Super-replication of financial derivatives via convex programming

Nabil Kahalé *

First version: November 7, 2012
This version: March 22, 2016

Abstract

We give a method based on convex programming to calculate the optimal super-replicating and sub-replicating prices and corresponding hedging portfolios of a financial derivative in terms of other financial derivatives in a discrete-time setting. Our method produces a model that matches the super-replicating (or sub-replicating) price within an arbitrary precision and is consistent with the other financial derivatives prices. Applications include robust replication in terms of call prices with various strikes and maturities of forward start options, volatility and variance swaps and derivatives, cliquets calls, barrier options, lookback and Asian options. Numerical examples show that, in some cases, the best super-replicating and/or sub-replicating prices are within 10% of the price obtained by a standard model, but considerably differ from it in other cases. Our method can incorporate bid-ask spreads, interest rates and dividends and various limitations to the diffusion model.

Keywords: Model risk, robust replication, robust hedging, convex programming, financial derivatives.

1 Introduction

Several models based on local volatility, stochastic volatility or jump diffusion (see, e.g., (Hull 2012)) have been used to price financial derivatives. However, even if these models exactly fit the current market prices of liquid financial products, such as vanilla call and put options, they may produce different prices for other products such as barrier options (Britten-Jones and A. Neuberger 2000). This gives rise to model risk in the pricing of financial derivatives, of which practitioners are well aware (see (Committee on Banking Supervision 2009, p. 29)). This model risk can be assessed through the calculation of model-independent bounds on the derivative price: the larger the gap between the upper and lower bound, the larger the risk. Another possible application of model-independent bounds comes from the observation that if a bank sells (resp. buys) a derivative at a model-independent upper-bound (resp. lower-bound) on its price, it can realize a non-negative net profit by using proper hedging, independently of the future underlying securities behavior. The range of possible arbitrage free prices for a given derivative security ψ is large in general (Carassus, Gobet and Temam 2007) but can be narrowed if the prices of liquid derivatives, such as call prices, are known at time 0. Bounds on option prices in terms of other financial derivatives prices have been derived in the literature in a static setting as well as in a dynamic setting, i.e. when dynamic trading in the stock is allowed at future times.

In a static setting, (Boyle and Lin 1997) have derived a semi-definite upper-bound on a call on the maximum of several assets in terms of their means and correlations. (Bertsimas and Popescu 2002, Gotoh and Konno 2002) have used semi-definite programming to optimally

*ESCP Europe and Labex ReFi, 79 avenue de la République, 75011 Paris, France; e-mail: nkahale@escpeurope.eu.

super-replicate call options in single-period markets on a single asset given several moments of the underlying asset. Tight bounds on prices of basket options (Hobson, Laurence and Wang 2005a, Hobson, Laurence and Wang 2005b, Laurence and Wang 2005, d'Aspremont and El Ghaoui 2006) and on spread options (Laurence and Wang 2008, Laurence and Wang 2009) have been established in terms of prices of vanilla options with the same maturity.

In a dynamic setting, optimal bounds on specific options have been previously derived under certain conditions. When interest rates are null, tight model-independent bounds on lookback and barrier options have been obtained (Hobson 1998, Brown, Hobson and Rogers 2001b, Brown, Hobson and Rogers 2001a, Hobson and Pedersen 2002, Cox and Obłój 2011a, Cox and Obłój 2011b, Galichon, Henry-Labordere and Touzi 2014, Henry-Labordere, Obloj, Spoida and Touzi 2013) in terms of call prices. Under a continuity assumption on the stock price, numerical schemes for robust hedging of financial derivatives have been given (Bonnans and Tan 2011, Tan and Touzi 2013) in terms of a continuum set of call prices, robust pricing of a continuously-monitored variance swap given prices for a finite number of co-maturing call options has been obtained (Davis, Obloj and Raval 2014), and optimal bounds on continuously-monitored variance options have been derived analytically (Cox and Wang 2013). Explicit bounds on discretely-monitored variance swaps (Kahalé 2014, Hobson and Klimmek 2012) have been established in terms of a continuum set of call prices with the same maturity and shown to be optimal when the number of time-steps goes to infinity. (Henry-Labordere and Touzi 2013) have derived optimal bounds on variance swaps monitored at t_i when call prices are known for maturities t_i, $1 \leq i \leq m$, and all strikes. A tight but non-constructive upper bound (Hobson and Neuberger 2012) on at the money forward-start options and, under certain conditions, a tight and explicit lower bound (Hobson and Klimmek 2015) have been obtained in terms of a continuum set of call prices.

Most previous derivations of explicit bounds construct explicitly associated hedging strategies. Also, they often identify explicit scenarios that lead to the extremum prices. However, explicit optimal bounds on path-dependent derivatives that we are aware of make at least one of the following simplifying assumptions: interest rates are null, call prices are known for a continuum set of strikes, or the underlying variables prices follow a continuous process. These assumptions do not hold in practice, though, and may introduce an important bias in the pricing of a financial derivative. For instance, the price of a continuously-monitored variance swap, when jumps in the stock prices are allowed, may significantly differ from the price obtained under the continuity assumption (Demeterfi, Derman, Kamal and Zou 1999, Broadie and Jain 2008, Carr and Wu 2009, Carr, Lee and Wu 2012). Also, in practice, variance swaps are discretely monitored, but it has been often been assumed in the literature that they are continuously monitored, which introduces another source of bias as shown in (Broadie and Jain 2008). Furthermore, optimal explicit bounds are only known for specific options and not known in general for a portfolio of options, even in two-periods markets. This can be explained by noting that robust super-replication is closely related (Henry-Labordere and Touzi 2013) to the optimal transportation problem, for which no explicit solutions are known in general.

This paper gives a unified methodology based on convex programming to calculate the best super-replicating and sub-replicating prices and corresponding hedging portfolios of a financial derivative ψ in terms of prices of a finite set of l liquid derivatives. Super-replicating (or sub-replicating) strategies for ψ consist of static positions in a zero-coupon bond and in the l derivatives at time 0 combined with dynamic trading in d underlying liquid securities. We assume that interest rates are deterministic and work in a multi-period setting but do not make any continuity assumptions on the underlying variables. Under certain conditions, we show the convex program can be solved efficiently for a large class of financial derivatives. Furthermore, our method produces a model that matches the super-replicating (or sub-replicating) price within an arbitrary precision and is consistent with the other financial derivatives prices. Applications include the calculation of the best super-replicating and sub-replicating prices in terms of call options with different strikes and maturities of a wide variety of derivatives such as

- discretely-monitored, standard and generalized variance swaps including cliquet calls and corridor variance swaps
- discretely-monitored volatility swaps and volatility derivatives
- discretely-monitored Asian options
- discretely monitored lookback and barrier options.

We are not aware of previous model-independent optimal bounds for any of these problems without the simplifying assumptions mentioned earlier. In particular, we give the first optimal bounds on discretely-monitored variance and volatility swaps in terms of a finite set of call prices without any continuity assumption on the stock, and the first optimal bounds on the price of arithmetic Asian options. Model-independent bounds for prices of Asian options have been previously derived (Simon, Goovaerts and Dhaene 2000, Albrecher, Mayer and Schoutens 2008), but they are not optimal. We also give the first optimal bounds on discretely monitored lookback and barrier options in terms of a finite set of call prices without assuming interest rates to be null. Finally, we give optimal robust bounds on forward start options. In (Beiglböck, Henry-Labordère and Penkner 2013), optimal model-independent bounds for forward-start options in terms of a finite set of call prices are calculated in a discrete-state setting using linear programming, but no bound is given on the discretization error. In addition, our convex program can be easily adapted to calculate the best model-free bounds on a financial derivative when bid and ask prices of the liquid financial derivatives are not equal, or the size of the jumps the assets can take are restricted.

A method based on linear programming with a large number of constraints is described in (Henry-Labordère 2013) to robust hedging of derivatives. However, this method requires the amount of basic securities held in a dynamic hedging position to be a function of a certain type (e.g. a polynomial) of the securities values. It also assumes the securities paths to belong to a certain grid generated by a Monte Carlo simulation scheme, and requires the knowledge of a continuum set of call options prices. No estimate on the errors due to these requirements is given in (Henry-Labordère 2013).

In line with previous literature on robust bounds on derivatives that depend on the underlyings' values on a finite set of time-steps (see, e.g., (Simon, Goovaerts and Dhaene 2000, Albrecher, Mayer and Schoutens 2008, Kahalé 2014, Hobson and Neuberger 2012, Hobson and Klimmek 2012, Hobson and Klimmek 2015, Beiglböck, Henry-Labordère and Penkner 2013, Henry-Labordere and Touzi 2013)), we use a discrete-time model with an infinite state-space without reference to a single prior measure. We only use risk-neutral probabilities with finite support, as these are sufficient to derive optimal model-independent bounds in our framework (see Theorem 3.2, and (Kahalé 2010) for related results). In our numerical applications, we discretize the state-space and show how to bound the discretization error.

The remainder of the paper is organised as follows. Definitions and preliminary results are given in Section 2. Section 3 presents our convex program and shows that, under suitable conditions, it can be solved using a subroutine described in Section 4. Several examples including numerical applications are discussed in Section 5. Our numerical examples assume the call prices on a stock are given by the Black-and-Scholes formula with a constant volatility for one or two maturities, but do not assume the stock to follow a log-normal process with a constant volatility. Section 6 discusses various extensions of our method. Section 7 contains concluding remarks. Most proofs are contained in the appendix.

2 Preliminaries

2.1 The modelling framework

We aim to establish model-independent bounds in a simple setting inspired from the classical theory of multi-period markets (see (Pliska 2005, Chapter 3) and (Follmer and Schied 2004, Ch. 5)). For simplicity, we assume for now that interest rates and dividends are null. Section 6 shows how to incorporate interest rates and dividends in our framework. An m-period market $\mathcal{M}$ consists of d basic securities on a non-empty sample space Ω. The securities prices $S_0^1, \ldots, S_0^d$ at time-step 0 are known constants. The price S_i^k of security k at time-step i, $0 \leq i \leq m$, is a function from Ω to $\mathbb{R}$. Let X_i be the price vector $(S_i^1, \ldots, S_i^d)$. An investor can buy ξ_i^k positions in security k, $1 \leq k \leq d$, at time-step $i-1$ and sell them at time-step i. Denote by ξ_i the d-dimensional vector (ξ_i^k), $1 \leq k \leq d$. The vector ξ_i is an arbitrary function of the past values of the d securities, i.e. it is a function of S_j^k, $1 \leq k \leq d$, and $1 \leq j \leq i-1$. The cumulative payoff of the investor, which we call *a gains function,* is $\sum_{i=1}^m \xi_i^T (X_i - X_{i-1})$. We assume that zero-coupon bonds maturing at time-step m are liquid at time 0. A financial derivative is a function from Ω to $\mathbb{R}$. A finite-support probability on Ω is a non-negative function on Ω that takes positive values on a finite number of elements and that sums up to 1.

Definition 2.1. *A risk-neutral probability is a finite-support probability P on Ω such that $E_P(g) = 0$ for any gains function g.*

2.2 The super-hedging cost

Consider a financial derivative ψ. We say that a financial derivative ψ' super-replicates ψ, and write $\psi \leq_g \psi'$, if there is a gains function g such that $\psi(\omega) \leq \psi'(\omega) + g(\omega)$ for $\omega \in \Omega$. Thus, the payoff of a properly hedged long position in ψ' is always no less than the payoff of ψ. Consider a portfolio of γ bonds that pay 1 at time-step m. The portfolio cost is γ. Define the super-hedging cost $c(\psi) = c(\psi; \mathcal{M})$ of ψ as

$$c(\psi) = \inf\{\gamma : \gamma \in \mathbb{R}, \psi \leq_g \gamma\}, \tag{2.1}$$

with the usual convention that $\inf(\emptyset) = \infty$. Thus $c(\psi)$ is the infimum price of a bond that super-replicates ψ. Let $\mathbb{P}$ be the set of risk-neutral probabilities. Since $E_P(\psi) \leq \gamma$ for any risk-neutral probability P and any γ such that $\psi \leq_g \gamma$,

$$\sup_{P \in \mathbb{P}} E_P(\psi) \leq c(\psi). \tag{2.2}$$

2.3 The super-replicating price

Assume now that $\phi = (\phi_1, \ldots, \phi_l)$ is a vector of financial derivatives that trade at prices $\pi_1, \ldots, \pi_l$ at time-step 0. Let $\beta = (\beta_1, \ldots, \beta_l)$ be a vector of length l. The portfolio $\beta^T \phi + \gamma$ consists of β_j derivatives ϕ_j, $1 \leq j \leq l$, and of γ bonds that pay 1 at time-step m. It has a cost of $\beta^T \pi + \gamma$, where $\pi = (\pi_1, \ldots, \pi_l)$. Define the best super-replicating price $\pi_{\text{sup}} = \pi_{\text{sup}}(\psi, \phi; \mathcal{M})$ as

$$\pi_{\text{sup}} = \inf_{(\beta,\gamma) \in V_0} \beta^T \pi + \gamma, \tag{2.3}$$

where

$$V_0 = V_0(\psi, \phi; \mathcal{M}) = \{(\beta, \gamma) \in \mathbb{R}^l \times \mathbb{R}, \psi \leq_g \beta^T \phi + \gamma\}.$$

In other words, π_{sup} is the infimum cost of a super-replicating portfolio composed of a bond and of positions in the derivatives ϕ_j, $1 \leq j \leq l$. Define $V = V(\psi, \phi; \mathcal{M})$ as

$$V = \{(\beta, \gamma) \in \mathbb{R}^l \times \mathbb{R} : c(\psi - \beta^T \phi) \leq \gamma\}. \tag{2.4}$$

If $(\beta,\gamma) \in V_0$ then $\psi - \beta^T\phi \leq_g \gamma$ and so, by (2.1), $(\beta,\gamma) \in V$. Conversely, it follows from (2.1) that, if $(\beta,\gamma) \in V$ and $\gamma' > \gamma$, then $\psi - \beta^T\phi \leq_g \gamma'$, and so $(\beta,\gamma') \in V_0$. Thus, $V_0 \subseteq V \subseteq \overline{V_0}$, where $\overline{V_0}$ is the closure of V_0, and (2.3) can be rewritten as

$$\pi_{\sup} = \inf_{(\beta,\gamma)\in V} \beta^T\pi + \gamma. \tag{2.5}$$

Note that

$$V \subseteq \{(\beta,\gamma) \in \mathbb{R}^l \times \mathbb{R} : E_P(\psi) \leq \beta^T E_P(\phi) + \gamma \text{ for } P \in \mathbb{P}\}. \tag{2.6}$$

This is because the RHS of (2.6) is closed and contains V_0, and so it contains V.

Example 2.1. *Assume $m = 2$, $d = l = 1$, and the basic security is a stock valued at S_i at time-step i, $0 \leq i \leq 2$, with $\phi_1 = \max(S_2 - K, 0)$ and $\psi = \max((S_1 + S_2)/2 - K, 0)$. As*

$$\psi \leq \frac{1}{2} 1_{S_1 \geq K}(S_1 - S_2) + \phi_1,$$

$\pi_{\sup}$ is upper-bounded by the price of ϕ_1.

Remark 2.1. *An optimal sub-replicating price and strategy for ψ can be obtained by negating an optimal super-replicating price and strategy for $-\psi$.*

Throughout the rest of the paper, the running time refers to the number of arithmetic operations. Denote by $\mathbf{0}_k$ the null vector and by $\mathbf{1}_k$ the all-one vector in $\mathbb{R}^k$, and denote by e_i the vector whose i-th coordinate is 1 and remaining coordinates are null.

3 Super-replication as a convex program

3.1 General assumptions

We assume throughout this section that we are given positive real numbers $\delta \leq 1/2$ and q such that

A1. $c(\psi - \beta^T\phi)$ is upper-bounded by q for $||\beta|| \leq \delta$.

A2. For any element π' of the set $\bigcup_{i=1}^{l}\{\pi + \delta e_i, \pi - \delta e_i\}$, there is a risk-neutral probability P with $E_P(\phi) = \pi'$ and $-q \leq E_P(\psi)$.

A3. For $\beta \in \mathbb{R}^l$, there is a risk-neutral probability P such that

$$c(\psi - \beta^T\phi) = E_P(\psi - \beta^T\phi), \tag{3.1}$$

and a subroutine that, on input β, calculates $c(\psi - \beta^T\phi)$ and $E_P(\phi)$ in finite time $\mathcal{T}$.

Assumptions A1, A2 and A3 will be used to bound the running time of our convex program (see Theorem 3.2). Each time we will use Theorem 3.2 in the examples of Section 5, we will first prove that these assumptions hold by discretizing the state-space and using techniques developed in Sections 4 and 5. Assumption A1 implies that $(\mathbf{0}_l, q) \in V$ and so, by (2.5), $\pi_{\sup} \leq q$. Note that if π' belongs to $\{\pi + \delta e_i, \pi - \delta e_i\}$, then π and π' have the same coordinates expect for the i-th coordinate, and $|\pi_i - \pi'_i| = \delta$. Assumption 2 holds under certain no-arbitrage conditions (see, e.g. (Kahalé 2010, Theorem 4.6)), but will be shown directly in our examples. On the other hand, Assumption A3 implies that there is no duality gap in (2.2) for the derivative $\psi - \beta^T\phi$. The performance of our super-replication algorithm depends to a large extent on $\mathcal{T}$ (see Theorem 3.2, and Section 5), both in theory and in practice. When the state-space is finite and under a geometric condition (see Assumption A4), Section 4 shows that there is no duality gap in (2.2) for a class of derivatives, and gives a construction of the subroutine in Assumption A3.

Remark 3.1. *If $c(-\psi)$ is upper-bounded by a real number q' then, by (2.2), $-q' \leq E_P(\psi)$ for $P \in \mathbb{P}$.*

3.2 The convex program

We first show that

$$V = \{(\beta, \gamma) \in \mathbb{R}^l \times \mathbb{R} : E_P(\psi) \leq \beta^T E_P(\phi) + \gamma \text{ for } P \in \mathbb{P}\}. \tag{3.2}$$

By (2.6), the RHS of (3.2) contains V. Conversely, assume that (β, γ) belongs to the RHS of (3.2). By Assumption A3, there is $P \in \mathbb{P}$ such that $c(\psi - \beta^T \phi) = E_P(\psi - \beta^T \phi)$, and so $c(\psi - \beta^T \phi) \leq \gamma$. Hence, $(\beta, \gamma) \in V$.

Thus V is a convex set and (2.5) is a convex program since it implies that π_{sup} is the infimum of a linear function over a convex set. Convex programs over bounded sets that admit a separation oracle can be solved efficiently under conditions stated in Subsection 3.4. Note that V is unbounded since $(\mathbf{0}_l, \gamma) \in V$ for $\gamma \geq q$, but (2.5) can be rewritten as

$$\pi_{\text{sup}} = \inf_{(\beta,\gamma)\in V'} \beta^T \pi + \gamma, \text{ where} \tag{3.3}$$

$$V' = \{(\beta, \gamma) \in V : \beta^T \pi + \gamma \leq q + 1\}.$$

Lemma 3.1 shows that V' is bounded. For the rest of the paper, let $R_0 = (2q + 1)/\delta$ and

$$R_1 = 4\frac{(1 + q)\sqrt{l}(1 + ||\pi||)}{\delta}.$$

Lemma 3.1. *Let*

$$\mathcal{E} = \{(\beta, \gamma) \in \mathbb{R}^l \times \mathbb{R} : -q \leq \beta^T (\pi \pm \delta e_i) + \gamma \textit{ for } 1 \leq i \leq l\}, \tag{3.4}$$

$$\mathcal{E}' = \mathcal{E}'(l, \pi, q, \delta) = \{(\beta, \gamma) \in \mathcal{E} : \beta^T \pi + \gamma \leq q + 1\}. \tag{3.5}$$

Then $V \subseteq \mathcal{E}$, $V' \subseteq \mathcal{E}'$ and, for $(\beta, \gamma) \in \mathcal{E}'$, $||\beta||_\infty \leq R_0$ and $||(\beta, \gamma)|| \leq R_1$. Furthermore, V' contains the ball of radius $\delta(1 + ||\pi||)^{-1}$ centered at $(\mathbf{0}_l, q + \delta)$.

3.3 A separation oracle for V'

A separation oracle for a convex set $C \subseteq \mathbb{R}^k$ is a subroutine with the following property. The oracle accepts as input any vector $y \in \mathbb{R}^k$. If $y \in C$ the oracle returns a "Yes", whereas if $y \notin C$ the oracle returns a vector $a \in \mathbb{R}^k - \{0\}$ such that $a^T y \leq a^T x$ for any $x \in C$.

Proposition 3.1. *V' admits a separation oracle that runs in $\mathcal{T} + O(1)$ time, where $\mathcal{T}$ is the running time of the subroutine in Assumption A3. On any input $(\beta, \gamma) \in \mathbb{R}^l \times \mathbb{R}$, the oracle either returns a "Yes", or returns one of the two vectors $-(\pi, 1)$ or $(E_P(\phi), 1)$, where $P \in \mathbb{P}$ is such that $\beta^T E_P(\phi) + \gamma \leq E_P(\psi)$.*

Proof. On input $(\beta, \gamma) \in \mathbb{R}^l \times \mathbb{R}$, the separation oracle for V' runs through the following steps.

1. Let $a_0 = (\pi, 1)$. If $a_0^T(\beta, \gamma) > q + 1$, the oracle returns $-a_0$. This is a valid return since $a_0^T(\beta', \gamma') \leq q + 1$ for $(\beta', \gamma') \in V'$, and $(\beta, \gamma) \notin V'$.

2. Else, calculate $c(\psi - \beta^T \phi)$ via the subroutine in Assumption A3. If $c(\psi - \beta^T \phi) \leq \gamma$ the oracle returns a "Yes".

3. Else, use the subroutine in Assumption A3 on input β to calculate $a = (E_P(\phi), 1)$, where P is a risk-neutral probability such that (3.1) holds. By (3.1), $\gamma \leq E_P(\psi - \beta^T \phi)$ and so $a^T(\beta, \gamma) \leq E_P(\psi)$. On the other hand, (3.2) implies that $E_P(\psi) \leq a^T(\beta', \gamma')$ for $(\beta', \gamma') \in V$. Hence $a^T(\beta, \gamma) \leq a^T(\beta', \gamma')$. The oracle returns a.

$\square$

3.4 Cutting plane algorithms

We now show how to solve the convex program (3.3) using a cutting plane algorithm. We assume throughout this subsection that C is a convex subset of $\mathbb{R}^k$ that contains a ball of radius r, is contained in the ball $B(0,R)$ of radius R centered at 0, and that C admits a separation oracle. We describe a generic cutting plane algorithm (see (Grötschel, Lovász and Schrijver 1981, Atkinson and Vaidya 1995, Vaidya 1996, Bertsimas and Vempala 2004) and references therein) for minimizing $a_0^T x$ for $x \in C$, where a_0 is a vector of $\mathbb{R}^k$. The algorithm takes as input r, R, a_0, a real number $\epsilon \in (0,1)$, and a separation oracle for C. It outputs a vector $y \in C$ such that $a_0^T y \leq a_0^T x + \epsilon$ for any $x \in C$. The algorithm makes N iterations and uses the following steps:

1. Let $\mathcal{R}_0 \subseteq \mathbb{R}^k$ be a bounded region that contains $B(0,R)$.
2. For $i = 1$ to N, choose a point $y_i \in \mathcal{R}_{i-1}$. Query the separation oracle for C on y_i. If the oracle returns a "Yes", let $a_i = -a_0$. We say in this case that i is a feasible index. Otherwise, let a_i be the vector returned by the oracle. In both cases, let $\mathcal{R}_i$ be a region such that
$$\mathcal{R}_i \supseteq \mathcal{R}_{i-1} \cap \{x \in \mathbb{R}^k : a_i^T y_i \leq a_i^T x\}. \tag{3.6}$$
3. Output y_j, where j is a feasible index such that $a_0^T y_j \leq a_0^T y_i$ for all feasible indices i, $1 \leq i \leq N$.

The choice of y_i and $\mathcal{R}_i$ depends on the specific cutting plane algorithm. In a basic analytic center cutting plane algorithm (see (Atkinson and Vaidya 1995) and references therein, and (Boyd, Vandenberghe and Skaf 2008) for a detailed description, which largely inspired our numerical implementation), $\mathcal{R}_0$ is the set of vectors x with $||x||_\infty \leq R$, $\mathcal{R}_i$ is the RHS of (3.6) for $i \geq 1$, and y_i is the analytic center of $\mathcal{R}_{i-1}$. Let

$$I = \{i \in [1,N] : a_i \neq -a_0\},$$

$$\tilde{C} = \{x \in \mathcal{R}_0 : a_i^T y_i \leq a_i^T x \text{ for } i \in I\}.$$

Note that $C \subseteq \tilde{C}$ since $C \subseteq \mathcal{R}_0$ and, for $i \in I$, a_i is the vector returned by the separation oracle for C on input y_i and so, for $x \in C$, $a_i^T y_i \leq a_i^T x$. A cutting plane algorithm based on the volumetric center yields the following result.

Theorem 3.1. *(Vaidya 1996, Sections 2 and 4). Assume $r \leq 1 \leq ||a_0||$ and $R \geq 1$. There is a cutting plane algorithm that finds a vector $y \in C$ such that $a_0^T y \leq a_0^T x + \epsilon$ for every $x \in \tilde{C}$. The algorithm makes N calls to the separation oracle for C, where*

$$N = O(k \ln(\frac{2kR^2||a_0||}{\epsilon r})), \tag{3.7}$$

and runs in $O(N(k^3 + \mathcal{T}))$ time, where $\mathcal{T}$ is the running time of the oracle.

Since $C \subseteq \tilde{C}$, $a_0^T y \leq a_0^T x + \epsilon$ for every $x \in C$. In the convex program (3.3), $a_0 = (\pi, 1)$, the number of variables is $k = l + 1$, $C = V'$ and, by Lemma 3.1, $r = \delta(1 + ||\pi||)^{-1}$ and $R = R_1$. The separation oracle for V' described in Proposition 3.1 can be used in Step 2 of the cutting plane algorithm.

3.5 A lower bound on π_{sup} via risk-neutral probabilities

Assume we use a generic cutting plane algorithm with N iterations to solve the convex program (3.3). For any $P \in \mathbb{P}$ such that $E_P(\phi) = \pi$, it follows from (2.5) and (3.2) that $E_P(\psi) \leq \pi_{\text{sup}}$. We calculate below a lower bound on π_{sup} by choosing P to be a weighted average of the risk-neutral probabilities defined in Assumption A2 or generated by the cutting plane algorithm. For $i \in I$, a_i is the vector returned by the separation oracle for V' on input $y_i = (\beta_i, \gamma_i)$ and so, by Proposition 3.1, there is $P_i \in \mathbb{P}$ such that

$$a_i = (E_{P_i}(\phi), 1) \text{ and } a'_i \leq E_{P_i}(\psi), \tag{3.8}$$

where $a'_i = a_i^T y_i$. For $1 \leq i \leq l$, set $a_{i+N} = (\pi + \delta e_i, 1)$ and $a_{i+l+N} = (\pi - \delta e_i, 1)$. Finally, for $N+1 \leq i \leq N+2l$, set $a'_i = -q$, and let $I' = I \cup [N+1, N+2l]$. By Assumption A2, for $N+1 \leq i \leq N+2l$, there is $P_i \in \mathbb{P}$ such that (3.8) holds. Let $a_0 = (\pi, 1)$ and consider the linear program:

$$b_0 = \max_{\lambda} \sum_{i \in I'} \lambda_i a'_i \tag{3.9}$$

subject to:

$$\lambda_i \geq 0 \text{ for } i \in I', \text{ and } \sum_{i \in I'} \lambda_i a_i = a_0. \tag{3.10}$$

This program has $|I'| + l + 1$ constraints and $|I'|$ variables, with $|I'| \leq N + 2l$. It is feasible since the constraints (3.10) are satisfied when $\lambda_i = 0$ for $i \in I$ and $\lambda_i = (2l)^{-1}$ for $i \in [N+1, N+2l]$. By (3.8), for any λ_i's satisfying (3.10), $\sum_{i \in I'} \lambda_i = 1$, $\sum_{i \in I'} \lambda_i E_{P_i}(\phi) = \pi$ and $\sum_{i \in I'} \lambda_i a'_i \leq \sum_{i \in I'} \lambda_i E_{P_i}(\psi)$. Thus $P = \sum_{i \in I'} \lambda_i P_i$ is a risk-neutral probability, $E_P(\phi) = \pi$ and $\sum_{i \in I'} \lambda_i a'_i \leq E_P(\psi)$. Hence $b_0 \leq E_P(\psi) \leq \pi_{\text{sup}}$. The following lemma will show the tightness of this lower bound on π_{sup}.

Lemma 3.2. *For any generic cutting plane algorithm,* $\inf_{x \in \tilde{C}} a_0^T x \leq b_0$.

Combining the preceding results yields the following.

Theorem 3.2. *Under Assumptions A1, A2 and A3, V' admits a separation oracle that runs in $\mathcal{T} + O(1)$ time, where $\mathcal{T}$ is the running time of the subroutine in Assumption A3, and $V' \subseteq B(0, R_1)$. By solving the convex program (3.3), we can calculate in $O(N(l^3 + \mathcal{T}))$ time a vector $(\beta^*, \gamma^*) \in V'$ such that $||\beta^*||_\infty \leq R_0$ and*

$$\pi_{\text{sup}} \leq {\beta^*}^T \pi + \gamma^* \leq \pi_{\text{sup}} + \epsilon, \tag{3.11}$$

where

$$N = O(l \ln(\frac{l(1+q)(1+||\pi||)}{\epsilon\delta})). \tag{3.12}$$

In addition, by solving a linear program with $O(N)$ variables and constraints, we can find weights $\lambda_i \geq 0$, $i \in I'$, that sum up to 1, with $E_P(\phi) = \pi$, where $P = \sum_{i \in I'} \lambda_i P_i$ is a risk-neutral probability, and

$${\beta^*}^T \pi + \gamma^* \leq E_P(\psi) + \epsilon. \tag{3.13}$$

Thus the algorithm calculates π_{sup} with precision ϵ and outputs a portfolio that super-replicates ψ at cost at most $\pi_{\text{sup}} + 2\epsilon$.

Proof. As shown in Subsection 3.4, we can solve the convex program (3.3) by using the cutting plane algorithm in (Vaidya 1996, Sections 2 and 4). The algorithms finds $y = (\beta^*, \gamma^*) \in V'$ such that (3.11) holds. Since $r = \delta(1 + ||\pi||)^{-1}$ and $R = R_1$, (3.12) follows from (3.7) after some calculations. Furthermore, since $(\beta^*, \gamma^*) \in V$, the discussion in subsection 2.3 shows that ψ is super-replicated by the portfolio ${\beta^*}^T \phi + \gamma^* + \epsilon$, whose cost ${\beta^*}^T \pi + \gamma^* + \epsilon$ is at most $\pi_{\text{sup}} + 2\epsilon$. Lemma 3.1 implies that $||\beta^*||_\infty \leq R_0$. On the other hand, by Lemma 3.2, if we solve the linear program (3.9) and (3.10), we find b_0 such that $\inf_{x \in \tilde{C}} a_0^T x \leq b_0 \leq E_P(\psi)$, where P is a risk-neutral probability such that $E_P(\phi) = \pi$. Since, by Theorem 3.1, $a_0^T y \leq a_0^T x + \epsilon$ for $x \in \tilde{C}$, $a_0^T y \leq E_P(\psi) + \epsilon$. Hence (3.13). □

4 Recursive calculation of the super-hedging cost

Theorem 3.2 shows how to calculate $\pi_{\sup}$ if Assumptions A1, A2 and A3 hold. We will see in Section 5 how to verify Assumptions A1 and A2. We show in this section that A3 holds for a class of financial derivatives when the state-space is finite and under a geometric condition described in Assumption A4. We also give a construction of the subroutine needed in A3 by describing an algorithm that calculates the super-hedging cost of a financial derivative Ψ by backward induction using concave envelopes, in the same spirit as in (Carassus, Gobet and Temam 2007), and applying the algorithm with $\Psi = \psi - \beta^T \phi$, for a given vector β.

Let f be a real-valued function defined on a subset W of $\mathbb{R}^d$ and bounded above by a linear function. Denote by $\widehat{W}$ the convex hull of W and by $\overline{f}$ the concave envelope of f, i.e. the smallest concave function on $\widehat{W}$ bounded below by f on W. If $x \in \widehat{W}$, denote by $\mathbb{Q}(x, W)$ the set of non-negative functions $\mathcal{Q}$ on W that take positive values on a finite number of elements, sum up to 1, with $\sum_{y \in W} \mathcal{Q}(y)y = x$. Note that $\mathcal{Q}$ defines a probability on W, endowed with its Borel σ-algebra. It can be shown (Boyd and Vandenberghe 2004, Exercise 3.30) that, for $x \in \widehat{W}$,

$$\overline{f}(x) = \sup_{\mathcal{Q} \in \mathbb{Q}(x,W)} E_{\mathcal{Q}}(f), \tag{4.1}$$

where

$$E_{\mathcal{Q}}(f) = \sum_{y \in W} \mathcal{Q}(y) f(y).$$

When W is finite, (4.1) can be rewritten as

$$\overline{f}(x) = \sup_{\mathcal{Q}} \{ \sum_{y \in W} \mathcal{Q}(y) f(y) | \mathcal{Q}(y) \geq 0 \text{ for } y \in W, \sum_{y \in W} \mathcal{Q}(y) = 1, \sum_{y \in W} \mathcal{Q}(y) y = x \}, \tag{4.2}$$

where $\mathcal{Q}$ is a real-valued function on W. By linear programming duality,

$$\overline{f}(x) = \min_{\eta_0 \in \mathbb{R}, \eta \in \mathbb{R}^d} \{ \eta_0 + \eta^T x | f(y) \leq \eta_0 + \eta^T y \text{ for } y \in W \}. \tag{4.3}$$

Let (η_0, η) be a pair that attains the RHS of (4.3). Then, for $y \in W$,

$$f(y) \leq \overline{f}(x) + \eta^T (y - x). \tag{4.4}$$

Thus, f is upper-bounded by a linear function valued at $\overline{f}(x)$ at x.

For $1 \leq i \leq m$, let $D_i = \text{Im}(X_1, \ldots, X_i)$ be the set of paths that X can follow from time-step 1 through i, and set $D_0 = \{\emptyset\}$. For $\theta \in D_i$, let

$$D(\theta) = \{ x \in \mathbb{R}^d : (\theta, x) \in D_{i+1} \} \tag{4.5}$$

be the set of possible values of X_{i+1} given that X has followed the path θ in the first i time-steps. By convention, if $i = 0$ and $\theta = \emptyset$, (θ, x) refers to x. If ζ is a real-valued function on D_{i+1}, denote by $\zeta(\theta, .)$ the function that maps x to $\zeta(\theta, x)$ for $x \in D(\theta)$. The rest of this section makes the following assumption. By convention, $x_0 = X_0$.

A4. Ω is finite and, if $0 \leq i \leq m-1$ and $\theta = (x_1, \ldots, x_i) \in D_i$, then $x_i \in \widehat{D(\theta)}$.

Assumption A4 implies that, for any given path that X has followed in the first i time-steps, X_i belongs to the convex hull of the set of possible values of X_{i+1}. Consider now a financial derivative Ψ of the form $\Psi = \Psi^*(X_1, \ldots, X_m)$, where Ψ^* is a real-valued function defined on D_m that can be calculated in finite time. When $m = 1$, Proposition 4.1 shows that,

$$c(\Psi) = \overline{\Psi^*}(X_0) = \sup_{\mathcal{Q} \in \mathbb{Q}(X_0, D_1)} E_{\mathcal{Q}}(\Psi^*),$$

and so $c(\Psi)$ is equal to the concave envelope of Ψ^* evaluated at X_0. Proposition 4.1 also shows how to calculate by backward induction the super-hedging cost of Ψ in an m-period market, for any integer m.

Proposition 4.1. *Define the functions Ψ_i^* by backward induction as follows: $\Psi_m^* = \Psi^*$ and $\Psi_i^*(\theta) = \overline{\Psi_{i+1}^*(\theta,.)}(x_i)$ for $0 \le i \le m-1$ and $\theta = (x_1,\dots,x_i) \in D_i$. Let Ψ_i be the financial derivative $\Psi_i^*(X_1,\dots,X_i)$. Then $c(\Psi) = \Psi_0$ and*

$$\Psi_i^*(\theta) = \sup_{\mathcal{Q}\in\mathbb{Q}(x_i,D(\theta))} E_{\mathcal{Q}}(\Psi_{i+1}^*(\theta,.)). \tag{4.6}$$

We can interpret Ψ_i as the super-hedging cost at time-step i of Ψ, and (4.6) as saying that Ψ_i is equal to the super-hedging cost of Ψ_{i+1} in the underlying single-period market at time-step i.

Proposition 4.2 below shows that, as in (Pliska 2005, Section 3.4), we can paste together risk-neutral probabilities in single-period markets by multiplying them along any given path to obtain a risk-neutral probability P in the m-period market. It also describes how to calculate $E_P(\eta)$, for a class of financial derivatives η.

Proposition 4.2. *For $\theta = (x_1,\dots,x_i) \in D_i$, choose a probability $\mathcal{P}(\theta) \in \mathbb{Q}(x_i,D(\theta))$. There is a risk-neutral probability P such that, for any deterministic function η^* on D_m, $E_P(\eta) = \eta_0^*(\emptyset)$, where $\eta = \eta^*(X_1,\dots,X_m)$, and the functions η_i^* are defined on D_i by backward induction as follows: $\eta_m^* = \eta^*$ and, for $0 \le i \le m-1$ and $\theta \in D_i$,*

$$\eta_i^*(\theta) = E_{\mathcal{P}(\theta)}(\eta_{i+1}^*(\theta,.)). \tag{4.7}$$

Furthermore, $c(\Psi) = E_P(\Psi)$ if all $\mathcal{P}(\theta)$ attain the RHS of (4.6).

We now use Propositions 4.1 and 4.2 to show that Assumption A3 holds and to give a generic construction of the subroutine needed in Assumption A3. Assume that $\psi = \psi^*(X_1,\dots,X_m)$ and $\phi = \phi^*(X_1,\dots,X_m)$, where ψ^* (resp. ϕ^*) is a function from D_m to $\mathbb{R}$ (resp. $\mathbb{R}^l$). On input β, the subroutine uses the following steps.

1. Let $\Psi = \psi - \beta^T\phi$, $\Psi_m^* = \psi^* - \beta^T\phi^*$ and $\phi_m^* = \phi^*$.
2. For $i = m-1$ down to 0 and $\theta = (x_1,\dots,x_i) \in D_i$, choose a probability $\mathcal{P}(\theta) \in \mathbb{Q}(x_i,D(\theta))$ that maximizes $E_{\mathcal{P}(\theta)}(\Psi_{i+1}^*(\theta,.))$. Set $\Psi_i^*(\theta) = E_{\mathcal{P}(\theta)}(\Psi_{i+1}^*(\theta,.))$ and $\phi_i^*(\theta) = E_{\mathcal{P}(\theta)}(\phi_{i+1}^*(\theta,.))$.
3. Output $c(\Psi) = \Psi_i^*(\emptyset)$ and $E_P(\phi) = \phi_0^*(\emptyset)$, where $P \in \mathbb{P}$ is obtained by pasting the probabilities $\mathcal{P}(\theta)$.

In other words, we find by backward induction, in every single-period market at time-step i, a probability $\mathcal{P}(\theta)$ that attains RHS of (4.6), and use the probabilities $\mathcal{P}(\theta)$ to calculate both $c(\Psi)$ and $E_P(\phi)$. By Proposition 4.2, $c(\Psi) = E_P(\Psi)$.

Finding $\mathcal{P}(\theta)$ can be done in finite time, as shown in Subsection 4.1, and so Assumption A3 holds. But, since $|D_{m-1}|$ is in general exponential in m, so is the running time of a naive implementation of Step 2. The running time can often be considerably reduced using Remarks 4.1 and 4.2 below and techniques used in the valuation of path-dependent derivatives via binomial trees (Hull 2012, Section 26.5). The optimized subroutine replaces the loop on θ in Step 2 by a loop on x_i and on at most one additional state variable. For instance, if ψ is a variance swap and ϕ consists of call options, the loop on θ is replaced by a loop on x_i. If ψ is a volatility swap (resp. Asian option), we can replace the loop on θ by a loop on x_i and on the current realized variance (resp. the price running sum). See Section 5 for details.

Remark 4.1. *Fix $i \in \{0,\dots,m\}$. Consider a financial derivative ζ which is a deterministic function of $X_0,\dots,X_m$, and a financial derivative ζ' which is a deterministic function of $X_0,\dots,X_i$. Let $\Psi = \zeta + \zeta'$. It can be shown by backward induction that $\Psi_i = \zeta_i + \zeta'$, where Ψ_i (resp. ζ_i) is the super-hedging cost of Ψ (resp. ζ) at time-step i.*

Remark 4.2. *Fix $i \in \{1, \dots, m-1\}$. Consider a financial derivative ζ of the form $\zeta = \zeta^*(X_1, \dots, X_m)$, where ζ^* is a real-valued function on D_m. Assume that $\zeta^*(x_1, \dots, x_m)$ and $D(x_1, \dots, x_j)$ depend on $(x_1, \dots, x_i)$ only through a (one or multi-dimensional) deterministic function $h(x_1, \dots, x_i)$ for $j \geq i$. It can be shown by backward induction that the super-hedging cost ζ_i of ζ at time-step i is a deterministic function of X_i and of $h(X_1, \dots, X_i)$.*

4.1 Convex hull calculation

By (4.2) thru (4.4), we can use linear programming to find $\mathcal{Q} \in \mathbb{Q}(x_i, D(\theta))$ that attains the RHS of (4.6), where $\theta = (x_1, \dots, x_i) \in D_i$. When $d = 1$, this can be done more efficiently via the following proposition.

Proposition 4.3 ((Andrew 1979))**.** *Given a finite set $W \subseteq \mathbb{R}$ whose elements $x_1, \dots, x_n$ are sorted in increasing order, a real number $x \in [x_1, x_n]$, and a function f that takes known values on W, we can calculate in $O(n)$ time a probability $\mathcal{Q} \in \mathbb{Q}(x, W)$ that maximizes $E_{\mathcal{Q}}(f)$, and a real number ξ^* such that, for $y \in W$,*

$$f(y) \leq E_{\mathcal{Q}}(f) + \xi^*(y - x). \tag{4.8}$$

Furthermore, $\mathcal{Q}$ is supported on two points.

The algorithm, due to (Andrew 1979), first calculates the ordered subset $U(j)$ of W, $2 \leq j \leq n$, recursively as follows. Let $U(2) = \{x_1, x_2\}$. Assume $U(j-1) = \{x'_1, \dots, x'_h\}$. Let k be the largest index such that $(x'_k, f(x'_k))$ is above the segment $[(x'_{k-1}, f(x'_{k-1})), (x_j, f(x_j))]$, if such an index exists, otherwise let $k = 1$. Set $U(j) = \{x'_1, \dots, x'_k, x_j\}$. Let x' and x'' be consecutive elements of $U(n)$ such that $x \in [x', x'']$. The algorithm outputs the probability $\mathcal{Q}$ that assigns $s = (x'' - x)/(x'' - x')$ to x' and $1 - s$ to x'', and

$$\xi^* = \frac{f(x'') - f(x')}{x'' - x'}. \tag{4.9}$$

5 Examples

Let S be a stock valued at S_i at time-step i, $0 \leq i \leq m$, where S_0 is known. For $1 \leq i \leq m$, we are given a (possibly empty) increasing sequence $K_{i,j}$ of positive strikes, $1 \leq j \leq l_i$, together with prices $c_{i,j}$ of calls with maturity t_i and strike $K_{i,j}$. By convention, $l_i = 0$ if no calls trade at time-step i. Let $\mathcal{K}$ be the set that consists of S_0 and of $K_{i,j}$, $1 \leq i \leq m$, $1 \leq j \leq l_i$. Given subsets $F_i \supseteq \mathcal{K}$ of $\mathbb{R}^+$ for $1 \leq i \leq m$, consider the market $\mathcal{M}(F_1, \dots, F_m)$ where S_i ranges in F_i for $1 \leq i \leq m$. We formally build $\mathcal{M}(F_1, \dots, F_m)$ by setting $\Omega = \{S_0\} \times F_1 \times \cdots \times F_m$, with $S_i(\omega) = x_i$ for $\omega = (x_0, \dots, x_m) \in \Omega$ and $0 \leq i \leq m$. We assume that $F_{i-1} \subseteq \widehat{F}_i$ for $2 \leq i \leq m$, so that Assumption A4 holds if the sets F_i, $1 \leq i \leq m$, are finite. If $F_i = F$ for $1 \leq i \leq m$, denote $\mathcal{M}(F_1, \dots, F_m)$ by $\mathcal{M}_m(F)$.

Definition 5.1. *Consider an ordered pair $\{k_*, k^*\}$ of fictitious strikes such that*

$$0 \leq k_* < \min(\mathcal{K}) \leq \max(\mathcal{K}) < k^* \leq \infty.$$

For $1 \leq i \leq m$, set

$$K_{i,0} = k_*, K_{i,l_i+1} = k^*, c_{i,0} = S_0 - k_* \text{ and } c_{i,l_i+1} = 0. \tag{5.1}$$

We say that $\{k_*, k^*\}$ is acceptable if, for $1 \leq i \leq i' \leq m$, $i \leq i'' \leq m$, $1 \leq j \leq l_i$, $0 \leq j' \leq l_{i'} + 1$, $0 \leq j'' \leq l_{i''} + 1$, if $K_{i',j'} \leq K_{i,j} \leq K_{i'',j''}$, $(i,j) \notin \{(i',j'), (i'',j'')\}$, and $K_{i',j'} < K_{i'',j''}$, then

$$\max(0, S_0 - K_{i,j}) < c_{i,j} < w c_{i',j'} + (1-w) c_{i'',j''}, \tag{5.2}$$

where $w = (K_{i'',j''} - K_{i,j})/(K_{i'',j''} - K_{i',j'})$. By convention, $w = 1$ if $K_{i'',j''} = \infty$.

(5.1) can be interpreted by noting that, if $F_m \subseteq [k_*, k^*]$, a call price with strike k_* (resp. k^*) must equal $S_0 - k_*$ (resp. 0). On the other hand, it is shown in (Davis and Hobson 2007) that, if $F_m \subseteq [k_*, k^*]$ and under no-arbitrage constraints, (5.2) holds if the strict inequalities are replaced with weak inequalities. Indeed, no-arbitrage constraints imply that call prices are convex with respect to the strike and non-decreasing with respect to time, and (5.2) combines a strict version of these two constraints.

For $x \geq 0$ and $1 \leq i \leq m$, define the vector $f_i(x) = (\max(x - K_{i,j}, 0))$, $1 \leq j \leq l_i$, and let $b_i \in \mathbb{R}^{l_i}$. Consider the vector of calls $\phi = (f_i(S_i))$, and set $\beta = (b_i)$, $1 \leq i \leq m$. Let ψ be a financial derivative that pays $\psi^*(S_1, \ldots, S_m)$, where ψ^* is a deterministic function, such that $c(\psi)$ and $c(-\psi)$ are finite. Consider the financial derivative $\Psi = \psi - \beta^T \phi$. We will use the following proposition to show that Assumptions A1 and A2 hold.

Proposition 5.1. *Assume that an acceptable pair $\{k_*, k^*\}$ is contained in F_i, for $1 \leq i \leq m$. Let $\delta_0 = \delta_0(k_*, k^*)$ be the minimum of $1/2$ and of the minimum difference between the right-hand sides and left-hand sides of the inequalities in (5.2). Then A1 and A2 hold in the market $\mathcal{M} = \mathcal{M}(F_1, \ldots, F_m)$ if $q \geq \max(c(\psi; \mathcal{M}), c(-\psi; \mathcal{M})) + \delta_0 \sqrt{l} S_0$ and $0 < \delta < \delta_0$. Furthermore,*

$$||\pi|| \leq S_0 \sqrt{l}. \tag{5.3}$$

In the following examples, we show that A3 holds as well. We then use Theorem 3.2 to calculate an optimal super-replicating price and a corresponding hedging portfolio for ψ via the call components of ϕ. An optimal sub-replicating price and a corresponding hedging portfolio can be derived as well using Remark 2.1. In our examples, we first derive discretized optimal model-independent bounds by assuming the F_i's to be finite, then give an empirical and a theoretical estimate of the discretization error. The empirical estimate is calculated using a large number of discretization points. Our numerical examples were obtained using an analytic center cutting plane algorithm with at most 200 iterations that calculated the model-independent price bound (or its discrete approximation) within an error of order 10^{-5}. The simulation experiments were performed on a desktop PC with an Intel Pentium 2.90 GHz processor and 4 Go of RAM, running Windows 7 Professional. The codes were written in the C++ programming language, and the compiler used was Microsoft Visual C++ 2013. The computing time is given in seconds.

5.1 Forward start options

Consider a forward start option ψ. For ease of exposition, assume the option is an at the money call that pays $\max(0, S_2 - S_1)$ at maturity. Extension to the general case is straightforward. The derivative Ψ pays $\Psi^*(S_1, S_2)$, where

$$\Psi^*(x_1, x_2) = \max(0, x_2 - x_1) - b_1^T f_1(x_1) - b_2^T f_2(x_2).$$

For $x_1 \in F_1$, let

$$\Psi_1^*(x_1; F_2) = \sup_{\mathcal{Q} \in \mathbb{Q}(x_1, F_2)} E_{\mathcal{Q}}(\Psi^*(x_1, .)). \tag{5.4}$$

If F_1 and F_2 are finite, by Proposition 4.1, in the market $\mathcal{M}(F_1, F_2)$, the super-hedging cost of Ψ at time-step 1 is $\Psi_1^*(S_1; F_2)$, and the super-hedging cost of Ψ at time-step 0 is

$$c(\Psi; \mathcal{M}(F_1, F_2)) = \sup_{\mathcal{Q}' \in \mathbb{Q}(S_0, F_1)} E_{\mathcal{Q}'}(\Psi_1^*(.; F_2)). \tag{5.5}$$

We illustrate the calculation of $c(\Psi; \mathcal{M}_2(F))$ when $S_0 = 100$ and $F = \{70, 80, \ldots, 130\}$. Assume ϕ consists of the calls maturing at time-step i with strike $K_{i,j} = 80 + 10j$ for $1 \leq i \leq 2$ and $1 \leq j \leq 3$. Let $b_1 = (-0.3, -0.2, -0.4)$ and $b_2 = (0.5, 0.4, 0.3)$. Proposition 4.3 shows how to calculate $\Psi_1^*(x_1; F)$ for $x_1 \in F$ and probabilities $\mathcal{P}(x_1)$ and $\mathcal{P}(\emptyset)$ that attain the RHS of (5.4) and (5.5), respectively. Using Proposition 4.2, we can then calculate $E_P(\phi)$, where P

is a risk-neutral probability such that $c(\Psi;\mathcal{M}_2(F)) = E_P(\Psi)$. Table 1 lists the values of Ψ when $S_1 = 90$. Fig. 1 plots the function $\Psi^*(90,.)$ and its concave envelope, which shows that $\Psi_1^*(90;F_1) = 2/3 \times 5 + 1/3 \times 0 = 10/3$. The probability that assigns a weight $2/3$ to 100 and $1/3$ to 70 maximises the RHS of (5.4) when $x_1 = 90$. We can perform a similar calculation for each $x_1 \in F$ and then calculate $c(\Psi;\mathcal{M}_2(F))$ via (5.5). Fig. 2 draws a tree by connecting each $x_1 \in F$ (resp. S_0) to the support of $\mathcal{P}(x_1)$ (resp. $\mathcal{P}(\emptyset)$), and Table 2 gives the probability p_u assigned to the largest element in the support of $\mathcal{P}(x_1)$ (resp. $\mathcal{P}(\emptyset)$). It also shows that $E_P(\eta) = \$7.5$, where η is the ATM vanilla call maturing at the second time-step. The expected value under P of any component of ϕ can be calculated in a similar manner.

We now show how to calculate the best super-replicating (resp. sub-replicating) price $\pi_{\sup}(\mathbb{R}^+)$ (resp. $\pi_{\inf}(\mathbb{R}^+)$) of ψ without restrictions on S_1 and S_2 other than being non-negative, and set $\Psi_1^*(x_1) = \Psi_1^*(x_1;\mathbb{R}^+)$ for $x_1 \geq 0$. It follows from (5.4) and the convexity of $\Psi^*(x_1,.)$ on any interval disjoint with $\mathcal{K}$ that $\Psi_1^*(x_1)$ depends only on the values of $\Psi^*(x_1,.)$ on $\mathcal{K} \cup \{0\}$ and on its asymptotic slope on $[\max(\mathcal{K}),\infty)$. Furthermore, it can be shown that Ψ_1^* is linear on $[\max(\mathcal{K}),\infty)$, and so we only need to know the values of Ψ_1^* on $[0,\max(\mathcal{K})]$ and its slope on $[\max(\mathcal{K}),\infty)$ to calculate $c(\Psi;\mathcal{M}_2(\mathbb{R}^+))$ via (5.5). This suggests that $\pi_{\sup}(\mathbb{R}^+)$ is well approximated by the best super-replicating price $\pi_{\sup}(F_1,F_2)$ of ψ in the market $\mathcal{M}(F_1,F_2)$, where

$$F_1 = \mathcal{K} \cup \{\max(\mathcal{K})\frac{j}{n}, 0 \leq j \leq n\} \cup \{\epsilon'^{-1}\max(\mathcal{K})\}, \text{ and} \tag{5.6}$$

$$F_2 = \mathcal{K} \cup \{0, \epsilon'^{-1}\max(\mathcal{K}), \epsilon'^{-2}\max(\mathcal{K})\}, \tag{5.7}$$

n is a positive integer, and $\epsilon' \in (0,1)$. Theorem 5.1 below shows how to calculate $\pi_{\sup}(F_1,F_2)$ and gives an upper bound on the discretization error. The constants behind the O notation in (5.8) and in Section F depend on l, S_0, k_0^*, the calls strikes, maturities and prices, but do not depend on n and ϵ'.

Theorem 5.1. *Assume that $\{0,k_0^*\}$ is acceptable, where k_0^* is a positive real number. Then, for $\epsilon' \in (0,\max(\mathcal{K})/k_0^*]$, $\pi_{\sup}(F_1,F_2)$ can be calculated in $O(N(l^3 + ln))$ total time with precision ϵ via the convex program* (3.3)*, where*

$$N = O(l\ln(\frac{l(1+S_0)}{\epsilon\delta_0(0,k_0^*)})).$$

V' has a separation oracle that runs in $O(ln)$ time. Furthermore,

$$\pi_{\sup}(F_1,F_2) - \pi_{\sup}(\mathbb{R}^+) = O(n^{-2} + \epsilon'). \tag{5.8}$$

Proof. Let $\epsilon' \in (0,\max(\mathcal{K})/k_0^*)$ and $k^* = \epsilon'^{-1}\max(\mathcal{K})$. Since $k^* \geq k_0^*$, it follows from Definition 5.1 that $\{0,k^*\}$ is acceptable. Furthermore, an easy calculation shows that $\delta_0(0,k^*) \geq \delta_0(0,k_0^*)$. On the other hand, since $-\psi \leq 0$ and $\psi \leq S_2$, the super-hedging costs of ψ and of $-\psi$ in the market $\mathcal{M}(F_1,F_2)$ are at most S_0. Thus, by Proposition 5.1, Assumptions A1 and A2 hold for $\mathcal{M}(F_1,F_2)$ if

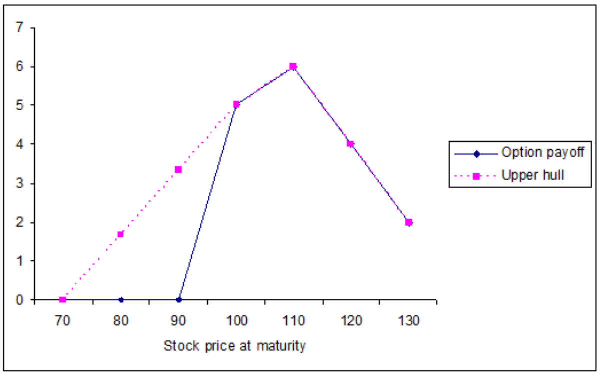
$$q = S_0(1+\sqrt{l}) \text{ and } \delta = \delta_0(0,k_0^*)/2. \tag{5.9}$$

Proposition 4.3 shows how to calculate $\Psi_1^*(x_1;F_2)$ for all $x_1 \in F_1$, $c(\Psi;\mathcal{M}(F_1,F_2))$, and probabilities $\mathcal{P}(x_1)$ and $\mathcal{P}(\emptyset)$ that attain the right-hand sides of (5.4) and (5.5), respectively, in total time $O(|F_1||F_2|)$. By Proposition 4.2, there is a risk-neutral probability P in $\mathcal{M}(F_1,F_2)$ such that $c(\Psi;\mathcal{M}(F_1,F_2)) = E_P(\Psi)$ and, for any financial derivative $\eta = \eta^*(S_1,S_2)$, if we set $\eta_1^*(x_1) = E_{\mathcal{P}(x_1)}(\eta^*(x_1,.))$ for $x_1 \in F_1$, then $E_{\mathcal{P}(\emptyset)}(\eta_1^*) = E_P(\eta)$. Thus, we can then calculate $E_P(\phi)$ in $O(l)$ time, and so A3 holds, with $\mathcal{T} = O(ln)$. The first part of the theorem then follows from Theorem 3.2 and (5.3). The proof of (5.8) is in Section F. $\square$

Table 1: The function $\Psi^*(90,.)$.

S_2	70	80	90	100	110	120	130
$\Psi^*(90,.)$	0	0	0	5	6	4	2

Figure 1: The solid line plots the payoff of Ψ when $S_1 = 90$ and the dotted line plots the corresponding upper hull.



It follows from Theorem 5.1 that, as ϵ goes to 0, we can calculate $\pi_{\text{sup}}(\mathbb{R}^+)$ with precision $O(\epsilon)$ in total time $O(\epsilon^{-1/2}\ln(\epsilon^{-1}))$ by setting $n = \epsilon^{-1/2}$ and $\epsilon' = \epsilon$. (Beiglböck, Henry-Labordère and Penkner 2013) calculate a discrete approximation of $\pi_{\text{sup}}(\mathbb{R}^+)$ similar to $\pi_{\text{sup}}(F_1, F_2)$ using linear programming with $\Theta(n)$ variables and constraints. Our algorithm is much faster for large values of n.

We calculate $\pi_{\text{inf}}(\mathbb{R}^+)$ in a similar manner using Remark 2.1 and setting

$$\Psi^*(x_1, x_2) = -\max(0, x_2 - x_1) - b_1^T f_1(x_1) - b_2^T f_2(x_2).$$

We first calculate the best sub-replicating price $\pi_{\text{inf}}(F_1', F_2')$ of ψ in the market $\mathcal{M}(F_1', F_2')$, where

$$F_1' = \{0, \epsilon'^{-1}(\max \mathcal{K})\} \cup \mathcal{K},$$

$$F_2' = F_1' \cup \{\epsilon'^{-2}(\max \mathcal{K})\},$$

and $0 < \epsilon' < 1$. The choice of F_i' can be motivated in a way similar to that of F_i in (5.6) and (5.7), by noting that $\Psi^*(x_1,.)$ is linear on any interval disjoint with $\mathcal{K} \cup \{x_1\}$, and that Ψ_1^* is convex on any interval disjoint with $\mathcal{K}$. The constant behind the O notation in (5.10) depends on l, S_0, the calls strikes, maturities and prices, but does not depend on ϵ'.

Theorem 5.2. *Assume that $\{0, k_0^*\}$ is acceptable, where k_0^* is a positive real number. Then, for $\epsilon' \in (0, \max(\mathcal{K})/k_0^*]$, $\pi_{\text{inf}}(F_1', F_2')$ can be calculated in $O(Nl^3)$ total time with precision ϵ via the convex program* (3.3)*, where*

$$N = O(l \ln(\frac{l(1 + S_0)}{\epsilon \delta_0(0, k_0^*)})).$$

V' has a separation oracle that runs in $O(l^2)$ time. Furthermore,

$$\pi_{\text{inf}}(\mathbb{R}^+) - \pi_{\text{inf}}(F_1', F_2') = O(\epsilon'). \tag{5.10}$$

Proof. The first part of the theorem can be shown as in the proof of Theorem 5.1. The proof of (5.10) is in Section G. □

Figure 2: Connecting each $x_1 \in F$ (resp. S_0) to the support of a probability that maximises the RHS of (5.4) (resp. (5.5)). For instance, the probability that assigns a weight 2/3 to 100 (represented by L) and 1/3 to 70 (represented by O) maximises the RHS of (5.4) when $x_1 = 90$ (represented by F).

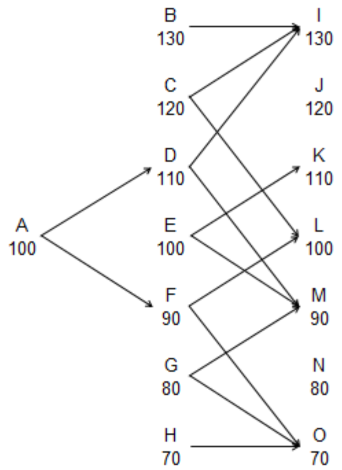


Table 2: The probabilities p_u of an "up" movement and the values of Ψ_i (resp. η_i^*) calculated via Proposition 4.1 (resp. Proposition 4.2) by backward induction inside the tree of Fig. 2, where i is the time-step of the corresponding node.

node	A	B	C	D	E	F	G	H
p_u	1/2	1	2/3	1/2	1/2	2/3	1/2	1
Ψ_i	1.1666	-12	-3.3333	-1	1	3.3333	5	0
η_i^*	7.5	30	20	15	5	0	0	0

Table 3: Optimal super-replication and corresponding hedging of the forward-start option via calls, with $n = 10^6$ and $\epsilon' = 10^{-5}$. The computing time is 33 seconds. The optimal super-replicating price is \$5.2756. The optimal sub-replicating price is \$1.9363, and is obtained in 0.4 seconds. The Black-and-Scholes price is \$3.9878.

Strike	70	80	90	100	110	120	130
b_1	0.1661	-0.2734	-0.4729	-0.4261	-0.4729	0.3096	0.0000
b_2	0.4716	0.4624	0.4366	0.4624	0.4366	0.4624	0.2461

Table 4: The discretization error for optimal super-replication of the forward-start option estimated using $n = 10^6$.

n	Error	Computing time
100	9.7×10^{-3}	0.4
200	1.2×10^{-3}	0.4
400	7.2×10^{-4}	0.4
800	1.8×10^{-4}	0.4
1600	1.3×10^{-5}	0.4

In our numerical example, $t_1 = 1/6$ (two months) and $t_2 = 5/12$ (five months), where t_i is the maturity of time-step i, and the market price $c_{i,j}$ of the call maturing at time-step i with strike $K_{i,j} = 60 + 10j$ is equal to the Black-and-Scholes price with the corresponding strike, maturity t_i and volatility $\sigma = 0.2$ for $1 \le i \le 2$ and $1 \le j \le 7$. We have checked numerically that $\{0, \epsilon'^{-1}(\max \mathcal{K})\}$ is acceptable. Table 3 gives the optimal super-replicating price and the amount of call positions in an optimal super-replicating portfolio. Table 4 gives the discretization error and computing time as a function of n.

5.2 Variance swaps

Consider a variance swap that pays at maturity T the amount

$$\psi = \sum_{i=1}^{m} H(S_{i-1}, S_i),$$

where H is a deterministic bivariate function. For instance, $H(x, y) = T^{-1} \ln^2(y/x)$ for standard variance swaps, $H(x, y) = T^{-1} \ln^2(y/x) 1_{y \in I}$ for a corridor variance swaps, where I is a specified interval of $\mathbb{R}^+$, and $H(x, y) = \max(0, y/x - K)$ for a cliquet call, where K is a constant. In practice, m is quite large and $l_i = 0$ for most values of i. Let $H_i(x, y) = H(x, y) - b_i^T f_i(y)$, so that

$$\Psi = \sum_{i=1}^{m} H_i(S_{i-1}, S_i). \tag{5.11}$$

Let $\pi_{\sup}(F)$ (resp. $\pi_{\inf}(F)$) denote the optimal super-replicating (resp. sub-replicating) price of a standard variance swap ψ in the market $\mathcal{M}_m(F)$. For standard variance swaps, F must be bounded away from 0 in order for the best super-replicating price to be finite. Consider now an interval $[L, M]$ that contains $\mathcal{K}$, with $L > 0$, and let

$$F_0 = \mathcal{K} \cup \{L(M/L)^{j/n}, j \in \{0, \ldots, n\}\}. \tag{5.12}$$

Given $i \in [0, m]$, consider the financial derivative $\zeta = \sum_{j=i+1}^{m} H_j(S_{j-1}, S_j)$. Let Ψ_i (resp. ζ_i) denote the super-hedging cost of Ψ (resp. ζ) at time-step i in $\mathcal{M}_m(F_0)$. Since ζ does not depend on $S_0, \ldots, S_{i-1}$, Remark 4.2 shows that $\zeta_i = c_i(S_i)$, where c_i is a deterministic function on F_0, with $c_m = 0$. Since $\Psi = \zeta + \sum_{j=1}^{i} H_j(S_{j-1}, S_j)$, Remark 4.1 shows that

$$\Psi_i = c_i(S_i) + \sum_{j=1}^{i} H_j(S_{j-1}, S_j), \tag{5.13}$$

and so, by (4.6), for $0 \le i \le m-1$,

$$c_i(S_i) = \sup_{\mathcal{Q} \in \mathbb{Q}(S_i, F_0)} E_{\mathcal{Q}}(c_{i+1} + H_{i+1}(S_i, .)). \tag{5.14}$$

We can interpret (5.14) by observing that the super-hedging cost of ζ at time-step $i+1$ is $c_{i+1}(S_{i+1}) + H_{i+1}(S_i, S_{i+1})$.

Theorem 5.3. *Assume that $\{L, M\}$ is acceptable. Let $B = mT^{-1}\ln^2(M/L)$. Then $\pi_{\sup}(F_0)$ and $\pi_{\inf}(F_0)$ can be calculated in $O(N(l^3 + n^2m + lmn))$ total time with precision ϵ via the convex program* (3.3)*, where*

$$N = O(l\ln(\frac{l(1+S_0)(1+B)}{\epsilon\delta_0(L,M)})),$$

and V' has a separation oracle that runs in $O(n^2m + lmn)$ time. Furthermore,

$$\pi_{\sup}([L,M]) - \pi_{\sup}(F_0) = O(\frac{M}{M-L}\ln^3(\frac{M}{L})\frac{m}{Tn^2}), \tag{5.15}$$

and

$$\pi_{\inf}([L,M]) - \pi_{\inf}(F_0) = O((\ln^2(\frac{M}{L}) + \ln^3(\frac{M}{L}))\frac{m}{Tn^2}). \tag{5.16}$$

Proof. Since $0 \le \psi \le B$, $\max(c(\psi; \mathcal{M}_m(F_0)), c(-\psi; \mathcal{M}_m(F_0))) \le B$. Thus, by Proposition 5.1, Assumptions A1 and A2 hold in $\mathcal{M}_m(F_0)$ if

$$q = B + S_0\sqrt{l} \text{ and } \delta = \delta_0(L,M)/2.$$

We extend the definition of $c_i(x)$ to all $x \in [L, M]$ by setting $c_m = 0$ and, for $0 \le i \le m-1$,

$$c_i(x) = \sup_{\mathcal{Q} \in \mathbb{Q}(x, F_0)} E_{\mathcal{Q}}(c_{i+1} + H_{i+1}(x, .)). \tag{5.17}$$

By Proposition 4.3, we can calculate by backward induction $c_i(x)$ for $x \in F_0$, $0 \le i \le m-1$, and probabilities $\mathcal{Q}(x, i)$ that maximise the RHS of (5.17) in $O(n^2m)$ total time. Using the same notation for θ and x_i as in Assumption A4, let P be a risk-neutral probability obtained by pasting the probabilities $\mathcal{P}(\theta) = \mathcal{Q}(x_i, i)$ in the market $\mathcal{M}_m(F_0)$. By Proposition 4.2, the super-hedging cost $c(\Psi)$ of Ψ in the market $\mathcal{M}_m(F_0)$ equals $E_P(\Psi)$. Given a call option η with maturity $k \le m$ and payoff $\eta^*(S_k)$, define the function η_i^* by backward induction on F_0, $0 \le i \le k$, by setting $\eta_k^* = \eta^*$ and, for $0 \le i \le k-1$ and $x \in F_0$,

$$\eta_i^*(x) = E_{\mathcal{Q}(x,i)}(\eta_{i+1}^*).$$

Using Proposition 4.2, it can be shown by backward induction that $\eta_0^*(S_0) = E_P(\eta)$. Hence, we can calculate in $O(n^2m + lmn)$ time $c(\Psi) = c_0(S_0)$ and $E_P(\phi)$. Thus A3 holds, with $\mathcal{T} = O(n^2m + lmn)$. The first part of the theorem then follows from Theorem 3.2 and (5.3). Section H contains the remainder of the proof. □

Table 5: Optimal super-replicating and sub-replicating prices of a variance swap maturing in one month, with $m = 20$. The discretization error is estimated using $n = 3200$.

n	$\sqrt{\pi_{\text{inf}}(F_0)}$	Computing time	Error	$\sqrt{\pi_{\text{sup}}(F_0)}$	Computing time	Error
50	19.33%	0.4	1.6×10^{-3}	21.67%	0.4	4.3×10^{-4}
100	19.05%	0.7	5.2×10^{-4}	21.72%	0.7	2.0×10^{-4}
200	18.93%	1.7	9.3×10^{-5}	21.75%	1.8	5.2×10^{-5}
400	18.92%	5.6	2.6×10^{-5}	21.76%	5.6	1.1×10^{-5}
800	18.91%	21	3.5×10^{-6}	21.76%	21	2.0×10^{-6}

Table 6: Optimal super-replicating and sub-replicating prices of a variance swap maturing in one month, with $m = 20$, and $n = 800$. The computing time for each price ranged between 21 and 22 seconds.

σ	10%	15%	20%	25%	30%	35%	40%
$\sqrt{\pi_{\text{sup}}(F_0)}$	12.43%	16.92%	21.76%	26.81%	32.01%	37.38%	42.94%
$\sqrt{\pi_{\text{inf}}(F_0)}$	8.68%	13.90%	18.91%	23.77%	28.52%	33.19%	37.78%

In our numerical example, we set $S_0 = \$100$, with $L = \$50$ and $M = \$200$. We first consider a variance swap with maturity T of one month and $m = 20$ daily observations. We assume the market price $c_{m,j}$ of the call maturing at T with strike $K_{m,j} = 65 + 5j$ is equal to the Black-and-Scholes price with the corresponding strike, maturity T and volatility σ for $1 \leq j \leq 13$ and that no other call prices are known. Thus $l_m = 13$ and $l_i = 0$ for $1 \leq i < m$. We have checked numerically that $\{L, M\}$ is acceptable. Table 5 gives the optimal super-replicating and sub-replicating prices, the computing time and the discretization error in terms of n when $\sigma = 0.2$. Table 6 lists the optimal robust bounds and the computing time in terms of the volatility, and table 7 gives the amount of call positions in optimal super-replicating and sub-replicating portfolios when $\sigma = 0.2$. Table 8 gives the optimal super-replicating and sub-replicating prices, the computing time and the discretization error in terms of n when $\sigma = 0.2$, the swap and the calls mature in one year with $m = 252$ daily observations.

As noted before, for standard variance swaps, since ψ is not upper-bounded on $(0, M]$, it follows from (2.3) that $\pi_{sup}((0, M]) = \infty$ for any $M > \max(\mathcal{K})$. In other words, if there is no positive lower bound on stock prices, the best super-replicating price of a standard variance swap is infinite. On the other hand, Section I shows how to calculate $\pi_{inf}((0, M])$ using techniques similar to those of Theorem 5.3, without assuming any positive lower bound on stock prices.

5.3 Volatility swaps

As is sometimes the case in market practice, we assume for simplicity that the volatility swap payment is capped at $\nu_{\max}\sqrt{T^{-1}}$, where $\nu_{\max}$ is a constant, and so the swap pays at maturity T

$$\psi = \sqrt{T^{-1}} \min(\nu_{\max}, \sqrt{\sum_{i=1}^{m} \ln^2(\frac{S_i}{S_{i-1}})}).$$

Table 7: Optimal super-replicating (b_{sup}) and sub-replicating (b_{inf}) portfolios for a variance swap when $\sigma = 0.2$, $m = 20$, and $n = 800$.

Strike	70	75	80	85	90	95	100	105	110	115	120	125	130
b_{sup}	0.434	0.054	0.030	0.023	0.018	0.014	0.012	0.012	0.011	0.011	0.010	0.010	0.057
b_{inf}	-0.19	-0.012	0.015	0.014	0.013	0.012	0.012	0.010	0.008	0.006	0.004	0.003	0.001

Table 8: Super-replicating and sub-replicating prices of a variance swap maturing in one year, with $m = 252$. The discretization error is estimated using $n = 3200$.

n	$\sqrt{\pi_{\inf}(F_0)}$	Computing time	Error	$\sqrt{\pi_{\sup}(F_0)}$	Computing time	Error
50	18.30%	3.1	4.9×10^{-4}	22.76%	3.4	4.9×10^{-4}
100	18.24%	8.6	2.6×10^{-4}	22.81%	8.1	2.4×10^{-4}
200	18.20%	26	1.1×10^{-4}	22.84%	25	1.1×10^{-4}
400	18.18%	93	4.5×10^{-5}	22.86%	87	4.1×10^{-5}
800	18.17%	343	1.2×10^{-5}	22.86%	327	1.3×10^{-5}

In order to construct the subroutine needed in Assumption A3, we will discretize both the stock price and the realized volatility. We show how to discretize the realized volatility at all time-steps. Let n' be an integer,

$$\Lambda = \{\frac{i\nu_{\max}}{n'} : 0 \le i \le n'\}, \text{ and}$$

$$\rho(z) = \max\{x \in \Lambda : x \le z\},$$

for $z \ge 0$. Define the financial derivative ν_i, $0 \le i \le m$, by induction by setting $\nu_0 = 0$ and

$$\nu_{i+1} = \rho(\sqrt{\nu_i^2 + \ln^2(\frac{S_{i+1}}{S_i})}). \tag{5.18}$$

Consider the financial derivative $\psi' = \nu_m$. Since the function $\nu \mapsto \sqrt{\nu^2 + a^2}$ is 1-Lipschitz for any constant a, it can be shown by induction that

$$0 \le \psi\sqrt{T} - \psi' \le \frac{m\nu_{\max}}{n'}. \tag{5.19}$$

Thus, ψ' is a discrete approximation of $\psi\sqrt{T}$. Let $\Psi' = \psi' - \beta^T\phi$. Given integer $i \in [0, m]$, let

$$\zeta = \psi' - \sum_{j=i+1}^{m} b_j^T f_j(S_j). \tag{5.20}$$

By (5.18), ψ' is a deterministic function of ν_i and of S_j, $j \ge i$. Let Ψ'_i (resp. ζ_i) denote the super-hedging cost of Ψ' (resp. ζ) at time-step i in the market $\mathcal{M}_m(F_0)$, where F_0 is given by (5.12). By Remark 4.2, it follows that $\zeta_i = c_i(\nu_i, S_i)$, where c_i is a deterministic function defined on $\Lambda \times F_0$. By (5.20), $c_m(\nu_m, S_m) = \nu_m$. Since $\Psi' = \zeta - \sum_{j=1}^{i} b_j^T f_j(S_j)$, Remark 4.1 implies that

$$\Psi'_i = c_i(\nu_i, S_i) - \sum_{j=1}^{i} b_j^T f_j(S_j).$$

Thus, by (4.6) and (5.18),

$$c_i(\nu_i, S_i) = \sup_{\mathcal{Q} \in \mathbb{Q}(S_i, F_0)} E_{\mathcal{Q}}(h_i(\nu_i, S_i, .)), \tag{5.21}$$

where, for $z > 0$,

$$h_i(\nu, x, z) = c_{i+1}(\rho(\sqrt{\nu^2 + \ln^2(z/x)}), z) - b_{i+1}^T f_{i+1}(z).$$

We can interpret (5.21) by observing that the super-hedging cost of ζ at time-step $i+1$ is $c_{i+1}(\nu_{i+1}, S_{i+1}) - b_{i+1}^T f_{i+1}(S_{i+1})$, which equals $h_i(\nu_i, S_i, S_{i+1})$.

Denote by $\pi_{\sup}([L,M])$ the optimal super-replicating price of ψ in the market $\mathcal{M}_m([L,M])$ and by $\pi_{\sup}(n,n')$ the optimal super-replicating price of ψ' in the market $\mathcal{M}_m(F_0)$. Define similarly $\pi_{\inf}([L,M])$ and $\pi_{\inf}(n,n')$. Theorem 5.4 shows how to calculate $\pi_{\sup}(n,n')$ and $\pi_{\inf}(n,n')$ and gives an upper bound on the discretization error. The constants behind the O notation in (5.22), (5.23) and Section J depend on l, S_0, the calls strikes, maturities and prices, $\nu_{\max}$, L and M, but do not depend on n, n', m and T.

Theorem 5.4. *Assume that $\{L,M\}$ is acceptable. Then $\pi_{\sup}(n,n')$ and $\pi_{\inf}(n,n')$ can be calculated in $O(N(l^3+nn'm(n+l)))$ total time with precision ϵ via the convex program* (3.3)*, where*

$$N = O(l\ln(\frac{l(1+S_0)(1+\nu_{\max})}{\epsilon\delta_0(L,M)})).$$

V' has a separation oracle that runs in $O(nn'm(n+l))$ time. Furthermore,

$$|\pi_{\sup}([L,M]) - \pi_{\sup}(n,n')\sqrt{T^{-1}}| \leq \frac{2m\nu_{\max}}{\sqrt{Tn'}} + O(\frac{m^2}{\sqrt{Tn}}), \quad (5.22)$$

$$|\pi_{\inf}([L,M]) - \pi_{\inf}(n,n')\sqrt{T^{-1}}| \leq \frac{2m\nu_{\max}}{\sqrt{Tn'}} + O(\frac{m^2}{\sqrt{Tn}}). \quad (5.23)$$

Proof. Since $0 \leq \psi' \leq \nu_{\max}$, $\max(c(\psi'), c(-\psi')) \leq \nu_{\max}$. Thus, by Proposition 5.1, Assumptions A1 and A2 hold for (ψ',ϕ) in the market $\mathcal{M}_m(F_0)$ if

$$q = \nu_{\max} + S_0\sqrt{l} \text{ and } \delta = \delta_0(L,M)/2. \quad (5.24)$$

Extend the definition of $c_i(\nu,x) = c_i(\nu,x,\beta)$ to $i \in [0,m]$, $\nu \in \Lambda$ and $x \in F_0$ by setting $c_m(\nu,x) = \nu$ and

$$c_i(\nu,x) = \sup_{\mathcal{Q}\in\mathbb{Q}(x,F_0)} E_{\mathcal{Q}}(h_i(\nu,x,.)). \quad (5.25)$$

By Proposition 4.3, we can calculate by backward induction $c_i(\nu,x)$ for $\nu \in \Lambda$, $x \in F_0$ and $0 \leq i \leq m-1$, and probabilities $\mathcal{Q}(\nu,x,i)$ that maximise the RHS of (5.25) in $O(n^2n'm)$ total time. Define the function with i variables ν_i^*, $0 \leq i \leq m$, by induction by setting $\nu_0^* = 0$ and

$$\nu_{i+1}^*(x_1,\ldots,x_{i+1}) = \rho(\sqrt{\nu_i^{*2}(x_1,\ldots,x_i) + \ln^2(\frac{x_{i+1}}{x_i})}).$$

Note that $\nu_i = \nu_i^*(S_1,\ldots,S_i)$. Using the same notation for θ and x_i as in Assumption A4, let P be a risk-neutral probability obtained by pasting the probabilities $\mathcal{P}(\theta) = \mathcal{Q}(\nu_i^*(\theta),x_i,i)$. By Proposition 4.2, $E_P(\Psi')$ is equal to the super-hedging cost $c(\Psi')$ of Ψ' in the market $\mathcal{M}_m(F_0)$. Given a call option η with maturity $k \leq m$ and payoff $\eta^*(S_k)$, define the function η_i^* by backward induction on $\Lambda \times F_0$, $0 \leq i \leq k$, by setting $\eta_k^*(\nu,x) = \eta^*(x)$ and, for $0 \leq i \leq k-1$, $\nu \in \Lambda$, and $x \in F_0$,

$$\eta_i^*(\nu,x) = E_{\mathcal{Q}(\nu,x,i)}(\eta_k^*(\nu,x,.)),$$

where $\eta_k^*(\nu,x,z) = \eta_{i+1}^*(\rho(\sqrt{\nu^2+\ln^2(z/x)}),z)$. Using Proposition 4.2, it can be shown by backward induction that $\eta_0^*(0,S_0) = E_P(\eta)$. Hence, we can calculate in $O(nn'm(n+l))$ time $c(\Psi') = c_0(0,S_0)$ and $E_P(\phi)$, and so A3 holds for (ψ',ϕ) in the market $\mathcal{M}_m(F_0)$, with $\mathcal{T} = O(nn'm(n+l))$. The first part of the theorem then follows from Theorem 3.2 and (5.3). Section J contains the proof of (5.22) and (5.23). □

Our numerical experiments use the same setting as in Subsection 5.2 and cap the volatility swap payoff at $\sigma\sqrt{2.5}$. Rather than rounding with respect to ν in the calculation of $c_i(\nu,x)$, we have used linear interpolation, as described in Section K, which performs much better in practice. When calculating the best super-replicating price, we set $n' = n$. When calculating the

Table 9: Optimal super-replicating and sub-replicating prices of a capped volatility swap maturing in one month. The discretization error is estimated using $n = 1600$ and $n' = 137$ for the sub–replicating price, and $n = n' = 800$ for the super-replicating price.

n	n'	π_{inf}	Computing time	Error	n	n'	π_{sup}	Computing time	Error
50	14	8.03%	3.5	3.7×10^{-3}	25	25	20.18%	2.5	1.1×10^{-2}
100	22	7.77%	16	9.9×10^{-4}	50	50	21.23%	13	6.3×10^{-5}
200	34	7.70%	82	3.2×10^{-4}	100	100	21.23%	67	6.5×10^{-6}
400	54	7.67%	525	5.2×10^{-6}	200	200	21.23%	466	4.6×10^{-6}

Table 10: Optimal super-replicating and sub-replicating prices of a capped volatility swap maturing in one month. The super-replicating prices were obtained with $n = n' = 200$. The sub-replicating prices were obtained with $n = 400$ and $n' = 54$.

σ	10%	15%	20%	25%	30%	35%	40%
π_{sup}	12.24%	16.59%	21.23%	26.01%	30.88%	35.82%	40.85%
Running time	471	477	469	473	485	477	494
π_{inf}	3.74%	5.55%	7.67%	9.61%	11.56%	13.60%	15.49%
Running time	502	531	525	545	521	537	534

best sub-replicating price, we first calculate a discrete risk-neutral probability consistent with the call prices at maturity T (see, e.g., (Davis and Hobson 2007)), then replace F_0 by another set of size $n' = n^{2/3}$ where, on each interval I delimited by consecutive elements in $\{L, M\} \cup \mathcal{K}$, the points are geometrically distributed and their number is proportional to the risk-neutral probability that S_m belongs to I. This resulted in significantly improved performance in all our numerical experiments. Tables 9 and 12 estimate the discretization error for maturities of one month and one year, respectively. Table 10 lists the optimal super-replicating and sub-replicating prices for the swap, and Table 11 gives the amount of call positions in optimal super-replicating and sub-replicating portfolios when $\sigma = 0.2$.

5.4 Discussion of results

A discretization error of ϵ can be achieved in Theorems 5.1, 5.2, 5.3, and 5.4 by using a cutting plane algorithm with $O(\epsilon^{-1/2})$, $O(1)$, $O(\epsilon^{-1})$, and $O(\epsilon^{-3})$ running time per iteration, respectively, and so a global error of ϵ on the robust prices in the infinite space markets can be achieved with $O(\epsilon^{-1/2}\ln(\epsilon^{-1}))$, $O(\ln(\epsilon^{-1}))$, $O(\epsilon^{-1}\ln(\epsilon^{-1}))$, and $O(\epsilon^{-3}\ln(\epsilon^{-1}))$ total time. Our numerical results confirm that the tradeoff between the discretization error and the running time of the robust pricing algorithms for forward start options, variance and volatility swaps is best for forward start options and worst for volatility swaps. Theorems 5.3 and 5.4 show that, for fixed N, l, n and (for volatility swaps) n', robust prices of variance and volatility swaps are computed in time asymptotically proportional to the number of periods m, as confirmed numerically in Section L. On the other hand, Tables 10 and 12 suggest that the model risk for volatility swaps is higher than that for variance swaps. This may be explained by the fact that, under a continuity assumption on the stock price, the price of a continuously-monitored variance swap can be exactly determined (Dupire 1993, Neuberger 1994) from the prices of a continuum set of co-maturing call options, which is not the case for continuously-monitored volatility swaps.

5.5 Other financial derivatives

Other financial derivatives such as lookback options, options on realized variance and realized volatility, single or double barrier options and Asian options can be handled in a similar fashion.

Table 11: Optimal super-replicating ($b_{\sup}$) and sub-replicating ($b_{\inf}$) portfolios for a capped volatility swap when $\sigma = 0.2$, using the same grids and in Table 10.

Strike	70	75	80	85	90	95	100	105	110	115	120	125	130
$b_{\sup}$	0.402	0.044	0.039	0.036	0.033	0.031	0.028	0.026	0.024	0.023	0.021	0.019	0.024
$b_{\inf}$	-0.016	-0.001	0.000	-0.001	-0.027	-0.008	0.042	-0.008	-0.027	-0.001	0.000	0.000	0.000

Table 12: Optimal super-replicating and sub-replicating prices of a capped volatility swap maturing in one year. The discretization error is estimated using $n = 800$ and $n' = 86$ for the sub–replicating price, and $n = n' = 400$ for the super-replicating price.

n	n'	$\pi_{\inf}$	Computing time	Error	n	n'	$\pi_{\sup}$	Computing time	Error
50	14	8.04%	37	1.2×10^{-2}	25	25	20.98%	29	7.0×10^{-4}
100	22	7.63%	179	7.5×10^{-3}	50	50	20.97%	146	5.4×10^{-4}
200	34	7.30%	967	4.2×10^{-3}	100	100	20.94%	903	2.8×10^{-4}
400	54	7.06%	6463	1.8×10^{-3}	200	200	20.92%	6147	4.3×10^{-5}

Consider for instance an Asian call ψ that pays

$$\max(\frac{S_1 + \cdots S_m}{m} - K, 0).$$

Let $F_0 = \mathcal{K} \cup \{jM/n : 0 \leq j \leq n\}$, where $M > S_0$ and integer n are fixed. Using the same techniques as before, we can show that the super-hedging cost of Ψ at time-step i in $\mathcal{M}_m(F_0)$ is

$$\Psi_i = c_i(\nu_i, S_i) - \sum_{j=1}^{i} b_j^T f_j(S_j),$$

where $\nu_i = \Sigma_{j=1}^{i-1} S_j$. The functions c_i can be calculated by backward induction for $x \in F_0$ and $\nu \in [0, mM]$, by setting $c_m(\nu, x) = \max((\nu + x)/m - K, 0)$ and

$$c_i(\nu, x) = \sup_{\mathcal{Q} \in \mathbb{Q}(x, F_0)} E_{\mathcal{Q}}(c_{i+1}(\nu + x, .) - b_{i+1}^T f_{i+1}).$$

In practice, we need to discretize the values that ν can take, in the same manner as we did for volatility swaps. We can also derive optimal replicating prices and strategies for the financial derivatives considered in this section in terms of prices of moments (or of the logarithmic function) of the stock price at different maturities rather than in terms of call prices.

6 Extensions

6.1 Taking interest rates and dividends into account

Interest rates can be taken into account in the usual way by replacing each security by its discounted value. We can also incorporate deterministic dividends as well as proportional dividends in our model by replacing each security by the value of the security plus reinvested dividends, using a technique similar to the one in (Pliska 2005, SubSec 3.2.3).

6.2 Taking bid-ask spreads into account

Assume that the financial derivatives ϕ_k, $1 \leq k \leq l$, have distinct bid and ask prices, and let π^b (resp. π^a) be the length l vector of bid (resp. ask) prices. Results in the preceding sections can be easily extended to this case. We have for instance the following.

Theorem 6.1. *If assumptions A1 and A3 hold and there is a risk-neutral probability P such that*

$$E_P(\phi) \in [\pi^b + \delta \mathbf{1}_l, \pi^a - \delta \mathbf{1}_l] \tag{6.1}$$

and $-q \le E_P(\psi)$, then $\pi_{\sup}$ can be calculated via the convex program with $2l+1$ variables

$$\pi_{\sup} = \inf_{(\beta^a, \beta^b, \gamma) \in V''} a_0^T(\beta^a, \beta^b, \gamma),$$

where $a_0 = (\pi^a, -\pi^b, 1)$, and

$$V'' = \{(\beta^a, \beta^b, \gamma) \in \mathbb{R}^{+l} \times \mathbb{R}^{+l} \times \mathbb{R} : a_0^T(\beta^a, \beta^b, \gamma) \le q + 1, (\beta^a - \beta^b, \gamma) \in V\}.$$

Furthermore, for $(\beta^a, \beta^b, \gamma) \in V''$,

$$||(\beta^a, \beta^b, \gamma)|| \le \frac{4(1+q)\sqrt{2l}(1 + ||(\pi^a, \pi^b)||)}{\delta}, \tag{6.2}$$

V'' contains a ball of radius $r = \delta ||a_0||^{-1} l^{-1/2}/3$ centered at $(r\mathbf{1}_{2l}, q+\delta)$, and admits a separation oracle that runs in $\mathcal{T} + O(1)$ time, where $\mathcal{T}$ is the running time of the subroutine in Assumption A3.

The vector β^a (resp. β^b) represents the amount of assets bought (resp. sold) in order to super-replicate ψ.

6.3 Limiting the jump sizes or the realized volatility

We can tighten the bounds on the price of ψ by limiting the underlying dynamics. Limitations to the up and/or down jumps can be achieved by limiting the set $D(\theta)$ in (4.5) to the securities values that respect these limitations. For instance, if the log-returns are constrained to be at most $3\sigma\sqrt{T/m}$ in absolute value at any time-step, $\mathcal{Q}$ must be supported on the set

$$\{z \in F_0 : |\ln(\frac{z}{x})| \le 3\sigma\sqrt{\frac{T}{m}}\}$$

in (5.25), and the best sub-replicating price for volatility swaps becomes 10.70% when $\sigma = 0.2$. We can upper-limit the realized volatility to $\sigma\sqrt{2.5}$ by setting the payoff at maturity to $-\infty$ if the realized volatility exceeds $\sigma\sqrt{2.5}$ and keeping the remaining calculations unchanged, which yields a best sub-replicating price of 13.09%. If both the log-returns and the realized volatility are required to satisfy the previous constrains, the best sub-replicating price becomes 13.95%, while the best super-replicating price remains essentially the same as in the un-constrained case. These prices have been calculated using the same setting as in Table 10. The running times for the three experiments were respectively 432, 614, and 463 seconds.

6.4 Portfolios of financial derivatives

Consider a portfolio of financial derivatives. Using (2.3), it can be shown that the optimal super-replicating (resp. sub-replicating) portfolio price given other derivatives prices is at most (resp. at least) the sum of the optimal super-replicating (resp. sub-replicating) prices of the portfolio components. Using the same setting as in Table 8 with $n = 800$, Table 13 lists the optimal super-replicating and sub-replicating prices for a standard variance swap, a variance swap where $H(x, y) = T^{-1}(y/x - 1)^2$, and the average of the two swaps. The gap between $\pi_{\inf}$ and $\pi_{\sup}$ for the average variance swap turned out to be less than that for each component. In general, our method allows the efficient calculation of optimal robust bounds on a portfolio of derivatives of the type considered in Section 5, such as barrier options with different barrier levels and strikes.

Table 13: Optimal super-replicating and sub-replicating prices for variance swaps based on Log-returns, returns, and their average, with $n = 800$. The running times ranged from 327 to 369 seconds.

	Log-returns	Returns	Average
$\sqrt{\pi_{\text{sup}}}$	22.86%	26.53%	22.90%
$\sqrt{\pi_{\text{inf}}}$	18.17%	16.98%	19.26%

7 Conclusion

We have shown that optimal super-replicating and sub-replicating prices and corresponding hedging portfolios can be calculated efficiently for a wide variety of exotic financial derivatives in terms of liquid financial derivatives in a discrete-time setting. The main novelty behind our approach is the use of convex programs, which are solved via cutting planes generated by risk-neutral probabilities. Super-hedging costs are calculated via recursive evaluations of concave envelopes. We have implemented our method using an analytic center cutting plane algorithm and an optimized convex hull algorithm. Numerical calculations of optimal super-replicating and sub-replicating prices in terms of call options were given for forward start options, variance and volatility swaps. In our examples, we have discretized the state-space and gave theoretical and empirical bounds on the discretization errors. These prices are close to those obtained by a standard model in some cases and differ considerably from them in other cases. Our method is much more flexible than known explicit methods: it can incorporate interest rates and dividends, bid-ask spreads, limitations to the jumps or to the realized volatility of the underlying assets, and can be used to calculate efficiently optimal prices on portfolio of options of certain type. It can be applied to multi-period financial derivatives on multiple assets but, in general, the corresponding running time is exponential in the number of assets. This is because, in general, the number of points needed to discretize the possible values of the assets vector is exponential in the number of assets.

8 Acknowledgments

This paper has been presented at the 30th French Finance Association Conference, the European Financial Management Association 2013 Annual Meetings, and the Advances in Financial Mathematics 2014 conference in Paris. The author thanks seminar participants, two anonymous referees, an anonymous associate editor, and Gustavo Manso (department editor), for helpful comments and suggestions. This work was achieved through the Laboratory of Excellence on Financial Regulation (Labex ReFi) under the reference ANR-10-LABX-0095. It benefitted from a French government support managed by the National Research Agency (ANR).

References

Albrecher, H., Mayer, P. A. and Schoutens, W. (2008). General lower bounds for arithmetic Asian option prices, *Applied Mathematical Finance* **15**(2): 123 – 149.

Andrew, A. M. (1979). Another efficient algorithm for convex hulls in two dimensions, *Inf. Process. Lett.* **9**(5): 216–219.

Atkinson, D. S. and Vaidya, P. M. (1995). A cutting plane algorithm for convex programming that uses analytic centers, *Mathematical Programming* **69**(1-3): 1–43.

Beiglböck, M., Henry-Labordère, P. and Penkner, F. (2013). Model-independent bounds for option prices-a mass transport approach., *Finance and Stochastics* **17**(3): 477 – 501.

Bertsimas, D. and Popescu, I. (2002). On the relation between option and stock prices: a convex optimization approach, *Operations Research* **50**(2): 358–374.

Bertsimas, D. and Vempala, S. (2004). Solving convex programs by random walks, *Journal of the ACM (JACM)* **51**(4): 540–556.

Bonnans, J. F. and Tan, X. (2011). A model-free no-arbitrage price bound for variance options, *Applied Mathematics & Optimization* pp. 1–31.

Boyd, S. and Vandenberghe, L. (2004). *Convex optimization*, Cambridge university press, UK.

Boyd, S., Vandenberghe, L. and Skaf, J. (2008). Analytic center cutting-plane method. Stanford University, California, USA, Lecture notes, http://www. stanford. edu/class/ee364b.

Boyle, P. P. and Lin, X. S. (1997). Bounds on contingent claims based on several assets, *Journal of Financial Economics* **46**(3): 383–400.

Britten-Jones, M. and A. Neuberger, A. (2000). Option prices, implied price processes, and stochastic volatility, *Journal of Finance* **55**(2): 839–66.

Broadie, M. and Jain, A. (2008). The effect of jumps and discrete sampling on volatility and variance swaps, *Int. J. Theor. Appl. Finance* **11**(8): 761–797.

Brown, H., Hobson, D. and Rogers, L. C. G. (2001a). The maximum maximum of a martingale constrained by an intermediate law, *Probab. Theory Related Fields* **119**(4): 558–578.

Brown, H., Hobson, D. and Rogers, L. C. G. (2001b). Robust hedging of barrier options, *Math. Finance* **11**(3): 285–314.

Carassus, L., Gobet, E. and Temam, E. (2007). A class of financial products and models where super-replication prices are explicit, *in* J. Akahori, S. Ogawa and S. Watanabe (eds), *Proc. Ritsumeikan Int. Symp. on Stoch. Proc. and Apps to Math. Fin.*, World Scientific, Kusatsu, Japan, pp. 67–84.

Carr, P., Lee, R. and Wu, L. (2012). Variance swaps on time-changed Levy processes, *Finance and Stochastics* **16**(2): 335–355.

Carr, P. P. and Wu, L. (2009). Variance Risk Premiums, *The Review of Financial Studies,* **22**(3): 1311–1341.

Committee on Banking Supervision, B. (2009). Revisions to the Basel II market risk framework, *Technical report*, Bank for International Settlements.

Cox, A. M. and Obłój, J. (2011a). Robust hedging of double touch barrier options, *SIAM Journal on Financial Mathematics* **2**(1): 141–182.

Cox, A. M. and Obłój, J. (2011b). Robust pricing and hedging of double no-touch options, *Finance and Stochastics* **15**(3): 573–605.

Cox, A. M. and Wang, J. (2013). Roots barrier: construction, optimality and applications to variance options, *The Annals of Applied Probability* **23**(3): 859–894.

d'Aspremont, A. and El Ghaoui, L. (2006). Static arbitrage bounds on basket option prices, *Math. Program.* **106**(3, Ser. A): 467–489.

Davis, M. H. A. and Hobson, D. G. (2007). The range of traded option prices, *Math. Finance* **17**(1): 1–14.

Davis, M., Obloj, J. and Raval, V. (2014). Arbitrage bounds for prices of weighted variance swaps, *Math. Finance* **24**(4): 821–854.

Demeterfi, K., Derman, E., Kamal, M. and Zou, J. (1999). A guide to volatility and variance swaps, *Journal of Derivatives* **6**(4): 9–32.

Dupire, B. (1993). Model art, *Risk* **6**: 118–120.

Follmer, H. and Schied, A. (2004). *Stochastic finance; An Introduction in Discrete Time, 2nd Edition*, Walter de Gruyter, Berlin.

Galichon, A., Henry-Labordere, P. and Touzi, N. (2014). A stochastic control approach to no-arbitrage bounds given marginals, with an application to Lookback options., *The Annals of Applied Probability* **24**(1): 312–336.

Gotoh, J.-y. and Konno, H. (2002). Bounding option prices by semidefinite programming: A cutting plane algorithm, *Management Science* **48**(5): 665–678.

Grötschel, M., Lovász, L. and Schrijver, A. (1981). The ellipsoid method and its consequences in combinatorial optimization, *Combinatorica* **1**(2): 169–197.

Henry-Labordère, P. (2013). Automated option pricing: Numerical methods, *International Journal of Theoretical and Applied Finance* **16**(8).

Henry-Labordere, P., Obloj, J., Spoida, P. and Touzi, N. (2013). Maximum maximum of martingales given marginals. Available at SSRN 2031461.

Henry-Labordere, P. and Touzi, N. (2013). An explicit martingale version of Brenier's theorem. Available at SSRN: http://ssrn.com/abstract=2218488.

Hobson, D. G. (1998). Robust hedging of the lookback option, *Finance and Stochastics* **2**(4): 329–347.

Hobson, D. G. and Pedersen, J. L. (2002). The minimum maximum of a continuous martingale with given initial and terminal laws, *Ann. Probab.* **30**(2): 978–999.

Hobson, D. and Klimmek, M. (2012). Model-independent hedging strategies for variance swaps, *Finance and Stochastics* **16**: 611–649.

Hobson, D. and Klimmek, M. (2015). Robust price bounds for the forward starting straddle, *Finance and Stochastics* **19**(1): 189–214.

Hobson, D., Laurence, P. and Wang, T.-H. (2005a). Static-arbitrage optimal subreplicating strategies for basket options, *Insurance Math. Econom.* **37**(3): 553–572.

Hobson, D., Laurence, P. and Wang, T.-H. (2005b). Static-arbitrage upper bounds for the prices of basket options, *Quant. Finance* **5**(4): 329–342.

Hobson, D. and Neuberger, A. (2012). Robust bounds for forward start options, *Math. Finance* **22**(1): 31–56.

Hull, J. (2012). *Options, Futures and Other Derivatives*, eigth edn, Pearson Education Limited, England.

Kahalé, N. (2010). Sparse calibrations of contingent claims, *Math. Finance* **20**(1): 105–115.

Kahalé, N. (2014). Model-independent lower bound on variance swaps. http://dx.doi.org/10.1111/mafi.12083. To appear in Mathematical Finance.

Laurence, P. and Wang, T.-H. (2005). Sharp upper and lower bounds for basket options, *Applied Mathematical Finance* **12**(3): 253–282.

Laurence, P. and Wang, T.-H. (2008). Distribution-free upper bounds for spread options and market-implied antimonotonicity gap, *The European Journal of Finance* **14**(8): 717 – 734.

Laurence, P. and Wang, T.-H. (2009). Sharp distribution free lower bounds for spread options and the corresponding optimal subreplicating portfolios, *Insurance: Mathematics and Economics* **44**(1): 35–47.

Neuberger, A. J. (1994). The log contract, *Journal of Portfolio Management* **20**(2): 74–80.

Pliska, S. (2005). *Introduction to Mathematical Finance*, Blackwell, Malden, MA.

Simon, S., Goovaerts, M. and Dhaene, J. (2000). An easy computable upper bound for the price of an arithmetic Asian option, *Insurance: Mathematics and Economics* **26**(2): 175–183.

Tan, X. and Touzi, N. (2013). Optimal transportation under controlled stochastic dynamics, *Ann. Probab.* **41**(5): 3201–3240.

Vaidya, P. M. (1996). A new algorithm for minimizing convex functions over convex sets, *Mathematical programming* **73**(3): 291–341.

A Proof of Lemma 3.1

By Assumption A2 and (3.2), $V \subseteq \mathcal{E}$, and so $V' \subseteq \mathcal{E}'$. For $(\beta, \gamma) \in \mathcal{E}'$ and $1 \leq i \leq l$, it follows from (3.4) that $-q \leq \beta^T \pi + \gamma \pm \delta\beta_i$. Thus, by (3.5), $-q \leq q + 1 \pm \delta\beta_i$, and so $|\beta_i| \leq R_0$. Hence $||\beta|| \leq R_0\sqrt{l}$. On the other hand, (3.4) implies that $-q \leq \beta^T\pi + \gamma$ which, together with (3.5), shows that

$$|\gamma| \leq q + 1 + |\beta^T \pi|.$$

By the Cauchy-Schwartz inequality, it follows that

$$|\gamma| \leq q + 1 + R_0||\pi||\sqrt{l}.$$

Since $||(\beta, \gamma)|| \leq ||\beta|| + |\gamma|$ and $q + 1 \leq R_0$, we conclude that $||(\beta, \gamma)|| \leq R_1$.

Consider now a vector $(\beta, \gamma) \in \mathbb{R}^l \times \mathbb{R}$ that belongs to the ball of radius $r = \delta(1 + ||\pi||)^{-1}$ centered at $(\mathbf{0}_l, q+\delta)$. Since $||(\beta, q+\delta-\gamma)|| \leq r \leq \delta$, $||\beta|| \leq \delta$ and $q \leq \gamma$. Thus, by Assumption A1 and (2.4), $(\beta, \gamma) \in V$. On the other hand, since

$$a_0^T(\beta, \gamma) = a_0^T(\beta, \gamma - q - \delta) + q + \delta,$$

where $a_0 = (\pi, 1)$, it follows from the Cauchy-Schwartz inequality that $a_0^T(\beta, \gamma) \leq r||a_0|| + q + \delta$. Since $||a_0|| \leq 1 + ||\pi||$, we conclude that $a_0^T(\beta, \gamma) \leq q + 1$. Hence $(\beta, \gamma) \in V'$. □

B Proof of Lemma 3.2

For shorthand, denote a vector $(\beta, \gamma) \in \mathbb{R}^l \times \mathbb{R}$ by x. Let

$$K = \{x \in \mathbb{R}^{l+1} : a_i' \leq a_i^T x \text{ for } i \in I'\}, \tag{B.1}$$

$$K' = \{x \in K : a_0^T x \leq q + 1\}. \tag{B.2}$$

By (3.9) and (3.10) and linear program duality,

$$b_0 = \inf_{x \in K} a_0^T x. \tag{B.3}$$

Note that the last $2l$ constraints in the RHS of (B.1) are the same as those in (3.4). Thus,

$$K = \{x \in \mathcal{E} : a_i' \le a_i^T x \text{ for } i \in I\},$$

$$K' = \{x \in \mathcal{E}' : a_i' \le a_i^T x \text{ for } i \in I\}.$$

But $\mathcal{E}' \subseteq \mathcal{R}_0$ since $B(0, R_1) \subseteq \mathcal{R}_0$ and, by Lemma 3.1, $\mathcal{E}' \subseteq B(0, R_1)$. As

$$\tilde{C} = \{x \in \mathcal{R}_0 : a_i' \le a_i^T x \text{ for } i \in I\},$$

we conclude that $K' = \mathcal{E}' \cap \tilde{C}$. Since $V' \subseteq \mathcal{E}'$ by Lemma 3.1 and $V' \subseteq \tilde{C}$, it follows that $V' \subseteq K'$ and so K' is non-empty. Thus, by (B.2) and (B.3), $b_0 = \inf_{x \in K'} a_0^T x$. Since $K' \subseteq \tilde{C}$, this concludes the proof. □

C Proof of Proposition 4.1

By Assumption A4, x_i is a convex combination of elements of $D(\theta)$, and so $\mathbb{Q}(x_i, D(\theta))$ is non-empty. (4.6) follows immediately from (4.1). By (2.1), for any $\epsilon > 0$, there is a gains function $g = \sum_{j=1}^m \xi_j^T (X_j - X_{j-1})$ such that Ψ is upper-bounded by $c(\Psi) + \epsilon + g$. Note that, for $j \in [1, m]$, there is a deterministic function ξ_j^* on D_{j-1} such that $\xi_j = \xi_j^*(X_1, \ldots, X_{j-1})$. We show by backward induction that

$$\Psi_i \le c(\Psi) + \epsilon + \sum_{j=1}^{i} \xi_j^T (X_j - X_{j-1}). \tag{C.1}$$

(C.1) clearly holds when $i = m$. If it holds for $i+1$ then, for $\theta = (x_1, \ldots, x_i) \in D_i$ and $x \in D(\theta)$,

$$\Psi_{i+1}^*(\theta, x) \le c(\Psi) + \epsilon + {\xi_{i+1}^*}^T (x - x_i) + \sum_{j=1}^{i} {\xi_j^*}^T (x_j - x_{j-1}), \tag{C.2}$$

where we denote $\xi_j^*(x_1, \ldots, x_{j-1})$ by ξ_j^* for shorthand. Since the RHS of (C.2) is a linear function of x, it upper bounds $\overline{\Psi_{i+1}^*(\theta, .)}(x)$ for $x \in \widehat{D(\theta)}$. Hence $\Psi_i^*(\theta)$ is well defined,

$$\Psi_i^*(\theta) \le c(\Psi) + \epsilon + \sum_{j=1}^{i} {\xi_j^*}^T (x_j - x_{j-1}),$$

and (C.1) holds for i. Thus $\Psi_0 \le c(\Psi)$.

Conversely, for $0 \le i \le m-1$, by the definition of Ψ_i^* and (4.4), there is a financial derivative ξ_{i+1} which is a function of $X_0, \ldots, X_i$ such that

$$\Psi_{i+1} \le \xi_{i+1}^T (X_{i+1} - X_i) + \Psi_i.$$

Hence there is a gains function g such that $\Psi \le g + \Psi_0$. Thus $c(\Psi) \le \Psi_0$, and so $\Psi_0 = c(\Psi)$. □

D Proof of Proposition 4.2

For $(x_1, \ldots, x_m) \in D_m$, choose a state $\omega \in \Omega$ such that $X_j(\omega) = x_j$ for $1 \le j \le m$, and set

$$P(\{\omega\}) = \Pi_{i=1}^m \mathcal{P}_{(x_1, \ldots, x_{i-1})}(\{x_i\}).$$

The reader can verify that P is a probability, and that $E_P(\eta) = \eta_0^*(\emptyset)$ if η^* is the indicator function of a path $\theta \in D_m$. By linearity of expectations, if follows that $E_P(\eta) = \eta_0^*(\emptyset)$ for any function η^*. Let $\eta = \sum_{j=1}^m \xi_j^T(X_j - X_{j-1})$ be a gains function. Using backward induction and (4.7), it can be shown that, for $0 \le i \le m$,

$$\eta_i^*(X_1, \ldots, X_i) = \sum_{j=1}^{i} \xi_j^T(X_j - X_{j-1}).$$

Thus $E_P(\eta) = 0$ and P is risk-neutral.
Assume now that $\mathcal{P}(\theta)$ attains the RHS of (4.6). It can be shown by backward induction that $\Psi_i^* = \bar{\Psi}_i$, where Ψ_i^* is defined as in Proposition 4.1 and $\bar{\Psi}_i$ is the sequence obtained by replacing η^* with Ψ^* in (4.7), and so $\Psi_0 = E_P(\psi)$. □

E Proof of Proposition 5.1

Assume that δ and q satisfy the conditions in Proposition 5.1. For any given integers $i_0 \in [1, m]$ and $j_0 \in [1, l_i]$, if c_{i_0,j_0} is replaced with $c_{i_0,j_0} \pm \delta$ and the other call prices remain unchanged, then (5.2) still holds. On the other hand, it follows from the proof of (Davis and Hobson 2007, Theorem 4.2) that, if (5.2) holds, there is a risk-neutral probability P supported on $\mathcal{K} \cup \{k_*, k^*\}$ such that $E_P(\max(0, S_i - K_{i,j})) = c_{i,j}$ for $1 \le i \le m$ and $1 \le j \le l_i$. Using Remark 3.1, we conclude that A2 holds. Furthermore, $c(\Psi) \le c(\psi) + ||\beta||_1 S_0$, since c is sub-additive and since the super-hedging cost of a call is at most S_0. Since, by the Cauchy-Schwartz inequality, $||\beta||_1 \le \sqrt{l}||\beta||$,

$$c(\Psi) \le c(\psi) + \sqrt{l}||\beta||S_0, \tag{E.1}$$

and so Assumption A1 holds as well. On the other hand, it follows from (5.2) that, since $c_{i,0}$ and c_{i,l_i+1} are upper-bounded by S_0, so is $c_{i,j}$. Hence (5.3). □

F Remainder of the proof of Theorem 5.1

Let $\mathcal{M}'$ be a market with sample space Ω'. We say that a market $\mathcal{M}$ is a sub-market of $\mathcal{M}'$ if it is obtained by restricting the basic securities of $\mathcal{M}'$ to a non-empty subset Ω of Ω'. To simplify the notation, if η is a derivative in $\mathcal{M}'$, the restriction of η to Ω is also denoted by η. To upper-bound the discretization errors in Theorems 5.1, 5.2, 5.3, and 5.4, we will use the following result, where $\mathcal{M}$ (resp. $\mathcal{M}'$) is a market with finite (resp. infinite) state-space, and prove (F.1) by extending a hedging scheme from $\mathcal{M}$ to $\mathcal{M}'$. The proofs of Theorems 5.1, 5.2, 5.3, and 5.4 thus allow us to construct hedging strategies in $\mathcal{M}'$ which are optimal, up to ϵ and the discretization error.

Proposition F.1. *Let ψ be a derivative in $\mathcal{M}'$, ϕ a vector of derivatives in $\mathcal{M}'$, and $\mathcal{M}$ a sub-market of $\mathcal{M}'$. Assume that Assumptions A1, A2 and A3 hold for (ψ, ϕ) in the market $\mathcal{M}$, and*

$$c(\psi - \beta^T\phi; \mathcal{M}') \le c(\psi - \beta^T\phi; \mathcal{M}) + \alpha, \tag{F.1}$$

for $||\beta||_\infty \le R_0$ and some constant α. Then

$$\pi_{\sup}(\psi, \phi; \mathcal{M}) \le \pi_{\sup}(\psi, \phi; \mathcal{M}') \le \pi_{\sup}(\psi, \phi; \mathcal{M}) + \alpha. \tag{F.2}$$

Proof. For a derivative η in $\mathcal{M}'$ and a constant γ, if $\eta \le_g \gamma$ in $\mathcal{M}'$, then $\eta \le_g \gamma$ in $\mathcal{M}$. Hence, $V_0(\psi, \phi; \mathcal{M}') \subseteq V_0(\psi, \phi; \mathcal{M})$. By (2.3), the first inequality in (F.2) follows. We now show the second one. By Theorem 3.2, for $\epsilon > 0$, there is $(\beta^*, \gamma^*) \in V(\psi, \phi; \mathcal{M})$ with $||\beta^*||_\infty \leq R_0$ such that

$$\beta^{*T}\pi + \gamma^* \leq \pi_{\sup}(\psi, \phi; \mathcal{M}) + \epsilon.$$

By (2.4), $c(\psi - \beta^{*T}\phi; \mathcal{M}) \leq \gamma^*$, and so $c(\psi - \beta^{*T}\phi; \mathcal{M}') \leq \gamma^* + \alpha$. Thus, $(\beta^*, \gamma^* + \alpha) \in V(\psi, \phi; \mathcal{M}')$. Since

$$\beta^{*T}\pi + \gamma^* + \alpha \leq \pi_{\sup}(\psi, \phi; \mathcal{M}) + \alpha + \epsilon,$$

we infer, by applying (2.5) in the market $\mathcal{M}'$, that $\pi_{\sup}(\psi, \phi; \mathcal{M}') \leq \pi_{\sup}(\psi, \phi; \mathcal{M}) + \alpha + \epsilon$. Letting ϵ go to 0 concludes the proof. □

For any finite subset F of $\mathbb{R}$, let $F^* = F \cup [\max(F), \infty)$. Denote by $g_{|F'}$ the restriction of g to a subset F' of $\mathbb{R}$. By convention, $\max(\emptyset) = -\infty$.

Proposition F.2. *Let F be a finite subset of $\mathbb{R}^+$ containing $\{0\}$, and g a function on $\mathbb{R}^+$ which is convex on any closed interval delimited by consecutive points in F. Then $\overline{g}(x) = \overline{g_{|F^*}}(x)$, for $x \geq 0$.*

Proof. The proposition follows by observing that a concave function upper-bounds g on $\mathbb{R}^+$ if and only if it upper-bounds g on F^*. □

Proposition F.3. *Let $w \geq 0$ and (g_j), $j \in J$, a finite family of continuous functions on an interval $[y, z]$ such that, for $j \in J$, g_j'' exists on (y, z) and is lower-bounded by $-w$. Let $f = max_{j\in J}(g_j)$. If $f(x) \leq g(x)$ for $x \in \{y, z\}$, where g is a linear function, then $f(x) \leq g(x) + w(z - y)^2/8$ for $x \in [y, z]$.*

Proof. The proposition follows by observing that the function $x \mapsto g_j(x) + w(x - y)(x - z)/2$ is convex and is upper-bounded by g on $\{y, z\}$. Thus it is thus upper-bounded by g on $[y, z]$, and so g_j is upper-bounded by $g + w(z - y)^2/8$ on $[y, z]$. □

Lemma F.1. *Let F be a finite subset of $\mathbb{R}^+$ with $|F| > 1$, and g a function on $[\min(F), \infty)$ which is convex on $[\max(F), \infty)$ and such that $g(u) = \lambda u + \lambda'$ for $u \geq x'$, for some constants $x' \geq \max(F)$, λ and λ'. Then, for $x \geq \min(F)$,*

$$\overline{g_{|F}}(x) = \max_{u,u'\in F, u\leq x\leq u', u\neq u'} \frac{(u' - x)g(u) + (x - u)g(u')}{u' - u}, \tag{F.3}$$

with the convention that $\overline{g_{|F}}(x) = -\infty$ if $x > \max(F)$. Furthermore,

$$\overline{g_{|F^*}}(x) = \max(\overline{g_{|F}}(x), \max_{u\in F, u\leq x} g(u) + \lambda(x - u)), \tag{F.4}$$

and there is a real number ξ^* such that, for $y \in F^*$,

$$g(y) \leq \overline{g_{|F^*}}(x) + \xi^*(y - x). \tag{F.5}$$

If $\min(F) \leq x \leq x'$ and $x' = \max(F)$, then

$$0 \leq \overline{g_{|F^*}}(x) - \overline{g_{|F}}(x) \leq \frac{x}{x'} \max_{u\in F, u<x}(|\lambda u + \lambda' - g(u)|). \tag{F.6}$$

Proof. By Proposition 4.3, we can assume the probabilities $\mathcal{Q}$ in (4.1) to be supported on two distinct points u and u' when $d = 1$, with $u \leq x \leq u'$, which gives (F.3). On the other hand, since g is convex on $[\max(F), \infty)$, for $\max(F) < y < u'$,

$$\frac{g(y) - g(\max(F))}{y - \max(F)} \leq \frac{g(y) - g(u')}{y - u'}.$$

Letting u' go to infinity shows that, for $y \geq \max(F)$,

$$g(y) \leq g(\max(F)) + \lambda(y - \max(F)). \tag{F.7}$$

Similarly, for $u, u' \in F^*$, $u \leq x \leq u', u \neq u'$,

$$\overline{g_{|F^*}(x)} \geq \frac{(u' - x)g(u) + (x - u)g(u')}{u' - u}.$$

Letting u' go to infinity implies that $\overline{g_{|F^*}(x)} \geq g(u) + \lambda(x - u)$. Hence $A(x) \leq \overline{g_{|F^*}}(x)$, where $A(x)$ is the RHS of (F.4).

Assume now $x \in [\min(F), \max(F)]$. By Proposition 4.3, there is a real number ζ such that, for $y \in F$,

$$g(y) \leq \overline{g_{|F}}(x) + \zeta(y - x).$$

Let $\xi^* = \max(\lambda, \zeta)$ and $y \in F$. Since $\overline{g_{|F}}(x) \leq A(x)$, if $y \geq x$ or if $\xi^* = \zeta$,

$$g(y) \leq A(x) + \xi^*(y - x). \tag{F.8}$$

Also, it follows from the definition of $A(x)$ that (F.8) holds if $y < x$ and $\xi^* = \lambda$. Thus, (F.8) holds for any $y \in F$.

Similarly, if $x \geq \max(F)$, let $\xi^* = \lambda$. It follows from the definition of $A(x)$ that (F.8) holds for $y \in F$. We conclude that, for $x \geq \min(F)$, there is $\xi^* \geq \lambda$ such that (F.8) holds for $y \in F$. In particular,

$$g(\max(F)) \leq A(x) + \xi^*(\max(F) - x),$$

which, by (F.7), implies that (F.8) holds for $y \in F^*$. Hence $\overline{g_{|F^*}}(x) \leq A(x)$. This implies that (F.4) and (F.5) hold.

We now prove (F.6). By (F.3) and (F.4),

$$\overline{g_{|F^*}}(x) = \max(\overline{g_{|F}}(x), \max_{u \in F, u < x} g(u) + \lambda(x - u)).$$

On the other hand, for $u \in F$ with $u < x \leq x'$,

$$\begin{aligned} g(u) + \lambda(x - u) &= \frac{(x' - x)g(u) + (x - u)g(x')}{x' - u} + \frac{x - u}{x' - u}(g(u) - \lambda u - \lambda') \\ &\leq \overline{g_{|F}}(x) + \frac{x}{x'} \max_{u' \in F, u' < x}(|\lambda u + \lambda' - g(u)|). \end{aligned}$$

This concludes the proof. □

We now prove the remainder of the theorem. Set $K_{2,0} = 0$, $M_0 = \max(\mathcal{K})$, and $M_i = \max(F_i)$ for $1 \leq i \leq 2$. Assume that $\beta = (b_1, b_2)$ is such that $||\beta||_\infty \leq R_0$. By (5.9), $R_0 = O(1)$. Since $|\max(x - K_{i,j}, 0) - x| \leq M_0$ for $x \geq 0$, $1 \leq i \leq 2$ and $1 \leq j \leq l_i$,

$$|b_i^T f_i(x) - (b_i^T \mathbf{1}_{l_i})x| \leq R_0 M_0 l,$$

and so, for $x_1 \geq 0$, $x_2 \geq 0$,

$$\Psi^*(x_1, x_2) = \max(0, x_2 - x_1) - b_2^T \mathbf{1}_{l_2} x_2 - b_1^T \mathbf{1}_{l_1} x_1 + O(1). \tag{F.9}$$

It follows from (5.4) that, for $x_1 \in F_1$,

$$\Psi_1^*(x_1; F_2) = \overline{\Psi^*(x_1, .)_{|F_2}}(x_1). \tag{F.10}$$

Similarly, for $x_1 \geq 0$, $\Psi_1^*(x_1) = \overline{\Psi^*(x_1, .)}(x_1)$. Since $g = \Psi^*(x_1, .)$ is convex on each interval $[K_{2,j}, K_{2,j+1}]$, $0 \leq j \leq l_2 - 1$, it follows by applying Proposition F.2 with $F = F_2$ that

$$\Psi_1^*(x_1) = \overline{\Psi^*(x_1, .)_{|F_2^*}}(x_1). \tag{F.11}$$

Since g is convex on $[M_2, \infty)$, it satisfies the conditions of Lemma F.1 with $F = F_2$, $x' = \max(M_2, x_1)$. Furthermore, by (F.9), $\lambda = O(1)$, $\lambda' = O(1 + x_1)$ and $\Psi^*(x_1, u) = O(1 + x_1)$ for $u \leq M_0$. Hence, by (F.6), (F.10) and (F.11), for $x_1 \in F_1$ and some constant $\kappa = O(1)$,

$$0 \leq \Psi_1^*(x_1) - \Psi_1^*(x_1; F_2) \leq \kappa x_1 \epsilon'. \tag{F.12}$$

Also, by (F.5) and (F.11), for $x_1 \geq 0$, there is a real number $\xi_{x_1}^*$ such that, for $x_2 \in F_2^*$,

$$\Psi^*(x_1, x_2) \leq \Psi_1^*(x_1) + \xi_{x_1}^*(x_2 - x_1). \tag{F.13}$$

Since g is convex on any interval delimited by consecutive points of F_2, (F.13) holds for $x_2 \geq 0$. On the other hand, the function g satisfies the conditions of Lemma F.1 with $F = \{0\} \cup \mathcal{K}$, $x' = \max(M_0, x_1)$, and $\lambda = 1 - b_2^T \mathbf{1}_{l_2}$. Let K'_j, $0 \leq j \leq l'$, denote the elements of F in increasing order, with $K'_{l'+1} = \infty$. (F.3), (F.4) and (F.11) imply that, for $x_1 \in [K'_j, K'_{j+1})$, $0 \leq j \leq l'$,

$$\Psi_1^*(x_1) = \max\Big(\max_{0 \leq j' \leq j < j'' \leq l_2} f_{j',j''}(x_1), \max_{0 \leq j' \leq j} \Psi^*(x_1, K'_{j'}) + (1 - b_2^T \mathbf{1}_{l_2})(x_1 - K'_{j'})\Big), \tag{F.14}$$

where

$$f_{j',j''}(x_1) = \frac{(K'_{j''} - x_1)\Psi^*(x_1, K'_{j'}) + (x_1 - K'_{j'})\Psi^*(x_1, K'_{j''})}{K'_{j''} - K'_{j'}}. \tag{F.15}$$

It follows that $\Psi_1^*(x_1) = O(1)$ for $x_1 \in [0, M_0]$, since, by (F.9), $\Psi^*(x_1, K'_{j'}) = O(1)$ for $0 \leq j' \leq l'$. Furthermore, since $\Psi^*(x_1, K'_{j'})$ and $\Psi^*(x_1, K'_{j''})$ are linear functions of x_1 with $O(1)$ slope on (K'_j, K'_{j+1}),

$$f''_{j',j''}(x_1) = O(1), \tag{F.16}$$

for $x_1 \in (K'_j, K'_{j+1})$. Moreover, by (F.9), $\Psi^*(., K'_{j'})$ is linear with slope $-b_1^T \mathbf{1}_{l_1}$ on $[M_0, \infty)$ and so, by (F.14), the function Ψ_1^* is also linear on $[M_0, \infty)$ and satisfies the conditions of Lemma F.1 with $F = F_1$, $x' = M_1$ and $\lambda = 1 - b_1^T \mathbf{1}_{l_1} - b_2^T \mathbf{1}_{l_2}$. Note that λ and $\lambda' = \Psi_1^*(M_0) - \lambda M_0$ are $O(1)$. Hence, by (F.6),

$$\overline{\Psi_{1|F_1^*}^*}(S_0) = \overline{\Psi_{1|F_1}^*}(S_0) + O(\epsilon'), \tag{F.17}$$

and there is a real number ξ_0^* such that, for $x_1 \in F_1^*$,

$$\Psi_1^*(x_1) \leq \overline{\Psi_{1|F_1^*}^*}(S_0) + \xi_0^*(x_1 - S_0). \tag{F.18}$$

Consider now a probability $\mathcal{Q} \in \mathbb{Q}(S_0, F_1)$. By (F.12),

$$E_{\mathcal{Q}}(\Psi_1^*) = E_{\mathcal{Q}}(\Psi_1^*(.; F_2)) + O(\epsilon').$$

Hence, by (4.1) and (5.5) and taking the supremum over $\mathcal{Q} \in \mathbb{Q}(S_0, F_1)$, it follows that

$$\overline{\Psi_{1|F_1}^*}(S_0) = c(\Psi; \mathcal{M}(F_1, F_2)) + O(\epsilon'),$$

and so, by (F.17),

$$\overline{\Psi_{1|F_1^*}^*}(S_0) = c(\Psi; \mathcal{M}(F_1, F_2)) + O(\epsilon'). \tag{F.19}$$

Since (F.18) holds for $x_1 \in \{M_0, M_1\}$ and Ψ_1^* is linear on $[M_0, M_1]$, (F.18) holds for $x_1 \in [M_0, M_1]$, and so it holds for $x_1 \in F_1 \cup [M_0, \infty)$. Using (F.14), (F.16) and Proposition (F.3), we conclude that, for $x_1 \geq 0$,

$$\Psi_1^*(x_1) \leq \overline{\Psi_{1|F_1^*}^*}(S_0) + \xi_0^*(x_1 - S_0) + O(n^{-2}).$$

Using (F.13) and (F.19), we infer that, for $x_1 \geq 0$ and $x_2 \geq 0$,

$$\Psi^*(x_1, x_2) \leq c(\Psi; \mathcal{M}(F_1, F_2)) + \xi_0^*(x_1 - S_0) + \xi_{x_1}^*(x_2 - x_1) + O(n^{-2} + \epsilon').$$

Thus, $c(\Psi; \mathcal{M}_2(\mathbb{R}^+)) \leq c(\Psi; \mathcal{M}(F_1, F_2)) + O(n^{-2} + \epsilon')$ which, by Proposition F.1, implies (5.8). □

G Remainder of the proof of Theorem 5.2

Since $\Psi^*(x_1, x_2)$ is a convex function of x_2 on any interval of positive real numbers disjoint with $\mathcal{K}$ and not containing x_1, it can be shown using arguments similar to those in the proof of Theorem 5.1 that, for $x_1 \in [K'_j, K'_{j+1})$, $0 \le j \le l'$,

$$\Psi_1^*(x_1) = \max(\Psi^*(x_1, x_1), \max_{0 \le j' \le j < j'' \le l_2} f_{j',j''}(x_1), \max_{0 \le j' \le j} \Psi^*(x_1, K'_{j'}) - (1 + b_2^T \mathbf{1}_{l_2})(x_1 - K'_{j'})),$$

where $f_{j',j''}$ is given by (F.15). Hence, on each interval I delimited by two consecutive points in $\{0\} \cup \mathcal{K}$, the function Ψ_1^* is convex, since $f_{j',j''}$ is convex on I, and Ψ_1^* is linear on $[\max \mathcal{K}, \infty)$. Thus, by Proposition F.2 and Lemma F.1, we can calculate $c(\Psi; \mathcal{M}_2(\mathbb{R}^+))$ in terms of the slope of Ψ_1^* on $[\max(\mathcal{K}), \infty)$ and the values of Ψ_1^* on $\{0\} \cup \mathcal{K}$. We can then show (5.10) by arguments similar to those in the proof of Theorem 5.1. □

H Remainder of the proof of Theorem 5.3

The following lemma shows that, under smoothness conditions on H, $\pi_{\sup}([L, M])$ is well approximated by $\pi_{\sup}(F_0)$.

Lemma H.1. *Assume that H is continuous on $[L, M]^2$ and there are functions a and b on (L, M), and $w \ge 0$, such that the following conditions hold. The function a is non-positive. For $x \in [L, M]$, the function $y \mapsto H(x, y)$ has a second derivative lower-bounded by $b(y)$ for $y \in (L, M)$. If y and z are two elements of F_0 and $y < x < z$, the second derivative of $(z - x)H(x, y) + (x - y)H(x, z)$ with respect to x is lower bounded by $(z - y)a(x)$. Furthermore, $(a(x) + b(x))(z - y)^2 \ge -w$, if y and z are consecutive elements of F_0. Then*

$$\pi_{\sup}(F_0) \le \pi_{\sup}([L, M]) \le \pi_{\sup}(F_0) + m\frac{w}{8}.$$

Proof. Let $i \in [1, m - 1]$ and $x \in [y, z]$, where y and z are consecutive elements of F_0. By Proposition 4.3, we can assume the probability $\mathcal{Q}$ in (5.17) to be supported by two points y' and z', and so

$$c_i(x) = \max_{y', z' \in F_0, y' \le y \le z \le z'} g_{i,y',z'}(x), \tag{H.1}$$

where

$$g_{i,y',z'}(x) = \frac{z' - x}{z' - y'}(c_{i+1}(y') + H_{i+1}(x, y')) + \frac{x - y'}{z' - y'}(c_{i+1}(z') + H_{i+1}(x, z')). \tag{H.2}$$

Note that (H.1) also holds if $i = m$ and $g_{m,y',z'}$ is the null function. Thus, for $1 \le i \le m$ and $x_0 \in [L, M]$,

$$c_i(x) + H_i(x_0, x) = \max_{y', z' \in F_0, y' \le y \le z \le z'} f_{i,x_0,y',z'}(x),$$

where

$$f_{i,x_0,y',z'}(x) = g_{i,y',z'}(x) + H_i(x_0, x).$$

On the other hand, by (4.8) and (5.17), there is a real number ξ^*_{i,x_0} be such that, for $x \in F_0$,

$$c_i(x) + H_i(x_0, x) \le c_{i-1}(x_0) + \xi^*_{i,x_0}(x - x_0).$$

Furthermore, by (H.2), $g_{i,y',z'}$ is continuous on $[y, z]$ and $g''_{i,y',z'}(x) \ge a(x)$ for $x \in (y, z)$. Hence $f_{i,x_0,y',z'}$ is continuous on $[y, z]$ and $f''_{i,x_0,y',z'}(x) \ge a(x) + b(x)$ for $x \in (y, z)$. Since $f''_{i,x_0,y',z'} \ge -w/(z - y)^2$ on (y, z), by Proposition F.3, we conclude that, for $x \in [y, z]$,

$$c_i(x) + H_i(x_0, x) \le c_{i-1}(x_0) + \xi^*_{i,x_0}(x - x_0) + \frac{w}{8}. \tag{H.3}$$

Thus, in the market $\mathcal{M}_{[L,M]}$,

$$c_i(S_i) + H_i(S_{i-1}, S_i) \le c_{i-1}(S_{i-1}) + \xi_i(S_i - S_{i-1}) + \frac{w}{8}, \tag{H.4}$$

where ξ_i is the financial derivative $\xi_i = \xi^*_{i,S_{i-1}}$. By (5.11) and summing (H.4) for $i \in [1, m]$, we infer that

$$\Psi \le c_0(S_0) + \sum_{i=1}^{m} \xi_i(S_i - S_{i-1}) + m\frac{w}{8}.$$

On the other hand, by applying (5.13) with $i = 0$, it follows that

$$c_0(S_0) = c(\Psi; \mathcal{M}_m(F_0)). \tag{H.5}$$

Thus,

$$c(\Psi; \mathcal{M}_m([L, M])) \le c(\Psi; \mathcal{M}_m(F_0)) + m\frac{w}{8}.$$

We conclude the proof using Proposition F.1. □

We now prove (5.15) and (5.16). Let $H(x, y) = T^{-1}\ln^2(y/x)$. Then

$$\frac{\partial^2 H}{\partial y^2}(x, y) = 2T^{-1}\frac{1 + \ln(x/y)}{y^2}, \tag{H.6}$$

and the second derivative of $((z - x)H(x, y) + (x - y)H(x, z))/(z - y)$ with respect to x is

$$\frac{2}{x^2}T^{-1}(1 - \frac{(z + x)\ln(x/y) + (x + y)\ln(z/x)}{(z - y)}).$$

By noting that $x + y \le x + z \le 2z$ if $y < x < z$, and that the function $z\ln(z/y)/(z - y)$ is increasing with respect to z and decreasing with respect to y for $0 < y < z$, a standard calculation shows that the conditions of Lemma H.1 hold with

$$a(x) = -\frac{4}{x^2}T^{-1}\frac{M\ln(M/L)}{M - L},$$

$$b(y) = -\frac{2}{y^2}T^{-1}\ln(M/L), \text{ and}$$

$$w = 6T^{-1}\frac{M\ln(M/L)}{M - L}((M/L)^{1/n} - 1)^2.$$

Hence (5.15). Similarly, (5.16) follows by noting that, when $H(x, y) = -T^{-1}\ln^2(y/x)$, the conditions of Lemma H.1 hold with

$$a(x) = -\frac{2}{x^2}T^{-1}, \tag{H.7}$$

$$b(y) = -\frac{2}{y^2}T^{-1}(1 + \ln(M/L)), \text{ and} \tag{H.8}$$

$$w = 2T^{-1}(2 + \ln(M/L))((M/L)^{1/n} - 1)^2.$$

I Calculation of $\pi_{\inf}((0, M])$ for Variance Swaps

We first show the following propositions.

Proposition I.1. *Assume that $\{0, M\}$ is acceptable. Then, for any $L \in [0, \delta_0(0, M)/2]$, $\{L, M\}$ is acceptable, and $\delta_0(L, M) \geq \delta_0(0, M)/2$.*

Proof. Using the same notation as in Definition 5.1, with $k_* = 0$ and $k^* = M$, for $1 \leq i \leq m$ and $1 \leq j \leq l_i$, by letting $i' = i'' = i$, $j' = 0$ and $j'' = l_i + 1$, (5.2) becomes

$$\max(0, S_0 - K_{i,j}) < c_{i,j} < wS_0,$$

and so

$$\delta_0(0, M) \leq S_0 - c_{i,j} \leq K_{i,j}. \tag{I.1}$$

On the other hand, for $0 < k_* < \delta_0(0, M)$ and $k^* = M$, it follows by inspection that k_* appears in (5.2) only when $j' = 0$, and an easy calculation shows that, when $j' = 0$,

$$\begin{aligned} \frac{\partial(wc_{i',j'} + (1-w)c_{i'',j''})}{\partial k_*} &= \frac{w}{K_{i'',j''} - k_*}(S_0 - k_* - c_{i'',j''}) - w \\ &\geq -1. \end{aligned}$$

The second equation follows from (I.1). Thus, for $L \in [0, \delta_0(0, M)/2]$, if the value of k_* changes from 0 to L, $wc_{i',j'} + (1-w)c_{i'',j''}$ diminishes by at most L. Hence, (5.2) stills holds if $k_* = L$ and $k^* = M$, and $\delta_0(L, M) \geq \delta_0(0, M) - L$. This concludes the proof. □

Proposition I.2. *Let F be a subset of $\mathbb{R}^+$, x_0 an element of F, and F' a finite subset of F that contains $F \cap [x_0, \infty)$. Let g be a function on F and λ a real number such that, for any $z \in F$,*

$$g(z) \leq g(x_0) + \lambda(z - x_0). \tag{I.2}$$

Then, for any element x of $F \cap (x_0, \infty)$, $\overline{g_{|F'}}(x) = \overline{g_{|F\cap[x_0,\infty)}}(x)$, and there is a real number ξ^* such that, for $z \in F$,

$$g(z) \leq \overline{g_{|F\cap[x_0,\infty)}}(x) + \xi^*(z - x). \tag{I.3}$$

Proof. Proof.Let x be any element of $F \cap (x_0, \infty)$. By Proposition 4.3, there is a real number ξ^* such that (I.3) holds for $z \in F \cap [x_0, \infty)$. In particular,

$$g(x_0) \leq \overline{g_{|F\cap[x_0,\infty)}}(x) + \xi^*(x_0 - x). \tag{I.4}$$

On the other hand, (I.2) implies that $\overline{g_{|F\cap[x_0,\infty)}}(x) \leq g(x_0) + \lambda(x - x_0)$, and so, by (I.4), $\xi^* \leq \lambda$. Thus, by (I.2), $g(z) \leq g(x_0) + \xi^*(z - x_0)$ for $z \in F \cap (-\infty, x_0]$ which, together with (I.4), implies that (I.3) holds for $z \in F \cap (-\infty, x_0]$. Hence, (I.3) holds for any $z \in F$, and so $\overline{g_{|F'}}(x) \leq \overline{g_{|F\cap[x_0,\infty)}}(x)$. But, since any concave function that upper-bounds g on F' also upper-bounds g on $F \cap [x_0, \infty)$, we conclude that $\overline{g_{|F'}}(x) = \overline{g_{|F\cap[x_0,\infty)}}(x)$.qed We now show how to calculate $\pi_{\inf}((0, M])$. Given $n > 0$, set

$$F' = \mathcal{K} \cup \{\frac{j}{n}M, j \in \{1, \ldots, n\}\}.$$

The constants behind the O notation in (I.5) and in the proof of Theorem I.1 depend on l, S_0, the calls strikes, maturities and prices, M, m, and T, but do not depend on n. The intuition behind the proof of Theorem I.1 is that the realized variance of any path containing very low stock prices is high. Thus $\pi_{\inf}((0, M])$ can be calculated by excluding such paths and applying the same techniques as in the proof of Theorem 5.3.

Theorem I.1. *Assume that $\{0, M\}$ is acceptable. Let $B = mT^{-1}\ln^2(4M/\delta_0(0,M))$. Then, for $n \geq 4M/\delta_0(0,M)$, $\pi_{\inf}(F')$ can be calculated in $O(N(l^3+n^2m+lmn))$ total time with precision ϵ via the convex program* (3.3)*, where*

$$N = O(l\ln(\frac{l(1+S_0)(1+B)}{\epsilon\delta_0(0,M)})),$$

and V' has a separation oracle that runs in $O(n^2m+lmn)$ time. Furthermore,

$$\pi_{\inf}((0,M]) - \pi_{\inf}(F') = O(n^{-2}). \tag{I.5}$$

Proof. Let $H(x,y) = -T^{-1}\ln^2(y/x)$. Since $\psi \leq 0$, it follows from (E.1) that Assumption A1 holds in the market $M_m(F')$, with

$$q = B + S_0\sqrt{l} \text{ and } \delta = \delta_0(0,M)/4.$$

Since the spacing between two consecutive elements of F' is at most $\delta_0(0,M)/4$, there is a real number $k_* \in F' \cap [\delta_0(0,M)/4, \delta_0(0,M)/2]$. By Proposition I.1, $\{k_*, M\}$ is acceptable and $\delta_0(k_*, M) \geq \delta_0(0,M)/2$. On the other hand, as $0 \leq |\psi| \leq B$ in the market $\mathcal{M}_m(F' \cap [\delta_0(0,M)/4, M])$, by Proposition 5.1, Assumption A2 hold in this market. Since any risk-neutral probability in a sub-market of $M_m(F')$ is also a risk-neutral probability in $M_m(F')$, we conclude that A2 holds in the market $M_m(F')$ as well. Finally, an argument similar to that in the proof of Theorem 5.3 shows that A3 holds in the market $M_m(F')$, with $\mathcal{T} = O(n^2m+lmn)$. This shows the first part of the theorem.

We now show (I.5). Let $L_i = \min(\mathcal{K})e^{-(m-i)\sqrt{lR_0M}}$ for $1 \leq i \leq m$, and let L_0 be the largest element of $F' \cap [0, \min(\mathcal{K})e^{-m\sqrt{lR_0M}}]$, which is non-empty if n is sufficiently large. Assume that $\beta = (b_1, \dots, b_m)$ is such that $||\beta||_\infty \leq R_0$. Define $\tilde{c}_i(x)$ for $x \in (0,M]$ by setting $\tilde{c}_m = 0$ and, for $0 \leq i \leq m-1$, $\tilde{c}_i(x) = 0$ if $x \in (0, L_0]$, and, for $x \in (L_0, M]$,

$$\tilde{c}_i(x) = \sup_{\mathcal{Q}\in\mathbb{Q}(x,F')} E_{\mathcal{Q}}(\tilde{c}_{i+1} + H_{i+1}(x,.)). \tag{I.6}$$

(I.6) is obtained from (5.17) by replacing c_i, c_{i+1}, and F_0 with $\tilde{c}_i$, $\tilde{c}_{i+1}$, and F' respectively. For $y > 0$ and $1 \leq i \leq m$, every component of $f_i(y)$ is at most y, and so $b_i^T f_i(y) \geq -l_iR_0y$. Hence $H_i(x,y) \leq l_iR_0y$, and so it follows by backward induction that, for $x \in (0,M]$ and $1 \leq i \leq m$, $\tilde{c}_i(x) \leq (\sum_{j=i+1}^m l_j)R_0x$. Thus, for $x \in (0,M]$,

$$\tilde{c}_i(x) - b_i^T f_i(x) \leq lR_0x. \tag{I.7}$$

We show by backward induction that, for $1 \leq i \leq m$, $x \in (0, L_{i-1}]$ and $z \in (0,M]$,

$$\tilde{c}_i(z) + H_i(x,z) \leq 0. \tag{I.8}$$

Assume first that $i = m$, and let $x \in (0, L_{m-1}]$. Then (I.8) holds if $z \leq L_m$ since f_m is null on $(0, L_m]$. On the other hand, if $z \in (L_m, M]$, then $H(x,z) \leq -lR_0M$, and so, by (I.7), (I.8) holds again in this case. Thus the induction hypothesis holds for $i = m$. Assume it holds for $i+1$. By (I.6), it follows that $\tilde{c}_i(x) \leq 0$ for $x \in (0, L_i]$. Assume now that $x \in (0, L_{i-1}]$. Since $H(x,z) \leq -lR_0M$ for $z \in [L_i, M]$, (I.7) implies that (I.8) holds for $z \in [L_i, M]$. On the other hand, (I.8) also holds for $z \in (0, L_i]$ since f_i is null on $(0, L_i]$. Thus the induction hypothesis holds for i.

Let $x \in (0, L_0] \cap F'$. It follows from (I.8) that the RHS of (I.6) is at most 0. But, by considering the probability $\mathcal{Q}$ supported on $\{x\}$, it can be seen that the RHS of (I.6) is non-negative, and is thus null. Hence (I.6) holds for x, and so it holds for any $x \in F'$. Thus, an argument similar to the proof of (H.5) shows that $\tilde{c}_0(S_0) = c(\Psi; \mathcal{M}_m(F'))$.

Set

$$L' = \frac{L_0}{e + \ln(M/L_0) + lR_0TL_0},$$

and let L be the largest element of $F' \cap [0, L']$, which is non-empty if n is sufficiently large. Fix $x \in (L_0, M]$ and $i \in [0, m-1]$, and let $g(z) = \tilde{c}_{i+1}(z) + H_{i+1}(x, z)$. We show that, for $z \in (0, M]$,

$$g(z) \leq g(L) + (z - L)g'(L). \tag{I.9}$$

Note that $g(z) = H(x, z)$, for $z \in (0, L_0]$, and so (I.9) holds in this case because, by (H.6), $H(x, .)$ is concave on $(0, L_0]$. On the other hand, for $L_0 < z \leq M$,

$$g(L) + (z - L)g'(L) = T^{-1}(2(\frac{z}{L} - 1) - \ln(\frac{x}{L}))\ln(\frac{x}{L}).$$

But

$$\begin{aligned} 2(\frac{z}{L} - 1) - \ln(\frac{x}{L}) &= 2(\frac{z}{L} - 1) - \ln(\frac{z}{L}) - \ln(\frac{x}{z}) \\ &\geq \frac{z}{L} - 1 - \ln(\frac{M}{L_0}) \\ &\geq lR_0Tz. \end{aligned}$$

The second equation follows from the inequality $\ln(y) \leq y - 1$, and the third equation holds for $z \geq L_0$ by standard calculations. Thus

$$g(L) + (z - L)g'(L) \geq lR_0z,$$

which, together with (I.7), implies (I.9). Thus (I.9) holds for $z \in (0, M]$.

Let $\tilde{F} = (0, L] \cup F'$. By (I.9), and since $\tilde{F} \cap [L, M] \subseteq F' \subseteq \tilde{F}$, we can apply Proposition I.2 with $F = \tilde{F}$, $x_0 = L$ and $\lambda = g'(L)$. Hence, by (I.6),

$$\tilde{c}_i(x) = \sup_{\mathcal{Q} \in \mathbb{Q}(x, F' \cap [L, M])} E_{\mathcal{Q}}(\tilde{c}_{i+1} + H_{i+1}(x, .)),$$

and there is a real number $\xi^*_{i,x}$ be such that, for $z \in \tilde{F}$,

$$\tilde{c}_{i+1}(z) + H_{i+1}(x, z) \leq \tilde{c}_i(x) + \xi^*_{i,x}(z - x).$$

Since $\tilde{c}_{i+1}$ is null on $[L, L_0]$, arguments similar to those used in the proof of Theorem 5.3 (see (H.3), (H.7) and (H.8)) show that, for $z \in (0, M]$,

$$\tilde{c}_{i+1}(z) + H_{i+1}(x, z) \leq \tilde{c}_i(x) + \xi^*_{i,x}(z - x) + w, \tag{I.10}$$

where $w \geq 0$ and $w = O(n^{-2})$.

On the other hand, by (I.8), (I.10) holds for $x \in (0, L_0]$ if we set $\xi^*_{i,x} = 0$, and so it holds for $x \in (0, M]$. We infer that, in the market $\mathcal{M}_m((0, M])$,

$$\Psi \leq \tilde{c}_0(S_0) + \sum_{i=1}^{m} \xi_i(S_i - S_{i-1}) + mw,$$

where ξ_i is the financial derivative $\xi_i = \xi^*_{i,S_{i-1}}$. Hence,

$$c(\Psi; \mathcal{M}_m(0, M]) \leq c(\Psi; \mathcal{M}_m(F')) + mw.$$

We conclude the proof using Proposition F.1. □

In practice, $\pi_{\inf}(F')$ can be calculated for any integer n such that $\{M/n, M\}$ is acceptable. Numerical calculations of $\pi_{\inf}(F')$, using the same setting as in Table 5, are given in Table 14. The prices in Table 14 are very close to $\pi_{\inf}(F_0)$ in Table 5.

Table 14: Optimal sub-replicating prices of a variance swap maturing in one month in $\mathcal{M}_m((0, M])$, with $M = 200$ and $m = 20$. The discretization error is estimated using $n = 3200$.

n	$\sqrt{\pi_{\inf}(F')}$	Computing time	Error
50	19.85%	0.4	3.7×10^{-3}
100	19.19%	0.7	1.1×10^{-3}
200	19.01%	1.6	3.7×10^{-4}
400	18.92%	5.1	3.4×10^{-5}
800	18.91%	20	1.1×10^{-5}

J Remainder of the proof of Theorem 5.4

We say that a real-valued function g defined on a set of real numbers W is κ-Lipschitz on W if $|g(x) - g(y)| \leq \kappa|x - y|$ for $x, y \in W$.

Lemma J.1. *Let g be a real-valued bivariate function defined on $[L, M]^2$ such that, for some constants κ_x and κ_y, for any $x, y \in [L, M]$, the functions $g(., y)$ and $g(x, .)$ are respectively κ_x-Lipschitz and κ_y-Lipschitz on $[L, M]$. Let W be a finite subset of $[L, M]$ that contains $\{L, M\}$ and w the maximum distance between consecutive elements of W. For $x \in [L, M]$, set*

$$h(x) = \sup_{\mathcal{Q} \in \mathbb{Q}(x,W)} E_{\mathcal{Q}}(g(x, .)). \tag{J.1}$$

Then h is $(\kappa_x+\kappa_y)$-Lipschitz on $[L, M]$. Furthermore, for any $x \in [L, M]$, there is a real number ξ^* such that, for $y \in [L, M]$,

$$g(x, y) \leq h(x) + \kappa_y w + \xi^*(y - x). \tag{J.2}$$

Proof. Fix $y, y' \in W$ with $y < y'$. For $x \in [y, y']$, let

$$F(x, y, y') = \frac{(y' - x)g(x, y) + (x - y)g(x, y')}{y' - y}.$$

Since, for $y < x < x' < y'$,

$$(y' - y)(F(x', y, y') - F(x, y, y')) = (x - y)(g(x', y') - g(x, y'))+ \\ (y' - x)(g(x', y) - g(x, y)) + (x' - x)(g(x', y') - g(x', y)),$$

it follows after some calculations that the function $x \mapsto F(x, y, y')$ is $(\kappa_x + \kappa_y)$-Lipschitz on $[y, y']$.

Fix now two consecutive elements y_0 and y_1 of W. By Proposition 4.3, the probability $\mathcal{Q}$ in (J.1) can be assumed to be supported on two points y and y' of W and so, for $x \in [y_0, y_1]$,

$$h(x) = \max_{y,y' \in W, y \leq y_0, y_1 \leq y'} F(x, y, y').$$

Since the maximum of a finite set of κ-Lipschitz functions on $[y_0, y_1]$ is κ-Lipschitz on $[y_0, y_1]$, it follows that h is $(\kappa_x + \kappa_y)$-Lipschitz on $[y_0, y_1]$. Thus, h is $(\kappa_x + \kappa_y)$-Lipschitz on any interval delimited by consecutive elements of W, and so it is $(\kappa_x + \kappa_y)$-Lipschitz on $[L, M]$.

By Proposition 4.3 and (4.9), there is a real number ξ^* such that $|\xi^*| \leq \kappa_y$ and, for $z \in W$,

$$g(x, z) \leq h(x) + \xi^*(z - x).$$

(J.2) follows since, for any $y \in [L, M]$, there is $z \in W$ within distance $w/2$ from y. □

For $\nu \geq 0$, $0 \leq i \leq m$, and $x, y \in [L, M]$, let $\zeta_\nu(x, y) = \sqrt{\nu^2 + \ln^2(y/x)}$. Define by backward induction the functions $\tilde{c}_i(\nu, x) = \tilde{c}_i(\nu, x, \beta)$, for $\nu \geq 0$, $x \in [L, M]$, and $0 \leq i \leq m$ by setting $\tilde{c}_m(\nu, x) = \min(\nu, \nu_{\max})$ and, for $0 \leq i \leq m-1$, $\nu \in \Lambda$,

$$\tilde{c}_i(\nu, x) = \sup_{\mathcal{Q} \in \mathbb{Q}(x, F_0)} E_{\mathcal{Q}}(g_{i+1,\nu}(x, .)), \tag{J.3}$$

where

$$g_{i,\nu}(x, y) = \tilde{c}_i(\zeta_\nu(x, y), y) - b_i^T f_i(y), \tag{J.4}$$

and, for $\nu \in \mathbb{R}^+ - \Lambda$,

$$\tilde{c}_i(\nu, x) = (1 - \lambda)\tilde{c}_i(\rho(\nu), x) + \lambda \tilde{c}_{i+1}(\rho(\nu + \frac{\nu_{\max}}{n'}), x), \tag{J.5}$$

where $\lambda = (\nu - \rho(\nu))/(\nu_{\max}/n')$. (J.3) can be thought as the limit of (5.25) as n' goes to infinity, while (J.5) says that $\tilde{c}_i(\nu, x)$ is calculated by linear interpolation for $\nu \in [0, \nu_{\max}] - \Lambda$, and is set to $\tilde{c}_i(\nu_{\max}, x)$ for $\nu \geq \nu_{\max}$. As the function $\nu \mapsto \sqrt{\nu^2 + a^2}$ is 1-Lipschitz for any constant a, and since the supremum and the weighted average of 1-Lipschitz functions is 1-Lipschitz, it can be shown by backward induction that $\tilde{c}_i(\nu, x)$ is 1-Lipschitz and non-decreasing with respect to ν and, for $\nu \in \Lambda$ and $x \in F_0$,

$$c_i(\nu, x) \leq \tilde{c}_i(\nu, x) \leq c_i(\nu, x) + \frac{(m - i)\nu_{\max}}{n'}. \tag{J.6}$$

We now prove the following.

Lemma J.2. *For $\nu \geq 0$, $x \in [L, M]$, $0 \leq i \leq m$, the function $g_{i,\nu}(x, .)$ is κ_i-Lipschitz on $[L, M]$, where*

$$\kappa_i = \frac{2(m - i) + 1}{L} + \Sigma_{j=i}^m ||b_j||_1.$$

Proof. We prove the lemma by backward induction. The induction hypothesis holds when $i = m$ since the derivative of $\zeta_\nu(x, y)$ with respect to y is at most $1/y$, and any call function is 1-Lipschitz. Assume it holds for $i+1$. Since $\tilde{c}_{i+1}(., y)$ is 1-Lipschitz and the derivative of $\zeta_\nu(x, y)$ with respect to x is at most $1/x$, for $y \in [L, M]$, the function $g_{i+1,\nu}(., y)$ is (L^{-1})-Lipschitz on $[L, M]$. Also, by the induction hypothesis, the function $g_{i+1,\nu}(x, .)$ is κ_{i+1}-Lipschitz on $[L, M]$, for $x \in [L, M]$. By Lemma J.1 and (J.3), it follows that $\tilde{c}_i(\nu, .)$ is $(L^{-1} + \kappa_{i+1})$-Lipschitz on $[L, M]$ if $\nu \in \Lambda$. By (J.5), the same conclusion holds for any $\nu \geq 0$. On the other hand, by (J.4),

$$|g_{i,\nu}(x, y') - g_{i,\nu}(x, y)| \leq |\tilde{c}_i(\zeta_\nu(x, y'), y') - \tilde{c}_i(\zeta_\nu(x, y), y')| + |\tilde{c}_i(\zeta_\nu(x, y), y') - \tilde{c}_i(\zeta_\nu(x, y), y)| + |b_i^T (f_i(y) - f_i(y'))|. \tag{J.7}$$

The first term in the RHS of (J.7) is at most $L^{-1}|y' - y|$ since $\tilde{c}_i(., y')$ is 1-Lipschitz and $\zeta_\nu(x, .)$ is (L^{-1})-Lipschitz. The second term is at most $(L^{-1} + \kappa_{i+1})|y' - y|$ since $\tilde{c}_i(\nu, .)$ is $(L^{-1} + \kappa_{i+1})$-Lipschitz. The last term is at most $||b_i||_1|y' - y|$ since any call function is 1-Lipschitz. Hence

$$|g_{i,\nu}(x, y') - g_{i,\nu}(x, y)| \leq \kappa_i |y' - y|,$$

and the induction hypothesis holds for i. □

Let $w = L((M/L)^{1/n} - 1)$. It follows from Lemmas J.1, J.2 and (J.3) that, for $x \in [L, M]$, $0 \leq i \leq m$, and $\nu \in \Lambda$, there is a real number $\xi^*_{i,\nu,x}$ such that, for $y \in [L, M]$,

$$g_{i+1,\nu}(x, y) \leq \tilde{c}_i(\nu, x) + \kappa_1 w + \xi^*_{i,\nu,x}(y - x).$$

Since $\tilde{c}_{i+1}(., y)$ is non-decreasing, it follows that, in the market $\mathcal{M}' = \mathcal{M}_m([L, M])$,

$$\tilde{c}_{i+1}(\nu_{i+1}, S_{i+1}) - b_{i+1}^T f_{i+1}(S_{i+1}) \leq \tilde{c}_i(\nu_i, S_i) + \kappa_1 w + \xi^*_{i,\nu_i,S_i}(S_{i+1} - S_i). \tag{J.8}$$

Define the financial derivative $\xi_{i+1} = \xi^*_{i,\nu_i,S_i}$. By summing up (J.8) for $i \in [0, m-1]$, we get

$$\Psi' \leq \tilde{c}_0(0, S_0) + m\kappa_1 w + \sum_{i=0}^{m-1} \xi_{i+1}(S_{i+1} - S_i).$$

Thus,

$$c(\Psi'; \mathcal{M}') \leq \tilde{c}_0(0, S_0) + m\kappa_1 w.$$

Assume now that $\beta = (b_1, \ldots, b_m)$ is such that $||\beta||_\infty \leq R_0$, and let $\kappa = 2L^{-1} + lR_0$, so that $\kappa_1 \leq m\kappa$. By (J.6) and, since $c_0(0, S_0) = c(\Psi'; \mathcal{M})$, where $\mathcal{M} = \mathcal{M}_m(F_0)$, it follows that

$$c(\Psi'; \mathcal{M}') \leq c(\Psi'; \mathcal{M}) + \frac{m\nu_{\max}}{n'} + m^2\kappa w.$$

Hence, by Proposition F.1,

$$|\pi_{\sup}(\psi', \phi; \mathcal{M}') - \pi_{\sup}(\psi', \phi; \mathcal{M})| \leq \frac{m\nu_{\max}}{n'} + m^2\kappa w.$$

On the other hand, by (5.19),

$$|\pi_{\sup}(\psi, \phi; \mathcal{M}')\sqrt{T} - \pi_{\sup}(\psi', \phi; \mathcal{M}')| \leq \frac{m\nu_{\max}}{n'}.$$

Thus,

$$|\pi_{\sup}(\psi, \phi; \mathcal{M}')\sqrt{T} - \pi_{\sup}(\psi', \phi; \mathcal{M})| \leq 2\frac{m\nu_{\max}}{n'} + m^2\kappa w.$$

Since, by (5.24), $R_0 = O(1)$, this implies (5.22). We can show (5.23) in a similar manner. □

K Robust pricing of Volatility swaps via linear interpolation

Even though $\tilde{c}_0(0, S_0) = \tilde{c}_0(0, S_0, \beta)$ does not appear to have a simple financial interpretation, we can use it rather than $c_0(0, S_0) = c(\psi' - \beta^T\phi; \mathcal{M})$ in order to calculate the super-replicating price of a volatility swap. This yields the following algorithm. By analogy to (2.4), let

$$\tilde{V}' = \{(\beta, \gamma) \in \mathbb{R}^l \times \mathbb{R} : \tilde{c}_0(0, S_0, \beta) \leq \gamma, \beta^T\pi + \gamma \leq q + 1\},$$

where q is given by (5.24), and

$$\tilde{\pi}_{\sup}(n, n') = \inf_{(\beta,\gamma)\in\tilde{V}'} \beta^T\pi + \gamma.$$

We use a cutting plane algorithm to minimize $\beta^T\pi+\gamma$ over $\tilde{V}'$. We construct a separation oracle for $\tilde{V}'$ by analogy to that of V' in Proposition 3.1, as described below. Using Proposition 4.3, we calculate by backward induction $\tilde{c}_i(\nu, x)$ for $\nu \in \Lambda$, $x \in F_0$ and $0 \leq i \leq m-1$, and probabilities $\tilde{\mathcal{Q}}(\nu, x, i)$ that maximise the RHS of (J.3). Given a call option η with maturity $k \leq m$ and payoff $\tilde{\eta}(S_k)$, define the function $\tilde{\eta}_i$ by backward induction on $\mathbb{R}^+ \times F_0$, $0 \leq i \leq k$, by setting $\tilde{\eta}_k(\nu, x) = \tilde{\eta}(x)$ and, for $0 \leq i \leq k-1$, $\nu \in \Lambda$, and $x \in F_0$,

$$\eta^*_i(\nu, x) = E_{\tilde{\mathcal{Q}}(\nu,x,i)}(\eta^*_{i+1}(\sqrt{\nu^2 + \ln^2(z/x)}, z)),$$

and, for $\nu \in \mathbb{R}^+ - \Lambda$,

$$\eta^*_i(\nu, x) = (1-\lambda)\eta^*_i(\rho(\nu), x) + \lambda\eta^*_i(\rho(\nu + \nu_{\max}/n'), x),$$

where $\lambda = (\nu - \rho(\nu))/(\nu_{\max}/n')$. We construct a separation oracle for $\tilde{V}'$ by replacing $c(\psi - \beta^T\phi)$ with $\tilde{c}_0(0, S_0)$ and $E_P(\phi)$ with the vector $(\eta^*_0(0, S_0))$, where η ranges over the components of ϕ, in the algorithm described in Proposition 3.1.

Table 15: Computing times for robust prices of a variance swap with $n = 400$ in terms of the number of periods m. The time-length of each period is $1/252$. The discretization error is estimated using $n = 3200$.

m	$\sqrt{\pi_{\inf}(F_0)}$	Computing time	Error	$\sqrt{\pi_{\sup}(F_0)}$	Computing time	Error
21	18.91%	6	2.5×10^{-5}	21.76%	6	1.2×10^{-5}
42	19.01%	12	2.4×10^{-5}	21.37%	12	1.8×10^{-5}
84	18.87%	24	3.1×10^{-5}	21.50%	24	2.9×10^{-5}
168	18.51%	48	4.0×10^{-5}	22.26%	46	3.8×10^{-5}

Table 16: Computing times for robust prices of a capped volatility swap in terms of the number of periods m, with $n = 100$ and $n' = 22$ for the sub–replicating price, and $n = n' = 50$ for the super-replicating price. The time-length of each period is $1/252$. The discretization error is estimated using $n = 800$ and $n' = 86$ for the sub–replicating price, and $n = n' = 400$ for the super-replicating price.

m	$\pi_{\inf}$	Computing time	Error	$\pi_{\sup}$	Computing time	Error
21	7.74%	17	1.0×10^{-3}	21.23%	12	6.3×10^{-5}
42	7.60%	32	3.8×10^{-3}	20.68%	25	2.7×10^{-4}
84	7.68%	61	6.0×10^{-3}	20.48%	50	3.6×10^{-4}
168	7.70%	127	7.3×10^{-3}	20.73%	99	4.5×10^{-4}

L Computing times for variance and volatility swaps in terms of the number of periods

This section uses the same settings as in Subsections 5.2 and 5.3. Tables 15 and 16 list the optimal super-replicating and sub-replicating prices for variance and volatility swaps, for fixed l, n and (for volatility swaps) n', with maturities ranging from 1 to 8 months. These prices are computed in time proportional to the number of traded days m, which is consistent with Theorems 5.3 and 5.4.

M Proof of Theorem 6.1

Let $(\beta^a, \beta^b, \gamma) \in V''$. By the assumptions of the theorem, there is a risk-neutral probability P that satisfies (6.1) and such that

$$\begin{aligned} -q &\leq E_P(\psi) \\ &\leq (\beta^a - \beta^b)^T E_P(\phi) + \gamma \\ &\leq (\beta^a)^T(\pi^a - \delta \mathbf{1}_l) - (\beta^b)^T(\pi^b + \delta \mathbf{1}_l) + \gamma \\ &= (\beta^a, \beta^b)^T((\pi^a, -\pi^b) - \delta \mathbf{1}_{2l}) + \gamma, \end{aligned}$$

where the second inequality follows from (2.6). Hence, for $1 \leq i \leq 2l$,

$$-q \leq (\beta^a, \beta^b)^T((\pi^a, -\pi^b) \pm \delta e_i) + \gamma,$$

and so $V'' \subseteq \mathcal{E}'(2l, (\pi^a, -\pi^b), q, \delta)$. Thus (6.2) follows by arguments similar to those in the proof of Lemma 3.1.

Consider now a vector $(\beta^a, \beta^b, \gamma) \in \mathbb{R}^{2l} \times \mathbb{R}$ that belongs to the ball of radius r centered at $u = (r\mathbf{1}_{2l}, q+\delta)$. Since $||\beta^a - r\mathbf{1}_l|| \leq r$, $\beta^a \in \mathbb{R}^{+l}$. Similarly, $||\beta^b - r\mathbf{1}_l|| \leq r$ and $\beta^b \in \mathbb{R}^{+l}$. By the triangular inequality, $||\beta^a - \beta^b|| \leq 2r \leq \delta$. Furthermore, $q \leq \gamma$ since $|q + \delta - \gamma| \leq r \leq \delta$. Thus, by Assumption A1 and (2.4), $(\beta^a - \beta^b, \gamma) \in V$. On the other hand, since $(\beta^a, \beta^b, \gamma - q - \delta) = (\beta^a, \beta^b, \gamma) - u + r(\mathbf{1}_{2l}, 0)$, $||(\beta^a, \beta^b, \gamma - q - \delta)|| \leq r + r\sqrt{2l}$. As

$$a_0^T(\beta^a, \beta^b, \gamma) = a_0^T(\beta^a, \beta^b, \gamma - q - \delta) + q + \delta,$$

it follows from the Cauchy-Schwartz inequality that $a_0^T(\beta^a, \beta^b, \gamma) \leq 3r\sqrt{l}||a_0|| + q + \delta$. Thus $a_0^T(\beta^a, \beta^b, \gamma) \leq q + 1$, and so $(\beta, \gamma) \in V''$. The rest of the proof is similar to that of Proposition 3.1. $\square$