A deep learning method for high dimensional PDE's

An application to N-agent games

by

Carlos Daniel Contreras Quiroz

Advisor: Mauricio Junca

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Master in
Mathematics

at the
Universidad de los Andes
2023

A deep learning method for high dimensional PDE's

An application to N-agent games

Carlos Daniel Contreras Quiroz

Abstract

Nada

Acknowledgements

A mi lulú y mi pancita.Nadita.

Contents

Chapter 1

Introduction

Las ecuaciones en derivadas parciales aparecen comúnmente como herramientas útiles para la modelación en múltiples disciplinas. Se encuentran frecuentemente aplicaciones en ciencias naturales como la física y biología, en diseño en ingeniería , y también en áreas como la economía y finanzas. Sin embargo, las propiedades matemáticas de las ecuaciones que aparecen son tan diversas como las áreas en que se aplican, y aunque se pueden clasificar parcialmente según algunas de sus características, no podría existir una teoría completa que describa nuestro conocimiento sobre estas.

Por otro lado, las soluciones analíticas a estos modelos generalmente no están a nuestro alcance, por lo que es necesario recurrir a métodos numéricos para obtener aproximaciones. Para esto, usualmente se recurre a métodos clásicos como diferencias finitas, elementos finitos, volúmenes finitos o métodos espectrales, para los cuales existe una amplia teoría que soporta y justifica rigurosamente su funcionamiento.

No obstante, la aplicación de estos métodos a problemas particulares a veces se restringe por propiedades especificas de la ecuación que se resuelve. Por ejemplo, los métodos mencionados sufren de la maldición de la dimensionalidad (*"the curse of dimentionality"*), esto es, su complejidad computacional escala exponencialmente en la dimensión del problema, por lo que su uso se restringe a problemas de dimensión baja ($n = 1, 2, 3, 4$). Lo anterior dificulta su implementación en aplicaciones como valoración en matemática financiera, donde la dimensión del problema está determinada por el número de activos considerados . También, su eficiencia computacional se reduce considerablemente conforme se aumenta la complejidad de los dominios en que se resuelven, o por las no-linealidades que aparecen, como es el caso de la ecuación de Navier-Stokes modelando flujos turbulentos.

Otra área en donde estos inconvenientes aparecen es en el análisis de datos y aprendizaje de maquinas. Por ejemplo, la complejidad de algunos modelos de regresión no lineal crece exponencialmente con el tamaño de los datos subyacentes. Para este tipo de problemas se han desarrollado herramientas poderosas que permiten modelar problemas en altas dimensiones y con posibles no linealidades. Entre estas, las redes neuronales han demostrado ser de gran utilidad como modelo para representar funciones con estas complejidades[1].

En consecuencia, intentando replicar el éxito obtenido con estas herramientas en aprendizaje de máquinas, recientemente han surgido nuevas perspectivas para aproximar soluciones de ecuaciones en derivadas parciales usando estas mismas herramientas. Entre estas se encuentran las PINNs (Physics Informed Neural Networks)[**PINNs**, **PINNS2**], FNO (Fourier Neural Operators)[2], y DGM (Deep Garlekin Method)[3]. La evidencia práctica muestra que estos métodos pueden proporcionar soluciones en casos donde los clásicos no [4, 5], a pesar de usualmente no competir con su eficiencia en las situaciones donde los últimos sí aplican. Además, se ha venido desarrollando un marco teórico riguroso que permite justificar su aplicación en situaciones específicas.

Chapter 2

Backward stochastic differential equations and PDEs

When addressing deterministic optimal control problems of dynamical systems, there are two approaches, one involving Bellman's dynamic programming principle, and the other relying on the Pontryagin's maximum principle. The former approach leads to a partial differential equation, the Hamilton-Jacobi-Bellman equation, to be solved for the value function and the optimal control of the process. The latter leads to a system of ordinary differential equations, one equation forward in time for the state and one backward in time for its adjoint.

The stochastic version of these problems is solved by methods analogous to those of the deterministic case. However, there are issues with desirable mathematical properties of solutions when we state them extending directly the ones proposed by deterministic methods. That is the case of the stochastic version of the Pontryagin's maximum principle, in which the backward differential equation cannot be stated directly as an SDE with terminal condition, as the solution is not guaranteed to be adapted to the filtration generated by the brownian motion.

The theory of backward stochastic differential equations (BSDEs) emerged in Bismut's [6] early work, and later generalized by Pardoux and Peng [7], as an attempt to formalize the application of the stochastic maximum principle. Here we give an introduction and compilation of results about them based on [8, 9, 10, 11], including its relation with a certain class of nonlinear parabolic partial differential equations, which will be the main tool for the method explained in the following chapters.

2.1 Backward stochastic differential equations

2.1.1 Motivation

Let's introduce the necessity for a different formulation of stochastic differential equations through an example [10]. In the usual setting for a stochastic differential equation (SDE), we specify the evolution of a $\mathbb{R}^d$-valued stochastic process X_t through its dynamics and an initial value $x_0 \in \mathbb{R}^d$(possibly random), in the form

$$X_t = x_0 + \int_0^t \mu(t, X_t)dt + \int_0^t \sigma(t, X_t)dW_t, \tag{2.1}$$

or equivalently,

$$\begin{aligned} dX_t &= \mu(t, X_t)dt + \sigma(t, X_t)dW_t \\ X_0 &= x_0, \end{aligned} \tag{2.2}$$

where W_t is a m-dimensional Brownian motion process and the stochastic integral is defined in the Ito sense.

We know that, under some Lipschitz and boundedness conditions for the drift μ and the volatility σ, the equation with initial condition (2.2) has a unique solution which is adapted with respect to the filtration $\mathbb{F} = (\mathcal{F}_t)_t$ generated by W_t.

Now, what happens if we consider the problem (2.2) with a terminal condition at time $T > 0$? Consider, for instance, the particular case with $\mu(t, X_t) = \sigma(t, X_t) = 0$, and a square-integrable $\mathcal{F}_T$-measurable random variable $\xi \in L^2(0, T)$ for which we try to solve the problem of finding a process Y_t such that

$$\begin{aligned} dY_t &= 0 \\ Y_t(T) &= \xi. \end{aligned} \tag{2.3}$$

This equation has a unique solution given by $Y(t) = \xi$, which is not necessarily $\mathcal{F}_t-$measurable for every $0 \leq t \leq T$, and therefore (2.3) may not have solution in the usual SDE sense.

Despite this, we can try to solve this problem reinterpreting the solution to (2.3) based on the following representation theorem.

Theorem 2.1.1 (Martingale representation theorem [12])**.** *Let $(M_t)_{0\leq t\leq T}$ be a continuous $\mathbb{R}^d$-valued square-integrable martingale with respect to $\mathcal{F}_t$, the augmented filtration generated by an m-dimensional Brownian motion $(W_t)_t$. Then, there is a unique $\mathbb{R}^{d\times m}$-valued $\mathcal{F}_t$-adapted stochastic process $f(s)$, with $\mathbb{E}[\int_0^T |f|^2 dt] < \infty$, such that*

$$M_t = M_0 + \int_0^t f(s)dW_s \quad \textit{for} \quad t \in [0, T], \tag{2.4}$$

where the uniqueness is interpreted in the mean squared norm.

We can intend to enforce the solution Y_t to be $\mathcal{F}_t-$measurable for every $0 \leq t \leq T$ by taking its conditional expectation with respect to the evolving σ-algebra

$$Y(t) := \mathbb{E}[\xi|\mathcal{F}_t], \tag{2.5}$$

which satisfies the terminal condition $Y(T) = \xi$, since ξ is $\mathbb{F}_T$-measurable. Thus, as a consequence of the Martingale representation theorem 2.1.1, we conclude that there exist a square-integrable $\mathcal{F}_t$-measurable process Z_t such that

$$Y(t) = Y(0) + \int_0^t Z_s dW_s \quad \text{for} \quad t \in [0, T], \tag{2.6}$$

which can be written as

$$\begin{aligned} dY_t &= Z_t dW_t \\ Y(T) &= \xi \end{aligned} \tag{2.7}$$

Therefore, problem (2.3) can be reinterpreted as in problem (2.7), that we will denote as a bacward stochastic differential equation (BSDE) , in which we seek a pair of processes (Y_t, Z_t) that will provide an adapted solution to our original problem. Indeed, the process Z_t will "steer" the system so that the process Y_t remains adapted, and is thus called a control process. It is not possible to revert time as $t \to T - t$ as the filtration goes only in one direction [13].

Finally, we can write this equation in another form. Note that (2.7) is a forward SDE problem, hence we can solve for $Y(0)$ in the integral form , and so we have

$$Y(0) = \xi - \int_0^T Z_s dW_s, \tag{2.8}$$

that is inserted in (2.6) to obtain

$$Y(t) = \xi - \int_0^T Z_s dW_s + \int_0^t Z_s dW_s = \xi - \int_t^T Z_s dW_s \quad \forall t \in [0, T], \tag{2.9}$$

which is the standard way to write the BSDE in integral form.

2.1.2 Some useful theorems

Now that we have motivated the use of BSDEs, we follow [14] to provide a formal definition and prove that under certain regularity conditions, we can ensure the existence of a solution for that kind of equation.

Let be $(\Omega, \mathcal{F}, \mathbb{P})$ a probability space and $T > 0$ a fixed horizon time. We consider a d-dimensional Brownian motion $W = (W_t)_{t\in[0,T]}$ and let $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ be the corresponding natural augmented filtration (i.e with the completeness and right continuity conditions).

Denote by $\mathbb{S}^2(0, T)$ the set of real-valued progressively measurable processes Y_t such that

$$\mathbb{E}\left[\sup_{0\le t\le T} |Y_t|^2\right] < \infty, \tag{2.10}$$

and by $\mathbb{H}^2(0, T)^d$ the set of $\mathbb{R}^d$-valued progressively measurable processes Z_t such that

$$\mathbb{E}\left[\int_0^T |Z_t|^2 dt\right] < \infty. \tag{2.11}$$

Here we consider the backward stochastic differential equation

$$\begin{aligned} dY_t &= -f(t, Y_t, Z_t)dt + Z_t \cdot dW_t \\ Y(T) &= \xi \end{aligned} \tag{2.12}$$

Definition 2.1.1. *A solution to the BSDE* (2.12) *is a pair* $(Y, Z) \in \mathbb{S}^2(0,T) \times \mathbb{H}^2(0,T)^d$ *such that*

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)\, ds - \int_t^T Z_s \cdot dW_s, \quad 0 \le t \le T \tag{2.13}$$

Now we establish an existence and uniqueness theorem for $\mathbb{R}$-valued process, which can be extended to $\mathbb{R}^d$-valued processes.

Assumptions 2.1.2. *Let* (ξ, f) *satisfy*

I. $\xi \in L^2(\Omega, \mathcal{F}_T, \mathbb{P}; \mathbb{R})$

II. $f : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ *such that*

a) $f(\cdot, t, y, z)$*, written* $f(t, y, z)$ *for simplicity, is progressively measurable for all* y, z

b) $f(t, 0, 0) \in \mathbb{H}^2[0,T]$

c) f *is uniformly Lipschitz in* (y, z)*, i.e ,there exist a constant* C_f *such that for all* $y_1, y_2 \in \mathbb{R} \times \mathbb{R}$ *and* $z_1, z_2 \in \mathbb{R}^d \times \mathbb{R}^d$ *we have*

$$|f(t, y_1, z_1) - f(t, y_2, z_2)| \le C_f (|y_1 - y_2| + |z_1 - z_2|) \quad a.s \tag{2.14}$$

Theorem 2.1.2 (Existence and uniqueness of solutions to BSDEs [14])**.** *Given a pair* (ξ, f)*, called the terminal condition and the driver of the BSDE, that satisfy the assumptions 2.1.2 , there exist a unique solution* (Y, Z) *to the backward stochastic differential equation* (2.12)*.*

To give a demonstration we will need the following inequalities about SDEs, whose proofs will be omitted.

Theorem 2.1.3 (Doob's martingale inequality [12])**.** *Let* $\{M_t\}_t \ge 0$ *be a* $\mathbb{R}^d$*-valued martingale in* $L^p(\Omega; \mathbb{R}^d)$*. Let* $[0,T]$ *be a bounded interval with* $T > 0$ *and let* $p > 1$*. Then*

$$\mathbb{E}\left[\sup_{0\le t\le T} |M_t|^p\right] \le \left(\frac{p}{p-1}\right)^p \mathbb{E}[|M_T|^p], \tag{2.15}$$

in particular, if $p = 2$*,*

$$\mathbb{E}\left[\sup_{0\le t\le T} |M_t|^2\right] \le 4\mathbb{E}[|M_T|^2]. \tag{2.16}$$

Theorem 2.1.4 (Burkholder-Davis-Gundy inequality [12])**.** *Let* $g \in L^2(\mathbb{R}^+; \mathbb{R}^{d\times m})$*. Define for* $t \ge 0$

$$x(t) = \int_0^t g(s) dW_s \quad \text{and} \quad A(t) = \int_0^t |g(s)|^2 ds$$

then, for every $p > 0$ *there exist universal positive constants* c_p, C_p*, depending only on* p*, such that the following inequalities hold*

$$c_p \mathbb{E}[|A(t)|^{\frac{p}{2}}] \le \mathbb{E}\left[\sup_{0\le s\le t} |x(s)|^p\right] \le C_p \mathbb{E}[|A(t)|^{\frac{p}{2}}], \tag{2.17}$$

in particular, if $p = 1$, we can take $c_p = \frac{1}{2}$ and $C_p = 4\sqrt{2}$

Proof of theorem 2.1.2 . Here we give a fixed point argument. To do it, lets consider a pair of process $(U, V) \in \mathbb{S}^2(0, T) \times \mathbb{H}^2(0, T)^d$ and, as in the motivation example, consider the martingale

$$M_t = \mathbb{E}\left[\xi + \int_0^T f(s, U_s, V_s) ds \middle| \mathcal{F}_t\right], \tag{2.18}$$

which is square-integrable under the hypothesis on (ξ, f). Using to the martingale representation theorem 2.1.1, we deduce the existence and uniqueness of a process $Z_s \in \mathbb{H}^2(0, T)^d$ such that

$$M_t = M_0 + \int_0^t Z_s \cdot dW_s. \tag{2.19}$$

Now, define the process Y_t for $0 \leq t \leq T$ as

$$\begin{aligned} Y_t &= \mathbb{E}\left[\xi + \int_t^T f(s, U_s, V_s) ds \middle| \mathcal{F}_t\right] = \mathbb{E}\left[\xi + \int_0^T f(s, U_s, V_s) ds - \int_0^t f(s, U_s, V_s) ds \middle| \mathcal{F}_t\right] \\ &= M_t - \int_0^t f(s, U_s, V_s) ds \end{aligned} \tag{2.20}$$

and note that from this and using (2.19), Y_t satisfies

$$\begin{aligned} Y_t &= M_0 + \int_0^t Z_s \cdot dW_s - \int_0^t f(s, U_s, V_s) ds \\ &= \xi + \int_t^T f(s, U_s, V_s) ds - \int_t^T Z_s \cdot dW_s. \end{aligned} \tag{2.21}$$

Thus, consider the function $\Phi : \mathbb{S}^2(0, T) \times \mathbb{H}^2(0, T)^d \to \mathbb{S}^2(0, T) \times \mathbb{H}^2(0, T)^d$ that maps the pair (U, V) to the pair (Y, Z) constructed as above, $\Phi(U, V) = (Y, Z)$. Note that it is well-defined as the Z process is unique, and by Doob's martingale inequality 2.1.3 we have

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} \left|\int_t^T Z_s \cdot dW_s\right|^2\right] \leq 4\mathbb{E}\left[\int_0^T |Z_s|^2 ds\right] < \infty, \tag{2.22}$$

and therefore, by assumptions I, $IIa)$ and $IIb)$, Y_t lies in $\mathbb{S}^2(0, T)$. Also note that a solution to the BSDE (2.12) is a fixed point of Φ. We will show that such fixed point exist by showing it is a contraction if we endow the $\mathbb{S}^2(0, T) \times \mathbb{H}^2(0, T)^d$ space with the metric

$$\|(Y, Z)\|_\beta = \left(\mathbb{E}\left[\int_0^T e^{\beta s}(|Y_s|^2 + |Z_s|^2) ds\right]\right)^{\frac{1}{2}}, \tag{2.23}$$

where $\beta > 0$ is a parameter to be chosen later.

To show that Φ is a contraction, let $(U, V), (U', V') \in \mathbb{S}^2(0, T) \times \mathbb{H}^2(0, T)^d$ and $(Y, Z) = \Phi(U, V)$, $(Y', Z') = \Phi(U', V')$. We denote $(\bar{U}, \bar{V}) = (U - U', V - V')$, $(\bar{Y}, \bar{Z}) = (Y - Y', Z - Z')$ and $\bar{f}_t = f(t, U_t, V_t) - f(t, U'_t, V'_t)$.

Using equation (2.21), we know that $\bar{Y}_s$ satisfies

$$\bar{Y}_s = -\int_0^t \bar{f}_s ds + \int_0^t \bar{Z}_s \cdot dW_s \tag{2.24}$$

So let's apply Ito's formula to the process $e^{\beta s}|\bar{Y}_s|^2$ between 0 and T to obtain

$$\begin{aligned} e^{\beta T}|\bar{Y_T}|^2 = |\bar{Y_0}|^2 &+ \int_0^T (\beta e^{\beta s}|\bar{Y_s}|^2 - 2e^{\beta s}\bar{Y}_s \cdot \bar{f}_s + e^{\beta s}|\bar{Z}_s|^2)ds \\ &+ \int_0^T 2e^{\beta s}\bar{Y}_s\bar{Z}_s \cdot dW_s. \end{aligned} \tag{2.25}$$

Observe that we can apply the Burkholder-Davis-Gundy inequality 2.1.4 with $p = 1$ to the following expectation of the supremum associated with the last term

$$\begin{aligned} \mathbb{E}\left[\sup_{0\le t\le T}\left|\int_0^t 2e^{\beta s}\bar{Y}_s\bar{Z}_s \cdot dW_s\right|\right] &\le 4\sqrt{2}\,\mathbb{E}\left[\left(\int_0^T 4e^{2\beta s}|\bar{Y}_s|^2|\bar{Z}_s|^2 ds\right)^{\frac{1}{2}}\right] \\ &\le 4\sqrt{2}e^{\beta T}\,\mathbb{E}\left[\sup_{0\le t\le T}|Y_t|^2 + \int_0^T |\bar{Z}_s|^2 ds\right] \\ &< \infty, \end{aligned} \tag{2.26}$$

which shows that the local martingale $\int_0^t 2e^{\beta s}\bar{Y}_s\bar{Z}_s \cdot dW_s$ is actually a uniformly integrable martingale and therefore its expected value remains constant zero. Also, note that $\bar{Y}_T = Y_T - Y'_T = \xi - \xi = 0$.

Using these facts, take the expected value to (2.25) and reorder terms to obtain

$$\begin{aligned} \mathbb{E}\left|\bar{Y_0}\right|^2 + \mathbb{E}\left[\int_0^T e^{\beta s}\left(\beta\left|\bar{Y_s}\right|^2 + \left|\bar{Z_s}\right|^2\right)ds\right] &= 2\mathbb{E}\left[\int_0^T e^{\beta s}\bar{Y}_s \cdot \bar{f}_s ds\right] \\ \le 2C_f\mathbb{E}\left[\int_0^T e^{\beta s}\left|\bar{Y_s}\right|\left(\left|\bar{U_s}\right| + \left|\bar{V_s}\right|\right)ds\right] &\quad \text{(by condition } IIc)) \\ \le 4C_f^2\mathbb{E}\left[\int_0^T e^{\beta s}\left|\bar{Y_s}\right|^2 ds\right] + \frac{1}{2}\mathbb{E}\left[\int_0^T e^{\beta s}\left(\left|\bar{U_s}\right|^2 + \left|\bar{V_s}\right|^2\right)ds\right], \end{aligned} \tag{2.27}$$

so if we choose $\beta = 1 + 4C_f^2$ and ignore the $\mathbb{E}\left|\bar{Y_0}\right|^2$ term, we obtain

$$\mathbb{E}\left[\int_0^T e^{\beta s}\left(\left|\bar{Y_s}\right|^2 + \left|\bar{Z_s}\right|^2\right)ds\right] \le \frac{1}{2}\mathbb{E}\left[\int_0^T e^{\beta s}\left(\left|\bar{U_s}\right|^2 + \left|\bar{V_s}\right|^2\right)ds\right], \tag{2.28}$$

which is $\|(\Phi(U, V))\|_\beta \le \frac{1}{2}\|(U, V)\|_\beta$, that means Φ is a contraction in a Banach space, as $\mathbb{S}^2(0, T) \times \mathbb{H}^2(0, T)^d$ is the product of Banach spaces, and therefore has a unique fixed point. ■

As in the every differential equation, there are cases where we can provide an explicit solution. The next theorem provides one for the BSDE with linear generator

Theorem 2.1.5 (Linear BSDEs [14]). *Let A_t,B_t be bounded progressively measurable processes with values in $\mathbb{R}$ and $\mathbb{R}^d$, C a process in $\mathbb{H}^2(0,T)$ and $\xi \in L^2(\omega, \mathcal{F}_T, \mathbb{P}, \mathbb{R})$. Then, the linear backward stochastic differential equation*

$$\begin{aligned} dY_t &= -(A_tY_t + Z_t \cdot B_t + C_t)dt + Z_t \cdot dW_t \\ Y_T &= \xi \end{aligned} \tag{2.29}$$

has a unique solution, and is given by the formula

$$\Gamma_t Y_t = E\left[\Gamma_T \xi + \int_t^T \Gamma_s C_s ds \mid \mathcal{F}_t\right], \tag{2.30}$$

where Γ_t is the solution to the adjoint process

$$\begin{aligned} d\Gamma_t &= \Gamma_t(A_tdt + B_t \cdot dW_t) \\ \Gamma_0 &= 1 \end{aligned} \tag{2.31}$$

Proof. First apply Ito's formula to $\Gamma_t Y_t$ to obtain

$$\begin{aligned} d(\Gamma_t Y_t) &= Y_t d\Gamma_t + \Gamma_t dY_t + d\Gamma_t dY_t \\ &= Y_t(\Gamma_t A_t dt + \Gamma_t B_t \cdot dW_t) + \Gamma_t(-(A_tY_t + Z_t \cdot B_t + C_t)dt + Z_t \cdot dW_t) \\ &\quad + \Gamma_t Z_t \cdot B_t dt \\ &= -\Gamma_t C_t dt + \Gamma_t(Z_t + Y_tB_t) \cdot dW_t, \end{aligned} \tag{2.32}$$

that can be written in integral form as

$$\Gamma_t Y_t + \int_0^t \Gamma_s C_s ds = Y_0 + \int_0^t \Gamma_s \left(Z_s + Y_s B_s\right) \cdot dW_s. \tag{2.33}$$

We will show, as in the proof of theorem 2.1.2, that the stochastic integral in the last expression, which is a local martingale, is in fact a uniformly integrable martingale. We have $\mathbb{E}\left[\sup_{0\leq t\leq T} |\Gamma_t|^2\right] < \infty$,since A_t and B_t are bounded. Also, let's denote b_∞ the upper bound on B_t, then the following inequalities hold

$$\begin{aligned} \mathbb{E}\left[\sup_{0\leq t\leq T}\left|\int_0^t \Gamma_s \left(Z_s + Y_s B_s\right) \cdot dW_s\right|\right] &\leq 4\sqrt{2}\,\mathbb{E}\left[\left(\int_0^T |\Gamma_s|^2 |Z_s + Y_s B_s|^2 ds\right)^{\frac{1}{2}}\right] \\ &\quad \text{(By BDG inequality 2.1.4)} \\ &\leq \frac{4\sqrt{2}}{2} E\left[\sup_{0\leq t\leq T} |\Gamma_t|^2 + 2\int_0^T |Z_t|^2\, dt + 2b_\infty^2 \int_0^T |Y_t|^2\, dt\right] \\ &< \infty. \end{aligned} \tag{2.34}$$

Consequently, the right-hand side of is a uniformly integrable martingale, and so, if we take expected values to the equality (2.33), we have

$$\begin{aligned} \Gamma_t Y_t + \int_0^t \Gamma_s C_s ds &= \mathbb{E}\left[\Gamma_T Y_T + \int_0^T \Gamma_s C_s ds \middle| \mathcal{F}_t\right] \\ &= \mathbb{E}\left[\Gamma_T \xi + \int_0^T \Gamma_s C_s ds \middle| \mathcal{F}_t\right] \end{aligned} \tag{2.35}$$

and, as $\int_0^t \Gamma_s C_s ds$ is $\mathcal{F}_t$-measurable, we obtain

$$\Gamma_t Y_t = \mathbb{E}\left[\Gamma_T \xi + \int_t^T \Gamma_s C_s ds \middle| \mathcal{F}_t\right], \tag{2.36}$$

that is what we wanted to prove. The control solution Z_t can be obtained by the martingale representation theorem 2.1.1 applied to this process. ∎

Finally, we state the next comparison principle for solution of BSDEs

Theorem 2.1.6 (Comparison principle for BSDEs [14])**.** *Let (ξ_1, f_1) and (ξ_2, f_2) two pairs of terminal conditions and generators satisfying assumptions 2.1.2,and let $(Y_{1,t}, Z_{1,t})$ and $(Y_{2,t}, Z_{2,t})$ the solutions to their corresponding BSDE. Suppose that*

1. $\xi_1 \leq \xi_2$ *a.s*
2. $f_1(t, Y_{1,t}, Z_{1,t}) \leq f_2(t, Y_{1,t}, Z_{1,t})$ $dt \times d\mathbb{P}$*-a.e*
3. $f_2(t, Y_{1,t}, Z_{1,t}) \in \mathbb{H}^2(0, T)$

Then $Y_{1,t} \leq Y_{2,t}$ for all $0 \leq t \leq T$, a.s. Furthermore, if $Y_{2,0} \leq Y_{1,0}$, then $Y_{1,t} = Y_{2,t}$ for $t \in [0, T]$. In particular, if $\mathbb{P}(\xi_1 < \xi_2) > 0$ or $f_1(t, \cdot, \cdot) < f_2(t, \cdot, \cdot)$ on a set with strictly positive measure $dt \times d\mathbb{P}$ then $Y_{1,0} < Y_{2,0}$.

Proof. To simplify notation, we give a proof with $d = 1$. We denote $\bar{Y}_t = Y_{2,t} - Y_{1,t}$ and $\bar{Z}_t = Z_{2,t} - Z_{1,t}$. Then $(\bar{Y}_t, \bar{Z}_t)$ satisfy the BSDE

$$\begin{aligned} d\bar{Y}_t &= -\left(\Delta_t^y \bar{Y}_t + \Delta_t^z \bar{Z}_t + \bar{f}_t\right) dt + \bar{Z}_t dW_t \\ \bar{Y}_T &= \xi_2 - \xi_1, \end{aligned} \tag{2.37}$$

where

$$\begin{aligned} \Delta_t^y &= \frac{f_2\left(t, Y_{2,t}, Z_{2,t}\right) - f_2\left(t, Y_{1,t}, Z_{2,t}\right)}{Y_{2,t} - Y_{1,t}} 1_{Y_{2,t} - Y_{1,t} \neq 0} \\ \Delta_t^z &= \frac{f_2\left(t, Y_{1,t}, Z_{2,t}\right) - f_2\left(t, Y_{1,t}, Z_{1,t}\right)}{Z_{2,t} - Z_{1,t}} 1_{Z_{2,t} - Z_{1,t} \neq 0} \\ \bar{f}_t &= f_2\left(t, Y_{1,t}, Z_{1,t}\right) - f_1\left(t, Y_{1,t}, Z_{1,t}\right). \end{aligned} \tag{2.38}$$

By assumption, f_2 is Lipschitz in y, z, hence Δ_t^y and Δ_t^z are bounded. Moreover, $\bar{f}_t \in \mathbb{H}^2(0, T)$. Therefore, the solution to (2.37) is given by theorem 2.1.5 as

$$\Gamma_t \bar{Y}_t = \mathbb{E}\left[\Gamma_T\left(\xi_2 - \xi_1\right) + \int_t^T \Gamma_s \bar{f}_s ds \middle| \mathcal{F}_t\right], \tag{2.39}$$

where Γ_t satisfies

$$\begin{aligned} d\Gamma_t &= \Gamma_t(\Delta_t^y dt + \Delta_t^z dW_t) \\ \Gamma_0 &= 1. \end{aligned} \tag{2.40}$$

Note that Γ_t is strictly positive. We conclude the stated result using that $\xi_2 - \xi_2 \geq 0$ by assumption 1), and $\bar{f}_t \geq 0$ by assumption 2).

∎

2.1.3 Forward-Backward stochastic differential equations

Now we consider a special case of backward stochastic differential equations in which the randomness of the drift enters through a process satisfying a forward stochastic differential equation. In its more general form, the problem is stated as find three processes $(X_t, Y_t, Z_t) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d}$ such that

$$\begin{aligned} dX_t &= \mu(t, X_t, Y_t, Z_t)dt + \sigma(t, X_t, Y_t, Z_t)dW_t \\ X_0 &= x, \end{aligned} \tag{2.41}$$

and

$$\begin{aligned} dY_t &= -f(t, X_t, Y_t, Z_t)dt + Z_t dW_t \\ Y_T &= g(X_T), \end{aligned} \tag{2.42}$$

where μ, σ and g are known functions.

This problem is rather difficult, as the coupling between the processes may forbid a solution to exist. There are conditions on μ, σ, g where we can establish the existence and uniqueness of solutions to the former system, but their detailed proof is very technical and thus is not presented here, see [8].

However, we can say something simpler about the decoupled case

$$\begin{aligned} dX_t &= \mu(t, X_t)dt + \sigma(t, X_t)dW_t \\ X_0 &= x, \\ dY_t &= -f(t, X_t, Y_t, Z_t)dt + Z_t dW_t \\ Y_T &= g(X_T). \end{aligned} \tag{2.43}$$

In this case, if μ and σ satisfy enough regularity conditions to ensure that a solution to the forward SDE in (2.43) exists, for example, if they are Lipschitz and bounded, then we can solve it for the process X_t and insert the solution into the backward equation in 2.43 and solve the now standard BSDE for which we have proven a solution exist. Let's establish this assertion in the following theorem, whose proof can be found in [8].

Assumptions 2.1.3. *Let* (μ, σ, f, g) *satisfy*

I. μ, σ, f, g *are uniformly Lipschitz continuous in* x

II. $\mu(\cdot, 0)$,$\sigma(\cdot, 0)$,$f(\cdot, 0, 0, 0)$ *and* $g(0)$ *are bounded*

Theorem 2.1.7. *Under assumptions 2.1.3, the uncoupled forward backward stochastic equation* (**??**) (**??**) *has a unique markovian solution* (X, Y, Z).

2.2 The Feynman-Kac formulas

Now we shall establish the connection between stochastic differential equations with parabolic linear partial differential equations and its non-linear generalization based on backward stochastic differential equations.

2.2.1 The linear Feynman-Kac formula

2.2.2 The non-linear Feynman-Kac formula

Chapter 3

Deep Learning Methods for PDEs

3.1 Deep BSDE method

3.2 Deep Backward Dynamic Programming

3.3 Deep Splitting

Chapter 4

Crowd motion modeling

Chapter 5

An application

Chapter 6

Conclusion

Appendix A

Neural Networks

Bibliography

[1] Catherine F. Higham and Desmond J. Higham. "Deep Learning: An Introduction for Applied Mathematicians". In: *SIAM Review* 61.3 (Jan. 2019), pp. 860–891. ISSN: 0036-1445, 1095-7200. DOI: 10.1137/18M1165748. URL: https://epubs.siam.org/doi/10.1137/18M1165748 (visited on 09/26/2022).

[2] Zongyi Li et al. *Fourier Neural Operator for Parametric Partial Differential Equations*. May 16, 2021. arXiv: 2010.08895[cs,math]. URL: http://arxiv.org/abs/2010.08895 (visited on 10/10/2022).

[3] Justin Sirignano and Konstantinos Spiliopoulos. "DGM: A deep learning algorithm for solving partial differential equations". In: *Journal of Computational Physics* 375 (Dec. 2018), pp. 1339–1364. ISSN: 00219991. DOI: 10.1016/j.jcp.2018.08.029. URL: https://linkinghub.elsevier.com/retrieve/pii/S0021999118305527 (visited on 10/10/2022).

[4] Salvatore Cuomo et al. *Scientific Machine Learning through Physics-Informed Neural Networks: Where we are and What's next*. June 7, 2022. arXiv: 2201.05624[physics]. URL: http://arxiv.org/abs/2201.05624 (visited on 10/05/2022).

[5] Jan Blechschmidt and Oliver G. Ernst. "Three ways to solve partial differential equations with neural networks — A review". In: *GAMM-Mitteilungen* 44.2 (June 2021). ISSN: 0936-7195, 1522-2608. DOI: 10.1002/gamm.202100006. URL: https://onlinelibrary.wiley.com/doi/10.1002/gamm.202100006 (visited on 09/26/2022).

[6] Jean-Michel Bismut. "Conjugate convex functions in optimal stochastic control". en. In: *Journal of Mathematical Analysis and Applications* 44.2 (Nov. 1973), pp. 384–404. ISSN: 0022-247X. DOI: 10.1016/0022-247X(73)90066-8. URL: https://www.sciencedirect.com/science/article/pii/0022247X73900668 (visited on 02/21/2023).

[7] E. Pardoux and S. G. Peng. "Adapted solution of a backward stochastic differential equation". en. In: *Systems & Control Letters* 14.1 (Jan. 1990), pp. 55–61. ISSN: 0167-6911. DOI: 10.1016/0167-6911(90)90082-6. URL: https://www.sciencedirect.com/science/article/pii/0167691190900826 (visited on 02/21/2023).

[8] Jianfeng Zhang. *Backward Stochastic Differential Equations*. en. Vol. 86. Probability Theory and Stochastic Modelling. New York, NY: Springer New York, 2017. ISBN: 978-1-4939-7254-8 978-1-4939-7256-2. DOI: 10.1007/978-1-4939-7256-2.

URL: http://link.springer.com/10.1007/978-1-4939-7256-2 (visited on 02/15/2023).

[9] Etienne Pardoux and Aurel R şcanu. *Stochastic Differential Equations, Backward SDEs, Partial Differential Equations*. en. Vol. 69. Stochastic Modelling and Applied Probability. Cham: Springer International Publishing, 2014. ISBN: 978-3-319-05713-2 978-3-319-05714-9. DOI: 10.1007/978-3-319-05714-9. URL: https://link.springer.com/10.1007/978-3-319-05714-9 (visited on 02/15/2023).

[10] Ricardo Romo Romero. "Maestro en ciencias con especialidad en probabilidad y estadística". es. In: ().

[11] Nizar Touzi. *Optimal Stochastic Control, Stochastic Target Problems, and Backward SDE*. en. Vol. 29. Fields Institute Monographs. New York, NY: Springer New York, 2013. ISBN: 978-1-4614-4285-1 978-1-4614-4286-8. DOI: 10.1007/978-1-4614-4286-8. URL: https://link.springer.com/10.1007/978-1-4614-4286-8 (visited on 02/15/2023).

[12] Xuerong Mao. *Stochastic differential equations and applications*. 2nd ed. OCLC: ocn176925635. Chichester: Horwood Pub, 2008. 422 pp. ISBN: 978-1-904275-34-3.

[13] Jared Chessari et al. *Numerical Methods for Backward Stochastic Differential Equations: A Survey*. Mar. 10, 2022. arXiv: 2101.08936[cs,math]. URL: http://arxiv.org/abs/2101.08936 (visited on 01/24/2023).

[14] Huyên Pham. *Continuous-time Stochastic Control and Optimization with Financial Applications*. Vol. 61. Stochastic Modelling and Applied Probability. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009. ISBN: 978-3-540-89499-5 978-3-540-89500-8. DOI: 10.1007/978-3-540-89500-8. URL: https://link.springer.com/10.1007/978-3-540-89500-8 (visited on 02/28/2023).