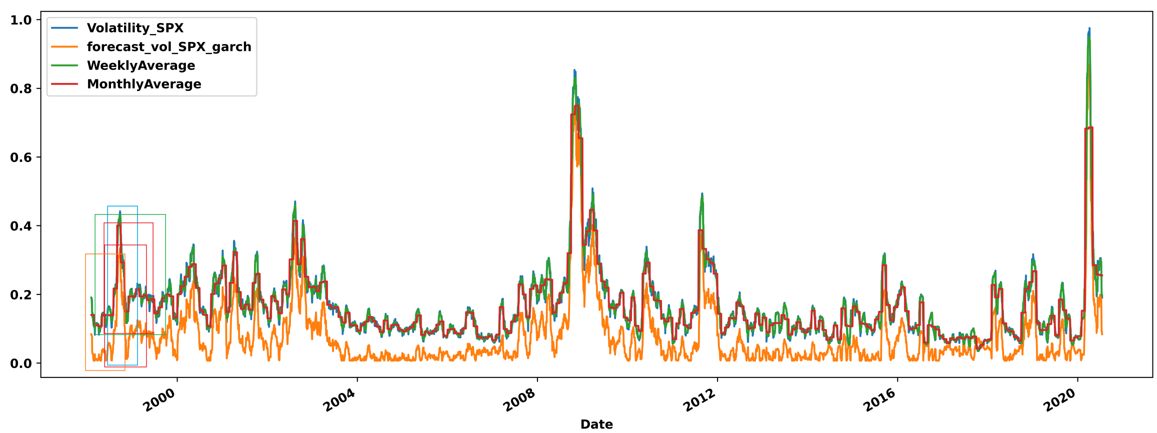


Figure 5.3: An extract from the Training Set for S& P 500

5.2.4 RNNs

Neural network models (NNs) are essentially inspired by the functioning of the human brain, in terms of mimicking the connections and transmissions of signals among neurons. Indeed, a hidden unit inside a NN essentially mimics the output transmission of real neurons to other neurons.

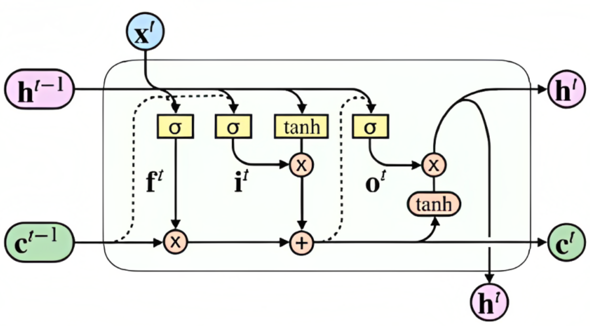
In terms of their applications within the machine learning scenario, the main advantage of neural networks lies in their ability to reconstruct non-linear functions while recognizing internal patterns within data.

Most used NNs-based forecasting solutions when dealing with time-series data are those belonging to the large family of the so-called Recurrent Neural Network (RNN) approaches, which are widely applied for stock market projections as well as for sales forecasting, see, e.g., [91, 176, 199, 215]. In this work, we focus on two specific RNN variants, i.e. LSTM and GRU.

Our work has been focused on exploiting both the GRU and the LSTM approach since they have been shown to be very effective when dealing with time series. For the sake of completeness, ANNs-based solutions have been also tried, which have been discarded because of poorly obtained results. We have planned to implement Temporal Convolutional NNs (TCNN), see, e.g., [53, 210], type solutions in future research, particularly trying to capture possible volatility clusters and huge variations well before they really realized within real markets.

LSTM hidden units

The Long Short-Term Memory (LSTM) units were designed in 1997 by S. Hochreiter and J. Schmidhuber [124] to address long-term dependence issues. The main idea behind the LSTM architecture is to replace the recurrent neural network's (RNN) hidden layer neurons with a single set of memory cells, called the "elaboration unit." Additionally, the LSTM model uses a gate structure to filter inputs and keep memory cells up to date. The gate structure includes input, forget, and output gates. Each memory cell is characterized by three sigmoid layers and a tanh layer, as shown in Figure 5.4.



- $f^t = \sigma(x^t W_f + h^{t-1} U_f + b_f)$
- $i^t = \sigma(x^t W_f + h_{t-1} U_i + b_i)$
- $o^t = \sigma(x^t W_0 + h_{t-1} U_0 + b_0)$
- $C^t = \sigma(f_t \times C^{t-1} + i^t \times \tilde{C}^t)$
- $h^t = \tanh(C^t) \times o^t$

Figure 5.4: Structure of a LSTM cell

In the context of a Long Short-Term Memory (LSTM) network, the notation is typically defined as follows: x^t represents the input vector at time t, f_t is the activation vector of the forget gate at time t, i^t is the activation vector of the input gate at time t, and o^t is the activation vector of the output gate at time t. The output vector of the LSTM unit at time t is denoted by h^t. Additionally, C^t represents the cell state vector at time t. The weight matrices and bias vector parameters of the network are denoted by W, U, and b, respectively.

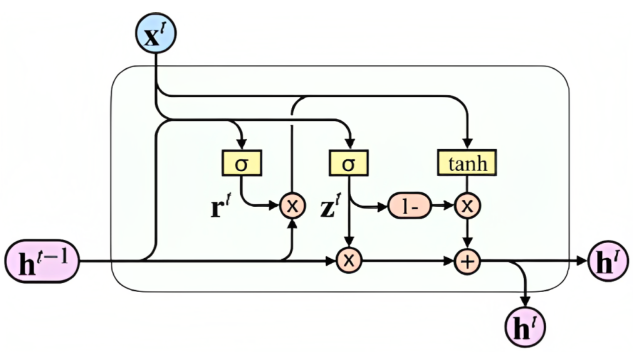
The number of parameters to be learned during the training process by a hidden layer with n LSTM units is given by:

$$\#_{parameters,LSTM} = 4 * ((n+m) * +n) * n \tag{5.13}$$

where we're considering 4 functions, namely: 3 sigmoids (f^t i^t, o^t) and 1 hyperbolic tangent (h^t), whit 1 bias parameter, while n, resp. m, represents the size of the input vector, resp. the dimension of the output vector, for a given hidden layer, see, e.g., [125, 164], for further details.

GRU hidden units

As a modification of the standard LSTM approach, Gated Recurrent Units (GRUs), as presented in [58], preserve the LSTM's ability to avoid the problem of vanishing gradients. However, they achieve this with a simpler internal structure, resulting in a more straightforward training phase, as fewer computations are required to update the internal states. The updated port controls how much of the past state information is retained in the current state, while the reset port determines whether the current state should be combined with previous information. Figure 5.5 illustrates the structure of the GRU memory cells, which are characterized as follows:



- $z^t = \sigma(x^t W^{z^t} + h^{t-1} U^z + b_z)$
- $r^t = \sigma(x^t W^{r_t} + h^{t-1} U^r + b_r)$
- $h^t = (1 - z^t) \times \tilde{h}^t + z^t \times h^{t-1}$.

Figure 5.5: Structure of a GRU cell

W^z,W^r and W being the weight matrices associated with the input vector U^z, U^r and U being the weight matrices from previous step, and b_r, b_z and b representing the biases. Moreover, in this model the logistic sigmoid function is represented by σ, the reset gate is given by r^t, and the update gate is denoted by z^t.

According to [80], the total number of parameters in the GRU-RNN for a given layer equals

$$\#_{parameters,GRU} = 3(n^2 + mn + 2n)\,, \tag{5.14}$$

where n, resp. m, represents the size of the input, resp. of the output, vector.

Neural Networks architectures

The key difference between GRU and LSTM is that a GRU unit has two gates, namely the reset and the update gate, whereas an LSTM has three gates, namely input, output and the forget gate. Nevertheless, it has been shown that the performance of a GRU is comparable to a LSTM if limited to specific applications, like, e.g., classification problems, see [159], or prediction tasks, see [164].

In this project, we consider the following NN architectures:

1. *LSTM-GARCH*: 2 layers with LSTM units plus a 1 dense layer;

2. *GRU-GARCH*: 2 layers with GRU hidden units plus 1 dense layer;

3. *LSTM-GRU-GARCH*: a mixed architecture with 3 layers combining different units: a first one with LSTM units followed by a layer with GRU units and a dense layer.

4. An *ENSEMBLE* model aggregating the results from all previous NNs.

We remark that the presence of *GARCH* into names would emphasize that the considered models are the hybrid ones since the presence of the GARCH(1,1) forecast into the NN inputs.

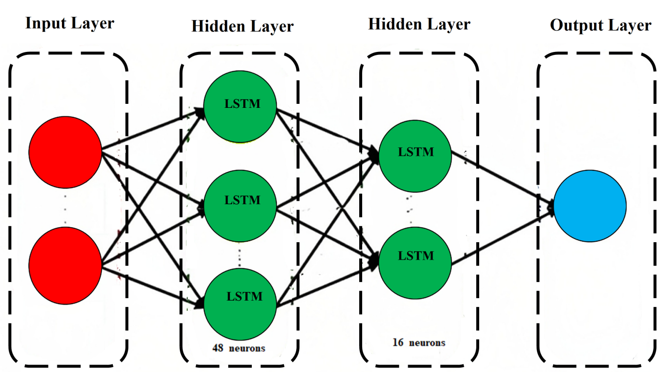


Figure 5.6: Graphical representation of the LSTM-GARCH model.

We selected the set of parameters in Table 1 through a trial-and-error procedure. Our goal was to optimize the model's predictive accuracy while also ensuring a comparable number of parameters across different models. This allowed us to perform a consistent analysis across all models.

The number of parameters is computed accordingly to Eq. (5.13), for layers with LSTM units, e.g. $4*((48+16)*16+16) = 4160)$, and via Eq. (5.14), for layers with GRU units, leading to $3*(16^2+16*64+2*16) = 3936$.

We recall that for Dense layer with n neurons, the number of parameters involved equals

$$\#_{parameters,DENSE} = n*(\#_{previouslayer}+1)$$

giving $85 = 5 * (16 + 1)$.

Table 5.1: Neural Network architecture and parameters

	LSTM-GARCH	GRU-GARCH	LSTM-GRU-GARCH
First Layer Units	LSTM	GRU	LSTM
First Layer # Neurons	48	48	48
First Layer # Hidden parameters	10176	7776	10176
Second Layer Hidden Units	LSTM	GRU	GRU
Second Layer # Neurons	16	16	16
Second Layer # Hidden parameters	4160	3168	3936
Dense Layer # Hidden parameters	85	85	85
Total # Hidden parameters	14421	11029	13429
Dropout	0.1	0.1	0.1
Learning Rate	0.01	0.01	0.01
Batch Size	32	32	32
Epochs	80	80	80

The number of parameters is selected empirically using a trial-and-error approach to optimize predictions and ensure consistency in the analysis across different models, considering a training dataset size of $22,604$ entries. It should be noted that there are alternative methods available for parameter selection. For example, grid search involves testing all possible combinations of predefined parameter values and selecting the best-performing set based on a performance metric. Another alternative is random search, which involves randomly selecting parameter values from predefined ranges and evaluating the model for each combination. Bayesian optimization uses probabilistic models to determine the most promising parameter values for improving performance. Gradient-based optimization involves iteratively updating parameter values using gradients of the objective function, while evolutionary algorithms generate a population of parameter sets and iteratively mutate the best-performing sets to create a new population. The choice of method depends on the specific problem and available resources. All the aforementioned approaches require proper procedures to be implemented, and we plan to use a selection of these methods in future research. The results for the present case are sufficient for our purposes, even though they required a time-consuming activity due to the specific exploratory nature of the research.

Measuring predictions errors

To evaluate the model prediction performances of the test error, we use the following error metrics

- Mean Square Error (MSE)

$$MSE = \frac{1}{n}\sum_{j=1}^{n}(\mathcal{Y}_j - \hat{\mathcal{Y}}_j)^2\,, \tag{5.15}$$

- Mean Absolute Error (MAE)

$$MAE = \frac{1}{n}\sum_{j=1}^{n}|\mathcal{Y}_j - \hat{\mathcal{Y}}_j|\,, \tag{5.16}$$

- Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{1}{n}\sum_{j=1}^{n}\frac{|\mathcal{Y}_j - \hat{\mathcal{Y}}_j|}{|\mathcal{Y}_j|}\,, \tag{5.17}$$

where n is the number of samples used in the test set, $\mathcal{Y}_j$ represents the real data value and $\hat{\mathcal{Y}}_j$ shows the forecast, i.e. the NN output from Eq. (5.11).

We considered two different loss functions during the training process, namely Mean Squared Error (MSE) and Mean Absolute Error (MAE). It is worth noting that using MSE implies giving more weight to larger errors than smaller ones, due to the squaring operation. This can lead to a skewed error estimate, where the contribution of outliers is overvalued. In contrast, MAE considers the absolute value of errors, so all errors are weighted on the same linear scale. This provides a more generic measure of error without overvaluing outliers. However, we observed that using MAE as the loss function generally results in higher test errors, since it does not penalize large discrepancies caused by outliers.

5.2.5 Ensemble of Neural Networks

An ensemble method relies on a learning paradigm where many NNs-based architectures are jointly used to solve a problem as, e.g., shown by Fig. 5.7, where we illustrate our (ensemble) forecasting scheme, see, e.g., [29], [193] and [202] for further details.

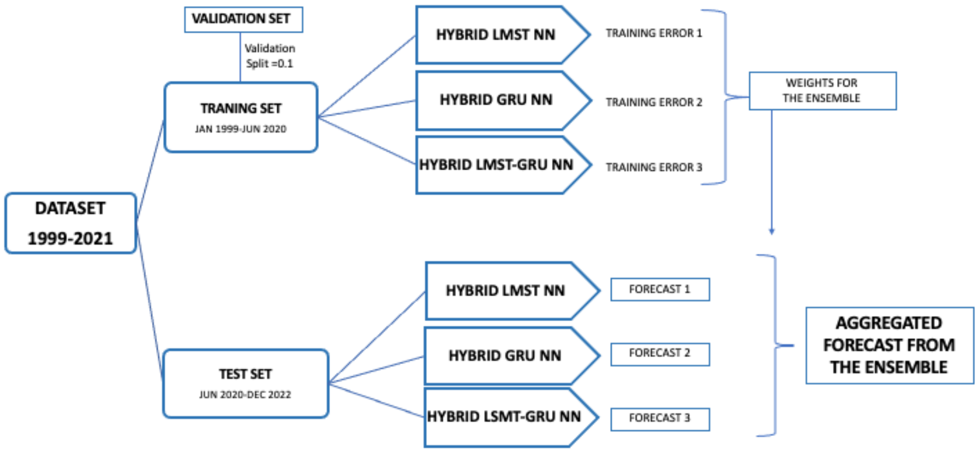


Figure 5.7: Scheme of the aggregated Ensemble forecast.

Extension of the Bates & Granger Technique. Based on the portfolio diversification theory, the Bates & Granger Technique ignores the correlation between forecasting models. The idea is to use the estimated mean squared forecast errors to obtain the combining weights, see, [126, 55] for further details. Following the latter approach, we consider a convex combination of forecasts obtained by different models, associated coefficients being weighted according to training loss, with a higher coefficient assigned to the network with lower test error. We use a power of grade 5, hence penalizing more terms with higher errors, by taking

$$\mathrm{W_{LG}} = \frac{(1/MSE_{LG})^2}{\text{Total Error}}, \quad \mathrm{W_{GG}} = \frac{(1/\mathrm{MSE_{GG}})^2}{\text{Total Error}}, \quad \mathrm{W_{GLG}} = \frac{(1/\mathrm{MSE_{GLG}})^2}{\text{Total Error}} \tag{5.18}$$

where LG stands for LSTM-GARCH, GG for GRU- GARCH, and GLG for GRU-LSTM-GARCH model, respectively, and

$$Total\ Error = \left(\frac{1}{MSE_{LG}}\right)^5 + \left(\frac{1}{MSE_{GLG}}\right)^5 + \left(\frac{1}{MSE_{GG}}\right)^5. \tag{5.19}$$

5.3 Data analysis and Results

In this section, we present the numerical results. We calibrate the model with real market data using the daily closing prices of 4 indexes: Standard and Poor's 500, STOXX Europe 600, Hang Seng Index and Nikkei 225.

During a time interval of 20 years, more precisely:

- *Training Set*: January 2000 - June 2020 with validation split set at 0.1. In Fig. 5.8 we can see the training performance for the NN model with GARCH preprocessing and GRU hidden units;
- *Test Set*: October 2020 - December 2022, corresponding also to the period of the historical simulation for portfolio ruled by the Risk Parity technique.

5.3.1 General Information about the Dataset

It is essential to emphasize that a Risk Parity strategy usually deals with portfolios made by indexes, bonds and commodities because of the different intrinsic risks of these instruments. In our setting, we focused on a 4-indexes, representing relevant both USA and international markets proxies, risk parity portfolio consisting of:

1. S&P500 index or Standard and Poor's 500 is the market capitalization-weighted index of 500 large companies listed on stock exchanges in the United States. It is widely recognized as one of the most accurate indicators of the performance of major American stocks, and by extension, the stock market as a whole.
2. STOXX Europe 600, or Stoxx600, is the European stock index comprising a fixed number of 600 components representing major, mid, and small-capitalization companies from 17 European nations, also with exposure to Great Britain, Switzerland and the Scandinavian countries. Given the wide market exposure, the STOXX Europe 600 index is often considered the European equivalent of the S&P 500 index;
3. Hang Seng Index as a proxy for Asia ex Japan (HSI) is a free-float-adjusted market capitalization-weighted index. It is the primary gauge of Hong Kong's stock market performance since it records and monitors the daily changes in the stock prices of the major businesses listed on the city's stock exchange;
4. Nikkei is an abbreviation for Japan's Nikkei 225 Stock Average, being a price-weighted index comprised of the Tokyo Stock Exchange's top 225 blue-chip companies, therefore the Nikkei is similar to the USA index called Dow Jones Industrial Average (DJIA).

5.3.2 Calibration results from the training set

For the sake of simplicity, both for the training results and for the predictions in Section 5.3.3, we present plots and results related to S&P 500 index, ignoring other indexes since they follow a similar pattern over years.

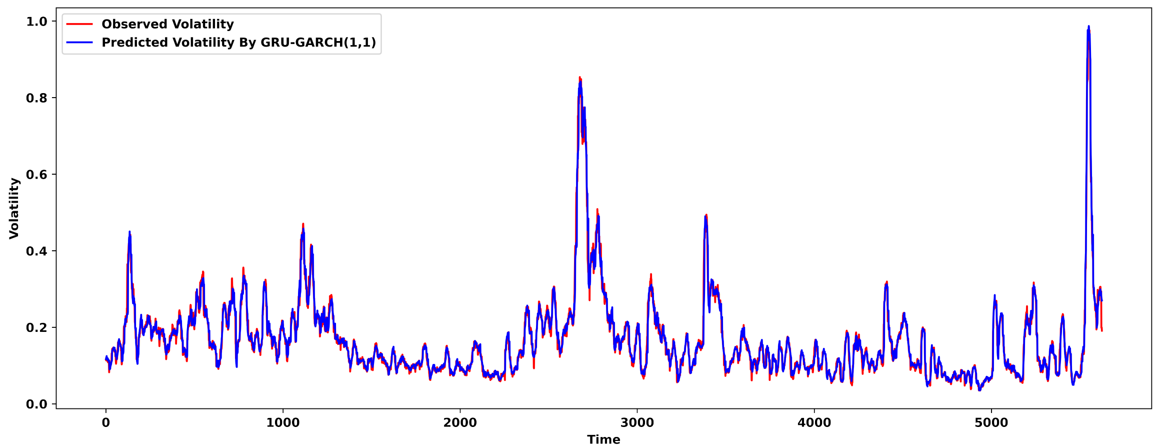


Figure 5.8: Calibration from the training set with GRU-GARCH model for the S&P 500 index.

We can see from the second and third pictures from Fig. 5.9, which correspond to LSTM-GRU-GARCH and GRU-GARCH, that we obtain comparable training and validation errors that ensure good behaviour for the model. Conversely, we obtain for LSTM-GARCH a training loss higher than the validation one as it appears from the first loss function in Fig. 5.9. This behaviour can be due to the fact that the validation set does not consider dropout.

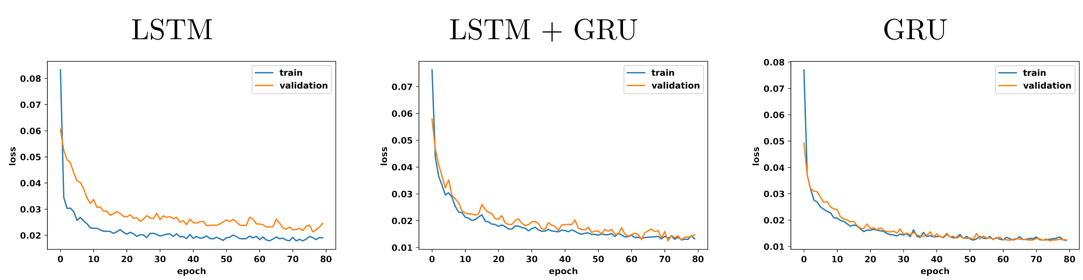


Figure 5.9: S& P500, Validation and Train errors for all models with 80 epochs

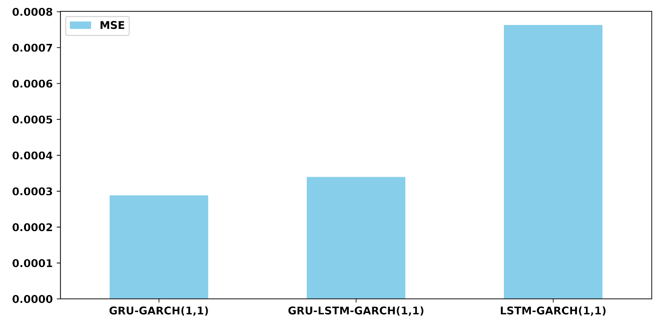


Figure 5.10: Ranked training MSE for training set of the S&P index

Models	**MSE**	**RMSE**	**MAE**	**MAPE** [%]
GARCH	0.00908571	0.0953190	0.09150850	345.56987
LSTM-GARCH	0.00076309	0.0953190	0.01614862	10.246528
GRU-GARCH	**0.0002882**	**0.095319**	**0.0915085**	**6.329703**
LSTM-GRU-GARCH	0.00033954	0.0953190	0.01105917	6.7951889

Table 5.2: Training Errors for 2000-2020 simulation, for S&P500

As we highlight within the previous table, on the training set the highest MSE error has been obtained by the GARCH model. The 345% MAPE in Table 5.2 is due to the bad-performance related to the GARCH forecast as clearly visible in the plot of Fig. 5.10. The predictive performance is really improved by adding NN with the best model being the GRU-GARCH model.

5.3.3 Forecast results of the hybrid algorithms for the S&P500 volatility

In this section, we present daily, weekly and monthly frequency forecasts for historical volatility for 2021 - July 2022.

Daily Forecast

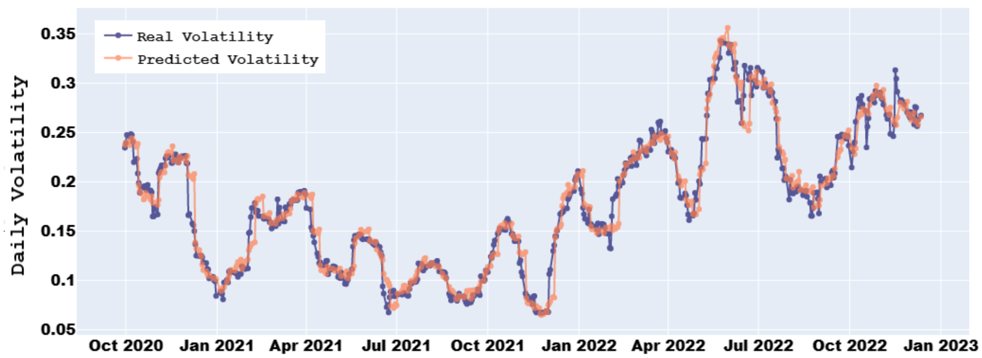


Figure 5.11: Daily 21 days RV vs predicted volatility for Ensemble model

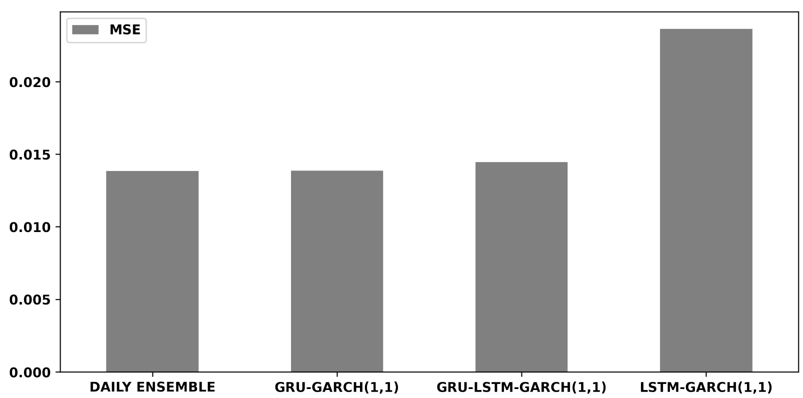


Figure 5.12: Ranked MSE errors on daily forecast for S&P500.

As highlighted in Table 5.3, on the test set for daily forecast, the highest MSE error has been obtained by the LSTM-GARCH model, and modifying this model by NN models significantly reduces the MSE error. Best performances have been obtained by the GRU-GARCH solution, the latter showing the lowest MSE error.

Model	MSE	RMSE	MAE	MAPE
LSTM-GARCH	0.023628595	0.153715	0.0177804	12.043454
GRU-GARCH	0.013872029	0.120729	**0.0013675**	**5.7662236**
LSTM-GRU-GARCH	0.014445018	0.120187	0.0059371	6.6963885
ENSEMBLE	**0.01384343**	**0.11765**	0.009488	5.922075

Table 5.3: ERRORS ON DAILY FORECAST, for S&P500

Weekly Forecast

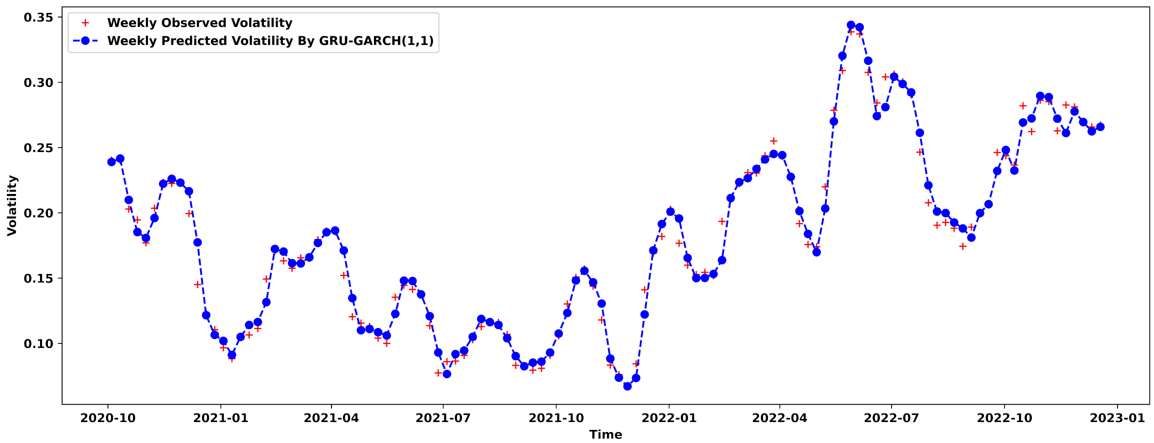


Figure 5.13: Weekly predictions for the GRU-GARCH(1,1) model

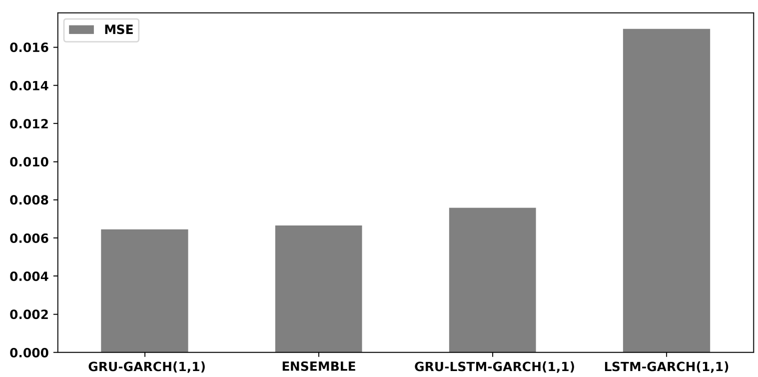


Figure 5.14: MSE errors on weekly forecast for S&P 500 index

Models	**MSE**	**RMSE**	**MAE**	**MAPE**
LSTM-GARCH	0.021117	0.145316	0.016951	10.92226
GRU-GARCH	**0.00890**	**0.09435**	**0.00644**	**4.03471**
LSTM-GRU-GARCH	0.009724	0.098611	0.007565	5.036304
ENSEMBLE	0.009004	0.094884	0.006649	4.258328

Table 5.4: ERRORS ON WEEKLY FORECAST, for S&P500

As highlighted in the Table 5.4, the highest MSE error has been obtained by the LSTM-GARCH model. Modifying the latter by NN models reduces the MSE error. In particular, the best performances have been those obtained by the GRU-GARCH solution showing the lowest MSE error.

Monthly Forecast

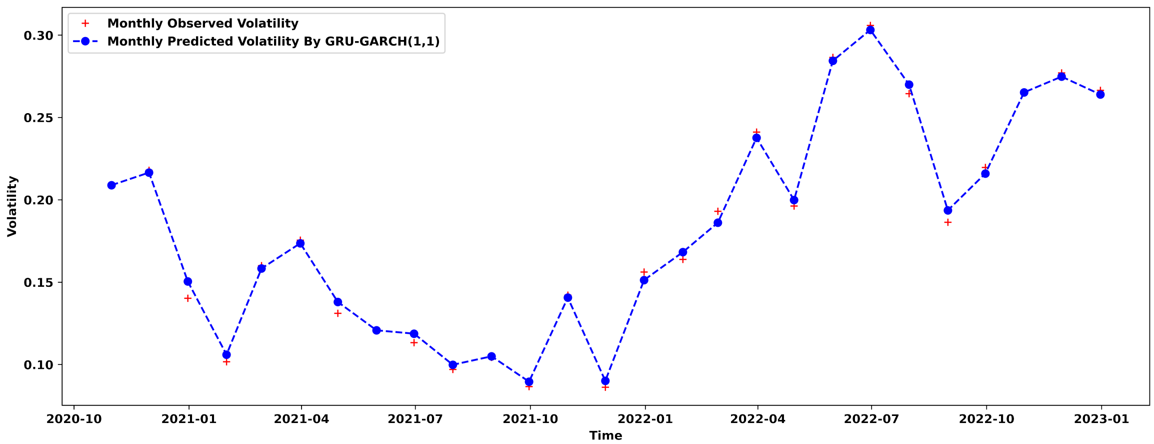


Figure 5.15: Monthly predictions for GARCH-GRU model

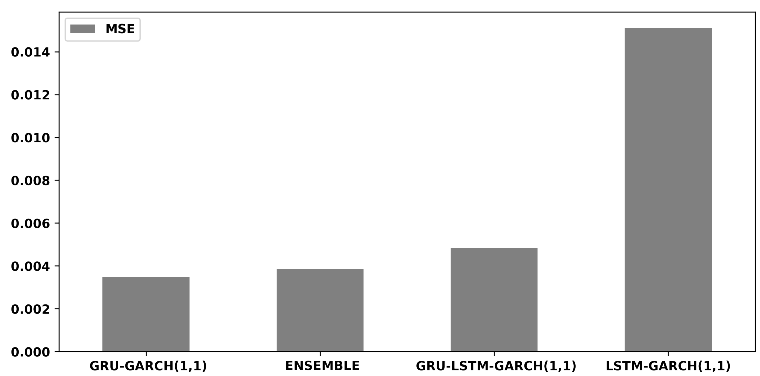


Figure 5.16: MSE errors on monthly forecast

Models	**MSE**	**RMSE**	**MAE**	**MAPE**
LSTM-GARCH	0.0177986	0.133411	0.01510	9.380881
GRU-GARCH	**0.004206**	**0.06845**	**0.00346**	**2.25129**
LSTM-GRU-GARCH	0.0054082	0.073540	0.004821	3.379187
ENSEMBLE	0.0044469	0.066685	0.006649	2.600597

Table 5.5: ERRORS ON MONTHLY FORECAST, for S&P 500

As highlighted in Table 5.5, the highest MSE on the test set for monthly forecast has been obtained by the LSTM-GARCH model, while the best results are related to the GRU-GARCH model, showing the lowest MSE error.

5.3.4 Discussion of the Results

The main improvements in terms of MSE-decreasing have been obtained by implementing the following:

- enlarging the space of features: the initial model only uses the daily volatility as input, while by including the weekly and monthly means as features, the model predictions ended up being smoother.

- Optimizing the *Dropout* hyperparameter: this regularization technique improved significantly the model performance on test data. It was noted that without it, the model performed much better on trained data, but worse in test ones.

- Adding a *Dense layer*: we note that adding as final layer a *dense layer*, whose neurons receive input from all neurons in the previous one, considerably improves the overall accuracy.

- Comparing LSTM and GRU performances: LSTMs remember longer sequences than GRUs and should outperform GRUs in tasks requiring modelling long-distance relations. On the other hand, GRUs train faster and perform better than LSTMs on less training data. Thus, it is crucial to select the opportune fraction of the dataset to have a proper *train-validation split*. The choice of this parameter (we have considered values from 0.08 to 0.2) directly affects the specific models giving the best predictions.

- Considering an ensemble algorithm: there are two main reasons to use an ensemble over a single model. The first one is related to performance: the ensemble has good predictions behind only the GARCH-GRU representing our benchmark, as shown in Figures 5.12 and 5.17. The second advantage deals with robustness since an ensemble reduces the spread or dispersion of the predictions and model performance, limiting the bias connected to a single model.

Moreover, comparing Tables 5.3 and 5.4, we can see how monthly predictions portfolios are better than the weekly ones in terms of overall percentage error, with a MAPE of $2,25\%$ and $4,03\%$ respectively. This behaviour may be explained by the fact that averaging over a monthly basis provides a smoothing effect over the noise that directly affects the volatility prediction.

To summarize, as we can clearly see from Tables 5.2, 5.3 and 5.5, reporting the forecast errors, and from the box plot in Figure 5.17, the GRU-GARCH solution outperforms other hybrid models in terms of overall MSE. We also obtain a good performance from the ensemble model, dealing with the aggregated forecast.

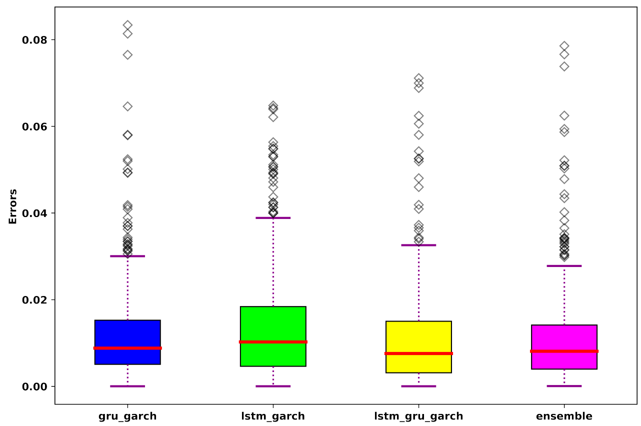


Figure 5.17: Box plot of the comparison for the MSE on the daily forecast for different NNs architecture with S&P500 data

We report the errors for different simulations in the aggregated results presented in the Appendix 6.6.

5.4 Risk parity portfolio

In this section, we compare an equally-weighted indexes (S& 500, STOXX Europe 600, Hang Seng, Nikkei) basket with a risk-parity weighted portfolio of the considered indexes over the same reference period October 2020 - December 2022.

We use a bottom-up data-driven approach to determine the weights for the portfolio assets. We refer to Section 5.2.1 for a detailed description of the re-balancing rules.
We begin by calculating the long-term historical volatility for each index, as discussed in Subsection 5.2.3. Volatility time series are then used as input to update the weights according to Eq. (5.5).

For the sake of simplicity, we do not consider transaction costs.

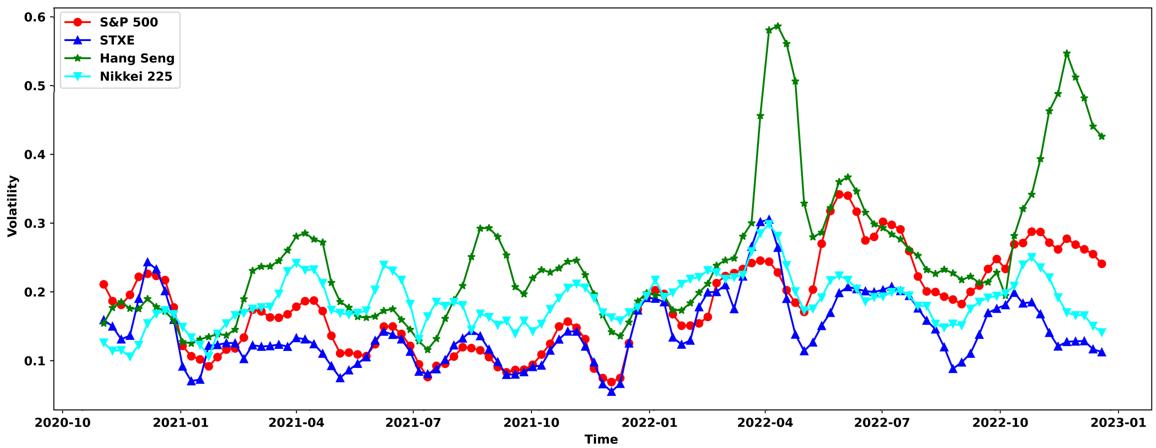


Figure 5.18: Weekly Forecast for the 4 considered indexes

The contract position weight within each asset class is proportional to the inverse of its standard deviation, see Equation (5.5).

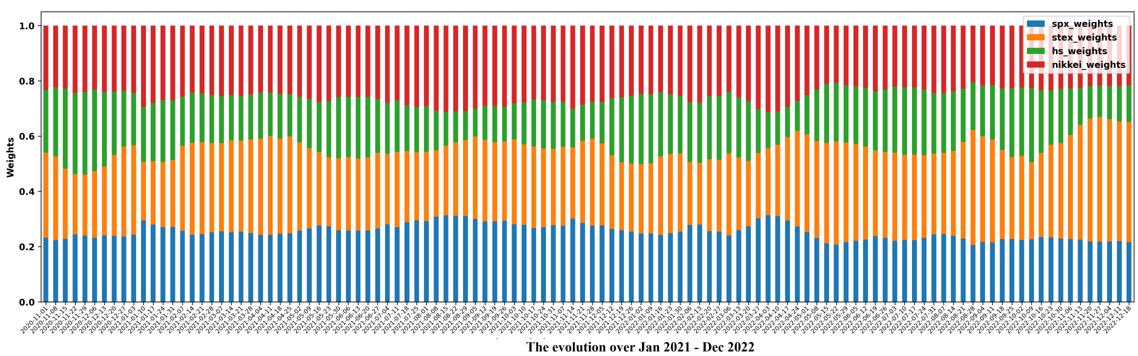


Figure 5.19: Weekly Weights for the 4 considered indexes

According to weights of Fig. 5.19, we compute the return in Fig. 5.20 of an investment over the considered horizon of 112 weeks (October 2020 - December 2022).

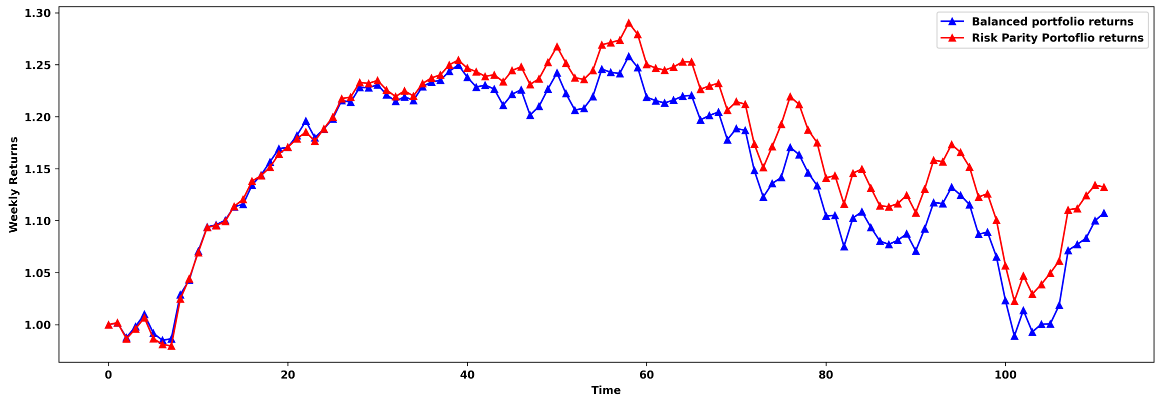


Figure 5.20: Weekly rebalanced portfolio

In Fig. 5.21, we report the monthly volatility used to compute the corresponding weights reported in Fig. 5.22.

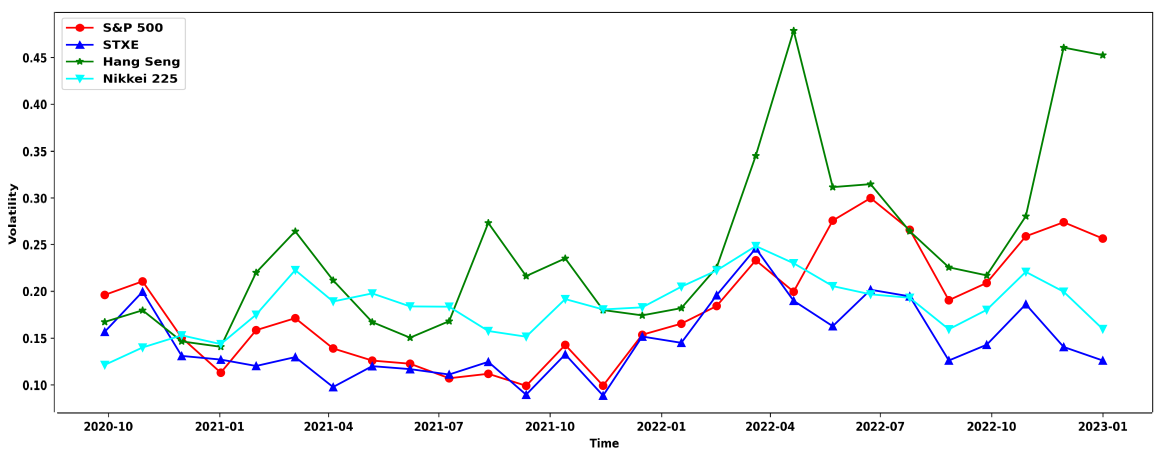


Figure 5.21: Monthly Forecast for the 4 considered indexes

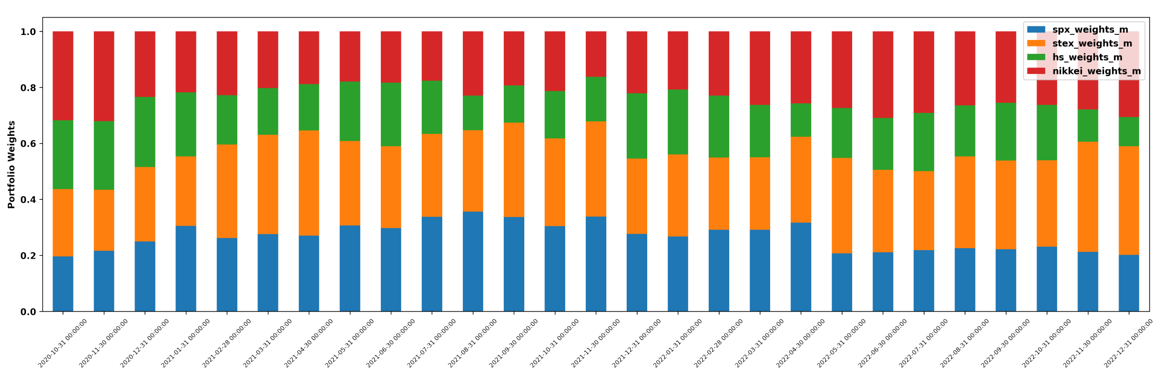


Figure 5.22: Monthly Weights for the 4 considered indexes

Fig. 5.23 reports a plot of returns for a monthly re-balanced portfolio over the considered 27 months. We also compare its performance to an equally balanced one.

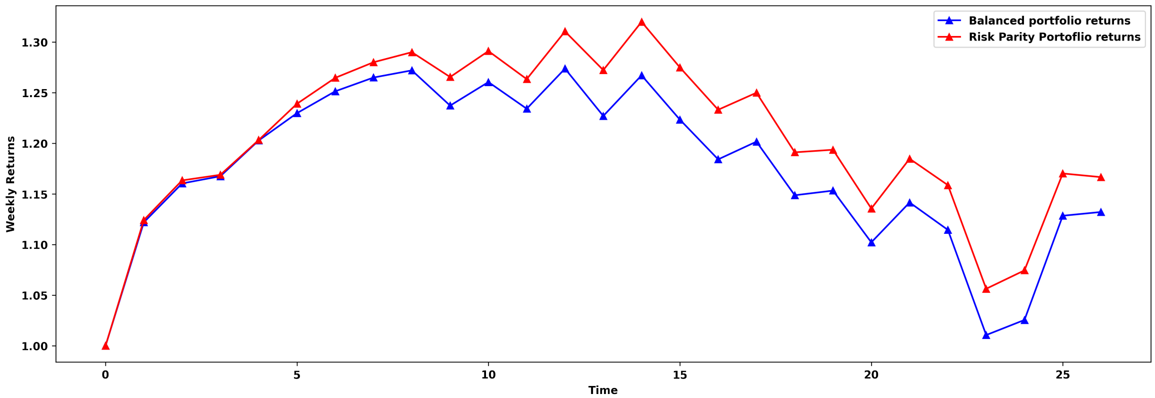


Figure 5.23: Monthly re-balanced portfolio

5.4.1 Performance Analysis

We report some financial considerations from the analysis of the portfolio performance:

- According to Fig. 5.20 and 5.23 where we consider the evolution of different portfolios, we can see how Risk Parity techniques outperformed equally balanced portfolios over the considered period October 2020 - December 2022. The latter suggests that a diversification based on equalizing risk is able to offer good protection for investments in periods of market turbulence.

- Figures 5.19 and 5.22 suggest a possible explanation of how the discrete allocation of the Risk Parity mechanism can benefit overall returns. The real advantages of risk parity appear only at the end of the first year. For example, we can see that during 2021 positions on the American and European indices are strengthened up to more than 70% of the allocation in December 2021. While for the previous period with the redistribution mechanism not rewarding any particular stocks, the difference in returns between the two portfolios does not grow considerably as it happens during the peaks of volatility within the 2022 year.

5.5 Conclusions

We presented a hybrid method to compute volatility forecast to perform a risk-controlled strategy over a portfolio of 4 indexes, namely S&P 500, STOXX Europe 600, Hang Seng, Nikkei, as a proxy of different international markets.

We prove how the version of Risk Parity proposed by [20], moving from Eq. (5.5), effectively produces encouraging results beyond its simplicity, also suggesting to development of more advanced techniques to include cross-correlation of assets in addition to individual volatility. Considering an ensemble method helps to achieve an improvement in terms of the accuracy of predictions.

We emphasize that our findings remain robust when dealing with different indexes, providing empirical evidence for a universal volatility mechanism among indexed.

Eventually, we present some future research directions:

- When dealing with real applications, the standard deviation is a dangerously limited estimate of the true risk of an asset class. We believe the model could benefit from using other risk measures (such as sampled entropy or cross-entropy) as well as improving the risk analysis;

- We would like to analyze other NN algorithms such as Temporal CNNs or Regression Trees to analyze the problem even if the considered models already give acceptable results.

- We assume a portfolio composed only of indexes. It is worth mentioning that dealing with different assets, e.g., bonds, futures, and options, may increase the effectiveness of the proposed methodology. It would also raise the potential diversification benefits of a Risk Parity strategy in terms of risk efficiency, namely the risk-adjusted returns.

- Finally, in [86] the authors consider a portfolio analysis focused on a VolTarget strategy that, moving from volatility predictions, iteratively balances a portfolio between a risky investment and a risk-less one. It might be worthy of further investigation running a comparison analysis between VolTarget and Risk Parity to understand analogies over long-term horizons and, moreover, to consider a mixed algorithm that includes both the 2 risk-controlled strategies.

6 Single Agent Reinforcement Learning for Bidding Strategy Optimization in the Day Ahead Energy Market

6.1 Introduction

The price of electricity in the European market can be very unstable since it is affected by various modes of production from different sources, each with its own constraints in terms of weather, production volume, and storage difficulty. From a financial, economic and ecological standpoint, forecasting market movements and the situation of the following day's market is now a critical issue in order to maximize power output and help the energy transition process. In general, day-ahead energy markets use a simple auction paradigm: operators on the supply and demand side submit bids that include a quantity (q) and a price per unit (p) for the trading period(s) (typically hours) of the following day. In this basic setup, the intersection point of the demand and supply curves is sought, which ensures a balance of consumption and production while determining the Marginal Clearing Price (MCP). To facilitate the integration of European power markets, a market-clearing algorithm called EUPHEMIA (Pan-European Hybrid Electricity Market Integration Algorithm) has been developed among European power exchanges. It is a strictly economic dispatch algorithm and provides several options and products to market participants.

The aim of this project is to develop an algorithm to solve the stochastic optimal control problem related to an operator aiming at maximizing her profit playing in an electrical market regulated by the Euphemia algorithm. The optimization problem is solved by Reinforcement Learning (RL) where an agent, i.e. the operator, interacts with a stochastic environment that is represented by the values of electrical prices. Rather than directly looking at her profit, the agent performs an action by selecting a proper offering curve and receives a reward based on the new state of the environment, i.e. the prices of the next day from which the reward is computed. The aim consists of computing the (deterministic) optimal policy that links to a given state (historical prices) the best action in terms of maximizing the cumulative discounted reward.

The chapter is organized as follows: in Sec. 6.2 we present the day ahead energy market and the related auction algorithm and we review some popular RL methods; in Sec. 6.3 we introduce the theoretical setting of the single agent; in Sec. 6.4 we analyse the Deep Deterministic Policy Gradient (DDPG) method and its adjustment into the electricity auction framework; in Sec. 6.5 we present some numerical simulations, we review some modelling assumptions and we report some considerations and limitations of the algorithm. We conclude the chapter with Sec. 6.6 sketching future directions that may employ this algorithm as a reference starting point.

6.2 Background and Literature review

Electricity auctions are a crucial feature of power markets because they make it possible for generators, retailers, and other market participants to trade electricity. The use of Reinforcement Learning (RL) approaches to boost market participants' performance and the general effectiveness of power auctions has gained popularity in recent years. We will explore the main issues and techniques while providing a brief summary of the state-of-the-art in RL for electricity auctions.

6.2.1 Energy market

Modern energy networks rely heavily on day-ahead electricity markets, which offer a framework for the effective purchase of electricity for delivery the next day. The unitary MCP of energy, also known as PUN (Prezzo Unitario Nazionale) in the Italian market, is established in these markets by a complicated bidding procedure that involves the interaction of several market participants, including producers, consumers and traders.

The unitary price represents the cost of producing the last unit of electricity needed to meet demand and is set at the intersection of the supply and demand curves.

Accurately predicting future unitary prices in day-ahead markets remains a challenging problem due to the dynamic nature of the electricity system and the complexity of the bidding process. Unitary prices set by electricity auctions can be determined in various ways, including the aforementioned day-ahead markets, real-time markets, and capacity markets. Participants submit bids and offers to buy or sell electricity, and the market operator clears the market by determining the market clearing price and the amount of electricity to be traded. The Euphemia algorithm, for instance, is widely used in European day-ahead markets to allocate cross-zonal capacities and determine market clearing prices [100].

Euphemia algorithm

The Euphemia algorithm, which stands for "European Price Formation Mechanism", is a pan-European algorithm used to determine electricity prices in the day-ahead market. It was developed as a collaboration between European Power Exchanges (PXs) and Transmission System Operators (TSOs) in response to the European Union's (EU) target to create a single, integrated energy market.
The day-ahead electricity market consists of a process aiming at matching demand and supply once the offers by the operators have been submitted. The amounts of energy bought/sold in the market may come from different zones, thus the optimization of electricity allocations takes into account also the transmission capacities of the network.
The Euphemia algorithm aims to optimize social welfare, which is the sum of consumer surplus, producer surplus and congestion revenues. It takes into account several factors, such as supply and demand, transmission capacities, and market constraints, to calculate the most efficient and cost-effective allocation of energy resources and cross-border capacities. Essentially, it could be described according to the following *workflow*:

1. the algorithm starts taking input data from various sources: bids and offers from market participants, transmission capacities between bidding zones and market constraints. Bids and offers contain information about the quantity of electricity, minimum and maximum prices, and the time at which the electricity will be delivered;

2. the Euphemia algorithm processes the bids and offers to create order books for each bidding zone. The order books list the bids and offers in descending order of price, thus creating a demand curve (bids) and a supply curve (offers);

3. the algorithm matches the demand and supply curves for each bidding zone: the intersection of these curves determines the MCP and the cleared volume of electricity for each zone;

4. later the algorithm optimally allocates the available cross-zonal capacities between bidding zones. This process takes into account the transmission capacities and other constraints to minimize congestion and maximize the efficient use of resources.

5. if there is congestion in the transmission network, the algorithm will re-optimize the allocation of resources and cross-zonal capacities. This may result in price differences between bidding zones, known as congestion rent, which is used to cover the costs of congestion management.

6. finally, the Euphemia algorithm calculates the day-ahead prices for each bidding zone based on the results of the clearing process and the allocated cross-zonal capacities

Its ability to optimally allocate resources and capacities helps to reduce market inefficiencies, lower energy costs for consumers, and promote the adoption of renewable energy sources across the continent.
To give a mathematical description of how the MCP/PUN is computed, we consider the be the set of P market participants. Each participant $P \in \mathcal{P}$ submits a series of bids or offers. Each one of these offers can be represented as a tuple $[q_P, s_P]$, where:

- q_P: Quantity of electricity in MWh
- s_P: Step price in €/MWh;

Let B be the set of all bidding zones, and C be the set of all connections between the bidding zones.
For each bidding zone $b \in B$, the algorithm creates a demand curve D_b and a supply curve S_b. The demand curve consists of bids sorted in descending order of price, and the supply curve consists of offers sorted in ascending order of price. Mathematically, the demand and supply curves can be represented as functions:

- $D_b(p)$: Price $\rightarrow$ Quantity (for bids)
- $S_b()$: Price $\rightarrow$ Quantity (for offers)

The market clearing price (MCP) and cleared volume (CV) for each bidding zone $b \in B$ are determined by finding the intersection of the demand and supply curves:

$$\begin{aligned} \text{MCP}_b &= p^* \text{ such that } D_b(p^*) = S_b(p^*) \\ \text{CV}_b &= D_b(\text{MCP}_b)\,, \end{aligned}$$

as we can see in Fig. 6.1.

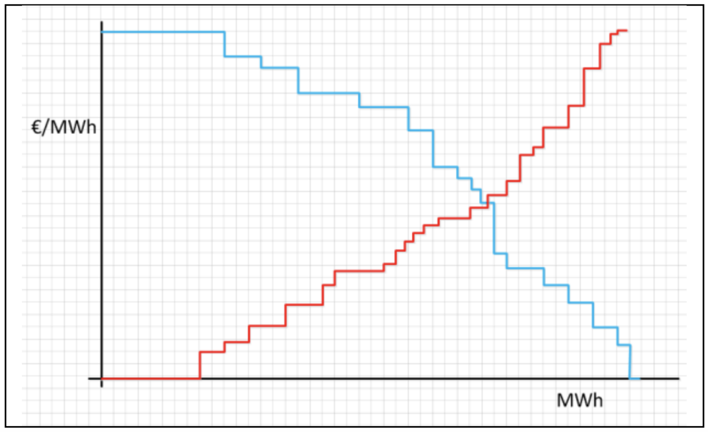


Figure 6.1: Demand and Offer stepwise aggregated curve (from [100]).

Let T_{bc} be the maximum transmission capacity from bidding zone b to c. The algorithm finds an optimal allocation of cross-zonal capacities (X_{bc}) that maximizes social welfare while adhering to the transmission capacities:

$$\max \sum_{b,c \in B} (\mathrm{MCP}_b - \mathrm{MCP}_c) \cdot X_{bc}$$

subject to:

- $0 \leq X_{bc} \leq T_{bc}, \forall b, c \in B$
- Flow conservation constraints

When there is congestion in the transmission network, the algorithm re-optimizes the allocation of resources and cross-zonal capacities. This may result in price differences between bidding zones, known as congestion rent:

$$\mathrm{CR}_{bc} = (\mathrm{MCP}_b - \mathrm{MCP}_c) \cdot X_{bc}$$

The final day-ahead prices for each bidding zone $b \in B$ are calculated based on the results of the clearing process and the allocated cross-zonal capacities:

$$\mathrm{Price}_b = \mathrm{MCP}_b$$

The initial simplification we introduce in our RL method is that we concentrate just on stages 1-3 of the workflow, which means that we will only take into account the supply/demand coupling problem without taking into account the subdivision into the cross-zonal sub-problems.

6.2.2 Reinforcement Learning

Reinforcement learning [204, 131] is a learning paradigm mapping situations to actions to maximize a numerical reward signal through repeated experience gained by interacting with the environment. The agent's objective is to develop a strategy that maximizes the expected cumulative reward over time by learning a policy that maps states to actions. The most common algorithms for RL include Q-learning, deep Q-networks (DQN), and policy gradient methods, such as REINFORCE and proximal policy optimization (PPO).
In a recent survey paper [120], the authors review model-free RL algorithms with infinite horizon and discounted reward, focusing on some classical value-based and policy-based methods.

Model-free algorithms: value-based vs policy-based approaches

Model-free algorithms do not require any knowledge about the underlying model and instead focus on directly optimizing the policy or other value parameters in a goal-oriented approach. They can be further divided into two categories: value-based approaches and policy-based approaches. The objective of value-based methods is to find accurate estimates of the state and/or state-action pair value functions $V(s)$ and $Q(s, a)$. One example of this approach is the well-known Q-learning algorithm. On the other hand, policy-based methods do not require the estimation of the value function, but instead use a parameterized policy that represents a probability distribution of actions over states $\pi_\theta = Pr[a|s]$ as a neural network. The policy is directly optimized by defining an objective function and using gradient ascent to reach an optimal point. An example of a policy-based method is the Actor-Critic algorithm.
Two networks are trained in the family of algorithms known as Actor-Critic. The critic evaluates the effectiveness of the action taken, i.e., it approximates the value function, whereas the actor approximates the policy and chooses which action to take.

ML Electricity forecast.

The problem of price electricity forecasting and energy market modelling have been tackled by a plethora of methods and predictive algorithms. We refer to [147] for a presentation and a comparison of several statistical and deep learning algorithms in the field of electricity price forecasting. In [208], the authors evaluate different Machine Learning models, such as

Support Vector Machine, Recurrent NN and Deep Learning, for predicting electricity price forecasting on the day-ahead market over different areas of Europe on separated test periods focusing on the impact of adding new predictive features. A similar perspective is followed in [157], where the novel CTSGAN (Conditional Time Series Generative Adversarial Network) is introduced proving its stability and consistency for forecasting day-ahead electricity prices in the Australian Electricity Market.

Methodologies for RL in Electricity Auctions

In recent years, there has been increasing interest in applying reinforcement learning techniques to the modeling of day-ahead electricity markets, with the aim of developing more accurate and effective strategies for market participants. In [218], the authors models the electricity auction market by means of a Q-learning algorithm considering each supplier bidding strategy as Markov Decision Problem where the agents learn from past experience an optimal bidding strategy to maximize its own payoff. Although there are certain limits in terms of application for actual case scenarios due to the usage of simple synthetic datasets and Q-tables with discrete pairings of actions-state, this work still serves as a reference point.
This section presents a selection of RL methodologies that have been applied to electricity auctions, along with their key contributions and limitations.

1. Q-learning is a popular model-free RL algorithm for learning optimal action-value functions in discrete state and action spaces [211]. In a discrete Q-learning setting, a simple data structure that we use to keep track of the states, actions and their expected rewards is represented by a table, known as Q-table: the Q-table maps a state-action pair to a Q-value that represents the quality (hence the estimated optimal future value) of the selected action given a particular space which the agent will learn. At the start of the Q-Learning algorithm, the Q-table is initialized to all zeros indicating that the agent does not know anything about the world. Using trial and error aiming to learn each state-action pair's expected reward and updating the Q-table with the new Q-value is called *exploration*. Conversely, explicitly choosing the best-known action at a state is called *exploitation*. In the context of electricity auctions, Q-learning has been used to learn bidding strategies for market participants, such as generators and retailers [174]. However, the discrete nature of Q-learning can limit its applicability to auctions with large or continuous state and action spaces.

2. Deep Q-Networks (DQN) extend Q-learning by using Deep Neural Networks (NNs) to approximate the action-value function, enabling RL in large or continuous state spaces [169]. DQN has been applied to electricity auctions for learning optimal bidding

strategies in various market settings, such as day-ahead markets and real-time markets. However, DQN still assumes discrete action spaces and can be computationally expensive due to the use of deep neural networks. The idea is to exploit a Neural Network mapping states to (action, Q-value) pairs to approximate the state-action value function. The success of Deep RL is based on the following features. The first one is the introduction of an experience replay mechanism in which every experience tuple e_t, composed of state transition, action selected and received reward, is stored in a dataset and then randomly batched avoiding the correlation between consecutive iterations. The second feature concerns the use of 2 NNs with the same architecture but different weights in the learning process. The first NN aims at approximating Q, the Q-network. Conversely, for every n steps, the parameters from the main network are copied to the target network that uses the following training loss function, defined as

$$L_i(\theta_i) = \mathbb{E}_{(s,a,r,s')\sim U(D)}\left[\left(r + \gamma \max_{a'} Q(s',a';\theta_i^-) - Q(s,a;\theta_i)\right)^2\right]$$

being θ_i^- and θ_i the parameters of the target network and of the Q-network at iteration i, respectively.

3. Policy gradient methods have been extended to make use of deep NNs, keeping the advantage of allowing for policies in the continuous action space. Policy gradient methods, such as REINFORCE and Proximal Policy Optimization (PPO), directly optimize the policy by estimating the gradient of the expected cumulative reward [213, 198]. These methods can handle continuous state and action spaces, making them suitable for electricity auctions with complex market dynamics. Applications of policy gradient methods in electricity auctions include learning optimal bidding strategies for generators and demand response aggregators [54]. One limitation of policy gradient methods is that they may require a large number of samples for stable learning.

4. Electricity auctions involve multiple agents with different objectives and learning dynamics, making it a natural fit for Multi-Agent Reinforcement Learning (MARL) approaches. MARL algorithms, such as Independent Q-learning, Multi-Agent Deep Deterministic Policy Gradient (MADDPG), and Centralized Critics with Decentralized Actors (CCDA), have been applied to learn coordinated bidding strategies for electricity auctions [220]. Although MARL can capture complex agent interactions, it may suffer from scalability issues and instabilities in the learning process.

5. Actor-critic methods combine the advantages of policy gradient methods and value function approximation to improve the learning process [140]. The actor is responsible

for generating actions based on the current policy, while the critic learns to evaluate the policy by estimating the value function. In electricity auctions, actor-critic methods have been used to learn bidding strategies and demand response management, offering a balance between exploration and exploitation [111].

6. Deep Deterministic Policy Gradient (DDPG) is an off-policy algorithm that extends the idea of the actor-critic method to continuous action spaces [155]. DDPG uses a deep neural network to approximate the policy and another deep neural network to approximate the value function. In the context of electricity auctions, DDPG has been applied to learn optimal bidding strategies for generators and energy storage systems in day-ahead markets and real-time markets [221].

7. Monte Carlo Tree Search (MCTS) is a tree search algorithm that uses Monte Carlo simulations to estimate the expected value of actions in a given state [32]. MCTS has been applied to electricity auctions to handle complex decision-making problems with large state spaces and uncertainty. For example, MCTS has been used to optimize bidding strategies in multi-stage electricity auctions, considering the uncertainty in future market conditions [59].

8. Inverse Reinforcement Learning (IRL) aims to learn the underlying reward function of an expert agent by observing its behavior [172]. In electricity auctions, IRL has been used to model the bidding behavior of market participants, allowing for the analysis of strategic interactions and the development of counter-strategies [206]. By learning the reward function of other market participants, IRL can provide insights into their objectives and decision-making processes.

6.3 The Single Agent RL setting

One important issue for electricity suppliers is how to optimally bid into the electricity auction market in order to maximize their profit. We model this problem into a single agent RL setting.

In general terms, a RL algorithm includes: a representation of a value function that provides a prediction of how good each state (or each state-action pair) is; a direct representation of the policy $\pi(s)$ or $\pi(s, a)$; a model of the environment by means of the transition function and the reward function in conjunction with a planning algorithm (any computational process that uses a model to create or improve a policy).

The first two components are related to what is called model-free RL. When the latter component is used, the algorithm is referred to as model-based RL.

The framework for this chapter falls into the setting of continuous control with DDPG first introduced in [155]. DDPG is a reinforcement learning algorithm that uses deep neural networks to approximate policy and value functions in a high dimensional, continuous actions space.
The reward $r : S \times A \times S \to \mathbb{R}$ is the reward function representing the instantaneous reward received, transitioning from (s,a) to s'.
At stage t, each seller $i \in N$ selects and executes an action, that in our setting corresponds to a step-wise energy/price function, depending on the learned policy $\pi^i : S \to P(A)$, which maps states from the state space $\mathcal{S}$ to a probability distribution over the action space $\mathcal{A}$, that determines the behaviour of an agent. In Fig. 6.2, we report an example of an offering curve that we obtain corresponding to an action that the agent can take at time t.

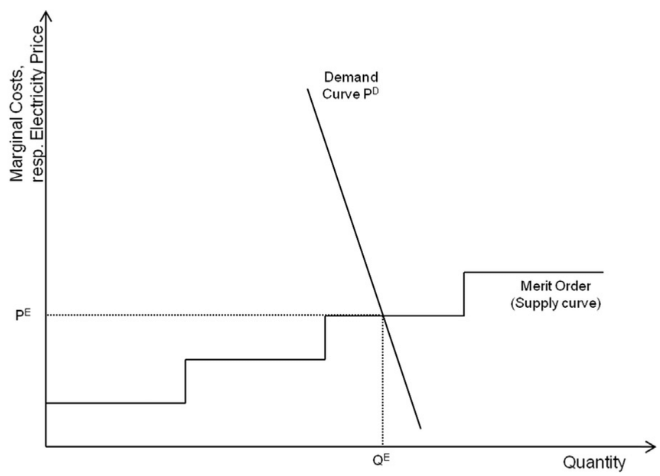


Figure 6.2: Offering Curve: the intersection with the demand (thus the reward) is computed with the next-day price of electricity.

.

The immediate reward is insufficient for providing insights into the long-term profit. Therefore, it is crucial to introduce the return value R_t, defined over a finite time horizon T. The return value R_t is given by the following expression:

$$R_t = r_{t+1}(s_t,a_t) + \gamma^{t+1} r_{t+2}(s_{t+1},a_{t+1}) + \dots + \gamma^{T-1} r_T(s_{T-1},a_{T-1}) = r_{t+1}(s_t,a_t) + \sum_{i=t+1}^{T-1} \gamma^i r_{i+1}(s_i,a_i) \tag{6.1}$$

In this equation, r_{t+k} denotes the instantaneous reward at time $t+k$, and $\gamma \in [0,1]$ is the discount factor. The discount factor determines the importance of future rewards compared to immediate ones, with lower values placing more emphasis on the short-term rewards. The

term $r(s_i, a_i)$ represents the reward obtained at time i when taking action a_i in state s_i. The return value R_t is the discounted sum of future rewards, allowing agents to optimize their actions for long-term profit.
Whereas the r_t indicates what is good in an immediate sense, R_t specifies what is good in the long run.
The goal of the agent is to take actions that will maximize its expected performance over the long term, in relation to an unknown transition function P. To achieve this, the agent learns a behaviour policy $\pi : \mathcal{S} \to \mathcal{P}(\mathcal{A})$ that optimizes its expected performance as it learns. The system progresses from state s_t under the joint action $\mathbf{a}_t \in \mathcal{A}$, based on the transition probability function P, to the next state s_{t+1}, providing updated information on the aggregated load, corresponding unitary loads, and new unitary price. Additionally, each agent receives R_i as immediate feedback for the state transition.
The goal of the agent is to optimize an objective in the form of R_t

$$J = \mathbb{E}_{r_i, s_i, a_i \sim \pi}[R_t] \tag{6.2}$$

that corresponds to learning a policy that maximizes the cumulative future payoff to be received starting from any given time t until the terminal time T.
The agent's value function associated with such a control problem is

$$V(s_n) = \max_{a_n \in \mathcal{A}} \mathbb{E}\left[R(s_n, a_n, s_{n+1})\right]$$

The dynamic programming principle implies that V satisfies the Bellman equation

$$V(s_n) = \max_{a_n \in \mathcal{A}} \mathbb{E}\left[r(s_n, a_n, s_{n+1}) + \gamma V(s_{n+1})\right]$$

In RL, it is useful to define an action–value function to measure the *quality* of taking a specific action a, and hence is called the Q-function $Q : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ defined as

$$Q(s_n, a_n) = R(s_n, a_n) + \gamma \max_{a} Q(s_{n+1}, a)$$

that solves a Bellman equation

$$Q(s_n, a_n) = R(s_n, a_n) + \gamma \max_{a_{n+1}} Q(s_{n+1}, a_{n+1}) \tag{6.3}$$

6.4 DDPG for elecrity auction

DDPG is a model-free actor-critic algorithm based on the deterministic policy gradient with a continuous action spaces.
Both actor and critic are approximated using deep Feed Forward Neural Networks (FFNN) with a second set of target FFNNs.
The Actor-Network approximates a parameterized actor function $\mu(s|\theta^{\mu})$ which deterministically describes the current policy and outputs the best offering curve directly given an observation of last 7 days prices. In Fig. 6.3, we sketch the Feed Forward NN we use to model the Actor Network.

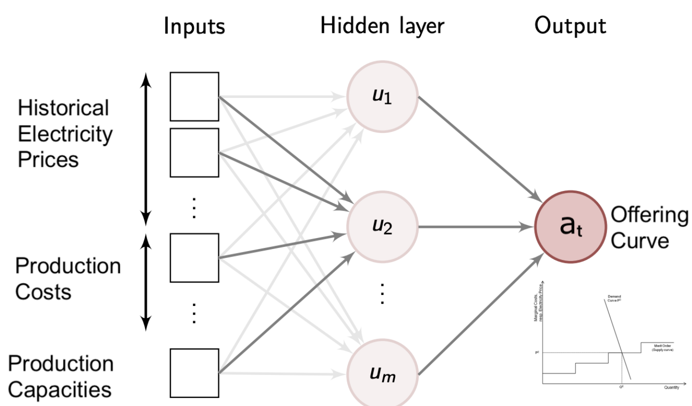


Figure 6.3: The Feed Forward Actor Network produces a vectorial output representing the Offering Curve.

.

The critic network is used to approximate the Q function, for a given state-action pair. On the other hand, the critic function $Q(s,a)$ approximates the Q-value value function given a (price-offering curve) pair estimating the expected return by approximating the Bellman Equation.

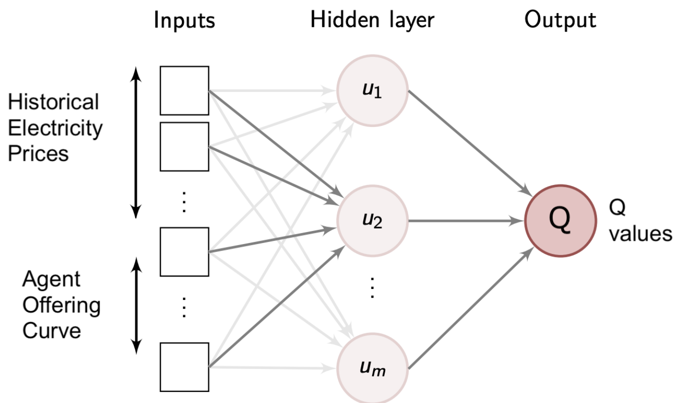


Figure 6.4: The Feed Forward Critic Network approximates the Bellman Equation given (s_t, a_t)

.

The actor is updated by applying the chain rule using the sampled policy gradient

$$\nabla_{\theta_\mu} J \sim \frac{1}{N} \sum_i \nabla_a Q(s,a|\theta_Q)\big|_{\{s=s_i, a=\mu(s_i)\}} \nabla_{\theta_\mu} \mu(s\,|\,\theta_\mu)\big|_{\{s=s_i\}} \tag{6.4}$$

that minimizes the distance between the current policy's actions and actions that maximize expected rewards. This technique involves estimating the gradient of the expected cumulative reward with respect to the policy parameters by using samples collected during interactions with the environment.

The DDPG algorithm consists of the following steps:

1. Initialize the actor network with weights θ_μ and the critic network with weights θ_Q.

2. Initialize the target networks for the actor and critic with the same weights: $\theta_\mu^{'} = \theta_\mu$ and $\theta_Q^{'} = \theta_Q$.

3. Sample a minibatch of transitions $(s, a, r, s^{'})$ from the replay buffer.

4. Update the critic network by minimizing the loss:

$$L = \frac{1}{N} \sum_i (y_i - Q(s_i, a_i|\theta_Q))^2 \tag{6.5}$$

   where $y_i = r_i + \gamma Q(s_i^{'}, \mu(s_i^{'}|\theta_\mu^{'})|\theta_Q^{'})$ is the target Q-value, and μ is the deterministic policy from the actor network. The critic loss (6.5) corresponds to Mean Squared Error (MSE) between the predicted Q-values and the actual rewards plus the discounted Q-value of the next state (i.e., the Bellman equation). This loss guides the critic network to approximate the true Q-values as accurately as possible.

5. Update the actor network using the sampled policy gradient introduced in (6.4);

6. Update the target networks using the soft update rule:

$$\theta_{\mu}^{'} \leftarrow \tau\theta_{\mu} + (1-\tau)\theta_{\mu}^{'}, \quad \theta_{Q}^{'} \leftarrow \tau\theta_{Q} + (1-\tau)\theta_{Q}^{'}, \tag{6.6}$$

where $\tau \ll 1$ is a small constant that controls the rate of the update.

using an exploration strategy based on noise processes, such as the Ornstein-Uhlenbeck one, to add temporally correlated noise to the actions. Hence, the noise process X_t is defined by the following SDE:

$$dX_t = -\theta(X_t - \mu)dt + \sigma dW_t \tag{6.7}$$

that we discretize by the following Euler-Maruyama method:

$$X_{t+\Delta t} = X_t - \theta(X_t - \mu)\Delta t + \sigma\sqrt{\Delta t}\xi_t \tag{6.8}$$

where ξ_t represents a random sample from a standard normal distribution.
Finally, the generated noise is added to the actions produced by the Actor network:

$$a_t = \mu(s_t|\theta_{\mu}) + X_t \tag{6.9}$$

where a_t is the action taken at time t, $\mu(s_t|\theta_{\mu})$ is the action produced by the actor network for the state s_t, and X_t is the noise generated by the Ornstein-Uhlenbeck process.
We report some tools used by DDPG in order to control the stability of the algorithm:

- employing a replay buffer to store past experiences and sample mini-batches of transitions for training. This technique helps to break correlations in the data, thus improving stability;

- using separate target networks for both the actor and critic to provide a more stable target for learning. The target networks are updated slowly, using a soft update rule with a small mixing factor, which helps to stabilize learning;

- adding regularization techniques, such as L_2 regularization, to the loss functions for the actor and critic networks can help prevent overfitting and improve the stability of the learned policy.

- carefully tuning the hyperparameters, such as the learning rates, discount factor, and soft update rate, can significantly impact the stability of DDPG. Conducting a systematic search or using optimization techniques can help find the best hyperparameters for a specific problem.
- clipping gradients during the backpropagation step can help mitigate exploding gradients and stabilize learning;

Moreover, in order to improve the stability of the algorithm a solution of the algorithm we work in mini-batches (the so-called experiences arrays) and we add common noise to randomize the actions.
Therefore, we derive the following scheme:

Algorithm 1 DDPG for electricity auctions

1: Initialize the Actor network $\mu(s_t|\theta_\mu)$ and the Critic network $Q(s_t, a_t|\theta_Q)$ with random weights θ_μ and θ_Q.

2: Initialize the target networks $\mu'(s_t|\theta'_\mu)$ and $Q'(s_t, a_t|\theta'_Q)$ with weights $\theta'_\mu \leftarrow \theta_\mu$ and $\theta'_Q \leftarrow \theta_Q$.

3: Initialize the Ornstein-Uhlenbeck noise process X_t.

4: For each episode:

Initialize the environment and obtain the initial state s_0.

For each time step t:

1: Select the action $a_t = \mu(s_t|\theta_\mu) + X_t$, where X_t is the noise generated by the Ornstein-Uhlenbeck process.

2: Execute the action a_t in the environment and observe the reward r_t and the next state s_{t+1}.

3: Store the transition (s_t, a_t, r_t, s_{t+1}) in the replay buffer.

4: Update the noise process $X_{t+\Delta t} = X_t - \theta(X_t - \mu)\Delta t + \sigma\sqrt{\Delta t}\xi_t$, where ξ_t is a random sample from a standard normal distribution.

5: If the replay buffer contains enough samples, sample a mini-batch of transitions (s_j, a_j, r_j, s_{j+1}) from the replay buffer.

6: Update the Critic network by minimizing the loss:

$$L(\theta_Q) = \frac{1}{m}\sum_{j=1}^{m}\Big(Q(s_j, a_j|\theta_Q) - (r_j + \gamma Q'(s_{j+1}, \mu'(s_{j+1}|\theta'_\mu)|\theta'_Q))\Big)^2 \tag{6.10}$$

7: Update the Actor Network using the sampled policy gradient:

$$\nabla_{\theta_\mu} J(\theta_\mu) \approx \frac{1}{m}\sum_{j=1}^{m} \nabla_a Q(s_j, a|\theta_Q)\big|_{a=\mu(s_j|\theta_\mu)} \nabla_{\theta_\mu}\mu(s_j|\theta_\mu) \tag{6.11}$$

8: Update the target networks using soft updates:

$$\theta'_\mu \leftarrow \tau\theta_\mu + (1-\tau)\theta'_\mu, \quad \theta'_Q \leftarrow \tau\theta_Q + (1-\tau)\theta'_Q, \tag{6.12}$$

9: Repeat until the desired level of performance is achieved or a maximum number of episodes is reached.

6.4.1 Selection of the DDPG hyperparameters

We briefly describe the hyperparameters we use for the DDPG algorithm while referring to Table 6.1 for the reference values we set for the implementation.

Table 6.1: DDPG Hyperparamters

Episodes	1000, 1500
Length of an Episode [Days]	15, 20, 30
Batch Size	64
Hidden Size	64
Actor Learning Rate	0.00001
Critic Learning Rate	0.00001
Discount Factor	0.99
Tau	0.01
Max Memory Size	50000

An episode refers to a complete sequence of interactions between the agent and the environment. The *Episodes* corresponds to the number of times the agent will engage with the environment to learn and improve its policy. The *Length of an Episode* refers to the number of time steps or interactions the agent experiences within a single episode. It defines how long the agent operates in the environment before the episode concludes. The *Batch Size* refers to the number of experiences (state-action-reward-next state tuples) sampled from the replay buffer at each iteration of the training process. The *Hidden Size* refers to the number of neurons or units in the hidden layers of the neural networks used in the DDPG algorithm. The *Actor Learning Rate* and the *Critic Learning Rate* control the step size of the gradient update at which the actor network's weights θ_μ and the critic network's weights θ_Q are updated during training. The *Discount Factor* parameter γ introduced in Eq. (6.1) represents the relative importance of future rewards compared to immediate rewards. A higher gamma value places more importance on long-term rewards, potentially encouraging the agent to consider the future consequences of its actions. The *Target Update Factor* τ is a relaxation factor defining how often the parameters are copied from the original networks to the target network. for the copied network parameters. The *Max Memory Size* determines the capacity of the replay buffer, which is a crucial component in DDPG. The replay buffer stores past experiences (state, action, reward, next state) that the agent uses to learn from.

Since DDPG hyperparameters are interconnected, finding the right tuning has a significant impact on the performance and stability of the algorithm. We calibrate the values of Table 6.1 by an experimental procedure by adjusting them in order to optimize the global reward.

We refer to next Subsection for a complete overview of the tuning procedure accompanied by some technical considerations.

6.5 Implementation of the Algorithm

We consider a single agent setting of an energy operator interacting with the market (environment) in discrete time steps.
At each time step t, the agent:

1. receives an observation x_t consisting in of 24-hour PUN array prices P_t of the last d days. We assume a fully-observable environment: the state of the environment is represented by the market electricity prices expressed in €/MWh;

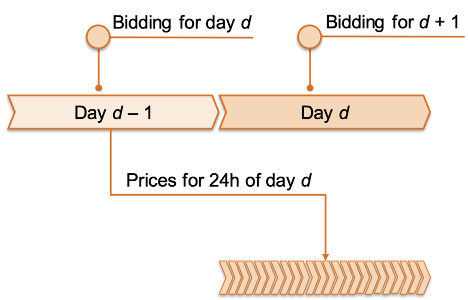


Figure 6.5: Bidding settlement in day-ahead auctions (from [161])

.

2. generates an action a_t, i.e. a step-wise curve modelling the offering curve, corresponding to the output of the Actor Network;

3. observes a feedback scalar reward r_t corresponding to the payoff of the agent.

At hour t, the reward reads

$$r_t = \sum_{i=1}^{K} \left(P_t - C^i \right) D_t^i \tag{6.13}$$

where P is the unitary electricity price, C^i is the production cost and D_t^i is the dispatched power.
Following the paradigm of a repeated day-ahead electricity auction market, the agent will attempt to maximize her profit in the long run. We assume a fixed generation capacity and different production costs corresponding to different sources of energy. More precisely, to calibrate our result, we consider a constant array of production costs and available power

for computing a value for the reward function defined in Eq. (6.13). In particular, we set the number of sources of production K to 3 and we set cost $C^i = [10, 30, 60]$ and production capacity $D^i = [30, 200, 800]$. From a modelling point of view, the source with a low marginal cost and low capacity, i.e. $(10, 30)$ represents a renewable source of production, the one with high marginal cost and capacity a conventional one (e.g. gas) plus an intermediate one.
Another presumption we make in the model is that the agent only knows about its own expenses, available resources, and historical electricity prices and nothing about its competitors. As a result, her bidding strategy can be represented as a stochastic process that adheres to a decision-making framework.
Each producer in the stochastic bidding strategy strives to maximize her profit over the long term while meeting the needs of the available

6.5.1 Numerical Results

The dataset consists of Italian electricity prices over 4 years. Precisely, we use hourly data for PUN from 2017 to 2020 that we download at `https://www.mercatoelettrico.org/it/Download/DatiStorici.aspx` plotted in Fig. 6.6.

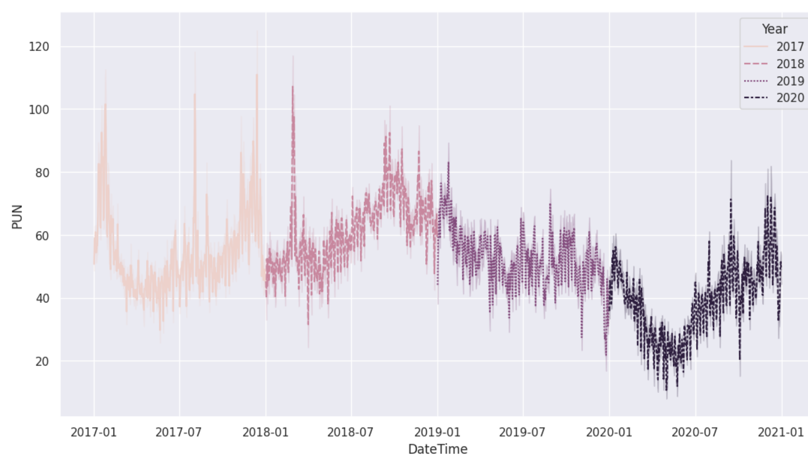


Figure 6.6: Italian hourly PUN from 01-Jan-2017 to 12-Dec-2020 for a total of 35064 data points

.

Rewards obtained from the environment might vary significantly in magnitude since they depend on the price time series that is highly not stationary, as we can see from Fig. 6.6. To address this issue, we include a normalization factor into the reward to consider the potential *normalized reward* that can be obtained at each time t.
The normalized reward, denoted as $\mathcal{R}_{norm}$, is calculated by dividing the actual reward r

introduced in Eq. (6.13) for a given time step by the maximum possible value of the reward r_{max} that can be achieved the particular time step conditions. When production costs, production capacity, and the matching pun for a specific time step are taken into account, the maximum reward reflects the most profit that may be realized. For more steady and effective learning, this normalization scales the reward values to a consistent range between 0 and 1: when the agent receives the maximum reward, the normalized reward will be 1. The normalized reward ranges from 0 to 1 if the agent performs below the maximum. This option enables the agent to focus on tailoring its strategy in response to the relative performance improvement, which improves the consistency of comparisons between various learning contexts.
Besides the normalized reward, we also plot two interesting metrics the Policy Loss and the Critic Loss. The agent learns to improve its policy by adjusting the actor network's weights to maximize the expected cumulative reward J, defined in Eq. (6.2). The policy loss measures the discrepancy between the actions chosen by the current policy and the actions that would lead to higher expected rewards and it is approximated by the sampled policy gradient introduced in (6.4). The Critic Loss introduced in Eq. (6.5) is associated with training the critic network and it measures the accuracy of the critic's Q-value predictions and guides the critic network to approximate Q-values that satisfy the Bellman equation (6.3).
We point out that the algorithm is really sensitive to the initialization of parameters such as (# episode, # days in the episode, production cost) that have to be carefully chosen.
For this reason, we perform several simulations for different values of *Episodes* and *Length of an Episodes* reporting the obtained results for 5 different simulations for which we obtain the following plots.

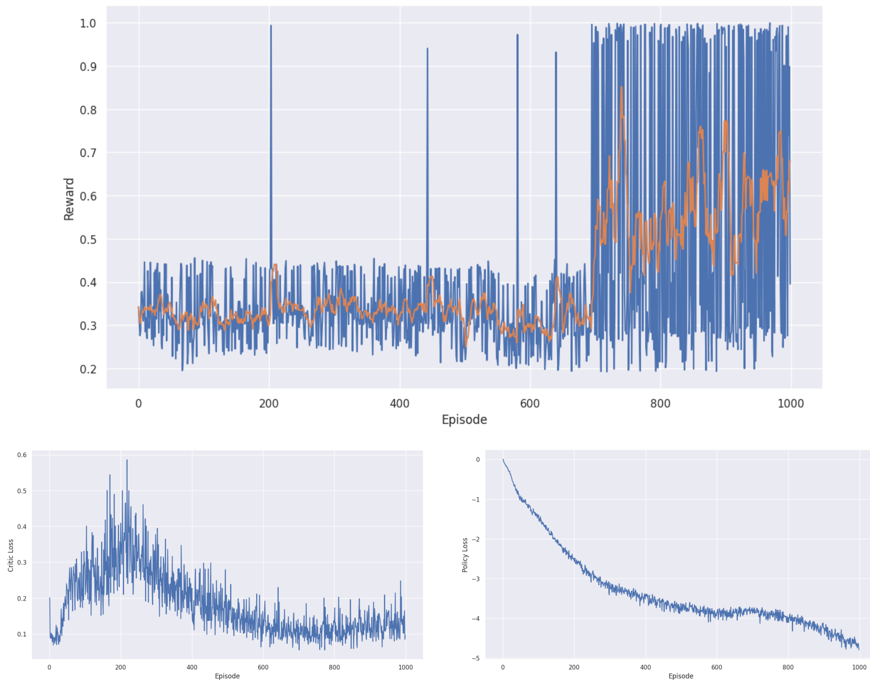


Figure 6.7: **Simulation 1**: Episodes = 1000, length of episode = 30

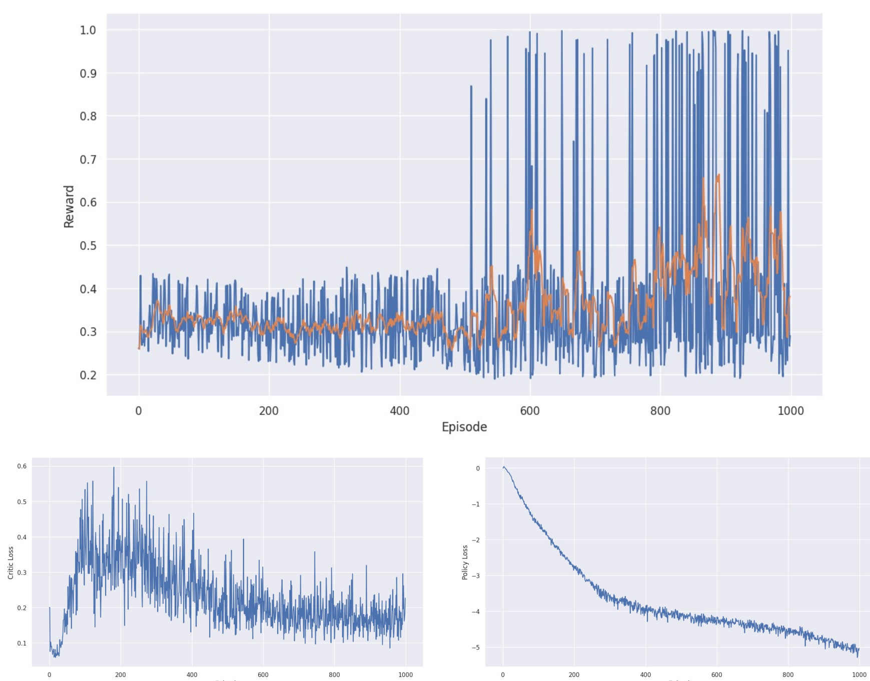


Figure 6.8: **Simulation 2**: Episodes = 1000, length of episode = 30

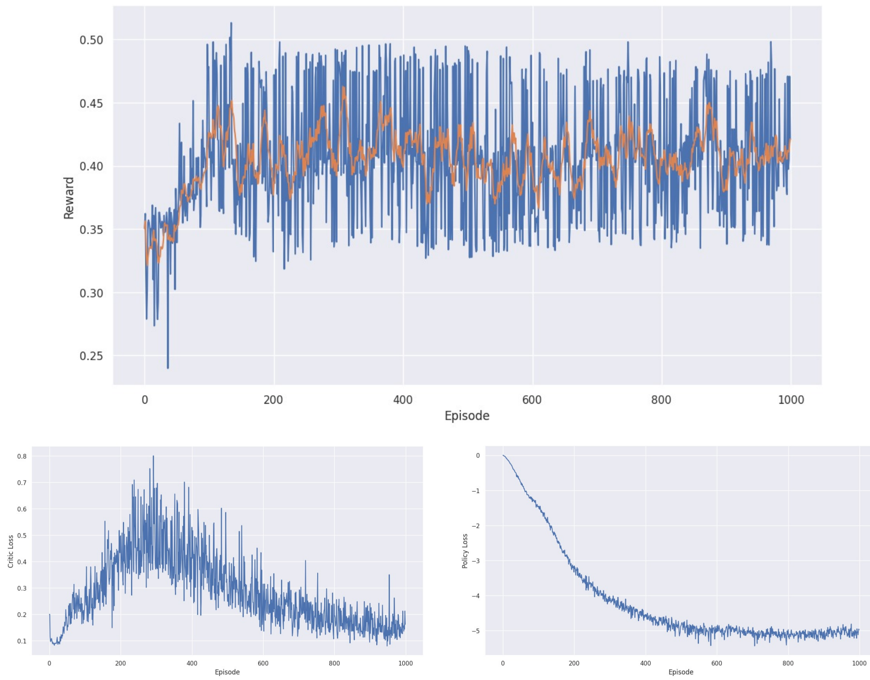


Figure 6.9: **Simulation 3**: Episodes = 1000, length of episode = 30

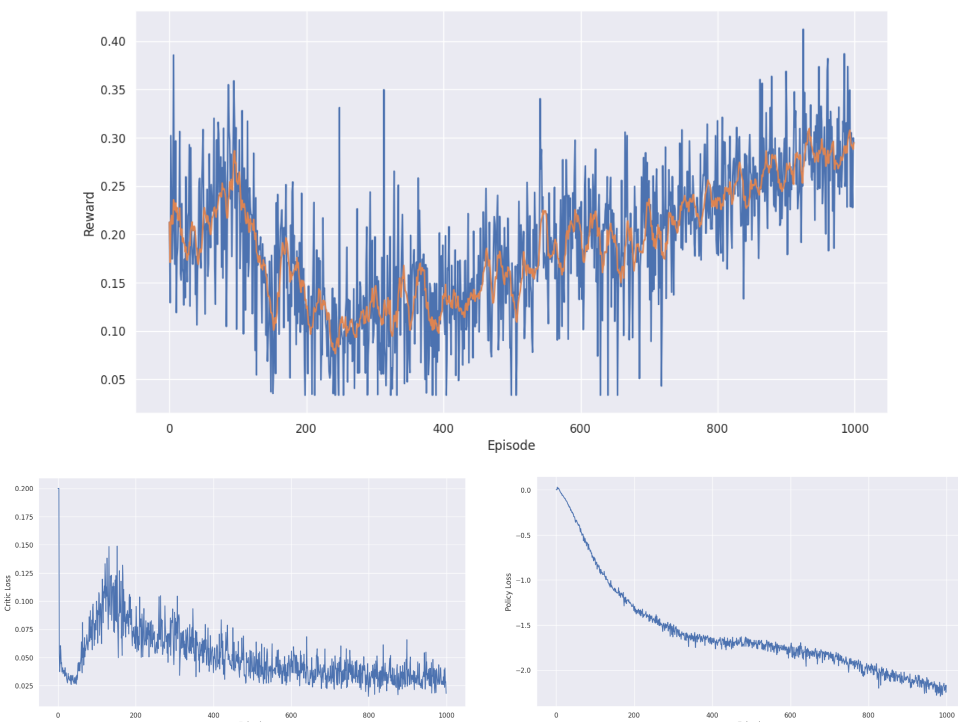


Figure 6.10: **Simulation 4**:Episodes = 1000, length of episode = 20

In all plots, we see that the policy loss consistently decreases over time, hence the actor is moving towards a better policy. Conversely, the critic loss shows an initial spike before gradually decreasing, thus the Q-value estimate improves only after approximately 300 *Episodes.* This behaviour is a common pattern in the training dynamics of RL algorithms that can be attributed to the interaction with a stochastic environment. Specifically, in the early stages of training, the agent's policy might be far from optimal, leading to higher Q-values and a larger critic loss. As the agent explores the environment and gathers more experiences, it gradually refines its policy, causing the critic loss to decrease. Moreover, this delayed learning is also linked to the target networks which are updated slowly leading to a larger initial critic loss. As training progresses and the target networks catch up, the critic loss starts to decrease. However, the gradual decrease in both policy and critic losses as training metrics as well as the increase of the normalized reward are indicative of the agent's convergence towards more optimal policies.

6.5.2 Considerations, Challenges and Limitations of the model

As previously encountered in other settings such as [191], in the single-agent framework, the market rewards may be highly stochastic and may be unable to converge to a fixed point, leaving to the oscillatory behaviours of market producers. This behavior is in a sense expected because generators do not account for the strategies of their competitors, which may evolve over time. In contrast, the gradual decline in policy loss and the initial spike and subsequent decline in critic loss are signs of the DDPG algorithm's learning process. They demonstrate the agent's evolution from initial exploration and poorly predicted Q-values to more precise value estimations and a more developed strategy. These actions indicate that the algorithm is learning from its training data and adjusting to its surroundings, which are generally expected behaviors.

The difficulties we run across while building the algorithm can be divided into two groups: those caused by the way the electrical auction markets are set up, and those that are specifically related to the RL algorithm. The first type consists of the action space's partial observability (the agent can learn about other people's actions and tactics only through time series of historical pricing), non-stationarity of the data, and high problem dimensionality. Other sources of challenges that arise from the stability of the DDPG are connected to inherent difficulties in learning, such as function approximation errors and non-stationary targets.

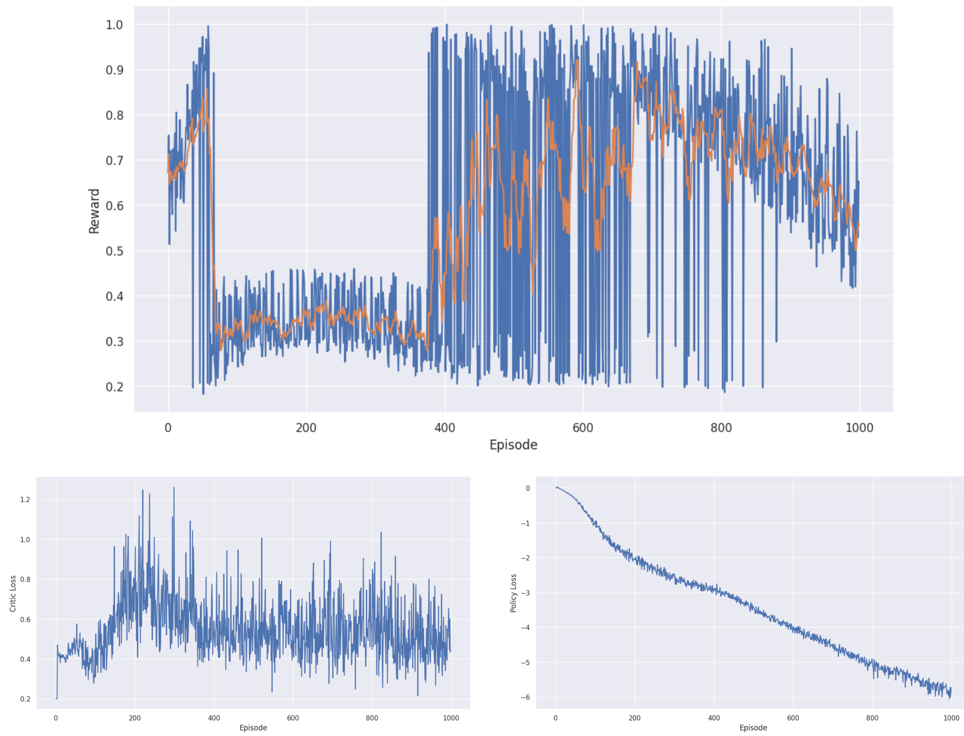


Figure 6.11: **Simulation 5**: Episodes = 1000, length of episode = 15

Additionally, when we monitor the algorithm's process, we notice the following observations, which are mostly concerned with the stability of the DDPG technique and the selection of the hyperparameters:

- A higher number of episodes allows the agent to explore the environment more extensively, potentially leading to better policy convergence. Episodes provide opportunities for the agent to encounter a diverse range of states and situations, which is crucial for learning a robust policy that performs well across different scenarios. Conversely, the learning process often exhibits diminishing returns with increasing episodes, see Fig. 6.11 that is trained with 15 days for episode. Initially, the agent might rapidly learn and improve its policy but over time, the rate of improvement might slow down as it explores less novel situations. Hence, the length of an episode appears to be a crucial parameter for the exploration-exploitation trade–off;

- Also *Length of an Episode* hyperparameter can influence the learning dynamics and the exploration strategy of the agent. A longer episode provides the agent more time to investigate the surroundings thoroughly and make more choices, which might result in further exploration. A shorter episode, on the other hand, would encourage the agent to focus more on using what it already knows, thereby limiting exploration. We identify a crucial value of 15 days for *Length of an Episode* below which we do not have convergence.

- By experimentation, we set an equal learning rate both for the critic and for the actor network updates to obtain a more stable training by preventing one network from significantly outpacing the other;
- A larger batch size can lead to more efficient learning as it allows the agent to learn from more experiences in parallel. However, we note that employing a very high batch size may increase the amount of noise and volatility in the learning process, resulting in less consistent training. As each update is based on a smaller subset of the overall experiences, we selected a batch size that is relatively modest to encourage the agent to explore more varied encounters.

6.6 Conclusions and possible future directions

The goal of this project concerns the development of an optimizing strategy for a single-agent setting.
Through an actor-critic architecture, DDPG can learn a deterministic policy while still making use of a Q-value function to guide its learning. This combination allows DDPG to effectively handle continuous action spaces, making it a powerful algorithm for a wide range of applications.
In literature there are studies developing RL algorithms with generated data such as [218] where the Q learning is first adopted to study the electricity bidding problem. Basically, we extend the result in [218] introducing NNs, specifically the DDPG method. Another source of novelty relies on a model that directly uses historical real prices during the decision–making process.
We conclude the chapter by presenting some extensions of this model that may be worth to further investigate:

- we solely test feed forward NNs for the implementation of this model. We think the algorithm's performance could be enhanced by choosing architectures that are better suited for time series, including recurrent NNs, such as Long Short Term Memory NN. the need for thorough data preprocessing in order to execute a recursive method across several episodes, as well as the growing computing cost of implementing recurrent NNs;
- we model production cost C_t and production capacity as deterministic constant quantity.

We think that the model may have real benefits by considering stochastic capacities to model the random fluctuations related to renewable modes of production.

- this project could act as a foundational building block for the creation of an algorithm for the Multi-Agent or Mean Field system. The construction of a distributed optimization system—an algorithm based on decentralized coordination such as local rewards or a consensus scheme—is the foundation for the extension of a single agent into a multi-agent context.

List of Figures

List of Tables

Appendix

Appendix: proof of Theorem 1.6

Proof. The existence and the uniqueness are obtained by the Banach fixed point theorem. We consider ϕ fixed in Λ and we define the map Γ on $\mathcal{A}$ with $\mathcal{A} := \mathcal{C}\big([0,T]\,;\mathbb{S}_0^2(\mathbb{R})\big)$.
For $R \in \mathcal{A}$, we define $\Gamma(R) := Y$, where, for $t \in [0,T]$, the triple of adapted processes $\big(Y^t(s), Z^t(s), U^t(s,z)\big)_{s\in[t,T]}$ is the unique solution of the following BSDE

$$\begin{aligned} Y^t(s) = h(X^{t,\phi}) + \int_s^T F(r, X^{t,\phi}, Y^t(r), Z^t(r), \tilde{U}^t(r), R_r^t)dr \\ - \int_s^T Z^t(r)dW(r) - \int_s^T \int_{\mathbb{R}} U^t(r,z)\tilde{N}(dr,dz)\,, \quad s \in [t,T]\ . \end{aligned} \tag{14}$$

For $s \in [0,t]$ we prolong the solution by taking $Y^t(s) := Y^s(s)$ and $Z^t(s) = U^{t,\phi}(s) := 0$.
Step 1. Let us first show that Γ takes values in the Banach spaces $\mathcal{A}$. We take $R \in \mathcal{A}$ and we will prove that $Y := \Gamma(R) \in \mathcal{A}$. Thus, for every $t \in [0,T]$ we have to show that

$$Y^t \in \mathbb{S}_0^2(\mathbb{R})\,,$$

and that the application

$$[0,T] \ni t \mapsto Y^t \in \mathbb{S}_0^2(\mathbb{R})\,,$$

is continuous.
Let $t \in [0,T]$ be fixed and $t' \in [0,T]$; with no loss of generality, we will suppose that $t < t'$ and $t' - t < \delta$.

Concerning the solution of the BSDE defined in (14), we obtain the following estimate

$$\begin{aligned}
&\mathbb{E}\big(\sup_{s\in[0,T]}|Y^t(s)-Y^{t'}(s)|^2\big)\\
&\le \mathbb{E}\big(\sup_{s\in[0,t']}|Y^t(s)-Y^{t'}(s)|^2\big)+\mathbb{E}\big(\sup_{s\in[t',T]}|Y^t(s)-Y^{t'}(s)|^2\big)\\
&\le 2\mathbb{E}\big(\sup_{s\in[t,t']}|Y^t(s)-Y^t(t)|^2\big)+2\mathbb{E}\big(\sup_{s\in[t,t']}|Y^t(t)-Y^s(s)|^2\big)\\
&\quad +\mathbb{E}\big(\sup_{s\in[t',T]}|Y^t(s)-Y^{t'}(s)|^2\big).
\end{aligned}$$

We start by proving that

$$\mathbb{E}\big[\sup_{s\in[t,t']}|Y^t(s)-Y^t(t)|^2\big]\to 0\,,$$

as $t'\to t$. By plugging the explicit solution and applying Doob's inequality, we get

$$\begin{aligned}
\mathbb{E}\big(\sup_{s\in[t,t']}|Y^t(s)-Y^t(t)|^2\big)&=\mathbb{E}\Big[\sup_{s\in[t,t']}\Big|\int_t^s F(r,X^{t,\phi},Y^t(r),Z^t(r),\tilde{U}^t(r),R^t_r)dr+\\
&\quad -\int_t^s Z^t(r)dW(r)-\int_t^s\int_{\mathbb{R}}U^t(r,z)\tilde{N}(dr,dz)\Big|^2\Big]\\
&\le 3\mathbb{E}\left[\int_t^{t'}\left|F(r,X^{t,\phi},Y^t(r),Z^t(r),\tilde{U}^t(r),R^t_r)\right|^2dr\right]+\\
&\quad +3\mathbb{E}\left[\sup_{s\in[t,t']}\left|\int_t^s Z^t(r)dW(r)\right|^2\right]+3\mathbb{E}\left[\sup_{s\in[t,t']}\left|\int_t^s\int_{\mathbb{R}}U^t(r,z)\tilde{N}(dr,dz)\right|^2\right]\\
&\le 3\mathbb{E}\left[\int_t^{t'}\left|F(r,X^{t,\phi},Y^t(r),Z^t(r),\tilde{U}^t(r),R^t_r)\right|^2dr\right]+\\
&\quad +12\mathbb{E}\left[\int_t^{t'}\left|Z^t(r)\right|^2dr\right]+12\mathbb{E}\left[\int_t^{t'}\int_{\mathbb{R}}\left|U^t(r,z)\right|^2\nu(dz)dr\right],
\end{aligned}$$

From the absolute continuity of the Lebesgue integral, we deduce that

$$\mathbb{E}\big[\sup_{s\in[t,t']}|Y^t(s)-Y^t(t)|^2\big]\to 0\,,$$

as $t'\to t$.
Concerning the term $\mathbb{E}\big(\sup_{s\in[t',T]}|Y^t(s)-Y^{t'}(s)|^2\big)$ let us denote for short, only throughout this step,

$$\begin{aligned}
\Delta Y(r)&:=Y^t(r)-Y^{t'}(r)\,, &\qquad \Delta Z(r)&:=Z^t(r)-Z^{t'}(r)\,,\\
\Delta U(r,z)&:=U^t(r,z)-U^{t'}(r,z)\,, &\qquad \Delta R_r(r)&:=R^t_r(r)-R^{t'}_r(r)
\end{aligned}$$

and

$$\Delta h:=h(X^{t,\phi})-h(X^{t',\phi}),$$
$$\Delta F(r):=F(r,X^{t,\phi},Y^t(r),Z^t(r),\tilde{U}^t(r),R^t_r)-F(r,X^{t',\phi},Y^t(r),Z^t(r),\tilde{U}^t(r),R^t_r).$$

We apply Itô's formula to $e^{\beta s}|\Delta Y(s)|^2$ and we derive, for any $\beta>0$ and any $s\in[t',T]$,

$$
\begin{aligned}
&e^{\beta s}|\Delta Y(s)|^2+\beta\int_s^T e^{\beta r}|\Delta Y(r)|^2dr+\int_s^T e^{\beta r}|\Delta Z(r)|^2dr+\int_s^T\int_{\mathbb{R}}e^{\beta r}|\Delta U(r,z)|^2\nu(dz)dr\\
&=e^{\beta T}|\Delta Y(T)|^2-2\int_s^T e^{\beta r}\Delta Y(r)\Delta Z(r)\cdot dW(r)-2\int_s^T\int_{\mathbb{R}}e^{\beta r}\Delta Y(r)\Delta U(r,z)\tilde{N}(dr,dz)\\
&+2\int_s^T e^{\beta r}\Delta Y(r)\left(F(r,X^{t,\phi}Y^t(r),Z^t(r),\tilde{U}^t(r),R^t_r)-F(r,X^{t',\phi},Y^{t'}(r),Z^{t'}(r),\tilde{U}^{t'}(r),R^{t'}_r)\right)dr\\
&-\int_s^T\int_{\mathbb{R}}e^{\beta r}|\Delta U(r,z)|^2\tilde{N}(dr,dz).
\end{aligned}
$$

We note that the following estimate

$$
\begin{aligned}
&\int_s^T e^{\beta r}\left(\int_{-\delta}^0\left(|\Delta R(r+\theta)|^2\right)\alpha(d\theta)\right)dr=\int_{-\delta}^0\left[\int_s^T e^{\beta r}\left(|\Delta R(r+\theta)|^2\right)dr\right]\alpha(d\theta)\\
&\le e^{\beta\delta}\cdot\int_{-\delta}^0\alpha(d\theta)\cdot\int_0^T e^{\beta r}\left(|\Delta R(r)|^2\right)dr\le Te^{\beta\delta}\sup_{r\in[0,T]}\left(e^{\beta r}|\Delta R(r)|^2\right),
\end{aligned}
$$

holds. From assumptions (A_3)–(A_5), we have for any $a>0$,

$$
\begin{aligned}
&2\int_s^T e^{\beta r}\Delta Y(r)\left(F(r,X^{t,\phi},Y^t(r),Z^t(r),\tilde{U}^t(r),R^t_r)-F(r,X^{t',\phi},Y^{t'}(r),Z^{t'}(r),\tilde{U}^{t'}(r),R^{t'}_r)\right)dr\\
&\le a\int_s^T e^{\beta r}|\Delta Y(r)|^2dr+\frac{3}{a}\int_s^T e^{\beta r}|\Delta F(r)|^2dr\\
&\quad+\frac{6L^2}{a}\int_s^T e^{\beta r}\left(|\Delta Y(r)|^2+|\Delta Z(r)|^2+\int_{\mathbb{R}}|\Delta U(r,z)|^2\lambda(z)\nu(dz)\right)dr\\
&\quad+\frac{3TKe^{\beta\delta}}{a}\sup_{r\in[0,T]}\left(e^{\beta r}|\Delta R(r)|^2\right).
\end{aligned}
$$

Therefore we have

$$
\begin{aligned}
&e^{\beta s}|\Delta Y(s)|^2+\left(\beta-a-\frac{6L^2}{a}\right)\int_s^T e^{\beta r}|\Delta Y(r)|^2dr\\
&\qquad+\left(1-\frac{6L^2}{a}\right)\int_s^T e^{\beta r}|\Delta Z(r)|^2dr+\left(1-\frac{6L^2}{a}\right)\int_s^T\int_{\mathbb{R}}e^{\beta r}|\Delta U(r,z)|^2\lambda(z)\nu(dz)dr\\
&\le e^{\beta T}|\Delta Y(T)|^2+\frac{3}{a}\int_s^T e^{\beta r}|\Delta F(r)|^2dr\\
&\quad-2\int_s^T e^{\beta r}\Delta Y(r)\Delta Z(r)\cdot dW(r)-2\int_s^T\int_{\mathbb{R}}e^{\beta r}\Delta Y(r)\Delta U(r,z)\tilde{N}(dr,dz)\\
&\quad+\frac{3TKe^{\beta\delta}}{a}\sup_{r\in[0,T]}e^{\beta r}|\Delta R(r)|^2.
\end{aligned}
$$

We now choose $\beta,a>0$ such that

$$
a+\frac{6L^2}{a}<\beta\quad\text{and}\quad\frac{6L^2}{a}<1,\tag{15}
$$

hence we obtain

$$\left(1-\frac{6L^2}{a}\right)\mathbb{E}\left[\int_s^T e^{\beta r}|\Delta Z(r)|^2dr+\int_s^T\int_{\mathbb{R}}e^{\beta r}|\Delta U(r,z)|^2\lambda(z)\nu(dz)dr\right]\leq \mathbb{E}\big(e^{\beta T}|\Delta h|^2\big)+\frac{3}{a}\mathbb{E}\int_s^T e^{\beta r}|\Delta F(r)|^2dr+\frac{3TKe^{\beta\delta}}{a}\mathbb{E}\big(\sup_{r\in[0,T]}e^{\beta r}|\Delta R(r)|^2\big). \tag{16}$$

By Burkholder–Davis–Gundy's inequality, we have

$$2\mathbb{E}\left[\sup_{s\in[t',T]}\left|\int_s^T e^{\beta r}\Delta Y(r)\,\Delta Z(r)\cdot dW(r)\right|\right] \leq \frac{1}{4}\mathbb{E}\left[\sup_{s\in[t',T]}e^{\beta s}|\Delta Y(s)|^2\right]+144\mathbb{E}\int_{t'}^T e^{\beta r}|\Delta Z(r)|^2dr.$$

and

$$2\mathbb{E}\left[\sup_{s\in[t',T]}\left|\int_s^T\int_{\mathbb{R}}e^{\beta r}\Delta Y(r)\,\Delta U(r)\,\tilde{N}(dr,dz)\right|\right] \leq \frac{1}{4}\mathbb{E}\left[\sup_{s\in[t',T]}e^{\beta s}|\Delta Y(s)|^2\right]+144\mathbb{E}\int_{t'}^T\int_{\mathbb{R}}e^{\beta r}|\Delta U(r,z)|^2\lambda(z)\nu(dz)dr.$$

which immediately implies

$$\frac{1}{2}\mathbb{E}\left[\sup_{s\in[t',T]}e^{\beta s}|\Delta Y(s)|^2\right]\leq \mathbb{E}\big(e^{\beta T}|\Delta h|^2\big)+\frac{3}{a}\mathbb{E}\int_{t'}^T e^{\beta r}|\Delta F(r)|^2dr +\frac{3TKe^{\beta\delta}}{a}\mathbb{E}\big(\sup_{r\in[0,T]}e^{\beta r}|\Delta R(r)|^2\big)+144\,\mathbb{E}\int_{t'}^T|\Delta Z(r)|^2dr+ +144\,\mathbb{E}\int_{t'}^T\int_{\mathbb{R}}e^{\beta r}|\Delta U(r,z)|^2\lambda(z)\nu(dz)dr\,.$$

Hence, we have

$$\frac{1}{2}\mathbb{E}\left(\sup_{s\in[t',T]}e^{\beta s}|\Delta Y(s)|^2\right)\leq \mathbb{E}\big(e^{\beta T}|\Delta h|^2\big)+\frac{3}{a}C_1\mathbb{E}\int_{t'}^T e^{\beta r}|\Delta F(r)|^2dr +\frac{3TKe^{\beta\delta}C_1}{a}\mathbb{E}\big(\sup_{r\in[0,T]}e^{\beta r}|\Delta R(r)|^2\big), \tag{17}$$

where

$$C_1:=1+\frac{144}{1-6L^2/a}\,.$$

Exploiting thus assumptions (A_3) and (A_5) together with the fact that $X^{\cdot,\phi}$ is continuous and bounded, we have

$$C_1\mathbb{E}(e^{\beta T}|\Delta h|^2)+\frac{3}{a}C_1\mathbb{E}\int_{t'}^{T}e^{\beta r}|\Delta F(r)|^2dr\to 0 \quad \text{as } t'\to t.$$

Since $R\in\mathcal{A}$, and therefore we have

$$\mathbb{E}\left[\sup_{r\in[0,T]}e^{\beta r}|\Delta R(r)|^2\right]\to 0,$$

as $t'\to t$, we have

$$\mathbb{E}\left[\sup_{s\in[t',T]}e^{\beta s}|\Delta Y(s)|^2\right]\to 0, \tag{18}$$

as $t'\to t$.

We are left to show that the term $\mathbb{E}\big(\sup_{s\in[t,t']}|Y^t(t)-Y^s(s)|^2\big)$ is also converging to 0 as $t'\to t$.

Since the map $t\mapsto Y^t(t)$ is deterministic, we have from equation (14),

$$\begin{aligned}
&Y^t(t)-Y^s(s)=\mathbb{E}\big[Y^t(t)-Y^s(s)\big]\\
&=\mathbb{E}\big[h(X^{t,\phi})-h(X^{s,\phi})\big]+\mathbb{E}\int_t^T F(r,X^{t,\phi},Y^t(r),Z^t(r),\tilde{U}^t(r),R^t_r)dr\\
&\quad-\mathbb{E}\int_s^T F(r,X^{s,\phi},Y^s(r),Z^s(r),\tilde{U}^s(r),R^s_r)dr\\
&=\mathbb{E}\big[h(X^{t,\phi})-h(X^{s,\phi})\big]+\mathbb{E}\int_t^s F(r,X^{t,\phi},Y^t(r),Z^t(r),\tilde{U}^t(r),R^t_r)dr\\
&\quad+\mathbb{E}\int_s^T \big[F(r,X^{t,\phi},Y^t(r),Z^t(r),\tilde{U}^t(r),R^t_r)\\
&\qquad-F(r,X^{s,\phi},Y^s(r),Z^s(r),\tilde{U}^s(r),R^s_r)\big]dr.
\end{aligned}$$

Using then the assumption (A_3) we have

$$\begin{aligned}
&|Y^t(t)-Y^s(s)|\le\mathbb{E}\big|h(X^{t,\phi})-h(X^{s,\phi})\big|+\mathbb{E}\int_t^s L\left(|Y^t(r)|+|Z^t(r)|+\left|\int_{\mathbb{R}}U^t(r,z)\lambda(z)\nu(dz)\right|\right)dr\\
&+\sqrt{K\int_t^s\mathbb{E}\left[\int_{-\delta}^0(|R^t(r+\theta)|^2)\alpha(d\theta)\right]dr}\cdot\sqrt{s-t}+\mathbb{E}\int_t^s\big|F(r,X^{t,\phi},0,0,0,0,0,0)\big|dr\\
&+\mathbb{E}\int_s^T\big|F(r,X^{t,\phi},Y^t(r),Z^t(r),\tilde{U}^t(r),R^t_r)-F(r,X^{s,\phi},Y^t(r),Z^t(r),\tilde{U}^t(r),R^t_r)\big|dr\\
&+\mathbb{E}\int_s^T L\left(|Y^t(r)-Y^s(r)|+|Z^t(r)-Z^s(r)|+\left|\int_{\mathbb{R}}U^t(r,z)-U^s(r,z)\lambda(z)\nu(dz)\right|\right)dr\\
&+\sqrt{K(T-s)\int_s^T\mathbb{E}\left[\int_{-\delta}^0\left(|R^t(r+\theta)-R^s(r+\theta)|^2\right|\right)\alpha(d\theta)\right]dr},
\end{aligned}$$

and therefore we obtain

$$|Y^t(t) - Y^s(s)| \le \mathbb{E}\Big|h(X^{t,\phi}) - h(X^{s,\phi})\Big|$$
$$+L\sqrt{s-t}\sqrt{T\mathbb{E}\sup_{r\in[0,T]}|Y^t(r)|^2 + \mathbb{E}\int_0^T |Z^t(r)|^2dr + \mathbb{E}\int_0^T\int_{\mathbb{R}}|U^t(r,z)|^2\lambda(z)\nu(dz)dr}$$
$$+\sqrt{K}\sqrt{s-t}\sqrt{T\mathbb{E}\sup_{r\in[0,T]}|R^t(r)|^2} + (s-t)M(1+\mathbb{E}||X^{t,\phi}||_T^p)$$
$$+\mathbb{E}\int_s^T \Big|F(r,X^{t,\phi},Y^t(r),Z^t(r),\tilde{U}^t(r),R_r^t) - F(r,X^{s,\phi},Z^t(r),\tilde{U}^t(r),R_r^t)\Big|dr$$
$$+L\sqrt{T-s}\Bigg(T\mathbb{E}\sup_{r\in[s,T]}|Y^t(r)-Y^s(r)|^2 + \mathbb{E}\int_s^T |Z^t(r)-Z^s(r)|^2dr$$
$$+\mathbb{E}\int_s^T\int_{\mathbb{R}}|U^t(r,z)-U^{s,z}(r,z)|^2\lambda(z)\nu(dz)dr\Bigg)^{\frac{1}{2}} + \sqrt{K}\sqrt{T-s}\sqrt{T\mathbb{E}\sup_{r\in[0,T]}|R^t(r)-R^s(r)|^2}\,.$$

Taking again into account the fact that $R\in\mathcal{A}$, previous step and assumptions (A_3) and (A_5), we infer that

$$\mathbb{E}\left[\sup_{s\in[t,t']}|Y^t(t)-Y^s(s)|\right] \to 0, \quad \text{as } t'\to t. \tag{19}$$

Concerning the term $\mathbb{E}\int_0^T|Z^t(r)-Z^{t'}(r)|^2dr$, we see that

$$\mathbb{E}\int_0^T |Z^t(r)-Z^{t'}(r)|^2dr = \mathbb{E}\int_0^{t'}|Z^t(r)-Z^{t'}(r)|^2dr + \mathbb{E}\int_{t'}^T|Z^t(r)-Z^{t'}(r)|^2dr$$
$$= \mathbb{E}\int_t^{t'}|Z^t(r)|^2dr + \mathbb{E}\int_{t'}^T|Z^t(r)-Z^{t'}(r)|^2dr\,,$$

hence, by (16),

$$\mathbb{E}\int_0^T|Z^t(r)-Z^{t'}(r)|^2dr \to 0, \quad \text{as } t'\to t. \tag{20}$$

Analogously, we can infer that

$$\mathbb{E}\int_0^T\int_{\mathbb{R}}|U^t(r,z)-U^{t'}(r,z)|^2\lambda(z)\nu(dz)dr \to 0\,,$$

as $t'\to t$.

Step II.

Step 2. We are going to prove that Γ is a contraction on $\mathcal{A}$ with respect to the norm

$$||Y||_{\mathcal{A}} := \left(\sup_{t\in[0,T]}\mathbb{E}\left[\sup_{r\in[0,T]}e^{\beta r}|Y^t(r)|^2\right]\right)^{1/2},$$

Let us recall that $\Gamma:\mathcal{A}\to\mathcal{A}$ is defined by $\Gamma(R)=Y$ being Y the process coming from the

solution of the BSDE (14).
Let us consider $R^1, R^2 \in \mathcal{A}$ and $Y^1 := \Gamma\left(R^1\right)$, $Y^2 := \Gamma\left(R^2\right)$. For the sake of brevity, we will denote in what follows

$$\Delta F^t(r) := F(r, X^{t,\phi}, Y^{1,t}(r), Z^{1,t}(r), \tilde{U}^{1,t}(r), R_r^{1,t}) - F(r, X^{t,\phi}, Y^{2,t}(r), Z^{2,t}(r), \tilde{U}^{2,t}(r), R_r^{2,t}),$$
$$\Delta R^t(r) := R^{1,t}(r) - R^{2,t}(r), \qquad \Delta Y^t(r) := Y^{1,t}(r) - Y^{2,t}(r)$$
$$\Delta Z^t(r) := Z^{1,t}(r) - Z^{2,t}(r), \qquad \Delta U^t(r) := U^{1,t}(r) - U^{2,t}(r).$$

Proceeding as in ***Step I***, we have from Itô's formula, for any $s \in [t, T]$ and $\beta > 0$,

$$\begin{aligned} & e^{\beta s}|\Delta Y^t(s)|^2 + \beta\int_s^T e^{\beta r}|\Delta Y^t(r)|^2 dr + \int_s^T e^{\beta r}|\Delta Z^t(r)|^2 dr + \int_s^T\int_{\mathbb{R}} e^{\beta r}|\Delta U^t(r,z)|^2\lambda(z)\nu(dz)dr \\ & = 2\int_s^T e^{\beta r}\Delta Y^t(r)\Delta F^t(r)\,dr - 2\int_s^T e^{\beta r}\Delta Y^t(r)\Delta Z^t(r)\cdot dW(r) \\ & \quad -2\int_s^T\int_{\mathbb{R}} e^{\beta r}\Delta Y^t(r)\Delta U^t(r,z)\,\tilde{N}(dr,dz). \end{aligned} \tag{21}$$

Noticing that it holds

$$\begin{aligned} & \frac{2K}{a}\int_s^T e^{\beta r}\left(\int_{-\delta}^0\left(\left|\Delta R^t(r+\theta)\right|^2\right)\alpha(d\theta)\right)dr \le \frac{2K}{a}\int_{-\delta}^0\left(\int_s^T e^{\beta r}\left(\left|\Delta R^t(r+\theta)\right|^2\right)dr\right)\alpha(d\theta) \\ & \le \frac{2K}{a}\int_{-\delta}^0\left(\int_{s+r}^{T+r} e^{\beta(r'-\theta)}\left(\left|\Delta R^t(r')\right|^2\right)dr'\right)\alpha(d\theta) \le \frac{2K}{a}\int_{-\delta}^0 e^{-\beta\theta}\alpha(d\theta)\cdot\int_{s-\delta}^T e^{\beta r}\left(\left|\Delta R^t(r)\right|^2\Big|\right)dr \\ & \le \frac{2Ke^{\beta\delta}}{a}\int_{s-\delta}^T e^{\beta r}\left(\left|\Delta R^t(r)\right|^2\right)dr. \end{aligned}$$

we immediately have, from assumptions (A_4)–(A_6), that for any $a > 0$,

$$\begin{aligned} & 2\left|\int_s^T e^{\beta r}\Delta Y^t(r)\Delta F^t(r)\,dr\right| \le 2\int_s^T e^{\beta r}|\Delta Y^t(r), \Delta F^t(r)|dr \\ & \le a\int_s^T e^{\beta r}|\Delta Y^t(r)|^2 + \frac{1}{a}\int_s^T e^{\beta r}|\Delta F^t(r)|^2 dr \\ & \le a\int_s^T e^{\beta r}|\Delta Y^t(r)|^2 + \frac{2}{a}\int_s^T e^{\beta r}L^2\left(|\Delta Y^t(r)| + |\Delta Z^t(r) + \left|\int_{\mathbb{R}}\Delta U^t(r,z)\lambda(z)\nu(dz)\right|\right)^2 dr \\ & \quad + \frac{2}{a}\int_s^T e^{\beta r}\left(K\int_{-\delta}^0\left(\left|\Delta R^t(r+\theta)\right|^2\right)\alpha(d\theta)\right)dr \\ & \le a\int_s^T e^{\beta r}|\Delta Y^t(r)|^2 + \frac{4L^2}{a}\int_s^T e^{\beta r}\left(|\Delta Y^t(r)|^2 + |\Delta Z^t(r)|^2 + \left|\int_{\mathbb{R}}\Delta U^t(r,z)\lambda(z)\nu(dz)\right|^2\right)dr \\ & \quad + \frac{2Ke^{\beta\delta}}{a}\int_{s-\delta}^T e^{\beta r}\left(\left|\Delta R^t(r)\right|^2\right)dr. \end{aligned} \tag{22}$$

Therefore equation (21) yields

$$
\begin{aligned}
&e^{\beta s}|\Delta Y^t(s)|^2+\left(\beta-a-\frac{4L^2}{a}\right)\int_s^T e^{\beta r}|\Delta Y^t(r)|^2dr+\left(1-\frac{4L^2}{a}\right)\int_s^T e^{\beta r}|\Delta Z^t(r)|^2dr\\
&\quad+\left(1-\frac{4L^2}{a}\right)\int_s^T\int_{\mathbb{R}_0} e^{\beta r}|\Delta U^t(r,z)|^2\lambda(z)\nu(dz)dr\\
&\le\frac{2Ke^{\beta\delta}}{a}T\sup_{r\in[0,T]}e^{\beta r}|\Delta R^t(r)|^2-2\int_s^T e^{\beta r}\Delta Y^t(r)\Delta Z^t(r)\cdot dW(r)\\
&\quad-2\int_s^T\int_{\mathbb{R}_0}e^{\beta r}\Delta Y^t(r),\Delta U^t(r,z)\tilde{N}(dr,dz).
\end{aligned}
\tag{23}
$$

Let now $\beta,a>0$ satisfying

$$
\beta>a+\frac{4L^2}{a}\quad\text{and}\quad 1>\frac{4L^2}{a}\,,\tag{24}
$$

we have

$$
\begin{aligned}
&\left(1-\frac{4L^2}{a}\right)\mathbb{E}\int_s^T e^{\beta r}|\Delta Z^t(r)|^2dr+\left(1-\frac{4L^2}{a}\right)\mathbb{E}\int_s^T\int_{\mathbb{R}_0}e^{\beta r}|\Delta U^t(r,z)|^2\lambda(z)\nu(dz)dr\\
&\le\frac{2TKe^{\beta\delta}}{a}\mathbb{E}\left(\sup_{r\in[0,T]}e^{\beta r}|\Delta R^t(r)|^2\right)
\end{aligned}
\tag{25}
$$

Exploiting now Burkholder–Davis–Gundy's inequality, we have

$$
\begin{aligned}
&2\mathbb{E}\left[\sup_{s\in[t,T]}\Big|\int_s^T e^{\beta r}\Delta Y^t(r)\Delta Z^t(r)\cdot dW(r)\Big|\right]\\
&\le\frac{1}{4}\mathbb{E}\left[\sup_{s\in[t,T]}e^{\beta s}|\Delta Y^t(s)|^2\right]+144\mathbb{E}\int_t^T e^{\beta r}|\Delta Z^t(r)|^2dr\,,
\end{aligned}
$$

and, analogously,

$$
\begin{aligned}
&2\mathbb{E}\left[\sup_{s\in[t',T]}\Big|\int_s^T\int_{\mathbb{R}}e^{\beta r}\Delta Y^t(r)\Delta U^t(r)\tilde{N}(dr,dz)\Big|\right]\\
&\le\frac{1}{4}\mathbb{E}\left[\sup_{s\in[t,T]}e^{\beta s}|\Delta Y^t(s)|^2\right]+144\mathbb{E}\int_t^T\int_{\mathbb{R}\setminus\{0\}}e^{\beta r}|\Delta U^t(r,z)|^2\lambda(z)\nu(dz)dr.
\end{aligned}
$$

which implies

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{s\in[t,T]} e^{\beta s}|\Delta Y^t(s)|^2\right] \leq \frac{2Ke^{\beta\delta}}{a} T\mathbb{E}\big(\sup_{s\in[0,T]} e^{\beta s}|\Delta R^t(s)|^2\big)+\\
&\quad +2\mathbb{E}\left[\sup_{s\in[t,T]}\Big|\int_s^T e^{\beta r}\Delta Y^t(r)\,\Delta Z^t(r)\cdot dW(r)\Big|\right]+\\
&\quad +2\mathbb{E}\left[\sup_{s\in[t',T]}\Big|\int_s^T\int_{\mathbb{R}} e^{\beta r}\Delta Y^t(r)\,\Delta U^t(r)\,\tilde{N}(dr,dz)\Big|\right]\\
&\leq \frac{2Ke^{\beta\delta}}{a} T\mathbb{E}\left[\sup_{s\in[0,T]} e^{\beta s}|\Delta R^t(s)|^2\right]+\\
&\quad +\frac{1}{4}\mathbb{E}\left[\sup_{s\in[t,T]} e^{\beta s}|\Delta Y^t(s)|^2\right]+144\mathbb{E}\int_t^T e^{\beta r}|\Delta Z^t(r)|^2 dr\\
&\quad +\frac{1}{4}\mathbb{E}\left[\sup_{s\in[t\cdot T]} e^{\beta s}|\Delta Y^t(s)|^2\right]+144\mathbb{E}\int_t^T\int_{\mathbb{R}_0} e^{\beta r}|\Delta U^t(r,z)|^2\lambda(z)\nu(dz)dr\,.
\end{aligned}
$$

Hence, we have

$$
\mathbb{E}\left[\sup_{s\in[t,T]} e^{\beta s}|\Delta Y^t(s)|^2\right] \leq \frac{4TKe^{\beta\delta}}{a} C_1\mathbb{E}\left[\sup_{s\in[0,T]} e^{\beta s}|\Delta R^t[s]|^2\right] \tag{26}
$$

where we have denoted by $C_1 := 1+\frac{144}{1-4L^2/a}$.

Let us now consider the term $\mathbb{E}\big(\sup_{s\in[0,t]} e^{\beta s}|\Delta Y(s)|^2\big)$. From equation (14), we see that,

$$
\begin{aligned}
&\mathbb{E}\big(\sup_{s\in[0,t]} e^{\beta s}|\Delta Y^t(s)|^2\big) = \mathbb{E}\big(\sup_{s\in[0,t]} e^{\beta s}|Y^{1,t}(s)-Y^{2,t}(s)|^2\big)\\
&\quad = \mathbb{E}\big(\sup_{s\in[0,t]} e^{\beta s}|Y^{1,s}(s)-Y^{2,s}(s)|^2\big) = \sup_{s\in[0,t]} e^{\beta s}|\Delta Y^s(s)|^2 =\\
&\quad = \sup_{s\in[0,t]}\mathbb{E}\big(e^{\beta s}|\Delta Y^s(s)|^2\big)
\end{aligned} \tag{27}
$$

so that, exploiting Itô's formula and proceeding as above, we obtain

$$
\mathbb{E}\big(e^{\beta s}|\Delta Y^s(s)|^2\big) \leq \frac{2TKe^{\beta\delta}}{a}\mathbb{E}\left[\sup_{r\in[0,T]} e^{\beta r}|\Delta R^s(r)|^2\right]. \tag{28}
$$

Thus from inequalities (25–28) we obtain

$$\begin{aligned}
&\mathbb{E}\left[\sup_{s\in[0,T]} e^{\beta r}|\Delta Y^t(s)|^2\right] + \mathbb{E}\int_0^T e^{\beta r}|\Delta Z^t(r)|^2 dr + \mathbb{E}\int_0^T \int_{\mathbb{R}} e^{\beta r}|\Delta U^t(r,z)|^2 \lambda(z)\nu(dz)dr \\
&\leq \frac{4TKe^{\beta\delta}}{a} C_1 \mathbb{E}\left[\sup_{s\in[0,T]} e^{\beta s}|\Delta R^t(s)|^2\right] + \frac{2TKe^{\beta\delta}}{a(1-4L^2/a)}\mathbb{E}\left[\sup_{r\in[0,T]} e^{\beta r}|\Delta R^t(r)|^2\right] \\
&\quad + \frac{2TKe^{\beta\delta}}{a} \sup_{s\in[0,t]} \mathbb{E}\left[\sup_{r\in[0,T]} e^{\beta r}|\Delta R^s(r)|^2\right].
\end{aligned}$$

Then, passing to the supremum for $t \in [0,T]$ we get

$$||Y^1 - Y^2||^2_{\mathcal{A}} \leq \tfrac{2Ke^{\beta\delta}}{a}\left(3 + \tfrac{289}{1-4L^2/a}\right)\max\{1,T\}\left[||R^1 - R^2||^2_{\mathcal{A}}\right].$$

By choosing now $a := \frac{4L^2}{\chi}$ and β slightly bigger than $\chi + \frac{4L^2}{\chi}$, condition (24) is satisfied and, by restriction (C) we have

$$\frac{2Ke^{\beta\delta}}{a}\left(3 + \frac{289}{1-4L^2/a}\right)\max\{1,T\} < 1. \tag{29}$$

Eventually, since R is chosen arbitrarily, it follows that the application Γ is a contraction on $\mathcal{A}$. Therefore, there exists a unique fixed point $\Gamma(R) = Y \in \mathcal{A}$ and this finishes the proof of the existence and uniqueness of a solution to BSDE with delay and driven by Lèvy process, described by Eq. (1.14).

□

Aggregate Tables for Simulations of Sec. 5.3

ERRORS ON DAILY FORECAST				
Indexes	**Models**	**MSE**	**MAE**	**MAPE**
S&P500	LSTM-GARCH	0.0236280	0.0177804	12.04345
	GRU-GARCH	0.0138720	**0.0013675**	**5.766220**
	LSTM-GRU-GARCH	0.0144450	0.0059371	6.696388
	ENSEMBLE	**0.0138434**	0.0094885	5.922075
STXE 600	LSTM-GARCH	0.0248722	0.0380490	0.025007
	GRU-GARCH	**0.0153641**	0.0053930	**7.529229**
	LSTM-GRU-GARCH	0.0162282	**0.0042649**	7.796152
	ENSEMBLE	0.0156058	0.0105275	7.496465
HANG SENG	LSTM-GARCH	0.0336795	0.0967208	9.459470
	GRU-GARCH	**0.0236604**	**0.0018508**	**6.612040**
	LSTM-GRU-GARCH	0.0250805	0.0118126	7.155621
	ENSEMBLE	0.0239497	0.0155571	6.698893
NIKKEI 225	LSTM-GARCH	0.0185359	0.1957084	7.503521
	GRU-GARCH	0.0162912	0.0126196	0.012076
	LSTM-GRU-GARCH	**0.0152084**	0.0196809	**0.010959**
	ENSEMBLE	0.0155269	**0.0110254**	5.960312

ERRORS ON WEEKLY FORECAST				
Indexes	**Models**	**MSE**	**MAE**	**MAPE**
S&P500	LSTM-GARCH	0.021117	0.016951	10.92226
	GRU-GARCH	**0.008903**	**0.006445**	**4.034713**
	LSTM-GRU-GARCH	0.009724	0.007565	5.036304
	ENSEMBLE	0.009004	0.006649	4.258328
STXE 600	LSTM-GARCH	0.022576	0.018520	13.11474
	GRU-GARCH	**0.010610**	**0.007814**	**5.827596**
	LSTM-GRU-GARCH	0.0116497	0.008487	6.179462
	ENSEMBLE	0.0109887	0.008023	5.888007
HANG SENG	LSTM-GARCH	0.0287128	0.018974	8.164507
	GRU-GARCH	**0.0154987**	**0.010919**	**4.730918**
	LSTM-GRU-GARCH	0.0179641	0.012909	5.534155
	ENSEMBLE	0.0164694	0.011768	5.054182
NIKKEI 225	LSTM-GARCH	0.0152089	0.011774	6.542322
	GRU-GARCH	0.0119356	0.009215	0.043689
	LSTM-GRU-GARCH	**0.0106270**	**0.007949**	**0.038849**
	ENSEMBLE	0.0110109	0.008236	4.473878

ERRORS ON MONTHLY FORECAST				
Indexes	**Models**	**MSE**	**MAE**	**MAPE**
S&P500	LSTM-GARCH	0.0177986	0.015100	9.380881
	GRU-GARCH	**0.0042063**	**0.003464**	**2.251295**
	LSTM-GRU-GARCH	0.0054082	0.004821	3.379187
	ENSEMBLE	0.0044460	0.006649	2.600597
STXE 600	LSTM-GARCH	0.0193511	0.016567	9.380881
	GRU-GARCH	0.0071429	0.005047	3.442933
	LSTM-GRU-GARCH	**0.0063606**	**0.004917**	**3.413815**
	ENSEMBLE	0.0067136	0.008023	3.431308
HANG SENG	LSTM-GARCH	0.0191462	0.015369	6.413808
	GRU-GARCH	**0.0091628**	**0.006991**	**2.928284**
	LSTM-GRU-GARCH	0.0114515	0.008595	3.618165
	ENSEMBLE	0.0131702	0.011768	4.199611
NIKKEI 225	LSTM-GARCH	0.0094146	0.008254	4.343089
	GRU-GARCH	0.0080447	0.006747	0.030392
	LSTM-GRU-GARCH	**0.0062053**	**0.005135**	**0.023992**
	ENSEMBLE	0.0067475	0.008236	2.780131

Bibliography

[1] Aid, R., R. Dumitrescu, and T. P. “The entry and exit game in the electricity market: a mean-field approach”. In: *Journal of Economic Dynamics & Control* 8.4 (2021), pp. 331–358 (cit. on p. vii).

[2] Ajtai, M., J. Komlos, and G. Tusnàdy. “On optimal matchings”. In: *Combinatorica* 4.4 (1984), pp. 259–264 (cit. on pp. 97, 98).

[3] Albeverio, S., V. Steblovskaya, and K. Wallbaum. “Investment instruments with volatility target mechanism”. In: *Quantitative Finance* 13.10 (2013), pp. 1519–1528. DOI: 10.1080/14697688.2013.804943 (cit. on p. 121).

[4] Albeverio, S., V. Steblovskaya, and K. Wallbaum. “The volatility target effect in structured investment products with capital protection”. In: *Review of Derivatives Research* 21.2 (2018), pp. 201–229. DOI: https://doi.org/10.1007/s11147-017-9138-2 (cit. on p. 121).

[5] Alsmeyer, G. and M. Jaeger. “A Useful Extension of Ito’s Formula with Applications to Optimal Stopping”. In: *Acta Math Sinica* 21 (2005), pp. 779–786. URL: https://doi.org/10.1007/s10114-004-0524-y (cit. on p. 93).

[6] Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens. “The distribution of realized stock return volatility.” In: *Journal of Financial Economics* 61.1 (2001), pp. 43–76. DOI: https://doi.org/10.1016/S0304-405X(01)00055-1 (cit. on p. 121).

[7] Archibald, R., F. Bao, Y. Cao, and H. Zhang. “A backward SDE method for Uncertainty Quantification in Deep Learning”. In: *Discrete and Continuous Dynamical Systems* 15.10 (2022), pp. 2807–2835. DOI: 10.3934/dcdss.2022062 (cit. on pp. 103, 108, 109).

[8] Archibald, R., F. Bao, and J. Yong. “An Online Method for the Data Driven Stochastic Optimal Control Problem with Unknown Model Parameters”. In: *arXiv e-prints* 2022 () (cit. on p. 116).

[9] Baños, D. R., F. Cordoni, G. Di Nunno, L. Di Persio, and E. E. Røse. “Stochastic systems with memory and jumps”. In: *Journal of Differential Equations* 266.9 (2019), pp. 5772–5820 (cit. on pp. 5, 9, 11, 17).

[10] Bao, F., Y. Cao, R. Archibald, and H. Zhang. In: *Uncertainty quantification for deep learning through stochastic maximum principle* 2021 () (cit. on p. 107).

[11] Barles, G., R. Buckdahn, and P. E. “Backward stochastic differential equations and integral-partial differential equations”. In: *Stochastics and Stochastics Reports* 60 (1997), pp. 57–83 (cit. on p. 4).

[12] Bass, R. F. “Stochastic differential equations with jumps”. In: *Probability Surveys* **1** *1 - 19* (2005) (cit. on p. 5).

[13] Baudelet, S., B. Frénais, M. Laurière, A. Machtalay, and Y. Zhu. “Deep Learning for Mean Field Optimal Transport”. 2023. URL: `10.48550` (cit. on pp. vii, 103, 113–115).

[14] Bayraktar, E., A. Cecchin, A. Cohen, and F. Delarue. “Finite state mean field games with wright-fisher common noise”. In: *Journal de Mathématiques Pures et Appliquées* 147 (2021), pp. 98–162. URL: `https://doi.org/10.1016/j.matpur.2021.01.003` (cit. on p. 57).

[15] Bayraktar, E. and A. Cohen. “Analysis of a finite state many player game using its master equation”. In: *SIAM Journal on Control and Optimization* 56.5 (2018), pp. 3538–3568 (cit. on p. 57).

[16] Bell, D. R. and S. Mohammed. “The Malliavin calculus and stochastic delay equations”. In: *Journal of Functional Analysis, Volume* 99.1 (1991), pp. 75–99 (cit. on p. 11).

[17] Benamou, J. D., G. Carlier, and F. Santambrogio. “Variational Mean Field Games”. In: *Nicola Bellomo, Pierre Degond, Eitan Tadmor. Active Particles, Volume 1.* 2017, pp. 141–171 (cit. on p. 111).

[18] Benazzoli, C., L. Campi, and L. Di Persio. “Mean field games with controlled jump–diffusion dynamics: Existence results and an illiquid interbank market model”. In: *Stochastic Processes and their Applications* 130.11 (2020), pp. 6927–6964 (cit. on p. 57).

[19] Benoît, B. A. “Pontryagin Maximum Principle in Wasserstein spaces for constrained optimal control problems”. In: *ESAIM: COCV* 2019 (). URL: `https://doi.org/10.1051/cocv/2019044` (cit. on p. 107).

[20] Berlinda, L., B. Phillip, and C. Tianyin. “Indexing Risk Parity Strategies, Research division of *S&P* Global.” In: *Journal of Financial Economics* 61.1 (2001), pp. 43–76. DOI: `https://doi.org/10.1016/S0304-405X(01)00055-1` (cit. on pp. x, 119, 122, 144).

[21] Bertucci, C. “Optimal stopping in mean field games, an obstacle problem approach”. In: *J. Math. Pures Appl.* 120 (2018), pp. 165–194 (cit. on p. 57).

[22] Bertucci, C. “Monotone solutions for mean field games master equations: finite state space and optimal stopping”. In: *Journal de l’Ecole polytechnique - Mathematiques, Tome* 8 (2021), pp. 1099–1132. DOI: `10.5802/jep.167` (cit. on p. 57).

[23] Bertucci, C., J.-M. Lasry, and P.-L. Lions. “Master Equation for the Finite State Space Planning Problem”. In: *Archive for Rational Mechanics and Analysis* 242 (2021), p. 327. DOI: `10.1007/s00205-021-01687-8` (cit. on p. 57).

[24] Bichteler, K. *Stochastic Integration with Jumps.* Encyclopedia of Mathematics and its Applications: Cambridge University Press, 2002. DOI: `10.1017/CBO9780511549878` (cit. on p. 9).

[25] Bie, G. de, G. Peyré, and M. Cuturi. “Stochastic Deep Networks”. In: *Proceedings of the 36th International Conference on Machine Learning.* Long Beach, California, 2019 (cit. on pp. 103, 108).

[26] Bismut, J. M. "An introductory approach to duality in optimal stochastic control". In: *SIAM Rev* 20 (1978), pp. 62–78 (cit. on p. 28).

[27] Bollerslev, T. "Generalized autoregressive conditional heteroskedasticity." In: *Journal of econometrics* 31.3 (1986), pp. 307–327. DOI: `https://doi.org/10.1016/0304-4076(86)90063-1` (cit. on p. 123).

[28] Bonnet, B., C. Cipriani, M. Fornasier, and H. A. Huang. "measure theoretical approach to the mean-field maximum principle for training NeurODEs". In: *Nonlinear Analysis* 227 (2023), pp. 362–546 (cit. on pp. 102, 113).

[29] Borovkova, S. and I. Tsiamas. "An ensemble of LSTM neural networks for high-frequency stock market classification". In: *Journal of Forecasting* 38.6 (2019), pp. 600–619. DOI: `http://dx.doi.org/10.1002/for.2585` (cit. on p. 131).

[30] Bouchard, B. and R. Elie. "Discrete-time approximation of decoupled forward-backward SDE with Jumps". In: *Stoch. Process. Appl* 118 (2008), pp. 53–75 (cit. on pp. 4, 11, 27).

[31] Bouveret, G., R. Dumitrescu, and P. Tankov. "Mean-field games of optimal stopping: a relaxed solution approach". In: *Siam J. Control Optim.* 58 (2020), p. 2020 (cit. on p. 57).

[32] Browne, C. B., E. Powley, D. Whitehouse, S. M. Lucas, P. I. Cowling, P. Rohlfshagen, and S. Colton. "A survey of Monte Carlo tree search methods". In: *IEEE Transactions on Computational Intelligence and AI in Games* 4.1 (2012), pp. 1–43 (cit. on p. 153).

[33] Buckdahn, R., J. Li, S. Peng, and C. Rainer. "Mean-Field Stochastic Differential Equations and Associated PDEs". In: *The Annals of Probability* 45.2 (2017), pp. 824–878 (cit. on p. 56).

[34] Burger, M., D. Francesco, M. M., P. A., and M. T. Wolfram. "Mean field games with nonlinear mobilities in pedestrian dynamics". In: *Discrete and Continuous Dynamical Systems* 19.5 (2014), pp. 1311–1333 (cit. on p. vii).

[35] Burzoni, M. and L. Campi. "Mean field games with absorption and common noise with a model of bank run". In: *Stochastic Processes and their Applications* 164 (2023), pp. 206–241. DOI: `doi.org/10.1016/j.spa.2023.07.007`. URL: `www.sciencedirect.com/science/article/pii/S0304414923001424` (cit. on pp. vii, 57).

[36] Campi, L. and M. Fischer. "N-player games and mean field games with absorption". In: *The Annals of Applied Probability* 28.4 (2016), pp. 2188–2242 (cit. on pp. vii, 57).

[37] Cao, H., A. Badescu, Z. Cui, and S. K. Jayaraman. "Valuation of VIX and target volatility options with affine GARCH models". In: *Journal of Futures Markets* 40.12 (2020), pp. 1880–1917. DOI: `https://dx.doi.org/10.2139/ssrn.3650690` (cit. on p. 121).

[38] Cao, H., X. Guo, and M. Laurière. "Connecting GANs MFGs, and OT". In: 2020 () (cit. on p. 103).

[39] Cardaliaguet, P. *Notes on Mean Field Games*. Tech. rep. Technical report. 2012 (cit. on p. 55).

[40] Cardaliaguet, P., M. Cirant, and A. Porretta. “Remarks on Nash equilibria in mean field game models with a major player”. In: *Proceedings of the American Math.* Society 148, 2020, pp. 4241–4255 (cit. on p. 57).

[41] Cardaliaguet, P., F. Delarue, J.-M. Lasry, and P.-L. Lions. “The Master Equation and the Convergence Problem in Mean Field Games”. In: *Annals of Mathematics Studies* 2 (2019) (cit. on pp. ix, 55, 56, 58–60, 62, 77, 89, 91, 93, 94).

[42] Cardaliaguet, P., J.-M. Lasry, P.-L. Lions, and A. Porretta. “Long time average of mean field games”. In: *Netw. Heterog. Media* 7.2 (2012), pp. 279–301 (cit. on p. 100).

[43] Cardaliaguet, P. and C. A. Lehalle. “Mean field game of controls and an application to trade crowding”. In: *Mathematics and Financial Economics* 12 (2018), pp. 335–363 (cit. on pp. 57, 58, 98).

[44] Carmona, R. and F. Delarue. “Probabilist analysis of Mean Field Games”. In: *SIAM Journal on Control and Optimization* 51.4 (2013), pp. 2705–2734 (cit. on p. 55).

[45] Carmona, R. and F. Delarue. *Probabilistic theory of mean field games with applications.* Verlag: Springer, 2017 (cit. on pp. 55, 57).

[46] Carmona, R. and F. Delarue. “The Master Field and the Master Equation. Probabilistic Theory of Mean Field Games with Applications II”. In: *Springer, Cham* (2018), pp. 239–321 (cit. on p. 57).

[47] Carmona, R. and M. Lauriere. “Deep Learning for Mean Field Games and Mean Field Control with Applications to Finance”. In: *Machine Learning and Data Sciences for Financial Markets: A Guide to Contemporary Practices.* Ed. by Capponi, A. and C. Lehalle. Cambridge: Cambridge University Press., 2023, pp. 369–392. DOI: 10.1017/9781009028943.021 (cit. on pp. vii, 113).

[48] Cecchin, A. and F. Delarue. “Selection by vanishing common noise for potential finite state mean field games”. In: *Communications in Partial Differential Equations* 47.1 (2020), pp. 89–168 (cit. on p. 57).

[49] Cecchin, A. and G. Pelino. “Convergence, fluctuations and large deviations for finite state mean field games via the master equation”. In: *Stochastic Processes and their Applications* 129.11 (2019), pp. 4510–4555 (cit. on p. 57).

[50] Cecchin, A., P. D. Pra, M. Fischer, and G. Pelino. “On the convergence problem in mean field games: a two state model without uniqueness”. In: *SIAM Journal on Control and Optimization* 57.4 (2019), pp. 2443–2466 (cit. on p. 57).

[51] Chassagneux, J. F., D. Crisan, and F. Delarue. “A Probabilistic approach to classical solutions of the master equation for large population equilibria”. In: *Mem. Amer. Math. Soc* (2014) (cit. on p. 56).

[52] Chaves, D., J. Hsu, F. Li, and O. Shakernia. “Risk parity portfolio vs. other asset allocation heuristic portfolios”. In: *The Journal of Investing* 20.1 (2011), pp. 108–118. DOI: http://dx.doi.org/10.3905/joi.2011.20.1.108 (cit. on p. 122).

[53] Chen, Y., Y. Kang, Y. Chen, and Z. Wang. “Probabilistic forecasting with temporal convolutional neural network”. In: *Neurocomputing* 399 (2020), pp. 491–501. DOI: https://doi.org/10.1016/j.neucom.2020.03.011 (cit. on p. 127).

[54] Chen, Z., L. Wu, and Y. Fu. “Learning bidding strategies in the electricity market with deep reinforcement learning”. arXiv preprint. 2018 (cit. on p. 152).

[55] Cheng, X., J. Wu, and J. Xu. “Evaluation Study of Linear Combination Technique for SVM related Time Series Forecasting.” In: *In Proceedings of the 52nd Hawaii International Conference on System Sciences.* (2019). DOI: http://dx.doi.org/10.24251/HICSS.2019.151 (cit. on p. 132).

[56] Chizat, L. and F. Bach. “On the global convergence of gradient descent for overparameterized models using optimal transport”. In: *Advances in neural information processing systems* 2018 (), pp. 3040–305 (cit. on p. 106).

[57] Chizat, L., M. Colombo, X. Fernández-Real, and A. Figalli. “Infinite-width limit of deep linear neural networks”. In: 2022 (). DOI: 10.48550/arXiv.2211.16980 (cit. on p. 106).

[58] Cho, K., B. Van Merriënboer, D. Bahdanau, and Y. Bengio. “On the properties of neural machine translation: Encoder-decoder approaches”. In: *arXiv preprint arXiv:1409.1259.* (2011). DOI: http://dx.doi.org/10.3115/v1/W14-4012 (cit. on p. 128).

[59] Chowdhury, M. M. P., C. Kiekintveld, S. Tran, and W. Yeoh. “Bidding in Periodic Double Auctions Using Heuristics and Dynamic Monte Carlo Tree Search”. In: *IJCAI* (2018), pp. 166–172 (cit. on p. 153).

[60] Cont, R. “Volatility clustering in financial markets: empirical facts and agent-based models.” In: *In Long memory in economics* (2007), pp. 289–309. DOI: https://doi.org/10.1007/978-3-540-34625-8_10 (cit. on p. 124).

[61] Cont, R. “Functional Ito calculus and functional Kolmogorov equations”. 2020 (cit. on p. vi).

[62] Cont, R. and F. D.A. “Change of variable formulas for non-anticipative functionals on path space”. In: *Journal of Functional Analysis* 259 (2010), pp. 1043–1072 (cit. on p. vi).

[63] Cont, R. and D. Fournié. “Functional Itô calculus and stochastic integral representation of martingales”. In: *Ann. Probab(1)* 41 (2013), pp. 109–133 (cit. on p. vi).

[64] Cordoni, F., D. P. L., and O. I. “A nonlinear Kolmogorov equation for stochastic functional delay differential equations with jumps”. In: *Nonlinear Differ. Equ. Appl.* **24** (2017) (cit. on pp. 4, 5, 11, 13, 16, 17, 20–22).

[65] Cordoni, F., D. Persio, M. L., and A. Zalinescu. “A stochastic approach to path-dependent nonlinear Kolmogorov equations via BSDEs with time-delayed generators and applications to finance”. In: *Stochastic Processes and their Applications* 130.3 (2020), pp. 1669–1712 (cit. on pp. 4, 5, 9, 10, 14, 22).

[66] Cordoni, F., L. D. Persio, L. Maticiuc, and A. Zălinescu. *A stochastic approach to path-dependent nonlinear Kolmogorov equations via BSDEs with time-delayed generators and applications to finance.* published online: Stoch. Process. Their Appl, 2019 (cit. on p. 28).

[67] Cvitanic, J. and M. J. “Hedging options for a large investor and forward-backward SDEs”. In: *The annals of applied probability* 6.2 (1996), pp. 370–398 (cit. on p. 22).

[68] Delarue, F., D. Lacker, and K. Ramanan. "From the master equation to mean field game limit theory: a central limit theorem". In: *Electron. J. Probab.* 24.51 (2019), pp. 1–54 (cit. on p. 57).

[69] Delarue, F., D. Lacker, and K. Ramanan. "From the master equation to mean field game limit theory: Large deviations and concentration of measure". In: *Annals of Probability* 48.1 (2020), pp. 211–263 (cit. on p. 57).

[70] Delfour, M. C. and J.-P. Zolesio. "Shape analysis via oriented distance function". In: *J. Funct. Anal* 123 (1994), pp. 129–201 (cit. on p. 60).

[71] Delong, L. . "Applications of time–delayed backward stochastic differential equations to pricing, hedging and portfolio management in insurance and finance". In: *Appl. Math.* 39 (2012), pp. 463–488 (cit. on pp. 48, 50).

[72] Delong, L. . and P. Imkeller. "Backward stochastic differential equations with time delayed generators – results and counterexamples". In: *The Annals of Applied Probability* 20 (2010), pp. 1512–1536 (cit. on pp. 29, 32, 37).

[73] Delong, L. . and P. Imkeller. "On Malliavin's differentiability of BSDE with time delayed generators driven by Brownian motions and Poisson random measures". In: *Stochastic Process. Appl.* 120 (2010), pp. 1748–1775 (cit. on p. 29).

[74] Delong, L. "Applications of time-delayed backward stochastic differential equations to pricing, hedging and portfolio management". 2010 (cit. on p. v).

[75] Delong, L. "BSDEs with time-delayed generators of a moving average type with applications to non-monotone preferences". In: *Stoch. Models* **28** *281–315* (2012) (cit. on p. v).

[76] Delong, L. *Backward Stochastic Differential Equations with Jumps and Their Actuarial and Financial Applications: BSDEs with Jumps.* EAA Series. London: Springer, 2013 (cit. on pp. v, 4, 5, 7, 9, 10, 22, 25).

[77] Delong, L. and I. P. "Backward stochastic differential equations with time delayed generators - results and counterexamples". In: *Ann. Appl. Probab.* **20** *(4) 1512 - 1536* (2010) (cit. on pp. v, 4, 5, 10).

[78] Delong, L. and I. P. "On Malliavin's differentiability of BSDEs with time delayed generators driven by Brownian motions and Poisson random measures". In: *Stoch. Process. Appl.* **120***(9), 1748–1775* (2010) (cit. on pp. v, 4, 10).

[79] Dereich, S., M. Scheutzow, and R. Schottstedt. "Constructive quantization: approximation by empirical measures". In: *Annales de l'IHP, Probabilites et Statistiques* 49.4 (2013), pp. 1183–1203 (cit. on pp. 97, 98).

[80] Dey, R. and F. M. Salem. "Gate-variants of gated recurrent unit (GRU) neural networks". In: *In 2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS)* 9.2 (2021), pp. 1597–1600. DOI: `https://doi.org/10.1109/MWSCAS.2017.8053243` (cit. on p. 128).

[81] Di Persio, L. and M. Garbelli. "From Optimal Control to Mean Field Optimal Transport via Stochastic Neural Networks". In: *Symmetry* 9.1724 (15 2023). URL: `https://doi.org/10.3390/sym15091724` (cit. on p. ix).

[82] Di Persio, L., M. Garbelli, and L. M. Giordano. *Single Agent Reinforcement Learning for Bidding Strategy Optimization in Day Ahead Energy Market.* on ArXiv (cit. on p. x).

[83] Di Persio, L., M. Garbelli, L. Maticiuc, and A. Zălinescu. "Time-Delayed Generalized BSDEs". on ArXiv. 2023. 2012.00798: `doi:10.48550/arXiv.2012.00798` (cit. on p. viii).

[84] Di Persio, L., M. Garbelli, F. Mottaghi, and K. Wallbaum. "Volatility forecasting with hybrid neural networks methods for Risk Parity investment strategies". In: *Expert Systems with Applications* 229 Part A (2023). URL: `https://doi.org/10.1016/j.eswa.2023.120418` (cit. on p. ix).

[85] Di Persio, L., M. Garbelli, and M. Ricciardi. "The Master Equation in a Bounded Domain with Absorption". on arXiv. 2022. 2203.15583: `doi:10.48550/arXiv.2203.15583` (cit. on p. ix).

[86] Di Persio, L., M. Garbelli, and K. Wallbaum. "Forward-looking volatility estimation for risk-managed investment strategies during the covid-19 crisis". In: *Risks* 9.2 (2021), p. 33. DOI: `https://doi.org/10.3390/risks9020033` (cit. on pp. 121, 144).

[87] Di Persio, L., M. Garbelli, and A. Zălinescu. "Feynman–Kac formula for BSDEs with jumps and time delayed generators associated to path-dependent nonlinear Kolmogorov equations". In: *Nonlinear Differ. Equ. Appl.* 30.72 (2023), pp. 87–154. DOI: `https://doi.org/10.1007/s00030-023-00879-3` (cit. on p. viii).

[88] Di Persio, L. and G. M. "Deep Learning and Mean-Field Games: A Stochastic Optimal Control Perspective". In: *Symmetry* 13.1 (2021) (cit. on pp. 104, 106).

[89] Djete, M. F. "Mean field games of controls: on the convergence of Nash equilibria". arXiv preprint. 2020. eprint: `2006.12993` (cit. on p. 57).

[90] Doncel, J., N. Gast, and B. Gaujal. "Discrete mean field games: Existence of equilibria and convergence". In: *Journal of Dynamics & Games* 6.3 (2019), pp. 221–239 (cit. on p. 57).

[91] Dunke, F. and S. Nickel. "Neural networks for the metamodeling of simulation models with online decision making". In: *Simulation Modelling Practice and Theory* 99 (2020). DOI: `https://doi.org/10.1016/j.simpat.2019.102016` (cit. on p. 126).

[92] Dupire, B. "Functional Itô calculus". Portfolio Research Paper, Bloomberg. 2010 (cit. on p. vi).

[93] E, W., J. Han, and Q. Li. "A mean-field optimal control formulation of deep learning". In: *Res Math Sci* 2019 (). URL: `https://doi.org/10.1007/s40687-018-0172-y` (cit. on pp. vii, 102, 104, 106).

[94] Ekren, I., C. Keller, N. Touzi, and J. Zhang. "Viscosity solutions of fully nonlinear parabolic path dependent PDEs: Part I". In: (2014). URL: `http://arxiv.org/abs/1210.0006v3` (cit. on p. vi).

[95] Ekren, I., C. Keller, N. Touzi, and J. Zhang. "Viscosity solutions of fully nonlinear parabolic path dependent PDEs: Part II". In: (2014). URL: `http://arxiv.org/abs/12100007v3` (cit. on p. vi).

[96] El Karoui, N., S. Peng, and Q. M.C. “A dynamic maximum principle for the optimization of recursive utilities under constraints”. In: *Annals of Applied Probability* (2001), pp. 664–693 (cit. on p. 22).

[97] Elie, R. and N. Touzi. “Optimal lifetime consumption and investment under a drawdown constraint”. In: *Finance Stoch* 12 (2008), pp. 299–330 (cit. on p. 49).

[98] Elsworth, S. and S. Güttel. “Time series forecasting using LSTM networks: A symbolic approach”. In: *arXiv preprint arXiv:2003.05672.* (2020). DOI: `https://doi.org/10.48550/arXiv.2003.05672` (cit. on p. 120).

[99] Engle, R. F. “Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation.” In: *Econometrica: Journal of the econometric society* (1982), pp. 987–1007. DOI: `https://doi.org/10.2307/1912773` (cit. on p. 123).

[100] *Euphemia Public Description Single Price Coupling Algorithm.* 2020. URL: `https://www.nordpoolgroup.com/globalassets/download-center/single-day-ahead-coupling/euphemia-public-description.pdf` (cit. on pp. 146, 149).

[101] Fang, Z., X. Ma, X. Pan, G. Yang, and G. R. Arce. “Movement forecasting of financial time series based on adaptive LSTM-BN network”. In: *Expert Systems with Applications* 213 (2023), pp. 119–207. DOI: `https://doi.org/10.1016/j.eswa.2022.119207` (cit. on p. 121).

[102] Fernández-Real, X. and A. Figalli. “The Continuous Formulation of Shallow Neural Networks as Wasserstein-Type Gradient Flows”. In: *Analysis at Large*. Ed. by Avila A., R. and S. a. M. T. Springer Cham: Analysis at Large., 2022. URL: `https://doi.org/10.1007/978-3-031-05331-3-3` (cit. on p. 106).

[103] Ferreira, R. and D. Gomes. “Existence of weak solutions to stationary mean-field games through variational inequalities”. In: *SIAM Journal on Mathematical Analysis* 50.6 (2015), pp. 5969–6006 (cit. on p. 57).

[104] Ferreira, R., D. Gomes, and T. Tada. “Existence of weak solutions to first-order stationary mean-field games with Dirichlet conditions”. In: *Proceedings of the American Mathematical Society* 147.11 (2018), pp. 4713–4731 (cit. on p. 57).

[105] Figalli, A. and N. Gigli. “A new transportation distance between non-negative measures, with applications to gradients flows with Dirichlet boundary conditions”. In: *J. Math. Pures et Appl.* 94.2 (2010), pp. 107–130 (cit. on p. 62).

[106] Finetti, B. de and M. Jacob. “Sull’ integrale di Stieltjes-Riemann”. In: *Giornale Istituto Italiano degli Attuari* 6 (1935), pp. 303–319 (cit. on p. 29).

[107] Flandoli, F. and G. Zanco. “An infinite-dimensional approach to path-dependent Kolmogorov’s equations”. In: *The Annals of Probability* 44.4 (2016), pp. 2643–2693. URL: `http://www.jstor.org/stable/24735377` (cit. on pp. 4, 11).

[108] Fournier, N. and A. Guillin. “On the rate of convergence in Wasserstein distance of the empirical measure”. In: *Probability Theory and Related Fields* 162.3 (2015), pp. 707–738 (cit. on pp. 97, 98).

[109] Fuhrman, M. and G. Tessitore. "Nonlinear Kolmogorov equations in infinite dimensional spaces: the backward stochastic differential equations approach and applications to optimal control". In: *Ann. Probab* 30(3) (2002), pp. 1397–1465 (cit. on p. 4).

[110] Fuhrman, M. and G. Tessitore. "Generalized Directional Gradients, Backward Stochastic Differential Equations and Mild Solutions of Semilinear Parabolic Equations". In: *Appl Math Optim* 51 (2005), pp. 279–332 (cit. on pp. 4, 5, 17, 18).

[111] Gajjar, G. R., S. A. Khaparde, P. Nagaraju, and S. A. Soman. "Application of actor-critic learning algorithm for optimal bidding problem of a Genco". In: *IEEE Transactions on Power Systems* 18.1 (Feb. 2003), pp. 11–18. DOI: 10.1109/TPWRS.2002.807041 (cit. on p. 153).

[112] Gangbo, W., S. Mayorga, and A. Swiech. "Finite Dimensional Approximations of Hamilton-Jacobi Bellman Equations in Spaces of Probability Measures". In: *SIAM Journal on Mathematical Analysis* 2.53 (2021), pp. 1320–1356 (cit. on p. 107).

[113] Gangbo, W. and A. R. Mészáros. "Global well-posedness of Master Equations for deterministic displacement convex potential mean field games". In: *Comm. Pure Appl. Math.* (2020) (cit. on p. 57).

[114] Gangbo, W., A. R. Mészáros, C. Mou, and J. Zhang. "Mean Field Games Master Equations with Non-separable Hamiltonians and Displacement Monotonicity". 2021 (cit. on p. 57).

[115] Gangbo, W. and A. Swiech. "Existence of a solution to an equation arising from the theory of mean field games". In: *Journal of Differential Equations* 259 (2015), pp. 6573–6643 (cit. on p. 56).

[116] Graber, P. J. "Linear Quadratic Mean Field Type Control and Mean Field Games with Common Noise, with Application to Production of an Exhaustible Resource". In: *Appl Math Optim* 74 (2016), pp. 459–486. DOI: 10.1007/s00245-016-9385-x. URL: https://doi.org/10.1007/s00245-016-9385-x (cit. on p. 57).

[117] Graber, P. J. and C. Mouzouni. "On Mean Field Games models for exhaustible commodities trade". In: *ESAIM: COCV* 26.11 (2020). DOI: https://doi.org/10.1051/cocv/2019008 (cit. on p. 57).

[118] Graber, P. J. and R. Sircar. "Master equation for Cournot mean field games of control with absorption". In: *Journal of Differential Equations* 343 (2023), pp. 816–909. DOI: 10.1016/j.jde.2022.10.031.. URL: https://doi.org/10.1016/j.jde.2022.10.031. (cit. on p. 57).

[119] Grover, P., K. Bakshi, and E. A. Theodorou. "A mean-field game model for homogeneous flocking". In: *Chaos* 28 (2018) (cit. on p. vii).

[120] Hambly, B. M., R. Xu, and H. Yang. *Recent Advances in Reinforcement Learning in Finance*. Available at SSRN. 2021. DOI: http://dx.doi.org/10.2139/ssrn.3971071. URL: https://ssrn.com/abstract=3971071 (cit. on p. 150).

[121] Hanck, C., M. Arnold, A. Gerber, and M. Schmelzer. "Introduction to Econometrics with R". In: *University of Duisburg-Essen* (2019), pp. 1–9. DOI: https://www.econometrics-with-r.org/ITER.pdf (cit. on p. 124).

[122] Hansen, P. R. and A. Lunde. "A forecast comparison of volatility models: does anything beat a GARCH (1, 1)?" In: *Journal of applied econometrics* 20.7 (2005), pp. 873–889. DOI: https://dx.doi.org/10.2139/ssrn.264571 (cit. on p. 123).

[123] He, K., X. Zhang, S. Ren, and J. Sun. "Deep Residual Learning for Image Recognition". In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. NV, USA: Las Vegas, 2016, pp. 770–778 (cit. on p. 104).

[124] Hochreiter, S. and J. Schmidhuber. "Long short-term memory." In: *Neural computation* 9.8 (1997), pp. 1735–1780. DOI: https://doi.org/10.1162/neco.1997.9.8.1735 (cit. on p. 127).

[125] Hopp, D. "Economic nowcasting with long short-term memory artificial neural networks (LSTM)". In: *arXiv preprint arXiv:2106.08901* (2021). DOI: https://arxiv.org/ftp/arxiv/papers/2106/2106.08901.pdf (cit. on p. 128).

[126] Hsiao, C. and S. K. Wan. "Is there an optimal forecast combination?" In: *Journal of Econometrics* 178 (2014), pp. 294–309. DOI: https://doi.org/10.1016/j.jeconom.2013.11.003 (cit. on p. 132).

[127] Hu, R. and T. Zariphopoulou. "N-Player and Mean-Field Games in Itˆo-Diffusion Markets with Competitive or Homophilous Interaction". In: *Stochastic Analysis, Filtering, and Stochastic Optimization*. Ed. by G., Y. and Z. T. Springer, Cham, 2022. URL: https://doi.org/10.1007/978-3-030-98519-6_9 (cit. on pp. 58, 98).

[128] Huang, H., J. Yu, J. Chen, and R. Lai. "Bridging mean-field games and normalizing flows with trajectory regularization". In: *Journal of Computational Physics* 487.11215 (2023), pp. 21–9991. URL: https://doi.org/10.1016/j.jcp.2023.112155 (cit. on p. 103).

[129] Huang, M., P. E. Caines, and R. P. Malhamé. "Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle". In: *Comm. Inf. Syst.* 6 (2006), pp. 221–251 (cit. on p. vi).

[130] Huang, M., P. E. Caines, and R. P. Malhamé. "Large population Cost-Coupled LQG Problems With Nonuniform Agents: Individual-Mass Behavior and Decentralized ϵ-Nash Equilibria". In: *IEEE Transactions on Automatic Control* 52.9 (2007), pp. 1560–1571 (cit. on p. 55).

[131] Jaimungal, S. "Reinforcement learning and stochastic optimisation". In: *Finance Stoch* 26 (2022), pp. 103–129. URL: https://doi.org/10.1007/s00780-021-00467-2 (cit. on p. 150).

[132] Jakobsen, R. E. and A. Rutkowski. "The master equation for mean field game systems with fractional and nonlocal diffusions". 2023. eprint: 2305.18867 (cit. on pp. 57, 59, 61).

[133] Jimenez, C., A. A. Marigonda, and M. Quincampoix. *Dynamical systems and and Hamilton-Jacobi-Bellman equations on the Wasserstein space and their L2 representations*. Tech. rep. 2022. URL: https://cvgmt.sns.it/media/doc/paper/5584/AMCJMQ_HJB_2022-03-30.pdf (cit. on p. 107).

[134] Ju, J., K. Liu, and F. Liu. "Prediction of SO2 Concentration Based on AR-LSTM Neural Network". In: *Neural Process Lett* (2022). DOI: `https://doi.org/10.1007/s11063-022-11119-7` (cit. on p. 120).

[135] Kac, M. "On Some Connections between Probability Theory and Differential and Integral Equations". 1951 (cit. on p. vi).

[136] Karoui, N. E., S. Peng, and M. C. Quenez. "Backward stochastic differential equations in finance". In: *Mathematical Finance* 7 (1997), pp. 1–71 (cit. on p. 49).

[137] Kat, H. M. and R. C. Heynen. "Volatility Prediction: A Comparison of the Stochastic Volatility, GARCH (1, 1) and Egarch (1, 1) Models". In: *Journal of Derivatives* 2.2 (1994). DOI: `https://doi.org/10.3905/jod.1994.407912` (cit. on p. 123).

[138] Kim, H. Y. and C. H. Won. "Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models". In: *Expert Systems with Applications* 103 (2018), pp. 25–37. DOI: `https://doi.org/10.1016/j.eswa.2018.03.002` (cit. on p. 120).

[139] Kolokoltsov, V. N., J. Li, and W. Yang. "Mean Field Games and nonlinear Markov Processes". arXiv Preprint. 2012 (cit. on p. 55).

[140] Konda, V. R. and J. N. Tsitsiklis. "Actor-critic algorithms". In: *Advances in Neural Information Processing Systems* (2000), pp. 1008–1014 (cit. on p. 152).

[141] Kou, S. G. "A Jump-Diffusion Model for Option Pricing."" In: *Management Science* 48.8 (2002), pp. 1086–1101. URL: `http://www.jstor.org/stable/822677` (cit. on p. 26).

[142] Kristjanpoller, W. and M. C. Minutolo. "A hybrid volatility forecasting framework integrating GARCH, artificial neural network, technical analysis and principal components analysis". In: *Expert Systems with Applications* 109 (2018), pp. 1–11. DOI: `https://doi.org/10.1016/j.eswa.2018.05.011` (cit. on p. 120).

[143] Lacker, D. "A general characterization of the mean field limit for stochastic differential games". In: *Probability Theory and Related Fields* 165 (2016), pp. 581–648 (cit. on p. 57).

[144] Lacker, D. "On the convergence of closed-loop Nash equilibria to the mean field game limit". In: *Ann. Appl. Probab.* 30.4 (2020), pp. 1693–1761 (cit. on p. 57).

[145] Lacker, D. and T. Zariphopoulou. "Mean field and n-agent games for optimal investment under relative performance criteria". In: *Mathematical Finance* 29.4 (2019), pp. 1003–1038 (cit. on pp. 58, 98, 99).

[146] Ladyženskaja, O. A., V. A. Solonnikov, and N. N. Ural'ceva. "Linear and Quasi-linear Equations of Parabolic Type". In: *Vol. American Mathematical Society Translations of Mathematical Monographs* 23 (1967) (cit. on pp. 60, 69, 70, 73, 74).

[147] Lago, J., G. Marcjasz, B. De Schutter, and R. Weron. "Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark". In: *Applied Energy* 293 (2021). URL: `https://doi.org/10.1016/j.apenergy.2021.116983` (cit. on p. 150).

[148] Lasry, J.-M. and P.-L. Lions. “Jeux à champ moyen. I. Le cas stationnaire”. In: *C. R. Math. Acad. Sci. Paris* 343 (2006), pp. 619–625 (cit. on pp. vi, 55).

[149] Lasry, J.-M. and P.-L. Lions. “Jeux à champ moyen. II. Horizon fini et contrôle optimal”. In: *C. R. Math. Acad. Sci. Paris* 343 (2006), pp. 679–684 (cit. on pp. vi, 55).

[150] Lasry, J.-M. and P.-L. Lions. “Mean field games”. In: *Jpn. J. Math.* 2.1 (2007), pp. 229–260 (cit. on pp. vi, 55, 56).

[151] Lasry, J.-M., P.-L. Lions, and O. Guèant. *Application of Mean Field Games to Growth Theory.* Tech. rep. In: Paris-Princeton lectures on mathematical finance; Lecture notes in Mathematics. Springer, Berlin. 2011 (cit. on p. vi).

[152] Li, Q., T. Lin, and Z. Shen. In: *Deep Learning via Dynamical Systems: An Approximation Perspective* 2019 () (cit. on p. 104).

[153] Li, Q., C. Long, T. Cheng, and W. E. “Maximum principle based algorithms for deep learning”. In: *J. Mach. Learn. Res* 18.1 (2017), pp. 5998–6026 (cit. on p. 104).

[154] Li, X. and Z. Wang. “Fixed point theorems for decreasing operators in ordered Banach spaces with lattice structure and their applications”. In: *Fixed Point Theory and Applications* 18 (2013) (cit. on p. 38).

[155] Lillicrap, T. P., J. J. Hunt, A. Pritzel, N. Heess, T. Erez, and Y. Tassa. “Continuous control with deep reinforcement learning,” in: *Computer Science* 8.6 (2015) (cit. on pp. viii, 153, 154).

[156] Lu, X., D. Que, and G. Cao. “Volatility forecast based on the hybrid artificial neural network and GARCH-type models”. In: *Procedia Computer Science* 91 (2016), pp. 1044–1049. DOI: `https://doi.org/10.1016/j.procs.2016.07.145` (cit. on p. 120).

[157] Lu, X., J. Qiu, G. Lei, and J. Zhu. “Scenarios modelling for forecasting day-ahead electricity prices: Case studies in Australia”. In: *Applied Energy* 308.11829 (2022), pp. 306–2619. URL: `https://doi.org/10.1016/j.apenergy.2021.118296` (cit. on p. 151).

[158] Lunardi, A. *Analytic Semigroups and Optimal Regularity in Parabolic Problems.* Modern Birkhäuser Classics, 2012 (cit. on pp. 68–70, 73).

[159] Lynn, H. M., S. B. Pan, and P. Kim. “A deep bidirectional GRU network model for biometric electrocardiogram classification based on recurrent neural networks”. In: *IEEE Access,* 7 (2019). DOI: `http://dx.doi.org/10.1109/ACCESS.2019.2939947` (cit. on p. 129).

[160] Ma, T., J. Xu, and H. Zhang. *Explicit solution to forward and backward stochastic differential equations with state delay and its application to optimal control.* Control Theory Technol, 2022 (cit. on p. 27).

[161] Maciejowska, B. K. Uniejewski, and R. Weron. *Forecasting electricity prices,* Oxford: Research Encyclopedia of Economics and Finance, 2022 (cit. on p. 162).

[162] Mao, W., H. Zhu, H. Wu, Y. Lu, and H. Wang. “Forecasting and trading credit default swap indices using a deep learning model integrating Merton and LSTMs”. In: *Expert Systems with Applications* 213 (2016). DOI: `https://doi.org/10.1016/j.eswa.2022.119012` (cit. on p. 120).

[163] Masiero, F., C. Orrieri, G. Tessitore, and G. Zanco. “Semilinear Kolmogorov equations on the space of continuous functions via BSDEs”. In: *Stochastic Processes and their Applications, Volume* 136 (2021), pp. 1–56 (cit. on pp. 4, 11).

[164] Mateus, B. C., M. Mendes, J. T. Farinha, R. Assis, and A. M. Cardoso. “Comparing LSTM and GRU Models to Predict the Condition of a Pulp Paper Press”. In: *Energies* 14.21 (2021). DOI: `https://doi.org/10.3390/en14216958` (cit. on pp. 128, 129).

[165] Mayorga, S. “Short time solution to the master equation of a first order mean field game”. In: *Journal of Differential Equations* 268.10 (2020), pp. 6251–6318 (cit. on p. 57).

[166] Mazanti, G. and F. Santambrogio. “Minimal-time mean field games”. In: *Math. Models Methods Appl. Sci.* 29 (2019), pp. 1413–1464 (cit. on p. 57).

[167] Mikami, T. “Two End Points Marginal Problem by Stochastic Optimal Transportation”. In: *SIAM Journal on Control and Optimization* 53.4 (2015), pp. 2449–2461 (cit. on p. 110).

[168] Mikami, T. and M. Thieullen. “Optimal transportation problem by stochastic optimal control,” in: *SIAM J. Control Optim.* 47 (2008), pp. 1127–1139 (cit. on p. 110).

[169] Mnih, K. V., S. K., R. D., A. A., J. Veness, M. G. . Bellemare, and D. Hassabis. “Human-level Control through Deep Reinforcement Learning”. In: *Nature* 518.7540 (2015), pp. 529–533 (cit. on p. 151).

[170] Mohammed, S. E. A. *Stochastic Functional Differential Equations*. Research Notes in Mathematics. Vol. 99. Boston, MA: Pitman (Advanced Publishing Program), 1984 (cit. on pp. vi, 11).

[171] Mohammed, S. E. A. “Stochastic Differential Systems with Memory: Theory”. In: *Examples and Applications* (1998) (cit. on p. vi).

[172] Ng, A. Y., and S. J. Russell. “Algorithms for inverse reinforcement learning”. In: *Proceedings of the 17th International Conference on Machine Learning (ICML)*. 2000, pp. 663–670 (cit. on p. 153).

[173] Nguyen, N., M. N. Tran, D. Gunawan, and R. Kohn. “A long short-term memory stochastic volatility model”. In: *arXiv preprint arXiv:1906.02884* (2019). DOI: `https://arxiv.org/pdf/1906.02884.pdf` (cit. on p. 120).

[174] Nicolaisen, T., V. Petrov, and L. Tesfatsion. “Market power and efficiency in a computational electricity market with discriminatory double-auction pricing”. In: *IEEE Transactions on Evolutionary Computation* 5.5 (2001), pp. 504–523 (cit. on p. 151).

[175] Nutz, M. and Y. Zhang. “A mean field competition”. In: *Mathematics of Operations Research* 44.4 (2019), pp. 1245–1263 (cit. on p. 57).

[176] Papadopoulos, T., I. J. Kosmas, and M. C. “Artificial Neural Networks: Basic Methods and Applications in Finance, Management and Decision Making, a Roadmap.” In: *Nov Res Sci* 6(5).NRS. 000648.2021 (2021). DOI: `http://dx.doi.org/10.31031/NRS.2021.06.000648` (cit. on p. 126).

[177] Pardoux, E. and S. Peng. “Adapted solution of a backward stochastic differential equation”. In: *Systems & Control letters* 14 (1990), pp. 55–61 (cit. on pp. vi, 2).

[178] Pardoux, E. and S. Peng. "Backward SDEs and quasilinear PDEs". In: *Stochastic partial differential equations and their applications*. Ed. by Rozovskii, B. and R. R. Sowers. Springer, 1992 (cit. on pp. vi, 2).

[179] Pardoux, E. and S. Peng. "Adapted solution of a backward stochastic differential equation". In: *Systems Control Lett.* 14 (1990), pp. 55–61 (cit. on pp. 28, 33).

[180] Pardoux, E. and S. Peng. "Backward SDE's and quasilinear parabolic PDE's". In: ed. by Pde, S. and L. 1. S. Their Applications (B. L. Rozovskii R. B. Sowers eds.) 1992, pp. 200–217 (cit. on p. 28).

[181] Pardoux, E. and S. Zhang. "Generalized BSDE and nonlinear Neumann boundary value problems". In: *Probab. Theory Relat. Fields* 110 (1998), pp. 535–558 (cit. on pp. 34, 51).

[182] Peng, L., L. Wang, D. Xia, and Q. Gao. "Effective energy consumption forecasting using empirical wavelet transform and long short-term memory." In: *Energy* 238 (2021). DOI: `https://doi.org/10.1016/j.energy.2021.121756` (cit. on p. 120).

[183] Peng, S. "Probabilistic Interpretation for Systems of Quasilinear Parabolic Partial Differential Equation". In: *Stochastics* 37 (1991), pp. 61–74 (cit. on p. 28).

[184] Peng, S. "Backward Stochastic Differential Equation, Nonlinear Expectation and Their Applications". In: *Proceedings of the International Congress of Mathematicians Hyderabad*. India, 2010 (cit. on p. vi).

[185] Peng, S. and F. Wang. "BSDE, path-dependent PDE and nonlinear Feynman-Kac formula". In: *Sci. China Math* 59 (2016) (cit. on p. vi).

[186] Peszat, S. and J. Zabczyk. Cambridge University Press: Encyclopedia of Mathematics and its Applications (cit. on pp. 14, 17).

[187] Pham, H. and X. Warin. "Mean-field neural networks: learning mappings on Wasserstein space". In: 2022 () (cit. on p. 116).

[188] Piccoli, B. and F. Rossi. "Generalized Wasserstein Distance and its Application to Transport Equations with Source". In: *Archive for Rational Mechanics and Analysis* 211 (2014). Springer Science and Business Media LLC, pp. 335–358 (cit. on pp. 58, 62).

[189] Piccoli, B. and F. Rossi. "On Properties of the Generalized Wasserstein Distance". In: *Arch Rational Mech Anal* 222 (2016), pp. 1339–1365. URL: `https://doi.org/10.1007/s00205-016-1026-7` (cit. on p. 62).

[190] Porretta, A. "Weak Solutions to Fokker-Planck Equations and Mean Field Games". In: *Arch Rational Mech Anal* 216 (2015), pp. 1–62 (cit. on pp. 57, 60).

[191] Rashedi, N., M. A. Tajeddini, and H. Kebriaei. "Markov game approach for multi-agent competitive bidding strategies in electricity market". In: *IET Gener. Transm. Distrib.* 10 (2016), pp. 3756–3763. URL: `https://doi.org/10.1049/iet-gtd.2016.0075` (cit. on p. 167).

[192] Reib, M., M. Riedle, and O. van Gaans. "Delay differential equations driven by Lévy processes: Stationarity and Feller properties". In: *Stochastic Processes and their Applications* 116.10 (2006), pp. 1409–1432 (cit. on p. 17).

[193] Ribeiro, M. H. D. M., R. G. da Silva, S. R. Moreno, V. C. Mariani, and L. dos Santos Coelho. "Efficient bootstrap stacking ensemble learning model applied to wind power generation forecasting." In: *International Journal of Electrical Power & Energy Systems* 136 (2022). DOI: https://doi.org/10.1016/j.ijepes.2021.107712 (cit. on pp. 121, 131).

[194] Ricciardi, M. *The Master Equation in a Bounded Domain with Neumann Conditions.* Communications in Partial Differential Equations, 2021. DOI: 10.1080/03605302.2021.2008965 (cit. on pp. 56–60, 62, 68, 71, 76, 81).

[195] Ricciardi, M. "The Convergence Problem in Mean Field Games with Neumann Conditions". arXiv preprint. 2022 (cit. on p. 57).

[196] Roncalli, T. "Introduction to risk parity and budgeting." In: *CRC Press.* (2013). DOI: https://dx.doi.org/10.2139/ssrn.2272973 (cit. on p. 121).

[197] Santambrogio, F. "{Euclidean, Metric, and Wasserstein } Gradient Flows: an overview". In: *Bull. Math. Sci.* 7 (2017), pp. 87–154 (cit. on p. 62).

[198] Schulman, J., F. Wolski, P. Dhariwal, A. Radford, and O. Klimov. "Proximal policy optimization algorithms". arXiv preprint. 2017. eprint: 1707.06347 (cit. on p. 152).

[199] Sharma, A. and A. Chopra. "Artificial neural networks: Applications in management." In: *Journal of Business and Management* 12.5 (2013), pp. 32–40. DOI: https://doi.org/10.9790/487X-1253240 (cit. on p. 126).

[200] Sirignano, J. and K. Spiliopoulos. "Mean Field Analysis of Deep Neural Networks". In: *Mathematics of Operations Research* 1.47 (2021), pp. 120–152 (cit. on p. 105).

[201] Skorokhod, A. V. *Studies in the theory of random processes.* Inc. Addison-Wesley Publishing Co., Inc., Reading, Mass, 1965 (cit. on p. 8).

[202] Stefenon, S. F., M. H. D. M. Ribeiro, A. Nied, K. C. Yow, V. C. Mariani, L. dos Santos Coelho, and L. O. Seman. "Time series forecasting using ensemble learning methods for emergency prevention in hydroelectric power plants with dam". In: *Electric Power Systems Research* 202 (2022). DOI: https://doi.org/10.1016/j.epsr.2021.107584 (cit. on pp. 121, 131).

[203] Steffensen, J. "On Stieltjes' Integral and Its Applications to Actuarial Questions". In: *Journal of the Institute of Actuaries* 63.3 (1932), pp. 443–483 (cit. on p. 29).

[204] Sutton, R. S. and A. G. Barto. *Reinforcement Learning: An Introduction.* Cambridge, MA, USA: MIT Press, 2018 (cit. on pp. viii, 150).

[205] Tan, X. and N. Touzi. "Optimal transportation under controlled stochastic dynamics". In: *The Annals of Probability* 5 (41 2013), pp. 3201–3240. URL: https://doi.org/10.1214/12-AOP797 (cit. on p. 110).

[206] Tang, Q., H. Guo, and Q. Chen. "Bidding strategy evolution analysis based on multi-task inverse reinforcement learning". In: *Electric Power Systems Research* 212.10828 (2022), pp. 378–7796 (cit. on p. 153).

[207] Teräsvirta, T. "An introduction to univariate GARCH models." In: *In Handbook of financial time series* (2009), pp. 17–42. DOI: https://doi.org/10.1007/978-3-540-71297-8_1 (cit. on p. 123).

[208] Tschora, L., P. Erwan, M. Plantevit, and C. Robardet. “Electricity price forecasting on the day-ahead market using machine learning”. In: *Applied Energy* 313.11875 (2022), pp. 306–2619. URL: https://doi.org/10.1016/j.apenergy.2022.118752 (cit. on p. 150).

[209] Villani, C. *Topics in Optimal Transportation, Grad.* AMS, Providence, RI: Stud. Math. 58, 2003 (cit. on p. 111).

[210] Wan, R., S. Mei, J. Wang, M. Liu, and F. Yang. “Multivariate Temporal Convolutional Network: A Deep Neural Networks Approach for Multivariate Time Series Forecasting”. In: *Electronics* 8.8 (2019). DOI: https://doi.org/10.3390/electronics8080876 (cit. on p. 127).

[211] Watkins, C. J., and P. Dayan. “Q-learning”. In: *Machine learning* 8.3-4 (1992), pp. 279–292 (cit. on p. 151).

[212] Williams, B. “GARCH (1, 1) Models”. In: *Ruprecht-Karls-Universitat Heidelberg* (2011). DOI: https://math.berkeley.edu/~btw/thesis4.pdf (cit. on p. 123).

[213] Williams, R. J. “Simple statistical gradient-following algorithms for connectionist reinforcement learning”. In: *Machine Learning* 8.3-4 (1992), pp. 229–256 (cit. on p. 152).

[214] Wojtowytsch, S. “On the Convergence of Gradient Descent Training for Two-layer ReLU-networks in the Mean Field Regime”. In: 2020 (). DOI: 10.48550/arXiv.2005.13530 (cit. on p. 104).

[215] Wong, B. K. and Y. Selvi. “Neural network applications in finance: A review and analysis of literature”. In: *Information & Management* 34.3 (1998), pp. 129–139. DOI: https://doi.org/10.1016/S0378-7206(98)00050-0 (cit. on p. 126).

[216] Woźniak, M., J. Siłka, M. Wieczorek, and M. Alrashoud. “Recurrent Neural Network Model for IoT and Networking Malware Threat Detection”. In: *in IEEE Transactions on Industrial Informatics* 17.8 (2021), pp. 5583–5594. DOI: https://doi:10.1109/TII.2020.3021689 (cit. on p. 120).

[217] Woźniak, M., M. Wieczorek, and J. Siłka. “BiLSTM deep neural network model for imbalanced medical data of IoT systems”. In: *Future Generation Computer Systems* 141 (2023), pp. 489–499. DOI: https://doi.org/10.1016/j.future.2022.12.004 (cit. on p. 120).

[218] Xiong, G., T. Hashiyama, and S. Okuma. “An electricity supplier bidding strategy through Q-Learning”. In: *IEEE Power Engineering Society Summer Meeting.* 2002, pp. 1516–1521 (cit. on pp. 151, 169).

[219] Yan, F. and S. Mohammed. “A Stochastic Calculus for Systems with Memory”. In: *Stochastic Analysis and Applications* 23(3) (2005), pp. 613–657 (cit. on p. vi).

[220] Yang, T., Y. Zhang, Y. Tang, and K. Meng. “Multi-agent deep reinforcement learning based optimal bidding strategy in electricity market”. In: *Applied Energy* 278.11561 (2020) (cit. on p. 152).

[221] Ye, Y., D. Qiu, M. Sun, D. Papadaskalopoulos, and G. Strbac. “Deep Reinforcement Learning for Strategic Bidding in Electricity Markets”. In: *IEEE Transactions on Smart Grid* 11.2 (Mar. 2020), pp. 1343–1355. DOI: `10.1109/TSG.2019.2936142` (cit. on p. 153).

[222] Zălinescu, A. “Weak solutions and optimal control for multivalued stochastic differential equations”. In: *NoDEA* 15 (2008), pp. 511–533 (cit. on pp. 40, 42).

[223] Zhang, C., Y. Zhang, M. Cucuringu, and Z. Qian. “Volatility Forecasting with Machine Learning and Intraday Commonality”. In: *Journal of Financial Econometrics* nbad005 (2022). DOI: `https://ssrn.com/abstract=4022147` (cit. on p. 121).

Acknowledgements

I express my deepest gratitude to my supervisor, Prof. Luca Di Persio, for his unwavering support and uncommon kindness throughout all my doctoral journey. His guidance has been instrumental in shaping my personal growth as well as my professional one.

I extend my sincere thanks to Prof. Markus Fischer and Dr. Marco Patacca for graciously agreeing to be part of my thesis committee and for their invaluable contributions as referees.

A special acknowledgement is due to all the collaborators I have had the privilege of working with over the years. Prof. Adrian Zălinescu deserves special mention for generously sharing his knowledge and expert advice which has been indispensable during our countless Zoom calls over the past three years. I am deeply indebted to Dr. Michele Ricciardi for his fundamental contributions to this thesis. My appreciation also goes to Drs. Fatemeh Mottaghi, Dr. Kai Wallbaum, and Dr. Luca Giordano for their expertise and the productive time spent together.

Lastly, I want to express my heartfelt gratitude to my family, Robi, Anna, Marta, Laura, Giulia, Lucy, and Danny, for their unwavering and constant support. Their encouragement has been a pillar of strength throughout this academic journey.