



Article

Pricing and Hedging American-Style Options with Deep Learning

Sebastian Becker ¹, Patrick Cheridito ^{1,*} and Arnulf Jentzen ²

¹ RiskLab, ETH Zurich, 8092 Zurich, Switzerland; sebastian.becker@math.ethz.ch

² Faculty of Mathematics and Computer Science, University of Münster, 48149 Münster, Germany; ajentzen@uni-muenster.de

* Correspondence: patrick.cheridito@math.ethz.ch

Received: 23 December 2019; Accepted: 15 July 2020; Published: 19 July 2020



Abstract: In this paper we introduce a deep learning method for pricing and hedging American-style options. It first computes a candidate optimal stopping policy. From there it derives a lower bound for the price. Then it calculates an upper bound, a point estimate and confidence intervals. Finally, it constructs an approximate dynamic hedging strategy. We test the approach on different specifications of a Bermudan max-call option. In all cases it produces highly accurate prices and dynamic hedging strategies with small replication errors.

Keywords: American option; Bermudan option; optimal stopping; lower bound; upper bound; hedging strategy; deep neural network

1. Introduction

Early exercise options are notoriously difficult to value. For up to three underlying risk factors, tree based and classical PDE approximation methods usually yield good numerical results; see, e.g., Forsyth and Vetzal (2002); Hull (2003); Reisinger and Witte (2012) and the references therein. To treat higher-dimensional problems, various simulation based methods have been developed; see, e.g., Tilley (1993); Barraquand and Martineau (1995); Carriere (1996); Andersen (2000); Longstaff and Schwartz (2001); Tsitsiklis and Van Roy (2001); García (2003); Broadie and Glasserman (2004); Bally et al. (2005); Kolodko and Schoenmakers (2006); Egloff et al. (2007); Broadie and Cao (2008); Berridge and Schumacher (2008); Jain and Oosterlee (2015). Haugh and Kogan (2004) as well as Kohler et al. (2010) have already used shallow¹ neural networks to estimate continuation values. More recently, in Sirignano and Spiliopoulos (2018) optimal stopping problems in continuous time have been solved by approximating the solutions of the corresponding free boundary PDEs with deep neural networks. In Becker et al. (2019a, 2019b), deep learning has been used to directly learn optimal stopping strategies. The main focus of these papers is to derive optimal stopping rules and accurate price estimates.

The goal of this article is to develop a deep learning method which learns the optimal exercise behavior, prices and hedging strategies from samples of the underlying risk factors. It first learns a candidate optimal stopping strategy by regressing continuation values on multilayer neural networks. Employing the learned stopping strategy on a new set of Monte Carlo samples gives a low-biased estimate of the price. Moreover, the candidate optimal stopping strategy can be used to construct an approximate solution to the dual martingale problem introduced by Rogers (2002) and Haugh and Kogan (2004), yielding a high-biased estimate and confidence intervals for the price. In the last step,

¹ meaning feedforward networks with a single hidden layer.

our method learns a dynamic hedging strategy in the spirit of Han et al. (2018) and Buehler et al. (2019). However, here, the continuation value approximations learned during the construction of the optimal stopping strategy can be used to break the hedging problem down into a sequence of smaller problems that learn the hedging portfolio only from one possible exercise date to the next. Alternative ways of computing hedging strategies consist in calculating sensitivities of option prices (see, e.g., Bally et al. 2005; Bouchard and Warin 2012; Jain and Oosterlee 2015) or approximating a solution to the dual martingale problem (see, e.g., Rogers 2002, 2010).

Our work is related to the preprints Lapeyre and Lelong (2019) and Chen and Wan (2019). Lapeyre and Lelong (2019) also use neural network regression to estimate continuation values. However, the networks are slightly different. While they work with leaky ReLU activation functions, we use tanh activation. Moreover, Lapeyre and Lelong (2019) study the convergence of the pricing algorithm as the number of simulations and the size of the network go to infinity, whereas we calculate a posteriori guarantees for the prices and use the estimated continuation value functions to implement efficient hedging strategies. Chen and Wan (2019) propose an alternative way of calculating prices and hedging strategies for American-style options by solving BSDEs.

The rest of the paper is organized as follows. In Section 2 we describe our neural network version of the Longstaff–Schwartz algorithm to estimate continuation values and construct a candidate optimal stopping strategy. In Section 3 the latter is used to derive lower and upper bounds as well as confidence intervals for the price. Section 4 discusses two different ways of computing dynamic hedging strategies. In Section 5 the results of the paper are applied to price and hedge a Bermudan call option on the maximum of different underlying assets. Section 6 concludes.

2. Calculating a Candidate Optimal Stopping Strategy

We consider an American-style option that can be exercised at any one of finitely² many times $0 = t_0 < t_1 < \dots < t_N = T$. If exercised at time t_n , it yields a discounted payoff given by a square-integrable random variable G_n defined on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P} = (\mathcal{F}_n)_{n=0}^N, \mathbb{P})$. We assume that \mathcal{F}_n describes the information available at time t_n and G_n is of the form $g(n, X_n)$ for a measurable function $g: \{0, 1, \dots, N\} \times \mathbb{R}^d \rightarrow [0, \infty)$ and a d -dimensional \mathbb{P} -Markov process³ $(X_n)_{n=0}^N$. We assume X_0 to be deterministic and \mathbb{P} to be the pricing measure. So that the value of the option at time 0 is given by

$$V = \sup_{\tau \in \mathcal{T}} \mathbb{E} G_\tau,$$

where \mathcal{T} is the set of all \mathbb{P} -stopping times $\tau: \Omega \rightarrow \{0, 1, \dots, N\}$. If the option has not been exercised before time t_n , its discounted value at that time is

$$V_{t_n} = \text{ess sup}_{\tau \in \mathcal{T}_n} \mathbb{E}[G_\tau \mid \mathcal{F}_n], \quad (1)$$

where \mathcal{T}_n is the set of all \mathbb{P} -stopping times satisfying $n \leq \tau \leq N$.

Obviously, $\tau_N \equiv N$ is optimal for $V_T = G_N$. From there, one can recursively construct the stopping times

$$\tau_n := \begin{cases} n & \text{if } G_n \geq \mathbb{E}[G_{\tau_{n+1}} \mid X_n] \\ \tau_{n+1} & \text{if } G_n < \mathbb{E}[G_{\tau_{n+1}} \mid X_n]. \end{cases} \quad (2)$$

Clearly, τ_n belongs to \mathcal{T}_n , and it can be checked inductively that

² This covers Bermudan options as well as American options that can only be exercised at a given time each day. Continuously exercisable options must be approximated by discretizing time.

³ That is, X_n is \mathcal{F}_n -measurable, and $\mathbb{E}[f(X_{n+1}) \mid \mathcal{F}_n] = \mathbb{E}[f(X_{n+1}) \mid X_n]$ for all $n \leq N - 1$ and every measurable function $f: \mathbb{R}^d \rightarrow \mathbb{R}$ such that $f(X_{n+1})$ is integrable.

$$V_{t_n} = \mathbb{E}[G_{\tau_n} | \mathcal{F}_n] = G_n \vee \mathbb{E}[V_{t_{n+1}} | X_n] \quad \text{for all } n \leq N-1.$$

In particular, τ_n is an optimizer of (1).

Recursion (2) is the theoretical basis of the Longstaff and Schwartz (2001) method. Its main computational challenge is the approximation of the conditional expectations $\mathbb{E}[G_{\tau_{n+1}} | X_n]$. It is well known that $\mathbb{E}[G_{\tau_{n+1}} | X_n]$ is of the form $c(X_n)$, where $c: \mathbb{R}^d \rightarrow \mathbb{R}$ minimizes the mean squared distance $\mathbb{E}[\{G_{\tau_{n+1}} - c(X_n)\}^2]$ over all Borel measurable functions from \mathbb{R}^d to \mathbb{R} ; see, e.g., Bru and Heinich (1985). The Longstaff–Schwartz algorithm approximates $\mathbb{E}[G_{\tau_{n+1}} | X_n]$ by projecting $G_{\tau_{n+1}}$ on the linear span of finitely many basis functions. However, it is also possible to project on a different subset. If the subset is given by $c^\theta(X_n)$ for a function family $c^\theta: \mathbb{R}^d \rightarrow \mathbb{R}$ parametrized by θ , one can apply the following variant⁴ of the Longstaff–Schwartz algorithm:

- (i) Simulate⁵ paths $(x_n^k)_{n=0}^N$, $k = 1, \dots, K$, of the underlying process $(X_n)_{n=0}^N$.
- (ii) Set $s_N^k \equiv N$ for all k .
- (iii) For $1 \leq n \leq N-1$, approximate $\mathbb{E}[G_{\tau_{n+1}} | X_n]$ with $c^{\theta_n}(X_n)$ by minimizing the sum

$$\sum_{k=1}^K \left(g(s_{n+1}^k, x_{s_{n+1}^k}^k) - c^{\theta}(x_n^k) \right)^2 \quad \text{over } \theta. \quad (3)$$

- (iv) Set

$$s_n^k := \begin{cases} n & \text{if } g(n, x_n^k) \geq c^{\theta_n}(x_n^k) \\ s_{n+1}^k & \text{otherwise.} \end{cases}$$

- (v) Define $\theta_0 := \frac{1}{K} \sum_{k=1}^K g(s_1^k, x_{s_1^k}^k)$, and set c^{θ_0} constantly equal to θ_0 .

In this paper we specify c^θ as a feedforward neural network, which in general, is of the form

$$a_I^\theta \circ \varphi_{q_{I-1}} \circ a_{I-1}^\theta \circ \dots \circ \varphi_{q_1} \circ a_1^\theta, \quad (4)$$

where

- $I \geq 1$ denotes the depth and q_0, q_1, \dots, q_I the numbers of nodes in the different layers;
- $a_1^\theta: \mathbb{R}^{q_0} \rightarrow \mathbb{R}^{q_1}, \dots, a_I^\theta: \mathbb{R}^{q_{I-1}} \rightarrow \mathbb{R}^{q_I}$ are affine functions;
- For $j \in \mathbb{N}$, $\varphi_j: \mathbb{R}^j \rightarrow \mathbb{R}^j$ is of the form $\varphi_j(x_1, \dots, x_j) = (\varphi(x_1), \dots, \varphi(x_j))$ for a given activation function $\varphi: \mathbb{R} \rightarrow \mathbb{R}$.

The components of the parameter θ consist of the entries of the matrices A_1, \dots, A_I and vectors b_1, \dots, b_I appearing in the representation of the affine functions $a_i^\theta x = A_i x + b_i$, $i = 1, \dots, I$. So, θ lives in \mathbb{R}^q for $q = \sum_{i=1}^I q_i(q_{i-1} + 1)$. To minimize (3) we choose a network with $q_I = 1$ and employ a stochastic gradient descent method.

⁴ The main difference between this algorithm and the one of Longstaff and Schwartz (2001) is the use of neural networks instead of linear combinations of basis functions. In addition, the sum in (3) is over all simulated paths, whereas in Longstaff and Schwartz (2001), only in-the-money paths are considered to save computational effort. While it is enough to use in-the-money paths to determine a candidate optimal stopping rule, we need accurate approximate continuation values for all $x \in \mathbb{R}^d$ to construct good hedging strategies in Section 4.

⁵ As usual, we simulate the paths (x_n^k) , $k = 1, \dots, K$, independently of each other.

3. Pricing

3.1. Lower Bound

Once $\theta_0, \theta_1, \dots, \theta_{N-1}$ have been determined, we set $\Theta = (\theta_0, \dots, \theta_{N-1})$ and define

$$\tau^\Theta := \min \left\{ n \in \{0, 1, \dots, N-1\} : g(n, X_n) \geq c^{\theta_n}(X_n) \right\}, \quad \text{where } \min \emptyset \text{ is understood as } N.$$

This defines a valid \mathbb{F} -stopping time. Therefore, $L = \mathbb{E} g(\tau^\Theta, X_{\tau^\Theta})$ is a lower bound for the optimal value V . However, typically, it is not possible to calculate the expectation exactly. Therefore, we generate simulations g^k of $g(\tau^\Theta, X_{\tau^\Theta})$ based on independent sample paths⁶ $(x_n^k)_{n=0}^N$, $k = K+1, \dots, K+K_L$, of $(X_n)_{n=0}^N$ and approximate L with the Monte Carlo average

$$\hat{L} = \frac{1}{K_L} \sum_{k=K+1}^{K+K_L} g^k.$$

Denote by $z_{\alpha/2}$ the $1 - \alpha/2$ quantile of the standard normal distribution and consider the sample standard deviation

$$\hat{\sigma}_L = \sqrt{\frac{1}{K_L - 1} \sum_{k=K+1}^{K+K_L} (g^k - \hat{L})^2}.$$

Then one obtains from the central limit theorem that

$$\left[\hat{L} - z_{\alpha/2} \frac{\hat{\sigma}_L}{\sqrt{K_L}}, \infty \right) \quad (5)$$

is an asymptotically valid $1 - \alpha/2$ confidence interval for L .

3.2. Upper Bound, Point Estimate and Confidence Intervals

Our derivation of an upper bound is based on the duality results of [Rogers \(2002\)](#); [Haugh and Kogan \(2004\)](#) and [Becker et al. \(2019a\)](#). By [Rogers \(2002\)](#) and [Haugh and Kogan \(2004\)](#), the optimal value V can be written as

$$V = \mathbb{E} \left[\max_{0 \leq n \leq N} (G_n - M_n) \right],$$

where $(M_n)_{n=0}^N$ is the martingale part of the smallest \mathbb{F} -supermartingale dominating the payoff process $(G_n)_{n=0}^N$. We approximate $(M_n)_{n=0}^N$ with the \mathbb{F} -martingale $(M_n^\Theta)_{n=0}^N$ obtained from the stopping decisions implied by the trained continuation value functions c^{θ_n} , $n = 0, \dots, N-1$, as in Section 3.2 of [Becker et al. \(2019a\)](#). We know from Proposition 7 of [Becker et al. \(2019a\)](#) that if $(\varepsilon_n)_{n=0}^N$ is a sequence of integrable random variables satisfying $\mathbb{E}[\varepsilon_n | \mathcal{F}_n] = 0$ for all $n = 0, 1, \dots, N$, then

$$U = \mathbb{E} \left[\max_{0 \leq n \leq N} (G_n - M_n^\Theta - \varepsilon_n) \right]$$

is an upper bound for V . As in [Becker et al. \(2019a\)](#), we use nested simulation⁷ to generate realizations m_n^k of $M_n^\Theta + \varepsilon_n$ along independent realizations $(x_n^k)_{n=0}^N$, $k = K+K_L+1, \dots, K+K_L+K_U$, of $(X_n)_{n=0}^N$ sampled independently of $(x_n^k)_{n=0}^N$, $k = 1, \dots, K$, and estimate U as

⁶ generated independently of $(x_n^k)_{n=0}^N$, $k = 1, \dots, K$

⁷ The use of nested simulation ensures that m_n^k are unbiased estimates of M_n^Θ , which is crucial for the validity of the upper bound. In particular, we do not directly approximate M_n^Θ with the estimated continuation value functions c^{θ_n} .

$$\hat{U} = \frac{1}{K_U} \sum_{k=K+K_L+1}^{K+K_L+K_U} \max_{0 \leq n \leq N} (g(n, x_n^k) - m_n^k).$$

Our point estimate of V is

$$\hat{V} = \frac{\hat{L} + \hat{U}}{2}.$$

The sample standard deviation of the estimator \hat{U} , given by

$$\hat{\sigma}_U = \sqrt{\frac{1}{K_U - 1} \sum_{k=K+K_L+1}^{K+K_L+K_U} \left(\max_{0 \leq n \leq N} (g(n, x_n^k) - m_n^k) - \hat{U} \right)^2},$$

can be used together with the one-sided confidence interval (5) to construct the asymptotically valid two-sided $1 - \alpha$ confidence interval

$$\left[\hat{L} - z_{\alpha/2} \frac{\hat{\sigma}_L}{\sqrt{K_L}}, \hat{U} + z_{\alpha/2} \frac{\hat{\sigma}_U}{\sqrt{K_U}} \right] \quad (6)$$

for the true value V ; see Section 3.3 of [Becker et al. \(2019a\)](#).

4. Hedging

We now consider a savings account together with $e \in \mathbb{N}$ financial securities as hedging instruments. We fix a positive integer M and introduce a time grid $0 = u_1 < u_2 < \dots < u_{NM}$ such that $u_{nM} = t_n$ for all $n = 0, 1, \dots, N$. We suppose that the information available at time u_m is described by \mathcal{H}_m , where $\mathbb{H} = (\mathcal{H}_m)_{m=0}^{NM}$ is a filtration satisfying $\mathcal{H}_{nM} = \mathcal{F}_n$ for all n . If any of the financial securities pay dividends, they are immediately reinvested. We assume that the resulting discounted⁸ value processes are of the form $P_{u_m} = p_m(Y_m)$ for measurable functions $p_m: \mathbb{R}^d \rightarrow \mathbb{R}^e$ and an \mathbb{H} -Markov process⁹ $(Y_m)_{m=0}^{NM}$ such that $Y_{nM} = X_n$ for all $n = 0, \dots, N$. A hedging strategy consists of a sequence $h = (h_m)_{m=0}^{NM-1}$ of functions $h_m: \mathbb{R}^d \rightarrow \mathbb{R}^e$ specifying the time- u_m holdings in $P_{u_m}^1, \dots, P_{u_m}^e$. As usual, money is dynamically deposited in or borrowed from the savings account to make the strategy self-financing. The resulting discounted gains at time u_m are given by

$$(h \cdot P)_{u_m} := \sum_{j=0}^{m-1} h_j(Y_j) \cdot (p_{j+1}(Y_{j+1}) - p_j(Y_j)) := \sum_{j=0}^{m-1} \sum_{i=1}^e h_j^i(Y_j) (p_{j+1}^i(Y_{j+1}) - p_j^i(Y_j)).$$

4.1. Hedging Until the First Possible Exercise Date

For a typical Bermudan option, the time between two possible exercise dates $t_n - t_{n-1}$ might range between a week and several months. In case of an American option, we choose $t_n = n\Delta$ for a small amount of time Δ such as a day. We assume τ^Θ does not stop at time 0. Otherwise, there is nothing to hedge. In a first step, we only compute the hedge until time t_1 . If the option is still alive at time t_1 , the hedge can then be computed until time t_2 and so on. To construct a hedge from time 0 to t_1 , we approximate the time- t_1 value of the option with $V_{t_1}^{\theta_1} = v^{\theta_1}(X_1)$ for the function $v^{\theta_1}(x) = g(1, x) \vee c^{\theta_1}(x)$, where $c^{\theta_1}: \mathbb{R}^d \rightarrow \mathbb{R}$ is the time- t_1 continuation value function estimated in

⁸ Discounting is done with respect to the savings account. Then, the discounted value of money invested in the savings account stays constant.

⁹ That is, Y_m is \mathcal{H}_m -measurable and $\mathbb{E}[f(Y_{m+1}) \mid \mathcal{H}_m] = \mathbb{E}[f(Y_{m+1}) \mid Y_m]$ for all $m \leq NM - 1$ and every measurable function $f: \mathbb{R}^d \rightarrow \mathbb{R}$ such that $f(Y_{m+1})$ is integrable.

Section 2. Then we search for hedging positions h_m , $m = 0, 1, \dots, M-1$, that minimize the mean squared error

$$\mathbb{E} \left[\left(\hat{V} + (h \cdot P)_{t_1} - V_{t_1}^{\theta_1} \right)^2 \right].$$

To do that we approximate the functions h_m with neural networks $h^\lambda: \mathbb{R}^d \rightarrow \mathbb{R}^e$ of the form (4) and try to find parameters $\lambda_0, \dots, \lambda_{M-1}$ that minimize

$$\sum_{k=1}^{K_H} \left(\hat{V} + \sum_{m=0}^{M-1} h^{\lambda_m}(y_m^k) \cdot (p_{m+1}(y_{m+1}^k) - p_m(y_m^k)) - v^{\theta_1}(y_M^k) \right)^2 \quad (7)$$

for independent realizations of $(y_m^k)_{m=0}^M$, $k = 1, \dots, K_H$ of $(Y_m)_{m=0}^M$. We train the networks $h^{\lambda_0}, \dots, h^{\lambda_{M-1}}$ together, again using a stochastic gradient descent method. Instead of (7), one could also minimize a different deviation measure. However, (7) has the advantage that it yields hedging strategies with an average hedging error close to zero¹⁰.

Once $\lambda_0, \dots, \lambda_{M-1}$ have been determined, we assess the quality of the hedge by simulating new¹¹ independent realizations $(y_m^k)_{m=0}^M$, $k = K_H + 1, \dots, K_H + K_E$ of $(Y_m)_{m=0}^M$ and calculating the average hedging error

$$\frac{1}{K_E} \sum_{k=K_H+1}^{K_H+K_E} \left(\hat{V} + \sum_{m=0}^{M-1} h^{\lambda_m}(y_m^k) \cdot (p_{m+1}(y_{m+1}^k) - p_m(y_m^k)) - v^{\theta_1}(y_M^k) \right) \quad (8)$$

and the empirical hedging shortfall

$$\frac{1}{K_E} \sum_{k=K_H+1}^{K_H+K_E} \left(\hat{V} + \sum_{m=0}^{M-1} h^{\lambda_m}(y_m^k) \cdot (p_{m+1}(y_{m+1}^k) - p_m(y_m^k)) - v^{\theta_1}(y_M^k) \right)^- \quad (9)$$

over the time interval $[0, t_1]$.

4.2. Hedging Until the Exercise Time

Alternatively, one can precompute the whole hedging strategy from time 0 to T and then use it until the option is exercised. In order to do that we introduce the functions

$$v^{\theta_n}(x) := g(n, x) \vee c^{\theta_n}(x), \quad C^{\theta_n}(x) := 0 \vee c^{\theta_n}(x), \quad x \in \mathbb{R}^d,$$

and hedge the difference $v^{\theta_n}(Y_{nM}) - C^{\theta_{n-1}}(Y_{(n-1)M})$ on each of the time intervals $[t_{n-1}, t_n]$, $n = 1, \dots, N$, separately. v^{θ_n} describes the approximate value of the option at time t_n if it has not been exercised before, and the definition of C^{θ_n} takes into account that the true continuation values are non-negative due to the non-negativity of the payoff function g . The hedging strategy can be computed as in Section 4.1, except that we now have to simulate complete paths $(y_m^k)_{m=0}^{NM}$ of $(Y_m)_{m=0}^{NM}$, $k = 1, \dots, K_H$, and then for all $n = 1, \dots, N$, find parameters $\lambda_{(n-1)M}, \dots, \lambda_{nM-1}$ which minimize

$$\sum_{k=1}^{K_H} \left(C^{\theta_{n-1}}(y_{(n-1)M}^k) + \sum_{m=(n-1)M}^{nM-1} h^{\lambda_m}(y_m^k) \cdot (p_{m+1}(y_{m+1}^k) - p_m(y_m^k)) - v^{\theta_n}(y_{nM}^k) \right)^2.$$

Once the hedging strategy has been trained, we simulate independent samples $(y_m^k)_{m=0}^{NM}$, $k = K_H + 1, \dots, K_H + K_E$, of $(Y_m)_{m=0}^{NM}$ and denote the realization of τ^Θ along each sample path $(y_m^k)_{m=0}^{NM}$ by τ^k . The corresponding average hedging error is given by

¹⁰ see Table 2 and Figure 1 below.

¹¹ independent of $(y_m^k)_{m=0}^M$, $k = 1, \dots, K_H$.

$$\frac{1}{K_E} \sum_{k=K_H+1}^{K_H+K_E} \left(\hat{V} + \sum_{m=0}^{\tau^k M-1} h^{\lambda_m}(y_m^k) \cdot (p_{m+1}(y_{m+1}^k) - p_m(y_m^k)) - g(\tau^k, X_{\tau^k}) \right) \quad (10)$$

and the empirical hedging shortfall by

$$\frac{1}{K_E} \sum_{k=K_H+1}^{K_H+K_E} \left(\hat{V} + \sum_{m=0}^{\tau^k M-1} h^{\lambda_m}(y_m^k) \cdot (p_{m+1}(y_{m+1}^k) - p_m(y_m^k)) - g(\tau^k, X_{\tau^k}) \right)^-. \quad (11)$$

5. Example

In this section we study¹² a Bermudan max-call option¹³ on d financial securities with risk-neutral price dynamics

$$S_t^i = s_0^i \exp\left([r - \delta_i - \sigma_i^2/2]t + \sigma_i W_t^i\right), \quad i = 1, 2, \dots, d,$$

for a risk-free interest rate $r \in \mathbb{R}$, initial values $s_0^i \in (0, \infty)$, dividend yields $\delta_i \in [0, \infty)$, volatilities $\sigma_i \in (0, \infty)$ and a d -dimensional Brownian motion W with constant instantaneous correlations¹⁴ $\rho_{ij} \in \mathbb{R}$ between different components W^i and W^j . The option has time- t payoff $(\max_{1 \leq i \leq d} S_t^i - K)^+$ for a strike price $K \in [0, \infty)$ and can be exercised at one of finitely many times $0 = t_0 < t_1 < \dots < t_N = T$. In addition, we suppose there is a savings account where money can be deposited and borrowed at rate r .

For notational simplicity, we assume in the following that $t_n = nT/N$ for $n = 0, 1, \dots, N$, and all assets have the same¹⁵ characteristics; that is, $s_0^i = s_0$, $\delta_i = \delta$ and $\sigma_i = \sigma$ for all $i = 1, \dots, d$.

5.1. Pricing Results

Let us denote $X_n = S_{t_n}$, $n = 0, 1, \dots, N$. Then the price of the option is given by

$$\sup_{\tau} \mathbb{E} \left[e^{-r \frac{\tau T}{N}} \left(\max_{1 \leq i \leq d} X_{\tau}^i - K \right)^+ \right],$$

where the supremum is over all stopping times $\tau: \Omega \rightarrow \{0, 1, \dots, N\}$ with respect to the filtration generated by $(X_n)_{n=0}^N$. The option payoff does not carry any information not already contained in X_n . However, the training of the continuation values worked more efficiently when we used it as an additional feature. So instead of X_n we simulated the extended state process $\hat{X}_n = (X_n^1, \dots, X_n^d, X_n^{d+1})$ for

$$X_n^{d+1} = e^{-r \frac{nT}{N}} \left(\max_{1 \leq i \leq d} X_n^i - K \right)^+$$

¹² The computations were performed on a NVIDIA GeForce RTX 2080 Ti GPU. The underlying system was an AMD Ryzen 9 3950X CPU with 64 GB DDR4 memory running Tensorflow 2.1 on Ubuntu 19.10.

¹³ Bermudan max-call options are a benchmark example in the literature on numerical methods for high-dimensional American-style options; see, e.g., Longstaff and Schwartz (2001); Rogers (2002); García (2003); Broadie and Glasserman (2004); Haugh and Kogan (2004); Broadie and Cao (2008); Berridge and Schumacher (2008); Jain and Oosterlee (2015); Becker et al. (2019a, 2019b).

¹⁴ That is, $\mathbb{E}[(W_t^i - W_s^i)(W_t^j - W_s^j)] = \rho_{ij}(t - s)$ for all $i \neq j$ and $s < t$.

¹⁵ Simulation based methods work for any price dynamics that can efficiently be simulated. Prices of max-call options on underlying assets with different price dynamics were calculated in Broadie and Cao (2008) and Becker et al. (2019a).

to train the continuation value functions c^{θ_n} , $n = 1, \dots, N - 1$. The network $c^{\theta}: \mathbb{R}^{d+1} \rightarrow \mathbb{R}$ was chosen of the form (4) with depth $I = 3$ (two hidden layers), $d + 50$ nodes in each hidden layer and activation function $\varphi = \tanh$. For training we used stochastic gradient descent with mini-batches of size 8192 and batch normalization (Ioffe and Szegedy 2015). At time $N - 1$ we used Xavier (Glorot and Bengio 2010) initialization and performed 6000 Adam (Kingma and Ba 2015) updating steps¹⁶. For $n \leq N - 2$, we started the gradient descent from the trained network parameters θ_{n+1} and made 3500 Adam updating steps¹⁶. To calculate \hat{L} we simulated $K_L = 4,096,000$ paths of $(X_n)_{n=0}^N$. For \hat{U} we generated $K_U = 2048$ outer and 2048×2048 inner simulations.

Our results for \hat{L} , \hat{U} , \hat{V} and 95% confidence intervals for different specifications of the model parameters are reported in Table 1. To achieve a pricing accuracy comparable to the more direct methods of Becker et al. (2019a, 2019b), the networks used in the construction of the candidate optimal stopping strategy had to be trained for a longer time. But in exchange, the approach yields approximate continuation values that can be used to break down the hedging problem into a series of smaller problems.

5.2. Hedging Results

Suppose the hedging portfolio can be rebalanced at the times $u_m = mT/(NM)$, $m = 0, 1, \dots, NM$, for a positive integer M . We assume dividends paid by shares of S^i held in the hedging portfolio are continuously reinvested in S^i . This results in the adjusted discounted security prices

$$P_{u_m}^i = s_0 \exp\left(\sigma W_{u_m}^i - \sigma^2 u_m / 2\right), \quad m = 0, 1, \dots, NM.$$

We set $Y_m^i = P_{u_m}^i$. To learn the hedging strategy, we trained neural networks $h^{\lambda_m}: \mathbb{R}^d \rightarrow \mathbb{R}^d$, $m = 0, \dots, NM - 1$, of the form (4) with depth $I = 3$ (two hidden layers), $d + 50$ nodes in each hidden layer and activation function $\varphi = \tanh$. As in Section 5.1, we used stochastic gradient descent with mini-batches of size 8192 and batch normalization (Ioffe and Szegedy 2015). For $m = 0, \dots, M - 1$, we initialized the networks according to Xavier (Glorot and Bengio 2010) and performed 10,000 Adam (Kingma and Ba 2015) updating steps¹⁶, whereas for $m \geq M$, we started the gradient trajectories from the trained network parameters λ_{m-M} and made 3000 Adam updating steps¹⁶.

Table 2 reports the average hedging errors (8) and (10) together with the empirical hedging shortfalls (9) and (11) for different numbers M of rebalancing times between two consecutive exercise dates t_{n-1} and t_n . They were computed using $K_E = 4,096,000$ simulations of $(Y_m)_{m=0}^{NM}$.

Figure 1 shows histograms of the total hedging errors

$$\hat{V} + \sum_{m=0}^{\tau^k M-1} h^{\lambda_m}(y_m^k) \cdot \left(p_{m+1}(y_{m+1}^k) - p_m(y_m^k)\right) - g(\tau^k, X_{\tau^k}), \quad k = K_H + 1, \dots, K_E,$$

for $d \in \{5, 10\}$ and $M \in \{12, 96\}$.

¹⁶ The hyperparameters $\beta_1, \beta_2, \varepsilon$ were chosen as in Kingma and Ba (2015). The stepsize α was specified as $10^{-1}, 10^{-2}, 10^{-3}$ and 10^{-4} according to a deterministic schedule.

Table 1. Price estimates for max-call options on 5 and 10 symmetric assets for parameter values of $r = 5\%$, $\delta = 10\%$, $\sigma = 20\%$, $\rho = 0$, $K = 100$, $T = 3$, $N = 9$. t_L is the number of seconds it took to train τ^Θ and compute \hat{L} . t_U is the computation time for \hat{U} in seconds. 95% CI is the 95% confidence interval (6). The last column lists the 95% confidence intervals computed in Becker et al. (2019a).

d	s_0	\hat{L}	t_L	\hat{U}	t_U	Point Est.	95% CI	DOS 95% CI
5	90	16.644	132	16.648	8	16.646	[16.628, 16.664]	[16.633, 16.648]
5	100	26.156	134	26.152	8	26.154	[26.138, 26.171]	[26.138, 26.174]
5	110	36.780	133	36.796	8	36.788	[36.758, 36.818]	[36.745, 36.789]
10	90	26.277	136	26.283	8	26.280	[26.259, 26.302]	[26.189, 26.289]
10	100	38.355	136	38.378	7	38.367	[38.335, 38.399]	[38.300, 38.367]
10	110	50.869	135	50.932	8	50.900	[50.846, 50.957]	[50.834, 50.937]

Table 2. Average hedging errors and empirical hedging shortfalls for 5 and 10 underlying assets and different numbers M of reheding times between consecutive exercise times t_{n-1} and t_n . The values of the parameters r , δ , σ , ρ , K , T and N were chosen as in Table 1. IHE is the intermediate average hedging error (8), IHS the intermediate hedging shortfall (9), HE the total average hedging error (10) and HS the total hedging shortfall (11). \hat{V} is our price estimate from Table 1. T1 is the computation time in seconds for training the hedging strategy from time 0 to $t_1 = T/N$. T2 is the number of seconds it took to train the complete hedging strategy from time 0 to T .

d	s_0	M	IHE	IHS	IHS/ \hat{V}	T1	HE	HS	HS/ \hat{V}	T2
5	90	12	0.007	0.190	1.1%	102	−0.001	0.676	4.1%	379
5	90	24	0.007	0.139	0.8%	129	−0.002	0.492	3.0%	473
5	90	48	0.007	0.104	0.6%	234	−0.001	0.367	2.2%	839
5	90	96	0.007	0.081	0.5%	436	−0.001	0.294	1.8%	1546
5	100	12	0.013	0.228	1.4%	102	0.006	0.785	4.7%	407
5	100	24	0.013	0.163	1.0%	131	0.006	0.569	3.4%	512
5	100	48	0.013	0.118	0.7%	252	0.007	0.423	2.5%	931
5	100	96	0.013	0.089	0.5%	470	0.006	0.335	2.0%	1668
5	110	12	0.002	0.268	1.6%	102	−0.012	0.881	5.3%	380
5	110	24	0.002	0.192	1.2%	130	−0.012	0.638	3.8%	511
5	110	48	0.002	0.139	0.8%	262	−0.013	0.474	2.9%	950
5	110	96	0.002	0.105	0.6%	471	−0.010	0.374	2.3%	1673
10	90	12	−0.015	0.192	0.7%	111	−0.010	0.902	3.4%	414
10	90	24	−0.014	0.147	0.6%	145	−0.011	0.704	2.7%	534
10	90	48	−0.015	0.136	0.5%	269	−0.011	0.611	2.3%	958
10	90	96	−0.015	0.121	0.5%	506	−0.012	0.551	2.1%	1792
10	100	12	0.008	0.230	0.9%	111	0.015	1.025	3.9%	414
10	100	24	0.008	0.176	0.7%	152	0.014	0.797	3.0%	531
10	100	48	0.008	0.150	0.6%	271	0.016	0.682	2.6%	978
10	100	96	0.008	0.132	0.5%	512	0.014	0.672	2.6%	1803
10	110	12	−0.029	0.249	1.0%	112	−0.026	1.146	4.4%	410
10	110	24	−0.029	0.189	0.7%	146	−0.027	0.908	3.5%	530
10	110	48	−0.029	0.160	0.6%	269	−0.026	0.782	3.0%	965
10	110	96	−0.029	0.151	0.6%	507	−0.024	0.666	2.5%	1777

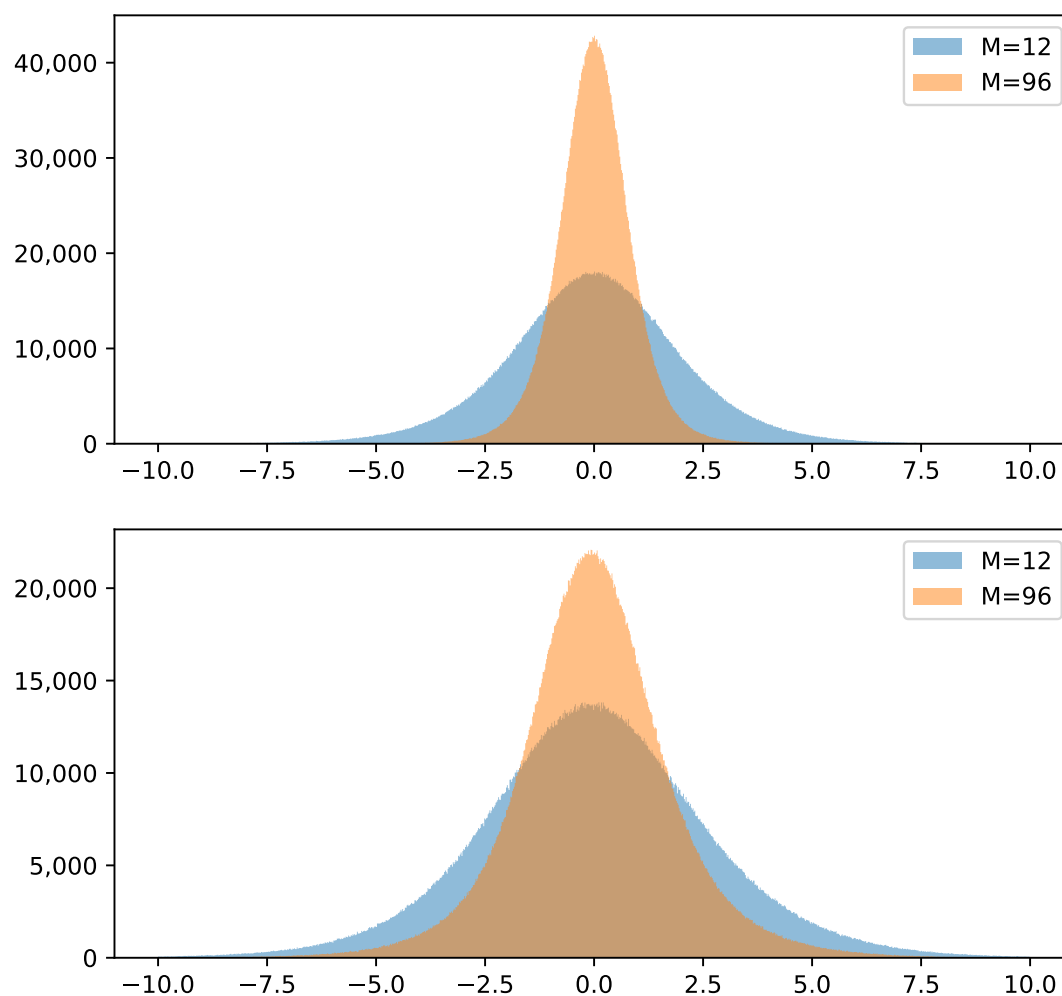


Figure 1. Total hedging errors for $s_0 = 100$, $M \in \{12, 96\}$, $d = 5$ (top) and $d = 10$ (bottom) along 4,096,000 sample paths of $(Y_m)_{m=0}^{NM}$. The values of the parameters $r, \delta, \sigma, \rho, K, T$ and N were as in Tables 1 and 2.

6. Conclusions

In this article, we used deep learning to price and hedge American-style options. In a first step our method employs a neural network version of the Longstaff–Schwartz algorithm to estimate continuation values and derive a candidate optimal stopping rule. The learned stopping rule immediately yields a low-biased estimate of the price. In addition, it can be used to construct an approximate solution of the dual martingale problem of Rogers (2002) and Haugh and Kogan (2004). This gives a high-biased estimate and confidence intervals for the price. To achieve the same pricing accuracy as the more direct approaches of Becker et al. (2019a, 2019b), we had to train the neural network approximations of the continuation values for a longer time. However, computing approximate continuation values has the advantage that they can be used to break the hedging problem into a sequence of subproblems that compute the hedge only from one possible exercise date to the next.

Author Contributions: S.B., P.C. and A.J. have contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Funding: A.J. acknowledges support from the DFG through Germany’s Excellence Strategy EXC 2044-390685587, Mathematics Münster: Dynamics - Geometry - Structure.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Andersen, Leif. 2000. A simple approach to the pricing of Bermudan swaptions in the multifactor LIBOR market model. *The Journal of Computational Finance* 3: 5–32. [\[CrossRef\]](#)
- Bally, Vlad, Gilles Pagès, and Jacques Printems. 2005. A quantization tree method for pricing and hedging multidimensional American options. *Mathematical Finance* 15: 119–68. [\[CrossRef\]](#)
- Barraquand, Jérôme, and Didier Martineau. 1995. Numerical valuation of high dimensional multivariate American securities. *The Journal of Financial and Quantitative Analysis* 30: 383–405. [\[CrossRef\]](#)
- Becker, Sebastian, Patrick Cheridito, and Arnulf Jentzen. 2019a. Deep optimal stopping. *Journal of Machine Learning Research* 20: 1–25.
- Becker, Sebastian, Patrick Cheridito, Arnulf Jentzen, and Timo Welti. 2019b. Solving high-dimensional optimal stopping problems using deep learning. *arXiv* arXiv:1908.01602.
- Berridge, Steffan J., and Johannes M. Schumacher. 2008. An irregular grid approach for pricing high-dimensional American options. *Journal of Computational and Applied Mathematics* 222: 94–111. [\[CrossRef\]](#)
- Bouchard, Bruno, and Xavier Warin. 2012. Monte-Carlo valuation of American options: Facts and new algorithms to improve existing methods. In *Numerical Methods in Finance*. Berlin and Heidelberg: Springer, pp. 215–255.
- Broadie, Mark, and Menghui Cao. 2008. Improved lower and upper bound algorithms for pricing American options by simulation. *Quantitative Finance* 8: 845–61. [\[CrossRef\]](#)
- Broadie, Mark, and Paul Glasserman. 2004. A stochastic mesh method for pricing high-dimensional American options. *Journal of Computational Finance* 7: 35–72. [\[CrossRef\]](#)
- Bru, Bernard, and Henri Heinrich. 1985. Meilleures approximations et médianes conditionnelles. *Annales de l'I.H.P. Probabilités et Statistiques* 21: 197–224.
- Buehler, Hans, Lukas Gonon, Josef Teichmann, and Ben Wood. 2019. Deep hedging. *Quantitative Finance* 19: 1271–91. [\[CrossRef\]](#)
- Carriere, Jacques F. 1996. Valuation of the early-exercise price for options using simulations and nonparametric regression. *Insurance: Mathematics and Economics* 19: 19–30. [\[CrossRef\]](#)
- Chen, Yangang, and Justin W.L. Wan. 2019. Deep neural network framework based on backward stochastic differential equations for pricing and hedging American options in high dimensions. *arXiv* arXiv:1909.11532.
- Egloff, Daniel, Michael Kohler and Nebojsa Todorovic. 2007. A dynamic look-ahead Monte Carlo algorithm for pricing Bermudan options. *Annals of Applied Probability* 17: 1138–71. [\[CrossRef\]](#)
- Forsyth, Peter A., and Ken R. Vetzal. 2002. Quadratic convergence for valuing American options using a penalty method. *SIAM Journal on Scientific Computing* 23: 2095–122. [\[CrossRef\]](#)
- García, Diego. 2003. Convergence and biases of Monte Carlo estimates of American option prices using a parametric exercise rule. *Journal of Economic Dynamics and Control* 27: 1855–79. [\[CrossRef\]](#)
- Glorot, Xavier, and Yoshua Bengio. 2010. Understanding the difficulty of training deep feedforward neural networks. Paper Presented at Thirteenth International Conference on Artificial Intelligence and Statistics, PMLR, Sardinia, Italy, May 13–15; vol. 9, pp. 249–256.
- Han, Jiequn, Arnulf Jentzen, and Weinan E. 2018. Solving high-dimensional partial differential equations using deep learning. *Proceedings of the National Academy of Sciences of the United States of America* 115: 8505–10. [\[CrossRef\]](#)
- Haugh, Martin B., and Leonid Kogan. 2004. Pricing American options: a duality approach. *Operations Research* 52: 258–70. [\[CrossRef\]](#)
- Hull, John C. 2003. *Options, Futures and Other Derivatives*. London: Pearson; Upper Saddle River: Prentice Hall.
- Ioffe, Sergey, and Christian Szegedy. 2015. Batch normalization: accelerating deep network training by reducing internal covariate shift. Paper presented at 32nd International Conference on Machine Learning, ICML 2015, Lille, France, July 6–11; vol. 37, pp. 448–456.
- Jain, Shashi, and Cornelis W. Oosterlee. 2015. The stochastic grid bundling method: efficient pricing of Bermudan options and their Greeks. *Applied Mathematics and Computation* 269: 412–31. [\[CrossRef\]](#)
- Kingma, Diederik P., and Jimmy Ba. 2015. Adam: A method for stochastic optimization. Paper Presented at International Conference on Learning Representations, San Diego, CA, USA, May 7–9.
- Kohler, Michael, Adam Krzyżak, and Nebojsa Todorovic. 2010. Pricing of high-dimensional American options by neural networks. *Mathematical Finance* 20: 383–410. [\[CrossRef\]](#)

- Kolodko, Anastasia, and John Schoenmakers. 2006. Iterative construction of the optimal Bermudan stopping time. *Finance and Stochastics* 10: 27–49. [[CrossRef](#)]
- Lapeyre, Bernard, and Jérôme Lelong. 2019. Neural network regression for Bermudan option pricing. *arXiv* arXiv:1907.06474.
- Longstaff, Francis A., and Eduardo S. Schwartz. 2001. Valuing American options by simulation: a simple least-squares approach. *The Review of Financial Studies* 14: 113–47. [[CrossRef](#)]
- Reisinger, Christoph, and Jan H. Witte. 2012. On the use of policy iteration as an easy way of pricing American options. *SIAM J. Financial Math.* 3: 459–78. [[CrossRef](#)]
- Rogers, Chris. 2002. Monte Carlo valuation of American options. *Mathematical Finance* 12: 271–86. [[CrossRef](#)]
- Rogers, Chris. 2010. Dual valuation and hedging of Bermudan options. *SIAM Journal on Financial Mathematics* 1: 604–8. [[CrossRef](#)]
- Sirignano, Justin, and Konstantinos Spiliopoulos. 2018. DGM: A deep learning algorithm for solving partial differential equations. *Journal of Computational Physics* 375: 1339–64. [[CrossRef](#)]
- Tilley, James A. 1993. Valuing American options in a path simulation model. *Transactions of the Society of Actuaries* 45: 83–104.
- Tsitsiklis, John N., and Benjamin Van Roy. 2001. Regression methods for pricing complex American-style options. *IEEE Transactions on Neural Networks* 12: 694–703. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).