AMSI SUMMER RESEARCH SCHOLARSHIPS 2024–25

Get a taste for Research this Summer

Implementing a numerical scheme for pricing of American options under model uncertainty

Billy Bourdaniotis

Supervised by Dr. Libo Li

University of New South Wales

Jane Street

AMSI Australian Mathematical Sciences Institute

Abstract

This report aims to construct and analyse the results of a numerical scheme for pricing American options using simulations. American options pricing is a well-studied topic, and this report will look at previous literature and synthesise concepts to derive a data-driven algorithm. The results obtained suggest that smoothening the indicator function used in the backward recursive relationship in pricing can change the variance of the point estimate significantly, and could be beneficial in reducing model risk in most cases.

Statement of Authorship

I confirm that this work is my original creation, except where explicit references have been made. All sources of information, ideas, and content used in the preparation of this document have been properly cited. I am grateful for my supervisor Dr. Libo Li for his suggestions and proofreading this report. Special thanks to Daniel Chee for his advice, clarification and guidance throughout the project.

Contents

1 Introduction

An American option gives the holder the right, with no obligation, to exercise the option and receive its payoff at any time before and including the termination date. American options are among the most complex financial instruments to price mathematically. To value American options, one needs to solve an optimal stopping problem. That is, for a sequence of random variables that can only be observed in order, which of the following actions is optimal: Accept the current observation based on past information and cannot resume the process, or reject the current observation permanently and go on to the next. The problem is given by the optimal stopping representation

$$V_n = \operatorname*{ess\,sup}_{n \leq \tau \leq N} E_{\mathbb{Q}}[P_\tau | \mathcal{F}_n],$$

where P some random payoff process. It is well-known that the price of an American option has no closed form solution and thus numerical methods are used to approximate a value. Recent research in this area has applied reinforcement learning methods to obtain estimates. In the work of Becker et al. (2019) [1] the authors propose a particular form of stopping times which can be approximated by training neural networks. The trained networks are used to estimate the functional form of the value of an option via Monte Carlo sampling reminiscent of Longstaff-Schwarz [4]. Another approach is explored in Dong (2023)[3] which applies entropy regularization to a PDE/HJB formulation of the problem. This casts the problem in a reinforcement learning framework whereby an optimal policy is derived. This can be interpreted as the probability of stopping related to the distance of the value from an optimal stopping boundary, which is then used as the basis for training neural networks and estimating an option value. This report will describe a data-driven numerical scheme to price American options using a Monte-Carlo algorithm leveraging neural networks, and applying regularisation techniques to reduce model risk. This report will outline a framework for the problem in Section 2 along with historical models. Section 3 will present two regularisation models and the general algorithm to price options, then Section 4 will discuss some examples and results. The main focus was on how smoothing the recursive policy algorithm can reduce model risk and applications to higher dimension problems.

2 Background

2.1 Optimal stopping problems, Snell envelopes and Superreplication

The payoff that the holder of an American option can receive is dependent on the state of an underlying asset process $S = (S_n)_{n=0}^N$. Simple payoff structures include calls, and puts,

$$P_\tau = \begin{cases} (S_\tau - K)^+, \text{ if call} \\ (K - S_\tau)^+, \text{ if put} \end{cases}$$

where $(x)^+ = \max(x, 0)$, K is the strike price and τ is exercise time.

Definition 2.1. A *random variable* X is a function that maps outcomes in a sample space Ω to $\mathbb{R}$ i.e. X maps $\omega \in \Omega$ to $X(\omega) \in \mathbb{R}$. A sequence of random variables $(X_n)_{n=0}^N$ is a *stochastic process.*

Definition 2.2. A a σ-algebra, $\mathcal{H}$, of Ω is a collection of subsets containing elements from Ω such that $\Omega, \emptyset \in \mathcal{H}$ and for any set of events $A_1, A_2, ... \in \mathcal{H}$, then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{H}$, or it is closed under countable union.

Definition 2.3. The natural filtration $(\mathcal{F}_n)_{n=0}^N$ is a sequence of increasing σ-algebras, meaning $\mathcal{F}_n \subset \mathcal{F}_m \subset \mathcal{F}$ for all $m > n$.

This is interpreted the information available to market participants up to a certain time, subject to the usual condition that $\bigcap_{k \geq n} \mathcal{F}_k = \mathcal{F}_n$, so one cannot gain an advantage by peeking between time steps in a discrete time setting.

Definition 2.4. A probability measure $\mathbb{P} : \mathcal{F} \to [0, 1]$ is a function that maps events in $\mathcal{F}$ to $[0, 1]$ such that $P(\Omega) = 1$, $P(\emptyset) = 0$ and $P(\cup_n A_n) = \sum_n P(A_n)$ for any disjoint sets $A_1, A_2, ...A_n \in \mathcal{F}$ i.e. $A_i \cap A_j = \emptyset, i \neq j$.

This report assumes a finite-discrete time horizon $\{0, 1, ...N\}$ and complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω represents the set of all spatial possibilities in the time horizon, with individual possible paths ω, and filtration $(\mathcal{F}_n)_{n=0}^N$.

Definition 2.5. A random variable X is *adapted*, or $\mathcal{F}$*-measurable* if for every $x \in \mathbb{R}$, the preimage of the interval $(-\infty, x]$ can be traced to an element in σ-algebra $\mathcal{F}$, or $X^{-1}((-\infty, x]) = \{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F}$.

Definition 2.6. Let X be $\mathcal{F}$-measurable random variable from Ω to $\mathbb{R}$, equipped with measure $\mathbb{P}$, then the *expected value function* is defined by

$$E_{\mathbb{P}}[X] = \int_{\Omega} X(\omega) \, d\mathbb{P}(\omega) = \int_{\mathbb{R}} x f(x) \, dx$$

where $f(x)$ is the probability density function of X.

Definition 2.7. The *conditional expectation* of X under sub σ-algebra $\mathcal{G} \subset F$ is a $\mathcal{G}$-measurable random variable $E_{\mathbb{P}}[X|\mathcal{G}] : \Omega \to \mathbb{R}$ such that for all events $G \in \mathcal{G}$

$$\int_G E_{\mathbb{P}}[X|\mathcal{G}](\omega) \, d\mathbb{P}(\omega) = \int_G X(w) \, d\mathbb{P}(\omega).$$

It framed as taking the expected value of X with partial information such that the random variable is different, has the same expected value, but not completely known.

Definition 2.8. We say that X is *integrable* if $E[|X|]$ is finite.

The *Snell Envelope* is the framework for problem-setting as suggested by Bensoussan (1984) [2], as it captures the optimal stopping valuation of the American option to be solved. We assume that there exists a probability measure $\mathbb{Q}$ (known as the arbitrage-free measure) that makes the discounted stock price $\frac{S_n}{B_n}$, a *martingale* (see Definition 2.10), according to an annual risk-free rate of continuous interest r, and $B_n = B_0 e^{rn}$.

Definition 2.9. The *Snell Envelope* $V = (V_n)_{n=0}^N$ which represents the value of the option, with respect to the measure $\mathbb{Q}$ of the payoff process $P = (P_n)_{n=0}^N$ is

$$\begin{cases} V_N = P_N \\ V_n = \max(P_n, E_{\mathbb{Q}}[V_{n+1}|\mathcal{F}_n]), \quad n \leq N-1 \end{cases}$$

Firstly, the terminal option value is the payoff at expiry because the holder of the option has no choice but to receive its payoff at time N if the option has survived until then. Any time before expiry, the maximum function is conveying that the value gained by holding the option depends on the stopping decision. The holder will exercise the option and receive P_n if it is greater than the expected value of keeping the option alive. This expected value is defined later as the continuation value and its calculation is one of the main challenges when pricing American options.

Definition 2.10. A *martingale* M, is an adapted and integrable stochastic process ($E[|M_n||\mathcal{F}_n] < \infty$) such that $E[M_m|\mathcal{F}_n] = M_n$ for all $m \geq n$. A *supermartingale* satisfies $E[M_m|\mathcal{F}_n] \leq M_n$ for all $m \geq n$.

Theorem 2.11. *The Snell envelope of P given by $V_n = \operatorname{ess\,sup}_{n \leq \tau \leq N} E_{\mathbb{Q}}[P_\tau|\mathcal{F}_n]$ is the smallest supermartingale that dominates P.*

Proof. Firstly, $V_n \geq E_{\mathbb{Q}}[V_{n+1}|\mathcal{F}_n]$ meaning it is a supermartingale, and $V_n \geq P_n$, thus V dominates P. Suppose U_n is some arbitrary super-martingale that dominates P. Then at time N, $U_N \geq P_N$. Assume that $U_N \geq V_N$. Then $U_{N-1} \geq E_{\mathbb{Q}}[U_N|\mathcal{F}_{N-1}] \geq E_{\mathbb{Q}}[V_N|\mathcal{F}_{N-1}]$ using the supermartingale property and induction assumption. Furthermore, $U_n \geq P_n$ since U dominates P. Hence $U_{N-1} \geq \max(P_{N-1}, E_{\mathbb{Q}}[V_N|\mathcal{F}_{N-1}]) = V_{N-1}$. Through backwards induction, any dominating super martingale of P also dominates V. □

Theorem 2.12. *For the process P, which is guided by the Snell Envelope V, largest optimal stopping time is the first instance when $V_\tau = P_\tau$ or N if this never happens.*

$$\tau = \min\{j \in \{1, ..., N\} : P_j = V_j\}.$$

Proof. The following proof is restated from a Yale handout on Optimal Stopping by David Pollard. [5]
The goal is to show, $E[P_\tau] \geq E[P_j]$ for all stopping times j.

Firstly, if $j > \tau$, then $E[P_{\min(\tau,j)} - P_j] = E[P_\tau - P_j] \geq E[V_\tau - V_j]$ since $V \geq P$ by definition and $P_\tau = V_\tau$. Also, $E[V_\tau - V_j] \geq 0$ since V is a supermartingale. Hence, $E[P_{\min(\tau,j)} - P_j] \geq 0$, thus $E[P_{\min(\tau,j)}] \geq E[P_j]$ for all stopping times j. This is an equality if $\tau \geq j$.

Next consider that $E[P_\tau] = E[V_\tau] = E[V_{\min(\tau,\tau)}]$ It can be show that the process $Y_j = V_{\min(\tau,j)}$ is a martingale. Hence $E[P_\tau] = E[V_{\min(\tau,\tau)}] = E[Y_\tau] = E[Y_j] \geq E[P_{\min(\tau,j)}]$ using the fact $V \geq P$, and it was shown previously $E[P_{\min(\tau,j)}] \geq E[P_j]$, so $E[P_\tau] \geq E[P_j]$. □

Using the Snell Envelope representation, if $V_\tau = \max(P_\tau, V_{\tau+}) = P_\tau$, then $P_\tau > V_{\tau+}$. In other words, the optimal stopping rule suggests stopping at the first time where the immediate reward is at least as good as the expected future reward. This is the largest time as it delays stopping as long as possible while still being optimal. What this implies for the long position, the holder of the option, is there is no point in exercising after τ if given the opportunity. This fact is the core idea in backwards recursion algorithms that estimate the continuation value, compare to the current payoff, and make an exercise decision.

Definition 2.13. The *continuation value* is $V_{n+} := E_{\mathbb{Q}}[V_{n+1}|\mathcal{F}_n]$, and represents the expected current reward for continuing the options life. Hence $V_n = \max(V_{n+}, P_n) = V_{n+} + (P_n - V_{n+})^+$ since the option will be exercised if the payoff is greater than the value of the option next period if not exercised.

Theorem 2.14. ***Doobs Decomposition*** *Let $V = (V_n)_{n=0}^N$ be an integrable, adapted supermartingale process. Then V admits a unique decomposition in the form*

$$V = M + A$$

where $M = (M_n)_{n=0}^N$ is a martingale and $A = (A_n)_{n=0}^N$ is an integrable, previsible process, meaning it is $\mathcal{F}_{n-1}$ measurable and $M_0 = A_0 = 0$.

Proof. Define A_n and M_n by:

$$A_n = \sum_{j=1}^{n}(V_{j-1+} - V_{j-1}), \quad M_n = V_0 + \sum_{j=1}^{n}(V_j - V_{j-1+}).$$

Since V_n is adapted, it is $\mathcal{F}_n$ measurable and integrable. Then A_{n+1} and M_n are $\mathcal{F}_n$ measurable. Furthermore, $E[|A_n|] \leq E[\sum |E[V_j|\mathcal{F}_{j-1}]|] + \sum E[|V_{j-1}|] < \infty$ by Triangle inequality and since V is integrable, so A_n is integrable.

Consider the conditional expectation $E[M_n|\mathcal{F}_{n-1}] = E[M_{n-1} + (V_n - E[V_n|\mathcal{F}_{n-1}])|\mathcal{F}_{n-1}] = E[M_{n-1}|\mathcal{F}_{n-1}] = M_{n-1}$. Hence M_n is a martingale. Using a similar argument to A, M can also be shown to be integrable.

Notice that in this unique decomposition, $A_n = \sum_{j=0}^{n-1}(V_{j+} - V_j) = -\sum_{j=0}^{n-1}(P_j - V_{j+})^+$. Hence, the Doobs Decomposition can be expressed as $V_n = M_n - \sum_{j=0}^{n-1}(P_j - V_{j+})^+$. □

The process $-A_n$ is non-decreasing and provides an indicator for when to exercise the option. Note that $(P_n - V_{n+})^+ = (P_n - V_{n+})\mathbb{1}_{\{P_n - V_{n+} > 0\}}$, so one would exercise at n if $P_n - V_{n+} > 0$, or the payoff granted exceeds the continuation option value. Hence the largest optimal stopping time can now be reframed as

$$\tau = \begin{cases} N, A_N = 0 \\ \min\{j \in \{1, ...N-1\} : A_{j+1} < 0\}, A_N < 0. \end{cases}$$

Theorem 2.15. *There exists a Martingale M, such that V and M satisfy*

$$V_n = P_N - (M_N - M_n) + \sum_{j=n}^{N-1}(P_j - V_{j+})^+.$$

Proof. From the Doobs Decomposition of V (Theorem 2.14) we have that there exists a unique Martingale M such that

$$V_n = M_n - \sum_{j=0}^{n-1}(P_j - V_{j+})^+, \quad \text{and} \quad V_N = P_N = M_N - \sum_{j=0}^{N-1}(P_j - V_{j+})^+.$$

Subtracting the first equation from the second yields $V_n - P_N = M_n - M_N + \sum_{j=n}^{N-1}(P_j - V_{j+})^+$. Hence, there exists a unique M that satisfies, $V_n = P_N - (M_N - M_n) + \sum_{j=n}^{N-1}(P_j - V_{j+})^+$. □

This representation of the option value dynamics is very versatile. It can be manipulated to bring different ways in evaluating the option price, and will be useful in proving results about these option values.

One of the purposes of holding options is to mitigate investment risks and uncertainty. The seller of the option may offset their position by investing in the underlying asset and a risk free asset in order to '*hedge*' their risk in most scenarios. Hedging refers to constructing a portfolio using a combination of assets related to the option to mimic its payoff at any time. In the case of American options, this is not enough.

Definition 2.16. *Self-Financing* is a term for an investment portfolio that does not require any additional capital beyond an initial investment of ϕ_1 worth of stock and ψ_1 worth of bonds at time 0. Mathematically, this means $\Delta V_n = \phi_{n-1}\Delta S_n + \psi_{n-1}\Delta B_n$, where Δ is the difference operator, $\Delta Y_n = Y_n - Y_{n-1}$.

Definition 2.17. A *Super-replicating portfolio* of P is self-financing strategy (ϕ, ψ) with initial investment x if for all $n \in \{0, ..., N\}$,

$$V_n = x + \sum_{k=1}^{n} \phi_k \Delta S_k + \sum_{k=1}^{n} \psi_k \Delta B_k \geq P_n$$

and for the discounted processes, V_n^*, P_n^*,

$$V_n^* = x + \sum_{k=1}^{n} \phi_k \Delta S_k^* \geq P_n^*.$$

Thus if π is the true price of the American object, then $\pi \leq x$. In other words, we construct a portfolio with value exceeding the option payoff at any time.

Definition 2.18. The *super-replicating price* of P is $\bar{\pi} = \inf\{x \in \mathbb{R} : \exists(\phi, \psi)$ that super-replicates $P\}$. Thus, $\bar{\pi}$ is the smallest amount needed to begin a superhedging portfolio that the seller of the option can use to service the payoff at any stopping time, optimal or not.

In the case of American options, the short position needs to superhedge (self-financing and super-replicating portfolio) because if the option is stopped optimally, an exact hedge may not be enough to cover the payoff, so the investor needs need to increase funds to cover shortfall between hedge and obligation to the holder. If the option is not stopped optimally, a part of the superhedge is consumed.

Theorem 2.19. *If the market is arbitrage-free and complete, meaning a unique risk-neutral martingale $\mathbb{Q}$ exists, then $\bar{\pi} = \operatorname{ess\,sup}_{\tau \leq N} E_{\mathbb{Q}}[P_\tau^*]$ where P^* is the discounted payoff process.*

This ensures the seller of the option, holds sufficient capital to hedge against the worst possible scenario of an optimal stopping by the option holder.

2.2 Existing Models

2.2.1 Binomial Model

The foundation was set by Cox, Ross and Rubinstein's Multinomial Tree Model where the underlying asset process has a chance to move constant fixed percentages at discrete intervals. It's simplicity was effective for pricing American options but the computational requirements increase exponentially with more time steps and higher dimensionality.

The Binomial Model assumes the option can be exercised at discrete times with intervals δt, so $t_{n+1} = t_n + \delta t$, and constant risk free rate r over this interval. The price of the underlying at time t_{n+1}, S_{n+1} is either $u_n S_n$ or $d_n S_n$ where $d_n < 1 \leq u_n$ for all n. The price of the American option can be expressed similar to the Snell Envelope where V_n is the value and P_n is the intrinsic value of the option at time t_n. The backwards recursive formula for the value of a call option where the underlying has moved up j times is

$$\begin{cases} V_{j,N} = (u^j d^{N-j} S_0 - K)^+ \\ V_{j,n} = \max((u^j d^{n-j} S_0 - K)^+, e^{-r\delta t} E_{\mathbb{Q}}[V_{n+1} | \mathcal{F}_{j,n}]), \quad n \leq N-1 \end{cases}$$

where $E_{\mathbb{Q}}[V_{j,n+1} | \mathcal{F}_n] = q_n V_{j+1,n+1} + (1-q_n) V_{j,n+1}$, and $q_n = \frac{e^r - d_n}{u_n - d_n}$.

2.2.2 Longstaff-Schwarz

Longstaff & Schwartz (2001) [4] take a least-squares Monte-Carlo approach to obtain an optimal exercise strategy using the continuation value. Sample paths train a model by regressing the next period discounted cashflow (Y) onto the current stock price (X) using a Laguerre polynomial basis in order to estimate the continuation value $(E[Y|X])$. This is recursively applied backwards to determine the optimal stopping time, and the discounted cashflow of each path collated to obtain a Monte-Carlo estimate of the option price, and compare their estimate to the finite difference method. Sample asset paths are generated using Geometric Brownian Motion described by the stochastic differential equation $dS_t = rS_t\,dt + vS_t\,dW_t$, with closed form $S_t = S_0 e^{(r - \frac{1}{2}\sigma^2)t + vW_t}$, where W_t is a Brownian Motion stochastic process. However, these sample paths are used for illustrative purposes, and this model can handle other sample forecasts in one dimension.

The intuition behind this approach is that a rational holder of an American option will try to maximise the value of their option by comparing the immediate exercise payoff with the expected discounted payoff from continuing the option's life. Hence, the task reduces to ways to estimate the continuation value using current or past data. The algorithm is as follows [4]:

1. Generate sample M paths for times $0, 1, ..., N$ (possibly under Geometric Brownian Motion).
2. Calculate time N payoffs, discount by one period to obtain a vector of discounted payoffs at time $N-1$, P_{N-1}.
3. Repeat the following for $n \in \{N-1, ...1\}$.

- Let X_n be the vector of stock prices at time n for ITM paths.
- Let Y_n be the vector of values from P_n for ITM paths.
- Regress Y_n onto X_n using a suitable basis to obtain the continuation value function $F(X_n) = E[Y_n|X_n]$ in terms of X_n.
- For path m, compare $F(x_{n,m})$ with the immediate payoff for that path. If immediate payoff is greater, replace the mth value in P_n with the current payoff.
- Discount the values of P_n by a single period and to obtain P_{n-1}.

4. Take the mean of P_0.

Furthermore, Step 3 assumes $E[X|Y]$ can be expressed as a countable basis of measurable L^2 basis functions, which is justifiable since L^2 is a Hilbert space, hence there exists a countable orthonormal basis of L^2. Thus, approximating $E[Y|X]$ involves a projection of the true functional form of discounted payoff onto some basis. By focusing on ITM options, the required estimation space is smaller with fewer basis function needed. By taking N (number of time steps) to be sufficiently large, the value of the option converges to that of an American option, and more sample paths results in the estimated regression to converge in mean-squared metric.

By proposition 1 of [4] V_0 is the true value of the American option estimated through this method, then

$$V_0 \geq \lim_{M\to\infty} \frac{1}{M} \sum_{i=1}^{M} D(P_{\tau_i}, 0, \tau_i).$$

The intuition is that the value of the American option is the maximum value attainable using any stopping strategy, so this stopping rule should converge to a value less than the value defined under Snell Envelope.

3 The Proposed Models

Similar to Longstaff-Schwarz, this model is data-driven, meaning any forecast of the underlying asset process can be used. The remaining pricing aspect is the method of estimating the continuation values to make a stopping decision. This will be handled by neural networks and optimisation techniques that require smooth loss. The deterministic stopping decision relies on the use of indicator functions which are discontinuous, so we shall remove this need.

3.1 Gaussian Model

For the first proposed model, instead of a binary stopping decision at each point of time, the indicator is smoothed by an auxiliary function that approximates the exact solution under limiting conditions. In this case, the Gaussian function $\Phi_\sigma(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-t^2}{2\sigma^2}}\, dt$ was chosen due to its link with the asymptotic distribution of Monte-Carlo samples. The smoothness enables optimisation and the existence of optimal network parameters. The additional flexibility in the function of choice adds randomness to the optimal stopping decision. The new stopping decision can be interpreted as being determined by a biased coin flip depending on the difference from the deterministic optimal stopping boundary ($V_{j+} = P_j$). We expect to see a change in sample variance and sensitivity to input parameters as new possible values in the stopping decision are allowed.

The backward stochastic relationship for the value process is

$$V_n = P_N + (M_N - M_n) + \sum_{j=n}^{N-1} (P_j - V_{j+}) \mathbb{1}_{\{P_j - V_{j+} > 0\}},$$

where the indicator function $\mathbb{1}_{\{x>0\}} = 1$ if $x > 0$, and $\mathbb{1}_{\{x>0\}} = 0$ if $x \leq 0$.

Consider the value process V^σ which satisfies

$$V_n^\sigma = P_N - (M_N^\sigma - M_n^\sigma) + \sum_{j=n}^{N-1} (P_j - V_{j+}^\sigma) \Phi_\sigma (P_j - V_{j+}^\sigma).$$

In essence, the indicator function is smoothly approximated by $\Phi_\sigma(x)$. As a heuristic argument to justify this choice, as $\sigma \to 0$, V^σ will approach V since $\Phi_\sigma(x) \to \mathbb{1}_{\{x>0\}}$. Many other sequence of continuous functions satisfy this convergence, but we use $\Phi_\sigma(x)$ for demonstration.

The first result we observe is that this method produces undervaluations of the theoretical option value at all dates before expiry.

Theorem 3.1. *For all $\sigma > 0$, then $V_n^\sigma < V_n$ for all $n < N$.*

Proof. Define $\xi_\sigma(x) := x\Phi_\sigma(x)$. For $x < 0$, $\Phi_\sigma(x) > 0$, so $\xi_\sigma(x) < 0$. For $x \geq 0$, $\Phi_\sigma(x) < 1$, so $\xi_\sigma(x) < x$. Thus for all x, $\xi_\sigma(x) < \max(x, 0) = x^+$.
It is known that $V_N = V_N^\sigma = P_N$, so, $V_{N-1+}^\sigma = V_{N-1+}$. Hence $V_{N-1}^\sigma = V_{N-1+}^\sigma + \xi_\sigma(P_{N-1} - V_{N-1+}^\sigma) < V_{N-1+}^\sigma + (P_{N-1} - V_{N-1+}^\sigma)^+ = V_{N-1+} + (P_{N-1} - V_{N-1+})^+$, so $V_{N-1}^\sigma < V_{N-1}$.
Using backward induction, assume that $V_k^\sigma < V_k$, so $V_{k-1+}^\sigma = E[V_k^\sigma|\mathcal{F}_{k-1}] < E[V_k|\mathcal{F}_{k-1}] = V_{k-1+}$. Now $V_{k-1}^\sigma = V_{k-1+}^\sigma + \xi_\sigma(P_{k-1} - V_{k-1+}^\sigma) < V_{k-1+}^\sigma + (P_{k-1} - V_{k-1+}^\sigma)^+ =$
$\max(V_{k-1+}^\sigma, P_n) \leq \max(V_{k-1+}, P_n) = V_{k-1}$ due to the inductive assumption and that $\xi_\sigma(x) < x^+$. □

The rate at which this approximation converges to the value in Doobs Decomposition can be studied through establishing an upper bound on the error function $E_\sigma(x) = |\mathbb{1}_{\{x\geq 0\}} - \Phi_\sigma(x)|$.

Theorem 3.2. *The error function $E_\sigma(x) = |\mathbb{1}_{\{x\geq 0\}} - \Phi_\sigma(x)|$ is bounded above by $\frac{\sigma}{|x|\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$ for non-zero x, and $E_\sigma(0) = \frac{1}{2}$.*

Proof. Firstly, $E_\sigma(0) = |1 - \Phi_\sigma(0)| = |1 - \frac{1}{2}| = \frac{1}{2}$.

For $x > 0$, $E_\sigma(x) = |1 - \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} \, dt|$. This is the upper tail Gaussian distribution $\frac{1}{\sigma\sqrt{2\pi}} \int_x^\infty e^{-\frac{t^2}{2\sigma^2}} \, dt \leq \frac{1}{\sigma\sqrt{2\pi}} \int_x^\infty \frac{t}{x} e^{-\frac{t^2}{2\sigma^2}} \, dt = \frac{\sigma}{x\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$. Hence, $E_\sigma(x) \leq \frac{\sigma}{x\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$ for $x > 0$. Further, $E_\sigma(x) = E_\sigma(-x)$, thus $E_\sigma(x) \leq \frac{\sigma}{|x|\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$ for all x. □

The following result also proves that the Gaussian Model converges to the true option value as σ approaches 0.

Theorem 3.3. *For $\sigma > 0$, $\sup(V_n - V_n^\sigma) \leq \frac{\sigma}{\sqrt{2\pi}}(N - n)E\left[\max_{n\leq k\leq N} e^{-\frac{P_k^2}{2\sigma^2}}\right] \leq \frac{\sigma}{\sqrt{2\pi}}(N - n)$.*

Proof. From Theorem 3.1 it can be shown $\xi_\sigma(x) \le x^+ \le \xi_\sigma(x) + \frac{\sigma}{\sqrt{2\pi}} e^{-\frac{x^2}{\sigma^2}}$ for all $\sigma > 0$. By taking the difference, $V_n - V_n^\sigma = M_N - M_n - (M_N^\sigma - M_n^\sigma) + \sum_{j=n}^{N-1} (P_j - V_{j+})^+ - \xi_\sigma(P_j - V_{j+}^\sigma) > 0$. Then, applying expectations and time n filtrations on both sides, $E\left[V_n - V_n^\sigma | \mathcal{F}_n\right] = V_n - V_n^\sigma$, and

$$E\left[\sum_{j=n}^{N-1} (P_j - V_{j+})^+ - \xi_\sigma(P_j - V_{j+}^\sigma) | \mathcal{F}_n\right] \le E\left[\sum_{j=n}^{N-1} (P_j - V_{j+}^\sigma)^+ - \xi_\sigma(P_j - V_{j+}^\sigma) | \mathcal{F}_n\right]$$

by Theorem 3.1 $V > V_\sigma$ and x^+ is a non-decreasing function. Applying the inequality we get this is less than

$$E\left[\sum_{j=n}^{N-1} \xi_\sigma(P_j - V_{j+}^\sigma) + \frac{\sigma}{\sqrt{2\pi}} e^{-\frac{(P_j - V_{j+}^\sigma)^2}{2\sigma^2}} - \xi_\sigma(P_j - V_{j+}^\sigma) | \mathcal{F}_n\right]$$

which is equal to $\frac{\sigma}{\sqrt{2\pi}} E\left[\sum_{j=n}^{N-1} e^{-\frac{(P_j - V_{j+}^\sigma)^2}{2\sigma^2}} | \mathcal{F}_n\right] \le \frac{\sigma}{\sqrt{2\pi}} (N-n) E\left[\max_{n \le k \le N} e^{-\frac{P_k^2}{2\sigma^2}} | \mathcal{F}_n\right] \le \frac{\sigma}{\sqrt{2\pi}} (N-n)$. Since $e^{-x^2} < 1$ for all x. □

This implies that convergence scales linearly in σ and depends on the remaining amount of exercise dates, since both policy values are identical at N, and diverge as the recursion algorithm is applied further.

3.2 Shannon-Entropy Regularisation Model

The second approach to remove the indicator function is through converting the backwards stochastic equation into a control problem and applying a regularising term to enforce a unique solution that can be studied. Firstly, we introduce a control process Γ that evolves discretely with increments $\Delta\Gamma_j$ that are in $[0, 1]$ to replace the indicator, and represent the weighting on the stopping decision. The value of the option must be maximised with this process.

$$V_n = P_N - (M_N - M_n) + \operatorname*{ess\,sup}_{\Delta\Gamma \in [0,1]} \sum_{j=n}^{N-1} (P_j - V_{j+}) \Delta\Gamma_j.$$

Suppose the assumption on the form of discretely times random variable increments $\Delta\Gamma_j$ can be relaxed with a discrete random variable X_j with support on a finite number of points on $[0, 1]$, $S = \{x_0, x_1, x_2, ...x_L\}$ with a given smooth probability mass function p. This constraint $\sum_{x \in S} p(x) = 1$ forces an explicit and unique solution to be found.

$$V_n = P_N - (M_N - M_n) + \operatorname*{ess\,sup}_{p} \sum_{j=n}^{N-1} \sum_{x \in S} (P_j - V_{j+}) x p_j(x).$$

The form of p, if it exists, is not well posed, so a regularising term is added. For example, Dong (2022) [3] used a regulariser in the form of $p_j(1 - \ln(p_j))$. We instead chose the Shannon-Entropy regulariser $p_j \ln(p_j)$. This penalty term adds randomness to the stopping decision similar to the Gaussian Model.

$$V_n = P_N - (M_N - M_n) + \operatorname*{ess\,sup}_{p} \sum_{j=n}^{N-1} \sum_{x \in S} \{(P_j - V_{j+}) x p_j(x) - \lambda p_j(x) \ln(p_j(x))\} \quad \text{(A)}$$

Intuitively, the double summation needs to be maximised in p, so $p_j(x) \ln(p_j(x))$ needs to be minimised. Across the interval $(0, 1]$, $\sup_{z \in (0,1]} z \ln(z) = 0$ when $z = 1$ or z approaches 0. Meanwhile the minimum of $z \ln(z)$ is

$-e^{-1}$, which occurs when $z = e^{-1} \approx 0.37$. Thus this condition punishes confident probability assignments of 0 and 1 which introduces greater spread of estimates. The higher λ is, the stronger the randomness.

Clearly, when $\lambda = 0$, we get the exact backwards equation for the option value, so we will observe the effects of taking the limit of λ closer to 0 on variance and the point estimate.

In order to obtain a stopping rule, we must look at the following integral to be maximised:

$$\sum_{x\in S}\{(P_j - V_{j+})xp_j(x) - \lambda p_j(x)\ln(p_j(x))\} = (P_j - V_{j+})\sum_{x\in S}\left(xp_j(x) - \frac{\lambda p_j(x)\ln(p_j(x))}{(P_j - V_{j+})}\right)$$
$$= (P_j - V_{j+})\sum_{x\in S}p_j(x)\left(\frac{x\frac{(P_j-V_{j+})}{\lambda} - \ln(p_j(x))}{\frac{(P_j-V_{j+})}{\lambda}}\right) = \frac{(P_j - V_{j+})}{y_j}\sum_{x\in S}p_j(x)\left(y_jx - \ln(p_j(x))\right)$$

where $y_j = \frac{(P_j-V_{j+})}{\lambda}$ for simplicity. The goal of maximising the integral in p is performed by method of Lagrangian multipliers under the constraint $\sum_{x\in S}p(x) = 1$, and objective function

$$\mathcal{L} = \sum_{x\in S}p_j(x)(y_jx - \ln(p_j(x))) + \mu\left(1 - \sum_{x\in S}p(x)\right).$$

The first order condition $\frac{\partial\mathcal{L}}{\partial\mu} = 0$ will yield the constraint. On the other hand,

$$\frac{\partial}{\partial p_j(x)}\{p_j(x)(y_jx - \ln(p_j(x))) + \mu(1 - p_j(x))\} = y_jx - \ln(p_j(x)) - 1 - \mu = 0$$
$$\ln(p_j(x)) = y_jx - \mu - 1$$
$$p_j(x) = e^{y_jx-\mu-1}$$
$$1 = e^{-\mu-1}\sum_{x\in S}e^{y_jx} \quad \text{by summing over the support}$$
$$\mu = \ln\left(\sum_{x\in S}e^{y_jx}\right) - 1$$
$$p_j(x) = \frac{e^{y_jx}}{\sum_{x'\in S}e^{y_jx'}} \qquad \ln(p_j(x)) = y_jx - \ln\left(\sum_{x'\in S}e^{y_jx'}\right)$$

Hence the integral simplifies to

$$\frac{(P_j - V_{j+})}{y_j}\sum_{x\in S}\frac{e^{y_jx}}{\sum_{x'\in S}e^{y_jx'}}\ln\left(\sum_{x''\in S}e^{y_jx''}\right) = \frac{(P_j - V_{j+})}{y_j}\ln\left(\sum_{x\in S}e^{y_jx}\right) = \lambda\ln\left(\sum_{x\in S}e^{\frac{P_j-V^\lambda_{j+}}{\lambda}x}\right).$$

Ultimately, the backward equation to solve for (A) is

$$V_n = P_N - (M_N - M_n) + \sum_{j=n}^{N-1}\lambda\ln\left(\sum_{x\in S}e^{\frac{P_j-V^\lambda_{j+}}{\lambda}x}\right).$$

In this equation, the integrand $\lambda\ln\left(\sum_{x\in S}e^{\frac{P_j-V^\lambda_{j+}}{\lambda}x}\right)$ acts as a replacement for $(P_j - V_{j+})^+$, so define the function $\Psi_{\lambda,S}(y) = \lambda\ln\left(\sum_{x\in S}e^{\frac{yx}{\lambda}}\right)$. If points in S are of the form $x_k = \frac{k}{L}$, then the sum inside the logarithm is a geometric progression, simplifying to $\Psi_{\lambda,L}(y) = \lambda\ln\left(\sum_{k=0}^{L}e^{\frac{kx}{\lambda L}}\right) = \lambda\ln\left(\frac{e^{\frac{x(L+1)}{\lambda L}}-1}{e^{\frac{x}{\lambda L}}-1}\right)$.

So, the recursive policy update algorithm is:

$$V^\lambda_n = V^\lambda_{n+} + \lambda\ln\left(\sum_{x\in S}e^{\frac{P_n-V^\lambda_{n+}}{\lambda}x}\right) = V^\lambda_{n+} + \lambda\ln\left(\frac{e^{(P_n-V^\lambda_{n+})\frac{(L+1)}{\lambda L}} - 1}{e^{\frac{P_n-V^\lambda_{n+}}{\lambda L}} - 1}\right) = V^\lambda_{n+} + \Psi(P_j - V_{j+})$$

Theorem 3.4. $|\Psi_{\lambda,L}(x) - x^+| \leq \lambda \ln(L+1)$ *for all* $x \neq 0$

Proof. If $x \leq 0$, then $|\Psi_{\lambda,L}(x) - 0| = |\lambda \ln\left(\sum_{k=0}^{L} e^{\frac{k}{L\lambda}x}\right)| \leq |\lambda \ln\left(\sum_{k=0}^{L} e^{\frac{k}{L\lambda}0}\right)| = \lambda \ln(L+1)$.

If $x > 0$, $|\Psi_{\lambda,L}(x) - x| = |\lambda \ln\left(\sum_{k=0}^{L} e^{\frac{k}{L\lambda}x}\right) - x| \leq |\lambda \ln\left(\sum_{k=0}^{L} e^{\frac{L}{L\lambda}x}\right) - x| = |\lambda \ln((L+1)e^{\frac{x}{\lambda}}) - x| = |\lambda(\ln(L+1) + \ln(e^{\frac{x}{\lambda}})) - x| = |\lambda(\ln(L+1) + \frac{x}{\lambda}) - x| = |\lambda \ln(L+1)|$. □

Theorem 3.5. *For all* $\lambda > 0$*, then* $V_n^\lambda > V_n$ *for all* $n < N$.

Proof. It is known that $V_N = V_N^\sigma = P_N$, so, $V_{N-1+}^\lambda = V_{N-1+}$. Hence $V_{N-1}^\lambda = V_{N-1+}^\lambda + \Psi(P_{N-1} - V_{N-1+}^\lambda; \lambda, L) > V_{N-1+}^\lambda + (P_{N-1} - V_{N-1+}^\lambda)^+ = V_{N-1+} + (P_{N-1} - V_{N-1+})^+ = V_{N-1}$, so $V_{N-1}^\lambda > V_{N-1}$.
Using backward induction, assume that $V_k^\lambda > V_k$, so $V_{k-1+}^\lambda = E[V_k^\lambda|\mathcal{F}_{k-1}] > E[V_k|\mathcal{F}_{k-1}] = V_{k-1+}$. Now $V_{k-1}^\lambda = V_{k-1+}^\lambda + \Psi_{\lambda,L}(P_{k-1} - V_{k-1+}^\lambda) > V_{k-1+}^\lambda + (P_{k-1} - V_{k-1+}^\lambda)^+ =$
$\max(V_{k-1+}^\lambda, P_n) \geq \max(V_{k-1+}, P_n) = V_{k-1}$ due to the inductive assumption and that $\Psi_{\lambda,L}(x) > x^+$. □

Theorem 3.6. *For* $0 < \lambda < 1$, $\sup(V_n - V_n^\lambda) \leq \lambda(N-n)\ln(L+1)$.

Proof. From Theorem 3.4 and Theorem 3.5 it can be shown $\Psi_{\lambda,L}(x) - \lambda \ln(L+1) \leq x^+ \leq \Psi_{\lambda,L}(x)$ for all $\lambda > 0$. Taking the difference, $V_n^\lambda - V_n = M_N^\lambda - M_n^\lambda - (M_N - M_n) + \sum_{j=n}^{N-1} \Psi_{\lambda,L}(P_j - V_{j+}^\lambda) - (P_j - V_{j+})^+ > 0$.
Taking expectations on both sides, $E\left[V_n - V_n^\lambda|\mathcal{F}_n\right] = V_n - V_n^\sigma$, and

$$E\left[\sum_{j=n}^{N-1} \Psi_{\lambda,L}(P_j - V_{j+}^\lambda) - (P_j - V_{j+})^+|\mathcal{F}_n\right] \leq E\left[\sum_{j=n}^{N-1} \Psi_{\lambda,L}(P_j - V_{j+}^\lambda) - (P_j - V_{j+}^\lambda)^+|\mathcal{F}_n\right]$$

by Theorem 3.5, $V^\lambda > V$ and x^+ is a non-decreasing function. Applying the inequality we get this is less than

$$E\left[\sum_{j=n}^{N-1} \Psi(P_j - V_{j+}^\lambda) - (\Psi_{\lambda,L}(P_j - V_{j+}^\lambda) - \lambda \ln(L+1))|\mathcal{F}_n\right]$$

which is equal to $E\left[\sum_{j=n}^{N-1} \lambda \ln(L+1)|\mathcal{F}_n\right] = \lambda(N-n)\ln(L+1)$ □

Similar to the Gaussian Model, this implies that convergence scales linearly in λ and depends on the remaining amount of exercise dates, since both policy values are identical at N, and diverge as the recursion algorithm is applied further. More divergence is observed the more granular the support points are.

3.3 Algorithm

The numerical scheme for pricing American options uses a backwards recursive Monte-Carlo framework and uses Neural Networks to approximate the functional form of continuation value in least-squares sense. The policy update formula has been adjusted to include a function g that smooths out the indicator.

Note that in the implementation of $V_n = V_{n+} + (P_n - V_{n+})g(P_n - V_{n+})$, the first two continuation values are different from the continuation value inside g, since the continuation value estimated by the model is used in the randomised stopping stopping decision only, and the continuation value via backwards recursion is the reward from continuing the options life. The Longstaff-Schwarz paper [4] notes that if the estimated continuation value were used to update policy values, then there will be an upward bias in the final value due to the convexity of the maximum operator. We will ignore this fact for the the Regularisation Model. Attempting to seperate the excess $(P_j - V_{j+})$ from the stopping decision, or the function approximating the indicator, results in a singularity around 0 which leads to numerical complications. However, as will be seen in the results section, this does not significantly skew the results.

Algorithm 1 Pricing Algorithm

Require: Number of sample paths M, risk-free rate r, volatility v, number of steps N, payoff function f, smoothing function g, network architecture

Ensure: Final value V_0

generate the sample paths X

set $V_N^m = f(X_N^m)$ for all $m = 1$ to M

for each $n = N - 1$ to 1 **do**

$V_n^m \leftarrow V_{n+1}^m e^{-r \cdot dt}$ for all $m = 1$ to M

optimise θ_n for objective function

$$\frac{1}{M} \sum_{m=1}^{M} \left(F^{\theta_n}(X_n^m) - V_n^m \right)^2$$

for each $m = 1$ to M **do**

$V_n^m \leftarrow V_n^m + (f(X_n^m) - V_n^m)\, g\left(f(X_n^m) - F^{\theta_n}(X_n^m)\right)$

end for

end for

$V_0^m \leftarrow V_1^m e^{-r \cdot dt}$

return $V_0 = \frac{1}{M} \sum_{m=1}^{M} V_0^m$

4 Numerical results

The algorithm in Section 3.3 was implemented in Python 3.10.11 along with numPy, pandas and PyTorch libraries. The sample paths for the asset are GBM simulations in order to compare to literature such as Becker [1]. That is, the process described by $dS_t = (r - \delta)\, dt + v\, dW_t, S_0 = s$. With all constants given, expiry date T and number of time steps N, the value of the asset at time t for path m is $S_t^m = se^{(r-\delta-\frac{v^2}{2})t + v\sum_{k=1}^{t} \Delta W_k^m}$ where $\Delta W_k^m = W_k^m - W_{k-1}^m$ are the Brownian motion increments, $W_k^m \sim N(0, \frac{T}{N})$ and $W_0^m = 0$ for all m. The neural network architecture could take d dimensions of inputs, produced 1 output and had two hidden layers of 32 and 16 neurons each, seperated by ReLu activation.

4.1 Sensitivity Testing

This section will compare the valuation of put options using the proposed models to the Longstaff Schwarz technique in both convergence of the point estimate and the change in variance of the Monte Carlo samples, as well as sensitivity to the change in input parameters.

The following figures will depict the effect of taking σ and λ to 0 in the Gaussian and Regularisation models respectively on the option value and sample path variance, using the Longstaff-Schwarz values as a benchmark. These results are highly dependent on the number of simulations used. Our implementation of the Longstaff-Schwarz model used $1,000,000$ simulations while the Gaussian and Regularisation Models used $100,000$ thousand.

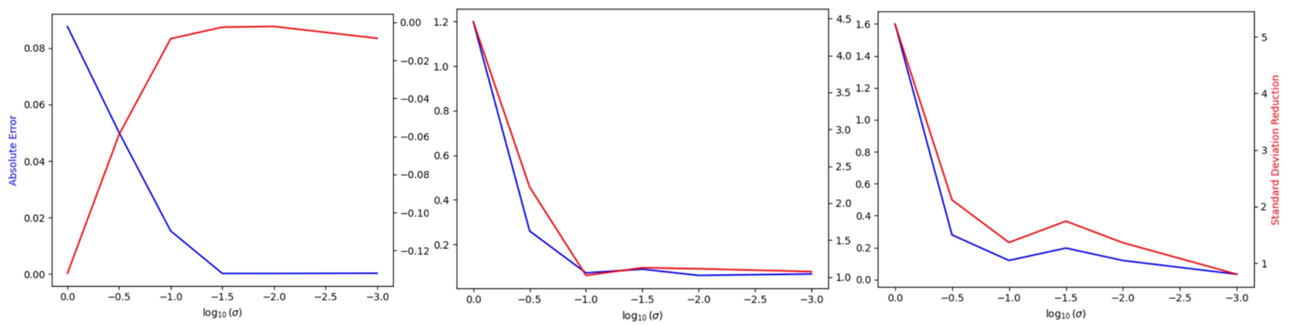
4.1.1 Volatility Changes

In the Gaussian Model (Figure 1), there is a reduction in variance compared to the Longstaff-Schwarz samples for the $v = 0.2$ and $v = 0.25$ cases, with this reduction getting smaller as bias decreases. This is the bias-variance trade off we were expecting to see for more scenarios. However when volatility is lower, $v = 0.1$, this variance reduction effect is flipped, and there is a slight increase in model error.

In all cases for the Regularisation Model (Figure 2), a different pattern is observed. As λ gets smaller, the option value converges. However, the reduction in variance is always negative, meaning model risk may worsen using the Regularisation model.

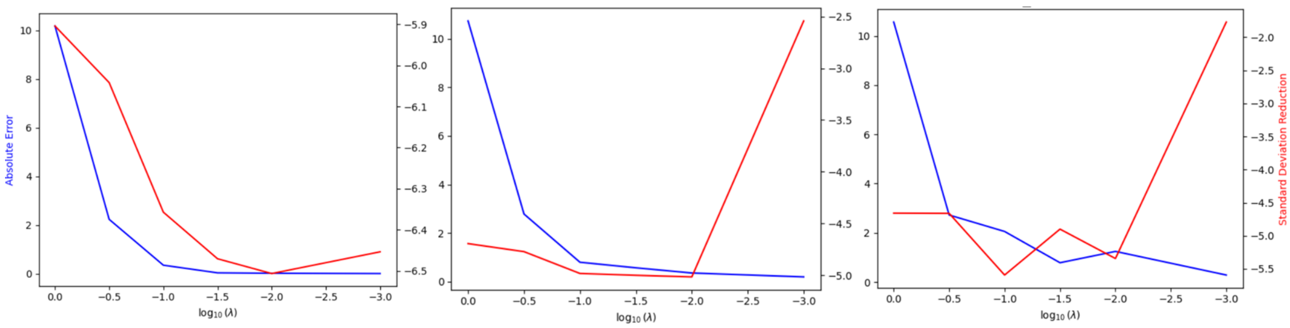
Table 1 describes the change in the option value for changes in volatility for each policy. This is applied to $v = 0.2$ to get a rough approximation for the partial derivative $\frac{\delta V_0}{\delta v}|_{v=v_0} \approx \frac{h_2(V_0(v_0)-V_0(v_0-h_1))-h_1(V_0(v_0+h_2)-V_0(v_0))}{h_1 h_2(h_1+h_2)}$.

The interpretation is that if v were to increase by 1, the option value would shift by $\frac{\delta V}{\delta v}$. A volatility adjustment of this magnitude would be unlikely in a real scenario, so more applicatively, dividing these entries by 100 would roughly describe the change with a 1% increase in volatility. We can observe the values in the Gaussian Model approach the Longstaff Schwarz from below, and the Regularisation from above, suggesting the Gaussian Model is less sensitive to changes in this input parameters.



$v = 0.1$ $v = 0.2$ $v = 0.25$

Figure 1: Put option, Gaussian Model, $s = 90, K = 100, r = 0.06, T = 1, N = 50$



$v = 0.1$ $v = 0.2$ $v = 0.25$

Figure 2: Put option, Regularisation Model, $s = 90, K = 100, r = 0.06, T = 1, N = 50$

σ/λ	1	$10^{-0.5}$	0.1	$10^{-1.5}$	0.01	0.001	Longstaff-Schwarz
Gaussian	15.774947	21.951162	22.624959	23.446612	22.968668	23.074128	24.336893
Regularisation	26.198337	28.831940	38.824398	23.050092	30.587422	26.174088	24.336893

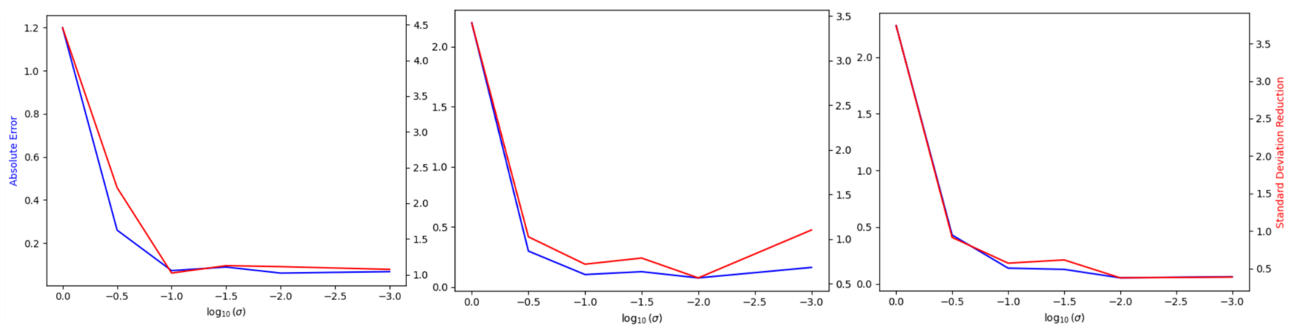
Table 1: Put option, $s = 90, K = 100, r = 0.06, v = 0.2, T = 1, N = 50$

4.1.2 Starting Price Changes

We see very similar results in these cases to the volatility sensitivity testing case $v = 0.2$, with the Gaussian Model (Figure 3) reducing variance and the Regularisation Model (Figure 4) increasing it.

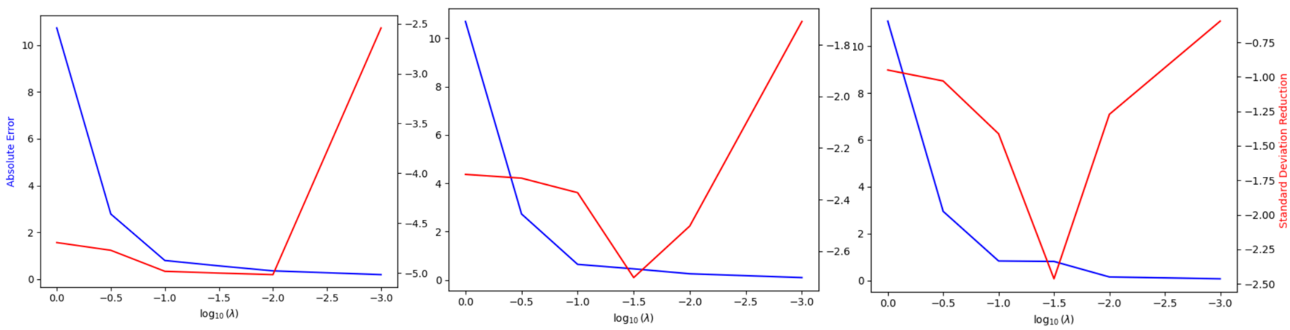
Table 2 depicts the change in the option value for changes in starting price for each policy. This is applied to $s = 100$ to get an approximation $\frac{\delta V_0}{\delta S_0}|_{S_0=s} = \frac{V_0(v_0+h) - V_0(v_0-h)}{2h}$.

The interpretation is that if s were to increase by \$1, the option value would shift by $\frac{\delta V}{\delta S_0}$. This is linked to a concept in Option Theory known as Delta. For put options, Delta is negative since a \$1 increase in the underlying would take the option less 'into-the-money', thus make its potential payoff distribution less.



$s = 90$ $s = 100$ $s = 110$

Figure 3: Put option, Gaussian Model, $K = 100, r = 0.06, v = 0.2, T = 1, N = 50$



$s = 90$ $s = 100$ $s = 110$

Figure 4: Put option, Regularisation Model, $K = 100, r = 0.06, v = 0.2, T = 1, N = 50$

σ/λ	1	$10^{-0.5}$	0.1	$10^{-1.5}$	0.01	0.001	Longstaff-Schwarz
Gaussian	-0.476775	-0.436857	-0.422981	-0.420807	-0.416171	-0.423750	-0.420494
Regularisation	-0.408675	-0.418091	-0.433940	-0.447188	-0.574733	-0.424451	-0.420494

Table 2: Put option, $s = 90, K = 100, r = 0.06, v = 0.2, T = 1, N = 50$

4.2 Multi-Dimensional Case

The algorithm can also handle option pricing requiring multiple input dimensions. An example is a Max-Call option which has payoff dependent on multiple underlying values $P_\tau = (\max_{i=1,\dots,d} S_\tau^i - K)^+$. Table 3 compares the Gaussian Model to Becker's [1] results for a symmetrical American max-call option, where all assets have the same input parameters.

S_0	V^1	$V^{10^{-0.5}}$	$V^{0.1}$	$V^{10^{-1.5}}$	$V^{0.01}$	$V^{0.001}$	$V^{0.0001}$	Becker
90	7.930149	8.152506	8.202082	8.165762	8.095579	8.214073	8.080862	8.074
100	13.680307	13.921888	13.873100	13.898158	13.942288	13.928569	13.932113	13.899
110	21.164794	21.319122	21.226748	21.203929	21.386157	21.340414	21.342699	21.349

Table 3: Symmetric Max-Call option, Gaussian Model, $K = 100, r = 0.05, \delta = 0.1, v = 0.2, T = 3, N = 9, d = 2$

5 Conclusion

Our results suggest the Gaussian Model is more reliable and reducing model risk relative to the Longstaff-Schwarz model, while the Regularisation Model is undesirable due to increasing model risk. This method of pricing options begs further investigation with many ways forward. Firstly, the control problem leading to the Regularisation Model was posed with with mass function of a discrete random variable, so changing this to a continuous random variable may yield different results. Experimenting with different regularising terms may generate a model more adept at reducing model risk. The algorithm could be expanded to accomodate exotics such as swing options, lookback options and game options. Further sensitivity testing for the simulated asset process could reveal more input parameter regions where variance reduction is observed. Moreover, the neural networks loss function could incorporate the fact the indicator function is substituted by a smooth approximator to allow optimisation methods such as stochastic gradient descent. Lastly, the models could be applied to real-time stock price data and compared to listed option values.

References

[1] Becker, S., Cheridito, P. and Jentzen, A. [2019], 'Deep optimal stopping', *Journal of Machine Learning Research* **20**, 1–25.

[2] Bensoussan, A. [1984], 'On the theory of option pricing.', *Acta Appl Math* **2**, 139–158.

[3] Dong, X. [2022], 'Randomized optimal stopping problem in continuous time and reinforcement learning algorithm', *Journal of Computational Finance* **28**, 45–72.

[4] Longstaff, F. and Schwartz, E. [2001], 'Valuing american options by simulation: A simple least-squares approach', *Review of Financial Studies* **14**, 113–147.

[5] Pollard, D. [2013], 'An introduction to optimal stopping', *Yale University Course Handout* .
URL: *http://www.stat.yale.edu/ pollard/Courses/251.spring2013/Handouts/optimal.pdf*