Deep Learning for the Multiple Optimal Stopping Problem

Mathieu Laurière* Mehdi Talbi†

November 30, 2025

Abstract

The aim of this paper is to present a numerical study of multiple optimal stopping problems, which arise in many applications (e.g., the optimal liquidation of a portfolio). Due to the possibly high dimensional nature of these problems, our approach relies on a combination of a neural network approximation of the value function and the dynamic programming principle. We analyze the convergence of the algorithm and present some numerical applications.

1 Introduction

Optimal stopping problems consist in choosing the best time to stop a stochastic process, in order to optimize a certain payoff that can be written as a function of this process at the chosen time. The decision must be non-anticipative, which implies that the time must be picked in the set of stopping times corresponding to the filtration generated by the reward process.

It is well known that the (first) optimal stopping time can be determined as the first time the reward processes reaches its so-called Snell envelope, defined as the smallest supermartingale larger than the reward process (see e.g. Shiryaev [16] or Karatzas and Shreve [10] for a general overview of the theory of optimal stopping). This can be seen as a consequence of the dynamic programming principle, which provides an algorithm approach to solve numerically optimal stopping problems. We may mention the celebrated Longstaff-Schwarz method [13], which consists in a recursive sequence of least square regressions to compute the price of American options, and also the stochastic mesh approach by Broadie and Glasserman [2], also in the context in American options. More recently, Becker, Cheridito, and Jentzen [1] proposed an approach combining dynamic programming and an approximation of the optimal stopping rule by a neural network. We also refer the interested reader to [14; 15; 20; 4; 6] for more numerical methods to solve stopping problems.

In the present paper, we are interested in numerically solving the multiple optimal stopping problem, in which one has the possibility to stop multiple stochastic processes by assigning each of them a stopping time that might be different from the others. Such problems have been introduced in a general continuous-time framework by Kobylanski, Quenez, and Rouy-Mironescu [11], who proved it can be reduced to a recursive sequence of single-agent optimal stopping problems with random horizon. An asymptotic version of this problem has been studied in [17; 18; 19]. We also refer to Carmona and Touzi [3] for the study of a special multipe stopping problem in the context of swing options, and to Grigorova, Quenez, and Yuan [5] for an extension of [11] to a class of nonlinear expectations.

*Shanghai Center for Data Science; NYU-ECNU Institute of Mathematical Sciences at NYU Shanghai; NYU Shanghai, Shanghai, People's Republic of China, mathieu.lauriere@nyu.edu

†Laboratoire de Probabilités, Statistiques et Modélisation, Université Paris-Cité, France, mtalbi@lpsm.paris

The literature is very limited when it comes to the numerical analysis of such problems. We may only refer to Han and Li [8], who extend the approach of [1] to multiple stopping problems arising in option pricing problems. In our contribution, we define the multiple stopping problem in a general discrete framework. After deriving a dynamic programming principle, we establish a backwards algorithm enabling the computation of the value function, which merely consists in maximizing at every time step a family of conditional expectations. It is well known that conditional expectations can be efficiently approximated with neural networks (see e.g. Györfi, Kohler, Krzyżak, and Walk [7, Chapter 11]). Here, our key idea consists in parametrizing all the family of conditional expectations by a single neural network, thanks to the addition of an extra-variable encapsulating the state of all the coordinates of the reward process, i.e. wether they are stopped or not. The approximated value function is then defined as the maximum of the neural network taken on this extra-argument, and the approximated optimal stopping policy as the argument reaching this maximum. Having in mind high dimensional multiple stopping (although the optimized algorithm for the study of the mean field problem is left for further research), we also propose an alternative algorithm with a reduced computation cost. We prove the convergence of both algorithms (in the spirit of Huré, Pham, Bachouch, and Langrené [9].

The paper is organized as follows. In Section 2, we define the problem and derive the dynamic programming principle. Section 3 contains the convergence results along with their proofs. In Section 4, we analyze the convergence error when the discrete-time process corresponds to the Euler scheme of a diffusion process, and illustrate this study with some numerical examples. Finally, the Appendix A contains numerical results.

2 The general discrete time setting

We consider a discrete time interacting particle system where each particle's dynamics are influenced by the other particles' states, until it stops. Let $p \in \mathbb{N}^*$ and $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space endowed with a filtration $\mathbb{F} := \{\mathcal{F}_n\}_{\{n\in[p]\}}$, with $\mathcal{F}_p = \mathcal{F}$. We denote by $\mathcal{T}_p$ the set of $[p]$-valued $\mathbb{F}$-stopping times, and by $\mathcal{T}_p^N$ the set of N-tuples of elements of $\mathcal{T}_p$.

Let $\{\varepsilon_n\}_{\{n\in[p]\}}$ be a sequence of i.i.d. random variables such that ε_{n+1} is independent of $\mathcal{F}_n$ for all $n \in [p-1]$. Given $\boldsymbol{\tau} := (\tau^1, \ldots, \tau^N) \in \mathcal{T}_p^N$, we denote by $\boldsymbol{I} := (I^1, \ldots, I^N)$ the corresponding vector of survival processes, defined by $I_n^k := \mathbf{1}_{n\le\tau^k}$ for all $n \in [p]$ and $k \in [N]^*$. We then consider the dynamics:

$$\boldsymbol{X}_{n+1} := \boldsymbol{X}_n + F_n(\boldsymbol{X}_n, \varepsilon_{n+1})\boldsymbol{I}_{n+1}, \;\; \boldsymbol{X}_0 \in \mathbb{R}^N, \tag{2.1}$$

where $F : \mathbb{R}^N \times \mathbb{R}^d \longrightarrow \mathbb{R}^N$, and the above product is understood as a Hadamard product.

2.1 The original problem

Let $g : \mathbb{R}^N \to \mathbb{R}$ be the reward function, and $c : [p] \times \mathbb{R}^N \times \{0,1\}^N \times \{0,1\}^N \to \mathbb{R}_+$ be a transaction cost function which is 0 is not particle is stopped, i.e.,

$$c_n(\boldsymbol{x}, \boldsymbol{i}, \boldsymbol{i}) = 0 \quad \text{for all } (n, \boldsymbol{x}, \boldsymbol{i}) \in [p] \times \mathbb{R}^N \times \{0,1\}^N.$$

The multiple optimal stopping problem consists in solving the following maximization problem:

$$V_0 := \sup_{\boldsymbol{\tau}\in\mathcal{T}_N^p} \mathbb{E}\Big[\sum_{n=0}^{N-1} c_n(\boldsymbol{X}_n, \boldsymbol{I}_n, \boldsymbol{I}_{n+1}) + g(X^1_{\tau^1}, \ldots, X^N_{\tau^N})\Big] = \sup_{\boldsymbol{\tau}\in\mathcal{T}_N^p} \mathbb{E}\Big[\sum_{n=0}^{N-1} c_n(\boldsymbol{X}_n, \boldsymbol{I}_n, \boldsymbol{I}_{n+1}) + g(\boldsymbol{X}_p)\Big]. \tag{2.2}$$

where the second equality comes from the fact that $\boldsymbol{X}$ is the vector of stopped processes, i.e., $X^k_{\tau^k} = X^k_p$ for all $k \in [N]^*$. This class of problems may be analyzed using the dynamic programming approach. For that, introduce the dynamic value function:

$$V_n(\boldsymbol{x}, \boldsymbol{i}) := \sup_{\boldsymbol{\tau} \in \mathcal{T}^N_{n,p}} \mathbb{E}\Big[\sum_{k=n}^{p-1} c_k(\boldsymbol{X}_k, \boldsymbol{I}_k, \boldsymbol{I}_{k+1}) + g(\boldsymbol{X}_p) \big| \boldsymbol{X}_n = \boldsymbol{x}, \boldsymbol{I}_n = \boldsymbol{i}\Big], \tag{2.3}$$

where $\mathcal{T}^N_{n,p}$ denotes the set of $\{n, \dots, p\}$-valued N-tuples of $\mathbb{F}$-stopping times. Based on the works of Kobylanski, Quenez, and Rouy-Mironescu [11] and Talbi, Touzi, and Zhang [19], we have:

Proposition 2.1. *Assume the functions c and g are continuous. Then:*
(i) *For all $(\boldsymbol{x}, \boldsymbol{i}) \in \mathbb{R}^N \times \{0,1\}^N$, we have:*

$$\begin{aligned} V_n(\boldsymbol{x}, \boldsymbol{i}) &= \max_{\{\boldsymbol{i}' \in \{0,1\}^N \ s.t. \ \boldsymbol{i}' \leq \boldsymbol{i}\}} \mathbb{E}[c_n(\boldsymbol{X}_n, \boldsymbol{I}_n, \boldsymbol{I}_{n+1}) + V_{n+1}(\boldsymbol{X}_{n+1}, \boldsymbol{I}_{n+1}) | \boldsymbol{X}_n = \boldsymbol{x}, \boldsymbol{I}_n = \boldsymbol{i}] \\ &= \max_{\{\boldsymbol{i}' \in \{0,1\}^N \ s.t. \ \boldsymbol{i}' \leq \boldsymbol{i}\}} \mathbb{E}[c_n(\boldsymbol{x}, \boldsymbol{i}, \boldsymbol{i}') + V_{n+1}(\boldsymbol{x} + F_n(\boldsymbol{x}, \varepsilon_{n+1})\boldsymbol{i}', \boldsymbol{i}')], \end{aligned}$$

where the inequality $\boldsymbol{i}' \leq \boldsymbol{i}$ is understood coordinate wise.
(ii) *There exists an optimal stopping strategy $\boldsymbol{\tau}^* \in \mathcal{T}^N_{n,p}$ for (2.2).*

Proof. To prove (i), let us introduce the set $\mathbb{I}^N$ of $\mathbb{F}$-predictable, decreasing, $\{0,1\}^N$-valued processes such that $\boldsymbol{I}_0 = \mathbf{1}$. It is clear that $\mathcal{T}^N_p$ and $\mathbb{I}^N$ are in bijection. Indeed, to each $\boldsymbol{\tau} = (\tau_1, \dots, \tau_N)$, we can associate $\boldsymbol{I} = (I^1, \dots, I^N)$ by setting $I^k_n := \mathbf{1}_{n \leq \tau_k} = 1 - \mathbf{1}_{n-1 \geq \tau_k}$, for $k \in [N]$. Since τ_k is a stopping time, it is clear that I^k, and therefore $\boldsymbol{I}$ is predictable. Conversely, given $\boldsymbol{I} \in \mathbb{I}^N$, we define the corresponding stopping times by setting $\tau_k := \min\{n \geq 0 : I^k_{n+1} = 0\} \wedge p$.
Now, let $\mathcal{A}^N$ be the set of $\mathbb{F}$-adapted processes taking values in $\{0,1\}^N$. Then $\mathbb{I}^N$ is in bijection with $\mathcal{A}^N$. Indeed, given $\boldsymbol{I}_n = (\mathbf{1}_{n \leq \tau_1}, \dots, \mathbf{1}_{n \leq \tau_n})$, we have, by setting $\alpha^k_n := \{\tau_k = n\}$, $k \in [p]$, $I^k_n = \prod_{j=1}^{n-1} \alpha^k_j$. Therefore, the multiple stopping problem (2.3) may be rewritten as a standard control problem on $\mathcal{A}$. Since g is continuous, the dynamic programming principle for this problem writes:

$$V_n(\boldsymbol{x}, \boldsymbol{i}) = \sup_{\boldsymbol{a} \in \{0,1\}^N} \mathbb{E}\big[c_n(\boldsymbol{x}, \boldsymbol{i}, \boldsymbol{I}_{n+1}) + V_{n+1}(\boldsymbol{X}_{n+1}, \boldsymbol{I}_{n+1})\big],$$

where $\boldsymbol{X}_{n+1} = \boldsymbol{x} + F_n(\boldsymbol{x}, \varepsilon_{n+1})\boldsymbol{I}_{n+1}$ and $\boldsymbol{I}_{n+1} = \boldsymbol{i}\boldsymbol{a}$. The desired result is obtained by observing that:

$$\{\boldsymbol{i}\boldsymbol{a} : \boldsymbol{a} \in \{0,1\}^N\} = \{\boldsymbol{i}' \in \{0,1\}^N : \boldsymbol{i}' \leq \boldsymbol{i}\},$$

and that this set is finite, and therefore the supremum is a maximum.
To prove (ii), we construct recursively the process $\boldsymbol{I}^*$ defined by

$$\boldsymbol{I}^*_n = \boldsymbol{i} \text{ and } \boldsymbol{I}^*_{m+1} \in \operatorname{argmax}_{\boldsymbol{i}' \leq \boldsymbol{I}^*_m} \mathbb{E}\big[c_n(\boldsymbol{x}, \boldsymbol{i}, \boldsymbol{i}') + V_{m+1}(\boldsymbol{X}_{m+1}, \boldsymbol{i}')\big].$$

$\boldsymbol{\tau}^*$ is then the N-tuple of stopping times associated to the N-tuple of survival processes $\boldsymbol{I}^*$. ∎

Intuitively, Proposition 2.1 means that the multiple optimal stopping problem can be reduced to a recursive sequence of standard stopping problems: at each time n, one decides which agents will be stopped. This decision is encapsulated in the vector $\boldsymbol{i}'$: if $i'_k = 0$, then the k-th agent is stopped; otherwise, they continue. Note that, at every time n, given a state vector $\boldsymbol{i} \in \{0,1\}^N$, one has to examine all the combinations of $\boldsymbol{i}' \leq \boldsymbol{i}$ to decide which particles should be stopped. As it will be seen in Section 3, this could induce a computational cost with exponential growth in N. We therefore examine an alternative multiple stopping problem which considerably reduces this cost.

2.2 The alternative problem

Introduce the following set of N-tuples of stopping times, in which two elements of the tuple cannot be equal unless they are equal to the terminal time:

$$\tilde{\mathcal{T}}_p^N := \Big\{\boldsymbol{\tau} = (\tau_1, \dots, \tau_N) \in \mathcal{T}_p^N : \tau_k = \tau_l \Rightarrow \tau_k = p \quad \text{for all } k \neq l\Big\}.$$

We then define the alternative multiple optimal stopping problem:

$$\tilde{V}_0 := \sup_{\boldsymbol{\tau} \in \tilde{\mathcal{T}}_N^p} \mathbb{E}\Big[\sum_{n=0}^{N-1} c_n(\boldsymbol{X}_n, \boldsymbol{I}_n, \boldsymbol{I}_{n+1}) + g(X^1_{\tau^1}, \dots, X^N_{\tau^N})\Big] = \sup_{\boldsymbol{\tau} \in \tilde{\mathcal{T}}_N^p} \mathbb{E}\Big[\sum_{n=0}^{N-1} c_n(\boldsymbol{X}_n, \boldsymbol{I}_n, \boldsymbol{I}_{n+1}) + g(\boldsymbol{X}_p)\Big] \tag{2.4}$$

In this problem, at each time $n \in [p-1]$, one can only stop (at most) one particle. Similarly to the original problem, we introduce a dynamical version of the value function in (2.4):

$$\tilde{V}_n(\boldsymbol{x}, \boldsymbol{i}) := \sup_{\tilde{\boldsymbol{\tau}} \in \tilde{\mathcal{T}}_{n,p}^N} \mathbb{E}\Big[\sum_{k=n}^{N-1} c_k(\boldsymbol{X}_k, \boldsymbol{I}_k, \boldsymbol{I}_{k+1}) + g(\boldsymbol{X}_p) | \boldsymbol{X}_n = \boldsymbol{x}, \boldsymbol{I}_n = \boldsymbol{i}\Big], \tag{2.5}$$

where $\tilde{\mathcal{T}}_{n,p}^N$ denotes the set of $\{n, \dots, p\}$-valued N-tuples of $\tilde{\mathcal{T}}_p^N$. We then have the following dynamic programming principle:

Proposition 2.2. *Assume g is continuous. Then:*
(i) *For every $(\boldsymbol{x}, \boldsymbol{i}) \in \mathbb{R}^N \times \{0,1\}^N$, we have:*

$$\tilde{V}_n(\boldsymbol{x}, \boldsymbol{i}) = \sup_{\ell \in [N]} \mathbb{E}\big[c_n(\boldsymbol{x}, \boldsymbol{i}, \boldsymbol{i}^{-\ell}) + \tilde{V}_{n+1}(\boldsymbol{x} + F_n(\boldsymbol{x}, \varepsilon_{n+1})\boldsymbol{i}^{-\ell}, \boldsymbol{i}^{-\ell})\big].$$

(ii) *There exists an optimal stopping strategy $\tilde{\boldsymbol{\tau}}^* \in \tilde{\mathcal{T}}_{n,p}^N$ for the problem (2.5).*

Proof. The proof follows the same path as the proof of Proposition 2.1, replacing $\mathcal{A}^N$ with the set $\tilde{\mathcal{A}}^N$ of $\mathbb{F}$-adapted processes taking their values in $\{\boldsymbol{1} - \boldsymbol{e}_\ell : \ell \in [N]\}$. ∎

According to Proposition 2.2, at every step n, we have to choose which particle to stop (if any). Then, it boils down to choose an index $\ell \in [N]$, with $\ell = 0$ standing for the case where we do not stop any particle. Compared with the original problem, we then trade a possible exponential cost in N with a linear cost of N. However, this cost reduction comes up with an additional error due to the fact that in the new problem, we cannot stop several particles at once. In Section 4, we analyze this error in the context of discretized diffusion processes.

3 Main results

3.1 The original algorithm

The first algorithm is based on Proposition 2.1, and approximates directly the original problem 2.2. The idea is the following: assuming the function V_{n+1} has been appropriately approximated at time $n+1$, we compute the function V_n in two steps:

1. First, we approximate the function $U_n : (\boldsymbol{x}, \boldsymbol{i}) \mapsto \mathbb{E}[V_n(\boldsymbol{x} + F_n(\boldsymbol{x}, \varepsilon_{n+1})\boldsymbol{i}, \boldsymbol{i})]$. For this, we approximate U_n with a dense neural network, and we use the classical least squares characterization of the conditional expectation joint with a Monte Carlo approach, for which we need M simulations $\{(X_n^{(m)}, I_n^{(m)}, \varepsilon_{n+1}^{(m)})\}_{\{1 \leq m \leq M\}}$ according to a distribution $\nu = \mu_X \otimes \mu_I \otimes \mu_\varepsilon$, that we assume to be independent from the time n for simplicity.

2. Then, given $\boldsymbol{i} \in \{0,1\}^N$, the function $V_n(\cdot, \boldsymbol{i})$ is defined as:

$$V_n(\cdot, \boldsymbol{i}) = \max_{\boldsymbol{i}' \leq \boldsymbol{i}} c_n(\boldsymbol{x}, \boldsymbol{i}, \boldsymbol{i}') + U_n(\boldsymbol{x}, \boldsymbol{i}').$$

In what follow, we denote by $\xi_M := \{(X_n^{(m)}, I_n^{(m)}, \varepsilon_{n+1}^{(m)}) : 1 \leq m \leq M, 0 \leq n \leq p-1\}$ the total set of simulations, and by $U_n^{\xi_M}$ and $V_n^{\xi_M}$ the neural network approximations of the functions U_n and V_n.

Algorithm 1.

1. Initialization: $\hat{V}_p^{\xi_M} = g$.

2. For $n \in [p-1]$:

   (a) Approximate the conditional expectation function $(\boldsymbol{x}, \boldsymbol{i}) \mapsto \mathbb{E}[V_{n+1}(\boldsymbol{X}_{n+1}, \boldsymbol{i}) | \boldsymbol{X}_n = \boldsymbol{x}] = \mathbb{E}[V_{n+1}(\boldsymbol{x} + F_n(\boldsymbol{x}, \varepsilon_{n+1})\boldsymbol{i}, \boldsymbol{i})]$:

   $$\hat{U}_n^{\xi_M} \in \underset{\phi \in \mathcal{V}}{\operatorname{argmin}} \frac{1}{M} \sum_{m=1}^{M} \left|\phi(\boldsymbol{X}_n^{(m)}, \boldsymbol{I}_n^{(m)}) - \hat{V}_{n+1}^{\xi_M}(\boldsymbol{X}_n + F_n(\boldsymbol{X}_n^{(m)}, \varepsilon_{n+1}^{(m)})\boldsymbol{I}_n^{(m)}, \boldsymbol{I}_n^{(m)})\right|^2.$$

   (b) Compute $\hat{V}_n$ as the increasing envelope of $\hat{U}_n$ with respect to $\boldsymbol{i}$, as well as an optimal strategy at time n:

   $$\hat{V}_n^{\xi_M}(\boldsymbol{x}, \boldsymbol{i}) := \max_{\boldsymbol{i}' \in \{0,1\}^N : \boldsymbol{i}' \leq \boldsymbol{i}} c_n(\boldsymbol{x}, \boldsymbol{i}, \boldsymbol{i}') + \hat{U}_n^{\xi_M}(\boldsymbol{x}, \boldsymbol{i}'). \tag{3.1}$$

   $$\boldsymbol{I}_{n+1}^{\xi_M}(\boldsymbol{x}, \boldsymbol{i}) \in \underset{\boldsymbol{i}' \in \{0,1\}^N : \boldsymbol{i}' \leq \boldsymbol{i}}{\operatorname{argmax}} c_n(\boldsymbol{x}, \boldsymbol{i}, \boldsymbol{i}') + \hat{U}_n^{\xi_M}(\boldsymbol{x}, \boldsymbol{i}'). \tag{3.2}$$

   (c) Given an initial condition $(\boldsymbol{X}_0, \boldsymbol{I}_0) := (\boldsymbol{x}, \boldsymbol{i})$, compute the optimal stopping policy $\boldsymbol{I}^{\xi_M}$ with:

   $$\boldsymbol{I}_{n+1}^{\xi_M} := \boldsymbol{I}_{n+1}^{\xi_M}(\boldsymbol{X}_n, \boldsymbol{I}_n^{\xi_M}), \quad \boldsymbol{X}_{n+1} = \boldsymbol{X}_n + F_n(\boldsymbol{X}_n, \varepsilon_{n+1})\boldsymbol{I}_{n+1}^{\xi_M}.$$

An important drawback of **Algorithm 1** is that (3.1) can be very costly, as we have:

$$\operatorname{Card}(\{\boldsymbol{i}' \in \{0,1\}^N : \boldsymbol{i}' \leq \boldsymbol{i}\}) = 2^{|\boldsymbol{i}|_1}.$$

Computing the maximum on this set therefore implies a complexity of order $\mathcal{O}(2^N)$. This motivates us to introduce an alternative algorithm.

3.2 The alternative algorithm

The second algorithm is based on Proposition 2.2, and approximates directly the alternative problem (2.4), and indirectly the original problem (2.2), see Proposition **??**. The idea is the following:

1. First, we approximate the function $U_n : (\boldsymbol{x}, \boldsymbol{i}) \mapsto \mathbb{E}[V_n(\boldsymbol{x} + F_n(\boldsymbol{x}, \varepsilon_{n+1})\boldsymbol{i}, \boldsymbol{i})]$ similarly to the previous algorithm.

2. Then, given $\boldsymbol{i} \in \{0,1\}^N$, the function $V_n(\cdot, \boldsymbol{i})$ is defined as:

$$V_n(\cdot, \boldsymbol{i}) = \max_{\ell \in [N]} c_n(\boldsymbol{x}, \boldsymbol{i}, \boldsymbol{i}(1 - \boldsymbol{e}_\ell)) + U_n(\boldsymbol{x}, \boldsymbol{i}(1 - \boldsymbol{e}_\ell)).$$

Algorithm 2.

1. Initialization: $\hat{V}_p^{\xi_M} = g$.
2. For $n \in [p-1]$:
   (a) Approximate the conditional expectation function $(\boldsymbol{x}, \boldsymbol{i}) \mapsto \mathbb{E}\big[V_{n+1}(\boldsymbol{X}_{n+1}, \boldsymbol{i})|\boldsymbol{X}_n = \boldsymbol{x}\big] = \mathbb{E}\big[V_{n+1}(\boldsymbol{x} + F_n(\boldsymbol{x}, \varepsilon_{n+1})\boldsymbol{i}, \boldsymbol{i})\big]$:
   $$\hat{U}_n^{\xi_M} \in \underset{\phi \in \mathcal{V}}{\operatorname{argmin}} \frac{1}{M} \sum_{m=1}^{M} \big|\phi(\boldsymbol{X}_n^{(m)}, \boldsymbol{I}_n^{(m)}) - \hat{V}_{n+1}^{\xi_M}(\boldsymbol{X}_n + F_n(\boldsymbol{X}_n^{(m)}, \varepsilon_{n+1}^{(m)})\boldsymbol{I}_n^{(m)}, \boldsymbol{I}_n^{(m)})\big|^2.$$
   (b) Compute $\hat{V}_n$ as the increasing envelope of $\hat{U}_n$ with respect to $\boldsymbol{i}$:
   $$\hat{V}_n^{\xi_M}(\boldsymbol{x}, \boldsymbol{i}) := \max_{\ell \in [N]} c_n(\boldsymbol{x}, \boldsymbol{i}, \boldsymbol{i}(1 - \boldsymbol{e}_\ell)) + \hat{U}_n^{\xi_M}(\boldsymbol{x}, \boldsymbol{i}(1 - \boldsymbol{e}_\ell)).$$
   $$\boldsymbol{I}_{n+1}^{\xi_M}(\boldsymbol{x}, \boldsymbol{i}) = \boldsymbol{i}(1 - \boldsymbol{e}_{\ell^{\xi_M}(\boldsymbol{x}, \boldsymbol{i})}), \text{ with } \ell^{\xi_M}(\boldsymbol{x}, \boldsymbol{i}) \in \underset{\boldsymbol{i}' \in \{0,1\}^N : \boldsymbol{i}' \leq \boldsymbol{i}}{\operatorname{argmax}} c_n(\boldsymbol{x}, \boldsymbol{i}, \boldsymbol{i}') + \hat{U}_n^{\xi_M}(\boldsymbol{x}, \boldsymbol{i}').$$
   (c) Given an initial condition $(\boldsymbol{X}_0, \boldsymbol{I}_0) := (\boldsymbol{x}, \boldsymbol{i})$, compute the optimal stopping policy $\boldsymbol{I}^{\xi_M}$ with:
   $$\boldsymbol{I}_{n+1}^{\xi_M} := \boldsymbol{I}_{n+1}^{\xi_M}(\boldsymbol{X}_n, \boldsymbol{I}_n^{\xi_M}), \quad \boldsymbol{X}_{n+1} = \boldsymbol{X}_n + F_n(\boldsymbol{X}_n, \varepsilon_{n+1})\boldsymbol{I}_{n+1}^{\xi_M}.$$

3.3 The convergence results

In order to analyze the convergence of the algorithms, we shall restrict the neural networks to the following class of functions:

$$\mathcal{N}_M := \Big\{ f : \mathbb{R}^N \times \mathbb{R}^{K_M(2+N)+1} \ni (\boldsymbol{x}, \theta) \mapsto \sum_{j=1}^{K_M} \alpha_j \sigma(\beta_j \cdot \boldsymbol{x} + \gamma_j) + \alpha_0, \text{ with } \theta := (\alpha_j, \beta_j, \gamma_j)_j \Big\},$$

where $\sigma : \mathbb{R} \to \mathbb{R}$ is some activation function, and $\{K_M\}_{M \geq 0}$ is such that

$$\delta_M := \frac{K_M}{M} \to 0 \text{ as } M \to \infty.$$

We shall also use the following notation: given to sequences of variables Y_M and Z_M, $M \geq 0$, we say that $Y_M = \mathcal{O}_{\mathbb{P}}(Z_M)$ if there exists a constant $C \geq 0$ such that $\mathbb{P}(|Y_M| \leq C|Z_M|) \to 0$ as $M \to \infty$. Since we mostly use neural networks to approximate conditional expectations, we also recall the following result from Kohler [12, Corollary 1], which will be useful to analyze the convergence of algorithm in the sense of $\mathcal{O}_{\mathbb{P}}$:

Lemma 3.1. *Let $(X_i, Y_i)_{1 \leq i \leq M}$ be a sequence of i.i.d. $\mathbb{R}^d \times \mathbb{R}$-valued random variables. Introduce the measurable functions:*

$$\mu(x) := \mathbb{E}[Y_1|X_1 = x], \quad \mu_M \in \underset{\phi \in \mathcal{N}_M}{\operatorname{argmin}} \sum_{i=1}^{M} |\phi(X_i) - Y_i|^2.$$

We assume that there exists two positive constants σ, λ such that:

$$\mathbb{E}\Big[\exp\Big(\frac{(Y_1 - m(X_1))^2}{\sigma^2}\Big)\Big|X_1\Big] \leq \lambda.$$

Then we have:

$$\mathbb{E}\Big[\Big|\mu_M(X_1) - \mu(X_1)\Big|^2\Big] = \mathcal{O}_{\mathbb{P}}\Big(\delta_M + \inf_{\phi \in \mathcal{N}_M} \Big\{\mathbb{E}\Big[\Big|\phi(X_1) - m(X_1)\Big|^2\Big]\Big\}\Big).$$

Before stating the main result, introduce the metric used to measure the error: for $f : \mathbb{R}^N \times \{0,1\}^N \to \mathbb{R}$, we denote:

$$\|f\|_{2,\infty}^{\xi_M} := \mathbb{E}\Big[\max_{\boldsymbol{i}\in\{0,1\}^N} \big|f(\boldsymbol{X}_0, \boldsymbol{i})\big|^2 \Big|\xi_M\Big]^{1/2},$$

where $\mathbb{P}\circ \boldsymbol{X}_0^{-1} = \mu$ and ξ_M is the set of all simulations used to train the neural network.

Assumption 3.2. *Assume that $\mathbb{P}\circ\boldsymbol{X}_n^{-1} = \mu$. Then, for all $\boldsymbol{i} \in \{0,1\}^N$, the random variable $\boldsymbol{X}_{n+1} = \boldsymbol{X}_n + F_n(\boldsymbol{X}_n, \varepsilon_{n+1})\boldsymbol{i}$ admits a bounded density with respect to μ conditionally on $\boldsymbol{X}_n$, denoted $\boldsymbol{x} \mapsto h_n(\boldsymbol{x}; \boldsymbol{X}_n, \boldsymbol{i})$.*

Theorem 3.3. *Let Assumption 3.2 hold.*
(i) *Let $\hat{V}^{\xi_M}$ be the function resulting from **Algorithm 1**. Then we have, as $M \to \infty$:*

$$\|\hat{V}_0^{\xi_M} - V_0\|_{2,\infty}^{\xi_M} = \mathcal{O}_{\mathbb{P}}\Big(\delta_M + \sup_{n\le k\le p}\inf_{\phi\in\mathcal{N}_M} |\phi - U_k|_{2,\infty}^{\xi_M}\Big),$$

for some $\delta_M \to 0$ as $M \to \infty$.
(ii) *Let $\hat{V}^{\xi_M}$ be the function resulting from **Algorithm 2**. Then we have, as $M \to \infty$:*

$$\|\hat{V}_0^{\xi_M} - \tilde{V}_0\|_{2,\infty}^{\xi_M} = \mathcal{O}_{\mathbb{P}}\Big(\delta_M + \sup_{n\le k\le p}\inf_{\phi\in\mathcal{N}_M} |\phi - U_k|_{2,\infty}^{\xi_M}\Big),$$

for some $\delta_M \to 0$ as $M \to \infty$.

Proof. We only write the proof of (i), as (ii) proceeds exactly from the same arguments. For $n \in [p]$, introduce the following functions:

$$\begin{aligned} U_n(\boldsymbol{x}, \boldsymbol{i}) &:= \mathbb{E}\big[V_{n+1}(\boldsymbol{x} + F_n(\boldsymbol{x}, \varepsilon_{n+1})\boldsymbol{i}, \boldsymbol{i})\big], \\ \bar{U}_n^{\xi_M}(\boldsymbol{x}, \boldsymbol{i}) &:= \mathbb{E}\big[\hat{V}_{n+1}^{\xi_M}(\boldsymbol{x} + F_n(\boldsymbol{x}, \varepsilon_{n+1})\boldsymbol{i}, \boldsymbol{i})\big], \end{aligned}$$

for all $(\boldsymbol{x}, \boldsymbol{i}) \in \mathbb{R}^N \times \{0,1\}^N$. Then we have, by definition of **Algorithm 1** and Proposition 2.1:

$$\begin{aligned} \|\hat{V}_n^{\xi_M} - V_n\|_{2,\infty}^{\xi_M} &\le \mathbb{E}\Big[\max_{\boldsymbol{i}\in\{0,1\}^N} \big|\max_{\boldsymbol{i'}\le\boldsymbol{i}} \hat{U}_n^{\xi_M}(X_n, \boldsymbol{i'}) - \max_{\boldsymbol{i'}\le\boldsymbol{i}} U_n(X_n, \boldsymbol{i'})\big|^2\Big|\xi_M\Big]^{1/2} \\ &\le \|\hat{U}_n^{\xi_M} - U_n\|_{2,\infty}^{\xi_M} \\ &\le \|\hat{U}_n^{\xi_M} - \bar{U}_n^{\xi_M}\|_{2,\infty}^{\xi_M} + \|\bar{U}_n^{\xi_M} - U_n\|_{2,\infty}^{\xi_M}. \end{aligned}$$

Now, introduce the set:

$$A_n^M := \Big\{\|\hat{U}_n^{\xi_M} - \bar{U}_n^{\xi_M}\|_{2,\infty}^{\xi_M} \le C_n\Big(\delta_M + \inf_{\phi\in\mathcal{N}_M}\|\phi - \bar{U}_n^{\xi_M}\|_{2,\infty}^{\xi_M}\Big)\Big\},$$

where $\delta_M \to 0$ and C_n is such that $\mathbb{P}(A_n^M) \to 1$ as $M \to \infty$, see Lemma 3.1. On this set, we have:

$$\begin{aligned} \|\hat{V}_n^{\xi_M} - V_n\|_{2,\infty}^{\xi_M} &\le C_n\Big(\delta_M + \inf_{\phi\in\mathcal{N}_M}\|\phi - \bar{U}_n^{\xi_M}\|_{2,\infty}^{\xi_M}\Big) + \|\bar{U}_n^{\xi_M} - U_n\|_{2,\infty}^{\xi_M} \\ &\le C_n\Big(\delta_M + \inf_{\phi\in\mathcal{N}_M}\|\phi - U_n\|_{2,\infty}\Big) + (1 + C_n)\|\bar{U}_n^{\xi_M} - U_n\|_{2,\infty}^{\xi_M}. \end{aligned} \tag{3.3}$$

Now, observe that:

$$\begin{aligned} \|\bar{U}_n^{\xi_M} - U_n\|_{2,\infty}^{\xi_M} &\le \mathbb{E}\Big[\max_{\boldsymbol{i}\in\{0,1\}^N} \big|\mathbb{E}[\hat{V}_{n+1}^{\xi_M}(\boldsymbol{X}_{n+1}, \boldsymbol{i})|\boldsymbol{X}_n, \xi_M] - \mathbb{E}[V_{n+1}(\boldsymbol{X}_{n+1}, \boldsymbol{i})|\boldsymbol{X}_n, \xi_M]\big|^2\Big|\xi_M\Big]^{1/2} \\ &\le \mathbb{E}\Big[\mathbb{E}\Big[\max_{\boldsymbol{i}\in\{0,1\}^N} \big|\hat{V}_{n+1}^{\xi_M}(\boldsymbol{X}_{n+1}, \boldsymbol{i}) - V_{n+1}(\boldsymbol{X}_{n+1}, \boldsymbol{i})\big|^2 |\boldsymbol{X}_n\Big]\Big|\xi_M\Big]^{1/2} \end{aligned}$$

with $\boldsymbol{X}_{n+1} = \boldsymbol{X}_n + F_n(\boldsymbol{X}_n, \varepsilon_{n+1})\boldsymbol{i}$. Now, using Assumption 3.2, we obtain:

$$\begin{aligned}
\|\bar{U}_n^{\xi_M} - U_n\|_{2,\infty}^{\xi_M} \leq & \mathbb{E}\Big[\int_{\mathbb{R}^N} \max_{\boldsymbol{i}\in\{0,1\}^N} \big|\hat{V}_{n+1}^{\xi_M}(\boldsymbol{x}, \boldsymbol{i}) - V_{n+1}(\boldsymbol{x}, \boldsymbol{i})\big|^2 h_n(\boldsymbol{x}; \boldsymbol{X}_n, \boldsymbol{i})\mu(d\boldsymbol{x})|\xi_M\Big]^{1/2} \\
\leq & \|h\|_\infty \mathbb{E}\Big[\int_{\mathbb{R}^N} \max_{\boldsymbol{i}\in\{0,1\}^N} \big|\hat{V}_{n+1}^{\xi_M}(\boldsymbol{x}, \boldsymbol{i}) - V_{n+1}(\boldsymbol{x}, \boldsymbol{i})\big|^2 \mu(d\boldsymbol{x})|\xi_M\Big]^{1/2} \\
= & \|h\|_\infty \|\hat{V}_{n+1}^{\xi_M} - V_{n+1}\|_{2,\infty}^{\xi_M}.
\end{aligned}$$

Plugging this into (3.3), we have on A_n^M:

$$\begin{aligned}
\|\hat{V}_n^{\xi_M} - V_n\|_{2,\infty}^{\xi_M} \leq & C_n\Big(\delta_M + \inf_{\phi\in\mathcal{N}_M} \|\phi - U_n\|_{2,\infty}\Big) + (1 + C_n)\|h\|_\infty \|\hat{V}_{n+1}^{\xi_M} - V_{n+1}\|_{2,\infty}^{\xi_M} \\
\leq & C\Big(\delta_M + \sup_{0\leq k\leq p} \inf_{\phi\in\mathcal{N}_M} \|\phi - U_k\|_{2,\infty}\Big) + (1 + C)\|h\|_\infty \|\hat{V}_{n+1}^{\xi_M} - V_{n+1}\|_{2,\infty}^{\xi_M},
\end{aligned}$$

with $C := \max_{n\in[p]} C_n$. Then, by induction, we have on $\cap_{n=0}^{p-1} A_n^M$, using the fact that $\hat{V}_N^{\xi_M} = g$:

$$\|\hat{V}_0^{\xi_M} - V_0\|_{2,\infty}^{\xi_M} \leq C(1 + C)^p \|h\|_\infty^p \Big(\delta_M + \sup_{0\leq k\leq p} \inf_{\phi\in\mathcal{N}_M} \|\phi - U_k\|_{2,\infty}\Big).$$

We conclude the proof by observing that $\mathbb{P}[(\cap_{n=0}^{p-1} A_n^M)^c] \leq \sum_{n=0}^p \mathbb{P}[(A_n^M)^c] \to 0$ as $M \to \infty$. ■

Theorem 3.4. (i) *Let $\boldsymbol{I}^{\xi_M}$ the stopping strategy provided by **Algorithm 1**, and denote for $n \in [p]$:*

$$J_n^{\xi_M} := \mathbb{E}\Big[\sum_{k=n}^{p-1} c_n(\boldsymbol{X}_k, \boldsymbol{X}_{k+1}, \boldsymbol{I}_{k+1}^{\xi_M}) + g(\boldsymbol{X}_p)\Big],$$

where the dynamics of $\boldsymbol{X}$ is controlled by $\boldsymbol{I}^M$. Then we have:

$$\|J_0^{\xi_M} - V_0\|_{2,\infty}^{\xi_M} = \mathcal{O}_{\mathbb{P}}(\delta_M + \sup_{0\leq k\leq p} \inf_{\phi\in\mathcal{N}_M} \|\phi - U_k\|_{2,\infty}),$$

with $\delta_M \to 0$ as $M \to \infty$.
(ii) *Let $\boldsymbol{I}^{\xi_M}$ the stopping strategy provided by **Algorithm 2**, and define $J_n^{\xi_M}$ as above. Then we have:*

$$\|J_0^{\xi_M} - \tilde{V}_0\|_{2,\infty}^{\xi_M} = \mathcal{O}_{\mathbb{P}}(\delta_M + \sup_{0\leq k\leq p} \inf_{\phi\in\mathcal{N}_M} \|\phi - U_k\|_{2,\infty}),$$

with $\delta_M \to 0$ as $M \to \infty$.

Proof. For simplicity, we write the proof for $c = 0$. We only detail the argument for (i), as (ii) is proved in the very same way. First, observe that:

$$\Delta_n := \|J_n^{\xi_M} - V_n\|_{2,\infty}^{\xi_M} \leq \|\hat{V}_n^{\xi_M} - V_n\|_{2,\infty}^{\xi_M} + \|\hat{V}_n^{\xi_M} - J_n^{\xi_M}\|_{2,\infty}^{\xi_M}. \tag{3.4}$$

Yet, we have:

$$\begin{aligned}
\|\hat{V}_n^{\xi_M} - J_n^{\xi_M}\|_{2,\infty}^{\xi_M} = & \mathbb{E}\Big[\max_{\boldsymbol{i}\in\{0,1\}^N} \Big|\hat{V}_n^{\xi_M}(\boldsymbol{X}_n, \boldsymbol{i}) - J_n^M(\boldsymbol{X}_n, \boldsymbol{i})\Big|^2 |\xi_M\Big]^{1/2} \\
= & \mathbb{E}\Big[\max_{\boldsymbol{i}\in\{0,1\}^N} \Big|\hat{U}_n^{\xi_M}(\boldsymbol{X}_n, \boldsymbol{I}_{n+1}^{\xi_M}) - \mathbb{E}\Big[J_{n+1}^{\xi_M}(\boldsymbol{X}_{n+1}, \boldsymbol{I}_{n+1}^{\xi_M})|\boldsymbol{X}_n\Big]\Big|^2 |\xi_M\Big]^{1/2} \\
\leq & \|\hat{U}_n^{\xi_M} - \bar{U}_n^{\xi_M}\|_{2,\infty}^{\xi_M} + \mathbb{E}\Big[\max_{\boldsymbol{i}\in\{0,1\}^N} \Big|\bar{U}_n^{\xi_M}(\boldsymbol{X}_n, \boldsymbol{I}_{n+1}^{\xi_M}) - \mathbb{E}\Big[J_{n+1}^{\xi_M}(\boldsymbol{X}_{n+1}, \boldsymbol{I}_{n+1}^{\xi_M})|\boldsymbol{X}_n\Big]\Big|^2 |\xi_M\Big]^{1/2}
\end{aligned}$$

$$\leq \|\hat{U}_n^{\xi_M} - \bar{U}_n^{\xi_M}\|_{2,\infty}^{\xi_M} + \mathbb{E}\Big[\max_{\boldsymbol{i}\in\{0,1\}^N} \Big|\bar{U}_n^{\xi_M}(\boldsymbol{X}_n, \boldsymbol{I}_{n+1}^{\xi_M}) - \mathbb{E}\Big[V_{n+1}(\boldsymbol{X}_{n+1}, \boldsymbol{I}_{n+1}^{\xi_M})|\boldsymbol{X}_n\Big]\Big|^2 |\xi_M\Big]^{1/2}$$

$$+ \mathbb{E}\Big[\max_{\boldsymbol{i}\in\{0,1\}^N} \Big|\mathbb{E}\Big[V_{n+1}(\boldsymbol{X}_{n+1}, \boldsymbol{I}_{n+1}^{\xi_M})|\boldsymbol{X}_n\Big] - \mathbb{E}\Big[J_{n+1}^{\xi_M}(\boldsymbol{X}_{n+1}, \boldsymbol{I}_{n+1}^{\xi_M})|\boldsymbol{X}_n\Big]\Big|^2 |\xi_M\Big]^{1/2}$$

$$\leq \|\hat{U}_n^{\xi_M} - \bar{U}_n^{\xi_M}\|_{2,\infty}^{\xi_M} + \mathbb{E}\Big[\max_{\boldsymbol{i}\in\{0,1\}^N} \mathbb{E}\Big[\Big|\hat{V}_{n+1}^{\xi_M}(\boldsymbol{X}_{n+1}, \boldsymbol{I}_{n+1}^{\xi_M}) - V_{n+1}(\boldsymbol{X}_{n+1}, \boldsymbol{I}_{n+1}^{\xi_M})\Big|^2 |\boldsymbol{X}_n\Big]|\xi_M\Big]^{1/2}$$

$$+ \mathbb{E}\Big[\max_{\boldsymbol{i}\in\{0,1\}^N} \mathbb{E}\Big[\Big|V_{n+1}(\boldsymbol{X}_{n+1}, \boldsymbol{I}_{n+1}^{\xi_M}) - J_{n+1}^{\xi_M}(\boldsymbol{X}_{n+1}, \boldsymbol{I}_{n+1}^{\xi_M})\Big|^2 |\boldsymbol{X}_n\Big]|\xi_M\Big]^{1/2}$$

$$\leq \|\hat{U}_n^{\xi_M} - \bar{U}_n^{\xi_M}\|_{2,\infty}^{\xi_M} + \|h\|_\infty \|\hat{V}_{n+1}^{\xi_M} - V_{n+1}\|_{2,\infty}^{\xi_M} + \|h\|_\infty \|\hat{V}_{n+1} - J_{n+1}^{\xi_M}\|_{2,\infty}^{\xi_M},$$

where we used Assumption 3.2 as in the proof of Theorem 3.3. Plugging this into (3.4), we obtain:

$$\Delta_n \leq \|\hat{V}_n^{\xi_M} - V_n\|_{2,\infty}^{\xi_M} + \|\hat{U}_n^{\xi_M} - \bar{U}_n^{\xi_M}\|_{2,\infty}^{\xi_M} + \|h\|_\infty \|\hat{V}_{n+1}^{\xi_M} - V_{n+1}\|_{2,\infty}^{\xi_M} + \|h\|_\infty \Delta_{n+1},$$

from which we deduce the desired result from Theorem 3.3, Lemma 3.1 and the fact that $\Delta_p = 0$. ∎

4 Application to diffusion processes

Our objective is to numerically compute the value function of the multiple optimal stopping problem for a N-dimensional stopped diffusion, as defined in Talbi, Touzi & Zhang []. More precisely, we are interested in the following problem:

$$\begin{aligned} V_0 &:= \sup_{\boldsymbol{\tau}\in\mathcal{T}_{[0,T]}^N} \mathbb{E}\Big[\sum_{0\leq s\leq T} \boldsymbol{c}_s(\boldsymbol{X}_s)\cdot(\boldsymbol{I}_s - \boldsymbol{I}_{s-}) + g(\boldsymbol{X}_T)\Big] \\ &= \sup_{\boldsymbol{\tau}\in\mathcal{T}_{[0,T]}^N} \mathbb{E}\Big[\sum_{k=1}^N c_{\tau_k}^k(\boldsymbol{X}_{\tau_k}) + g(X_{\tau_1}^1, \dots, X_{\tau_N}^N)\Big]. \end{aligned} \tag{4.1}$$

where $\mathcal{T}_{[0,T]}^N$ denotes the set of $[0,T]$-valued N-tuples of stopping times and $\boldsymbol{X} := (X^1, \dots, X^N)$ is the system of interacting stopped diffusions:

$$dX_t^k = I_t^k\big(b_k(t, \boldsymbol{X}_t)dt + \sigma_k(t, \boldsymbol{X}_t)dW_t^k + \sigma_0(t, \boldsymbol{X}_t)dW_t^0\big) \tag{4.2}$$

for all $k \in [N] := \{1, \dots, N\}$, with $I_t^k := \mathbf{1}_{\tau^k > t}$ and where the standard Brownian motions $W^0, \dots, W^N$ are independent. The following assumption will be in force throughout this Section:

Assumption 4.1. *For $\phi \in \{\beta_k, \sigma_k, \sigma_0, c, g\}, k \in [N]$, ϕ, there exists a nonnegative constant $C \geq 0$ and $\beta \in (0,1]$ such that:*

$$\begin{aligned} |\phi(t, \boldsymbol{x}) - \phi(t, \boldsymbol{x}')| &\leq C|\boldsymbol{x} - \boldsymbol{x}'|, \\ |\phi(t, \boldsymbol{x}) - \phi(s, \boldsymbol{x})| &\leq C|t - s|^\beta, \end{aligned}$$

for all $(t,s) \in [0,T]^2$ and $(\boldsymbol{x}, \boldsymbol{x}') \times \mathbb{R}^N \times \mathbb{R}^N$.

4.1 Discrete-time approximation

Let $\pi := \{t_0, \dots, t_p\}$ be a partition of $[0,T]$, with $p \in \mathbb{N}^*$. For simplicity, we assume that all the subintervals have the same size, i.e., $t_k = k\frac{T}{p}$ for all $k \in [p]$. In what follows, we shall denote by $h := \frac{T}{p}$ the step of the partition π. In this paragraph, we denote by $\mathcal{T}_p$ the set of π-valued stopping times, and by $\mathcal{T}_p^N$ the set of N-tuples of π-valued stopping times. We then introduce the discrete time multiple optimal stopping problem:

$$V_0^h := \sup_{\boldsymbol{\tau}\in\mathcal{T}_p^N} \mathbb{E}\Big[\sum_{k=1}^N c_{\tau_k}^k(\boldsymbol{X}_{\tau_k}^h) + g(X_{\tau_1}^{1,h},\ldots,X_{\tau_N}^{N,h})\Big], \tag{4.3}$$

with $\boldsymbol{X}^h := (X^{1,h},\ldots,X^{N,h})$ denotes the Euler scheme of $\mathbf{X}$ for the partition π, that is:

$$\begin{aligned} X_t^{k,h} = x_k + \int_0^t I_{t^p(s)+h}^k b_k(t^p(s),\boldsymbol{X}_{t^p(s)}^h)ds + \int_0^t I_{t^p(s)+h}^k \sigma_k(t^p(s),\boldsymbol{X}_{t^p(s)}^h)dW_s^k \\ + \int_0^t I_{t^p(s)+h}^k \sigma_0(t^p(s),\boldsymbol{X}_{t^p(s)}^h)dW_s^0 \end{aligned} \tag{4.4}$$

for all $k\in[N]$ and $t\in[0,T]$, with $t_p(s)$ is the largest $t'\in\pi$ such that $t'\le s$. This dynamics correspond to the general dynamics (2.1) with:

$$F_n(\boldsymbol{x},\varepsilon_{n+1}) = \begin{pmatrix} b_1(n,\boldsymbol{x}) \\ \ldots \\ b_N(n,\boldsymbol{x})\end{pmatrix} h + \begin{pmatrix} \sigma_1(n,\boldsymbol{x})\varepsilon_{n+1}^{(1)} \\ \ldots \\ \sigma_N(n,\boldsymbol{x})\varepsilon_{n+1}^{(N)}\end{pmatrix}, \quad \text{and} \quad \varepsilon_{n+1}^{(k)} := W_{t_{n+1}}^k - W_{t_n}^k \text{ for all } k\in[N].$$

In the following Lemma (whose proof is relegated to the appendix), we estimate the error between $\boldsymbol{X}$ and its Euler scheme, for stopping times taking their values in the discrete set $\{t_0,\ldots,t_p\}$:

Lemma 4.2. *Let Assumption 4.1 hold. For all $\boldsymbol{\tau}\in\mathcal{T}_p^N$, we have:*

$$\sup_{\boldsymbol{\tau}\in\mathcal{T}_p^N} \mathbb{E}\Big[\sup_{t\in[0,T]} |\boldsymbol{X}_t - \boldsymbol{X}_t^h|^2\Big] \le C_{T,N} h^{2\beta\wedge 1},$$

where the constant $C_{T,N}$ only depends on T and N.

This estimate implies the following estimate between the two value functions:

Proposition 4.3. *Let Assumption 4.1 hold. Then we have:*

$$|V_0 - V_0^h| \le C_{T,N} h^{\beta\wedge 1/2},$$

where the constant $C_{T,N}$ only depends on T and N.

Proof. Introduce the value function of the optimal stopping problem of $\boldsymbol{X}$ on the set of π-valued stopping times:

$$V_0^p := \sup_{\boldsymbol{\tau}\in\mathcal{T}_p^N} \mathbb{E}\Big[\sum_{k=1}^N c_{\tau_k}^k(\boldsymbol{X}_{\tau_k}) + g(X_{\tau_1}^1,\ldots,X_{\tau_N}^N)\Big].$$

We have:

$$|V_0 - V_0^h| \le |V_0 - V_0^p| + |V_0^p - V_0^h|.$$

Assumption 4.1 and Lemma 4.2 imply that:

$$|V_0^p - V_0^h| \le C h^{\beta\wedge 1/2}.$$

for some constant $C\ge 0$. The inequality $V_0^p \le V_0$ is clear, as it simply comes from the fact that $\mathcal{T}_p^N \subset \mathcal{T}_{[0,T]}^N$. Now, for $\varepsilon>0$, let $\boldsymbol{\tau}^\varepsilon \in \mathcal{T}_{[0,T]}^N$ be an ε-optimal policy for the problem (4.1). We then define $\bar{\boldsymbol{\tau}}^\varepsilon = (\bar\tau_1^\varepsilon,\ldots,\bar\tau_p^\varepsilon)$ as follows: for each $k\in[p]$, $\bar\tau_k^\varepsilon$ is the smallest $t_m\in\pi$ such that $\tau_k^\varepsilon \le t_m$. Since $\{\bar\tau_k^\varepsilon \le t_m\} = \{\tau_k^\varepsilon \in (t_{m-1},t_m]\} \in \mathcal{F}_{t_m}$, $\bar\tau_k^\varepsilon$ is a stopping time for the filtration $\mathbb{F}^p$, and therefore

$\bar{\boldsymbol{\tau}}^\varepsilon \in \mathcal{T}_p^N$. Then, by Assumption 4.1 and the fact that g is Lipschitz-continuous, Lemma A.2 and the fact that $\bar{\boldsymbol{\tau}} \in \mathcal{T}_p^N$, we have:

$$
\begin{aligned}
V_0 \leq & \mathbb{E}\Big[\sum_{k=1}^N c^k_{\tau_k^\varepsilon}(\boldsymbol{X}_{\tau_k^\varepsilon}) + g\big(X^1_{\tau_1^\varepsilon}, \ldots, X^N_{\tau_N^\varepsilon}\big)\Big] + \varepsilon \leq \mathbb{E}\Big[\sum_{k=1}^N c^k_{\bar{\tau}_k^\varepsilon}(\boldsymbol{X}_{\bar{\tau}_k^\varepsilon}) + g\big(X^1_{\bar{\tau}_1^\varepsilon}, \ldots, X^N_{\bar{\tau}_N^\varepsilon}\big)\Big] + C_{T,N} h^{\beta\wedge 1/2} + \varepsilon \\
\leq & V_0^h + C_{T,N} h^{\beta\wedge 1/2} + \varepsilon.
\end{aligned}
$$

We conclude by arbitrariness of $\varepsilon > 0$, and remarking that we typically have $h \leq 1$. ∎

Remark 4.4. *If g only depends on the empirical measure of $\boldsymbol{X}$, one can show that the constant $C(N,T)$ in fact does not depend on N.*

4.2 Error for the alternative problem

In this paragraph, we consider the alternative problem

$$
\tilde{V}_0^h := \sup_{\boldsymbol{\tau} \in \tilde{\mathcal{T}}_p^N} \mathbb{E}\Big[\sum_{k=1}^N c^k_{\tau_k}(\boldsymbol{X}^h_{\tau_k}) + g(X^{1,h}_{\tau_1}, \ldots, X^{N,h}_{\tau_N})\Big], \tag{4.5}
$$

where agents may only be stopped one by one. As it is more challenging to compare $\tilde{V}_0^h$ to V_0 directly, we start by comparing $\tilde{V}_0^h$ to V_0^h:

Proposition 4.5. *Let Assumption 4.1 hold. Then we have, for some constant $C_{T,N}$ depending on T and N only:*

$$
V_0^h - C_N h^{\beta\wedge 1/2} \leq \tilde{V}_0^h \leq V_0^h.
$$

Proof. The inequality $\tilde{V}_0^h \leq V_0^h$ is clear, since $\tilde{\mathcal{T}}_p^N \subset \mathcal{T}_p^N$. For the other inequality, consider an optimal strategy $\boldsymbol{\tau}^* = (\tau_1^*, \ldots, \tau_N^*)$ for (4.3), whose existence is granted by Proposition 2.1, that is:

$$
V_0^h = \mathbb{E}\Big[\sum_{k=1}^N c_{\tau_k^*}(\boldsymbol{X}^{h,*}_{\tau_k^*}) + g(\boldsymbol{X}^{h,*}_T)\Big],
$$

where $(\boldsymbol{X}^{*,h}, \boldsymbol{I}^*)$ denotes the process (4.4) controlled by $\boldsymbol{\tau}^*$. By Lemma A.1, we may construct a strategy $\tilde{\boldsymbol{\tau}} = (\tilde{\tau}_1, \ldots, \tilde{\tau}_N) \in \tilde{\mathcal{T}}_p^N$ such that

$$
|\tau_k^* - \tilde{\tau}_k| \leq Nh, \text{ a.s.},
$$

and we denote by $\tilde{\boldsymbol{X}}^h$ the Euler scheme controlled by $\tilde{\boldsymbol{\tau}}$. By Assumption 4.1, we have the estimates:

$$
\begin{aligned}
V_0^h \leq & \mathbb{E}\Big[L\Big(\sum_{k=1}^N (1 + |\boldsymbol{X}^{h,*}_{\tau_k^*}|)|\tau_k^* - \tilde{\tau}_k|^\beta + |\boldsymbol{X}^{h,*}_{\tau_k^*} - \tilde{\boldsymbol{X}}^h_{\tilde{\tau}_k}| + |\boldsymbol{X}^{h,*}_T - \tilde{\boldsymbol{X}}^h_T|\Big) + \sum_{k=1}^N c_{\tilde{\tau}_k}(\tilde{\boldsymbol{X}}^h_{\tilde{\tau}_k}) + g(\tilde{\boldsymbol{X}}^h_T)\Big] \\
\leq & L\Big((1 + \sup_{n \in [p]} \mathbb{E}[|\boldsymbol{X}^{h,*}_{t_n}|])h^\beta + C'_{T,N}\sqrt{\mathbb{E}[|\boldsymbol{\tau} - \tilde{\boldsymbol{\tau}}|]}\Big) + \tilde{V}_0^h \\
\leq & C_{T,N} h^{\beta\wedge 1/2} + \tilde{V}_0^h,
\end{aligned}
$$

where we used Lemma A.3 for the second inequality. ∎

Combining Propositions 4.3 and 4.5, we immediately deduce the following result:

Proposition 4.6. *Let Assumption 4.1 hold. Then we have:*

$$
|V_0 - \tilde{V}_0^h| \leq C_{T,N} h^{\beta\wedge 1/2},
$$

where the constant $C_{T,N}$ only depends on T and N.

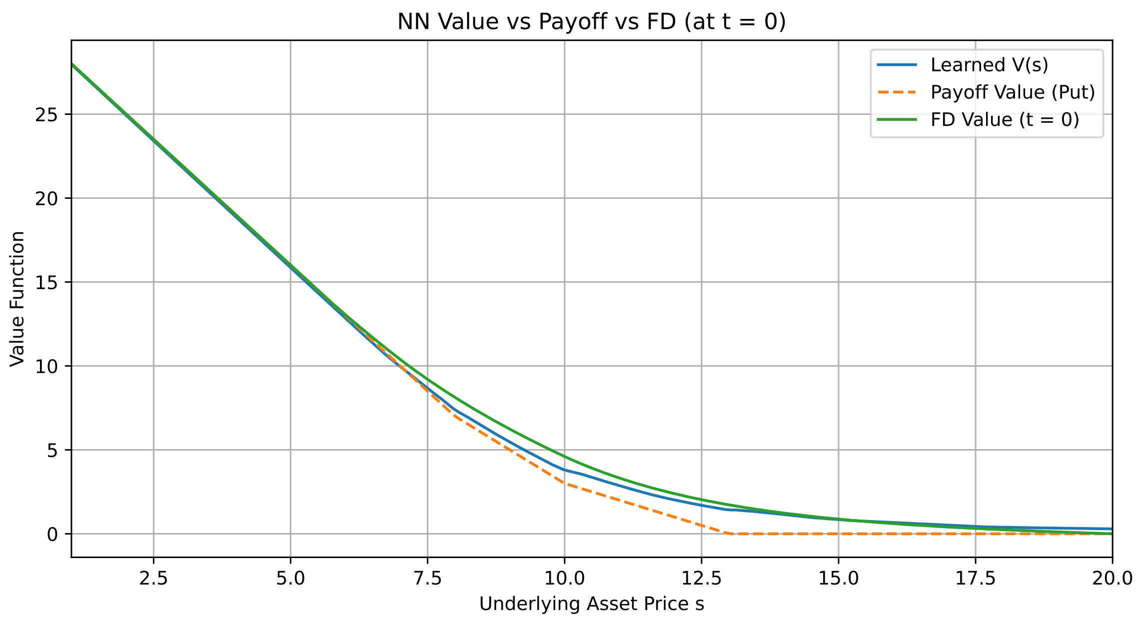

Figure 1: Multi American Put

4.3 Numerical implementation

Multipe American Put. We first test our algorithms on the following problem, which corresponds to the price of a basket of many American Puts, defined for all $\boldsymbol{x} = (x_1, \dots, x_N) \in \mathbb{R}_+^N$ by:

$$V_0(\boldsymbol{x}) := \sup_{\boldsymbol{\tau} \in \mathcal{T}_{[0,T]}^N} \mathbb{E}\Big[\sum_{j=1}^{N} (K_j - S_{\tau_j}^j)_+\Big],$$

where and $K_j \geq 0$, $S_t^j = x \exp\left([\mu_j - \sigma_j^2/2]t + \sigma_j W_t^j\right)$, $j \in [N]$, with $W^1, \dots, W^N$ independent Brownian motions. Denoting $S^{j,h}$ the Euler scheme of S^j, we also introduce:

$$V_0^h(\boldsymbol{x}) := \sup_{\boldsymbol{\tau} \in \mathcal{T}_p^N} \mathbb{E}\Big[\sum_{j=1}^{N} (K_j - S_t^{j,h})_+\Big].$$

Our objective is to compute V_0^h through our deep learning algorithm and to compare to V_0, which can be easily approximated by finite difference once we observe that $V_0(\boldsymbol{x}) = \sum_{j=1}^N V_0^j(x_j)$, with:

$$V_0^j(x_j) := \sup_{\boldsymbol{\tau} \in \mathcal{T}_{[0,T]}^N} \mathbb{E}\Big[(K_j - S_{\tau_j}^j)_+\Big],$$

which corresponds to a single-agent optimal stopping problem.
On Figure 1, we compare the prices of the multiple American Put ($N = 3$) respectively given by our neural network and by finite differences. More precisely, we draw the "diagonal function", that is, $x \mapsto V_0(x, \dots, x)$.

A nonlinear example. The following example is a sannity check to verify that our algorithms also works for "non-separable" utilities and to visualize the impact of the extra error added by

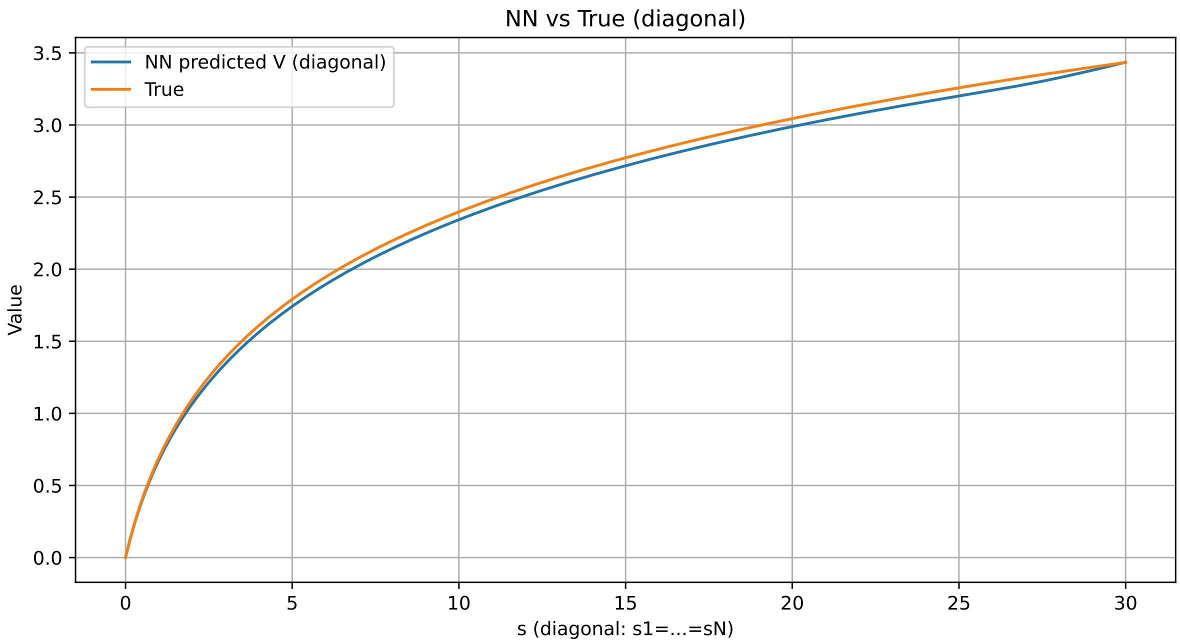

Figure 2: Multi American Put

Algorithm 2. We define:

$$V_0(\boldsymbol{x}) := \sup_{\boldsymbol{\tau} \in \mathcal{T}^N_{[0,T]}} \mathbb{E}\Big[\log\Big(1 + \frac{1}{N}\sum_{j=1}^N S^j_{\tau_j}\Big)\Big],$$

where the processes $S^1, \ldots, S^N$ are defined as the in previous example, with the extra condition that $\mu_j \le 0$ for all $j \in [N]$. We also introduce:

$$V_0^h(\boldsymbol{x}) := \sup_{\boldsymbol{\tau} \in \mathcal{T}^N_p} \mathbb{E}\Big[\log\Big(1 + \frac{1}{N}\sum_{j=1}^N S^{j,h}_{\tau_j}\Big)\Big],$$

Note that V_0 can be computed explicitely, as by Jensen inequality:

$$V_0(x) \le \sup_{\boldsymbol{\tau} \in \mathcal{T}^N_{[0,T]}} \log\Big(\mathbb{E}\Big[1 + \frac{1}{N}\sum_{j=1}^N S_{\tau_j}\Big]\Big) \le \log\Big(1 + \frac{1}{N}\sum_{j=1}^N x_j\Big),$$

coming from the fact that the processes $\{S^j\}_{j\in[N]}$ are supermartingales under the condition $\mu_j \le 0$. Since the above upper bound is reached for $\tau_j = 0$ for all $j \in [N]$, we conclude that $V_0(\boldsymbol{x}) = \log\Big(1 + \frac{1}{N}\sum_{j=1}^N x_j\Big)$. We may proceed the same way with V_0^h, and we have in fact $V_0 = V_0^h$. The goal of this example is then to visualize the error due to **Algorithm 2**, and we therefore introduce:

$$\tilde{V}_0^h(\boldsymbol{x}) := \sup_{\boldsymbol{\tau} \in \tilde{\mathcal{T}}^N_p} \mathbb{E}\Big[\log\Big(1 + \frac{1}{N}\sum_{j=1}^N S^{j,h}_{\tau_j}\Big)\Big].$$

Figure 2, compares the actual value function and $\tilde{V}_0$, also on the diagonal $\boldsymbol{x} = (x, \ldots, x)$ and for $N = 3$.

A Technical results

Lemma A.1. *Let $\boldsymbol{\tau} = (\tau_1, \dots, \tau_N) \in \mathcal{T}_p^N$. There exists $\tilde{\boldsymbol{\tau}} \in \tilde{\mathcal{T}}_p^N$ such that $|\boldsymbol{\tau} - \tilde{\boldsymbol{\tau}}|_\infty \leq N\Delta$.*

Proof. Recall the following notation: $I_t^k = \mathbf{1}_{t \leq \tau_k}$ for $k \in [N]$ and $t \in \{t_1, \dots, t_p\}$. For every time t_n, we introduce a set $\mathcal{J}_{t_n}$, which can be interpreted as the set of indices left to stop at time t_n. At every time, only the smallest element of the set, that is, $\min \mathcal{J}_{t_n}$, is stopped. More precisely, we define the set-valued random sequence $(\mathcal{J}_{t_n})_{n \in [p]}$ as follows:

$$\mathcal{J}_{t_0} := \{k \in [N] : I_{t_0}^k - I_{t_1}^k = 1\},$$

and, for $n \in [p-1]^*$,

$$\mathcal{J}_{t_n} = \left(\mathcal{J}_{t_{n-1}} \backslash \{\min \mathcal{J}_{t_{n-1}}\}\right) \cup \{k \in [N] : I_{t_n}^k - I_{t_{n+1}}^k = 1\}.$$

We also set $\mathcal{J}_p = \emptyset$. Then, for all $k \in [N]$, we introduce:

$$\tilde{\tau}_k := \min\{t_n \geq 0 : \min \mathcal{J}_{t_n} = k\} \wedge t_p. \tag{A.1}$$

We now verify that $\tilde{\boldsymbol{\tau}} := \{\tilde{\tau}_1, \dots, \tilde{\tau}_N\}$ satisfies the desired properties.

First, since I^k is a predictable process, by induction we see that the set-valued process $t \mapsto \mathcal{J}_t$ is adapted to filtration $\mathbb{F}^N$. By (A.1), it is then clear that $\tilde{\tau}_k$ is a $\mathbb{F}^N$-stopping time. Moreover, for $l \neq k$ and $n \in [p-1]$, we have:

$$\begin{aligned}\{\tilde{\tau}_k = \tilde{\tau}_l\} = \bigcup_{n=0}^{p} \{\tilde{\tau}_k = \tilde{\tau}_l = t_n\} = \Big(\bigcup_{n=0}^{p-1} \{k = \min \mathcal{J}_{t_n} = l\}\Big) \cup \{\tilde{\tau}_k = \tilde{\tau}_l = t_p\} \\ = \{\tilde{\tau}_k = \tilde{\tau}_l = t_p\},\end{aligned}$$

which implies that $\tilde{\boldsymbol{\tau}} \in \tilde{\mathcal{T}}_p^N$. Finally, we estimate $|\tau_k - \tilde{\tau}_k|$. Observe that:

$$\tau_k = \min\{t_n \geq t_0 : k \in \mathcal{J}_{t_n}\} \wedge t_p.$$

Note that we also have:

$$\tilde{\tau}_k = \min\{t_n \geq t_0 : k \in \mathcal{J}_{t_n} \text{ and } k \notin \mathcal{J}_{t_{n+1}}\} \wedge t_p,$$

from which we deduce that:

$$|\tau_k - \tilde{\tau}_k| = |\{n \in [p] : k \in \mathcal{J}_{t_n}\}|\Delta. \tag{A.2}$$

Now, observe that if $\mathcal{J}_{t_n}$ is nonempty, then at every time $m \geq n$, we keep removing one index from $\mathcal{J}_{t_m}$ until it is empty, although new indices may be added. However, no index $k \in [N]$ can return to $\mathcal{J}$ after exiting it. Therefore, since there is at most N indices to through $\mathcal{J}$, we have, $\mathbb{P}$-almost surely:

$$\mathcal{J}_{t_n} \neq \emptyset \Rightarrow \mathcal{J}_{(t_n + N\Delta) \wedge t_p} = \emptyset.$$

This implies that $|\{n \in [p] : k \in \mathcal{J}_{t_n}\}| \leq N$, and by (A.2) we conclude the proof. ■

Proof of Lemma 4.2 Without loss of generality, we assume that $\sigma_0 = 0$. Fix $\boldsymbol{\tau} \in \mathcal{T}_p^N$. For all $k \in [N]$ and $t \in [0, T]$, we have:

$$\mathbb{E}\big[\sup_{u \leq t} |X_u^k - X_u^{k,h}|^2\big] \leq 2\mathbb{E}\Big[T\int_0^t |b_k(s, \boldsymbol{X}_s) - b_k(t_p(s), \boldsymbol{X}_{t_p(s)}^h)|^2 ds + \int_0^t |\sigma_k(s, \boldsymbol{X}_s) - \sigma_k(t_p(s), \boldsymbol{X}_{t_p(s)}^h)|^2 ds\Big]$$

$$
\begin{aligned}
\leq & 2\mathbb{E}\Big[\int_0^t \big(T|b_k(t_p(s),\boldsymbol{X}_{t_p(s)}) - b_k(t_p(s),\boldsymbol{X}^h_{t_p(s)})|^2 + |\sigma_k(t_p(s),\boldsymbol{X}_{t_p(s)}) - \sigma_k(t_p(s),\boldsymbol{X}^h_{t_p(s)})|^2\big)ds\Big] \\
& + 2\mathbb{E}\Big[\int_0^t \big(T|b_k(s,\boldsymbol{X}_s) - b_k(t_p(s),\boldsymbol{X}^h_{t_p(s)})|^2 + |\sigma_k(s,\boldsymbol{X}_s) - \sigma_k(t_p(s),\boldsymbol{X}_{t_p(s)})|^2\big)ds\Big] \\
\leq & 2L(T+1)\int_0^t \mathbb{E}\big[\sup_{u\leq s}|\boldsymbol{X}_u - \boldsymbol{X}^h_u|^2 ds\big] \\
& + 2L(T+1)\int_0^t \mathbb{E}\big[|s - t^p(s)|^{2\beta} + |\boldsymbol{X}_s - \boldsymbol{X}_{t^p(s)}|^2 ds\big],
\end{aligned}
$$

where we successively used BDG inequality, the fact that $|I^k| \leq 1$, the Lipschitz-continuity of the coefficients b_k and σ_k in $\boldsymbol{x}$ and their β-Hölder-continuity in t. We then deduce from Gronwall's lemma that:

$$
\mathbb{E}\big[\sup_{t\leq T}|\boldsymbol{X}_t - \boldsymbol{X}^h_t|^2\big] \leq C_{T,N}\Big(h^{2\beta} + \int_0^t \mathbb{E}\big[|\boldsymbol{X}_s - \boldsymbol{X}_{t^p(s)}|^2\big]ds\Big). \tag{A.3}
$$

Now observe that:

$$
\begin{aligned}
\mathbb{E}\big[|X^k_s - X^k_{t^p(s)}|^2\big] \leq & 2\mathbb{E}\Big[\int_{t^p(s)}^s \big(T|b_k(r,\boldsymbol{X}_r)|^2 + |\sigma_k(r,\boldsymbol{X}_r)|^2\big)dr\Big] \\
\leq & hC_T\Big(1 + \mathbb{E}\Big[\sup_{t\leq T}|\boldsymbol{X}_t|^2\Big]\Big) \leq hC_{T,N}
\end{aligned}
$$

Using again BDG inquality and the fact that $|I^k| \leq 1$, we have:

$$
\mathbb{E}\Big[\sup_{u\leq t}|X^k_u|^2\Big] \leq C_T\int_0^t \big(1 + \mathbb{E}\big[\sup_{u\leq s}|X_u|^2\big]\big)ds,
$$

from which we deduce by Gronwall's lemma again that the second order moment of $\boldsymbol{X}$ are bounded independently from the stopping policy $\boldsymbol{\tau}$. Therefore:

$$
\mathbb{E}\big[|X^k_s - X^k_{t^p(s)}|^2\big] \leq hC_{T,N}
$$

We finally obtain the desired result by plugging the above estimate into (A.3) and by observing that none of the constants involves $\boldsymbol{\tau}$. ■

Lemma A.2. *Let $\boldsymbol{\tau}, \tilde{\boldsymbol{\tau}} \in \mathcal{T}^N_{0,T}$. We denote by $\boldsymbol{X} := (X^1,\dots,X^N)$ (resp. $\tilde{\boldsymbol{X}} := (\tilde{X}^1,\dots,\tilde{X}^N)$ the dynamics (2.1) controlled by $\boldsymbol{\tau}$ (resp. $\tilde{\boldsymbol{\tau}}$). Then there exists $C_{T,N} \geq 0$ (depending on N and T) such that:*

$$
\mathbb{E}\Big[\sup_{t\in[0,T]}|\boldsymbol{X}_t - \tilde{\boldsymbol{X}}_t|^2\Big] \leq C_{T,N}\sum_{k=1}^N \sqrt{\mathbb{E}\big[|\tau^k - \tilde{\tau}^k|\big]}
$$

Proof. Without loss of generality, we write the proof for $\sigma_0 = 0$. Fix $t \in [0,T]$ and $k \in [N]$. Using convexity and Burkholder-Davis-Gundy inequalities, we have:

$$
\mathbb{E}\Big[\sup_{u\leq t}|X^k_u - \tilde{X}^k_u|^2\Big] \leq C\Big(\int_0^t \mathbb{E}\Big[|b^k_s I^k_s - \tilde{b}^k_s\tilde{I}^k_s|^2\Big]ds + \int_0^t \mathbb{E}\Big[|\sigma^k_s I^k_s - \tilde{\sigma}^k_s\tilde{I}^k_s|^2\Big]ds\Big),
$$

where we denote $\varphi^k_s := \varphi_k(s,\boldsymbol{X}_s)$, $\tilde{\varphi}^k_s := \varphi_k(s,\tilde{\boldsymbol{X}}_s)$ for $\varphi \in \{b,\sigma\}$, $I^k_s := \mathbf{1}_{s<\tau_k}$ and $\tilde{I}^k_s := \mathbf{1}_{s<\tilde{\tau}_k}$. Now observe that:

$$
\begin{aligned}
\mathbb{E}\Big[|b^k_s I^k_s - \tilde{b}^k_s\tilde{I}^k_s|^2\Big] \leq & \mathbb{E}\Big[|b^k_s|^2|I^k_s - \tilde{I}^k_s|\Big] + \mathbb{E}\Big[|b^k_s - \tilde{b}^k_s|^2\Big] \\
\leq & C\mathbb{E}\Big[(1 + |X^k_s|^2)|I^k_s - \tilde{I}^k_s|\Big] + \mathbb{E}\Big[|X^k_s - \tilde{X}^k_s|^2\Big]
\end{aligned}
$$

$$\leq C\sqrt{\mathbb{E}\Big[(1+|X_s^k|^4)\Big]}\sqrt{\mathbb{E}\Big[|I_s^k-\tilde{I}_s^k|\Big]+\mathbb{E}\Big[|X_s^k-\tilde{X}_s^k|^2\Big]}.$$

Using estimates similar to the proof of Lemma 4.2, we can show that $\mathbb{E}\Big[|X_s^k|^4\Big]$ is bounded by a constant depending on N and T only. Then, writing the same inequalities for the term in σ, observing that

$$\int_0^t \sqrt{\mathbb{E}\Big[|I_s^k-\tilde{I}_s^k|\Big]} \leq \sqrt{T\mathbb{E}\Big[\int_0^T |I_s^k-\tilde{I}_s^k|ds\Big]} = \sqrt{T\mathbb{E}\Big[|\tau_k-\tilde{\tau}_k|\Big]},$$

and using Gronwall's Lemma, we derive the desired result. ∎

Lemma A.3. *Let Assumption 4.1 holds. Let $\boldsymbol{\tau},\tilde{\boldsymbol{\tau}} \in \mathcal{T}_p^N$ and denote by $(\boldsymbol{X}^h,\boldsymbol{I})$ and $(\tilde{\boldsymbol{X}}^h,\tilde{\boldsymbol{I}})$ the corresponding processes defined by (4.4), with $\boldsymbol{X}_0=\tilde{\boldsymbol{X}}_0$. Then we have:*

$$\mathbb{E}\Big[\max_{n\in[p]}|\boldsymbol{X}_n^h-\tilde{\boldsymbol{X}}_n^h|^2\Big] \leq C_{p,N}\sqrt{\mathbb{E}[|\boldsymbol{\tau}-\tilde{\boldsymbol{\tau}}|]} \quad \textit{for all } n\in[p].$$

Proof. The result can be proven by adapting the same estimates as in the proof of Lemma A.2 to the dynamics (4.4). ∎

References

[1] Sebastian Becker, Patrick Cheridito, and Arnulf Jentzen. Deep optimal stopping. *Journal of Machine Learning Research*, 20(74):1–25, 2019.

[2] Mark Broadie and Paul Glasserman. A stochastic mesh method for pricing high-dimensional american options. *Journal of Computational Finance*, 7:35–72, 2004.

[3] René Carmona and Nizar Touzi. Optimal multiple stopping and valuation of swing options. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 18(2):239–268, 2008.

[4] Chengfan Gao, Siping Gao, Ruimeng Hu, and Zimu Zhu. Convergence of the backward deep BSDE method with applications to optimal stopping problems. *SIAM Journal on Financial Mathematics*, 14(4):1290–1303, 2023.

[5] Miryana Grigorova, Marie-Claire Quenez, and Peng Yuan. The non-linear multiple stopping problem: between the discrete and the continuous time. *arXiv preprint arXiv:2504.13503*, 2025.

[6] Ivan Guo, Nicolas Langrené, and Jiahao Wu. Simultaneous upper and lower bounds of American-style option prices with hedging via neural networks. *Quantitative Finance*, 25(4): 509–525, 2025.

[7] László Györfi, Michael Kohler, Adam Krzyżak, and Harro Walk. *A distribution-free theory of nonparametric regression*. Springer, 2002.

[8] Yuecai Han and Nan Li. A new deep neural network algorithm for multiple stopping with applications in options pricing. *Communications in Nonlinear Science and Numerical Simulation*, 117:106881, 2023.

[9] Côme Huré, Huyên Pham, Achref Bachouch, and Nicolas Langrené. Deep neural networks algorithms for stochastic control problems on finite horizon: convergence analysis. *SIAM Journal on Numerical Analysis*, 59(1):525–557, 2021.

[10] Ioannis Karatzas and Steven E Shreve. *Methods of mathematical finance*, volume 39. Springer, 1998.

[11] Magdalena Kobylanski, Marie-Claire Quenez, and Elisabeth Rouy-Mironescu. Optimal multiple stopping time problem. *The Annals of Applied Probability*, 21(4):1365 – 1399, 2011.

[12] Michael Kohler. Nonparametric regression with additional measurement errors in the dependent variable. *Journal of statistical planning and inference*, 136(10):3339–3361, 2006.

[13] Francis A Longstaff and Eduardo S Schwartz. Valuing american options by simulation: A simple least-squares approach. *The review of financial studies*, 14(1):113–147, 2001.

[14] Nicolai Meinshausen and Ben M Hambly. Monte Carlo methods for the valuation of multiple-exercise options. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 14(4):557–583, 2004.

[15] John Schoenmakers. A pure martingale dual for multiple stopping. *Finance and Stochastics*, 16(2):319–334, 2012.

[16] Albert N Shiryaev. *Optimal stopping rules*. Springer, 2008.

[17] Mehdi Talbi, Nizar Touzi, and Jianfeng Zhang. Dynamic programming equation for the mean field optimal stopping problem. *SIAM Journal on Control and Optimization*, 61(4):2140–2164, 2023.

[18] Mehdi Talbi, Nizar Touzi, and Jianfeng Zhang. Viscosity solutions for obstacle problems on Wasserstein space. *SIAM Journal on Control and Optimization*, 61(3):1712–1736, 2023.

[19] Mehdi Talbi, Nizar Touzi, and Jianfeng Zhang. From finite population optimal stopping to mean field optimal stopping. *The Annals of Applied Probability*, 34(5):4237–4267, 2024.

[20] Jiefei Yang and Guanglian Li. A deep primal-dual BSDE method for optimal stopping problems. *arXiv preprint arXiv:2409.06937*, 2024.