On The Relation Between Option and Stock Prices: A Convex Optimization Approach

Dimitris Bertsimas * and Ioana Popescu †

July, 1999; Revised July, 2000

Abstract

The idea of investigating the relation of option and stock prices just based on the no-arbitrage assumption, but without assuming any model for the underlying price dynamics has a long history in the financial economics literature. We introduce convex, and in particular semidefinite, optimization methods, duality and complexity theory to shed new light to this relation. For the single stock problem, given moments of the prices of the underlying assets, we show that we can find best possible bounds on option prices with general payoff functions efficiently, either algorithmically (solving a semidefinite optimization problem) or in closed form. Conversely, given observable option prices, we provide best possible bounds on moments of the prices of the underlying assets, as well as on the prices of other options on the same asset by solving linear optimization problems. For options that are affected by multiple stocks either directly (the payoff of the option depends on multiple stocks) or indirectly (we have information on correlations between stock prices), we find non-optimal bounds using convex optimization methods. However, we show that it is NP-hard to find best possible bounds in multiple dimensions. We extend our results to incorporate transactions costs.

*Boeing Professor of Operations Research, Sloan School of Management, Rm. E53-363, Massachusetts Institute of Technology, Cambridge, Mass. 02139. Research partially supported by NSF grant DMI-9610486 and the MIT-Singapore Alliance.

†Department of Mathematics and Operations Research Center, Massachusetts Institute of Technology, Cambridge, Mass. 02139.

1 Introduction.

A central question in financial economics is to find the price of a derivative security given information on the underlying asset. Under the assumption that the price of the underlying asset follows a geometric Brownian motion and using the no-arbitrage assumption, the Black-Scholes formula provides an explicit and insightful answer to this question. Natural questions arise, however, when making no assumptions on the underlying price dynamics, but only using the no-arbitrage assumption:

(a) *What are the best possible bounds for the price of a derivative security with a general payoff function based on the k moments of the price of the underlying asset?*

(b) *Conversely, given observable option prices, what are the best possible bounds that we can derive on the moments of the underlying asset?*

(c) *Given observable option prices, what are the best possible bounds that we can derive on prices of other derivatives on the same asset?*

(d) *How can we derive best possible bounds on derivative securities that are based on multiple underlying assets, either directly (the payoff of the option depends on multiple stocks) or indirectly (we have information on correlations between stock prices) , when partial information on the asset prices and their correlations is given?*

(e) *What is the effect of transaction costs in the above instances?*

The idea of investigating the relation of option and stock prices just based on the no-arbitrage assumption, but without assuming any model for the underlying price dynamics has a long history in the financial economics literature. Cox and Ross [4] and Harrison and Kreps [10] show that the no-arbitrage assumption is equivalent with the existence of a probability distribution π (the so-called martingale measure) such that that option prices become martingales under π. The idea that it is possible in principle to infer the martingale measure from option prices has been introduced by Ross [19]. The idea of using optimization to infer the martingale measure based on option prices is present in the work of Rubinstein [20] who, extending earlier work of Longstaff [16], introduces the idea of deducing the martingale measure from observed European call prices by solving a quadratic optimization problem that measures the closeness of the martingale measure to the lognormal distribution. For related work, see Dupire [6] and Derman and Kani [5]. Closer to the theme of this paper are the papers by Lo [15], who derives best possible closed form bounds on the price of a European call option given the mean and variance of the underlying stock price, by Grundy [9], who extended Lo's work for the case when the first and the kth moments of the stock price are known, and by Boyle and Lin [3], who use semidefinite optimization to find an upper bound on the price of a European call option on the maximum of a number of assets, given the means, variances and covariances of these assets.

Our overall objective in this paper is to shed new light to the relation of option and stock prices, and to demonstrate that the natural way to address this relation, without making distributional assumptions for the underlying price dynamics, but only using the no-arbitrage assumption, is the use of convex optimization methods expanding upon the work in Boyle and Lin [3]. In particular, we give concrete answers to the previous questions (a)-(e) using convex, and in particular semidefinite, optimization techniques, duality, and complexity theory.

In order to motivate our overall approach we formulate the problem of deriving optimal bounds on the price of a European call option given the mean and variance of the underlying stock price. Following Cox and Ross [4] and Harrison and Kreps [10], the no-arbitrage assumption is equivalent with the existence of a probability distribution π (the so-called martingale measure) of the asset price X, such that the price of any European call option with strike price k is given by

$$q(k) = E_\pi[\max(0, X - k)],$$

where the expectation is taken over the unknown distribution π. Note that we have assumed, without loss of generality, that the risk free interest rate is zero. Moreover, given that the mean and variance of the underlying asset are observable:

$$E_\pi[X] = \mu, \qquad \text{and} \qquad Var_\pi[X] = \sigma^2,$$

the problem of finding the best possible upper bound on the call price, written as

$$\max_{X \sim (\mu,\sigma^2)^+} E_\pi[\max(0, X - k)],$$

(where the + operation means that X is defined on $[0, \infty)$) can be formulated as follows:

$$\begin{aligned} \text{maximize} \quad & E_\pi[\max(0, X-k)] \\ \text{subject to} \quad & E_\pi[X] = \mu \\ & Var_\pi[X] = \sigma^2. \\ & \int_0^\infty \pi(x)dx = 1 \\ & \pi(x) \geq 0. \end{aligned}$$

The closed form solution for this optimization problem is due to Scarf [21], in the context of an inventory control problem. Lo [15] observed the direct application of Scarf's result to option pricing. Grundy [9] introduced as open problems several of the problems that we solve here: using known option prices, find sharp upper and lower bounds on the moments of the stock price, and on the price of an option with a different strike price. These problems can be formulated as follows:

$$\begin{aligned} \max/\min \quad & E_\pi[X]\ , \text{ or } \ E_\pi[X^2]\ , \text{ or } \ E_\pi[\max(0, X-k)] \\ \text{subject to} \quad & E_\pi[\max(0, X-k_i)] = q_i\ ,\ i = 1, \ldots, n. \\ & \int_0^\infty \pi(x)dx = 1 \\ & \pi(x) \geq 0. \end{aligned}$$

For a multidimensional example, suppose we have observed the price q_1 of a European call option with strike k_1 for stock 1, and the price q_2 of a European call option with strike k_2 for stock 2. In addition, we have estimated the means μ_1, μ_2, the variances σ_1^2, σ_2^2 and the covariance σ_{12} of the prices of the two underlying stocks. Suppose, in addition, we are interested in obtaining an upper bound on the price of a European call option with strike k for stock 1. Intuition suggests that since the prices of the two stocks are correlated, the price of a call option on stock 1 with strike k might be affected by the available information regarding stock 2. We can find an upper bound on the price of a call option on stock 1 with strike k, by solving the following problem:

$$
\begin{array}{rl}
\text{maximize} & E_\pi[\max(0, X_1 - k)] \\
\text{subject to} & E_\pi[\max(0, X_1 - k_1)] = q_1 \\
& E_\pi[\max(0, X_2 - k_2)] = q_2 \\
& E_\pi[X_1] = \mu_1 \\
& E_\pi[X_2] = \mu_2 \\
& E_\pi[X_1^2] = \sigma_1^2 + \mu_1^2 \\
& E_\pi[X_2^2] = \sigma_2^2 + \mu_2^2 \\
& E_\pi[X_1 X_2] = \sigma_{12} + \mu_1\mu_2 \\
& \int_0^\infty \int_0^\infty \pi(x_1, x_2) dx_1 dx_2 = 1 \\
& \pi(x_1, x_2) \geq 0.
\end{array}
\tag{1}
$$

More generally, questions (a)-(d) above are special cases of the following general optimization problem:

$$
\begin{array}{rll}
\max / \min & E_\pi[\phi(X)] & \\
\text{subject to} & E_\pi[f_i(X)] = q_i\,, \; i = 0, 1, \ldots, n. & \\
& \pi(x) \geq 0, & x \in R_+^m,
\end{array}
\tag{2}
$$

where $X = (X_1, \ldots, X_m)$ is a multivariate random variable, and $\phi : R_+^m \to R$ is a real-valued objective function, $f_i : R_+^m \to R\,,\; i = 1, \ldots, n$ are also real-valued, so-called *moment functions* whose expectations $q_i \in R$, referred to as *moments*, are known and finite. We assume that $f_0(x) = 1$ and $q_0 = E_\pi[f_0(X)] = 1$, corresponding to the implied probability-mass constraint. Questions (a)-(d) introduced earlier can be formulated as follows:

(a) Question (a) for European call options can be formulated as Problem (2) with

$$\phi(x) = \max(0, x - k),\;\; f_i(x) = x^i,\;\; i = 1, \ldots, k,$$

where q_i is the ith moment of the price of the underlying asset.

(b) Question (b) for European call options can be formulated as Problem (2) with

$$\phi(x) = x, \text{ or } \phi(x) = x^2, \text{ and } f_i(x) = \max(0, x - k_i), \ i = 1, \ldots, n.$$

(c) Question (c) for European call options can be formulated as Problem (2) with

$$\phi(x) = \max(0, x - k), \text{ and } f_i(x) = \max(0, x - k_i), \ i = 1, \ldots, n.$$

(d) Question (d) for a general option with payoff $\phi(x_1, \ldots, x_m)$ that is based on m underlying assets can be formulated as Problem (2) with

$$f_i(x) = x_i, \ i = 1, \ldots, m, \ f_{ij}(x) = x_i x_j, \ i, j = 1, \ldots m, \ q_i = \mu_i, \ q_{ij} = \sigma_{ij}^2 + \mu_i \mu_j.$$

When $\phi(x) = \chi_S$ in Problem (2) is the indicator function of a convex set S, and f_i are power functions, then Problem (2) models the problem of finding the best possible bounds on the probability that a multidimensional random variable X belongs in the convex set S, given some joint moments on X. In this context, Problem (2) has received a lot of attention in the 1950s and 1960s. The major duality results from this period are due to Isii [11] and Karlin (see Karlin and Studden [14], p. 472) for the univariate case, and by Isii [12] for the multivariate case. The interested reader is referred to the book of Karlin and Studden [14] for a comprehensive coverage, to Bertsimas and Popescu [2] for a modern treatment, and to Smith [22] for applications in decision analysis.

The contributions and structure of this paper are as follows:

1. We provide in Section 2 an efficient (polynomial time) algorithm for question (a) for a general payoff function $\phi(x)$ by solving a single semidefinite optimization problem, thus generalizing earlier work of Lo [15], Grundy [9] and Boyle and Lin [3]. This result leads to an unexpected connection between finance and semidefinite optimization.

2. We derive in Section 3 closed form optimal bounds on call and put prices given prices

of other calls and puts on the same stock, thus answering question (b).

3. We derive in Section 4 best possible bounds on the mean and variance of the underlying stock price, when prices of options on this stock are given, thus answering question (c).

4. We extend in Section 5 the previous results by taking into account transaction costs, thus answering question (e).

5. We present in Section 6 an efficient (polynomial time) algorithm to provide bounds (although not best possible ones) for options that are affected by multiple stocks using convex, and in particular semidefinite, optimization methods, thus answering question (d). We also show that it is NP-hard to find optimal bounds in multiple dimensions.

2 Bounds On Option Prices Given Moment Information.

We are given the n first moments $(q_1, q_2, \dots, q_n)$, (we let $q_0 = 1$) of the price of an asset, and we are interested in finding the best possible bounds on the price of an option with payoff $\phi(x)$. An example is a European call option with payoff $\phi(x) = \max(0, x - k)$. In Section 2.1, we propose an efficient algorithmic solution for general payoff functions, while in Section 2.2 we provide a new proof based on duality of the closed form upper bound of the price of a European call option derived by Lo [15].

2.1 Bounds Based on Semidefinite Optimization.

As we discussed in the previous section the problem of finding the best upper bound on the price of a European call option with strike k can be formulated as follows.

$$\begin{array}{lll} \text{maximize} & E_\pi[\max(0, X - k)] = \displaystyle\int_0^\infty \max(0, x - k)\pi(x)dx & \\ \text{subject to} & E_\pi[X^i] = \displaystyle\int_0^\infty x^i \pi(x)dx = q_i & i = 0, 1, \dots, n, \\ & \pi(x) \geq 0. & \end{array} \tag{3}$$

In the spirit of linear programming theory (see Smith [22] and Bertsimas and Popescu

[2]), we write the dual of Problem (3) by associating a vector of dual variables $y = (y_0, y_1, \ldots, y_n)$ to each of the constraints in Problem (3). We obtain the following problem:

$$\begin{array}{ll} \text{minimize} & \sum_{r=0}^{n} y_i q_i \\ \text{subject to} & \sum_{r=0}^{n} y_r x^r \geq \max(0, x-k), \quad \forall x \in R_+. \end{array} \tag{4}$$

Isii [11] shows that strong duality holds, i.e., the optimal solution values of Problems (3) and (4) are equal. Thus, by solving Problem (4), we obtain the desired sharp bound. In this section, we show that the general problem (3) can be reformulated as a semidefinite optimization problem for which very efficient algorithmic solutions are known, both from a theoretical (see Nesterov and Nemirovski [18] and Vandenberghe and Boyd [23]) and practical standpoint (see Fujisawa, Kojima and Nakata [7]). The results in the following proposition are inspired by Ben-Tal and Nemirovski [1], p.140-142. The proofs are in the appendix.

Proposition 1

(a) *The polynomial* $g(x) = \sum_{r=0}^{2k} y_r x^r$ *satisfies* $g(x) \geq 0$ *if and only if there exists a positive semidefinite matrix* $X = [x_{ij}]_{i,j=0,\ldots,k}$, *such that*

$$y_r = \sum_{i,j:\ i+j=r} x_{ij}, \qquad r = 0, \ldots, 2k, \qquad X \succeq 0. \tag{5}$$

(b) *The polynomial* $g(x) = \sum_{r=0}^{n} y_r x^r$ *satisfies* $g(x) \geq 0$ *for all* $x \in [0, a]$ *if and only if there exists a positive semidefinite matrix* $X = [x_{ij}]_{i,j=0,\ldots,n}$, *such that*

$$\begin{aligned} 0 &= \sum_{i,j:\ i+j=2l-1} x_{ij}, \qquad && l = 1, \ldots, n, \\ \sum_{r=0}^{l} y_r \binom{k-r}{l-r} a^r &= \sum_{i,j:\ i+j=2l} x_{ij}, \qquad && l = 0, \ldots, n, \\ & X \succeq 0. \end{aligned} \tag{6}$$

(c) *The polynomial* $g(x) = \sum_{r=0}^{n} y_r x^r$ *satisfies* $g(x) \geq 0$ *for all* $x \in [a, \infty)$ *if and only if there exists a positive semidefinite matrix* $X = [x_{ij}]_{i,j=0,\dots,n}$, *such that*

$$\begin{aligned} 0 &= \sum_{i,j:\ i+j=2l-1} x_{ij}, && l = 1, \dots, n, \\ \sum_{r=l}^{k} y_r \binom{r}{l} a^r &= \sum_{i,j:\ i+j=2l} x_{ij}, && l = 0, \dots, n, \\ X &\succeq 0. \end{aligned} \tag{7}$$

(d) *The polynomial* $g(x) = \sum_{r=0}^{k} y_r x^r$ *satisfies* $g(x) \geq 0$ *for all* $x \in [a, b]$ *if and only if there exists a positive semidefinite matrix* $X = [x_{ij}]_{i,j=0,\dots,n}$, *such that*

$$\begin{aligned} 0 &= \sum_{i,j:\ i+j=2l-1} x_{ij}, && l = 1, \dots, n, \\ \sum_{m=0}^{l} \sum_{r=m}^{k+m-l} y_r \binom{r}{m} \binom{k-r}{l-m} a^{r-m} b^m &= \sum_{i,j:\ i+j=2l} x_{ij}, && l = 0, \dots, n, \\ X &\succeq 0. \end{aligned} \tag{8}$$

The next theorem shows that Problem (3) can be solved as a semidefinite optimization problem.

Theorem 1 *The best upper bound on the price of a European call option with strike* k *given the* n *first moments* $(q_1, \dots, q_n)$ $(q_0 = 1)$ *of the underlying stock is given by the solution of* *the following semidefinite optimization problem:*

$$
\begin{array}{lll}
\text{minimize} & \displaystyle\sum_{r=0}^{n} y_i q_i & \\
\text{subject to} & 0 = \displaystyle\sum_{i,j:\ i+j=2l-1} x_{ij}, & l = 1, \ldots, n, \\
& \displaystyle\sum_{r=0}^{l} y_r \binom{k-r}{l-r} k^r = \displaystyle\sum_{i,j:\ i+j=2l} x_{ij}, & l = 0, \ldots, n, \\
& 0 = \displaystyle\sum_{i,j:\ i+j=2l-1} z_{ij}, & l = 1, \ldots, n, \\
& (y_0 + k) + (y_1 - 1)k + \displaystyle\sum_{r=2}^{k} y_r k^r = z_{00}, & \\
& (y_1 - 1)k + \displaystyle\sum_{r=2}^{k} y_r r k^r = \displaystyle\sum_{i,j:\ i+j=2} z_{ij}, & \\
& \displaystyle\sum_{r=l}^{k} y_r \binom{r}{l} k^r = \displaystyle\sum_{i,j:\ i+j=2l} z_{ij}, & l = 2, \ldots, n, \\
& X, Z \succeq 0. &
\end{array}
\tag{9}
$$

Proof: We note that the feasible region of Problem (4) can be written as

$$
\begin{aligned}
\sum_{r=0}^{n} y_r x^r &\geq 0 \quad \text{for all } x \in [0, k], \\
(y_0 + k) + (y_1 - 1)x + \sum_{r=2}^{n} y_r x^r &\geq 0 \quad \text{for all } x \in [k, \infty).
\end{aligned}
$$

by applying Proposition 1 (b), (c) we reformulate Problem (4) as the semidefinite optimization Problem (9). □

We next consider an option with payoff function given as follows:

$$
\phi(x) = \begin{cases}
\phi_0(x), & x \in [0, k_1], \\
\phi_1(x), & x \in [k_1, k_2], \\
\vdots & \vdots \\
\phi_{d-1}(x), & x \in [k_{d-1}, k_d], \\
\phi_d(x), & x \in [k_d, \infty),
\end{cases}
\tag{10}
$$

where the functions $\phi_r(x)$, $r = 0, 1, \ldots, d$ are polynomials. Given the generality of the payoff function (10), we can approximate the payoff of any option using the payoff function (10). In this case the dual problem becomes:

$$\begin{array}{ll} \text{minimize} & \displaystyle\sum_{r=0}^{n} y_i q_i \\ \text{subject to} & \displaystyle\sum_{r=0}^{n} y_r x^r \geq \begin{cases} \phi_0(x), & x \in [0, k_1], \\ \phi_1(x), & x \in [k_1, k_2], \\ \quad\vdots & \quad\vdots \\ \phi_{d-1}(x), & x \in [k_{d-1}, k_d], \\ \phi_d(x), & x \in [k_d, \infty), \end{cases} \end{array} \tag{11}$$

The next theorem shows that the problem of finding best possible bounds on an option with a general piecewise polynomial payoff function is efficiently solvable both practically and theoretically as a semidefinite optimization problem.

Theorem 2 *The best possible bounds for the price of an option with a piecewise polynomial payoff function $\phi(x)$ shown in (10), given moments of the underlying asset, can be solved efficiently as a semidefinite optimization problem.*

Proof:

The constraint set for Problem (11) can be written as follows:

$$\sum_{r=0}^{n} y_r x^r \geq \phi_i(x), \qquad x \in [k_{i-1}, k_i], \quad i = 1, \ldots, d+1,$$

with $k_0 = 0$, $k_{d+1} = \infty$. Let $\phi_i(x) = \displaystyle\sum_{r=0,\ldots,m_i} a_{ir} x^r$, and assume without loss of generality that $m_i \leq n$. Then, the constraint set for Problem (11) can be equivalently written as

$$\sum_{r=0}^{m_i} (y_r - a_{ir}) x^r + \sum_{r=m_i+1}^{n} y_r x^r \geq 0, \qquad x \in [k_{i-1}, k_i], \quad i = 1, \ldots, d+1.$$

For the the interval $[k_0, k_1]$ we apply Proposition 1(b), for the intervals $[k_{i-1}, k_i]$, $i = 2, \ldots, d$, we apply Proposition 1(d), and for the interval $[k_d, \infty)$, we apply Proposition 1(c), to express Problem (11) as a semidefinite optimization problem. $\square$

2.2 Closed Form Bounds.

In this section, we provide a new proof from first principles of a closed form optimal bound of the price of a European call option with strike k.

Theorem 3 (Optimal upper bound on option prices, Lo [15]) *The optimal upper bound on the price of an option with strike k, on a stock whose price at maturity has a known mean μ and variance σ^2, is computed by:*

$$\max_{X \sim (\mu, \sigma^2)^+} E[\max(0, X-k)] = \begin{cases} \frac{1}{2}\left[(\mu - k) + \sqrt{\sigma^2 + (\mu-k)^2}\right], & \text{if } k \geq \frac{\mu^2+\sigma^2}{2\mu}, \\ \mu - k + k\frac{\sigma^2}{\mu^2+\sigma^2}, & \text{if } k < \frac{\mu^2+\sigma^2}{2\mu}. \end{cases}$$

Proof:

The optimal upper bound on the price of a European call option with strike k is given as the solution of Problem (4), which in this case is formulated by associating dual variables y_0, y_1, y_2 with the probability-mass, mean and respectively, variance constraints. We obtain the following dual formulation:

$$\begin{array}{rl} \text{minimize} & (\mu^2 + \sigma^2)\, y_2 + \mu\, y_1 + y_0 \\ \text{subject to} & g(x) = y_2\, x^2 + y_1\, x + y_0 \ \geq \ \max(0, x-k)\,, \ \ \forall x \geq 0\,. \end{array}$$

A dual feasible function $g(\cdot)$ is any quadratic function that, on the positive orthant, is nonnegative and lies above the line $(x-k)$. In an optimal solution, such a quadratic should be tangent to the line $(x-k)$, so we can write $g(x) - (x-k) = a(x-b)^2$, for some $a \geq 0$. The non-negativity constraint on $g(\cdot)$ can be expressed as $a(x-b)^2 + x - k \geq 0\,, \ \forall x \geq 0$. Let $x_0 = b - \dfrac{1}{2a}$ be the point of minimum of this quadratic. Depending whether x_0 is nonnegative or not, either the inequality at $x = x_0$ or at $x = 0$ is binding in an optimal solution. We have two cases:

(a) If $b \geq \dfrac{1}{2a}$, then $-\dfrac{1}{4a} + b - k = 0$ (binding constraint at $x = x_0$);

Substituting $a = \dfrac{1}{4(b-k)}$ in the objective, we obtain:

$$\begin{aligned}\max_{X\sim(\mu,\sigma^2)^+} E[\max(0, X-k)] &= \min_b \frac{((\mu-k)+(b-k))^2+\sigma^2}{4(b-k)} \\ &= \frac{1}{2}\left[(\mu-k)+\sqrt{\sigma^2+(\mu-k)^2}\right],\end{aligned}$$

achieved at $b_0 = \dfrac{\mu^2+\sigma^2}{\mu}$. Let $a_0 = \dfrac{1}{4(b_0-k)}$.

This bound is valid whenever $b_0 \geq \dfrac{1}{2a_0} = 2(b_0-k)$, that is $\dfrac{\mu^2+\sigma^2}{2\mu} \leq k$.

(b) If $b < \dfrac{1}{2a}$, then $ab^2 - k = 0$ (binding constraint at $x = 0$).

Substituting $a = \dfrac{k}{b^2}$ in the objective, we obtain:

$$\max_{X\sim(\mu,\sigma^2)^+} E[\max(0, X-k)] = \min_b \frac{k}{b^2}(\mu^2+\sigma^2) - 2\frac{k}{b}\mu + \mu = \mu - k\frac{\mu^2}{\mu^2+\sigma^2}.$$

achieved at $b_0 = \dfrac{\mu^2+\sigma^2}{\mu}$. Let $a_0 = \dfrac{k}{b_0^2}$. This bound is valid whenever $b_0 < \dfrac{1}{2a_0} = \dfrac{b_0^2}{2k}$, that is $\dfrac{\mu^2+\sigma^2}{2\mu} > k$. $\square$

3 Bounds On Option Prices Given Other Option Prices.

In this section, we derive closed form optimal upper and lower bounds on the price of a European call option on a single stock, when prices of other options with the same exercise date but different strikes on the same stock are known. For simplicity and without loss of generality, we assume that the risk-free interest rate is zero. In this section, we ignore transaction costs, the effect of which will be discussed in Section 5.

3.1 Bounds on Call Prices.

Let X be the random variable that represents the price of the underlying stock. We are given prices $q(k_i) = q_i = E[\max(0, X-k_i)]$ of call options on the same stock with strikes

$0 \leq k_1 \leq k_2 \leq \ldots \leq k_n$ and the same exercise date, and we want to compute optimal upper and lower bounds on $q(k) = E[\max(0, X - k)]$, for a given strike price k.

For notation purposes we set $k_0 = 0$ and $q_0 = q(0) = E[\max(0, X - 0)] = E[X]$. In some cases it may also be useful to assume an upper bound K on the price X of the stock at the time of maturity of the calls. This information can be easily integrated in this framework by defining $k_{n+1} = K$ and $q_{n+1} = q(K) = E[\max(0, X - K)] = 0$. If no such upper bound is assumed, then we assume $k_{n+1} = \infty$. We say that a given function $q(\cdot)$ is a valid call pricing function if there exists a distribution of the stock price X, such that $q(k) = E[\max(0, X - k)]$, $\forall\, k \geq 0$.

Theorem 4 (Optimal Bounds on Call Prices)
Given valid prices $q_i = q(k_i) = E[\max(0, X - k_i)]$ of call options with strikes $0 \leq k_1 \leq k_2 \leq \ldots \leq k_n$, on a stock X, the range of possible valid prices for a call option with strike price k, where $k \in (k_j, k_{j+1})$, for some $j = 0, \ldots, n$ is $[\, q^-(k)\,,\ q^+(k)\,]$, where:

$$
\begin{aligned}
q^-(k) &= \max\left(q_j \frac{k - k_{j-1}}{k_j - k_{j-1}} + q_{j-1} \frac{k_j - k}{k_j - k_{j-1}}\,,\quad q_{j+1} \frac{k_{j+2} - k}{k_{j+2} - k_{j+1}} + q_{j+2} \frac{k - k_{j+1}}{k_{j+2} - k_{j+1}} \right) \\
q^+(k) &= q_j \frac{k_{j+1} - k}{k_{j+1} - k_j} + q_{j+1} \frac{k - k_j}{k_{j+1} - k_j}\,.
\end{aligned}
\tag{12}
$$

In order to obtain some intuition on the nature of these bounds we note that for a given function $q(\cdot)$ to be a valid call pricing function, we need the existence of a nonnegative random variable X such that $q(k) = E[\max(0, X - k)]$, $\forall\, k \geq 0$. Clearly, $q(\cdot)$ is decreasing and convex. What Theorem 4 proves is that the necessary and sufficient conditions for $q(\cdot)$ to define a valid call pricing function is for it to be decreasing and convex. In particular, the values of $q^-(k)$ and $q^+(k)$ given above are precisely determined by the monotonicity and convexity of the call pricing function $q(\cdot)$. Figure 1 depicts the construction of the bounds $q^-(k)$ and $q^+(k)$ geometrically in a concrete example. Moreover, the range of prices of a call option with strike price $k \in (k_j, k_{j+1})$ is constrained only by the prices q_{j-1}, q_j of the two options with the closest strikes to the left of k and to the right of k, q_{j+1} and q_{j+2}.

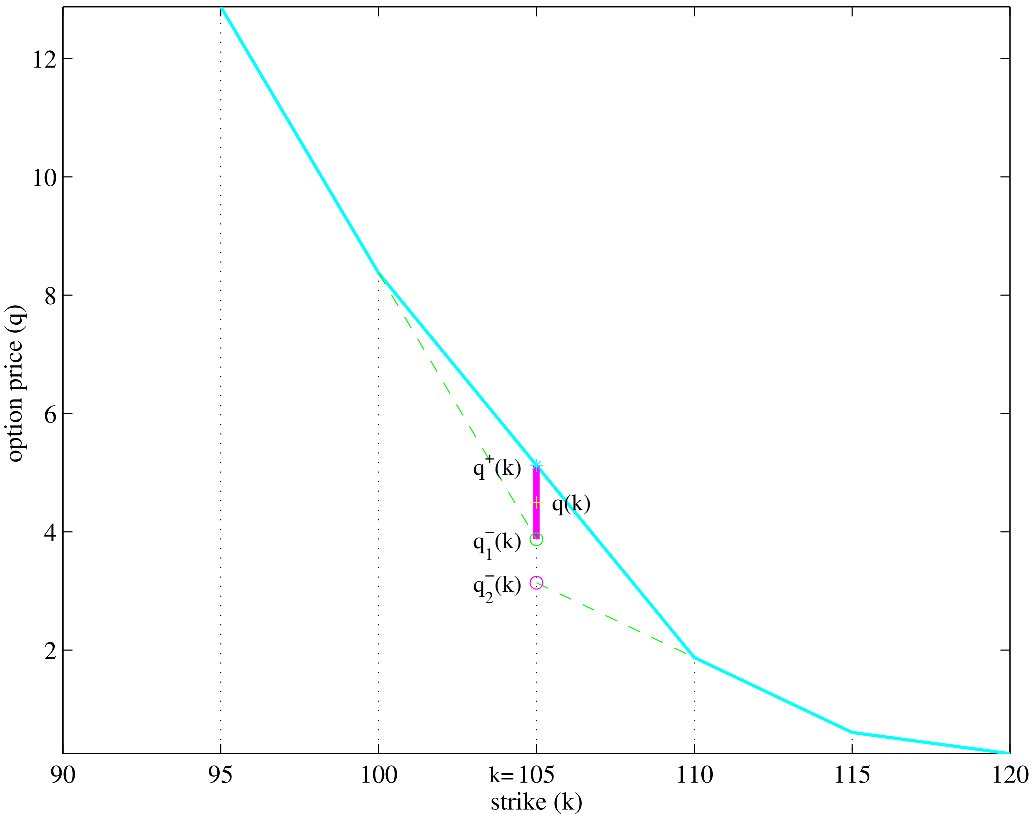

Figure 1: The optimal upper and lower bounds on the price of a call option, given prices of calls on the same stock, with different strikes and the same maturity date. (actual data quoted from The Wall Street Journal, July 7, 1998: Microsoft July '98 call options with: $k_i = [95, 100, 110, 115, 120]$, $q_i = [12\frac{7}{8}, 8\frac{3}{8}, 1\frac{7}{8}, \frac{5}{8}, \frac{1}{4}]; k = 105, q(k) = 4\frac{1}{2})$. Note that the bounds are derived by the convexity and monotonicity of the the price function $q(\cdot)$.

The bounds (12) are only relevant when the given options are correctly priced. Interestingly, this is not always the case, as one can see in the actual examples given in Figure 2, where some of the call pricing functions are clearly non-convex, and so the upper bounds computed by Theorem 4 may turn out smaller than the respective lower bounds (see the explanation in the caption of Figure 2).

Proof of Theorem 4:

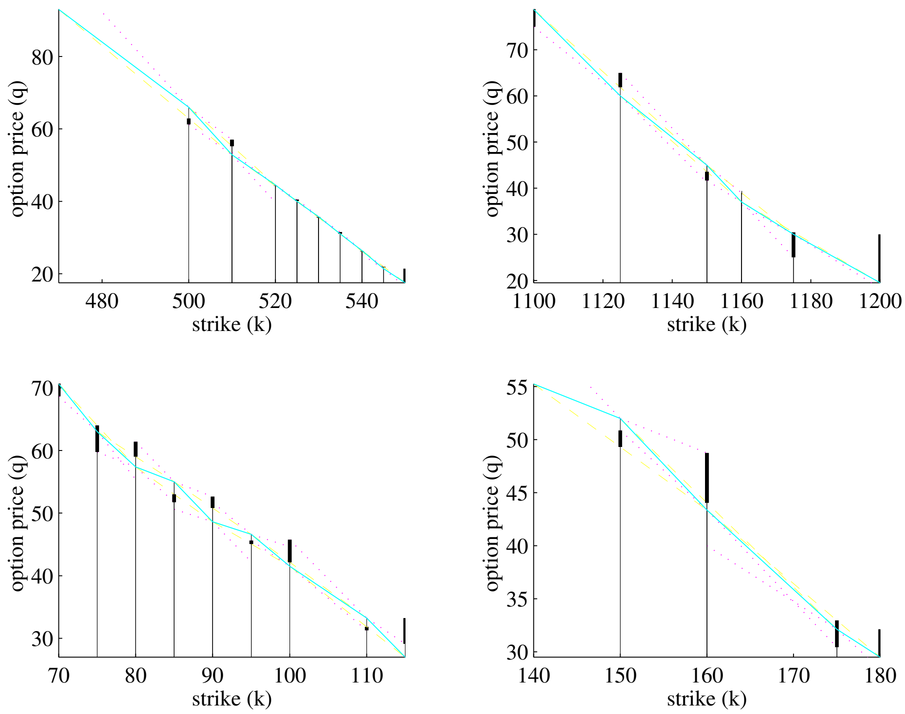

Figure 2: The optimal upper and lower bounds for each call price, determined by the prices of the neighboring calls. Clockwise: $S\&P100$ July '98, $S\&P500$ Sep '98, Yahoo Aug '98, Amazon July '98. Call prices from The Wall Street Journal July 7, 1998. This apparent mispricing can be explained by noting that these are closing prices, so these prices might not all be present simultaneously. Moreover, transaction costs are ignored.

The Lower Bound Problem. We first consider the lower bound problem and formulate it as a continuous optimization problem over all feasible stock price densities $\pi(x)$, as follows:

$$
\begin{aligned}
q^-(k) = \text{minimize} \quad & \int_k^\infty (x-k)\pi(x)dx \\
\text{subject to} \quad & \int_{k_i}^\infty (x-k_i)\pi(x)dx = q_i, \quad i=1,\ldots,n, \\
& \int_0^\infty \pi(x)dx = 1, \\
& \pi(x) \geq 0, \qquad\qquad\qquad \forall\, x \in R_+.
\end{aligned}
\tag{13}
$$

If we restrict our horizon to stock price distributions $p_x = P(X=x)$ over a **discrete** range of values $S \subseteq R_+$, that include the strike prices $k_i \in S$, $i=1,\ldots,n$, we can formulate the restricted problem as:

$$
\begin{aligned}
q_R^-(k) = \text{minimize} \quad & \sum_{x\geq k}(x-k)p_x \\
\text{subject to} \quad & \sum_{x\geq k_i}(x-k_i)p_x = q_i, \quad i=1,\dots,n, \\
& \sum_{x\geq 0} p_x = 1, \\
& p_x \geq 0, \qquad\qquad \forall\, x\in S.
\end{aligned}
\tag{14}
$$

Clearly $q_R^-(k) \geq q^-(k)$, since the minimization in Problem (14) is over a restricted set of distributions. We will show that $q_R^-(k) = q^-(k)$. We construct the corresponding dual problems by associating a dual variable $u_i,\ i=1,\dots,n$ with each of the first n constraints, and a dual variable v for the probability mass constraint. The dual of Problem (13) is:

$$
\begin{aligned}
q_D^-(k) = \text{maximize} \quad & v + \sum_{i=1}^{n} q_i u_i \\
\text{subject to} \quad & g(x) \ = \ v + \sum_{i\,|\,k_i\leq x}(x-k_i)u_i \ \leq \begin{cases} 0, & 0\leq x<k, \\ x-k, & x\geq k. \end{cases}
\end{aligned}
\tag{15}
$$

The dual problem of the restricted problem (14) is the same as (15), except the constraints need only hold on the discrete set of points $x\in S$, where X ranges. We denote its optimal solution values as $q_{RD}^-(k)$. Notice that for both problems, the dual feasible function $g(x)$ is piecewise linear, in which the slope changes at the points k_i, $i=1,\dots,n$, and therefore it is sufficient to solve each problem with constraints only at the points k_i. Thus the two dual problems are equivalent to:

$$
\begin{array}{rrlllll}
q_{RD}^-(k) = \text{maximize} & & v & + \sum_{i=1}^{n} q_i u_i & & & \\
\text{subject to} & & & & & & \\
& g(k_1) = & v & & & & \leq 0 \\
& g(k_2) = & v & + (k_2 - k_1)u_1 & & & \leq 0 \\
& \vdots & & \vdots & & & \vdots \\
& g(k_j) = & v & + (k_j - k_1)u_1 & + \ldots + & (k_j - k_{j-1})u_{j-1} & \leq 0 \\
& g(k) = & v & + (k - k_1)u_1 & + \ldots + & (k - k_j)u_j & \leq 0 \\
& g(k_{j+1}) = & v & + (k_{j+1} - k_1)u_1 & + \ldots + & (k_{j+1} - k_j)u_j & \leq k_{j+1} - k \\
& \vdots & & \vdots & \vdots & \vdots & \vdots \\
& g(k_n) = & v & + (k_n - k_1)u_1 & + \ldots + & (k_n - k_{n-1})u_{n-1} & \leq k_n - k \\
& & u_1 & + u_2 & + \ldots + & u_n & \leq 1,
\end{array}
\tag{16}
$$

where the last constraint is meant to capture the limiting situation as $x \to \infty$. We have $q_D^-(k) = q_{RD}^-(k)$, and weak duality holds for both primal-dual pairs, which means: $q_R^-(k) \geq q^-(k) \geq q_D^-(k)$. Moreover, strong duality holds for the discretized version (14), since these are linear optimization problems, and therefore, $q_R^-(k) = q_{RD}^-(k) = q_D^-(k)$. This shows that

$$
q^-(k) = q_R^-(k) = q_D^-(k). \tag{17}
$$

Moreover, there exists a discrete stock-price distribution that achieves the bound $q^-(k)$.

We next proceed to solve Problem (16). This is a linear optimization problem with $n+2$ constraints and $n+1$ variables whose optimum, if it exists, is achieved at a basic feasible solution. In an optimal basic feasible solution, $n+1$ of the constraints must be binding, including the one at k, that is the constraint $g(k) \leq 0$. In this case, the constraints $g(k_j) \leq 0$ and $g(k_{j+1}) \leq k_{j+1} - k$ cannot be simultaneously binding. We have two cases:

Case 1. Constraint $g(k_j) \le 0$ is not binding: In this case we obtain the following optimal solution:

$$g(x) = \begin{cases} 0\,, & x \le k_{j-1}, \\ \dfrac{k_j - k}{k_j - k_{j-1}}(x - k_{j-1})\,, & x \in (k_{j-1}, k_j), \\ x - k\,, & x \ge k_j, \end{cases}$$

that is the corresponding dual variables are:

$$\begin{aligned} u_{j-1} &= \frac{k_j - k}{k_j - k_{j-1}}, \\ u_j &= \frac{k - k_{j-1}}{k_j - k_{j-1}}, \\ u_i &= 0\,,\ \forall\ i \ne j-1, j, \\ v &= 0\,. \end{aligned}$$

The corresponding dual optimal objective value in this case is:

$$q_1^-(k) = q_j \frac{k - k_{j-1}}{k_j - k_{j-1}} + q_{j-1} \frac{k_j - k}{k_j - k_{j-1}}.$$

Case 2. Constraint $g(k_{j+1}) \le k_{j+1} - k$ is not binding: In this case we obtain the following optimal solution:

$$g(x) = \begin{cases} 0\,, & x \le k_{j+1}, \\ \dfrac{k_{j+2} - k}{k_{j+2} - k_{j+1}}(x - k_{j+1})\,, & x \in (k_{j+1}, k_{j+2}), \\ x - k\,, & x \ge k_{j+2}\,. \end{cases}$$

The corresponding dual variables are:

$$\begin{aligned}
u_{j+1} &= \frac{k_{j+2} - k}{k_{j+2} - k_{j+1}}, \\
u_{j+2} &= \frac{k - k_{j+1}}{k_{j+2} - k_{j+1}}, \\
u_i &= 0\,,\ \forall\, i \neq j+1, j+2, \\
v &= 0\,.
\end{aligned}$$

The corresponding dual optimal objective value in this case is:

$$q_2^-(k) = q_{j+1} \frac{k_{j+2} - k}{k_{j+2} - k_{j+1}} + q_{j+2} \frac{k - k_{j+1}}{k_{j+2} - k_{j+1}}.$$

The desired optimal lower bound is given by: $q^-(k) = \max\left(q_1^-(k), q_2^-(k)\right)$, which leads to the lower bound expression in Eq. (12). Note that an extremal distribution of the stock price X that achieves this bound is given by the corresponding optimal solution of the discretized primal problem.

The Upper Bound Problem. Using the same procedure, we formulate the optimal upper bound problem as a continuous optimization problem over all feasible stock price densities $\pi(x)$:

$$\begin{aligned}
q^+(k) = \text{maximize} \quad & \int_k^\infty (x-k)\pi(x)dx \\
\text{subject to} \quad & \int_{k_i}^\infty (x-k_i)\pi(x)dx = q_i, \quad i = 1, \ldots, n, \\
& \int_0^\infty \pi(x)dx = 1, \\
& \pi(x) \geq 0, \qquad\qquad\qquad \forall x \in R_+,
\end{aligned} \tag{18}$$

and solve the corresponding dual problem:

$$
\begin{aligned}
q_D^+(k) = \text{minimize} \quad & v + \sum_{i=1}^{n} q_i u_i \\
\text{subject to} \quad & g(x) = v + \sum_{i \,|\, k_i \le x} (x - k_i) u_i \le \begin{cases} 0, & 0 \le x < k, \\ x - k, & x \ge k. \end{cases}
\end{aligned} \tag{19}
$$

Similarly to the lower bound problem we prove that strong duality holds, the primal is equivalent to its discretized version, and it is sufficient to solve the dual problem with constraints only at the points $k_i,\ i = 1, \ldots, n$. We obtain that

$$
g(x) = \begin{cases} 0, & x \le k_j, \\ \dfrac{k_{j+1} - k}{k_{j+1} - k_j}\,(x - k_j), & x \in (k_j, k_{j+1}), \\ x - k, & x \ge k_{j+1}, \end{cases}
$$

and the corresponding dual variables are:

$$
\begin{aligned}
u_j &= \frac{k_{j+1} - k}{k_{j+1} - k_j}, \\
u_{j+1} &= \frac{k - k_j}{k_{j+1} - k_j}, \\
u_i &= 0,\ \forall\ i \ne j, j+1, \\
v &= 0.
\end{aligned}
$$

The corresponding dual optimal objective value in this case is:

$$
q^+(k) = q_j \frac{k_{j+1} - k}{k_{j+1} - k_j} + q_{j+1} \frac{k - k_j}{k_{j+1} - k_j}. \qquad \square
$$

3.2 Bounds on Prices of Mixed Options.

We now extend the results of the previous section to put options and combinations of calls and puts. Let $p(k)$ be the price of a put option with strike k and the same exercise date. Then $p(k) = E[\max(0, k - X)]$. Since the price of a call with the same strike satisfies

$q(k) = E[\max(0, X - k)]$, then

$$p(k) - q(k) = k - E[X].$$

Clearly the function $p(\cdot)$ is increasing and concave. Similarly to the case of calls, these conditions are in fact necessary and sufficient for put pricing functions to be valid. In other words, if the prices p_i of only puts are known with strikes k_i, then the best possible bounds for the price of a put with strike $k \in (k_j, k_{j+1})$ are $p^-(k) < p(k) < p^+(k)$ with

$$\begin{aligned} p^-(k) &= p_j \frac{k_{j+1} - k}{k_{j+1} - k_j} + p_{j+1} \frac{k - k_j}{k_{j+1} - k_j}, \\ p^+(k) &= \min\left(p_j \frac{k - k_{j-1}}{k_j - k_{j-1}} + p_{j-1} \frac{k_j - k}{k_j - k_{j-1}}\ , \ \ p_{j+1} \frac{k_{j+2} - k}{k_{j+2} - k_{j+1}} + p_{j+2} \frac{k - k_{j+1}}{k_{j+2} - k_{j+1}} \right). \end{aligned} \tag{20}$$

Suppose that we are given prices of call and put options with various strikes k_i, and we want to find optimal bounds on prices of a call or put option with strike k. Notice that if we know the prices of both a call and of a put option with a certain strike k_i, then we can derive the expected stock price from the put-call parity formula $E[X] = q_i - p_i + k_i$. Now, if we know the expected price of the stock $E[X]$, the problem can be directly reduced to the one we solved in Section 3.1, by simply writing all option prices in terms of call prices, using the put-call parity result.

Finally, suppose we are given prices q_i of calls with strikes $k_i, i = 1, \ldots, n$ and prices p_i of puts with strikes $c_i, i = 1, \ldots, m$, such that $c_i \neq k_j$ for all i, j. We are interested in finding the best possible bounds for a call with strike k. In this case we cannot determine $E[X]$ uniquely from the put-call parity. By the put-call parity we can transform the given put prices to corresponding call prices q_i' given by:

$$q_i' = p_i + E[X] - c_i.$$

We sort the strikes k_i and c_i. We want the call prices q_i' to be consistent. We apply Theorem 4 for the calls with strikes c_i as well as the call with strike k. This leads to linear inequalities involving the two unknowns $q(k)$ and $E[X]$. Solving the resulting linear optimization problem with the objective of maximizing or minimizing $q(k)$ gives optimal upper and lower bounds on $q(k)$.

4 Bounds On The Variance Of The Stock Price, Given Option Prices.

In this section, we determine optimal bounds on the variance of a stock price X, when prices of options with different strikes and same exercise date T on that stock are known.

4.1 Bounds on the Mean.

Typically the risk-neutral mean of the price of a stock is known (it is the present price). For completeness, we derive, however, bounds in case The present price is not observable.

Call Options. The bounds on the expected stock price given call prices information are easy to derive, since we can interpret $E[X] = E[\max(0, X - 0)] = q(0)$ as the price of a call option with zero strike. The result of Theorem 4 can be applied in this case to find the following optimal bounds on $E[X] = q(0) \in [\, q^-(0),\ q^+(0)\,]$:

$$\begin{aligned} q^-(0) &= \frac{q_1 k_2 - q_2 k_1}{k_2 - k_1} &&= M^-, \\ q^+(0) &= q_1 + k_1 &&= M^+. \end{aligned} \tag{21}$$

Notice that these bounds only depend on the prices of the two calls with smallest strikes.

Mixed Put And Call Options. Now suppose we are given miscellaneous prices of either calls or puts with different strikes. The optimal bounds M^- and M^+ on the expected stock price $E[X]$ can be determined by converting the put prices into call prices by the put- call parity result, and then constraining the call pricing function to be decreasing and convex, using the bounds from Theorem 4.

4.2 Upper Bounds on the Variance.

In order to compute bounds on the variance we need to assume a finite upper bound K on the stock price X at time T. We incorporate this information by introducing a call option with strike K with price equal to zero, i.e., $k_{n+1} = K$ and $q_{n+1} = q(k_{n+1}) = 0$.

Call Options. Suppose first our information consists of call prices only. We want to determine optimal upper bounds on the variance $Var[X] = E[X^2] - E[X]^2$ of a stock price X, when prices q_i of call options with strikes $k_i, i = 1, \ldots, n$ on that stock are known.

We can formulate this as an optimization problem as follows:

$$\begin{aligned} V^+ = \text{maximize} \quad & \int_0^K x^2\pi(x)dx - \left(\int_0^K x\pi(x)dx\right)^2 & \\ \text{subject to} \quad & \int_{k_i}^K (x-k_i)\pi(x)dx = q_i, & i = 1,\ldots,n, \\ & \int_0^K \pi(x)dx = 1 & \\ & \pi(x) \geq 0, & x \in [0,K]. \end{aligned} \tag{22}$$

Theorem 5 (Optimal upper bound on the variance of the stock price.)

(a) *Given prices q_i of European calls with strikes k_i, and assuming that the mean M of the stock price is known, the optimal upper bound on the variance of the stock price is given by:*

$$V^+(M) = \sum_{i=1}^{n}(k_{i+1} - k_{i-1})q_i + k_1 M - M^2. \tag{23}$$

(b) *If the mean price is not known, then the optimal upper bound is given by:*

$$V^+ = \begin{cases} \displaystyle\sum_{i=1}^{n}(k_{i+1} - k_{i-1})q_i + \frac{k_1^2}{4}, & \text{if} \quad \displaystyle\frac{q_1k_2 - q_2k_1}{k_2 - k_1} < \frac{k_1}{2}, \\ \displaystyle\sum_{i=1}^{n}(k_{i+1} - k_{i-1})q_i + k_1\frac{q_1k_2 - q_2k_1}{k_2 - k_1} - \left(\frac{q_1k_2 - q_2k_1}{k_2 - k_1}\right)^2, & \text{if} \quad \displaystyle\frac{q_1k_2 - q_2k_1}{k_2 - k_1} \geq \frac{k_1}{2}. \end{cases} \tag{24}$$

Proof:

(a) We first solve Problem (22) for a fixed value $M = E[X]$ obtaining an optimal value $V^+(M)$, which we then optimize over all feasible values of $M \in [M^-, M^+]$. We solve the following auxiliary problem with arbitrary fixed expected stock price $E[X] = M$.

$$
\begin{aligned}
Z^+(M) = V^+(M) + M^2 = \text{maximize} \quad & \int_0^K x^2 \pi(x) dx \\
\text{subject to} \quad & \int_0^K x\pi(x) dx = M = q_0 \\
& \int_{k_i}^K (x - k_i)\pi(x) dx = q_i, \quad i = 1, \ldots, n \\
& \int_0^K \pi(x) dx = 1 \\
& \pi(x) \geq 0, \qquad\qquad x \in [0, K] \,.
\end{aligned}
$$

We consider the corresponding dual problem:

$$
\begin{aligned}
Z_D^+(M) = \text{minimize} \quad & v + u_0 M + \sum_{i=1}^n u_i\, q_i \\
\text{subject to} \quad & g(x) = v + u_0 x + \sum_{i \,|\, k_i \leq x} (x - k_i) u_i \geq x^2, \qquad 0 \leq x \leq K.
\end{aligned}
$$

Again, the optimum is obtained by forcing the constraints to be binding at the points k_i: $g(k_i) = k_i^2$, $i = 0, \ldots, n+1$. The corresponding dual solution is $v = k_0 = 0$, $u_0 = k_1$, $u_i = k_{i+1} - k_{i-1}$, $i = 1, \ldots, n$, and the optimal objective value is $Z^+(M) = Z_D^+(M) = k_1 M + \sum_{i=1}^n (k_{i+1} - k_{i-1}) q_i$. Thus, given the mean price M, the optimal upper bound on the variance of X is given by Eq. (23).

(b) We next optimize over all feasible values of $M \in [M^-, M^+]$ to determine the upper bound on the variance:

$$\begin{aligned} V^+ = \text{maximize} \quad & \sum_{i=1}^{n}(k_{i+1} - k_{i-1})q_i + k_1 M - M^2 \\ \text{subject to} \quad & M^- = \frac{q_1 k_2 - q_2 k_1}{k_2 - k_1} \ \leq \ M \ \leq \ M^+ = q_1 + k_1. \end{aligned}$$

This is a concave quadratic optimization problem that can be solved in closed form leading to the closed form bound (24). □

Mixed Put And Call Options.

Suppose now that we are given either call or put prices for various strikes $q_i = q(k_i)$ for all $i \in Q$, and $p_i = p(k_i)$ for all $i \in P$, where P and Q are two sets of indices so that $P \cup Q = \{1, \ldots, n+1\}$ and $n+1 \in Q$, since we assumed $q_{n+1} = q(k_{n+1}) = q(K) = 0$. We transform the puts to corresponding calls with prices:

$$q_i' = p_i - k_i + M, \qquad i \in P.$$

We sort the strikes in $P \cup Q$, and apply the bound (23) for the sequence of calls with prices q_i, $i \in Q$ and q_i', $i \in P$. Note that if P and Q are not disjoint, then we can determine the value of $M = E[X]$ from the put-call parity result for a pair (p_j, q_j) with $j \in P \cap Q$. When the sets P and Q are disjoint, we can obtain an interval $[M^-, M^+]$, in which M lies using the technique of Section 4.1. Applying the bound (23), we will find that a optimal upper bound for the variance of the price given M is a concave quadratic function $V^+(M)$ of M. Then, the optimal upper bound on the variance given M is given by

$$V^+ = \max_{M \in [M^-, M^+]} V^+(M).$$

4.3 Lower Bounds on the Variance.

Call Options. Suppose for now that the available information consists of call prices only. We denote $q_0 = M$ and $k_0 = 0$. We prove the following result:

Theorem 6 (Optimal lower bound on the variance of the stock price)

(a) *Given prices q_i of European calls with strikes k_i, and assuming the mean $M = q_0$ of the stock price is known, the optimal lower bound on the variance of the price is given by*

$$V^-(M) = \sum_{i=0}^{n+1} t_i k_i^2 - Z(M) - M^2, \tag{25}$$

where $Z(M)$ is the objective value of the following quadratic network flow problem:

$$\begin{aligned} Z(M) = \text{minimize} \quad & \sum_{i=0}^{n+1} t_i\, d_i^2 \\ \textit{subject to} \quad & d_i + d_{i-1} \geq c_i, \quad i = 1, \ldots, n+1 \\ & d_i \geq 0, \qquad\qquad i = 0, \ldots, n+1 \end{aligned} \tag{26}$$

with $t_i = T_{i-1} - T_i \geq 0$, $T_i = -\dfrac{q_{i+1} - q_i}{k_{i+1} - k_i}$, $i = 0, \ldots, n$, ($T_{-1} = 1, T_{n+1} = 0$), and $c_i = k_i - k_{i-1} \geq 0$.

(b) *If the mean price is not known, then the optimal lower bound is given by:*

$$V^- = \min_{M \in [M^-, M^+]} V^-(M), \tag{27}$$

with $M^- = \dfrac{q_1 k_2 - q_2 k_1}{k_2 - k_1}$ and $M^+ = q_1 + k_1$.

Proof:

(a) Formulating a minimization optimization problem analogous to the maximization problem (22), and taking the dual, we obtain that given $M = q_0$, the lower bound is given by:

$$\begin{aligned} V^-(M) + M^2 = \text{maximize} \quad & v + \sum_{i=0}^{n} q_i u_i \\ \text{subject to} \quad & g(x) = v + \sum_{i \mid k_i \le x} (x - k_i) u_i \le x^2, \qquad 0 \le x \le K. \end{aligned}$$

For the upper bound problem, we just needed the constraints at the points k_i to be binding, namely: $g(k_i) = k_i^2$, $i = 0, \ldots, n+1$. This is not sufficient for the lower bound. To insure feasibility, we also need to make sure that the line segment $(g(k_i), g(k_{i+1}))$ lies below the quadratic x^2, on each interval $x \in [k_i, k_{i+1}]$. This can be interpreted geometrically as follows. Consider the line tangent from the point $(k_i, g(k_i))$ to the quadratic x^2. The constraint says that if the tangency point occurs within the interval (k_i, k_{i+1}), then the line segment connecting $g(k_i)$ and $g(k_{i+1})$ has to lie below the tangent.

In order to express this algebraically, notice that the constraints at the points k_i : $g(k_i) \le k_i^2$, can be formulated by denoting $d_i^2 = k_i^2 - g(k_i) \ge 0$, $i = 0, \ldots, n+1$, with $d_i \ge 0$. Then the x-coordinate of the tangency point equals $k_i + d_i$, hence the slope of the tangent is $2(k_i + d_i)$. The constraint for the interval (k_i, k_{i+1}) can be expressed as follows:

$$\text{if } k_i + d_i \le k_{i+1}\,, \text{ then } k_{i+1}^2 - d_{i+1}^2 \le k_i^2 - d_i^2 + 2(k_i + d_i)(k_{i+1} - k_i).$$

The last inequality can be written as $d_{i+1}^2 \ge (d_i - (k_{i+1} - k_i))^2$. By definition $d_i \ge 0$, $i = 0, \ldots, n+1$, so the constraints can be restated as $d_{i+1} \ge \max(0, k_{i+1} - k_i - d_i)$, $\forall i = 0, \ldots, n$.

In terms of d_i's, we can write:

$$u_i = \frac{(k_{i+1}^2 - d_{i+1}^2) - (k_i^2 - d_i^2)}{k_{i+1} - k_i} + \frac{(k_i^2 - d_i^2) - (k_{i-1}^2 - d_{i-1}^2)}{k_i - k_{i-1}} \text{ for all } \; i = 1, \ldots, n\,,$$

$$u_0 = \frac{(k_1^2 - d_1^2) - (k_0^2 - d_0^2)}{k_1 - k_0} \;\text{ and }\; v = k_0^2 - d_0^2 = -d_0^2\,.$$

By regrouping the terms in the objective, we can write the dual problem in terms of the

d_i's as follows:

$$V^-(M) + M^2 = \text{maximize} \quad \sum_{i=0}^{n+1} (k_i^2 - d_i^2)(T_{i-1} - T_i)$$

$$\text{subject to} \quad d_i \geq \max(0, k_i - k_{i-1} - d_{i-1}) \quad , \; i = 1, \ldots, n+1,$$

where $T_i = -\dfrac{q_{i+1} - q_i}{k_{i+1} - k_i}$, $i = 0, \ldots, n$, $T_{-1} = 1$, and $T_{n+1} = 0$. The optimal bound $V^-(M)$ can thus be rewritten as

$$V^-(M) = \sum_{i=0}^{n+1} t_i k_i^2 - Z(M) - M^2,$$

where $Z(M)$ is the objective value of the following quadratic network flow problem:

$$\begin{aligned} Z(M) = \text{minimize} \quad & \sum_{i=0}^{n+1} t_i \, d_i^2 \\ \text{subject to} \quad & d_i + d_{i-1} \geq c_i \quad , \; i = 1, \ldots, n+1 \\ & d_i \geq 0 \qquad\qquad , \; i = 0, \ldots, n+1, \end{aligned}$$

where we denoted $t_i = T_{i-1} - T_i \geq 0$, $c_i = k_i - k_{i-1} \geq 0$.

(b) In order to compute the overall optimal lower bound on the variance, independent from the mean, it remains to minimize $V^-(M)$ over all feasible values $M \in [M^-, M^+]$, where M^-, M^+ are given from the bounds in Eq. (21). □

Problem (26) is a separable quadratic optimization problem over network flow constraints. Because of its special structure it can be solved by the following dynamic programming algorithm:

Choose d_0 and let:

$$d_{i+1} = \begin{cases} k_{i+1} - k_i - d_i, & \text{if } k_{i+1} \geq k_i + d_i, \\ 0, & \text{if } k_{i+1} < k_i + d_i. \end{cases}$$

To obtain the optimal solution, one has to optimize over all initial choices of $d_0 \geq 0$. A heuristic solution, that performs very well in practice, starts with $d_0 = 0$.

Geometrically, this iterative construction can be visualized (see Figure 3) as follows: at each step $(i+1)$, from the point $(k_i, g(k_i))$ draw the positive slope tangent to the curve $f(x) = x^2$. The x-coordinate of the tangency point equals $k_i + d_i$, and according to whether or not this falls within the next interval $[k_i, k_{i+1}]$, we have two cases:

- If $k_i + d_i \leq k_{i+1}$, then draw the next segment of $g(x)$, $x \in [k_i, k_{i+1}]$ to be the tangent.

- If $k_i + d_i > k_{i+1}$, then draw the next segment of $g(x)$, $x \in [k_i, k_{i+1}]$ so that $g(k_{i+1}) = k_{i+1}^2$.

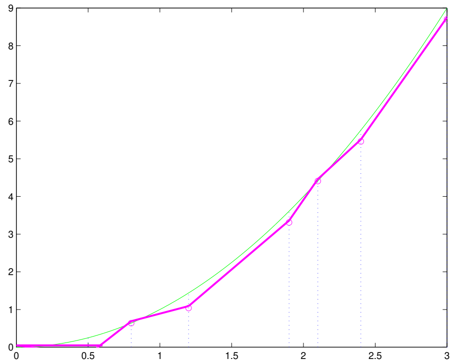

Figure 3: The dynamic programming algorithm.

Mixed Put And Call Options. In the case when both call and put prices are given, we first transform all information in terms of calls, and find the best possible bounds M^-, M^+ on the mean M, using Theorem 4. In order to find a optimal lower bound on the variance, we solve again the Problem in Theorem 6(b).

4.4 Computational Results.

In this section, we discuss the quality of the upper and lower bounds on the mean and variance of a stock price. In Table 1, we report results on the January '99 Microsoft stock price, computed using information on European call prices from The Wall Street Journal of

July 7, 1998. The current stock price is $S = 107\frac{13}{16}$ and the listed call options have strikes $k = [80, 95, 100, 110, 120, 140]$, and are sold at closing for $q = [31, 19, 16, 10, 6, 2.25]$. We also incorporate in our calculations the listed risk-free interest rate, listed as $r = .0557$. Call prices are given by:

$$q(k) = e^{-r(T-t)}E[\max(0, X - k)],$$

where $T - t$ is the time to maturity, measured in years. In this case $T - t = 0.5$.

Using the Black-Scholes option pricing formula, we estimate from the data an implied volatility of $s_{BS} = 0.3241$. The corresponding estimates for the mean and standard deviation of the forward stock price, under the risk neutral valuation, are $M_{BS} = S\,e^{r(T-t)} = 110.8573$ and $\sigma_{BS} = Se^{r(T-t)}\sqrt{e^{S^2(T-t)} - 1} = 26.7394$.

Using $M = 110.8573$ and assuming an upper bound $K = 160$ on the stock price, we apply the upper bound given in Eq. (23) and the lower bound given in Eq. (25) to obtain that the standard deviation of the stock price σ belongs in the interval: $\sigma \in (26.6677, 26.9851)$. If from the standard deviation we were to compute the implied volatility, as implied by the Black-Scholes formula, we would obtain $s \in (0.3223, 0.3259)$, which indeed is very close to the direct Black-Scholes forecasts.

If we do not use any information from the Black and Scholes model, but we only apply the bounds on the mean given by Eqs. (21) we obtain that the average stock price M is in the interval $M \in [M^-, M^+]$ with $M^- = 97.6829$, and $M^+ = 111.8755$. In Table 1, we vary the mean M in the interval [97.6829, 111.8755], and report the corresponding interval $[\sigma^-, \sigma^+]$ of the standard deviation of the stock price. For a given M, we observe that the bounds we derive on the standard deviation are extremely tight. As M varies in the interval [97.6829, 111.8755], we obtain that the standard deviation is within $23.7660 \leq \sigma \leq 49.2101$.

5 Bounds With Transaction Costs.

Up until now we have assumed a frictionless economy, and developed our results based on the theory of asset pricing under the no-arbitrage assumption, ignoring transaction costs. In this section, we derive bounds in the presence of transaction costs, using the no-arbitrage

M	σ^-	σ^+
97.6829	49.1709	49.2101
98.1723	48.5859	48.6306
98.6617	47.9888	48.0391
99.1511	47.3790	47.4352
99.6405	46.7563	46.8184
100.1299	46.1199	46.1882
100.6193	45.4693	45.5440
101.1087	44.8040	44.8852
101.5981	44.1232	44.2111
102.0875	43.4262	43.5212
102.5769	42.7122	42.8145
103.0663	41.9804	42.0903
103.5557	41.2298	41.3476
104.0451	40.4594	40.5854
104.5345	39.6679	39.8027
105.0239	38.8542	38.9980
105.5133	38.0168	38.1702
106.0027	37.1540	37.3175
106.4921	36.2642	36.4384
106.9815	35.3451	35.5307
107.4709	34.3946	34.5924
107.9603	33.4099	33.6207
108.4497	32.3878	32.6128
108.9391	31.3248	31.5651
109.4285	30.2165	30.4736
109.9179	29.0577	29.3333
110.4073	27.8422	28.1384
110.8967	26.5621	26.8815
111.3861	25.2076	25.5535
111.8755	23.7660	24.1427

Table 1: Optimal bounds on the standard deviation for various values of the mean M given option prices.

assumption. When transaction costs are taken into account the price of an option is within an interval defined by the bid-ask spread. A call pricing function is then defined as a pair: $q_{tc} : R_+ \to R_+^2,\ q_{tc}(k) = (q_{bid}(k) = q^-(k), q_{ask}(k) = q^+(k))$.

In a frictionless market, the asset pricing theory of Harrison and Kreps [10] insures the existence of a risk-neutral martingale measure that uniquely determines a valid linear pricing rule for all assets. Suppose we are given n call options with strikes k_i, and bid-ask prices $q_i^-, q_i^+,\ i = 1, \ldots, n$. In the presence of transaction costs, Jouini and Kallal [13] show that there is no arbitrage if and only if there exists a probability measure π such that

$$q_i^- \leq E_\pi[\max(0, X - k_i)] \leq q_i^+,\ i = 1, \ldots, n.$$

By Theorem 4, this is equivalent to the existence of a convex decreasing function $q^* : R_+ \to R_+$, such that $q^*(k_i) \in [q_i^-, q_i^+]$, for all $i = 1, \ldots, n$. If no convex decreasing function can be fitted between the bid (q^-) and ask (q^+) processes, then the given set of bid-ask spreads is not valid, and an arbitrage opportunity exists. This provides an easy test for arbitrage opportunities in a market with transaction costs.

The next theorem extends the results from the two previous sections by replacing the equality constraints in each respective primal problem by:

$$q_i^- \leq \int_{k_i}^{\infty} (x - k_i)\pi(x)dx \leq q_i^+,\ i = 1, \ldots, n\ .$$

We thus introduce corresponding dual variables u_i^-, u_i^+, which are non-negative for upper bound problems and non-positive for the lower bounds. The corresponding dual function becomes

$$g(x) = v + \sum_{i \mid k_i \leq x} (x - k_i)u_i,$$

where $u_i = u_i^+ - u_i^-$. With the notation $q_i = q_i^+ - q_i^- (\geq 0)$, we can write (for all problems) the dual objective as:

$$v+\sum_{i=1}^{n}(q_i^+ u_i^+ - q_i^- u_i^-) = v+\sum_{i=1}^{n} q_i^+ u_i + \sum_{i=1}^{n} q_i u_i^- = v+\sum_{i=1}^{n} q_i^- u_i + \sum_{i=1}^{n} q_i u_i^+.$$

By optimizing the corresponding dual, and using very similar techniques as in Theorems 4, 5, and 6 we prove the following result.

Theorem 7 (Bounds under transaction costs) *Given bid and ask prices q_i^- and q_i^+ for European calls with strikes k_i, $i=1,\ldots,n$, then:*

(a) *The optimal bounds on a call with strike k is given by:*

$$\begin{aligned} q^-(k) &= \max\left(q_j^- \frac{k-k_{j-1}}{k_j-k_{j-1}} + q_{j-1}^- \frac{k_j-k}{k_j-k_{j-1}}\ ,\ q_{j+1}^- \frac{k_{j+2}-k}{k_{j+2}-k_{j+1}} + q_{j+2}^- \frac{k-k_{j+1}}{k_{j+2}-k_{j+1}} \right) \\ q^+(k) &= q_j^+ \frac{k_{j+1}-k}{k_{j+1}-k_j} + q_{j+1}^+ \frac{k-k_j}{k_{j+1}-k_j} \end{aligned}$$

(b) *The optimal bounds on the mean stock price are:*

$$\begin{aligned} M^- &= \frac{q_1^- k_2 - q_2^- k_1}{k_2-k_1} \\ M^+ &= q_1^+ + k_1 \end{aligned}$$

(c) *The optimal lower bound on the variance is:*

$$V^-(M) = \sum_{i=0}^{n+1} t_i k_i^2 - Z(M) - M^2,$$

where $Z(M)$ is the objective value of the following quadratic network flow problem:

$$\begin{aligned} Z(M) = \text{minimize} \quad & \sum_{i=0}^{n+1} t_i\, d_i^2 \\ \textit{subject to} \quad & d_i + d_{i-1} \geq c_i, \quad i=1,\ldots,n+1 \\ & d_i \geq 0, \qquad\qquad i=0,\ldots,n+1 \end{aligned}$$

with $t_i = T_{i-1} - T_i \geq 0$, $T_i = -\dfrac{q^-_{i+1} - q^-_i}{k_{i+1} - k_i}$, $i = 0, \ldots, n$, *(*$T_{-1} = 1, T_{n+1} = 0$*), and* $c_i = k_i - k_{i-1} \geq 0$.

The optimal upper bound on the variance is:

$$V^+(M) = \sum_{i=1}^{n}(k_{i+1} - k_{i-1})q_i^+ + k_1 M - M^2.$$

6 Bounds in Multiple Dimensions.

In this section, we consider generalizations of the bounds we considered in earlier sections when we have information about a set of m different stocks. In particular, we have an option with payoff function $\phi(x)$, $\phi : R^m_+ \to R$, and a vector of n moment functions $f = (f_1, \ldots, f_n)$ (we let $f_0(x) = 1$), $f_i : R^m_+ \to R$, $i = 0, 1, \ldots, n$, and the corresponding vector of moments $q = (q_1, \ldots, q_n)$ (we let $q_0 = 1$). We address in this section the upper bound problem (2):

$$\begin{array}{lll} \text{maximize} & E_\pi[\phi(X)] & \\ \text{subject to} & E_\pi[f_i(X)] = q_i\ ,\ i = 1, \ldots, n. & \\ & \displaystyle\int_0^\infty \pi(x)dx = 1 & \\ & \pi(x) \geq 0, & x \in R^m_+, \end{array} \tag{28}$$

where the expectation is taken over all martingale measures defined on R^m_+. We can solve the lower bound problem by changing the sign of the objective function ϕ in Problem (28).

In Theorem 9 we show that solving Problem (28) is NP-hard. For this reason, we find a weaker bound by optimizing over all martingale measures defined on R^m as opposed to

R^m_+. For this reason we consider the following problem:

$$
\begin{array}{ll}
\text{maximize} & E_\pi[\phi(X)] \\
\text{subject to} & E_\pi[f_i(X)] = q_i \ , \ i = 1, \ldots, n. \\
& \displaystyle\int_{-\infty}^{\infty} \pi(x)dx = 1 \\
& \pi(x) \geq 0, \qquad\qquad\qquad x \in R^m,
\end{array}
\tag{29}
$$

and its dual:

$$
\begin{array}{ll}
\text{minimize} & y_0 + \displaystyle\sum_{i=1}^n y_i q_i \\
\text{subject to} & y_0 + \sum_{i=1}^n y_i f_i(x) \geq \phi(x), \quad \forall x \in R^m.
\end{array}
\tag{30}
$$

Isii [12](see also Karlin [14], p.472, or Smith [22]) shows that under weak conditions[1] on the moment vector q implies that strong duality holds, i.e., the optimal solution values of Problems (29) and (30) are equal.

The best possible upper bound corresponds to the optimal solution value of Problem (28). Since Problem (29) is a relaxation of Problem (28), we obtain an upper bound, although not necessarily the optimal one, by solving Problem (29), and by strong duality, Problem (30). In the next theorem we identify cases under which we can solve Problem (30), efficiently.

Theorem 8 *An upper bound on Problem (28) can be solved in polynomial time in the following cases:*

(a) *If ϕ and f_i, $i = 1, \ldots, n$ are quadratic or linear functions of the form*

$$
\begin{aligned}
\phi(x) &= x'Ax + b'x + c \\
f_i(x) &= x'A_ix + b_i'x + c_i, \qquad i = 1, \ldots, n.
\end{aligned}
\tag{31}
$$

then Problem (30), and thus Problem (29), can be solved in polynomial time by solving

[1] An example of such a condition is as follows: If the vector of moments q is interior to the feasible moment set $\mathcal{M} = \{E[f(X)] \mid X$ arbitrary multivariate distribution $\}$, then strong duality holds.

the following semidefinite optimization problem:

$$
\begin{array}{ll}
\text{minimize} & \sum_{i=1}^{n} y_i q_i \\
\textit{subject to} & \begin{bmatrix} \sum_{i=1}^{n} y_i c_i + y_0 - c & \left(\sum_{i=1}^{n} y_i b_i - b\right)'/2 \\ \left(\sum_{i=1}^{n} y_i b_i - b\right)/2 & \sum_{i=1}^{n} y_i A_i - A \end{bmatrix} \succeq 0.
\end{array}
\tag{32}
$$

(b) *If ϕ and f_i, $i = 1, \ldots, n$, are quadratic or piecewise linear functions of the form*

$$
\begin{aligned}
\phi(x) &= x'Ax + b_k'x + c_k \qquad x \in D_k, \quad k = 1, \ldots, d, \\
f_i(x) &= x'A_ix + b_{ik}'x + c_{ik}, \quad x \in D_k, \quad i = 1, \ldots, n, \ k = 1, \ldots, d,
\end{aligned}
\tag{33}
$$

over the d disjoint polyhedra $D_1, \ldots, D_d$ that form a partition of R^m, and d is a polynomial in n, m, then Problem (30), and thus Problem (29), can be solved in polynomial time.

Proof:

(a) We consider first the case when all the functions ϕ and f_i are are quadratic or linear as in Eq. (31). In this case, Problem (30) becomes:

$$
\begin{array}{ll}
\text{minimize} & y_0 + \sum_{i=1}^{n} y_i q_i \\
\text{subject to} & g(x) \geq 0, \quad \forall x \in R^m,
\end{array}
$$

where

$$
g(x) = y_0 + \sum_{i=1}^{n} y_i f_i(x) - \phi(x) = x'\hat{A}x + \hat{b}'x + \hat{c},
$$

with

$$
\hat{A} = \sum_{i=1}^{n} y_i A_i - A, \qquad \hat{b} = \sum_{i=1}^{n} y_i b_i - b, \qquad \hat{c} = \sum_{i=1}^{n} y_i c_i + y_0 - c.
$$

Thus, the constraints $g(x) \geq 0$ are equivalent to

$$x'\hat{A}x + \hat{b}'x + \hat{c} \geq 0, \quad \forall x \in R^m,$$

or equivalently

$$\begin{pmatrix} 1 \\ x \end{pmatrix}' \begin{bmatrix} \hat{c} & \hat{b}/2 \\ \hat{b}/2 & \hat{A} \end{bmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix} \geq 0, \quad \forall x \in R^m. \tag{34}$$

Eq. (34) holds if and only if

$$\begin{bmatrix} \hat{c} & \hat{b}'/2 \\ \hat{b}/2 & \hat{A} \end{bmatrix} \succeq 0,$$

i.e., the matrix $\begin{bmatrix} \hat{c} & \hat{b}'/2 \\ \hat{b}/2 & \hat{A} \end{bmatrix}$ is positive semidefinte. Thus, Problem (30) is equivalent to the semidefinite optimization problem (32), which is solvable in polynomial time (see for example Nesterov and Nemirovski [18] and Vandenberghe and Boyd [23]).

(b) If the functions ϕ or $f_i, i = 1, \ldots, n$ are as in (33), then Problem (30) can be expressed as

$$\begin{aligned} \text{minimize} \quad & y_0 + \sum_{i=1}^{n} y_i q_i \\ \text{subject to} \quad & g_k(x) = x'\hat{A}x + \hat{b}_k'x + \hat{c}_k \geq 0, \qquad \forall\ x \in D_k,\ k = 1, \ldots, d, \end{aligned} \tag{35}$$

where

$$\hat{A} = \sum_{i=1}^{n} y_i A_i - A, \qquad \hat{b}_k = \sum_{i=1}^{n} y_i b_{ik} - b_k, \qquad \hat{c}_k = \sum_{i=1}^{n} y_i c_{ik} + y_0 - c_k.$$

By the equivalence of separation and optimization (see Grötschel, Lovász and Schrijver [8]), Problem (35) can be solved in polynomial time if and only if the following separation problem can be solved in polynomial time.

The Separation Problem:

Given an arbitrary $y = (y_0, y_1, \ldots, y_n)$, check whether $g_k(x) \geq 0$, for all $x \in D_k$, $k = 1, \ldots, n$

and if not, find a violated inequality.

We show next that solving the separation problem reduces to checking whether the matrix $\hat{A}$ is positive semidefinite, and in this case solving the convex quadratic problems

$$\min_{x \in D_k} g_k(x), \qquad k = 1, \ldots, d.$$

This can be done in polynomial time using ellipsoid algorithm (see Grötschel, Lovász and Schrijver [8]). The following algorithm solves the separation problem in polynomial time:

Algorithm A:

1. If $\hat{A}$ is not positive semidefinite, we construct a vector x_0 so that $g_k(x_0) < 0$ for some $k = 1, \ldots, n$. We decompose $\hat{A} = Q'\Lambda Q$, where $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ is the diagonal matrix of eigenvalues of $\hat{A}$. Let $\lambda_i < 0$ be a negative eigenvalue of $\hat{A}$. Let u be a vector with $u_j = 0$, for all $j \neq i$, and u_i selected as follows: Let v_k be the largest root of each polynomial if it exists. Let $u_i = \max_k v_k + 1$. If all the polynomials do not have real roots, then u_i can be chosen arbitrarily. Then

$$\lambda_i u_i^2 + (Q\hat{b}_k)_i u_i + \hat{c}_k < 0, \qquad \forall\ k = 1, \ldots, d.$$

Let $x_0 = Q'u$. Since the polyhedra D_k form a partition of R^m, then $x_0 \in D_{k_0}$ for some k_0. Then,

$$\begin{aligned} g_{k_0}(x_0) &= x_0'\hat{A}x_0 + \hat{b}_{k_0}' x_0 + \hat{c}_{k_0} \\ &= u'QQ'\Lambda QQ'u + \hat{b}_{k_0}' Q'u + \hat{c}_{k_0} \\ &= u'\Lambda u + (Q\hat{b}_{k_0})'u + \hat{c}_{k_0} \\ &= \sum_{j=1}^{n} \lambda_j u_j^2 + \sum_{j=1}^{n} (Q\hat{b}_{k_0})_j u_j + \hat{c}_{k_0} \\ &= \lambda_i u_i^2 + (Q\hat{b}_{k_0})_i u_i + \hat{c}_{k_0} < 0. \end{aligned}$$

This produces a violated inequality.

2. Otherwise, if $\hat{A}$ is positive semidefinite, then we test if $g_k(x) \geq 0$, $\forall\ x \in D_k$ by solving

d convex quadratic optimization problems:

$$\min_{x \in D_k} x'\hat{A}x + \hat{b}_k'x + \hat{c}_k\,, \quad \text{for} \quad k = 1, \ldots, d. \tag{36}$$

We denote by x_k^* an optimal solution of Problem (36), and $z_k = g_k(x_k^*)$ the optimal value of Problem (36). If $z_k \geq 0$ for all $k = 1, \ldots, d$, then there is no violated inequality. Otherwise, if $z_{k_0} < 0$ for some k_0, then we find x_0^k such that $g(x_k^*) < 0$, which represents a violated inequality.

Thus, Algorithm A solves the separation problem in polynomial time. It follows that Problem (30), and hence Problem (29), can be solved in polynomial time. □

6.1 Examples.

Suppose we have observed the price q_1 of a European call option with strike k_1 for stock 1, and the price q_2 of a European call option with strike k_2 for stock 2. In addition, we have estimated the means μ_1, μ_2, the variances σ_1^2, σ_2^2 and the covariance σ_{12} of the prices of the two underlying stocks. Suppose, in addition, we are interested in obtaining an upper bound on the price of a European call option with strike k for stock 1. Intuition suggests that since the prices of the two stocks are correlated, the price of a call option on stock 1 with strike k might be affected by the available information regarding stock 2. We can find an upper bound on the price of a call option on stock 1 with strike k, by solving the problem we formulated in (1), which is a special case of Problem (28), with $m = 2$, $n = 7$. From Theorem 8(b), Problem (1) can be solved efficiently. In this case, there are six sets D_k as follows:

$$\begin{aligned}
D_1 &= \{(x_1, x_2) \mid x_1 \geq k, x_2 \geq k_2\}, \\
D_2 &= \{(x_1, x_2) \mid x_1 \geq k, x_2 \leq k_2\}, \\
D_3 &= \{(x_1, x_2) \mid k_1 \leq x_1 \geq k, x_2 \geq k_2\}, \\
D_4 &= \{(x_1, x_2) \mid k_1 \leq x_1 \geq k, x_2 \leq k_2\}, \\
D_5 &= \{(x_1, x_2) \mid x_1 \leq k, x_2 \geq k_2\}, \\
D_6 &= \{(x_1, x_2) \mid x_1 \leq k, x_2 \leq k_2\}.
\end{aligned}$$

As another example, suppose we are interested to find an upper bound on the price of an option with payoff

$$\phi(x) = \max(0, a_1'x - k_1, a_2'x - k_2).$$

This option allows its holder to buy at maturity two stock indices: the first one (given by the vector a_1) at price k_1, and the second one (given by the vector a_2) at price k_2. Suppose we have estimated the mean and covariance matrix of the underlying securities. Again, Theorem 8(b) applies. In this case there are three sets D_k that form a polyhedral partition of R^m:

$$\begin{aligned} D_1 &= \{x \in R^m \mid a_1'x - k_1 \leq 0, a_2'x - k_2 \leq 0\}, \\ D_2 &= \{x \in R^m \mid 0 \leq a_1'x - k_1, a_2'x - k_2 \leq a_1'x - k_1\}, \\ D_3 &= \{x \in R^m \mid 0 \leq a_2'x - k_1, a_1'x - k_2 \leq a_2'x - k_1\}. \end{aligned}$$

Note that if $x \in D_1$, $\phi(x) = 0$, while if $x \in D_2$, $\phi(x) = a_1'x - k_1$. Finally, if $x \in D_3$, $\phi(x) = a_2'x - k_2$.

6.2 The Complexity of Optimal Bounds.

Theorem 8 provides optimal bounds in polynomial time if we optimize over R^m, but not over R^m_+. The next theorem shows that it is NP-hard to find optimal bounds over R^m_+.

Theorem 9 (Complexity of finding optimal bounds) *The problem of finding the optimal bound*

$$\max_{X \sim (M,\Gamma)^+} E[\phi(X)] \tag{37}$$

is NP-hard even if $\phi(x) = f'x$.

Proof:

The dual of Problem (37) is

$$\begin{aligned} \text{minimize} \quad & y_0 + \sum_{i=1}^{n} y_i \mu_i + \sum_{i=1}^{n} \sum_{j=1}^{i} y_{ij} (\sigma_{ij} + \mu_i \mu_j) \\ \text{subject to} \quad & x' Y x + y' x + y_0 \geq f' x, \ \forall \ x \geq 0, \end{aligned}$$

and the corresponding separation problems becomes:

Separation problem:

Given (Y, y, y_0), check if $\min\limits_{x \geq 0} \ x'Yx + y'x + y_0 - f'x \geq 0$, otherwise find a violated inequality.

The separation problem is NP-hard, as it reduces to verifying that the matrix Y is co-positive, which is an NP-hard problem (see Murty and Kabadi [17]). Therefore, by the equivalence of separation and optimization (see Grötschel, Lovász and Schrijver [8]), it is NP-hard to solve Problem (37). □

7 Concluding Remarks.

We have demonstrated that convex optimization is the natural way to address the relation between option and stock prices without making distributional assumptions for the underlying price dynamics, but only using the no-arbitrage assumption. For the single stock problem, we have shown that we can find optimal bounds on option prices efficiently, either algorithmically (solving a semidefinite optimization problem) or in closed form. For options that are affected by multiple stocks either directly (the payoff of the option depends on multiple stocks) or indirectly (we have information on correlations between stock prices), we can find bounds (but not optimal ones) using convex optimization methods. However, it is NP-hard to find optimal bounds in multiple dimensions.

8 Appendix: Proof of Proposition 1

(a) Suppose (5) holds. Let $e_x = (1, x, x^2, \ldots, x^k)'$. Then

$$\begin{aligned} g(x) &= \sum_{r=0}^{2k} \sum_{i+j=r} x_{ij} x^r \\ &= \sum_{i=0}^{k} \sum_{j=0}^{k} x_{ij} x^i x^j \\ &= e_x' X e_x \\ &\geq 0, \end{aligned}$$

since $X \succeq 0$.

Conversely, suppose that the polynomial $g(x)$ of degree $2k$ is nonnegative for all x. Then, the real roots of $g(x)$ should have even multiplicity, otherwise $g(x)$ would alter its sign in a neighborhood of a root. Let λ_i, $i = 1, \ldots, r$ be its real roots with corresponding multiplicity $2m_i$. Its complex roots can be arranged in conjugate pairs, $a_j + ib_j$, $a_j - ib_j$, $j = 1, \ldots, h$. Then,

$$g(x) = y_{2k} \prod_{i=1}^{r} (x - \lambda_i)^{2m_i} \prod_{j=1}^{h} \left((x - a_j)^2 + b_j^2 \right).$$

Note that the leading coefficient y_{2k} needs to be positive. Thus, by expanding the terms in the products, we see that $g(x)$ can be written as a sum of squares of polynomials, of the form

$$\begin{aligned} g(x) &= \sum_{i=0}^{k} \left(\sum_{j=0}^{k} x_{ij} x^j \right)^2 \\ &= e_x' X e_x, \end{aligned}$$

with X positive semidefinite, from where Equation (5) follows.

(b) We observe that $g(x) \geq 0$ for $x \in [0, a]$ if and only if

$$(1 + t^2)^k g\left(\frac{at^2}{1 + t^2} \right) \geq 0, \qquad \text{for all } t.$$

Since

$$
\begin{aligned}
(1+t^2)^k g\left(\frac{at^2}{1+t^2}\right) &= \sum_{r=0}^{k} y_r a^r t^{2r}(1+t^2)^{k-r} \\
&= \sum_{r=0}^{k} y_r a^r \sum_{l=0}^{k-r} \binom{k-r}{l} t^{2(l+r)} \\
&= \sum_{j=0}^{k} t^{2j} \left(\sum_{r=0}^{j} y_r \binom{k-r}{j-r} a^r\right),
\end{aligned}
$$

by applying part (a) we obtain (6).

(c) We observe that $g(x) \geq 0$ for $x \in [a, \infty)$ if and only if

$$
g(a(1+t^2)) \geq 0, \qquad \text{for all } t.
$$

Since

$$
\begin{aligned}
g(a(1+t^2)) &= \sum_{r=0}^{k} y_r a^r (1+t^2)^r \\
&= \sum_{r=0}^{k} y_r a^r \sum_{l=0}^{r} \binom{r}{l} t^{2l} \\
&= \sum_{l=0}^{k} t^{2l} \left(\sum_{r=l}^{k} y_r \binom{r}{l} a^r\right),
\end{aligned}
$$

by applying part (a) we obtain (7).

(d) We observe that $g(x) \geq 0$ for $x \in [a, b]$ if and only if

$$
(1+t^2)^k g\left(a + (b-a)\frac{t^2}{1+t^2}\right) \geq 0, \qquad \text{for all } t.
$$

Since

$$
\begin{aligned}
(1+t^2)^k g\left(a+(b-a)\frac{t^2}{1+t^2}\right) &= \sum_{r=0}^{k} y_r (a+bt^2)^r (1+t^2)^{k-r} \\
&= \sum_{r=0}^{k} y_r \sum_{m=0}^{r} \binom{r}{m} a^{r-m} b^m t^{2m} \sum_{j=0}^{k-r} \binom{k-r}{j} t^{2j} \\
&= \sum_{l=0}^{k} t^{2l} \left(\sum_{m=0}^{l} \sum_{r=m}^{k+m-l} y_r \binom{r}{m} \binom{k-r}{l-m} a^{r-m} b^m \right),
\end{aligned}
$$

by applying part (a) we obtain (8). $\square$

Acknowledgements.

We would like to thank Andrew Lo, Stephen Ross, and Jiang Wang for insightful discussions, and pointers to the literature.

References

[1] A. Ben-Tal and A. Nemirovski. Convex optimization in engineering: Modeling, analysis, algorithms. Research Report, Technion, Israel, 1998.

[2] D. Bertsimas and I. Popescu. Optimal inequalities in probability: A convex programming approach. Working paper, Operations Research Center, MIT, 1999.

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

[4] J. Cox and S. Ross. The valuation of options for alternative stochastic processes. *Journal of Financial Economics*, 3:145–166, 1976.

[5] E. Derman and I. Kani. Riding on the smile. *RISK*, 7:32–39, 1994.

[6] B. Dupire. Pricing with a smile. *RISK*, 7:18–20, 1994.

[7] K. Fujisawa, M. Kojima, and K. Nakata. Sdpa (semidefinite programming algorithm) user's manual, version 4.10. Research Report on Mathematical and Computing Sciences, Tokyo Institute of Technology, Tokyo, Japan, 1998.

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

[9] B. Grundy. Option prices and the underlying asset's return distribution. *Journal of Finance*, 46(3):1045–1070, July 1991.

[10] M. Harrison and D. Kreps. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20:381–408, 1979.

[11] K. Isii. The extrema of probability determined by generalized moments (i) bounded random variables. *Ann. Inst. Stat. Math.*, 12:119–133, 1960.

[12] K. Isii. On the sharpness of Chebyshev-type inequalities. *Ann. Inst. Stat. Math.*, 14:185–197, 1963.

[13] E. Jouini and H. Kallal. Martingales and arbitrage in securities markets with transaction costs. *Journal of Economic Theory*, 66:178–197, 1995.

[14] S. Karlin and W.J. Studden. *Tchebycheff Systems: with Applications in Analysis and Statistics*. Pure and Applied Mathematics, A Series of Texts and Monographs. Interscience Publishers, John Wiley and Sons, 1966.

[15] A. Lo. Semiparametric upper bounds for option prices and expected payoffs. *Journal of Financial Economics*, 19:373–388, 1987.

[16] F. Longstaff. Martingale restriction tests of option pricing models, version 1. Working paper, University of California, Los Angeles, 1990.

[17] K. G. Murty and S. N. Kabadi. Some np-complete problems in quadratic and nonlinear programming. *Mathematical Programming*, 39:117–129, 1987.

[18] Y. Nesterov and A. Nemirovskii. Interior point polynomial algorithms for convex programming. *Studies in Applied Mathematics*, 13, 1994.

[19] S. Ross. Options and efficiency. *Quarterly Journal of Economics*, 90:75–89, 1976.

[20] M. Rubinstein. Implied binomial trees. *Journal of Finance*, 49(3):771–819, 1994.

[21] H. Scarf. A min-max solution of an inventory problem. In *Studies in the mathematical theory of inventory and production*, pages 201–209, Stanford University Press, Stanford, CA, 1958. K.J. Arrow, S. Karlin and H. Scarf, eds.

[22] J. Smith. Generalized chebyshev inequalities: Theory and applications in decision analysis. *Operations Research*, 43(5):807–825, 1995.

[23] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38:49–95, 1996.