MONASH
CENTRE FOR
QUANTITATIVE
FINANCE AND
INVESTMENT
STRATEGIES

Pricing Bounds for Volatility Derivatives via Duality and Least Squares Monte Carlo

Monash CQFIS working paper

2017 – 8

Ivan Guo
School of Mathematical Sciences,
Monash University and Centre for
Quantitative Finance and Investment
Strategies
ivan.guo@monash.edu

Grégoire Loeper
School of Mathematical Sciences,
Monash University and Centre for
Quantitative Finance and Investment
Strategies
gregoire.loeper@monash.edu

Abstract

Derivatives on the Chicago Board Options Exchange volatility index have gained significant popularity over the last decade. The pricing of volatility derivatives involves evaluating the square root of a conditional expectation which cannot be computed by direct Monte Carlo methods. Least squares Monte Carlo methods can be used, but the sign of the error is difficult to determine. In this paper, we propose a new model-independent technique for computing upper and lower pricing bounds for volatility derivatives. In particular, we first present a general stochastic duality result on payoffs involving convex (or concave) functions. This result also allows us to interpret these contingent claims as a type of chooser options. It is then applied to volatility derivatives along with minor adjustments to handle issues caused by the square root function. The upper bound involves the evaluation of a variance swap, while the lower bound involves estimating a martingale increment corresponding to its hedging portfolio. Both can be achieved simultaneously using a single linear least square regression. Numerical results show that the method works very well for futures, calls and puts under a wide range of parameter choices.

Pricing Bounds for Volatility Derivatives via Duality and Least Squares Monte Carlo

Ivan Guo[1,2] and Gregoire Loeper[1,2]

[1]School of Mathematical Sciences, Clayton Campus, Monash University, VIC, 3800, Australia

[2]Centre for Quantitative Finance and Investment Strategies*, Monash University, Australia

Abstract

Derivatives on the Chicago Board Options Exchange volatility index have gained significant popularity over the last decade. The pricing of volatility derivatives involves evaluating the square root of a conditional expectation which cannot be computed by direct Monte Carlo methods. Least squares Monte Carlo methods can be used but the sign of the error is difficult to determine. In this paper, we propose a new model independent technique for computing upper and lower pricing bounds for volatility derivatives. In particular, we first present a general stochastic duality result on payoffs involving convex (or concave) functions. This result also allows us to interpret these contingent claims as a type of chooser options. It is then applied to volatility derivatives along with minor adjustments to handle issues caused by the square root function. The upper bound involves the evaluation of a variance swap, while the lower bound involves estimating a martingale increment corresponding to its hedging portfolio. Both can be achieved simultaneously using a single linear least square regression. Numerical results show that the method works very well for futures, calls and puts under a wide range of parameter choices.

Mathematics Subject Classification (2010): 91G20, 91G60
Keywords: VIX derivatives, convex conjugate, least squares Monte Carlo, pricing bounds

1 Introduction

The Chicago Board Options Exchange volatility index, commonly known as VIX, measures the volatility of the S&P500 index. Formally, the VIX is the square root of the expected integrated variance (often called the realised variance) over a 30-day period, multiplied by an annualisation factor. In practice, it is calculated using a weighted sum of options on the S&P500 index and it coincides with the square root of the par variance swap rate. The VIX itself is not a tradable asset, but VIX derivatives such as futures and options are. VIX futures began trading in 2004 while VIX options began in 2006. Since then, VIX derivatives have gained significant popularity as they allow traders to gain direct exposure to the volatility of the S&P500 index without having to hold options on the index.

***Acknowledgements** The Centre for Quantitative Finance and Investment Strategies has been supported by BNP Paribas.

In literature, there have been many theoretical approaches to the pricing of VIX derivatives. In earlier works, the authors focussed on finding analytical pricing formulae for volatility derivatives under particular volatility dynamics. Some examples include Whaley [1] (geometric Brownian motion), Grünbichler and Longstaff [2] (square root process), Detemple and Osakwe [3] (log-normal Ornstein-Ulenbeck process). By only considering volatility futures and options as opposed to VIX derivatives, these works do not explicitly deal with the integrated variance term. This is rectified by Zhang and Zhu [4] who derived an analytical formula for the price of VIX futures under the Heston model. Furthermore, they supplemented their work with empirical analyses by calibrating the model against historical VIX data. This pricing result was further generalised by Lian and Zhu [5] to the Heston model with jumps via a characteristic function approach. Further progress was made for cases where the variance process follows a square root process with jumps (Sepp [6]) and a 3/2 process with jumps (Baldeaux and Badran [7]). Finally, some authors undertook an alternative approach which directly models the variance swaps instead of the volatility. This allows for the consistent modelling and the simultaneous calibration of both index options and VIX derivatives. See Cont and Kokholm [8] for an example of this approach.

In terms of numerical methods, PDE methods work well but only if the underlying dynamic is Markovian and resides in a low dimensional space. Due to the non-linearity of the square root function in the definition of the VIX, the price of VIX futures is highly model-dependent and cannot be inferred from direct Monte Carlo simulations. Instead, the evaluation of the conditional expectation of the realised variance can be handled by nested simulations or least squares regressions. Nested Monte Carlo has good accuracy, but it is computationally expensive. Least square Monte Carlo approaches, popularised by Carriere [9] as well as Longstaff and Schwartz [10] for Bermudan options, are much faster. Although the results are asymptotically unbiased, it is usually difficult to determine the sign of the error, which can be a useful piece of information in risk management. Rogers [11] as well as Haugh and Kogan [12] proposed a stochastic duality result which produces an upper bound to Bermudan option prices, complementing the original least squares Monte Carlo method which naturally provides a lower bound via suboptimal exercise policies. The quality of the upper bound relies on the identification of a martingale which majorises the price process. Andersen and Broadie [13] suggested to estimate the martingale using nested Monte Carlo. Later on more efficient approaches were found in various works such as Schoenmakers et al. [14]. An overview of these upper bound methods without using nested simulations can be found in Joshi and Tang [15].

More recently, De Marco and Henry-Labordere [16] as well as Guyon et al. [17] worked on the robust hedging of VIX derivatives. The focus is on the super and sub-replication of VIX derivatives using vanilla options and VIX futures in a model-free setting. This problem is a variant of the classical martingale transport problem where the marginal distributions of the underlying are known at two different dates, while the distribution of the VIX future satisfies a particular constraint. Analytical pricing bounds are provided for VIX derivatives and illustrated through numerical experiments. In comparison to the pricing bounds of our paper, the results of [16] and [17] are quite different in nature and purpose. The works of [16] and [17] address the problem of finding analytical robust hedging pricing bounds under model uncertainty, whereas we focus on methods for computing sharper numerical bounds under a given model.

In this paper, we present a new application of the stochastic duality and the least squares Monte Carlo methods to VIX derivatives, resulting in true upper and lower pricing bounds. Our results are applicable to any pricing problem where the payoff contains a convex (or concave) function of a conditional expectation. One bound naturally arises from convexity, while the other can be found using convex (or concave) conjugates, which transforms the contingent claim into a type of chooser options (see Remark 4). These results are then applied to VIX derivatives. Numerically, by using techniques similar to Schoenmakers et al. [14], we perform a single least squares Monte Carlo to compute the required conditional expectation and martingale increment, which are used to evaluate the pricing bounds. The main results of the paper are Theorems 3 and 5. Theorem 3 presents a general stochastic duality result on payoffs involving concave functions. Theorem 5 applies it to VIX derivatives, with minor adjustments to handle issues caused by the square root functions. Even though much of the paper is focussing on VIX derivatives in the local-stochastic volatility model, the main techniques and results (as shown in Section 2) are in fact completely model independent and directly applicable to many other valuation problems in various settings.

The paper is organised as follows. The model independent duality bounds are first presented in Section 2. Then Section 3 introduces the VIX in a general local-stochastic volatility (LSV) framework and derives the pricing bounds for VIX derivatives, along with techniques to handle the square root function specifically. Section 4 describes the Monte Carlo algorithm in detail while Section 5 provides some numerical examples. Finally, Section 6 contains some concluding remarks.

2 Duality Bounds

We begin by presenting our bounding results in a model independent framework. Let $(\Omega, \mathbb{F}, \mathbb{P})$ be a filtered probability space, where the filtration $\mathbb{F} = \{\mathcal{F}_t : t \in [0,T]\}$ represents the information flow available to market participants, $\mathbb{P}$ is a pricing measure and $T > 0$ is a fixed time horizon. Let H be an $\mathcal{F}_T$-measurable random variable and g be a concave function.

If we want to numerically evaluate $\mathbb{E}(g(\mathbb{E}(H \,|\, \mathcal{F}_t)))$ via Monte Carlo simulations, the main challenge is the computation of the inner conditional expectation. Due to the non-linearity of g, a standard Monte Carlo simulation is insufficient. Instead, it requires a nested simulation or a least square Monte Carlo method. In this section, we assume that the exact value of $g(\mathbb{E}(H \,|\, \mathcal{F}_t))$ is unavailable, and propose a new Monte Carlo approach which produces true upper and lower bounds. This approach is similar to the well-known duality bounds for Bermudan and American options.

We will first briefly describe the duality bounds for a Bermudan or American option. For a more detailed exposition, the readers are referred to Rogers [11] or Haugh and Kogan [12]. Suppose that the discounted payoff process of the option is Z. If the option is alive at time $t \in [0,T]$, the holder of the option chooses $\tau \in \mathcal{T}_t$ where $\mathcal{T}_t$ is the set of stopping times with values in $[t,T]$, corresponding to the available exercise opportunities (discrete in Bermudan, continuous in American). For any chosen τ, the holder receives Z_τ at time τ. It is well-known that at time t the discounted price of the option is given by $V_t = \operatorname{ess\,sup}_{\tau \in \mathcal{T}_t} \mathbb{E}(Z_\tau \,|\, \mathcal{F}_t)$, and that the price process V is a supermartingale. It is clear that a lower bound of the option price V_0 can be found by selecting any sub-optimal stopping time τ' and computing $\mathbb{E}Z_{\tau'}$, with equality being achieved if $\tau' = \tau^*$ is the optimal stopping time. For an upper bound, let M be an arbitrary martingale and consider $M_0 + \mathbb{E}(\sup_t Z_t - M_t)$, where the supremum inside the expectation is taken path-wise. The validity of this upper bound can be checked by exchanging the expectation with the supremum and applying the optional sampling theorem. Equality is reached if the martingale M is taken from the Doob-Meyer decomposition of the price process V, which can also be interpreted as the hedging portfolio. To summarise, bounds for the option price V_0 are given by

$$\mathbb{E}Z_\tau \le V_0 \le M_0 + \mathbb{E}\left(\sup_{t\in[0,T]} Z_t - M_t \right),$$

where τ is an arbitrary stopping time and M is an arbitrary martingale.

A similar technique will be applied to obtain bounds for $\mathbb{E}(g(\mathbb{E}(H\,|\,\mathcal{F}_t)))$. First, let us briefly recall the following well-known definition and properties of the convex (and concave) conjugate. For more details, the readers are referred to Section 12 of Rockafellar [18], in particular, Theorem 12.2 and the comments preceding it.

Definition 1. (i) Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a function such that $f \not\equiv +\infty$. Then the *convex conjugate* of f is the function $f^* : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ defined by

$$f^*(y) := \sup_{x \in \mathbb{R}^d} (x \cdot y - f(x)).$$

The convex conjugate is also often known as the *Legendre-Fenchel transform*.

(ii) Let $g : \mathbb{R}^d \to \mathbb{R} \cup \{-\infty\}$ be a function such that $g \not\equiv -\infty$. Then the *concave conjugate* of g is the function $g_* : \mathbb{R}^d \to \mathbb{R} \cup \{-\infty\}$ defined by

$$g_*(y) := \inf_{x \in \mathbb{R}^d} (x \cdot y - g(x)).$$

In particular, $-g_*(y) = (-g)^*(-y)$.

Even though Definition 1 is formulated for functions defined on $\mathbb{R}^d$, it can actually be applied to functions defined on an arbitrary domain $\mathcal{D} \subset \mathbb{R}^d$. This can be done by simply setting $f(x) = +\infty$ in (i) or $g(x) = -\infty$ in (ii) for $x \in \mathbb{R}^d \setminus \mathcal{D}$.

Proposition 2. *Let f, f^*, g, g_* be as defined in Definition 1. Then we have the following properties:*
(i) f^* is convex and g_* is concave;
(ii) $f^{**} \equiv f$ if and only if f is convex and lower semi-continuous; $g_{**} \equiv g$ if and only if g is concave and upper semi-continuous.

Proof. In this proof, we will reference results from Rockafellar [18]. Only the convex case will be addressed, as the concave case can be dealt with analogously. Recall that the function $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ satisfies $f \not\equiv +\infty$.

(i) Theorem 12.2 in [18] shows that f^* is convex when f is convex. The comments preceding the theorem, as well as Corollary 12.1.1 in [18], extend this fact to an arbitrary function f.

(ii) Theorem 12.2 in [18] shows that $f^{**} = f$ if and only if f is a closed convex function. Theorem 7.1 in [18] and the remarks following it show that f is a closed convex function if and only if it is convex and lower semi-continuous. The combination of these statements implies the required result. □ □

Now we are in the position to state the main result of the paper. Note that Theorem 3 is formulated for concave functions since they are more readily applicable to VIX derivatives. Analogous results for convex functions can be easily obtained by negating the concave functions.

Theorem 3. *Fix $t \in [0, T]$. Let $g : \mathbb{R}^d \times \Omega \to \mathbb{R} \cup \{-\infty\}$ be an $\mathcal{F}_t$-measurable function such that $g(\cdot, \omega)$ is concave and upper semi-continuous almost surely. Let H be an $\mathcal{F}_T$-measurable, $\mathbb{R}^d$-valued random variable such that both H and $g(\mathbb{E}(H\,|\,\mathcal{F}_t))$ are integrable.*

(i) There exists an $\mathcal{F}_t$-measurable function $g_* : \mathbb{R}^d \times \Omega \to \mathbb{R} \cup \{-\infty\}$ such that $g_*(\cdot, \omega)$ is the concave conjugate of $g(\cdot, \omega)$ almost surely. Furthermore,

$$\mathbb{E}\Big(g(\mathbb{E}(H\,|\,\mathcal{F}_t))\Big) = \inf_{Y \in \mathcal{Y}_t} \mathbb{E}\Big(Y \cdot H - g_*(Y)\Big), \tag{1}$$

where $\mathcal{Y}_t$ is the set of $\mathcal{F}_t$-measurable, $\mathbb{R}^d$-valued random variables.

(ii) We also have the equality

$$\mathbb{E}\Big(g(\mathbb{E}(H \,|\, \mathcal{F}_t))\Big) = \sup_{M \in \mathcal{M}_t} \mathbb{E}\Big(g(H - M_T)\Big), \tag{2}$$

where $\mathcal{M}_t$ is the set of $\mathbb{R}^d$-valued martingales which vanish at time t.

Proof. (i) Since $g(\mathbb{E}(H \,|\, \mathcal{F}_t))$ is integrable, we must have $g(\cdot, \omega) \not\equiv -\infty$ almost surely, which implies the existence of g_*. From Definition 1 and Proposition (2), we almost surely have the equality

$$g(x, \omega) = \inf_{y \in \mathbb{R}^d} (x \cdot y - g_*(y, \omega)). \tag{3}$$

Since g, g_* and $x := \mathbb{E}(H \,|\, \mathcal{F}_t)$ are all $\mathcal{F}_t$-measurable, (3) implies

$$\mathbb{E}\Big(g(\mathbb{E}(H \,|\, \mathcal{F}_t))\Big) = \mathbb{E}\Big(\inf_{y \in \mathbb{R}^d} y \cdot \mathbb{E}(H \,|\, \mathcal{F}_t) - g_*(y)\Big) \leq \inf_{Y \in \mathcal{Y}_t} \mathbb{E}\Big(Y \cdot H - g_*(Y)\Big).$$

To show that equality can be achieved, first note that the integrability of $g(\mathbb{E}(H \,|\, \mathcal{F}_t))$ implies that $g(\mathbb{E}(H \,|\, \mathcal{F}_t)) > -\infty$ almost surely. For each $n \in \mathbb{N}^*$, choose Y_n so that

$$Y_n \cdot \mathbb{E}(H \,|\, \mathcal{F}_t) - g_*(Y_n) \leq g(\mathbb{E}(H \,|\, \mathcal{F}_t)) + \frac{1}{n}.$$

Again, since g, g_* and $\mathbb{E}(H \,|\, \mathcal{F}_t)$ are all $\mathcal{F}_t$-measurable, Y_n can also be chosen to be $\mathcal{F}_t$-measurable. Thus

$$\inf_{Y \in \mathcal{Y}_t} \mathbb{E}\Big(Y \cdot H - g_*(Y)\Big) \leq \mathbb{E}\Big(Y_n \cdot H - g_*(Y_n)\Big) \leq \mathbb{E}\Big(g(\mathbb{E}(H \,|\, \mathcal{F}_t))\Big) + \frac{1}{n},$$

giving the required result as $n \to \infty$.

(ii) For any $M \in \mathcal{M}_t$, recall that $M_t = 0$ by definition. Now, using Jensen's inequality and the $\mathcal{F}_t$-measurability of g and $\mathbb{E}(H \,|\, \mathcal{F}_t)$, we have

$$\begin{aligned} \mathbb{E}\Big(g(H - M_T)\Big) &= \mathbb{E}\Big(\mathbb{E}(g(H - M_T) \,|\, \mathcal{F}_t)\Big) \\ &\leq \mathbb{E}\Big(g(\mathbb{E}(H - M_T \,|\, \mathcal{F}_t))\Big) = \mathbb{E}\Big(g(\mathbb{E}(H \,|\, \mathcal{F}_t))\Big). \end{aligned}$$

Furthermore, equality can be achieved by choosing the martingale defined by $M_s := \mathbb{E}(H - \mathbb{E}(H \,|\, \mathcal{F}_t) \,|\, \mathcal{F}_s)$. Thus (2) is established. □

□

Remark 4. We may interpret $\mathbb{E}(g(\mathbb{E}(H \,|\, \mathcal{F}_t)))$ as the value of a contingent claim with a discounted payoff of $g(\mathbb{E}(H \,|\, \mathcal{F}_t))$ at time t. Unlike Bermudan or American options, this contingent claim does not involve any stopping strategies. Nevertheless, (1) in Theorem 3 (i) shows that the contingent claim can be interpreted as a type of chooser options in the following way. The seller of the option may select a vector $y \in \mathbb{R}^d$ at time t, and then must pay the holder the discounted payoff $y \cdot H - g_*(y)$ at time T. If the seller chooses optimally, i.e., minimising the expected payoff at time t, then the value of the option coincides with the original contingent claim which pays $g(\mathbb{E}(H \,|\, \mathcal{F}_t))$ at time t.

3 Bounds for VIX Derivatives

The bounding techniques and results of this section are model independent, in the sense that they can be applied to any model of the form

$$dS_t = \mu_t S_t dt + \sigma_t S_t dW_t,$$

where S_t is the price of a stock or an index, μ_t and σ_t are $\mathbb{F}$-adapted processes and W_t is a standard Brownian motion. To simplify our notation, the interest rate is set to be zero. Before continuing, let us again emphasise that the main results of the paper, Theorems 3 and 5, are directly applicable to large families of models, including models with stochastic interest rates, high dimensional cases, jump processes, and so on.

Let $0 \leq t_0 \leq T$. The *realised variance* of S_t during the time period $[t_0, T]$ is defined to be

$$AF \sum_{i=1}^{n} \left(\log \frac{S_{t_i}}{S_{t_{i-1}}} \right)^2,$$

where $t_0 < t_1 < \cdots < t_n = T$ are observation dates of S_t and AF is an annualisation factor. For example, if t_i corresponds to daily observations then $AF = 100^2 \times 252/n$ and the realised variance is expressed in basis points per annum. As the mesh of the partition $\pi^n := \{t_i : 0 \leq i \leq n\}$ tends to zero, the realised variance $R = R(t_0, T)$ can be represented as the quadratic variation of $\log S_t$, given by

$$R = R(t_0, T) := \lim_{n\to\infty} AF \sum_{t_i \in \pi^n} \left(\log \frac{S_{t_i}}{S_{t_{i-1}}} \right)^2 = \frac{100^2}{T - t_0} \int_{t_0}^{T} \sigma_t^2 \, dt. \tag{4}$$

The *volatility index (VIX)* $I = I(t_0, T)$ is defined to be the square root of the expected realised variance over a time period of one month,

$$\begin{aligned} I = I(t_0, T) &:= \sqrt{\mathbb{E}(R(t_0, T) \,|\, \mathcal{F}_{t_0})} \qquad (5) \\ &= 100 \times \sqrt{\frac{1}{T - t_0} \mathbb{E}\left(\int_{t_0}^{T} \sigma_t^2 \, dt \,\middle|\, \mathcal{F}_{t_0} \right)}. \end{aligned}$$

However, this quantity is not directly observable. In practice, the VIX is instead defined using the value of the log-contract according to the following identity [1],

$$\mathbb{E}\left(\frac{1}{2} \int_{t_0}^{T} \sigma_t^2 \, dt \,\middle|\, \mathcal{F}_{t_0} \right) = -\mathbb{E}\left(\log \frac{S_T}{F} \,\middle|\, \mathcal{F}_{t_0} \right) \tag{6}$$

$$= \int_{0}^{F} \frac{\mathbb{E}((k - S_T)^+ \,|\, \mathcal{F}_{t_0})}{k^2} \, dk + \int_{F}^{\infty} \frac{\mathbb{E}((S_T - k)^+ \,|\, \mathcal{F}_{t_0})}{k^2} \, dk, \tag{7}$$

where $F = \mathbb{E}(S_T \,|\, \mathcal{F}_{t_0})$ is the forward price (e.g., see Lian and Zhu [5]). As shown by (7), the value of the log-contract (and hence the VIX) can actually be inferred if the market prices of call and put options are available over a continuous spectrum of strikes. Since the methods described in this paper are able to numerically compute $\mathbb{E}(\sqrt{\mathbb{E}(H \,|\, \mathcal{F}_{t_0})})$ for any $\mathcal{F}_T$-measurable random variable H, they are equally applicable to both definitions of the VIX. For simplicity, we have chosen to work with the realised variance formulation given by (5).

Common derivatives on the VIX include futures, swaps, call options and put options. We will mostly focus on the pricing of VIX futures and the VIX caps, which involves the computation of the following expectations:

$$u^f := \mathbb{E}(I(t_0, T)) = \mathbb{E}(\sqrt{\mathbb{E}(R(t_0, T) \,|\, \mathcal{F}_{t_0})}), \tag{8}$$

$$u^c := \mathbb{E}(\min(I(t_0, T), K)) = \mathbb{E}(\min(\sqrt{\mathbb{E}(R(t_0, T) \,|\, \mathcal{F}_{t_0})}, K)). \tag{9}$$

[1] The first equality (6) holds only when the drift μ_t is deterministic. In this case, the realised variance is equivalent to the log-contract. For cases with stochastic drifts, an adjustment term is needed. The second equality (7) always holds.

Many other derivatives such as swaps, calls and puts can then be simply written in terms of u^f and u^c:

$$u^{swap} := \mathbb{E}(I(t_0,T) - K) = u^f - K, \tag{10}$$
$$u^{call} := \mathbb{E}(I(t_0,T) - K)^+ = u^f - u^c, \tag{11}$$
$$u^{put} := \mathbb{E}(K - I(t_0,T))^+ = K - u^c. \tag{12}$$

Note that if we were working in a model with stochastic interest rates, then forward prices will be used instead of futures in (10)–(12).

Since $\sqrt{x}$ and $\min(\sqrt{x}, c)$ are concave functions, Theorem 3 provides natural bounds for VIX futures and caps. The quality of the bounds depends on the exact choice of Y in (1) and M in (2). However, there is a problem with the lower bound

$$u^f \geq \mathbb{E}\left(\sqrt{R - M_T}\right), \tag{13}$$

since for many choices of M, M_T would exceed R with non-zero probability, which then leads to the unusable lower bound of $-\infty$. This issue is resolved by the following theorem.

Theorem 5. *Denote the realised variance over $[t_0, T]$ by $R = R(t_0, T)$. Let X be any positive $\mathcal{F}_{t_0}$-measurable random variable and M be any martingale with $M_{t_0} = 0$. Then we have the following inequalities.*

(i) The VIX future price $u^f = \mathbb{E}(\sqrt{\mathbb{E}(R \,|\, \mathcal{F}_{t_0})})$ satisfies

$$\begin{aligned} \mathbb{E}\left(\frac{R}{2\sqrt{X}} + \frac{\sqrt{X}}{2}\right) &\geq u^f \\ &\geq \mathbb{E}\left(\sqrt{(R - M_T)^+}\right) - \sqrt{\mathbb{E}\left(\sqrt{\max(R, M_T)} - \sqrt{R}\right)^2}, \end{aligned} \tag{14}$$

where $x^+ = \max(x, 0)$. Equalities are achieved when $X = \mathbb{E}(R \,|\, \mathcal{F}_{t_0})$ and $M_T = R - \mathbb{E}(R \,|\, \mathcal{F}_{t_0})$.

(ii) Fix $K > 0$, the VIX cap price $u^c = \mathbb{E}(\min(\sqrt{\mathbb{E}(R \,|\, \mathcal{F}_{t_0})}, K))$ satisfies

$$\begin{aligned} \mathbb{E}\left(\left(\frac{R}{2\sqrt{X}} + \frac{\sqrt{X}}{2}\right)\mathbb{1}(X \leq K^2) + K\mathbb{1}(X > K^2)\right) &\geq u^c \\ &\geq \mathbb{E}\left(\min\left(\sqrt{(R - M_T)^+}, K\right)\right) - \sqrt{\mathbb{E}\left(\sqrt{\max(R, M_T)} - \sqrt{R}\right)^2}. \end{aligned} \tag{15}$$

Equalities are again achieved when $X = \mathbb{E}(R \,|\, \mathcal{F}_{t_0})$ and $M_T = R - \mathbb{E}(R \,|\, \mathcal{F}_{t_0})$.

Proof. (i) The concave conjugate of $\sqrt{x}$ gives the following representation,

$$\sqrt{x} = \inf_{y>0}\left(xy + \frac{1}{4y}\right), \tag{16}$$

where the infimum is achieved by $y^* = \frac{1}{2\sqrt{x}}$. Then by Theorem 3 (i), for any positive $\mathcal{F}_{t_0}$-measurable random variable Y, we have

$$\mathbb{E}\left(\sqrt{\mathbb{E}(R \,|\, \mathcal{F}_{t_0})}\right) \leq \mathbb{E}\left(RY + \frac{1}{4Y}\right).$$

The upper bound in (14) follows from the substitution $Y = \frac{1}{2\sqrt{X}}$.

For the lower bound, first note the identity

$$\mathbb{E}(\max(R, M_T) \,|\, \mathcal{F}_{t_0}) = \mathbb{E}((R - M_T)^+ + M_T \,|\, \mathcal{F}_{t_0}) = \mathbb{E}((R - M_T)^+ \,|\, \mathcal{F}_{t_0}). \tag{17}$$

Now the required bound can be derived as follows,

$$\mathbb{E}(\sqrt{\mathbb{E}(R\,|\,\mathcal{F}_{t_0})}) \tag{18}$$

$$\geq \mathbb{E}\left(\sqrt{\mathbb{E}(\max(R, M_T)\,|\,\mathcal{F}_{t_0})} - \sqrt{\mathbb{E}((\sqrt{\max(R, M_T)} - \sqrt{R})^2\,|\,\mathcal{F}_{t_0})}\right) \tag{19}$$

$$= \mathbb{E}\left(\sqrt{\mathbb{E}((R - M_T)^+\,|\,\mathcal{F}_{t_0})} - \sqrt{\mathbb{E}((\sqrt{\max(R, M_T)} - \sqrt{R})^2\,|\,\mathcal{F}_{t_0})}\right) \tag{20}$$

$$\geq \mathbb{E}\left(\sqrt{(R - M_T)^+}\right) - \sqrt{\mathbb{E}\left(\sqrt{\max(R, M_T)} - \sqrt{R}\right)^2}. \tag{21}$$

The first inequality (18) follows from Minkowski's inequality with conditional expectations, while the last inequality (21) follows from Jensen's inequality. Note that we switched from $\max(R, M_T)$ to $(R - M_T)^+$ in (20) since the latter typically has lower variance for desirable choices of M_T (i.e., for $M_T \approx R - \mathbb{E}(R\,|\,\mathcal{F}_{t_0})$), leading to a tighter Jensen's inequality. The equality cases can be easily checked via substitution.

(ii) The VIX cap case is similar to (i) with a few adjustments. The concave conjugate of the function $\min(\sqrt{x}, K)$ gives the following representation,

$$\min(\sqrt{x}, K) = \inf_{y \geq 0} \left(xy + h(y)\right) \tag{22}$$

where

$$h(y) := \begin{cases} \dfrac{1}{4y}, & \text{if } y \geq \dfrac{1}{2K}, \\ K - K^2 y, & \text{if } y < \dfrac{1}{2K}. \end{cases}$$

The infimum in (22) is achieved by $y^* = \frac{1}{2\sqrt{x}}\mathbb{1}(x \leq K^2)$.

Again applying Proposition 3 (i), we have the upper bound

$$\mathbb{E}\left(\min(\sqrt{\mathbb{E}(R\,|\,\mathcal{F}_{t_0})}, K)\right) \leq \mathbb{E}\left(RY + h(Y)\right).$$

This simplifies to the required upper bound in (15) after the substitution $Y = \frac{1}{2\sqrt{X}}\mathbb{1}(X \leq K^2)$.

The lower bound can be established by using the same argument as (18)–(21) in (i), combined with the inequality

$$\sqrt{\mathbb{E}(\max(R, M_T)\,|\,\mathcal{F}_{t_0})} - \sqrt{\mathbb{E}(R\,|\,\mathcal{F}_{t_0})}$$
$$\geq \min(\sqrt{\mathbb{E}(\max(R, M_T)\,|\,\mathcal{F}_{t_0})}, K) - \min(\sqrt{\mathbb{E}(R\,|\,\mathcal{F}_{t_0})}, K).$$

Note that we have used the fact that $\max(R, M_T) \geq R$. Finally, the equality conditions can be checked by substitution. □ □

A key feature of the upper and lower bounds presented in Theorem 5 is that they can all be computed using a standard Monte Carlo simulation. The lower bound in (14) can be computed even if $\mathbb{P}(M_t > R) > 0$. In the case where $R \geq M_t$ holds almost surely, it reduces to the simpler bound in (13), $\mathbb{E}(\sqrt{R - M_T})$.

Remark 6. As an immediate consequence of Jensen's inequality, the value of the VIX future is bounded between the volatility swap and the square root of the variance swap, both evaluated at time 0,

$$\mathbb{E}(\sqrt{R}) \leq u^f \leq \sqrt{\mathbb{E}R}.$$

Both of these bounds can be seen as special cases of Theorem 5 (i), by setting X to be the variance swap $\mathbb{E}R$ evaluated at time 0 and by setting M_T to zero. Also, it is noteworthy that equality is reached in Theorem 5 when X is the variance swap evaluated at time t_0 and M is hedging portfolio of the same variance swap during $[t_0, T]$. In practical implementations, if M_T is poorly estimated and $M_t > R$ occurs frequently, it may be more advantageous to simply use $\mathbb{E}(\sqrt{R})$ as a lower bound instead.

4 Least Squares Monte Carlo

In this section, we shall describe the empirical Monte Carlo algorithm used to compute bounds for VIX derivatives. The algorithm utilises a variant of the least squares Monte Carlo proposed by Schoenmakers et al. [14] which simultaneously estimates the conditional expectation as well as the martingale increment. We refer the readers to Schoenmakers et al. [14] for results regarding stability and convergence of the method, as well as Joshi and Tang [15] for an overview of related methods.

For the least squares Monte Carlo algorithm, let us focus on the following local-stochastic volatility (LSV) model,

$$\begin{aligned} dS_t &= \mu(t, S_t) S_t dt + \sigma(t, S_t, V_t) S_t dW_t^S, \\ dV_t &= a(t, V_t) dt + b(t, V_t) dW_t^V, \\ \langle dW_t^S, dW_t^V \rangle &= \rho(t, S_t, V_t) dt, \end{aligned}$$

where W_t^S and W_t^V are standard Brownian motions. Suppose that the time interval $[t_0, T]$ is partitioned into $t_0 < t_1 < \cdots < t_n = T$. Recall that, by Theorem 5, in order to obtain good quality bounds on VIX derivatives, it is important to find good approximations to the conditional expectation $X = \mathbb{E}(R \,|\, \mathcal{F}_{t_0})$ and the martingale increment $M_T = R - \mathbb{E}(R \,|\, \mathcal{F}_{t_0})$. We postulate that X and M_T can be approximated in terms of the state variables in the following way:

$$X = \mathbb{E}(R \,|\, \mathcal{F}_{t_0}) \approx \Psi(S_{t_0}, V_{t_0}), \tag{23}$$

$$M_T = R - \mathbb{E}(R \,|\, \mathcal{F}_{t_0}) = \sum_{l=0}^{n-1} \mathbb{E}(R \,|\, \mathcal{F}_{t_{l+1}}) - \mathbb{E}(R \,|\, \mathcal{F}_{t_l}) \tag{24}$$

$$\approx \sum_{l=0}^{n-1} \Phi_{t_l}(S_{t_l}, V_{t_l}) \cdot \Delta W_{t_l}, \tag{25}$$

where

$$\Psi(s, v) := \sum_{j=1}^{p} \beta_j \psi_j(s, v), \tag{26}$$

$$\Phi_{t_l}(s, v) := \sum_{j=1}^{q} \gamma_{j,l} \phi_j(s, v). \tag{27}$$

Here $\psi_j : \mathbb{R}^2 \to \mathbb{R}$ and $\phi_j : \mathbb{R}^2 \to \mathbb{R}^2$ are basis functions chosen beforehand. Note that $\Delta W_{t_l} := (W_{t_{l+1}}^S - W_{t_l}^S, W_{t_{l+1}}^V - W_{t_l}^V)'$. In practice ΔW_{t_l} can be replaced by any other suitable martingale increment with the predictable representation property.

Remark 7. Due to the Markovian properties of the model and the predictable representation theorem, if the space spanned by the basis functions is rich enough, the conditional expectation can be matched exactly while the martingale increment will be replicated as the mesh of the partition goes to 0,

$$\mathbb{E}(R \,|\, \mathcal{F}_{t_0}) = \Psi(S_{t_0}, V_{t_0}), \quad R - \mathbb{E}(R \,|\, \mathcal{F}_{t_0}) = \int_{t_0}^{T} \Phi_t(S_t, V_t) \cdot dW_t.$$

Now we will describe the numerical algorithm. First simulate N trajectories S^i and V^i for $i = 1, \dots, N$, and compute the corresponding realised variances R^i. Next, the coefficients

$$B := (\beta_j : j = 1, \dots, p), \quad \Gamma := (\gamma_{j,l} : j = 1, \dots, q; l = 1, \dots, n)$$

are estimated in the linear least squares regression problem:

$$\begin{aligned}(\hat{B}, \hat{\Gamma}) &= \underset{B \in \mathbb{R}^p, \Gamma \in \mathbb{R}^{q \times n}}{\arg\min} \sum_{i=1}^{N} \left(R^i - \Psi(S^i_{t_0}, V^i_{t_0}) - \sum_{l=0}^{n-1} \Phi_{t_l}(S^i_{t_l}, V^i_{t_l}) \cdot \Delta W^i_{t_l} \right)^2 \\ &= \underset{B \in \mathbb{R}^p, \Gamma \in \mathbb{R}^{q \times n}}{\arg\min} \sum_{i=1}^{N} \left(R^i - \sum_{j=1}^{p} \beta_j \psi_j(S^i_{t_0}, V^i_{t_0}) \right. \\ &\qquad \left. - \sum_{l=0}^{n-1} \sum_{j=1}^{q} \gamma_{j,l} \phi_j(S^i_{t_l}, V^i_{t_l}) \cdot \Delta W^i_{t_l} \right)^2 .\end{aligned}$$

Let us denote the estimated functions by

$$\hat{\Psi}(s, v) = \sum_{j=1}^{p} \hat{\beta}_j \psi_j(s, v), \quad \hat{\Phi}_{t_l}(s, v) = \sum_{j=1}^{q} \hat{\gamma}_{j,l} \phi_j(s, v).$$

Remark 8. Since we don't require any additional conditional expectations aside from the one at t_0, we have chosen to perform one single global regression, instead of one regression per time step as described in Schoenmakers et al. [14]. In test cases, the global regression performed slightly better for our purposes, although it would be more computationally expensive if the number of time steps and basis functions are very large.

In order to compute true upper and lower bounds, we generate a new set of $\tilde{N}$ trajectories $\tilde{S}^i$ and $\tilde{V}^i$ for $i = 1, \dots, \tilde{N}$. This is performed to avoid the foresight bias caused by reusing the original trajectories. A detailed explanation of the foresight bias can be found in Fries [19]. Our new path-wise estimates of the conditional expectation and the martingale increment are

$$\hat{X}^i = \hat{\Psi}(\tilde{S}^i_{t_0}, \tilde{V}^i_{t_0}), \quad \hat{M}^i_T = \sum_{l=0}^{n-1} \hat{\Phi}_{t_l}(\tilde{S}_{t_l}, \tilde{V}_{t_l}) \cdot \Delta \tilde{W}^i_{t_l}.$$

At this point we apply Theorem 5 on the estimates $\hat{X}^i$ and $\hat{M}^i_T$ to produce bounds for the VIX future and cap. Specifically, for VIX futures we have

$$\overline{u}^f = \frac{1}{\tilde{N}} \sum_{i=1}^{\tilde{N}} \left(\frac{\tilde{R}^i}{2\sqrt{\hat{X}^i}} + \frac{\sqrt{\hat{X}^i}}{2} \right), \tag{28}$$

$$\underline{u}^f = \frac{1}{\tilde{N}} \sum_{i=1}^{\tilde{N}} \left(\sqrt{(\tilde{R}^i - \hat{M}^i_T)^+} \right) - \sqrt{\frac{1}{\tilde{N}} \sum_{i=1}^{\tilde{N}} \left(\sqrt{\max(\tilde{R}^i, \hat{M}^i_T)} - \sqrt{\tilde{R}^i} \right)^2}, \tag{29}$$

while for VIX caps we have

$$\overline{u}^c = \frac{1}{\tilde{N}} \sum_{i=1}^{\tilde{N}} \left(\left(\frac{\tilde{R}^i}{2\sqrt{\hat{X}^i}} + \frac{\sqrt{\hat{X}^i}}{2} \right) \mathbb{1}(\hat{X}^i \leq K^2) + K \mathbb{1}(\hat{X}^i > K^2) \right), \tag{30}$$

$$\underline{u}^c = \frac{1}{\tilde{N}} \sum_{i=1}^{\tilde{N}} \left(\min \left(\sqrt{(\tilde{R}^i - \hat{M}^i_T)^+}, K \right) \right) \tag{31}$$

$$- \sqrt{\frac{1}{\tilde{N}} \sum_{i=1}^{\tilde{N}} \left(\sqrt{\max(\tilde{R}^i, \hat{M}^i_T)} - \sqrt{\tilde{R}^i} \right)^2}. \tag{32}$$

Note that the realised variances $\tilde{R}^i$ are directly computed from $\tilde{S}^i$ and $\tilde{V}^i$. Bounds for other derivatives such as swaps, calls and puts can now be easily computed as follows:

$$\overline{u}^{swap} = \overline{u}^f - K, \quad \underline{u}^{swap} = \underline{u}^f - K, \tag{33}$$

$$\overline{u}^{call} = \overline{u}^f - \underline{u}^c, \quad \underline{u}^{call} = \underline{u}^f - \overline{u}^c, \tag{34}$$

$$\overline{u}^{put} = K - \underline{u}^c, \quad \underline{u}^{put} = K - \overline{u}^c. \tag{35}$$

Remark 9. At a first glance, the term $\sqrt{\hat{X}^i}$ in the upper bound calculation could cause problems since $\hat{X}^i$ may be negative. In practical implementations, a floor is often imposed on the instantaneous volatility. It is then natural to enforce the same floor on $\hat{X}^i$,

$$\hat{X}^i = \max(\hat{\Psi}(\tilde{S}^i_{t_0}, \tilde{V}^i_{t_0}), h).$$

The result will still be a true upper bound. This is in contrast to the lower bound term $\sqrt{R - M_T}$ where the sign of $R - M_T$ is more difficult to control. A simple floor on $R - M_T$ will violate the validity of the lower bound. Hence Theorem 5 was necessary to overcome this issue. In general, these issues can also be alleviated by using more and better basis functions, thus improving the least squares fit.

5 Numerical Results

For our numerical example, we choose the following variant of the CEV-Heston LSV model with volatility caps and floors:

$$\begin{aligned} dS_t &= \sigma(S_t, V_t) S_t dW_t^S, \\ dV_t &= \kappa(\theta - V_t)dt + \eta\sqrt{V_t} dW_t^V, \\ \sigma(S_t, V_t) &= f(\sqrt{V_t}(S_t/S_0)^{\alpha - 1}), \\ f(x) &= \max(\min(x, 10), 0.01), \\ \langle dW_t^S, dW_t^V \rangle &= \rho dt. \end{aligned}$$

This is essentially the same as the usual CEV-Heston model, but the effective volatility is bounded between 0.01 and 10. Recall that the interest rate is assumed to be zero. Table 1 contains our chosen parameter values as well as their interpretations.

We will be employing the algorithm described in Section 4 to compute bounds for VIX futures, caps, calls and puts. The simulation scheme used will be the standard Euler scheme with full truncation. Although there are better numerical schemes for this model, the differences would be minimal as the time step used is very small ($\Delta t = 1/120$). Antithetic variables are used for variance reduction. During the regression step, the following basis functions are used:

$$\begin{aligned} \Psi(s, v) &:= \sum_{j=1}^{p} \beta_j \psi_j(\log s, \sqrt{v}), \\ \Phi_{t_l}(s, v) &:= \sum_{j=1}^{p} \gamma_{j,l} \phi_j(\log s, \sqrt{v}) \left(\sigma(s, v) s \frac{d}{ds} \log s, \eta\sqrt{v} \frac{d}{dv} \sqrt{v} \right)' \\ &= \sum_{j=1}^{p} \gamma_{j,l} \phi_j(\log s, \sqrt{v}) \left(\sigma(s, v), \frac{\eta}{2} \right)', \end{aligned}$$

where ψ_j and ϕ_j are bivariate polynomials. Two cases are examined: lower degree polynomials where ψ_j and ϕ_j have degrees 3 and 2 respectively, and higher degree polynomials where

Table 1: Parameter values and interpretations

Parameter	Value	Interpretation
S_0	100	initial stock price
α	0.8	leverage between stock and volatility
$\sigma(S_0, V_0)$	0.3	initial volatility
V_0	0.09	initial variance
κ	0.6	mean-reversion speed
θ	0.09	long term variance
η	0.4	vol of vol
ρ	-0.5	correlation between stock and variance
t_0	1	VIX start date
T	1+1/12	VIX end date
Δt	1/120	time increment
N	100000	paths for regression
$\tilde{N}$	500000	paths for bound calculation

ψ_j and ϕ_j have degrees 4 and 3 respectively. During the computation of the upper bounds, the volatility cap and floor function (i.e., f) is also applied to $\hat{X}$. In the computation of the lower bounds, the martingale increments can be interpreted as the delta and vega hedging strategies.

As a benchmark, we will also be showing the results of a nested Monte Carlo. In this simulation, 500000 trajectories are generated up to time t_0. Then for each of these trajectories, a sub-simulation of 5000 trajectories is carried out on the time interval $[t_0, T]$ to compute the conditional expectation $\mathbb{E}(R \,|\, \mathcal{F}_{t_0})$ path-wise. The number of paths in the sub-simulation is chosen based on a comparison of the sample variances generated during the time intervals $[0, t_0]$ and $[t_0, T]$. Finally, the prices of the VIX derivatives are computed by averaging the relevant payoffs over all trajectories. In order to check the correctness of our bounds, the same 500000 paths on $[0, t_0]$ from the nested Monte Carlo will also be used in the second simulation of our least squares Monte Carlo. After that, the behaviour of the paths on $[t_0, T]$ are generated independently for the different methods. This allows us to compare the relative sizes of the results without the effects of variances due to simulation. In terms of computation times, the least squares Monte Carlo method took 17 seconds, which is more than 1200 times faster than the nested method which took 21154 seconds or 5.876 hours.

The results for lower degree polynomials are found in Tables 2 while the higher degree polynomials results are found in Table 3. In terms of notations, for VIX futures: u^f is the value computed using the nested Monte Carlo; $\hat{u}^f$ is the result of the classic least squares Monte Carlo by simply averaging the square root of the regression fit $\hat{\Psi}(\tilde{S}^i_{t_0}, \tilde{V}^i_{t_0})$; $\underline{u}^f$ and $\overline{u}^f$ are the lower and upper bounds computed as described in (28) and (29). For completeness, we have also included estimates for the volatility swap $\mathbb{E}\sqrt{R}$ and the square root of the variance swap $\sqrt{\mathbb{E}R}$. Similar notations are used for calls and puts over a range of strikes K. All confidence intervals are computed as 1.96 times the standard deviation. All values have also been annualised accordingly.

As shown in Table 2, even with lower degree polynomials, our method produces tight bounds across all VIX derivatives and at all strike levels. In many cases the classical least squares Monte Carlo estimates actually fall outside of our bounds. Our bounds are also clearly superior when compared to the bounds given by the volatility and variance swaps (see Remark 6). In the higher degree polynomials case shown in Table 3, the convergence of our method is verified by the fact that all four estimates are extremely close. In fact, the difference between the estimates is much smaller than the corresponding confidence intervals.

Table 2: Lower degree polynomials results for VIX futures, calls and puts, including nested Monte Carlo results, least square Monte Carlo estimates, as well as lower and upper bounds

	u^f	$\hat{u}^f$	$\underline{u}^f$	$\overline{u}^f$
	27.3728 ± 0.0445	27.3500 ± 0.0446	27.3582 ± 0.0445	27.4607 ± 0.0454
			$\mathbb{E}\sqrt{R} = 27.1018$	$\sqrt{\mathbb{E}R} = 31.7342$
K	u^{call}	$\hat{u}^{call}$	$\underline{u}^{call}$	$\overline{u}^{call}$
15	13.7302 ± 0.0404	13.7417 ± 0.0403	13.6269 ± 0.0676	13.8326 ± 0.0415
20	10.2909 ± 0.0367	10.2887 ± 0.0367	10.1892 ± 0.0817	10.3932 ± 0.0380
25	7.4785 ± 0.0324	7.4672 ± 0.0324	7.3775 ± 0.0939	7.5809 ± 0.0339
30	5.2738 ± 0.0280	5.2600 ± 0.0280	5.1716 ± 0.1021	5.3763 ± 0.0297
35	3.6176 ± 0.0236	3.6065 ± 0.0236	3.5156 ± 0.1052	3.7202 ± 0.0257
40	2.4230 ± 0.0196	2.4169 ± 0.0196	2.3208 ± 0.1034	2.5256 ± 0.0221
45	1.5912 ± 0.0160	1.5900 ± 0.0161	1.4887 ± 0.0978	1.6938 ± 0.0191
K	u^{put}	$\hat{u}^{put}$	$\underline{u}^{put}$	$\overline{u}^{put}$
15	1.3575 ± 0.0079	1.3916 ± 0.0083	1.2686 ± 0.0091	1.3719 ± 0.0079
20	2.9181 ± 0.0131	2.9386 ± 0.0134	2.8310 ± 0.0141	2.9326 ± 0.0131
25	5.1057 ± 0.0185	5.1172 ± 0.0187	5.0193 ± 0.0193	5.1203 ± 0.0185
30	7.9010 ± 0.0236	7.9100 ± 0.0238	7.8134 ± 0.0245	7.9157 ± 0.0236
35	11.2449 ± 0.0282	11.2565 ± 0.0284	11.1574 ± 0.0291	11.2595 ± 0.0283
40	15.0502 ± 0.0322	15.0669 ± 0.0323	14.9625 ± 0.0331	15.0650 ± 0.0322
45	19.2184 ± 0.0354	19.2400 ± 0.0355	19.1305 ± 0.0363	19.2331 ± 0.0354

Table 3: Higher degree polynomials results for VIX futures, calls and puts, including nested Monte Carlo results, least square Monte Carlo estimates, as well as lower and upper bounds

	u^f	$\hat{u}^f$	$\underline{u}^f$	$\overline{u}^f$
	27.3728 ± 0.0445	27.3739 ± 0.0445	27.3707 ± 0.0445	27.3751 ± 0.0455
			$\mathbb{E}\sqrt{R} = 27.1018$	$\sqrt{\mathbb{E}R} = 31.7342$
K	u^{call}	$\hat{u}^{call}$	$\underline{u}^{call}$	$\overline{u}^{call}$
15	13.7302 ± 0.0404	13.7313 ± 0.0404	13.7265 ± 0.0673	13.7346 ± 0.0415
20	10.2909 ± 0.0367	10.2921 ± 0.0367	10.2883 ± 0.0814	10.2952 ± 0.0380
25	7.4785 ± 0.0324	7.4793 ± 0.0324	7.4756 ± 0.0936	7.4828 ± 0.0338
30	5.2738 ± 0.0280	5.2741 ± 0.0280	5.2701 ± 0.1018	5.2781 ± 0.0296
35	3.6176 ± 0.0236	3.6173 ± 0.0236	3.6140 ± 0.1050	3.6220 ± 0.0255
40	2.4230 ± 0.0196	2.4224 ± 0.0196	2.4191 ± 0.1034	2.4274 ± 0.0218
45	1.5912 ± 0.0160	1.5904 ± 0.0160	1.5871 ± 0.0978	1.5955 ± 0.0187
K	u^{put}	$\hat{u}^{put}$	$\underline{u}^{put}$	$\overline{u}^{put}$
15	1.3575 ± 0.0079	1.3574 ± 0.0079	1.3558 ± 0.0088	1.3595 ± 0.0079
20	2.9181 ± 0.0131	2.9182 ± 0.0131	2.9176 ± 0.0140	2.9202 ± 0.0131
25	5.1057 ± 0.0185	5.1054 ± 0.0185	5.1049 ± 0.0194	5.1077 ± 0.0185
30	7.9010 ± 0.0236	7.9001 ± 0.0236	7.8994 ± 0.0246	7.9031 ± 0.0236
35	11.2449 ± 0.0282	11.2434 ± 0.0282	11.2432 ± 0.0292	11.2469 ± 0.0282
40	15.0502 ± 0.0322	15.0484 ± 0.0322	15.0484 ± 0.0332	15.0523 ± 0.0322
45	19.2184 ± 0.0354	19.2165 ± 0.0354	19.2164 ± 0.0364	19.2205 ± 0.0354

This indicates that, in terms of the bias-variance trade-off, most of the error comes from the variance caused by the simulation, while our method with higher degree polynomials has very little bias due to an excellent regression fit. Figure 1 further illustrates this by plotting the VIX future bounds as the number of simulation paths varies. As the number of paths increases, the bounds stabilise towards their limits. The higher degree polynomials results are noticeably better than the lower degree polynomials results, especially in the upper bounds which benefited greatly from the degree of ψ_j increasing from 3 to 4.

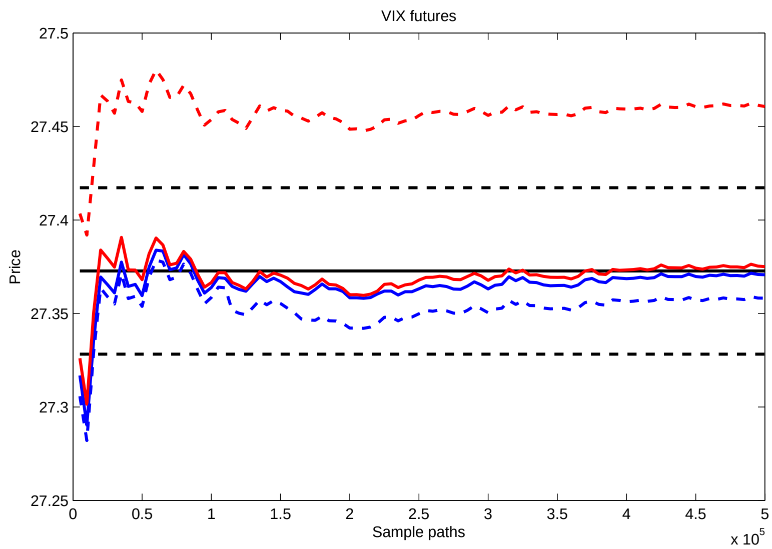

Figure 1: Plot of VIX future bounds for different number of simulation paths. The solid black line is the nested Monte Carlo result, with its confidence interval indicated by the dashed black lines. The solid red and blue lines are the upper and lower bounds using higher degree polynomials. The dashed red and blue lines are the upper and lower bounds using lower degree polynomials.

Now we examine the effect of varying a single parameter on VIX futures. The other parameters are kept as in Table 1 and lower degree polynomials are used. First of all, we vary the correlation coefficient ρ. As shown in Table 4, larger values of ρ lead to lower VIX future prices. Our method works very well in all cases, especially for higher correlations. This is due to the fact that a negative correlation combined with a leverage coefficient satisfying $\alpha < 1$ will lead to larger variations in the realised variance.

Table 4: VIX futures for different correlation values

ρ	$\hat{u}^f$	$\underline{u}^f$	$\overline{u}^f$
-0.8	27.6457 ± 0.0471	27.6439 ± 0.0469	27.7969 ± 0.0478
-0.6	27.4260 ± 0.0453	27.4411 ± 0.0452	27.5687 ± 0.0461
-0.4	27.2558 ± 0.0439	27.2629 ± 0.0439	27.3397 ± 0.0448
-0.2	27.0511 ± 0.0423	27.0549 ± 0.0423	27.0924 ± 0.0432
0.0	26.8643 ± 0.0409	26.8646 ± 0.0408	26.8888 ± 0.0418
0.2	26.6778 ± 0.0395	26.6776 ± 0.0395	26.6911 ± 0.0404
0.4	26.4776 ± 0.0381	26.4784 ± 0.0381	26.4867 ± 0.0390
0.6	26.2928 ± 0.0368	26.2955 ± 0.0368	26.3014 ± 0.0377
0.8	26.1158 ± 0.0355	26.1201 ± 0.0355	26.1235 ± 0.0365

Next, we vary the vol of vol η in Table 5. As η increases, the VIX future price decreases. For small values of η, the upper and lower bounds are essentially the same value. For extremely large values of η, the quality of the lower bound deteriorates substantially.

Table 5: VIX futures for different vol of vol

η	$\hat{u}^f$	$\underline{u}^f$	$\overline{u}^f$
0.1	30.2498 ± 0.0140	30.2497 ± 0.0140	30.2500 ± 0.0143
0.2	29.6859 ± 0.0245	29.6864 ± 0.0245	29.6864 ± 0.0250
0.3	28.6812 ± 0.0349	28.6876 ± 0.0349	28.6972 ± 0.0356
0.4	27.3309 ± 0.0445	27.3381 ± 0.0444	27.4319 ± 0.0453
0.5	25.9168 ± 0.0531	25.8331 ± 0.0530	26.2101 ± 0.0541
0.6	24.5632 ± 0.0605	24.2752 ± 0.0607	25.0365 ± 0.0621
0.7	23.5007 ± 0.0668	22.5781 ± 0.0676	24.4188 ± 0.0719
0.8	22.5599 ± 0.0726	21.2048 ± 0.0742	23.5922 ± 0.0806

Finally, Table 6 examines the effect of varying the leverage coefficient α. The bounds deteriorate somewhat for small values of α. This is due to the negative correlation ρ, which creates more extreme values of the realised variance for small values of α. The reverse would be true if ρ was positive.

Table 6: VIX futures for different leverage coefficients

α	$\hat{u}^f$	$\underline{u}^f$	$\overline{u}^f$
0.7	27.9683 ± 0.0492	26.5554 ± 0.0492	28.5024 ± 0.0506
0.8	27.3445 ± 0.0445	27.3564 ± 0.0445	27.4246 ± 0.0453
0.9	26.8659 ± 0.0414	26.8684 ± 0.0414	26.8716 ± 0.0424
1.0	26.4738 ± 0.0392	26.4725 ± 0.0392	26.4735 ± 0.0402
1.1	26.1138 ± 0.0373	26.1141 ± 0.0373	26.1145 ± 0.0383
1.2	25.8070 ± 0.0360	25.8049 ± 0.0360	25.8162 ± 0.0370
1.3	25.5235 ± 0.0350	25.5291 ± 0.0349	25.5564 ± 0.0358
1.4	25.2905 ± 0.0342	25.2924 ± 0.0341	25.3685 ± 0.0350
1.5	25.0769 ± 0.0335	25.0721 ± 0.0335	25.1879 ± 0.0344

Even though lower degree polynomials are used in Tables 4, 5 and 6, our method generally works very well. In fact, in many cases the bounds are even better than the direct estimates $\hat{u}^f$ obtained from the classical least squares regression approach. Since the tightness of our bounds depends on the quality of the regression fit, the method understandably performs worse when there are extreme variations in the realised variance. This is particularly noticeable for the lower bounds since a poor regression fit often leads to frequent occurrences of $M_T > R$. In these extreme cases, the results can be improved by using better basis functions. Alternatively, one may also use the volatility swap $\mathbb{E}\sqrt{R}$ as a replacement lower bound.

6 Conclusions

We have introduced a new model independent technique for the computation of true upper and lower bounds for VIX derivatives. Theorem 3 includes a general stochastic duality result on payoffs involving convex (or concave) functions. In particular, convex (or concave) conjugates are used to transform the contingent claims into a type of chooser options. This is then applied to VIX derivatives in Theorem 5, along with minor adjustments to handle issues caused by the square root function. The upper bound involves the evaluation of a variance swap, while the lower bound involves estimating a martingale increment corresponding to its hedging portfolio. Our bounding technique is particularly useful in complex models where it is difficult to directly compute VIX derivative prices. Numerically, a single linear least squares Monte Carlo method is used to simultaneously compute the upper and lower bounds. The method is shown to work very well for VIX futures, calls and puts under a wide range of parameter choices.

References

[1] Whaley, R. E.: Derivatives on market volatility: Hedging tools long overdue. The Journal of Derivatives, 1(1), 71–84 (1993)

[2] Grünbichler, A., Longstaff, F. A.: Valuing futures and options on volatility. Journal of Banking & Finance, 20(6), 985–1001 (1996)

[3] Detemple, J., Osakwe, C.: The valuation of volatility options. European Finance Review, 4(1), 21–50 (2000)

[4] Zhang, J. E., Zhu, Y.: VIX futures. Journal of Futures Markets, 26(6), 521–531 (2006)

[5] Lian, G. H., Zhu, S. P.: Pricing VIX options with stochastic volatility and random jumps. Decisions in Economics and Finance, 36(1), 71–88 (2013)

[6] Sepp, A.: VIX option pricing in a jump-diffusion model. Risk magazine, 84–89 (2008)

[7] Baldeaux, J., Badran, A.: Consistent modelling of VIX and equity derivatives using a 3/2 plus jumps model. Applied Mathematical Finance, 21(4), 299–312 (2014)

[8] Cont, R., Kokholm, T.: A consistent pricing model for index options and volatility derivatives. Mathematical Finance, 23(2), 248–274 (2013)

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

[10] Longstaff, F. A., Schwartz, E. S.: Valuing American options by simulation: a simple least-squares approach. Review of Financial studies, 14(1), 113-147 (2001)

[11] Rogers, L. C.: Monte Carlo valuation of American options. Mathematical Finance, 12(3), 271-286 (2002)

[12] Haugh, M. B., Kogan, L.: Pricing American options: a duality approach. Operations Research, 52(2), 258-270 (2004)

[13] Andersen, L., Broadie, M.: Primal-dual simulation algorithm for pricing multidimensional American options. Management Science, 50(9), 1222-1234 (2004)

[14] Schoenmakers, J., Zhang, J., Huang, J.: Optimal dual martingales, their analysis, and application to new algorithms for Bermudan products. SIAM Journal on Financial Mathematics, 4(1), 86–116 (2013)

[15] Joshi, M., Tang, R.: Effective sub-simulation-free upper bounds for the Monte Carlo pricing of callable derivatives and various improvements to existing methodologies. Journal of Economic Dynamics and Control, 40, 25–45 (2014)

[16] De Marco, S., Henry-Labordere, P.: Linking vanillas and VIX options: a constrained martingale optimal transport problem. SIAM Journal on Financial Mathematics, 6(1), 1171–1194 (2015)

[17] Guyon, J., Menegaux, R., Nutz, M.: Bounds for VIX futures given S&P 500 smiles. to appear in Finance and Stochastics (2016)

[18] Rockafellar, R. T.: Convex analysis. Princeton university press (2015)

[19] Fries, C. P.: Foresight Bias and Suboptimality Correction in MonteCarlo Pricing of Options with Early Exercise. In: Progress in Industrial Mathematics at ECMI 2006 (pp. 645–649). Springer Berlin Heidelberg (2008)