INSEAD
R & D

OPTIMAL INEQUALITIES IN PROBABILITY THEORY: A CONVEX OPTIMIZATION APPROACH

by

D. BERTSIMAS*
and
I. POPESCU**

2000/62/TM

* Boeing Professor of Operations Research, Sloan School of Management, Rm. E53-363, Massachusetts Institute of Technology, Cambridge, Mass. 02139, USA.

** Assistant Professor of Decision Sciences at INSEAD, Boulevard de Constance, 77305 Fontainebleau Cedex, France.

A working paper in the INSEAD Working Paper Series is intended as a means whereby a faculty researcher's thoughts and findings may be communicated to interested readers. The paper should be considered preliminary in nature and may require revision.

Printed at INSEAD, Fontainebleau, France.

Optimal Inequalities in Probability Theory: A Convex Optimization Approach

Dimitris Bertsimas * and Ioana Popescu †

September, 2000

Abstract

We characterize the complexity of deriving tight moment inequalities for $P(X \in S)$, for an arbitrary set S and a multivariate random variable X that has a given collection of moments. For the univariate case, we provide an optimal inequality for $P(X \in S)$, when the first k moments of X are known, as a solution of a semidefinite optimization problem in $k+1$ dimensions. For the n-variate case, given only the means of X, we show that we can find tight bounds by solving n convex optimization problems when the set S is convex, and we provide a polynomial time algorithm when S is a union of convex sets. When moments up to second order are given, the work of Marshal and Olkin [36] shows that we can find tight bounds by solving a single convex optimization problems if the set S is convex and $\Omega = R^n$. We finally show that it is NP-hard to find tight bounds when moments of third or higher order are given, or if moments of second order are given and the domain of X is R^n_+.

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

†Assistant Professor of Decision Sciences, INSEAD, Fontainebleau Cedex 77305, FRANCE. ioana.popescu@insead.fr.

1 Introduction.

The problem of deriving bounds on the probability that a certain random variable belongs in a set, given information on some of the moments of this random variable, has a very rich and interesting history, which is very much connected with the development of probability theory in the twentieth century. The inequalities due to Markov, Chebyshev and Chernoff are some of the classical and widely used results of modern probability theory. Natural questions, however, that arise are:

1. *Are such bounds "best possible", i.e., do there exist distributions that match them?*
2. *Can such bounds be generalized in multivariate settings, and in what circumstances can they be explicitly and/or algorithmically computed ?*
3. *Is there a general theory based on optimization methods to address in a unified manner moment-inequality problems in probability theory?*

In order to answer these questions we first define the notion of a *feasible moment sequence.*

Definition 1 *A sequence* $\bar{\sigma} : (\sigma_{k_1 \cdots k_n})_{k_1+\cdots+k_n \leq k}$ *is a feasible* (n, k, Ω)*-moment vector (or sequence), if there is a random variable* $X = (X_1, \ldots, X_n)$ *with domain* $\Omega \subseteq R^n$*, whose moments are given by* $\bar{\sigma}$*, that is* $\sigma_{k_1 \cdots k_n} = E\big[X_1^{k_1} \cdots X_n^{k_n}\big]$, $\forall\, k_1 + \cdots + k_n \leq k$ *. We say that any such random variable* X *has a* $\bar{\sigma}$*-feasible distribution and denote this as* $X \sim \bar{\sigma}$*.*

We denote by $\mathcal{M} = \mathcal{M}(n, k, \Omega)$ the set of feasible (n, k, Ω)-moment vectors. For the univariate case ($n = 1$), the problem of deciding if $\bar{\sigma} = (M_1, M_2, \ldots, M_k)$ is a feasible $(1, k, \Omega)$-moment vector is the classical moment problem. This problem has been completely characterized by necessary and sufficient conditions by Stieltjes [56], [57] in 1894-95, who adopts the "moment" terminology from mechanics (see also Karlin and Shapley [28], Akhiezer [1], Siu, Sengupta & Lind [51] and Kemperman [31]). For univariate, nonnegative random variables ($\Omega = R_+$), these conditions can be expressed by the semidefiniteness of the following matrices:

$$R_{2n} = \begin{pmatrix} 1 & M_1 & \ldots & M_n \\ M_1 & M_2 & \ldots & M_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ M_n & M_{n+1} & \ldots & M_{2n} \end{pmatrix} \succeq 0,$$

$$R_{2n+1} = \begin{pmatrix} M_1 & M_2 & \dots & M_{n+1} \\ M_2 & M_3 & \dots & M_{n+2} \\ \vdots & \vdots & \ddots & \vdots \\ M_{n+1} & M_{n+2} & \dots & M_{2n+1} \end{pmatrix} \succeq 0.$$

For univariate random variables with $\Omega = R$, the necessary and sufficient condition given by Hamburger [19], [20] in 1920-21 for a vector $\bar{\sigma} = (M_1, M_2, \dots, M_k)$ to be a feasible $(1, k, R)$-moment sequence is that $R_{2\lfloor \frac{k}{2} \rfloor} \succeq 0$. In the multivariate case, the formulation of the problem can be traced back to Haviland [21], [22] in 1935-36 (see also Godwin [17]). To date, the sufficiency part of the moment problem has not been completely resolved in the multivariate case.

Suppose that $\bar{\sigma}$ is a feasible moment sequence and X has a $\bar{\sigma}$-feasible distribution. We now define the central problem that this paper addresses:

The $(\mathbf{n}, \mathbf{k}, \Omega)$-Bound Problem.

Given a sequence $\bar{\sigma}$ of up to kth order moments

$$\sigma_{k_1 k_2 \cdots k_n} = E\big[X_1^{k_1} X_2^{k_2} \cdots X_n^{k_n}\big], \qquad k_1 + k_2 + \cdots + k_n \leq k,$$

of a multivariate random variable $X = (X_1, X_2, \dots, X_n)$ on $\Omega \subseteq R^n$, find the "best possible" or "tight" upper and lower bounds on $P(X \in S)$, for arbitrary events $S \subseteq \Omega$.

The term "best possible" or "tight" upper (and by analogy lower) bound above is defined as follows.

Definition 2 *We say that α is a tight upper bound on $P(X \in S)$ if $\alpha = \sup_{X \sim \bar{\sigma}} P(X \in S)$.*

Note that a bound can be tight without necessarily being *exactly achievable* (i.e. there is a random variable $\tilde{X} \sim \bar{\sigma}$ for which $P(\tilde{X} \in S) = \alpha$), but only *asymptotically.*

The well known inequalities due to Markov, Chebyshev and Chernoff, which are widely used if we know the first moment, the first two moments, and all moments (i.e., the generating function) of a random variable, respectively, are feasible but not necessarily optimal solutions to the (n, k, Ω)-bound problem, i.e., they are not necessarily tight bounds.

Literature and Historical Perspective.

The history of the developments in the area of (n, k, Ω)-bound problems, sometimes referred to as Chebyshev type inequalities, can be traced back to the work of Gauss, Cauchy, Chebyshev, Markov

etc, and has witnessed an unexpected evolution. The problem of finding bounds on univariate distributions under moment constraints has actually been proposed and formulated without proof initially by Chebychev [10] in 1874 and resolved ten years later by his student Markov [34] in his PhD thesis, using continued fractions techniques. In the 1950s and 1960s there has been a revival of the interest in this area, that resulted in a large literature on the topic of generalized Chebyshev inequalities. Surveys of early literature can be found in Shohat and Tamarkin [53] and Godwin [16], [17].

The idea that optimization methods and duality theory can be used to address moment-type inequalities in probability first appeared in 1960, and is due independently and simultaneously to Isii [23] and Karlin (lecture notes at Stanford, see [29], p.472), who show that certain types of Chebyshev inequalities for univariate random variables are sharp, via strong duality results. Isii [24] extends these duality results for multivariate random variables on complete regular spaces.

Marshall and Olkin in 1961 [38] give a game theoretic proof of the sharpness of Chebyshev type inequalities with first and second order moment constraints, as well as with trigonometric moments. The same authors [36], [37] were the first to actually compute tight, explicit bounds on probabilities given first and second order moments (the $(n, 2, \Omega)$ problem in our context), thus generalizing Chebyshev's inequality to a multivariate setting. A detailed, unified account of the evolution of *Chebyshev Systems* is given by Karlin and Studden [29] in their 1966 monograph (see in particular Chapters 12 and 13 that deal with (n, k, Ω)-type bounds).

Not for the first time in its history, "the problem of moments lay dormant for more than 20 years."[1] It revives briefly in the 1980s, with the book on *Probability Inequalities and Multivariate Distributions* of Y.L. Tong [58] in 1980, who also publishes a monograph on probability inequalities in 1984. The latter notably contains, among others, a generalization of Markov's inequality for multivariate tails, due to Marshall [35], and an application of moment inequalities for computing error bounds in stochastic programming, by Birge and Wets [3]. A volume on *Moments in Mathematics* edited by Landau in 1987 includes a background survey by the same author [33], as well as relevant papers of Kemperman [32] and Diaconis [12].

Thirty two years after Isii's [24] original multivariate proof, Smith [55] rederived the same duality results and proposed new interesting applications in decision analysis, dynamic programming, statistics and finance. Bonnans and Shapiro [6] provide an in depth account of strong duality and sensitivity analysis for a general class of semi-infinite programming problems from a modern perspective (see section 5.4). Shapiro provides a rigorous discussion of necessary topological conditions for strong duality to hold, in that sense relaxing the compactness assumptions underlying Isii's proof.

[1] Shohat and Tamarkin [53], p.10.

Another line of research loosely connected to our research, is the work of Pitowski [42], [43] who makes use of duality results to prove general theorems in probability (weak and strong laws of large numbers, approximate central limit, the Linial- Nissan theorem etc.). The author uses different linear programming formulations to define and study geometric and complexity properties of correlation polytopes, which arise naturally in probability and logic. A similar type of problem is addressed by Bukscár [9] and Prekopa [48], who study probability bounds on finite unions of events by means of specialized boolean tree structures.

Prekopa uses a different mathematical programming approach to study probability inequalities given multivariate moments for discrete distributions [45] and programming with probabilistic constraints [46], [47] for the case of discrete distributions. He derives upper and lower bounds on moment constrained problems when the objective function obeys higher order convexity conditions, and presents applications to Bonferroni inequalities. He also investigates applications of moment-constrained problems to stochastic programming [49]. In fact, there has been a significant amount of work related to moment type problems in the context of stochastic programming with incomplete distributional information. Important contributions include Birge and Wets [3], [4], [5], and Dupačová and Prekopa [13]. The interested reader is referred also to Cipra [11], Kall [25], [26], [27], Ermoliev, Gaivoronski and Nedeva [14].

For a broader investigation of the optimization framework underlying this type of problems, we refer the interested reader to Borwein and Lewis [7], [8] who provide an in depth analysis of partially finite convex programming.

Despite its long and scattered history, the common belief among researchers is still that "the theory [of moment problems] is not up to the demands of applications" (Diaconis [12], p. 129). The same author suggests that one of the reasons could be the high complexity of the problem: "numerical determination ... is feasible for a small number of moments, but appears to be quite difficult in general cases". Another reason is identified by Kemperman ([32], p.20) as being the lack of a general algorithmic approach:

> "...a deep study of algorithms has been rare so far in the theory of moments, except for certain very specific practical applications, for instance, to crystallography, chemistry and tomography. No doubt, there is a considerable need for developing reasonably good numerical procedures for handling the great variety of moment problems which do arise in pure and applied mathematics and in the sciences in general...".

In an attempt to address Kemperman's criticism, Smith [55] actually introduced a computational procedure for the (n, k, R^n)-bound problem, although he does not refer to it in this way. Unfortunately,

the procedure is far from an actual algorithm, as there is no proof of convergence, and no investigation (theoretical or experimental) of its efficiency. It is fair to say that understanding of the algorithmic aspects and of the complexity of the (n, k, Ω)-bound problem is still lacking.

Yet a stronger criticism brought by Smith is the lack of simple, closed form solutions for the (n, k, R^n)-bound problem: "the bounds given by Chebychev's inequalities ... are quite loose. The more general versions are rarely used because of the lack of simple closed-form expressions for the bounds"([55], p.808).

Goals and Contributions.

The previous discussion motivates our desire in the present paper to understand the complexity of deriving tight moment inequalities, search for efficient algorithms in a general framework, and, when possible, derive simple closed-form bounds. In particular, we provide a full characterization of which (n, k, Ω)-bound problems are efficiently solvable and which are NP-hard.

More concretely, the contributions and structure of the paper are as follows:

1. In Section 3, we investigate in detail the univariate case, i.e., the $(1, k, \Omega)$-bound problem for $\Omega = R, R_+$. Perhaps the most important new contribution of the paper is that we show that optimal bounds can be computed efficiently by solving **a single semidefinite optimization problem**. We also derive optimal bounds for tail probability events in closed form when up to three moments are given. For $k = 1$ we recover the Markov inequality, which also shows that the Markov inequality is best possible. For $k = 2$ we recover a strict improvement of the Chebyshev inequality that retains the simplicity of the bound. This inequality dates back at least to Uspensky's book ([59], p.198) from 1937, who proposes it as an exercise. Despite its simplicity, the bound has been strangely ignored in the recent literature and textbooks. For $k = 3$ we derive new closed form tight bounds.

2. In Section 4, we address the $(n, 1, \Omega)$ problem. We show that if the set S in the definition of the $(n, 1, R^n)$-bound problem is convex, we find best possible bounds for $P(X \in S)$ explicitly as a solution of n convex optimization problems. We also provide a polynomial time algorithm to solve the $(n, 1, \Omega)$ problem, when the set S is the union of disjoint convex sets. These bounds represent natural extensions and improvements of Markov's inequality, and of the bounds derived by Marshall [35] for multivariate tails.

3. In Section 5, we review the work of Marshal and Olkin [36] who showed that the $(n, 2, R^n)$-bound problem can be solved as a solution of a single convex optimization problem.

4. In Section 6, we show that the $(n, 2, R^n_+)$, (n, k, R^n)-bound problems for $k \geq 3$ are NP-hard. This is in sharp contrast with the results established in prior sections, i.e., that the $(1, k, \Omega)$, $(n, 1, \Omega)$ and $(n, 2, R^n)$-bound problems are solvable in polynomial time as convex optimization problems.

2 Primal and Dual Formulations of The $(\mathbf{n}, \mathbf{k}, \Omega)$-Bound Problem.

In this section, we formulate the (n, k, Ω)-upper bound problem as an optimization problem, where Ω is the domain of the random variables we consider. We examine the corresponding dual problem and present weak and strong duality results that permit us to develop algorithms for the problem. The same approach and results apply to the (n, k, Ω)-lower bound problem.

The (n, k, Ω)-upper bound problem can be formulated as the following optimization problem (P):

$$\begin{aligned} (P) \quad Z_P = \text{maximize} \quad & \int_S f(\bar{z}) d\bar{z} \\ \text{subject to} \quad & \int_\Omega z_1^{k_1} \cdots z_n^{k_n} f(\bar{z}) d\bar{z} = \sigma_{k_1 \ldots k_n}, \quad \forall\, k_1 + \cdots + k_n \leq k, \\ & f(\bar{z}) = f(z_1, \ldots, z_n) \geq 0, \qquad \forall\, \bar{z} = (z_1, \ldots, z_n) \in \Omega. \end{aligned}$$

Notice that if Problem (P) is feasible, then $\bar{\sigma}$ is a feasible moment sequence, and any feasible distribution $f(\bar{z})$ is a $\bar{\sigma}$-feasible distribution. The feasibility problem is exactly the classical multidimensional moment problem.

In the spirit of linear programming duality theory, we associate a dual variable $u_{k_1 \ldots k_n}$ with each equality constraint of the primal. We can identify the vector of dual variables with a k-degree, n- variate dual polynomial:

$$g(x_1, \ldots, x_n) = \sum_{k_1 + \cdots + k_n \leq k} u_{k_1 \ldots k_n} x_1^{k_1} \cdots x_n^{k_n}.$$

The dual objective translates to finding the smallest value of:

$$\sum u_{k_1 \ldots k_n} \sigma_{k_1 \ldots k_n} = \sum u_{k_1 \ldots k_n} E[X_1^{k_1} \cdots X_n^{k_n}] = E[g(X)],$$

where the expected value is taken over any $\bar{\sigma}$-feasible distribution. In this framework, the Dual Problem (D) corresponding to Problem (P) can be written as:

$$
\begin{aligned}
(D) \quad Z_D = \text{minimize} \quad & E[g(X)] \\
\text{subject to} \quad & g(x) \text{ } k\text{-degree, } n\text{-variate polynomial,} \\
& g(x) \geq \chi_S(x), \ \forall\, x \in \Omega,
\end{aligned}
$$

where $\chi_S(x)$ is the indicator function of the set S, defined by:

$$
\chi_S(x) = \begin{cases} 1, & \text{if } x \in S, \\ 0, & \text{otherwise.} \end{cases}
$$

Notice that in general the optimum may not be achievable. Whenever the primal optimum is achieved, we call the corresponding distribution **an extremal distribution**. We next establish weak duality.

Theorem 1 (Weak duality) $Z_P \leq Z_D$.

Proof: Let $f(\bar{z})$ be a primal optimal solution and let $g(\bar{z})$ be any dual feasible solution. Then:

$$
Z_P = \int_S f(\bar{z}) d\bar{z} = \int_\Omega \chi_S(\bar{z}) f(\bar{z}) d\bar{z} \leq \int_\Omega g(\bar{z}) f(\bar{z}) d\bar{z} = E[g(X)],
$$

and hence $Z_P \leq \inf_{g(\cdot) \geq \chi_S(\cdot)} E[g(X)] = Z_D$. □

Theorem 1 indicates that by solving the Dual Problem (D) we obtain an upper bound on the primal objective and hence on the probability we are trying to bound. Under some mild restrictions on the moment vector $\bar{\sigma}$, the dual bound turns out to be tight. This strong duality result follows from a univariate result due to Karlin and Isii in 1960 (see Karlin and Studden [29], p.472), and generalized by Isii in 1963 [24] for multivariate random variables defined on complete regular spaces. Shapiro [52] provides a modern proof of duality under a simple topological structure. The following theorem is a consequence of their work, and holds for arbitrary distributions X over a metric space Ω endowed with its Borel sigma algebra:

Theorem 2 (Strong Duality and Complementary Slackness)
If the moment vector $\bar{\sigma}$ is an interior point of the set $\mathcal{M}$ of feasible moment vectors, then the following results hold:

1. *Strong Duality: $Z_P = Z_D$.*

2. *Complementary Slackness: If the dual is bounded, there exists a dual optimal solution* $g_{opt}(\cdot)$ *and a discrete extremal distribution concentrated on points* x*, where* $g_{opt}(x) = \chi_S(x)$*, that achieves the bound.*

It can also be shown that if the dual is unbounded, then the primal is infeasible, i.e., the multi-dimensional moment problem is infeasible. Moreover, if $\bar{\sigma}$ is a boundary point of $\mathcal{M}$, then it can be shown that the $\bar{\sigma}$-feasible distributions are concentrated on a subset Ω_0 of Ω, and strong duality holds provided we relax the dual to Ω_0 (see Isii [24], p.190 or Smith [55], p. 824). These authors also prove that it is equivalent to optimize only over discrete distributions that are concentrated on $m+2$ points, where m is the number of moment constraints (in our case $m = \frac{n(n+1)}{2}$).

In the univariate case, Isii [23] proves that if $\bar{\sigma}$ is a boundary point of $\mathcal{M}$, then exactly one $\bar{\sigma}$-feasible distribution exists. Kemperman [30] proves that for almost every $\bar{\sigma}$ such that the dual is finite, the dual solution is unique. In Chapter 5.4 of their recent book, Bonnans and Shapiro [6] provide necessary and sufficient conditions for the uniqueness of the optimal distribution, as well as sensitivity results.

If strong duality holds, then by optimizing over Problem (D) we obtain a **tight** bound on $P(X \in S)$. On the other hand, solving Problem (D) is equivalent to solving the corresponding separation problem, under certain technical conditions (see Grötschel, Lovász and Schrijver [18]).

3 The $(1, k, \Omega)$-Bound Problem.

In this section, we restrict our attention to univariate random variables. Given the first k moments $M_1, \ldots, M_k$ (we let $M_0 = 1$) of a real random variable X with domain Ω, we are interested in deriving tight bounds on $P(X \in S)$, that is we address the $(1, k, \Omega)$-bound problem. Our main result in this section is that optimal bounds can be derived as a solution to a single semidefinite optimization problem. We also derive closed form tight bounds when up to the first three momens are given.

3.1 Tight Bounds as Semidefinite Optimization Problems.

From Section 2, given the first k moments of X with domain Ω, we can find a tight bound for $P(X \in S)$ by solving the following problem

$$\begin{aligned} \text{minimize} \quad & \sum_{r=0}^{k} y_r M_r \\ \text{subject to} \quad & \sum_{r=0}^{k} y_r x^r \geq 1, \qquad \forall\ x \in S \\ & \sum_{r=0}^{k} y_r x^r \geq 0, \qquad \forall\ x \in \Omega. \end{aligned} \tag{1}$$

Since S and Ω are intervals in the real line we show in the next proposition that the feasible region of Problem (1) can be expressed using semidefinite constraints. Semidefinite optimization problems are efficiently solvable using interior point methods. For a review of semidefinite optimization see Vandenberghe and Boyd [60]. The results and the proofs in the following proposition are inspired by Ben-Tal and Nemirovski [2], p.140-142.

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*

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

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

$$\begin{aligned} 0 &= \sum_{i,j:\ i+j=2l-1} x_{ij}, \qquad l = 1, \ldots, k, \\ y_l &= \sum_{i,j:\ i+j=2l} x_{ij}, \qquad l = 0, \ldots, k, \\ & X \succeq 0. \end{aligned} \tag{3}$$

(c) *The polynomial* $g(x) = \sum_{r=0}^{k} y_r x^r$ *satisfies* $g(x) \geq 0$ *for all* $x \in [0, a]$ *if and only if there exists a*

positive semidefinite matrix $X = [x_{ij}]_{i,j=0,\ldots,k}$, *such that*

$$\begin{aligned} 0 &= \sum_{i,j:\ i+j=2l-1} x_{ij}, \qquad l = 1, \ldots, k, \\ \sum_{r=0}^{l} y_r \binom{k-r}{l-r} a^r &= \sum_{i,j:\ i+j=2l} x_{ij}, \qquad l = 0, \ldots, k, \\ X &\succeq 0. \end{aligned} \tag{4}$$

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

$$\begin{aligned} 0 &= \sum_{i,j:\ i+j=2l-1} x_{ij}, \qquad l = 1, \ldots, k, \\ \sum_{r=l}^{k} y_r \binom{r}{l} a^{r-l} &= \sum_{i,j:\ i+j=2l} x_{ij}, \qquad l = 0, \ldots, k, \\ X &\succeq 0. \end{aligned} \tag{5}$$

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

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

(f) *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 $X = [x_{ij}]_{i,j=0,\ldots,k}$, *such that*

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

Proof:

(a) Suppose (2) holds. Let $x_{(k)} = (1, x, x^2, \ldots, x^k)'$. Then

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

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

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

with X positive semidefinite, from where Equation (2) follows.

(b) We observe that $g(x) \geq 0$ for $x \geq 0$ if and only if $g(t^2) \geq 0$ for all t. Since

$$g(t^2) = y_0 + 0 \cdot t + y_1 t^2 + 0 \cdot t^3 + y_2 t^4 + \cdots + y_k t^{2k},$$

we obtain (3) by applying part (a).

(c) 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 (4).

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

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

Since

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

by applying part (a) we obtain (5).

(e) We observe that $g(x) \geq 0$ for $x \in (-\infty, a]$ if and only if

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

Since

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

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

(f) 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 (7). □

We next show that Problem (1) can be written as a semidefinite optimization problem.

Theorem 3 *Given the first k moments $(M_1, \ldots, M_k)$ (we let $M_0 = 1$) of a random variable X defined on Ω we obtain the following tight upper bounds:*

(a) *If $\Omega = R^+$, the tight upper bound on $P(X \geq a)$ is given as the solution of the semidefinite opti-*

mization problem

$$
\begin{aligned}
\text{minimize} \quad & \sum_{r=0}^{k} y_r M_r \\
\text{subject to} \quad & 0 = \sum_{i,j:\ i+j=2l-1} x_{ij}, && l = 1, \ldots, k, \\
& (y_0 - 1) + \sum_{r=1}^{k} y_r a^r = x_{00}, \\
& \sum_{r=l}^{k} y_r \binom{r}{l} a^{r-l} = \sum_{i,j:\ i+j=2l} x_{ij}, && l = 1, \ldots, k, \\
& 0 = \sum_{i,j:\ i+j=2l-1} z_{ij}, && l = 1, \ldots, k, \\
& \sum_{r=0}^{l} y_r \binom{k-r}{l-r} a^r = \sum_{i,j:\ i+j=2l} z_{ij}, && l = 0, \ldots, k, \\
& X, Z \succeq 0.
\end{aligned} \tag{8}
$$

If $\Omega = R$, then the tight bound on $P(X \geq a)$ is as above with the next to last equation in (8) replaced by

$$
(-1)^l \sum_{r=l}^{k} y_r \binom{r}{l} a^{r-l} = \sum_{i,j:\ i+j=2l} z_{ij}, \qquad l = 0, \ldots, k.
$$

(b) *If $\Omega = R^+$, the tight upper bound on $P(a \leq X \leq b)$ is given as the solution of the semidefinite optimization problem*

$$
\begin{aligned}
\text{minimize} \quad & \sum_{r=0}^{k} y_r M_r \\
\text{subject to} \quad & 0 = \sum_{i,j:\ i+j=2l-1} x_{ij}, && l = 1, \ldots, k, \\
& \sum_{m=0}^{l} \sum_{r=m}^{k+m-l} y_r \binom{r}{m} \binom{k-r}{l-m} a^{r-m} b^m = \binom{k}{l} + \sum_{i,j:\ i+j=2l} x_{ij}, && l = 0, \ldots, k, \\
& 0 = \sum_{i,j:\ i+j=2l-1} z_{ij}, && l = 1, \ldots, k, \\
& y_l = \sum_{i,j:\ i+j=2l} z_{ij}, && l = 0, \ldots, k, \\
& X, Z \succeq 0.
\end{aligned} \tag{9}
$$

If $\Omega = R$, then the tight upper bound on $P(a \leq X \leq b)$ is as above with the next to last equation in (9)

replaced by

$$(-1)^l \sum_{r=l}^{k} y_r \binom{r}{l} a^{r-l} = \sum_{i,j:\ i+j=2l} z_{ij}, \qquad l = 0, \ldots, k,$$

and the following equations added

$$\begin{aligned} 0 &= \sum_{i,j:\ i+j=2l-1} u_{ij}, \qquad l = 1, \ldots, k, \\ \sum_{r=l}^{k} y_r \binom{r}{l} b^{r-l} &= \sum_{i,j:\ i+j=2l} u_{ij}, \qquad l = 0, \ldots, k, \\ U &\succeq 0. \end{aligned}$$

Proof:

(a) The feasible region of Problem (1) for $S = [a, \infty)$ and $\Omega = R_+$, becomes:

$$g(x) = \sum_{r=0}^{k} y_r x^r \geq 1, \ \forall\ x \in [a, \infty), \qquad \text{and } g(x) \geq 0, \ \forall\ x \in [0, a).$$

By applying Proposition 1(c),(d) we obtain (8). If $\Omega = R$, we apply Proposition 1(d),(e).

(b) The feasible region of Problem (1) for $S = [a, b]$ and $\Omega = R_+$, becomes:

$$g(x) = \sum_{r=0}^{k} y_r x^r \geq 1, \ \forall\ x \in [a, b], \qquad \text{and } g(x) \geq 0, \ \forall\ x \in [0, \infty).$$

By applying Proposition 1(b),(f) we obtain (9). If $\Omega = R$, we apply Proposition 1(d), (e), (f). □

3.2 Closed form bounds

In this section, we find closed form bounds when up to the first three first moments are given. We define the squared coefficient of variation: $C_M^2 = \dfrac{M_2 - M_1^2}{M_1^2}$, and the third order coefficient of variation $D_M^2 = \dfrac{M_1 M_3 - M_2^2}{M_1^4}$. Let $\delta > 0$.

Theorem 4 *The following bounds in Table 1 are tight for $k = 1, 2, 3$.*

$(k,\ \Omega)$	$P(X > (1+\delta)M_1)$	$P(X < (1-\delta)M_1)$	$P(\lvert X - M_1\rvert > \delta M_1)$
$(1, R_+)$	$\frac{1}{1+\delta}$	1	1
$(2, R)$	$\frac{C_M^2}{C_M^2+\delta^2}$	$\frac{C_M^2}{C_M^2+\delta^2}$	$\min\left(1, \frac{C_M^2}{\delta^2}\right)$
$(3, R_+)$	$f_1(C_M^2, D_M^2, \delta)$	$f_2(C_M^2, D_M^2, \delta)$	$f_3(C_M^2, D_M^2, \delta)$

Table 1: Tight Bounds for the $(1, k, \Omega)$-problem for $k \leq 3$.

The following definitions are used:

$$f_1(C_M^2, D_M^2, \delta) = \begin{cases} \min\left(\frac{C_M^2}{C_M^2+\delta^2}, \frac{1}{1+\delta}\cdot\frac{D_M^2}{D_M^2+(C_M^2-\delta)^2}\right), & \text{if } \delta > C_M^2, \\ \frac{1}{1+\delta}\cdot\frac{D_M^2+(1+\delta)(C_M^2-\delta)}{D_M^2+(1+C_M^2)(C_M^2-\delta)}, & \text{if } \delta \leq C_M^2, \end{cases}$$

$$f_2(C_M^2, D_M^2, \delta) = 1 - \frac{(C_M^2+1)^3}{(D_M^2+(C_M^2+1)(C_M^2+\delta))(D_M^2+(C_M^2+\delta)^2)},$$

$$f_3(C_M^2, D_M^2, \delta) = \min\left(1, 1+3^3\frac{d_M^2+C_M^4-\delta^2}{4+3(1+3\delta^2)+2(1+3\delta^2)^{\frac{3}{2}}}\right).$$

The proof of the theorem is given in Popescu [44].

4 The $(\mathbf{n}, \mathbf{1}, \Omega)$-Bound Problem.

In this section, we address the $(n, 1, \Omega)$-bound problem. We present tight bounds as solutions to n convex optimization problems for the $(n, 1, R_+^n)$-bound problem for the case when the event S is convex. We present a polynomial time algorithm for more general sets.

In this case, we are given a vector M that represents the vector of means of a random variable X defined in R_+^n, and we would like to find tight bounds on $P(X \in S)$ for a convex set S. Marshall [35] derived a tight bound for the case that $S = \{x_i > (1+\delta_i)M_i,\ i = 1, \ldots, n\}$ (see Theorem 14 below). For general convex sets S, we believe the following result is new.

Theorem 5 *The tight $(n, 1, R^n_+)$-upper bound for an arbitrary convex event S is given by:*

$$\sup_{X \sim M} P(X \in S) = \min\left(1, \max_{i=1,\ldots,n} \frac{M_i}{\inf\limits_{x \in S_i} x_i}\right), \tag{10}$$

where $S_i = S \cap (\cap_{j \neq i}\{x \in R^n_+ |\ M_i x_j - M_j x_i \leq 0\})$.

Proof: Problem (D) can be written as follows for this case:

$$\begin{aligned} Z_D = \text{minimize} \quad & a'M + b \\ \text{subject to} \quad & a'x + b \geq 1, \quad \forall x \in S, \\ & a'x + b \geq 0, \quad \forall x \in R^n_+. \end{aligned}$$

If the optimal solution (a_0, b_0) satisfies $\min\limits_{x \in S} a_0' x + b_0 = \alpha > 1$, then the solution $\left(\frac{a_0}{\alpha}, \frac{b_0}{\alpha}\right)$ has value $Z_D/\alpha < Z_D$. Therefore, $\inf\limits_{x \in S} a'x + b = 1$. By a similar argument we have that $b_0 \leq 1$. Moreover, since $a'x + b \geq 0,\ \forall x \in R^n_+$, $a \geq 0$, and $b \geq 0$. We thus obtain:

$$\begin{aligned} Z_D = \text{minimize} \quad & a'M + b \\ \text{subject to} \quad & \inf_{x \in S} a'x = 1 - b. \\ & a \geq 0,\ 0 \leq b \leq 1. \end{aligned}$$

Without loss of generality we let $a = \lambda v$, where λ is a nonnegative scalar, and v is a nonnegative vector with $||v|| = 1$. Thus, we obtain:

$$\begin{aligned} Z_D = \text{minimize} \quad & (1 - b)\frac{v'M}{\inf\limits_{x \in S} v'x} + b \\ \text{subject to} \quad & v \geq 0,\ ||v|| = 1,\ 0 \leq b \leq 1. \end{aligned}$$

Thus,

$$
\begin{aligned}
Z_D &= \min\left(1, \min_{||v||=1, v\geq 0} \frac{v'M}{\inf_{x\in S} v'x}\right) \\
&= \min\left(1, \min_{||v||=1, v\geq 0} \sup_{x\in S} \frac{v'M}{v'x}\right) \\
&= \min\left(1, \sup_{x\in S} \min_{||v||=1, v\geq 0} \frac{v'M}{v'x}\right) \qquad (11) \\
&= \min\left(1, \sup_{x\in S} \min_{i=1,\ldots,n} \frac{M_i}{x_i}\right) \qquad (12) \\
&= \min\left(1, \max_{i=1,\ldots,n} \frac{M_i}{\inf_{x\in S_i} x_i}\right), \qquad (13)
\end{aligned}
$$

where $S_i = S \cap (\cap_{j\neq i}\{x \in R_+^n |\ M_i x_j - M_j x_i \leq 0\})$ is a convex set. Note that in Eq. (11) we exchanged the order of min and sup (see Rockafellar [50], p. 382). In Eq. (12), we used $\min_{||v||=1, v\geq 0} \frac{v'M}{v'x}$ is attained at $v = e_j$, where

$$\frac{M_j}{x_j} = \min_{i=1,\ldots,n} \frac{M_i}{x_i}.$$

In order to understand Eq. (13), we let $\phi(x) = \min_{i=1,\ldots,n} \frac{M_i}{x_i}$. Note that $\phi(x) = \frac{M_i}{x_i}$, when $x \in \{x \in R_+^n |\ M_i x_j - M_j x_i \leq 0\}$. Then, we have

$$\sup_{x\in S} \phi(x) = \max_{i=1,\ldots,n} \sup_{x\in S_i} \phi(x) = \max_{i=1,\ldots,n} \sup_{x\in S_i} \frac{M_i}{x_i} = \max_{i=1,\ldots,n} \frac{M_i}{\inf_{x\in S_i} x_i}.$$

□

When we specialize Bound (10) for the set $S = \{x |\ x_i \geq (1+\delta_i)M_i\ ,\ \forall i = 1, \ldots, n\}$, we obtain that

$$\sup_{X\sim M^+} P(X_1 \geq (1+\delta_1)M_1, \ldots, X_n \geq (1+\delta_n)M_n) = \min_{i=1,\ldots,n} \frac{1}{1+\delta_i}, \qquad (14)$$

which represents a multidimensional generalization of Markov's inequality due to Marshall [35]. In particular, for a nonnegative univariate random variable, in the case that $S = [(1+\delta)M, \infty)$, the bound

(14) is exactly Markov inequality:

$$\sup_{X \sim M^+} P(X \geq (1+\delta)M) = \frac{1}{1+\delta}.$$

4.1 Extremal Distributions for The $(\mathbf{n}, \mathbf{1}, \mathbf{R_+^n})$-Bound Problem.

In this section, we construct a distribution that achieves Bound (10). We will say that the Bound (10) is achievable, when there exists an $x^* \in S$ such that

$$\min\left(1, \max_{i=1,\ldots,n} \frac{M_i}{\inf\limits_{x \in S_i} x_i}\right) = \frac{M_i}{x_i^*} < 1.$$

In particular, the bound is achievable when the set S is closed and $M \notin S$.

Theorem 6 **(a)** *If $M \in S$ or if the Bound (10) is achievable, then there is an extremal distribution that exactly achieves it.*

(b) *Otherwise, there is a sequence of distributions defined on R_+^n with mean M, that asymptotically achieve it.*

Proof: **(a)** If $M \in S$, then the extremal distribution is simply $P(X = M) = 1$. Now suppose that $M \notin S$ and the Bound (10) is achievable. We assume without loss of generality that the bound equals $\frac{M_1}{x_1^*} < 1$, and it is achieved at $x^* \in S$. Therefore, $\frac{M_1}{x_1^*} = \min\limits_{i=1,\ldots,n} \frac{M_i}{x_i^*}$. We consider the following random variable X defined on R_+^n:

$$X = \begin{cases} x^*, & \text{with probability} \quad p = \dfrac{M_1}{x_1^*}, \\ v = \dfrac{x_1^* M - M_1 x^*}{x_1^* - M_1}, & \text{with probability} \quad 1 - p = 1 - \dfrac{M_1}{x_1^*}. \end{cases}$$

Note that $E[X] = M$, and $v_i = \frac{M_i x_1^* - M_1 x_i^*}{x_1^* - M_1} \geq 0$ for all $i = 1, \ldots, n$, since $\frac{M_1}{x_1^*} = \min\limits_{i=1,\ldots,n} \frac{M_i}{x_i^*}$. Moreover, $v \notin S$, or else by the convexity of S, we have that $M = px^* + (1-p)v \in S$, a contradiction. Therefore,

$$P(X \in S) = P(X = x^*) = \frac{M_1}{x_1^*}.$$

(b) If $M \notin S$ and the Bound (10) is not achievable, then we construct a sequence of nonnegative distributions with mean M that approach it. Suppose without loss of generality that $\max\limits_{i=1,\ldots,n} \frac{M_i}{\inf\limits_{x \in S_i} x_i}$

equals $\frac{M_1}{x_1^*}$, for $x^* \in \overline{S}_1$ (the closure of S_1), so Bound (10) is equal to $\min\left(1, \frac{M_1}{x_1^*}\right)$. Consider a sequence $x^k \in S_1$, $x^k \to x^*$, so that $\lim_{k\to\infty} \min_{i=1,\ldots,n} \frac{M_i}{x_i^k} = \frac{M_1}{x_1^*}$, and a sequence p_k, $0 < p_k < \min\left(1, \frac{M_1}{x_1^k}\right)$ so that $p_k \to \min\left(1, \frac{M_1}{x_1^*}\right)$. Consider the sequence of distributions:

$$X_k = \begin{cases} x^k, & \text{with probability} \quad p_k, \\ v^k = \dfrac{M - p_k x^k}{1 - p_k}, & \text{with probability} \quad 1 - p_k. \end{cases}$$

Clearly, the random variables X_k are nonnegative with mean $E[X_k] = M$. Also $v^k \notin S$ or else $M \in S$, so $P(X_k \in S) = P(X_k = x^k) = p_k \to \min\left(1, \frac{M_1}{x_1^*}\right)$. This shows that the sequence of nonnegative, distributions X_k with mean M asymptotically achieve the Bound (10). □

4.2 A Polynomial Time Algorithm for Unions of Convex Sets.

In this section, we present a polynomial time algorithm that computes a tight $(n, 1, \Omega)$- bound for any event S that can be decomposed as a disjoint union of a polynomial (in n) number of convex sets. We further assume that the set Ω can be decomposed as a disjoint union of a polynomial (in n) number of convex sets. Our overall strategy is to formulate the problem as an optimization problem, consider its dual and exhibit an algorithm that solves the corresponding separation problem in polynomial time.

In this case we are given the mean-vector $M = (M_1, \ldots, M_n)$ of an n-dimensional random variable X with domain Ω that can be decomposed in a polynomial (in n) number of convex sets, and we want to derive tight bounds on $P(X \in S)$. Problem (D) can be written as follows:

$$\begin{aligned} Z_D = \text{minimize} \quad & u'M + u_0 \\ \text{subject to} \quad & g(x) = u'x + u_0 \geq \chi_S(x), \ \forall x \in \Omega. \end{aligned} \tag{15}$$

The separation problem associated with Problem (15) is defined as follows: Given a vector a and a scalar b we want to check whether $g(x) = a'x + b \geq \chi_S(x)$, $\forall x \in \Omega$, and if not, we want to exhibit a violated inequality. The following algorithm achieves this goal.

Algorithm A:

1. Solve the problem $\inf_{x\in\Omega} g(x)$ (note that the problem involves a polynomial number of convex optimization problems; in particular if Ω is polyhedral, this is a linear optimization problem). Let z_0 be the optimal solution value and let $x_0 \in \Omega$ be an optimal solution.

2. If $z_0 < 0$, then we have $g(x_0) = z_0 < 0$: this constitutes a violated inequality;

3. Otherwise, we solve $\inf_{x \in S} g(x)$ (again, the problem involves a polynomial number of convex optimization problems, while if S is polyhedral, this is a linear optimization problem). Let z_1 be the optimal solution value and let $x_1 \in S$ be an optimal solution.

   (a) If $z_1 < 1$, then for $x_1 \in S$ we have $g(x_1) = z_1 < 1$: this constitutes a violated inequality.

   (b) If $z_1 \geq 1$, then a, b are feasible.

The above algorithm solves the separation problem in polynomial time, since we can solve any convex optimization problem in polynomial time (see Nesterov and Nemirowskii [41], Nemhauser and Wolsey [40]). Therefore, the $(n, 1, \Omega)$-upper bound problem is polynomially solvable.

5 The $(\mathbf{n}, \mathbf{2}, \mathbf{R^n})$-Bound Problem.

In this section, we address the $(n, 2, R^n)$-bound problem. Rather than assuming that $E[X]$ and $E[XX']$ are known, we assume equivalently that the vector $M = E[X]$ and the covariance matrix $\Gamma = E[(X - M)(X-M)']$ are known. Given a set $S \subseteq R^n$, we find tight upper bounds, denoted by $\sup_{X \sim (M,\Gamma)} P(X \in S)$, on the probability $P(X \in S)$ for all multivariate random variables X defined on R^n with mean $M = E[X]$ and covariance matrix $\Gamma = E[(X - M)(X - M)']$. If the set S is convex, we review the work of Marshall and Olkin [36], who solved the problem as a convex optimization problem. If the set S is a union of convex sets, we provide a polynomial time algorithm.

5.1 The $(\mathbf{n}, \mathbf{2}, \mathbf{R^n})$-Bound Problem for a Convex Set.

The following result is due to Marshall and Olkin [36], who give a constructive proof. An alternative, optimization based proof is provided in Popescu [44].

Theorem 7 *The tight $(n, 2, R^n)$-upper bound for an arbitrary convex event S is given by:*

$$\sup_{X \sim (M,\Gamma)} P(X \in S) = \frac{1}{1 + d^2}, \tag{16}$$

where $d^2 = \inf_{x \in S} (x - M)' \Gamma^{-1} (x - M)$, is the squared distance from M to the set S, under the norm induced by the matrix Γ^{-1}.

The actual formulation provided by Marshall and Olkin [36] is the following :

$$\sup_{X\sim(0,\Gamma)} P(X \in S) = \inf_{a\in S^{\perp}} \frac{1}{1+(a'\Gamma a)^{-1}}\,, \tag{17}$$

where $S^{\perp} = \{a \in R^n |\ a'x \geq 1,\ \forall\, x \in S\}$, is the so-called "antipolar" of S (a.k.a "blocker", or "upper-dual"). The above result is with zero mean, but can be easily extended for nonzero mean by a simple transformation (see [36] p.1013, eqs. (7.8)-(7.9)). Given that $(a'\Gamma a)(x'\Gamma^{-1}x) \geq (a'x)^2 \geq 1\ \forall\, x \in S, a \in S^{\perp}$, one can easily see that the Bound (16) is at least as tight as Bound (17). Equality follows from nonlinear Gauge duality principles (see Freund [15]).

Marshall and Olkin [36] construct an extremal distribution of a random variable $X \sim (M, \Gamma)$, so that $P(X \in S) = 1/(1+d^2)$ with $d^2 = \inf_{x\in S}(x-M)'\Gamma^{-1}(x-M)$. We will say that the bound d is achievable, when there exists an $x^* \in S$ such that $d^2 = (x^*-M)'\Gamma^{-1}(x^*-M)$. In particular, d is achievable if the set S is closed.

Theorem 8 (Marshall and Olkin [36]) (a) *If $M \notin S$ and if $d^2 = \inf_{x\in S}(x-M)'\Gamma^{-1}(x-M)$ is achievable, then there is an extremal distribution that exactly achieves the Bound (16).*
(b) *Otherwise, if $M \in S$ or if d^2 is not achievable, then there is a sequence of (M,Γ)-feasible distributions that asymptotically approach the Bound (16).*

Bounds on Tail Probabilities.

Given a vector $M = (M_1, \ldots, M_n)'$, and an $n \times n$ positive definite, full rank matrix Γ, we next derive tight bounds on the following upper, lower, and two-sided tail probabilities of a multivariate random variable $X = (X_1, \ldots, X_n)'$ with mean $M = E[X]$ and covariance matrix $\Gamma = E[(X-M)(X-M)']$:

$$\begin{aligned} P(X > M_{e+\delta}) &= P(X_i > (1+\delta_i)M_i\,,\ \forall i = 1, \ldots, n), \\ P(X < M_{e-\delta}) &= P(X_i < (1-\delta_i)M_i\,,\ \forall i = 1, \ldots, n), \\ P(X > M_{e+\delta} \text{ or } X < M_{e-\delta}) &= P(|X_i - M_i| > \delta_i M_i\,,\ \forall i = 1, \ldots, n), \end{aligned}$$

where $\delta = (\delta_1, \ldots, \delta_n)'$, and we denote by $M_\delta = (\delta_1 M_1, \ldots, \delta_n M_n)'$. In order to obtain nontrivial bounds we require that not all $\delta_i M_i \leq 0$, which expresses the fact that the tail event does not include the mean vector.

The One-Sided Chebyshev Inequality.

In the following, we find a tight bound for $P(X > M_{e+\delta})$. The bound immediately extends to $P(X < M_{e-\delta})$.

Proposition 2 (a) *The tight multivariate one-sided* $(n, 2, R^n)$*-Chebyshev bound is*

$$\sup_{X\sim(M,\Gamma)} P(X > M_{e+\delta}) = \frac{1}{1+d^2}, \tag{18}$$

where d^2 *is given by:*

$$\begin{aligned} d^2 = \text{minimize} \quad & x'\Gamma^{-1}x \\ \text{subject to} \quad & x \geq M_\delta, \end{aligned} \tag{19}$$

or alternatively d^2 *is given by the Gauge dual problem of (19):*

$$\begin{aligned} \frac{1}{d^2} = \text{minimize} \quad & x'\Gamma x \\ \text{subject to} \quad & x'M_\delta = 1 \\ & x \geq 0. \end{aligned} \tag{20}$$

(b) *If* $\Gamma^{-1}M_\delta \geq 0$*, then the tight bound is expressible in closed form:*

$$\sup_{X\sim(M,\Gamma)} P(X > M_{e+\delta}) = \frac{1}{1+M_\delta'\Gamma^{-1}M_\delta}. \tag{21}$$

Proof: (a) Applying the Bound (16) for $S = \{x|\ x_i > (1+\delta_i)M_i\ ,\ \forall i = 1, \ldots, n\}$, and changing variables we obtain Eq. (18). The alternative expression (20) for d^2 follows from elementary Gauge duality theory (see Freund [15]).

(b) The Kuhn-Tucker conditions for Problem (19) are as follows:

$$2\Gamma^{-1}x - \lambda = 0, \qquad \lambda \geq 0, \qquad x \geq M_\delta, \qquad \lambda'(x - M_\delta) = 0.$$

The choice $x = M_\delta$, $\lambda = 2\Gamma^{-1}M_\delta \geq 0$ (by assumption) satisfies the Kuhn-Tucker conditions, which are sufficient (this is a convex quadratic optimization problem). Thus, $d^2 = M_\delta'\Gamma^{-1}M_\delta$, and hence, Eq. (21) follows. □

The Two-Sided Chebyshev Inequality.

The following result provides a tight bound for $P(X > M_{e+\delta} \text{ or } X < M_{e-\delta})$.

Proposition 3 (Marshall and Olkin [36])

(a) *The tight multivariate two-sided* $(n, 2, R^n)$*-Chebyshev bound is*

$$\sup_{X \sim (M,\Gamma)} P(X > M_{e+\delta} \text{ or } X < M_{e-\delta}) = \min(1, t^2), \tag{22}$$

where

$$\begin{aligned} t^2 = \text{minimize} \quad & x'\Gamma x \\ \text{subject to} \quad & x'M_\delta = 1 \\ & x \geq 0. \end{aligned} \tag{23}$$

(b) *If* $\Gamma^{-1}M_\delta \geq 0$*, then the tight bound is expressible in closed form:*

$$\sup_{X \sim (M,\Gamma)} P(X > M_{e+\delta} \text{ or } X < M_{e-\delta}) = \min\left(1, \frac{1}{M_\delta' \Gamma^{-1} M_\delta}\right). \tag{24}$$

This formulation follows via Gauge duality principles from the original zero-mean result of Marshall and Olkin [37].

In the univariate case $M_\delta = \delta M$ and $\Gamma = \sigma^2$. Therefore, $\Gamma^{-1}M_\delta = \dfrac{\delta M}{\sigma^2} \geq 0$, and the closed form bound applies, i.e.,

$$P(X > (1+\delta)M) \leq \frac{C_M^2}{\delta^2 + C_M^2}, \tag{25}$$

where $C_M^2 = \dfrac{\sigma^2}{M^2}$ is the coefficient of variation of the random variable X. The usual Chebyshev inequality is given by $P(X > (1+\delta)M) \leq \frac{C_M^2}{\delta^2}$. Inequality (25) is always stronger. Moreover, there exist extremal distributions that satisfy it with equality (see Theorem 8). The original univariate result can be traced back to the 1937 book of Uspensky [59], and is mentioned later by Marshall and Olkin (1960) [36], [37], but has not received much attention in modern probability textbooks.

5.2 The Tight $(\mathbf{n}, \mathbf{2}, \mathbf{R^n})$-Bound for Unions of Convex Sets.

We are given first and second order moment information (M, Γ) on the n- dimensional random variable X, and we would like to compute $\sup_{X \sim (M,\Gamma)} P(X \in S)$. Recall that the corresponding dual problem can be written as:

$$\begin{aligned} Z_D = \text{minimize} \quad & E[g(X)] \\ \text{subject to} \quad & g(x) = x'Hx + c'x + d \geq \chi_S(x), \qquad \forall\, x \in R^n. \end{aligned} \tag{26}$$

The separation problem corresponding to Problem (26) can be stated as follows: Given a matrix H, a vector c and a scalar d, we need to check whether $g(x) = x'Hx + c'x + d \geq \chi_S(x),\ \forall x \in R^n$, and if not, find a violated inequality. Notice that we can assume without loss of generality that the matrix H is symmetric.

The following algorithm solves the separation problem in polynomial time.

Algorithm B:

1. If H is not positive semidefinite, then we find a vector x_0 so that $g(x_0) < 0$. We decompose $H = Q'\Lambda Q$, where $\Lambda = \text{ diag}(\lambda_1, \ldots, \lambda_n)$ is the diagonal matrix of eigenvalues of H. Let $\lambda_i < 0$ be a negative eigenvalue of H. Let y be vector with $y_j = 0$, for all $j \neq i$, and y_i large enough so that $\lambda_i y_i^2 + (Qc)_i y_i + d < 0$. Let $x_0 = Q'y$. Then,

$$\begin{aligned} g(x_0) &= x_0' H x_0 + c' x_0 + d \\ &= y' Q Q' \Lambda Q Q' y + c' Q' y + d \\ &= y' \Lambda y + c' Q' y + d \\ &= \sum_{j=1}^{n} \lambda_j y_j^2 + \sum_{j=1}^{n} (Qc)_j y_j + d \\ &= \lambda_i y_i^2 + (Qc)_i y_i + d < 0. \end{aligned}$$

   This produces a violated inequality.

2. Otherwise, if H is positive semidefinite, then:

   (a) We test if $g(x) \geq 0,\ \forall x \in R^n$ by solving the convex optimization problem:

$$\inf_{x \in R^n} g(x).$$

Let z_0 be the optimal value. If $z_0 < 0$, we find x_0 such that $g(x_0) < 0$, which represents a violated inequality. Otherwise,

(b) We test if $g(x) \geq 1, \ \forall x \in S$ by solving a polynomial collection of convex optimization problems

$$\inf_{x \in S} g(x).$$

Let z_1 be the optimal value. If $z_1 \geq 1$, then $g(x) \geq 1, \ \forall x \in S$, and thus (H, c, d) is feasible. If not, we exhibit an x_1 such that $g(x_1) < 1$, and thus we identify a violated inequality.

Since we can solve the separation problem in polynomial time, we can also solve (within ϵ) the $(n, 2, R^n)$-bound problem in polynomial time (in the problem data and $\log \frac{1}{\epsilon}$).

6 The Complexity of $(\mathbf{n}, \mathbf{2}, \mathbf{R^n_+})$, $(\mathbf{n}, \mathbf{k}, \mathbf{R^n})$-Bound Problems.

In this section, we show that the separation problem associated with Problem (D) for the cases $(n, 2, R^n_+)$, (n, k, R^n)-bound problems are NP-hard for $k \geq 3$. By the equivalence of optimization and separation (see Grötschel, Lovász and Schrijver [18]), solving Problem (D) is NP-hard as well. Finally, because of Theorem 2, solving the $(n, 2, R^n_+)$, (n, k, R^n)-bound problems with $k \geq 3$ is NP-hard.

6.1 The Complexity of The $(\mathbf{n}, \mathbf{2}, \mathbf{R^n_+})$-Bound Problem.

The separation problem can be formulated as follows in this case:

Problem 2SEP: Given a multivariate polynomial $g(x) = x'Hx + c'x + d$, and a set $S \subseteq R^n_+$, does there exist $x \in S$ such that $g(x) < 0$?

If we consider the special case $c = 0$, $d = 0$, and $S = R^n_+$, Problem 2SEP reduces to the question whether a given matrix H is co-positive, which is NP-hard (see Murty and Kabadi [39]).

6.2 The Complexity of The $(\mathbf{n}, \mathbf{k}, \mathbf{R^n})$-Bound Problem for $k \geq 3$.

For $k \geq 3$, the separation problem can be formulated as follows:

Problem 3SEP: Given a multivariate polynomial $g(\cdot)$ of degree $k \geq 3$, and a set $S \subseteq R^n$, does there exist $x \in S$ such that $g(x) < 0$?

We show that problem $3SEP$ is NP-hard by performing a reduction from $3SAT$ (see Sipser [54]).

Theorem 9 *Problem $3SAT$ polynomially reduces to $3SEP$.*

Proof: For an arbitrary $3SAT$ instance ϕ (a 3CNF boolean formula in n variables), we consider the following arithmetization $g_\phi(\cdot)$ of ϕ: we replace each boolean variable x_i by the monomial $1 - x_i$, its negation $\bar{x}_i$ by x_i, and we convert $\bigwedge$'s into additions and $\bigvee$'s to multiplications. For example, the arithmetization of the formula $\phi = (\bar{x}_1 \vee x_2 \vee \bar{x}_3) \wedge (x_1 \vee x_3 \vee \bar{x}_4)$, is: $g_\phi(x) = x_1(1-x_2)x_3 + (1-x_1)(1-x_3)x_4$.

As motivation for the proof, note that $g_\phi(\cdot)$ is a 3-degree polynomial in n variables, evaluating to zero at any satisfying assignment of ϕ. Also note that $g_\phi(x)$ is a nonnegative integer for any boolean assignment $x \in \{0,1\}^n$. Thus if ϕ is unsatisfiable, then $g_\phi(x) \geq 1$ for any boolean assignment $x \in \{0,1\}^n$.

Starting with an instance ϕ of 3SAT with n variables and m clauses, we construct an instance $(g(\cdot), S)$ of $3SEP$ as follows:

$$g(x) = 2g_\phi(x) + (24m)^2 \sum_{i=1}^{n} x_i(1 - x_i) - 1, \qquad S = [0,1]^n.$$

Note that the construction can be done in polynomial time. We next show that formula ϕ is satisfiable if and only if there exists $x \in S$ such that $g(x) < 0$.

Clearly if ϕ is satisfiable, there exists a satisfying assignment corresponding to a vector $x_0 \in \{0,1\}^n$. Clearly $g_\phi(x_0) = 0$, and thus $g(x_0) = -1 < 0$.

Conversely, suppose ϕ is not satisfiable. We will show that for all $x \in S = [0,1]^n$, $g(x) \geq 0$. Let $\epsilon = \dfrac{1}{24m}$. For any $x \in S = [0,1]^n$, there are two possibilities:

(a) There exists a boolean vector $y \in \{0,1\}^n$ such that $|x_i - y_i| < \epsilon, \ \forall i$.

If we expand the term in $g_\phi(\cdot)$ corresponding to each of the m clauses of ϕ as a polynomial, we obtain a sum of at most one monomial of degree three, three monomials of degree two, three monomials of degree one, and one monomial of degree zero. Let S_k be the set k-tuples corresponding to the monomials of degree k, $k = 1, 2, 3$. Then, $|S_1| \leq 3m$, $|S_2| \leq 3m$, $|S_3| \leq m$. Matching corresponding monomials for x and y, canceling constants, and applying the triangle inequality, we obtain:

$$|g_\phi(x) - g_\phi(y)| \leq \sum_{(i,j,k) \in S_3} |x_i x_j x_k - y_i y_j y_k| + \sum_{(i,j) \in S_2} |x_i x_j - y_i y_j| + \sum_{i \in S_1} |x_i - y_i|.$$

Since $|x_i - y_i| < \epsilon, \ \forall i$, we obtain:

$$\begin{aligned} |x_i x_j x_k - y_i y_j y_k| &\leq |x_i x_j x_k - y_i x_j x_k| + |y_i x_j x_k - y_i y_j x_k| + |y_i y_j x_k - y_i y_j y_k| \\ &= |x_i - y_i| x_j x_k + y_i |x_j - y_j| x_k + y_i y_j |x_k - y_k| \leq 3\epsilon, \end{aligned}$$

since $x, y \in [0,1]^n$. Similarly, $|x_i x_j - y_i y_j| \leq 2\epsilon$. Therefore,

$$|g_\phi(x) - g_\phi(y)| \leq 3\epsilon|S_3| + 2\epsilon|S_2| + \epsilon|S_1| \leq 12m\epsilon = \frac{1}{2}.$$

Thus, $g_\phi(x) > g_\phi(y) - \frac{1}{2}$. Since ϕ is not satisfiable, we have $g_\phi(y) \geq 1$, for any boolean vector $y \in \{0,1\}^n$. Thus, $g_\phi(x) \geq 1 - \frac{1}{2} = \frac{1}{2}$, and hence $g(x) \geq 2g_\phi(x) - 1 \geq 0$.

(b) There exists at least one i for which $\epsilon \leq x_i \leq 1 - \epsilon$. This implies $x_i(1 - x_i) \geq \epsilon^2$, and, since $g_\phi(x) \geq 0$, $\forall\, x \in S$, it follows that $g(x) \geq (24m)^2\epsilon^2 - 1 = 0$.

Therefore, if ϕ is not satisfiable, then all $x \in S = [0,1]^n$, satisfy $g(x) \geq 0$, and the theorem follows. □

7 Concluding Remarks.

In this paper we characterized the complexity of the (n, k, Ω)-bound problem, by providing polynomial time algorithms, via convex optimization methods, for the $(1, k, \Omega)$, $(n, 1, \Omega)$, $(n, 2, R^n)$-bound problems, and by showing that the $(n, 2, R^n_+)$, (n, k, R^n)-bound problems for $k \geq 3$ are NP-hard.

Acknowledgents

We would like to thank Professors Guillelmo Gallego, James Primbs, Alexander Shapiro and James Smith for bringing to our attention some of the early and recent literature on the subject and many insightful comments on an earlier version of the paper.

References

[1] N. I. Akhiezer. *The Classical Moment Problem.* Hafner, New York, NY, 1965.

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

[3] J.R. Birge and R. J.-B. Wets. Approximations and error bounds in stochastic programming. In Y.L. Tong, editor, *Inequalities in Statistics and Probability*, volume 5 of *IMS Lecture Notes Monograph Series*, pages 178–186. 1984.

[4] J.R. Birge and R. J.-B. Wets. Designing approximation schemes for stochastic optimization problems, in particular for stochastic programs with recourse. *Math. Prog. Stud.*, 27:54–102, 1986.

[5] J.R. Birge and R. J.-B. Wets. Computing bounds for stochastic programming problems by means of a generalized moment problem. *Math. Oper. Res.*, 12(1):149–162, 1987.

[6] F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems*. Springer, New York, 2000.

[7] J.M. Borwein and A.S. Lewis. Partially finite convex programming, part i: quasi relative interiors and duality theory. *Math. Prog.*, 57:15–48, 1992.

[8] J.M. Borwein and A.S. Lewis. Partially finite convex programming, part ii: explicit lattice models. *Math. Prog.*, 57:48–83, 1992.

[9] J. Bukscár. Probability bounds with multitrees. RUTCOR Research Report, 1999.

[10] P. Chebychev. Sur les valeurs limites des intégrales. *Jour. Math. Pur. Appl.*, 19(2):157–160, 1874.

[11] T. Cipra. Moment problems with given covariance structure in stochastic programming. *Ekonom. Mat. Obzor.*, 21:66–77, 1985.

[12] P. Diaconis. Application of the method of moments in probability and statistics. In ed. H.J. Landau, editor, *Moments in Mathematics*, number 37 in Proc. Sympos. Appl. Math., Series III, pages 125–139, Providence, RI, 1987. Amer. Math. Soc.

[13] J. Dupaçová, A. Prekopa, J.R. Birge, and R. J.-B. Wets. Computing bounds for stochastic programming problems by means of a generalized moment problem. *Math. Opns. Res.*, 12:149–162, 1987.

[14] Yu. Ermoliev, A.A. Gaivoronski, and C. Nedeva. Stochastic programming problems with incomplete information on distribution functions. *SIAM J. Cont. and Optim.*, 23:697–716, 1985.

[15] R. Freund. Dual gauge programs, with applications to quadratic programming and the minimum-norm problem. *Mathematical Programming*, 38:47–67, 1987.

[16] H.J. Godwin. On generlizations of tchebycheff's inequality. *Am. Stat. Assoc. J.*, 50:923–945, 1955.

[17] H.J. Godwin. Inequalities on distribution functions. In ed. M.G. Kendall, editor, *Griffin's Statistical Monographs and Courses*, number 16. Charles Griffin and Co., London, 1964.

[18] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization.* Algorithms and Combinatorics. Springer, Berlin-Heidelberg, 1988.

[19] H. Hamburger. Ueber eine erweiterung der stieltjes'schen momentenproblems. *Mathematische Annalen*, 81:235–319, 1920.

[20] H. Hamburger. Ueber eine erweiterung der stieltjes'schen momentenproblems. *Mathematische Annalen*, 82:120–164, 168–187, 1921.

[21] E.K. Haviland. On the momentum problem for distributions in more than one dimension. *American Journal of Mathematics*, 57:562–568, 1935.

[22] E.K. Haviland. On the momentum problem for distribution functions in more than one dimension. *American Journal of Mathematics*, 58:164–168, 1936.

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

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

[25] P. Kall. Stochastic programms with recourse: An upper bound and the related moment problem. *Z. Opns. Res.*, 31:A119–A141, 1987.

[26] P. Kall. Stochastic programming with recourse: Upper bounds and moment problems, a review. In J. Guddat et al. (eds.), editor, *Advances in Mathematical Optimization.* Akademie-Verlag, Berlin, 1988.

[27] P. Kall. An upper bound for slp using first and total second moments. *Ann. Opns. Res.*, 30:267–276, 1991.

[28] S. Karlin and L.S. Shapley. Geometry of moment spaces. *Memoirs Amer. Math. Soc.*, 12, 1953.

[29] 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.

[30] J.H.B. Kemperman. The general moment problem, a geometric approach. *Annals. Math. Statist.*, (39):93–122, 1968.

[31] J. H. B. Kempermann. On the role of duality in the theory of moments. In *Lecture Notes in Econom. and Math. Systems*, volume 215, pages 63–92, Berlin-New York, 1983. Springer.

[32] J. H. B. Kempermann. Geometry of the moment problem. In ed. H.J. Landau, editor, *Moments in Mathematics*, number 37 in Proc. Sympos. Appl. Math., Series III, pages 16–53, Providence, RI, 1987. American Mathematical Society.

[33] H.J. Landau. Classical background on the moment problem. In ed. H.J. Landau, editor, *Moments in Mathematics*, number 37 in Proc. Sympos. Appl. Math., Series III, pages 1–15, Providence, RI, 1987. American Mathematical Society.

[34] A. Markov. *On Certain Applications of Algebraic Continued Fractions*. PhD thesis, St. Petersburg, 1884. in Russian.

[35] A. Marshall. Markov's inequality for random variables taking values in a linear topological space. In Y.L. Tong, editor, *Inequalities in Statistics and Probability*, volume 5 of *IMS Lecture Notes Monograph Series*, pages 104–108. 1984.

[36] A. Marshall and I. Olkin. Multivariate chebyshev inequalities. *Ann. Math. Stat.*, 31:1001–1014, 1960.

[37] A. Marshall and I. Olkin. A one-sided inequality of the chebyshev type. *Ann. Math. Stat.*, 31:488–491, 1960.

[38] A. Marshall and I. Olkin. A game theoretic proof that chebyshev inequalities are sharp. *Pacific. J. Math.*, 11:1421–1429, 1961.

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

[40] G.L. Nemhauser and L.A. Wolsey. *Integer and Combinatorial Optimization*. Wiley, New York, 1988.

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

[42] I. Pitowski. Correlation polytopes: Their geometry and complexity. *Mathematical Programming*, 50:395–414, 1991.

[43] I. Pitowski. Correlation polytopes ii: The geometry of limit laws in probability. preprint, 1996.

[44] I. Popescu. Applications of optimization in probability, finance and revenue management. Ph.D. Dissertation, Applied Mathematics Department and Operations Research Center, MIT, Cambridge, Massachusetts, 1999.

[45] A. Prekopa. Bounds on probabilities and expectations using multivariate moments of discrete distributions. RUTCOR Research Report, 1993.

[46] A. Prekopa. Programming under probabilistic constraints and maximizing a probability under constraints. RUTCOR Research Report, 1993.

[47] A. Prekopa. Programming under probabilistic constraints with discrete random variables. RUTCOR Research Report, 1996.

[48] A. Prekopa. Probability bounds with cherry trees. RUTCOR Research Report, 1998.

[49] A. Prekopa. The use of discrete moment bounds in probabilistic constrained stochastic programming models. RUTCOR Research Report, 1998.

[50] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, NJ, 1970.

[51] S. Sengupta, W. Siu, and N.C. Lind. A linear programming formulation of the problem of moments. *Z. Angew. Math. Mecg.*, 10:533–537, 1979.

[52] A. Shapiro. On duality theory of conic linear problems. working paper, 2000.

[53] J.A. Shohat and J.D. Tamarkin. The problem of moments. In *Mathematical Surveys*, number 1. American Mathematical Society, New York, 1943.

[54] M. Sipser. *Introduction to the Theory of Computation*. PWS, Boston, MA, 1996.

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

[56] T.J. Stieltjes. Recherches sur les fractions continues. *Annales de la Faculté de Sciences de Toulouse*, 8(1):1–122, 1894.

[57] T.J. Stieltjes. Recherches sur les fractions continues. *Annales de la Faculté de Sciences de Toulouse*, 9(1):5–47, 1895.

[58] Y.L. Tong. *Probability Inequalities in Multivariate Distributions*. New York, Academic Press, 1980.

[59] J.V. Uspensky. *Introduction to Mathematical Probability*. McGraw-Hill, New York, 1937.

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