

# A deep learning method for high dimensional PDE's

An application to crowd motion

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 crowd motion

**Carlos Daniel Contreras Quiroz**

## **Abstract**

Nada

# Acknowledgements

A mi lulú y mi pancita.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Backward stochastic differential equations and PDEs</b>	<b>3</b>
2.1	Backward stochastic differential equations . . . . .	3
2.1.1	Motivation . . . . .	3
2.1.2	An existence and uniqueness theorem . . . . .	5
2.1.3	Forward-Backward stochastic differential equations . . . . .	7
2.2	The Feynman-Kac formulas . . . . .	7
2.2.1	The linear Feynman-Kac formula . . . . .	7
2.2.2	The non-linear Feynman-Kac formula . . . . .	7
<b>3</b>	<b>The Deep BSDE method</b>	<b>8</b>
<b>4</b>	<b>Crowd motion modeling</b>	<b>9</b>
<b>5</b>	<b>An application</b>	<b>10</b>
<b>6</b>	<b>Conclusion</b>	<b>11</b>
<b>A</b>	<b>Neural Networks</b>	<b>12</b>

## 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 específicas de la ecuación que se resuelve. Por ejemplo, los métodos mencionados sufren de la maldición de la dimensionalidad (*"the curse of dimensionality"*), 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 máquinas. 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 Galerkin 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

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

or equivalently,

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

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  $\mu$  and the volatility  $\sigma$ , 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

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

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*

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

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  $\sigma$ -algebra

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

which satisfies the terminal condition  $Y(T) = \xi$ , since  $\xi$  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

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



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 backward 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  $t \rightarrow T - t$  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 An existence and uniqueness theorem

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

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

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

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

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

**Theorem 2.1.2** (Existence and uniqueness of solutions to BSDEs [14]). *Given a pair  $(\xi, f)$ , called the terminal condition and the driver of the BSDE, that satisfy the following assumptions*

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

II.  $f : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$  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

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

There exist a unique solution  $(Y, Z)$  to the 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  $\{M_t\}_t \geq 0$  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  $p > 0$ . Then*

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

in particular, if  $p = 2$ ,

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} |M_t|^2 \right] \leq 4\mathbb{E}[|M_T|^2]. \quad (2.16)$$

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

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

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

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

*Proof of theorem 2.1.2.* Here we give a fixed point argument. ■

### 2.1.3 Forward-Backward stochastic differential equations

## 2.2 The Feynman-Kac formulas

### 2.2.1 The linear Feynman-Kac formula

### 2.2.2 The non-linear Feynman-Kac formula

# Chapter 3

The Deep BSDE method

# 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](https://doi.org/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\[cs,math\]](https://arxiv.org/abs/2010.08895). 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](https://doi.org/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\[physics\]](https://arxiv.org/abs/2201.05624). 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](https://doi.org/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](https://doi.org/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](https://doi.org/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](https://doi.org/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 ȃscanu. *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](https://doi.org/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](https://doi.org/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](https://arxiv.org/abs/2101.08936)[cs,math]. URL: <http://arxiv.org/abs/2101.08936> (visited on 01/24/2023).
  - [14] Huy  n 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](https://doi.org/10.1007/978-3-540-89500-8). URL: <https://link.springer.com/10.1007/978-3-540-89500-8> (visited on 02/28/2023).