Neural network regression for Bermudan option pricing

Bernard Lapeyre, Jérôme Lelong

▶ To cite this version:

Bernard Lapeyre, Jérôme Lelong. Neural network regression for Bermudan option pricing. 2019. hal-02183587v1

Neural network regression
for Bermudan option pricing

Bernard Lapeyre[*]   Jérôme Lelong [†]

July 15, 2019

Abstract

The pricing of Bermudan options amounts to solving a dynamic programming principle, in which the main difficulty, especially in large dimension, comes from the computation of the conditional expectation involved in the continuation value. These conditional expectations are classically computed by regression techniques on a finite dimensional vector space. In this work, we study neural networks approximation of conditional expectations. We prove the convergence of the well-known Longstaff and Schwartz algorithm when the standard least-square regression is replaced by a neural network approximation.

Key words: Bermudan options, optimal stopping, regression methods, deep learning, neural networks.

1 Introduction

We fix some finite time horizon $T > 0$ and a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \le t \le T}, \mathbb{P})$ modeling a financial market. We assume that the short interest rate is modeled by an adapted process $(r_t)_{0 \le t \le T}$ with values in $\mathbb{R}_+$ and that $\mathbb{P}$ is an associated risk neutral measure. We consider a Bermudan option with exercising dates $0 = t_0 \le T_1 < T_2 < \cdots < T_N = T$ and discounted payoff $\tilde{Z}_{T_n}$ if exercised at time T_n. For convenience, we add 0 and T to the exercising dates. This is definitely not a requirement of the method we propose here but it makes notation lighter and avoids to deal with the purely European part involved in the Bermudan option. We assume that the discrete time discounted payoff process $(Z_{T_n})_{0 \le n \le N}$ is adapted to the filtration $(\mathcal{F}_{T_n})_{0 \le n \le N}$ and that $\mathbb{E}[\max_{0 \le n \le N} |Z_{T_n}|^2] < \infty$.

_____

[*]Université Paris-Est, Cermics (ENPC), INRIA, F-77455 Marne-la-Vallée, France
email: `bernard.lapeyre@enpc.fr`
[†]Univ. Grenoble Alpes, CNRS, Grenoble INP, LJK, 38000 Grenoble, France.
email: `jerome.lelong@univ-grenoble-alpes.fr`

Standard arbitrage pricing theory defines the discounted value $(U_n)_{0 \leq n \mathbb{N}}$ of the Bermudan option at times $(T_n)_{0 \leq n \mathbb{N}}$ by

$$\begin{cases} U_{T_N} &= Z_{T_N} \\ U_{T_n} &= \max\left(Z_{T_n}, \mathbb{E}[U_{T_{n+1}}|\mathcal{F}_{T_n}]\right) \end{cases} \tag{1}$$

Using the Snell enveloppe theory, the sequence U can be proved to be given by

$$U_{T_n} = \sup_{\tau \in \mathcal{T}_{T_n, T}} \mathbb{E}[Z_\tau|\mathcal{F}_{T_n}]. \tag{2}$$

Solving the backward recursion (1) known as the dynamic programming principle has been a challenging problem for years and various approaches have been proposed to approximate its solution. The real difficulty lies in the computation of the conditional expectation $\mathbb{E}[U_{T_{n+1}}|\mathcal{F}_{T_n}]$ at each time step of the recursion. If we were to classify the different approaches, we could say that there are regression based approaches (see Tilley [1993], Carriere [1996], Tsitsiklis and Roy [2001], Broadie and Glasserman [2004]) and quantization approaches (see Bally and Pages [2003], Bronstein et al. [2013]). We refer to Bouchard and Warin [2012] and Pagès [2018] for a survey of the different techniques to price Bermudan options.

Among all the available algorithms to compute U using the dynamic programming principle, the one proposed by Longstaff and Schwartz [2001] has the favour of practitioners. Their approach is based on iteratively selecting the optimal policy. Let τ_n be the smallest optimal policy after time T_n — the smallest stopping time reaching the supremum in (2) — then

$$\begin{cases} \tau_N = T_N \\ \tau_n = T_n \mathbf{1}_{\left\{Z_{T_n} \geq \mathbb{E}[Z_{\tau_{n+1}}|\mathcal{F}_{T_n}]\right\}} + \tau_{n+1} \mathbf{1}_{\left\{Z_{T_n} < \mathbb{E}[Z_{\tau_{n+1}}|\mathcal{F}_{T_n}]\right\}} \end{cases} \tag{3}$$

All these methods based on the dynamic programming principle either as value iteration (1) or policy iteration (3) require a Markovian setting to be implemented such that the conditional expectation knowing the whole past can be replaced by the conditional expectation knowing only the value of a Markov process at the current time.

We assume that the discounted payoff process writes $Z_{T_n} = \phi_n(X_{T_n})$, for any $0 \leq n \leq N$, where $(X_t)_{0 \leq t \leq T}$ is an adapted Markov process taking values in $\mathbb{R}^r$. Hence, the conditional expectation involved in (3) simplifies into $\mathbb{E}[Z_{\tau_{n+1}}|\mathcal{F}_{T_n}] = \mathbb{E}[Z_{\tau_{n+1}}|X_{T_n}]$ and can therefore be approximated by a standard least square method.

In local volatility models, the process X is typically defined as $X_t = (r_t, S_t)$, where S_t is the price of an asset and r_t the instantaneous interest rate (only $X_t = S_t$ when the interest rate is deterministic). In the case of stochastic volatility models, X also includes the volatility process σ, $X_t = (r_t, S_t, \sigma_t)$. Some path dependent options can also fit in this framework at the expense of increasing the size of the process X. For instance, in the case of an Asian option with payoff $(\frac{1}{T}A_T - S_T)_+$ with $A_t = \int_0^t S_u du$, one can define X as $X_t = (r_t, S_t, \sigma_t, A_t)$ and then the Asian option can be considered as a vanilla option on the two dimensional but non tradable assets (S, A).

Once the Markov process X is identified, the conditional expectations can be written

$$\mathbb{E}[Z_{\tau_{n+1}}|\mathcal{F}_{T_n}] = \mathbb{E}[Z_{\tau_{n+1}}|X_{T_n}] = \psi_n(X_{T_n}) \tag{4}$$

where ψ_n solves the following minimization problem

$$\inf_{\psi \in L^2(\mathcal{L}(X_{T_n}))} \mathbb{E}\left[\left|Z_{\tau_{n+1}} - \psi(X_{T_n})\right|^2\right]$$

with $L^2(\mathcal{L}(X_{T_n}))$ being the set of all measurable functions f such that $\mathbb{E}[f(X_{T_n})^2] < \infty$. The real challenge comes from properly approximating the space $L^2(\mathcal{L}(X_{T_n}))$ by a finite dimensional space: one typically uses polynomials or local bases (see Gobet et al. [2005], Bouchard and Warin [2012]) and in any case it is always a linear regression. In this work, we use neural networks to approximate ψ_n in (4). The main difference between neural networks and the regression approaches commonly used comes from the non linearity of neural networks, which also make their strength. Note that the set of neural networks with a fixed number of layers and neurons is obviously not a vector space and not even convex. Through neural networks, this paper investigates the effects of using non linear approximations of conditional expectations in the Longstaff Schwartz algorithm.

Kohler et al. [2010] already used neural networks to approximate American options but using equation (1) instead of (3) leading to a Tsitsiklis and Roy [2001]-type algorithm. Moreover they use new samples of the whole path of the underlying process X at each time step n to prove the convergence. In our approach, we use a neural network modification of the popular Longstaff-Schwartz algorithm and we draw a set of M samples with the distribution of $(X_{T_0}, X_{T_1}, \ldots, X_{T_N})$ before starting and we use these very same samples at each time step. Therefore, we save a lot of computational time by avoiding a very costly resimulation at each time step, which very much improves the efficiency of our approach. Deep learning was also used in the context of optimal stopping by Becker et al. [2018] to parametrize the optimal policy.

The paper is organized as follows. In Section 2, we start with some preliminaries on neural networks and recall the universal approximation theorem. Then, in Section 3, we describe our algorithm, whose convergence is studied in Section 4.

2 Preliminaries on deep neural network

Deep Neural networks (DNN) aim to approximate (complex non linear) functions defined on finite-dimensional space, and in contrast with the usual additive approximation theory built via basis functions, like polynomial, they rely on composition of layers of simple functions. The relevance of neural networks comes from the universal approximation theorem and the Kolmogorov-Arnold representation theorem (see Arnold [2009], Kolmogorov [1956], Cybenko [1989], Hornik [1991]), and this has shown to be successful in numerous practical applications.

We consider the feed forward neural network — also called multilayer perceptron — for the approximation of the continuation value at each time step. From a mathematical point view, we can model a DNN by a non linear function

$$x \in \mathcal{X} \subset \mathbb{R}^r \longmapsto \Phi(x; \theta) \in \mathbb{R}$$

where Φ typically writes as function compositions. Let $L \geq 2$ be an integer, we write

$$\Phi = A_L \circ \sigma \circ A_{L-1} \circ \cdots \circ \sigma \circ A_1 \tag{5}$$

where for $\ell = 1, \ldots, L$, $A_\ell : \mathbb{R}^{d_{\ell-1} \to \mathbb{R}^{d_\ell}}$ are affine functions

$$A_\ell(x) = \mathcal{W}_\ell x + \beta_\ell \quad \text{for } x \in \mathbb{R}^{d_{\ell-1}},$$

with $\mathcal{W}_\ell \in \mathbb{R}^{d_\ell \times d_{\ell-1}}$, and $\beta_\ell \in \mathbb{R}^{d_\ell}$. In our setting, we have $d_1 = r$ and $d_L = 1$. The function σ is often called the *activation function* and is applied component wise. The number d_ℓ of rows of the matrix $\mathcal{W}_\ell$ is usually interpreted as the number of neurons of the layer ℓ. The parameter θ embeds the parameters of all the different layers and we set $\theta = (\mathcal{W}_\ell, \beta_\ell)_{\ell=1,\ldots,L} \in \mathbb{R}^{N_d}$ with $N_d = \sum_{\ell=1}^{L} d_\ell(1 + d_{\ell-1})$.

Let $L > 0$ be fixed in the following, we introduce the set $\mathcal{NN}_\infty$ of all DNN of the above form. Now, we need to restrict the maximum number of neurons per layer. Let $p \in \mathbb{N}$, $p > 1$, we denote by $\mathcal{NN}_p$ the set of neural networks with at most p neurons per layer and $L - 1$ layers and bounded parameters. More precisely, we pick an increasing sequence of positive real numbers $(\gamma_p)_p$ such that $\lim_{p \to \infty} \gamma_p = \infty$. We introduce the set

$$\Theta_p = \{\theta \in \mathbb{R} \times \mathbb{R}^p \times (\mathbb{R}^p \times \mathbb{R}^{p \times p})^{L-2} \times \mathbb{R}^p \times \mathbb{R}^{p \times p} \ : \ |\theta| \leq \gamma_p\}. \tag{6}$$

Then, $\mathcal{NN}_p$ is defined by

$$\mathcal{NN}_p = \{\Phi(\cdot; \theta) \ : \ \theta \in \Theta_p\}$$

and we have $\mathcal{NN}_\infty = \cup_{p \in \mathbb{N}} \mathcal{NN}_p$. An element of $\mathcal{NN}_p$ with be denoted by $\Phi_p(\cdot; \theta)$ with $\theta \in \Theta_p$. Note that the space $\mathcal{NN}_p$ is not a vector space, nor a convex set and therefore finding the element of $\mathcal{NN}_p$ that best approximates a given function cannot be simply interpreted as an orthogonal projection.

The use of DNN as function approximations is justified by the fundamental results of Hornik [1991]

Theorem 2.1 (Universal Approximation Theorem) *Assume that the function σ is non constant and bounded. Let μ denote a probability measure on $\mathbb{R}^r$, then for any $L \geq 2$, $\mathcal{NN}_\infty$ is dense in $L^2(\mathbb{R}^r, \mu)$.*

Theorem 2.2 (Universal Approximation Theorem) *Assume that the function σ is a non constant, bounded and continuous function, then, when $L = 2$, $\mathcal{NN}_\infty$ is dense into $C(\mathbb{R}^r)$ for the topology of the uniform convergence on compact sets.*

Remark 2.3 *We can rephrase Theorem* (2.1) *in terms of approximating random variables. Let Y be a real valued random variable s.t. $\mathbb{E}[Y^2] < \infty$. Let X be a random variable taking values in $\mathbb{R}^r$ and $\mathcal{G}$ the smallest $\sigma-$algebra such that X is $\mathcal{G}$ measurable. Then, there exists a sequence $(\theta_p)_{p \geq 2} \in \prod_{p=2}^{\infty} \Theta_p$, such that $\lim_{p \to \infty} \mathbb{E}[|Y - \Phi_p(X; \theta_p)|^2] = 0$. Therefore, if for every $p \geq 2$, $\alpha_p \in \Theta_p$ solves*

$$\inf_{\theta \in \Theta_p} \mathbb{E}[|\Phi_p(X; \theta) - Y|^2],$$

then the sequence $(\Phi_p(X; \alpha_p))_{p \geq 2}$ converges to $\mathbb{E}[Y|X]$ in $L^2(\Omega)$ when $p \to \infty$.

3 The algorithm

3.1 Description of the algorithm

We aim at solving the following dynamic programming equation on the optimal policy

$$\begin{cases} \tau_N = T_N \\ \tau_n = T_n \mathbf{1}_{\left\{ Z_{T_n} \geq \mathbb{E}[Z_{\tau_{n+1}} | \mathcal{F}_{T_n}] \right\}} + \tau_{n+1} \mathbf{1}_{\left\{ Z_{T_n} < \mathbb{E}[Z_{\tau_{n+1}} | \mathcal{F}_{T_n}] \right\}}, \text{ for } 1 \leq n \leq N-1 \end{cases} \tag{7}$$

Then, the time-0 price of the Bermudan option writes

$$U_0 = \max(Z_0, \mathbb{E}[Z_{T_1}]).$$

The difficulty in solving this dynamic programming equation comes from the computation of the conditional expectation at each time step. The idea proposed by Longstaff and Schwartz [2001] was to approximate the conditional expectation by a regression problem on a well chosen set of functions. In this work, we use a DNN to perform this approximation.

$$\begin{cases} \tau_N^p = T_N^p \\ \tau_n^p = T_n \mathbf{1}_{\left\{ Z_{T_n} \geq \Phi_p(X_{T_n}; \theta_n^p) \right\}} + \tau_{n+1} \mathbf{1}_{\left\{ Z_{T_n} < \Phi_p(X_{T_n}; \theta_n^p) \right\}}, \text{ for } 1 \leq n \leq N-1 \end{cases} \tag{8}$$

where θ_n^p solves the following optimization problem

$$\inf_{\theta \in \Theta_p} \mathbb{E}\left[\left| \Phi_p(X_{T_n}; \theta) - Z_{\tau_{n+1}^p} \right|^2 \right]. \tag{9}$$

Since the conditional expectation operator is an orthogonal projection, we have

$$\mathbb{E}\left[\left| \Phi_p(X_{T_n}; \theta) - Z_{\tau_{n+1}^p} \right|^2 \right] = \mathbb{E}\left[\left| \Phi_p(X_{T_n}; \theta) - \mathbb{E}\left[Z_{\tau_{n+1}^p} | \mathcal{F}_{T_n} \right] \right|^2 \right]$$
$$+ \mathbb{E}\left[\left| Z_{\tau_{n+1}^p} - \mathbb{E}\left[Z_{\tau_{n+1}^p} | \mathcal{F}_{T_n} \right] \right|^2 \right].$$

Therefore, any minimizer in (9) is also a solution to the following minimization problem

$$\inf_{\theta \in \Theta_p} \mathbb{E}\left[\left| \Phi_p(X_{T_n}; \theta) - \mathbb{E}\left[Z_{\tau_{n+1}^p} | \mathcal{F}_{T_n} \right] \right|^2 \right]. \tag{10}$$

The standard approach is to sample a bunch of paths of the model $X_{T_0}^{(m)}, X_{T_1}^{(m)}, \ldots, X_{T_N}^{(m)}$ along with the corresponding payoff paths $Z_{T_0}^{(m)}, Z_{T_1}^{(m)}, \ldots, Z_{T_N}^{(m)}$, for $m = 1, \ldots, M$. To compute the τ_n's on each path, one needs to compute the conditional expectations $\mathbb{E}[Z_{\tau_{n+1}} | \mathcal{F}_{T_n}]$ for $n = 1, \ldots, N-1$. Then, we introduce the final approximation of the backward iteration policy, in which the truncated expansion is computed using a Monte Carlo approximation

$$\begin{cases} \widehat{\tau}_N^{p,(m)} = T_N \\ \widehat{\tau}_n^{p,(m)} = T_n \mathbf{1}_{\left\{ Z_{T_n}^{(m)} \geq \Phi_p(X_{T_n}^{(m)}; \widehat{\theta}_n^{p,M}) \right\}} + \widehat{\tau}_{n+1}^{p,(m)} \mathbf{1}_{\left\{ Z_{T_n}^{(m)} < \Phi_p^{(m)}(X_{T_n}^{(m)}; \widehat{\theta}_n^{p,M}) \right\}}, \text{ for } 1 \leq n \leq N-1 \end{cases}$$

where $\widehat{\theta}_n^{p,M}$ solves the following optimization problem

$$\inf_{\theta \in \Theta_p} \frac{1}{M} \sum_{m=1}^{M} \left| \Phi_p(X_{T_n}^{(m)}; \theta) - Z_{\tau_{n+1}^{p,(m)}}^{(m)} \right|^2. \tag{11}$$

Then, we finally approximate the time-0 price of the option by

$$U_0^{p,M} = \max \left(Z_0, \frac{1}{M} \sum_{m=1}^{M} Z_{\widehat{\tau}_1^{p,(m)}}^{(m)} \right). \tag{12}$$

4 Convergence of the algorithm

We start this section on the study of the convergence by introducing some bespoke notation following from Clément et al. [2002].

4.1 Notation

First, it is important to note that the paths $\tau_1^{p,(m)}, \ldots, \tau_N^{p,(m)}$ for $m = 1, \ldots, M$ are identically distributed but not independent since the computations of θ_n^p at each time step n mix all the paths. We define the vector ϑ of the coefficients of the successive expansions $\vartheta^p = (\theta_1^p, \ldots, \theta_{N-1}^p)$ and its Monte Carlo counterpart $\widehat{\vartheta}^{p,M} = (\widehat{\theta}_1^{p,M}, \ldots, \widehat{\theta}_{N-1}^{p,M})$.

Now, we recall the notation used by Clément et al. [2002] to study the convergence of the original Longstaff Schwartz approach.

Given a deterministic parameter $t^p = (t_1^p, \ldots, t_{N-1}^p)$ in ${\Theta_p}^{N-1}$ and deterministic vectors $z = z_1, \ldots, z_N$ in $\mathbb{R}^N$ and $x = (x_1, \ldots, x_N)$ in $(\mathbb{R}^r)^N$, we define the vector field $F = F_1, \ldots, F_N$ by

$$\begin{cases} F_N(t^p, z, x) &= z_N \\ F_n(t^p, z, x) &= z_n \mathbf{1}_{\left\{ z_n \geq \Phi_p(x_n; t_n^p) \right\}} + F_{n+1}(t^p, z, x) \mathbf{1}_{\left\{ z_n < \Phi_p(x_n; t_n^p) \right\}}, \text{ for } 1 \leq n \leq N-1. \end{cases}$$

Note that $F_n(t, z, x)$ does not depend on the first $n-1$ components of t^p, ie $F_n(t^p, z, x)$ depends only $t_n^p, \ldots, t_{N-1}^p$. Moreover,

$$\begin{aligned} F_n(\vartheta^p, Z, X) &= Z_{\tau_n^p}, \\ F_n(\widehat{\vartheta}^{p,M}, Z^{(m)}, X^{(m)}) &= Z_{\widehat{\tau}_n^{p,(m)}}^{(m)}. \end{aligned}$$

Moreover, we clearly have that for all $t^p \in {\Theta_p}^{N-1}$

$$|F_n(t^p, Z, X)| \leq \max_{n \geq n} |Z_{T_n}| \tag{13}$$

4.2 Deep neural network approximations of conditional expectations

Proposition 4.1 *Assume that* $\mathbb{E}[\max_{0 \leq n \leq N} |Z_{T_n}|^2]$. *Then,* $\lim_{p \to \infty} \mathbb{E}[Z_{\tau_n^p}|\mathcal{F}_{T_n}] = \mathbb{E}[Z_{\tau_n}|\mathcal{F}_{T_n}]$ *in* $L^2(\Omega)$ *for all* $1 \leq n \leq N$.

Proof. We proceed by induction. The result is true for $n = N$ as $\tau_N = \tau_N^p = T$. Assume it holds for $n+1$ ($0 \leq n \leq N-1$), we will prove it is true for n.

$$\mathbb{E}[Z_{\tau_n^p} - Z_{\tau_n}|\mathcal{F}_{T_n}]$$

$$= Z_{T_n}\left(\mathbf{1}_{\left\{Z_{T_n} \geq \Phi_p(X_n;\theta_n^p)\right\}} - \mathbf{1}_{\left\{Z_{T_n} \geq \mathbb{E}[Z_{\tau_{n+1}}|\mathcal{F}_{T_n}]\right\}}\right)$$

$$+ \mathbb{E}\left[Z_{\tau_{n+1}^p}\mathbf{1}_{\left\{Z_{T_n} < \Phi_p(X_n;\theta_n^p)\right\}} - Z_{\tau_{n+1}}\mathbf{1}_{\left\{Z_{T_n} < \mathbb{E}[Z_{\tau_{n+1}}|\mathcal{F}_{T_n}]\right\}}\Big|\mathcal{F}_{T_n}\right]$$

$$= (Z_{T_n} - \mathbb{E}[Z_{\tau_{n+1}}|\mathcal{F}_{T_n}])\left(\mathbf{1}_{\left\{Z_{T_n} \geq \Phi_p(X_n;\theta_n^p)\right\}} - \mathbf{1}_{\left\{Z_{T_n} \geq \mathbb{E}[Z_{\tau_{n+1}}|\mathcal{F}_{T_n}]\right\}}\right)$$

$$+ \mathbb{E}\left[Z_{\tau_{n+1}^p} - Z_{\tau_{n+1}}|\mathcal{F}_{T_n}\right]\mathbf{1}_{\left\{Z_{T_n} < \Phi_p(X_n;\theta_n^p)\right\}}.$$

By the induction assumption, the term $\mathbb{E}\left[Z_{\tau_{n+1}^p} - Z_{\tau_{n+1}}|\mathcal{F}_{T_n}\right]$ goes to zero in $L^2(\Omega)$ as p goes to infinity. So, we just have to prove that

$$A_n^p = (Z_{T_n} - \mathbb{E}[Z_{\tau_{n+1}}|\mathcal{F}_{T_n}])\left(\mathbf{1}_{\left\{Z_{T_n} \geq \Phi_p(X_n;\theta_n^p)\right\}} - \mathbf{1}_{\left\{Z_{T_n} \geq \mathbb{E}[Z_{\tau_{n+1}}|\mathcal{F}_{T_n}]\right\}}\right)$$

converges to zero in $L^2(\Omega)$ when $p \to \infty$.

$$|A_n^p| \leq \left|Z_{T_n} - \mathbb{E}[Z_{\tau_{n+1}}|\mathcal{F}_{T_n}]\right|\left|\mathbf{1}_{\left\{Z_{T_n} \geq \Phi_p(X_n;\theta_n^p)\right\}} - \mathbf{1}_{\left\{Z_{T_n} \geq \mathbb{E}[Z_{\tau_{n+1}}|\mathcal{F}_{T_n}]\right\}}\right|$$

$$\leq \left|Z_{T_n} - \mathbb{E}[Z_{\tau_{n+1}}|\mathcal{F}_{T_n}]\right|\left|\mathbf{1}_{\left\{\mathbb{E}[Z_{\tau_{n+1}}|\mathcal{F}_{T_n}] > Z_{T_n} \geq \Phi_p(X_n;\theta_n^p)\right\}} - \mathbf{1}_{\left\{\Phi_p(X_n;\theta_n^p) > Z_{T_n} \geq \mathbb{E}[Z_{\tau_{n+1}}|\mathcal{F}_{T_n}]\right\}}\right|$$

$$\leq \left|Z_{T_n} - \mathbb{E}[Z_{\tau_{n+1}}|\mathcal{F}_{T_n}]\right|\mathbf{1}_{\left\{\left|Z_{T_n} - \mathbb{E}[Z_{\tau_{n+1}}|\mathcal{F}_{T_n}]\right| \leq \left|\Phi_p(X_n;\theta_n^p) - \mathbb{E}[Z_{\tau_{n+1}}|\mathcal{F}_{T_n}]\right|\right\}}$$

$$\leq \left|\Phi_p(X_n;\theta_n^p) - \mathbb{E}[Z_{\tau_{n+1}}|\mathcal{F}_{T_n}]\right|$$

$$\leq \left|\Phi_p(X_n;\theta_n^p) - \mathbb{E}[Z_{\tau_{n+1}^p}|\mathcal{F}_{T_n}]\right| + \left|\mathbb{E}[Z_{\tau_{n+1}^p}|\mathcal{F}_{T_n}] - \mathbb{E}[Z_{\tau_{n+1}}|\mathcal{F}_{T_n}]\right|. \tag{14}$$

As the conditional expectation is an orthogonal projection, we clearly have that

$$\mathbb{E}\left[\left|\mathbb{E}[Z_{\tau_{n+1}^p}|\mathcal{F}_{T_n}] - \mathbb{E}[Z_{\tau_{n+1}}|\mathcal{F}_{T_n}]\right|^2\right] \leq \mathbb{E}\left[\left|\mathbb{E}[Z_{\tau_{n+1}^p}|\mathcal{F}_{T_{n+1}}] - \mathbb{E}[Z_{\tau_{n+1}}|\mathcal{F}_{T_{n+1}}]\right|^2\right].$$

Then, the induction assumption for $n+1$ yields that the second term on the r.h.s of (14) goes to zero in $L^2(\Omega)$ when $p \to \infty$.

To deal with the first term on the r.h.s of (14), we introduce for any $p \in \mathbb{N}$, $\tilde{\theta}_n^p \in \Theta_p$ defined as a solution to

$$\inf_{\theta \in \Theta_p} \mathbb{E}\left[\left|\Phi_p(X_{T_n};\theta) - \mathbb{E}\left[Z_{\tau_{n+1}}|\mathcal{F}_{T_n}\right]\right|^2\right]. \tag{15}$$

As θ_n^p solves (10), we clearly have that

$$\mathbb{E}\left[\left|\Phi_p(X_n; \theta_n^p) - \mathbb{E}[Z_{\tau_{n+1}^p}|\mathcal{F}_{T_n}]\right|^2\right]$$

$$\leq \mathbb{E}\left[\left|\Phi_p(X_n; \tilde{\theta}_n^p) - \mathbb{E}[Z_{\tau_{n+1}^p}|\mathcal{F}_{T_n}]\right|^2\right]$$

$$\leq 2\mathbb{E}\left[\left|\Phi_p(X_n; \tilde{\theta}_n^p) - \mathbb{E}[Z_{\tau_{n+1}}|\mathcal{F}_{T_n}]\right|^2\right] + 2\mathbb{E}\left[\left|\mathbb{E}[Z_{\tau_{n+1}}|\mathcal{F}_{T_n}] - \mathbb{E}[Z_{\tau_{n+1}^p}|\mathcal{F}_{T_n}]\right|^2\right] \quad (16)$$

Using the induction assumption for $n + 1$, the second term on the r.h.s of (16) goes to zero in $L^2(\Omega)$

$$\lim_{p\to\infty} \mathbb{E}\left[\left|\mathbb{E}[Z_{\tau_{n+1}}|\mathcal{F}_{T_n}] - \mathbb{E}[Z_{\tau_{n+1}^p}|\mathcal{F}_{T_n}]\right|^2\right] = 0.$$

From the universal approximation theorem (see Theorem 2.2 and Remark (2.3)), we deduce that

$$\lim_{p\to\infty} \mathbb{E}\left[\left|\Phi_p(X_n; \tilde{\theta}_n^p) - \mathbb{E}[Z_{\tau_{n+1}}|\mathcal{F}_{T_n}]\right|^2\right] = 0.$$

Then, we conclude that $\lim_{p\to\infty} \mathbb{E}[|A_n^p|^2] = 0$. ∎

Remark 4.2 *Note that in the proof of Proposition 4.1, there is no need for the sets Θ_p to be compact for every p. We could have chosen $\gamma_p = \infty$. However, this assumption will be required in the following section, so to work with the same approximations over the whole paper, we have decided to impose compactness on Θ_p for every p.*

4.3 Convergence of the Monte Carlo approximation

In the following, we assume that p is fixed and we study the convergence with respect to the number of samples M. Before studying the convergence of our algorithm, we recall some important results on the convergence of the solution of a sequence of optimization problems whose cost functions converge.

4.3.1 Convergence of optimization problems

Consider a sequence of real valued functions $(f_n)_n$ defined on a compact set $K \subset \mathbb{R}^d$. Define,

$$v_n = \inf_{x\in K} f_n(x)$$

and let x_n be a sequence of minimizers

$$f_n(x_n) = \inf_{x\in K} f_n(x).$$

From [Rubinstein and Shapiro, 1993, Chap. 2], we have the following result.

Lemma 4.3 *Assume that the sequence $(f_n)_n$ converges uniformly on K to a continuous function f. Let $v^\star = \inf_{x \in K} f(x)$ and $\mathcal{S}^\star = \{x \in K : f(x) = v^\star\}$. Then $v_n \to v^\star$ and $d(x_n, \mathcal{S}^\star) \to 0$ a.s.*

In the following, we will also make heavy use of the following result, which is a restatement of the law of large numbers in Banach spaces, see [Ledoux and Talagrand, 1991, Corollary 7.10, page 189] or [Rubinstein and Shapiro, 1993, Lemma A1].

Lemma 4.4 *Let $(\xi_i)_{i \geq 1}$ be a sequence of i.i.d. $\mathbb{R}^m$-valued random vectors and $h : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}$ be a measurable function. Assume that*

- *a.s., $\theta \in \mathbb{R}^d \mapsto h(\theta, \xi_1)$ is continuous,*

- *$\forall C > 0,\ \mathbb{E}\left[\sup_{|\theta| \leq C} |h(\theta, \xi_1)|\right] < +\infty$.*

Then, a.s. $\theta \in \mathbb{R}^d \mapsto \frac{1}{n} \sum_{i=1}^n h(\theta, \xi_i)$ converges locally uniformly to the continuous function $\theta \in \mathbb{R}^d \mapsto \mathbb{E}[h(\theta, \xi_1)]$, ie

$$\lim_{n \to \infty} \sup_{|\theta| \leq C} \left| \frac{1}{n} \sum_{i=1}^n h(\theta, \xi_i) - \mathbb{E}[h(\theta, \xi_1)] \right| = 0 \ a.s.$$

4.3.2 Strong law of large numbers

To prove a strong law of large numbers we will need the following assumptions.

($\mathcal{H}$-1) For every $p \in \mathbb{N},\ p > 1$, there exist $q \geq 1$ and $\kappa_p > 0$ s.t.

$$\forall x \in \mathbb{R}^r,\ \forall \theta \in \Theta_p, \quad |\Phi_p(x, \theta)| \leq \kappa_p (1 + |x|^q).$$

Moreover, for all $1 \leq n \leq N - 1$, a.s. the random functions $\theta \in \Theta_p \longmapsto \Phi_p(X_{T_n}, \theta)$ are continuous. Note that as Θ_p is a compact set, the continuity automatically yields the uniform continuity.

($\mathcal{H}$-2) For q defined in ($\mathcal{H}$-1), $\mathbb{E}[|X_{T_n}|^{2q}] < \infty$ for all $0 \leq n \leq N$.

($\mathcal{H}$-3) For all $p \in \mathbb{N},\ p > 1$ and all $1 \leq n \leq N - 1$, $\mathbb{P}\left(Z_{T_n} = \Phi_p(X_{T_n}; \theta_n^p)\right) = 0$.

We introduce the notation

$$\mathcal{S}_n^p = \arg \inf_{\theta \in \Theta_p} \mathbb{E}\left[\left|\Phi_p(X_{T_n}; \theta) - Z_{\tau_{n+1}^p}\right|^2\right].$$

Note that $\mathcal{S}_n^p$ is a non void compact set.

($\mathcal{H}$-4) For every $p \in \mathbb{N},\ p > 1$ and every $1 \leq n \leq N$, for all $\theta^1, \theta^2 \in \mathcal{S}_n^p$,

$$\Phi_p(X_{T_n}; \theta^1) = \Phi_p(X_{T_n}; \theta^2) \quad \text{a.s.}$$

9

Remark 4.5 *Assumption (H-1) is clearly satisfied for the classical activation functions ReLU $\sigma(x) = (x)_+$, sigmoid $\sigma(x) = (1 + e^{-x})^{-1}$ and $\sigma(x) = \tanh(x)$. When the law of X_{T_n} has a density with respect to the Lebesgue measure, the continuity assumption stated in (H-1) is even satisfied by the binary step activation function $\sigma(x) = \mathbf{1}_{\{x \geq 0\}}$.*

Remark 4.6 *Considering the natural symmetries existing in a neural network, it is clear that the set $\mathcal{S}_n^p$ will hardly ever be reduced to a singleton. So, none of the parameters $\widehat{\theta}_n^{p,M}$ or θ_n^p is unique. Here, we only require the function described by neural network approximation to be unique but not its representation, which is much weaker and more realistic in practice. We refer to Albertini et al. [1993], Albertini and Sontag [1994] for characterization of symetries of neural networks and to Williamson and Helmke [1995] for results on existence and uniqueness of an optimal neural network approximation (but not its parameters).*

To start, we prove the convergence of the neural network approximation.

Proposition 4.7 *Assume that Assumptions (H-1)-(H-4) hold. Then, for every $n = 1, \ldots, N$, $\Phi(X_{T_n}^{(1)}; \widehat{\theta}_n^{p,M})$ converges to $\Phi_p(X_{T_n}^{(1)}; \theta_n^p)$ a.s. as $M \to \infty$.*

Lemma 4.8 *For every $n = 1, \ldots, N - 1$,*

$$|F_n(a, Z, X) - F_n(b, Z, X)| \leq \left(\sum_{i=n}^{N} |Z_{T_i}| \right) \left(\sum_{i=n}^{N-1} \mathbf{1}_{\left\{ |Z_{T_i} - \Phi_p(X_{T_i}; b_i)| \leq |\Phi_p(X_{T_i}; a_i) - \Phi_p(X_{T_i}; b_i)| \right\}} \right)$$

Proof (Proof of Proposition 4.7). We proceed by induction. For $n = N - 1$, $\widehat{\theta}_{N-1}^{p,M}$ solves

$$\inf_{\theta \in \Theta_p} \sum_{m=1}^{M} \left| \Phi_p(X_{T_{N-1}}^{(m)}; \theta) - Z_{T_N}^{(m)} \right|^2.$$

We aim at applying Lemma 4.4 to the sequence of i.i.d. random functions $h_m(\theta) = \left| \Phi_p(X_{T_{N-1}}^{(m)}; \theta) - Z_{T_N}^{(m)} \right|^2$. From Assumptions (H-1) and (H-2), we deduce that

$$\mathbb{E}\left[\sup_{\theta \in \Theta_P} |h_m(\theta)| \right] \leq 2\kappa_p \mathbb{E}[|X_{T_{N-1}}|^{2q}] + \mathbb{E}[(Z_T)^2] < \infty.$$

Then, Lemma 4.4 implies that a.s. the function

$$\theta \in \Theta_p \longmapsto \frac{1}{M} \sum_{m=1}^{M} \left| \Phi_p(X_{T_{N-1}}^{(m)}; \theta) - Z_{T_N}^{(m)} \right|^2$$

converges uniformly to $\mathbb{E}[|\Phi_p(X_{T_{N-1}}; \theta) - Z_{T_N}|^2]$. Hence, we deduce from Lemma 4.3 that $d(\widehat{\theta}_{N-1}^{p,M}, \mathcal{S}_{N-1}^p) \to 0$ a.s. when $M \to \infty$. Hence, there exists a sequence of random variables

$(\xi_{N-1}^{p,M})_M$ taking values in S_{N-1}^p such that $\left|\widehat{\theta}_{N-1}^{p,M} - \xi_{N-1}^{p,M}\right| \to 0$. a.s when $M \to \infty$. Using that a.s. the random functions $\Phi(X_{T_{N-1}}^{(1)}; \cdot)$ are uniformly continuous ($\mathcal{H}$-1), we deduce that

$$\Phi(X_{T_{N-1}}^{(1)}; \widehat{\theta}_{N-1}^{p,M}) - \Phi(X_{T_{N-1}}^{(1)}; \xi_{N-1}^{p,M}) \to 0 \quad a.s.$$

Then, we conclude from Assumption ($\mathcal{H}$-4), that $\Phi(X_{T_{N-1}}^{(1)}; \widehat{\theta}_{N-1}^{p,M}) \to \Phi(X_{T_{N-1}}^{(1)}; \theta_{N-1}^p)$ a.s.

Choose $n \le N - 2$ and assume that the convergence result holds for $n + 1, \ldots, N - 1$, we aim at proving this is true for n. We recall that $\widehat{\theta}_n^{p,M}$ solves

$$\inf_{\theta \in \Theta_p} \sum_{m=1}^M \left|\Phi_p(X_{T_n}^{(m)}; \theta) - F_{n+1}(\widehat{\vartheta}^{p,M}, Z^{(m)}, X^{(m)})\right|^2.$$

We introduce the two random functions for $\theta \in \Theta_p$

$$\hat{v}^M(\theta) = \frac{1}{M} \sum_{m=1}^M \left|\Phi_p(X_{T_n}^{(m)}; \theta) - F_{n+1}(\widehat{\vartheta}^{p,M}, Z^{(m)}, X^{(m)})\right|^2$$

$$v^M(\theta) = \frac{1}{M} \sum_{m=1}^M \left|\Phi_p(X_{T_n}^{(m)}; \theta) - F_{n+1}(\vartheta^p, Z^{(m)}, X^{(m)})\right|^2.$$

The function v^M clearly writes as the sum of i.i.d. random variables. Moreover, by combining (13) and Assumptions ($\mathcal{H}$-1) and ($\mathcal{H}$-2), we obtain

$$\mathbb{E}\left[\sup_{\theta \in \Theta_p} |\Phi_p(X_{T_n}; \theta) - F_{n+1}(\vartheta^p, Z, X)|^2\right] \le 2\kappa \mathbb{E}[1 + |X_{T_n}|^{2q}] + \mathbb{E}\left[\max_{\ell \ge n+1} (Z_{T_\ell})^2\right] < \infty.$$

Then, the sequence of random functions v^M a.s. converges uniformly to the continuous function v defined for $\theta \in \Theta_p$ by

$$v(\theta) = \mathbb{E}\left[|\Phi_p(X_{T_n}; \theta) - F_{n+1}(\vartheta^p, Z, X)|^2\right].$$

It remains to prove that $\sup_{\theta \in \Theta_p} \left|\hat{v}^M(\theta) - v^M(\theta)\right| \to 0$ a.s. when $M \to \infty$.

$$\left|\hat{v}^M(\theta) - v^M(\theta)\right|$$

$$\le \frac{1}{M} \sum_{m=1}^M \left|2\Phi_p(X_{T_n}^{(m)}; \theta) - F_{n+1}(\widehat{\vartheta}^{p,M}, Z^{(m)}, X^{(m)}) - F_{n+1}(\vartheta^p, Z^{(m)}, X^{(m)})\right|$$

$$\left|F_{n+1}(\widehat{\vartheta}^{p,M}, Z^{(m)}, X^{(m)}) - F_{n+1}(\vartheta^p, Z^{(m)}, X^{(m)})\right|$$

$$\le \frac{1}{M} \sum_{m=1}^M 2\left(\kappa(1 + |X_{T_n}^{(m)}|^q) + \max_{\ell \ge n+1} |Z_{T_\ell}|\right)$$

$$\left|F_{n+1}(\widehat{\vartheta}^{p,M}, Z^{(m)}, X^{(m)}) - F_{n+1}(\vartheta^p, Z^{(m)}, X^{(m)})\right|$$

where we have used (13) and Assumptions ($\mathcal{H}$-1) and ($\mathcal{H}$-2). Then from Lemma 4.8, we can write

$$
\left| \hat{v}^M(\theta) - v^M(\theta) \right|
$$

$$
\leq \frac{1}{M} \sum_{m=1}^{M} 2 \left(\kappa(1 + |X_{T_n}^{(m)}|^q) + \max_{\ell \geq k+1} |Z_{T_\ell}| \right)
$$

$$
\left(\sum_{i=n+1}^{N} \left| Z_{T_i}^{(m)} \right| \right) \left(\sum_{i=n+1}^{N-1} \mathbf{1}_{\left\{ \left| Z_{T_i}^{(m)} - \Phi_p(X_{T_i}^{(m)}; \theta_i^p) \right| \leq \left| \Phi_p(X_{T_i}^{(m)}; \widehat{\theta}_i^{p,M}) - \Phi_p(X_{T_i}^{(m)}; \theta_i^p) \right| \right\}} \right)
$$

$$
\leq \frac{1}{M} \sum_{m=1}^{M} C \left(\kappa_p(1 + |X_{T_n}^{(m)}|^{2q}) + \sum_{i=n+1}^{N} \left| Z_{T_i}^{(m)} \right|^2 \right)
$$

$$
\left(\sum_{i=n+1}^{N-1} \mathbf{1}_{\left\{ \left| Z_{T_i}^{(m)} - \Phi_p(X_{T_i}^{(m)}; \theta_i^p) \right| \leq \left| \Phi_p(X_{T_i}^{(m)}; \widehat{\theta}_i^{p,M}) - \Phi_p(X_{T_i}^{(m)}; \theta_i^p) \right| \right\}} \right)
$$

where C is a generic constant only depending on κ_p, n and N.

Let $\varepsilon > 0$. Using the induction assumption and the strong law of large numbers, we have

$$
\limsup_M \sup_{\theta \in \Theta_p} \left| \hat{v}^M(\theta) - v^M(\theta) \right|
$$

$$
\leq \limsup_M \frac{1}{M} \sum_{m=1}^{M} C \left((1 + |X_{T_n}^{(m)}|^{2q}) + \sum_{i=n+1}^{N} \left| Z_{T_i}^{(m)} \right|^2 \right)
$$

$$
\left(\sum_{i=n+1}^{N-1} \mathbf{1}_{\left\{ \left| Z_{T_i}^{(m)} - \Phi_p(X_{T_i}^{(m)}; \theta_i^p) \right| \leq \varepsilon \right\}} \right)
$$

$$
\leq C\mathbb{E} \left[\left((1 + |X_{T_n}|^{2q}) + \sum_{i=n+1}^{N} |Z_{T_i}|^2 \right) \left(\sum_{i=n+1}^{N-1} \mathbf{1}_{\left\{ \left| Z_{T_i} - \Phi_p(X_{T_i}; \theta_i^p) \right| \leq \varepsilon \right\}} \right) \right]
$$

From ($\mathcal{H}$-3), we deduce that $\lim_{\varepsilon \to 0} \mathbf{1}_{\left\{ \left| Z_{T_i} - \Phi_p(X_{T_i}; \theta_i^p) \right| \leq \varepsilon \right\}} = 0$ a.s. and we conclude that a.s. $\hat{v}^M - v^M$ converges to zero uniformly. As we have already proved that a.s. v^M converges uniformly to the continuous function v, we deduce that a.s. $\hat{v}^M$ converges uniformly to v. From Lemma 4.3, we conclude that $d(\widehat{\theta}_n^{p,M}, S_n^p) \to 0$ a.s. when $M \to \infty$. Hence, there exists a sequence of random variables $(\xi_n^{p,M})_M$ taking values in S_n^p such that $\left| \widehat{\theta}_n^{p,M} - \xi_n^{p,M} \right| \to 0$. a.s when $M \to \infty$. Using that a.s. the random functions $\Phi(X_{T_{N-1}}^{(1)}; \cdot)$ are uniformly continuous ($\mathcal{H}$-1)

$$
\Phi(X_{T_{N-1}}^{(1)}; \widehat{\theta}_n^{p,M}) - \Phi(X_{T_n}^{(1)}; \xi_n^{p,M}) \to 0 \quad a.s. \text{ when } M \to \infty.
$$

Then, we conclude from Assumption ($\mathcal{H}$-4), that $\Phi(X_{T_n}^{(1)}; \widehat{\theta}_n^{p,M}) \to \Phi(X_{T_n}^{(1)}; \theta_n^p)$ a.s. when $M \to \infty$. ∎

Now that the convergence of the expansion is established, we can study the convergence of $U_0^{p,M}$ to U_0^p when $M \to \infty$.

Theorem 4.9 *Assume that Assumptions ($\mathcal{H}$-1)-($\mathcal{H}$-4) hold. Then, for $\alpha = 1, 2$ and every $n = 1, \ldots, N$,*

$$\lim_{M \to \infty} \frac{1}{M} \sum_{m=1}^{M} \left(Z_{\widehat{\tau}_n^{p,(m)}}^{(m)} \right)^\alpha = \mathbb{E}\left[(Z_{\tau_n^p})^\alpha \right] \quad a.s.$$

Proof. Note that $\mathbb{E}[(Z_{\tau_n^p})^\alpha] = \mathbb{E}[F_n(\vartheta^p, Z, G)^\alpha]$ and by the strong law of large numbers

$$\lim_{M \to \infty} \frac{1}{M} \sum_{m=1}^{M} F_n(\vartheta^p, Z^{(m)}, X^{(m)})^\alpha = \mathbb{E}[F_n(\vartheta^p, Z, X)^\alpha] \quad a.s.$$

Hence, we have to prove that

$$\Delta F_M = \frac{1}{M} \sum_{m=1}^{M} \left(F_n(\widehat{\vartheta}^{p,M}, Z^{(m)}, X^{(m)})^\alpha - F_n(\vartheta^p, Z^{(m)}, X^{(m)})^\alpha \right) \xrightarrow[M \to \infty]{a.s} 0.$$

For any $x, y \in \mathbb{R}$, and $\alpha = 1, 2$, $|x^\alpha - y^\alpha| = |x - y| |x^{q-1} + y^{q-1}|$. Using Lemma 4.8 and that $|F_n(\gamma, z, g)| \le \max_{n \le j \le N} |z_j|$, we have

$$|\Delta F_M| \le \frac{1}{M} \sum_{m=1}^{M} \left| F_n(\widehat{\vartheta}^{p,M}, Z^{(m)}, X^{(m)})^\alpha - F_n(\vartheta^p, Z^{(m)}, G^{(m)})^\alpha \right|$$

$$\le 2 \frac{1}{M} \sum_{m=1}^{M} \sum_{i=n}^{N} \max_{n \le j \le N} \left| Z_{T_j}^{(m)} \right| \left| Z_{T_{i+1}}^{(m)} \right| \left(\sum_{i=n}^{N-1} \mathbf{1}_{\left\{ \left| Z_{T_i}^{(m)} - \Phi_p(X_{T_i}^{(m)};\theta_i^p) \right| \le \left| \Phi_p(X_{T_i}^{(m)};\widehat{\theta}_i^{,M}) - \Phi_p(X_{T_i}^{(m)};\theta_i^p) \right| \right\}} \right)$$

Using Proposition 4.7, for all $i = n, \ldots, N - 1$, $\left| \Phi_p(X_{T_i}^{(1)};\widehat{\theta}_i^{p,M}) - \Phi_p(X_{T_i}^{(1)};\theta_i^p) \right| \to 0$ when $M \to \infty$. Then for any $\varepsilon > 0$,

$$\limsup_M |\Delta F_M|$$

$$\le 2 \limsup_M \frac{1}{M} \sum_{m=1}^{M} \sum_{i=n}^{N} \max_{n \le j \le N} \left| Z_{T_j}^{(m)} \right| \left| Z_{T_{i+1}}^{(m)} \right| \left(\sum_{i=k}^{N-1} \mathbf{1}_{\left\{ \left| Z_{T_i}^{(m)} - \Phi_p(X_{T_i}^{(m)};\theta_i^p) \right| \le \varepsilon \right\}} \right)$$

$$\le 2\mathbb{E} \left[\sum_{i=n}^{N} \max_{n \le j \le N} \left| Z_{T_j} \right| \left| Z_{T_{i+1}} \right| \left(\sum_{i=n}^{N-1} \mathbf{1}_{\left\{ \left| Z_{T_i} - \Phi_p(X_{T_i}^{(m)};\theta_i^p) \right| \le \varepsilon \right\}} \right) \right]$$

where the last inequality follows from the strong law of larger numbers as $\mathbb{E}[\max_{n \le j \le N} \left| Z_{T_j} \right|^2] < \infty$. We conclude that $\limsup_M |\Delta F_M| = 0$ by letting ε go to 0 and by using ($\mathcal{H}$-3). ∎

The case $\alpha = 1$ proves the strong law of large numbers for the algorithm. Considering that all the paths are actually mixed through the neural network approximation, it is unlikely that the estimators $\frac{1}{M} \sum_{m=1}^{M} Z_{\widehat{\tau}_n^{p,(m)}}^{(m)}$ for $1 \leq n \leq N$ are unbiased. We recall that $U_n^{p,M} = \frac{1}{M} \sum_{m=1}^{M} F_n(\widehat{\vartheta}^{p,M}, Z^{(m)}, G^{(m)})$ and $Z_{\tau_n^p} = F_n(\vartheta^p, Z, X)$. Then,

$$
\mathbb{E}\left[U_n^{p,M}\right] - \mathbb{E}\left[Z_{\tau_n^p}\right] = \mathbb{E}\left[\frac{1}{M} \sum_{m=1}^{M} \left(F_n(\widehat{\vartheta}^{p,M}, Z^{(m)}, X^{(m)}) - F_n(\vartheta^p, Z^{(m)}, X^{(m)})\right)\right]
$$

$$
= \mathbb{E}\left[F_n(\widehat{\vartheta}^{p,M}, Z^{(1)}, X^{(1)}) - F_n(\vartheta^p, Z^{(1)}, X^{(1)})\right]
$$

where we have used that all the random variables have the same distribution.

References

F. Albertini and E. D. Sontag. Uniqueness of weights for recurrent nets. *MATHEMATICAL RESEARCH*, 79:599–599, 1994.

F. Albertini, E. D. Sontag, and V. Maillot. Uniqueness of weights for neural networks. *Artificial Neural Networks for Speech and Vision*, pages 115–125, 1993.

V. I. Arnold. On functions of three variables. *Collected Works: Representations of Functions, Celestial Mechanics and KAM Theory, 1957–1965*, pages 5–8, 2009.

V. Bally and G. Pages. A quantization algorithm for solving multidimensional discrete-time optimal stopping problems. *Bernoulli*, 9(6):1003–1049, 2003.

S. Becker, P. Cheridito, and A. Jentzen. Deep optimal stopping. *arXiv preprint arXiv:1804.05394*, 2018.

B. Bouchard and X. Warin. Monte-carlo valuation of american options: Facts and new algorithms to improve existing methods. In R. A. Carmona, P. Del Moral, P. Hu, and N. Oudjane, editors, *Numerical Methods in Finance*, volume 12 of *Springer Proceedings in Mathematics*, pages 215–255. Springer Berlin Heidelberg, 2012.

M. Broadie and P. Glasserman. A stochastic mesh method for pricing high-dimensional american options. *Journal of Computational Finance*, 7:35–72, 2004.

A. L. Bronstein, G. Pagès, and J. Portès. Multi-asset american options and parallel quantization. *Methodology and Computing in Applied Probability*, 15(3):547–561, 2013.

J. F. Carriere. Valuation of the early-exercise price for options using simulations and nonparametric regression. *Insurance: mathematics and Economics*, 19(1):19–30, 1996.

E. Clément, D. Lamberton, and P. Protter. An analysis of a least squares regression method for american option pricing. *Finance and Stochastics*, 6(4):449–471, 2002.

G. Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2(4):303–314, Dec 1989.

E. Gobet, J. Lemor, and X. Warin. A regression-based Monte Carlo method to solve backward stochastic differential equations. *Annals of Applied Probability*, 15(3):2172–2202, 2005.

K. Hornik. Approximation capabilities of multilayer feedforward networks. *Neural Networks*, 4 (2):251 – 257, 1991.

M. Kohler, A. Krzyżak, and N. Todorovic. Pricing of high-dimensional american options by neural networks. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 20(3):383–410, 2010.

A. Kolmogorov. On the representation of continuous functions of several variables as superpositions of functions of smaller number of variables. In *Soviet. Math. Dokl*, volume 108, pages 179–182, 1956.

M. Ledoux and M. Talagrand. *Probability in Banach spaces*, volume 23 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1991. ISBN 3-540-52013-9. Isoperimetry and processes.

F. Longstaff and R. Schwartz. Valuing American options by simulation : A simple least-square approach. *Review of Financial Studies*, 14:113–147, 2001.

G. Pagès. *Numerical Probability: An Introduction with Applications to Finance*. Springer, 2018. doi: 10.1007/978-3-319-90276-0.

R. Y. Rubinstein and A. Shapiro. *Discrete event systems*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Ltd., Chichester, 1993. ISBN 0-471-93419-4. Sensitivity analysis and stochastic optimization by the score function method.

J. A. Tilley. Valuing american options in a path simulation model. *Transactions of the Society of Actuaries*, 45(83):104, 1993.

J. Tsitsiklis and B. V. Roy. Regression methods for pricing complex American-style options. *IEEE Trans. Neural Netw.*, 12(4):694–703, 2001.

R. C. Williamson and U. Helmke. Existence and uniqueness results for neural network approximations. *IEEE Transactions on Neural Networks*, 6(1):2–13, 1995.