ORIGINAL PAPER

CrossMark

An iterative algorithm to bound partial moments

Sander Muns[1] ⓘ

Abstract

This paper presents an iterative algorithm that bounds the lower and upper partial moments of an arbitrary univariate random variable X by using the information contained in a sequence of finite moments of X. The obtained bounds on the partial moments imply bounds on the moments of the transformation $f(X)$ for a certain function $f : \mathbb{R} \to \mathbb{R}$. Two examples illustrate the performance of the algorithm.

Keywords Moment problem · Bounds · Censored distributions · Iteration convergence

1 Introduction

Statistical computations for insurance policy, inventory management, Bayesian point estimation, and other areas often involve the computation of partial moments (Winkler et al. 1972). Bawa and Lindenberg (1977) derive a capital asset pricing model from utility functions based on lower partial moments. The i-th upper (lower) partial moment of a univariate real-valued random variable X with cumulative distribution function F is defined as the i-th moment in excess of (below) a certain threshold $a \in \mathbb{R}$,

$$\mu_i^+(a) := \int_a^\infty (x - a)^i \, dF(x) \qquad \mu_i^-(a) := \int_{-\infty}^a (a - x)^i \, dF(x). \qquad (1)$$

This paper presents an algorithm that bounds the partial moments of X using information contained in a sequence of (full) moments of X. This approach is strongly related to the moment problem.

The moment problem originally considered whether a certain sequence of moments corresponds to at least one univariate probability measure. This problem has been extensively discussed in the mathematical literature (Akhiezer 1965; Kreĭn and Nudelman 1977; Shohat and Tamarkin 1943; Stoyanov 2013). The admissible support of the

✉ Sander Muns
   s.muns@uvt.nl

[1] Tilburg University, Netspar, P.O. Box 90153, 5000 LE Tilburg, Netherlands

probability measure splits the moment problem in three different subproblems: The admissible support can be unrestricted (Hamburger moment problem), restricted to the positive half-line (Stieltjes), or restricted to a bounded interval (Haussdorff). The algorithm presented in this paper fits into the Hamburger moment problem as it allows for X with support at both tails.

In case a sequence of moments is known to correspond to at least one probability measure, an important follow-up question involves what information on the distribution is contained in this sequence. Research on this question has focused on several characteristics:

(i) class of probability measure (Berg 1995; Gut 2002; Lin 1997; Pakes et al. 2001; Stoyanov 2000)
(ii) tail of the distribution (Goria and Tagliani 2003; Lindsay and Basak 2000)
(iii) mode (Gavriliadis 2008)
(iv) cumulative distribution function (Mnatsakanov 2008b; Mnatsakanov and Hakobyan 2009)
(v) density function (Gavriliadis and Athanassoulis 2009; Mnatsakanov 2008a)
(vi) Shannon entropy (Milev et al. 2012)

The above-mentioned literature can differ in the admissible support of the probability measure.

By using a sequence of moments, this paper adds to the literature on the moment problem by presenting an iterative algorithm that bounds the partial moments (1) at $a = 0$. More specifically, the algorithm bounds the partial moments from the positive part, $\max(X, 0)$, and the negative part, $\max(-X, 0)$, of a univariate real-valued random variable X by using information contained in known (ratios of) subsequent finite moments of X. In case one is interested in the partial moments at $a \neq 0$, obtain moments of $X_a := X - a$ from the sequence of moments of X and bound the moments of $\max(X_a, 0)$ and $\max(-X_a, 0)$.

The bounds on the partial moments imply bounds on unobserved higher order moments of X. Using a Taylor expansion, one can obtain bounds on the moments of the transformation $f(X)$ for a certain function $f : \mathbb{R} \to \mathbb{R}$.

Where applicable, procedures from related literature can be added to our algorithm. If the random variable X belongs to the class of Pearson distributions,[1] a recursive relationship on partial moments is in Winkler et al. (1972). In case X has a known finite support, the information on the support implies additional bounds (Barnett et al. 2002) and suitable numerical optimization routines can be invoked (Dokov and Morton 2005; Frauendorfer 1988; Kall 1991). Because the sequence of known moments of X typically consists of more than six moments, an unrestricted support can hamper a numerical optimization routine that restricts the sequence of moments.

The paper proceeds as follows. Preliminaries are in Sect. 2. Section 3 presents the iterative algorithm that bounds partial moments, and how these bounds imply bounds on other expressions. In Sect. 4, the algorithm is demonstrated for two example distributions. Section 5 is reserved for concluding remarks and points of discussion.

---

[1] The density function of a Pearson distribution is proportional to $\exp\left(-\int \frac{x-a}{b_0+b_1 x+b_2 x^2}\, dx\right)$ with constants a, b_0, b_1, and b_2. Example distributions include the normal, beta, uniform, Gamma, Chi-squared, Student's t, and Pareto distribution.

2 Preliminaries

Let $\mathbb{R}_0^+ := [0, \infty)$ denote the set of nonnegative real numbers in the set $\mathbb{R}$ of real numbers. The sets $\mathbb{N}_0^+ = \{0, 1, \ldots\}$ and $\mathbb{N}^+ = \{1, 2, \ldots\}$ denote sets of natural numbers. Let $X : \Omega \to \mathbb{R}$ be a univariate real-valued random variable defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with unknown cumulative distribution function (cdf) $F : \mathbb{R} \to [0, 1]$.

The known moments $\mu_1, \ldots, \mu_n$ of X are finite. Assume $\mu_2 > 0$ to avoid trivial outcomes where all moments are zero. Denote the positive and negative part of X by $X^+ := \max(X, 0)$ and $X^- := \max(-X, 0)$, respectively. Let μ_i^+ and μ_i^- represent the i-th moment of X^+ and X^-, respectively ($i \in \mathbb{N}^+$). Thus, $\mu_i^+ := \mu_i^+(0)$ and $\mu_i^- := \mu_i^-(0)$ denote the partial moments of X at $a = 0$ in (1). Set $\mu_0^+ := \mathbb{P}(X > 0)$, $\mu_0^- := \mathbb{P}(X < 0)$, and $\mu_0 := \mathbb{P}(X \neq 0)$. The latter definition is a convention that enables tighter bounds.

For a quasi-degenerate random variable X, there exists $x \in \mathbb{R}$ with $\mathbb{P}(X \in \{0, x\}) = 1$. A nonquasi-degenerate (nqd) random variable X is a random variable that is not quasi-degenerate, thus $\mathbb{P}(X \in \{0, x\}) < 1$ for all $x \in \mathbb{R}$. A degenerate X is necessarily quasi-degenerate, while a nqd X is necessarily non-degenerate.

Define the i-th moment ratio of X as $m_i := \mu_i/\mu_{i-1}$ with $m_i = \infty$ if $\mu_{i-1} = 0$ ($i \in \mathbb{N}^+$). Let m_i^+ and m_i^- refer to the i-th moment ratio of the positive part X^+ and the negative part X^-, respectively. Both ratios can be referred to as a partial moment ratio of X. Lower bounds and upper bounds are denoted by underlines and bars, respectively. For instance, $\underline{\mu_i^+}$ is a lower bound on μ_i^+ and $\bar{m}_i^-$ is an upper bound on $m_i^- := \mu_i^-/\mu_{i-1}^-$. The negation of $s \in \{+, -\}$ is denoted by $-s$. The operation $s\,x$ denotes x if $s = +$ and $-x$ if $s = -$. The operation s/x means $1/x$ if $s = +$ and $-1/x$ if $s = -$. The equivalence $X \equiv Y$ means that $X = Y$ holds almost surely: $\mathbb{P}(X = Y) = 1$.

3 The algorithm

This section outlines the algorithm in a number of lemmas and a theorem. By using some simple constraints, the first lemma provides bounds on partial moments and partial moment ratios. This enables an initialization of the bounds.

Lemma 1 [Initialization I] *For $s \in \{+, -\}$,*

$$0 \leq \mu_{2n}^s \leq \mu_{2n} \qquad\qquad\qquad\qquad n \in \mathbb{N}_0^+ \qquad (2)$$

$$\mu_{2n-1}^s \geq s\,\mu_{2n-1} \qquad m_{2n-1}^s \geq s\,m_{2n-1} \qquad \frac{1}{m_{2n}^s} \geq \frac{s}{m_{2n}} \qquad n \in \mathbb{N}^+. \qquad (3)$$

(i) *The nonnegativity constraint in (2) holds with equality if and only if $X^s \equiv 0$.*
   The constraint $\mu_{2n}^s \leq \mu_{2n}$ in (2) holds with equality if and only if $X^{-s} \equiv 0$.
(ii) *Each of the constraints in (3) holds with equality if and only if $X^{-s} \equiv 0$.*

Consider the Christoffel function associated with cdf F of a non-degenerate X:

$$\lambda_k(x) := \frac{1}{\sum_{j=0}^{k} \left[P_j(x) \right]^2} \qquad\qquad k \in \mathbb{N}^+,$$

where $P_0(x) \equiv 1$ and

$$P_j(x) := \frac{D_j(x)}{\sqrt{H_{2j} H_{2j-2}}} \qquad\qquad D_j(x) := \begin{vmatrix} 1 & \mu_1 & \cdots & \mu_j \\ \mu_1 & \mu_2 & \cdots & \mu_{j+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{j-1} & \mu_j & \cdots & \mu_{2j-1} \\ 1 & x & \cdots & x^j \end{vmatrix} \qquad H_0 := 1$$

$$H_{2j} := \begin{vmatrix} 1 & \mu_1 & \cdots & \mu_j \\ \mu_1 & \mu_2 & \cdots & \mu_{j+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_j & \mu_{j+1} & \cdots & \mu_{2j} \end{vmatrix} \qquad\qquad j \in \mathbb{N}^+.$$

Let

$$L_j^{(i)} := 1 - \sum_{l=i}^{j} \lambda_j \left(x_j^{(l)} \right) \quad U_j^{(i)} := \sum_{l=1}^{i} \lambda_j \left(x_j^{(l)} \right) \quad i = 1, 2, \ldots, j \quad j \in \mathbb{N}^+,$$

where $x_j^{(1)} \leq \cdots \leq x_j^{(j)}$ are the j zeros of the j-th degree polynomial P_j.

Using the definitions above, the following lemma from Gavriliadis and Athanassoulis (2009) bounds the distribution function at certain points.

Lemma 2 *Suppose the moments $\mu_1, \ldots, \mu_{2n}$ of a non-degenerate X are known ($n \in \mathbb{N}^+$). The cumulative distribution function F satisfies at the j zeros $x_j^{(1)}, \ldots, x_j^{(j)}$ of polynomial P_j,*

$$L_j^{(i)} \leq F \left(x_j^{(i)} \right) \leq U_j^{(i)} \qquad\qquad j = 1, \ldots, n.$$

Let $\tilde{x}^{(1)} < \cdots < \tilde{x}^{(N)}$ represent the strictly increasing sequence of the N distinct zeros of the set of polynomials $\{P_j\}_{j=1}^{n}$:

$$\bigcup_{j=1}^{N} \tilde{x}^{(j)} = \bigcup_{j=1}^{n} \bigcup_{i=1}^{j} x_j^{(i)}$$

Define the increasing step functions

$$L(x) = \begin{cases} 0 & x < \tilde{x}^{(1)} \\ \max_{i,j} \{ L_j^{(i)} : x_j^{(i)} \leq x \} & x \geq \tilde{x}^{(1)} \end{cases}$$

$$U(x) = \begin{cases} \min_{i,j}\{U_j^{(i)} : x_j^{(i)} \geq x\} & x \leq \tilde{x}^{(N)} \\ 1 & x > \tilde{x}^{(N)}. \end{cases}$$

Notice that for $L(x)$ (respectively $U(x)$), one can simply take for each j the largest (smallest) i that satisfies $x_j^{(i)} \leq x$ ($x_j^{(i)} \geq x$) since $L_j^{(i)}$ ($U_j^{(i)}$) is nondecreasing with i. It follows from Lemma 2 that $L(x) \leq F(x) \leq U(x)$ holds for all $x \in \mathbb{R}$. Using these bounds on the cdf $F(x)$, Lemma 3 produces lower bounds on the partial moments:

Lemma 3 (Initialization II) *Given the notation above,*

$$\mu_i^+ \geq \sum_{j=j_0^+}^{N} \left(\tilde{x}^{(j)}\right)^i dU_j \qquad \mu_i^- \geq \sum_{j=1}^{j_0^+ - 1} \left(-\tilde{x}^{(j)}\right)^i dL_j \qquad i \in \mathbb{N}_0^+, \quad (4)$$

where $j_0^+ = \arg\min_{j \in \{1,\dots,N\}}\{\tilde{x}^{(j)} > 0\}$, $dL_1 = L(\tilde{x}^{(1)})$, $dU_N = 1 - U(\tilde{x}^{(N)})$, *and*

$$dL_j = L(\tilde{x}^{(j)}) - L(\tilde{x}^{(j-1)}) \quad dU_{j-1} = U(\tilde{x}^{(j)}) - U(\tilde{x}^{(j-1)}) \quad j = 2, \dots, N.$$

Upper bounds on the partial moments μ_i^+ and μ_i^- are not available along the lines in Lemma 3. For instance, the summation $\sum_{j=j_0^+}^{N} \left(\tilde{x}^{(j)}\right)^i dL_j$ is not an upper bound on μ_i^+ since dL_N is not a bound for dF on $(x^{(N)}, \infty)$. The iterative algorithm we develop derives upper bounds on μ_i^+ and μ_i^- under certain assumptions. In addition, the algorithm can sharpen the lower bounds obtained in Lemma 3.

The next lemma is useful to apply in the iterative algorithm as well as for proving some inequalities.

Lemma 4 (Moment ordering)

(i) *For nonnegative X and $n \geq 2$, or for arbitrary X and even $n \geq 2$,*

$$\mu_{n-1}^2 \leq \mu_n \mu_{n-2}.$$

(ii) *For nonnegative X with $\mu_1 > 0$ the moment ratio $m_i = \mu_i/\mu_{i-1}$ increases with $i \in \mathbb{N}^+$:*

$$0 < m_1 \leq m_2 \leq m_3 \leq \cdots.$$

The inequalities are all strict if and only if X is nqd.

Lemma 5 produces bounds on the moment ratio of certain summations of nonnegative random variables.

Lemma 5 *Consider $Y = \sum_{j \in J} Y_j$ where each $Y_j : \Omega \to \mathbb{R}_0^+$ is a nonnegative random variable with strictly positive first moment. Let $i \in \mathbb{N}^+$.*

(i) *Suppose the following two conditions both hold*

(a) *Y is a mixture*
(b) *for each $j \in J$, $m_i(Y_j) = g_i h_j$ and $m_{i+1}(Y_j) = g_{i+1} h_j$ where $g_i, g_{i+1}, h_j \in$*
   $\mathbb{R}$,
   then

$$\max_{j \in J} \frac{m_{i+1}(Y_j)}{m_i(Y_j)} \leq \frac{m_{i+1}(Y)}{m_i(Y)}. \tag{5}$$

(ii) *If Y is a mixture, i.e., for almost all outcomes $\omega \in \Omega$ at most one $Y_j(\omega)$ is nonzero, then*

$$m_i(Y) \leq \max_{j \in J} m_i(Y_j). \tag{6}$$

(iii) *If each Y_j is independent,*

$$m_i(Y) \leq \sum_{j \in J} m_i(Y_j). \tag{7}$$

(iv) *Suppose the following two conditions both hold*

(a) *each Y_j is independent*
(b) *each Y_j is logconcave,*

then for each nonempty $I \subseteq J$, $\sum_{j \in I} Y_j$ is logconcave and

$$m_i \left(\sum_{j \in I} Y_j \right) \leq m_i(Y). \tag{8}$$

Sufficient condition (*b*) in (*i*) reflects that the moment ratio $m_i(Y_j)$ must be separable into two parts with a common functional form across i and j. Nonnegative distributions with moment ratios that satisfy this condition include the Gamma distribution and distributions with a single, scalar parameter such as the half-normal distribution and uniform distributions of the type $U[0, b]$ with $b > 0$. The moment ratio of the lognormal distribution has not a separable form.

It is natural to inquire about the necessity of the sufficient conditions in Lemma 5. We provide for each condition a counterexample where the sufficient condition is not satisfied and the corresponding inequality fails to hold.

(i) (a) Distributions that violate sufficient condition (*a*) may not satisfy (5) as one can verify using $Y = Y_1 + Y_2$ with independent $Y_1, Y_2 \sim \exp(1)$.
    (b) For instance, the mixture Y with $\mathbb{P}(Y = Y_1) = \mathbb{P}(Y = Y_2) = \frac{1}{2}$ where $Y_1 \equiv 1$ and $Y_2 \sim \exp(1)$ satisfies (a), but violates (b). Inequality (5) is invalid here since $m_2(Y_2)/m_1(Y_2) = 2 > \frac{3}{2} = m_2(Y)/m_1(Y)$.
(ii) Equation (6) can be invalid if the mixture condition on Y is dropped. This follows for the case where $Y_1 \equiv Y_2 \equiv 1$.

(iii) Dependence between different Y_j can make inequality (7) invalid. For instance, the case where $\mathbb{P}\big((Y_1, Y_2) = \big(0, \frac{1}{2}\big)\big) = \mathbb{P}\big((Y_1, Y_2) = (1, 1)\big) = \frac{1}{2}$ leads to $m_3(Y_1) + m_3(Y_2) = 1.9 < 1.91\cdots = m_3(Y)$.

(iv) (a) Consider $Y = Y_1 + Y_2$ with the log-concave distributions $Y_1 \sim \text{Exp}(1)$ and $Y_2 \sim \text{Exp}\big(\frac{1}{2}\big)$.

To violate (a), impose a perfect negative correlation on the quantiles $F_{Y_1}(Y_1)$ and $F_{Y_2}(Y_2)$ by letting $F_{Y_1}(Y_1) = 1 - F_{Y_2}(Y_2)$. In other words, $(Y_1, Y_2) \sim \{(-\ln(1-U), -2\ln(U)) : U \sim U[0,1]\}$ such that $\mathbb{E}\big[Y^2\big] = \int_0^1 \ln^2\big(u^2(1-u)\big)\,\mathrm{d}u = 18 - \frac{2}{3}\pi^2$. It follows that $m_2(Y) = 6 - \frac{2}{9}\pi^2 < 4 = m_2(Y_2)$.

(b) For the independent, non-logconcave case where $\mathbb{P}(Y_1 = 0) = \frac{9}{10}$, $\mathbb{P}(Y_1 = 10) = \frac{1}{10}$, and $Y_2 \equiv 1$: $m_2(Y_1 + Y_2) = \frac{13}{2} < 10 = m_2(Y_1)$.

The following lemma can be supplemented with bounds from Lemma 5. This is helpful in the examples in Sect. 4.

Lemma 6 (Initialization III) *Let $i \in \mathbb{N}^+$.*

(i) *Suppose $X = \sum_{j \in I} X_j$ is a mixture of random variables $X_j : \Omega \to \mathbb{R}$, i.e., for almost all outcomes $\omega \in \Omega$ at most one $X_j(\omega)$ is nonzero, then*

$$m_i^+ := m_i(X^+) \le \max_{j \in I} m_i(X_j^+). \tag{9}$$

(ii) *If $X = Y - Z$ with $Y = \sum_j Y_j$ and the following conditions hold*
(a) Y and Z are independent and nonnegative random variables
(b) each Y_j is independent
(c) each Y_j is log-concave,[2]
then

$$m_i(X^+) \le m_i(Y). \tag{10}$$

(iii) *If $X = Y - Z$ is a mixture of nonnegative random variables Y and Z, i.e., $\mathbb{P}(\{Y = 0\} \cup \{Z = 0\}) = 1$, then $X^+ \equiv Y$ and*

$$m_i(X^+) = m_i(Y). \tag{11}$$

Similar relations apply to $m_i^- := m_i(X^-)$ and $m_i(Z)$.

Condition (c) in Lemma 6(ii) on logconcave distributions is satisfied by many univariate distributions. Examples include the exponential distribution, the normal distribution, the uniform distribution, and the Gamma distribution with shape parameter at least one. The Gamma distribution with shape parameter less than one, the Student's t-distribution, and the log-normal distribution are not log-concave.

We provide for each sufficient condition in Lemma 6 an example where the condition is not satisfied and the corresponding inequality is violated.

---

[2] The density function of a log-concave distribution is proportional to $\exp(\psi(y))$ with ψ a concave function on the domain of the distribution. Log-concave distributions are necessarily unimodal.

(i) Inequality (9) fails to hold for the nonmixture case where $X \equiv 2$ with $X_1 \equiv X_2 \equiv 1$. Here, X, is not a mixture and $m_i(X^+) = 2 > 1 = \max(m_i(X_1), m_i(X_2))$ where $i \in \mathbb{N}^+$.

(ii) (a) Suppose $Y \sim U[0, 1]$ and $Z = Y 1_{Y < \frac{1}{2}}$. The independence assumption (a) fails to hold, while assumptions (b) and (c) are satisfied. Inequality (10) is not satisfied: $m_2(X^+) = \frac{7}{9} > \frac{2}{3} = m_2(Y)$.

(b) Consider again the counterexample of Lemma 5(iv) where $Y = Y_1 + Y_2$ with the log-concave distributions $Y_1 \sim \text{Exp}(1)$ and $Y_2 \sim \text{Exp}\left(\frac{1}{2}\right)$, and let $Z \equiv 5$. This satisfies conditions (a) and (c). To violate (b), impose a perfect negative correlation on the quantiles $F_{Y_1}(Y_1)$ and $F_{Y_2}(Y_2)$ by letting $F_{Y_1}(Y_1) = 1 - F_{Y_2}(Y_2)$. It can be verified that $m_2(X^+) = 3.88 \cdots > 3.81 \cdots = 6 - \frac{2}{9}\pi^2 = m_2(Y)$.

(c) Suppose $\mathbb{P}(Y = 1) = \frac{9}{10}$, $\mathbb{P}(Y = 9) = \frac{1}{10}$, and $Z \equiv 1$. While conditions (a) and (b) are satisfied, condition (c) on logconcavity of each Y_j fails to hold. Inequality (10) also fails to hold: $m_2(X^+) = 8 > 5 = m_2(Y)$.

(iii) The nonmixture case with $Y \equiv 2$ and $Z \equiv 1$ corresponds for each $i \in \mathbb{N}^+$ to $X \equiv X^+ \equiv 1$ and $m_i^+ = 1 \neq 2 = m_i(Y)$.

The next theorem presents the theoretical novelty of the paper. For a nqd X, the moment μ_i^s is bounded in terms of the bounds on m_i^s, which are in turn in the proof derived from the bounds on the moments of X^{-s}. The requirement of a nqd X can be easily verified, since the moments of a quasi-degenerate distribution are uniquely characterized by finite $m_1 = m_2 = \ldots$ and $\mu_i = (\mu_n)^{i/n}$ ($i \in \mathbb{N}_0^+$). To identify a quasi-degenerate X, it therefore suffices if $\mu_i = (\mu_n)^{i/n}$ holds for some odd i and even n. The requirements on i and n exclude X that are restricted to the set $\{-x, 0, x\}$ for certain $x \in \mathbb{R}$.

Theorem 1 (μ_i^s in terms of the bounds on m_i^s) *Consider a nqd X with known moments* $\mu_{n-2}, \mu_{n-1}, \mu_n$ *with* $n \geq 2$ *even. Define*

$$a := \mu_n \mu_{n-2} - \mu_{n-1}^2 \qquad\qquad b := \sqrt{\frac{\mu_n}{\mu_{n-2}}} \qquad (12)$$

with constants ($s \in \{+, -\}$)

$$u^s(x) := \frac{a}{\mu_n/x - 2s\,\mu_{n-1} + x\mu_n} \qquad v^s(c, d) := \begin{cases} u^s(b) & d < b < c \\ \min\{u^s(c), u^s(d)\} & c \leq b \leq d \\ u^s(d) & b < \min(c, d) \\ u^s(c) & b > \max(c, d). \end{cases} \qquad (13)$$

Assume for the bounds on the partial moment ratios

(Lemma 1)	$\underline{m}_{n-1}^s \geq s\, m_{n-1}$	$1/\bar{m}_n^s \geq s/m_n$ (14)
(Lemma 4(ii))	$\underline{m}_{n-1}^s \leq \underline{m}_n^s$	$\bar{m}_{n-1}^s \leq \bar{m}_n^s.$ (15)

The moments of X^s are bounded by

$$\mu_{n-2}^s \leq \frac{1}{\underline{m}_{n-1}^s} v^s(\underline{m}_{n-1}^s, \underline{m}_n^s) \quad \mu_{n-1}^s \leq v^s(\bar{m}_{n-1}^s, \underline{m}_n^s) \quad \mu_n^s \leq \bar{m}_n^s v^s(\bar{m}_{n-1}^s, \bar{m}_n^s).$$

(16)

The next lemma bounds μ_i^s using previously obtained bounds.

Lemma 7 (μ_i^s *in terms of the bounds on* m_j^s, μ_j^s, *and* μ_j^{-s}) *For* $i \in \mathbb{N}^+$ *and* $s \in \{+, -\}$,

$$\mu_{i+1}^s \geq \frac{\left(\underline{\mu}_i^s\right)^2}{\bar{\mu}_{i-1}^s}$$

(17)

$$\mu_i^s \geq \max_{j<i} \left(\underline{\mu}_j^s\right)^{i/j} \qquad \mu_i^s \geq \max\left(\frac{\left(\underline{\mu}_{i+1}^s\right)^2}{\bar{\mu}_{i+2}^s}, \underline{m}_i^s \underline{\mu}_{i-1}^s, \frac{\underline{\mu}_{i+1}^s}{\bar{m}_{i+1}^s}\right)$$

(18)

$$\mu_i^s \leq \min_{j>i} \left(\bar{\mu}_j^s\right)^{i/j} \qquad \mu_i^s \leq \min\left(\sqrt{\bar{\mu}_{i-1}^s \bar{\mu}_{i+1}^s}, \bar{m}_i^s \bar{\mu}_{i-1}^s, \frac{\bar{\mu}_{i+1}^s}{\underline{m}_{i+1}^s}\right)$$

(19)

$$\mu_{2i-1}^s \geq \underline{\mu}_{2i-1}^{-s} + s\mu_{2i-1} \qquad \mu_{2i-1}^s \leq \bar{\mu}_{2i-1}^{-s} + s\mu_{2i-1}$$

(20)

$$\mu_{2i}^s \geq \mu_{2i} - \bar{\mu}_{2i}^{-s} \qquad \mu_{2i}^s \leq \mu_{2i} - \underline{\mu}_{2i}^{-s}.$$

(21)

Lemma 8 bounds the moment ratios in terms of the bounds on the moments.

Lemma 8 (m_i^s *in terms of the bounds on* μ_i^s)

$$\frac{\underline{\mu}_i^s}{\bar{\mu}_{i-1}^s} \leq m_i^s \leq \frac{\bar{\mu}_i^s}{\underline{\mu}_{i-1}^s} \qquad\qquad i \in \mathbb{N}^+$$

(22)

$$\sqrt{\frac{\underline{\mu}_i^s}{\bar{\mu}_{i-2}^s}} \leq m_i^s \leq \sqrt{\frac{\bar{\mu}_{i+1}^s}{\underline{\mu}_{i-1}^s}} \qquad\qquad i = 2, 3, \ldots.$$

(23)

The bounds in (22) *hold with equality if and only if the bounds on* μ_{i-1} *and* μ_i *hold with equality. The bounds in* (23) *hold with equality if and only if the bounds on* μ_j^s ($j \in \{i-2, \ldots, i+1\}$) *hold with equality and* X^s *is quasi-degenerate.*

The algorithm below is the iterative algorithm that bounds the partial moments μ_i^+ and μ_i^- and the corresponding moment ratios $m_i^+ = \mu_i^+ / \mu_{i-1}^+$ and $m_i^- = \mu_i^- / \mu_{i-1}^-$.

Algorithm 1 (*Partial moments*) The moments $\mu_1, \ldots, \mu_n$ of the random variable X are known. Before proceeding to the next step, each step is executed for $i = 1, \ldots, n$ [and also for $i = 0$ in (iii)–(iv)]. The index j denotes an index that may depend on i.

(i) Initialize bounds on m_i^+ and m_i^- (Lemma 1, 3, and 6).
(ii) Bound m_i^+ and m_i^- in terms of the bounds on m_j^+ and m_j^- (Lemma 4(ii)).
(iii) Bound μ_i^+ and μ_i^- in terms of the bounds on m_j^+ and m_j^- (Theorem 1).

(iv) Bound μ_i^+ and μ_i^- in terms of the bounds on μ_j^+, μ_j^-, m_j^+, and m_j^- (Lemma 7).

(v) Bound m_i^+ and m_i^- in terms of the bounds on μ_j^+ and μ_j^- (Lemma 8).

(vi) Stop if the change in the bounds since step (ii) is smaller than some predetermined tolerance, otherwise go to step (ii).

Within each step, apply all applicable inequalities. The different assumptions in particularly Lemmas 5 and 6 must be checked in advance. Some may be valid for a given random variable X, while some others may not.

The following algorithm applies Lemma 6(iii) and is helpful for bounding unobserved higher order moments.

Algorithm 2 (*Extrapolation*) The moments $\mu_1, \ldots, \mu_n$ of $X = Y - Z$ are known with $\mathbb{P}(\{Y = 0\} \cup \{Z = 0\}) = 1$. Execute Algorithm 1 to obtain bounds on $\mu_0^s, \ldots, \mu_n^s$ ($s \in \{+, -\}$). For a certain $v \in \mathbb{N}^+$, construct bounds on μ_i^s using bounds on m_i^s where $i = n + 1, \ldots, n + v$:

(i) *Lower bounds* If (a) X^s is a mixture, and (b) $m_i(X_j^s) = g_i h_j$ where $i > n$ and $g_i, h_j \in \mathbb{R}$: apply Lemma 5(i)
   else apply Lemma 4(ii).

(ii) *Upper bounds* If X^s is a mixture: apply Lemma 5(ii)
   else if $X^s = \sum_j X_j^s$ with each X_j^s independent: apply Lemma 5(iii).

Conditions on X^s are needed for an upper bound on partial moments μ_i^s ($i > n$) in Algorithm 2(ii). To see this, the tail of any X^s can admit $\mathbb{P}(X^s = x) = 1/x^{i-\varepsilon}$ for arbitrarily large x and $\varepsilon \in (0, i)$. By Markov's inequality, the probability mass at x of such X^s has unbounded impact on moments $i \in \mathbb{N}^+$ and higher if $x \to \infty$:

$$\mathbb{E}\left[\left(X^s\right)^i\right] \geq x^i \, \mathbb{P}\left(\left(X^s\right)^i \geq x^i\right) \geq \frac{x^i}{x^{i-\varepsilon}} = x^\varepsilon \to \infty$$

The bounds on the moments of X^+ and X^- from Algorithm 2 provide bounds on the unobserved higher order moments of X through $\mu_i = \mu_i^+ + (-1)^i \mu_i^-$ with $i = n+1, n+2, \ldots$. This can be used in a Taylor expansion of the transformation $f(X)$. For an expansion of $f(X)$ around a nonzero a, one can consider using $\max(X_a, 0)$ and $\max(-X_a, 0)$ with $X_a := X - a$.

4 Examples

The example in Sect. 4.1 reports bounds on partial moments for a case using Algorithm 1. Section 4.2 contains an example that uses Algorithm 1 and bounds a Taylor series expansion by extrapolating the results to higher order moments.

4.1 Example 1: a sum of random variables

Consider $X = \frac{1}{2}T - U + Z$ with T an exponential distribution with intensity one, U a uniform distribution on $[0, 1]$, and Z a standard normal distribution. The variables

T, U, and Z are independent. Since the odd moments of Z are zero, the moments μ_i of X can be obtained from the expansion

$$\mathbb{E}\left[X^i\right] = \sum_{j=0}^{\lfloor i/2 \rfloor} \sum_{k=0}^{i-2j} \binom{i}{i-2j-k,\, k,\, 2j} \mathbb{E}\left[\left(\frac{1}{2}T\right)^{i-2j-k}\right] \mathbb{E}\left[(-U)^k\right] \mathbb{E}\left[Z^{2j}\right] \tag{24}$$

where $i \in \mathbb{N}_0^+$, $\lfloor i/2 \rfloor$ denotes the largest natural number not greater than $i/2$ ($\lfloor i/2 \rfloor = \max_{j \in \mathbb{N}_0^+}\{j : j \leq i/2\}$), and

$$\mathbb{E}\left[T^i\right] = i! \qquad \mathbb{E}\left[U^k\right] = \frac{1}{1+k} \qquad \mathbb{E}\left[Z^{2j}\right] = \frac{(2j)!}{2^j (j!)}. \tag{25}$$

We obtain the i-th moment of X from (24) for $i = 0, 1, \ldots, 16$.

We are interested in bounding the moments of $X^+ := \max(X, 0)$ and $X^- := \max(-X, 0)$. Write $X = \left(\frac{1}{2}T + Z^+\right) - (U + Z^-)$, where the components T, U, Z^+, and Z^- are logconcave distributions. Notice that (i) $X^+ \not\equiv \frac{1}{2}T + Z^+$ and $X^- \not\equiv U + Z^-$ and (ii) the convolutions $\frac{1}{2}T + Z^+$ and $U + Z^-$ are logconcave (Lemma 5(iv)).

Consider the independent distributions (i) $T_1, T_2 \overset{d}{=} T$, (ii) $U_1, U_2 \overset{d}{=} U$, and (iii) $Z_1, Z_2 \overset{d}{=} Z$. Since $\mathbb{P}(Z \geq 0) = \frac{1}{2}$, the distribution of X is equal in distribution to a mixture of two distributions:

$$X \overset{d}{=} \begin{cases} \left(\frac{1}{2}T_0 + |Z_0|\right) - U_0 & \text{with probability } \frac{1}{2} \\ \frac{1}{2}T_1 - (U_1 + |Z_1|) & \text{with probability } \frac{1}{2} \end{cases} \tag{26}$$

Both components in (26) consist of three independent log-concave distributions. Upper bounds on $m_i^+ := m_i^+(X)$ and $m_i^- := m_i^-(X)$ are initialized by applying Lemma 6(i)–(ii) and then Lemma 5(iv),

$$m_i^+ \leq \max\left(m_i\left(\frac{1}{2}T + |Z|\right), m_i\left(|Z|\right)\right) = m_i\left(\frac{1}{2}T + |Z|\right) \tag{27}$$

$$m_i^- \leq \max\left(m_i\left(U\right), m_i\left(U + |Z|\right)\right) = m_i\left(U + |Z|\right). \tag{28}$$

The initial moment ratios in (27)–(28) follow from

$$\mathbb{E}\left[\left(\frac{1}{2}T + |Z|\right)^i\right] = \sum_{j=0}^{i} \binom{i}{j} \mathbb{E}\left[\left(\frac{1}{2}T\right)^{i-j}\right] \mathbb{E}\left[|Z|^j\right]$$

$$\mathbb{E}\left[(U + |Z|)^i\right] = \sum_{j=0}^{i} \binom{i}{j} \mathbb{E}\left[U^{i-j}\right] \mathbb{E}\left[|Z|^j\right], \tag{29}$$

where the moments of T and U are in (25), and

$$\mathbb{E}\left[|Z|^j\right] = \sqrt{\frac{2^j}{\pi}}\Gamma\left(\frac{j+1}{2}\right).$$

As a measure of convergence after iteration k, consider the change in the width of the error bounds:

$$\varepsilon_k := \max_{\substack{i=0,\dots,16 \\ s\in\{+,-\}}}\left\{1 - \frac{\bar{\mu}_{i,k}^s - \underline{\mu}_{i,k}^s}{\bar{\mu}_{i,k-1}^s - \underline{\mu}_{i,k-1}^s}\right\} \quad k\in\mathbb{N}^+, \tag{30}$$

where the first subscript i of each moment bound refers to the considered order of the partial moments, and the second subscript k refers to the iteration number with iteration 0 immediately after the initialization. A small ε_k indicates a small improvement in the error bounds.

Using $\varepsilon_k < 10^{-10}$ as a stopping condition, Algorithm 1 stops after 83 iterations and 151 ms of CPU time.[3] Table 1 reports the bounds on the moments μ_i^+ of X^+ and μ_i^- of X^-. By (20)–(21), the difference $\bar{\mu}_i^s - \underline{\mu}_i^s$ is independent of the sign $s\in\{+,-\}$ (3rd column). This difference relative to the moment μ_i tends to decrease with the order i of the moment (4th column). This can be helpful for bounding Taylor expansions because the moments of the highest observed order are most important for bounding unobserved higher order moments of X (here, the moments greater than order sixteen).

The difference between the bounds relative to $\underline{\mu}_i^+$ is also decreasing with the order (7th column). In contrast, the difference between the bounds relative to $\underline{\mu}_i^-$ increases for the higher order moments to 43.7% (rightmost column). This higher percentage may reflect that the sequence of 16 moments cannot uniquely pin down the characteristics of particularly X^-. The reason is that X^+ tends to be larger than X^-, as indicated by the moment bounds. As such, the higher order moments of X are mainly determined by the higher order moments of X^+.

The convergence by iteration is for the partial moments of order 10 in Table 2. Here, 25 iterations suffice for having each of the bounds with a difference less than 0.03% of the bounds after the final iteration on the 10-th moment. The final bounds on μ_{10}^+ and μ_{10}^- differ from each other by 2.2 and 22.0%, respectively. Table 3 indicates that for this example, executing 25 iterations gives bounds with a difference of at most 0.14% to the final bounds on the 16-th partial moments of X. The relative difference between the final bounds on μ_{16}^+ is, by coincidence, also 0.14%.

Figure 1 depicts the series ε_k as a function of the iteration k. The linear trend in ε_k suggests that for some $\tilde{a}, \tilde{b}\in\mathbb{R}$,

$$\log(\varepsilon_k) \approx \tilde{a} + \tilde{b}k. \tag{31}$$

A small number of numerical experiments suggest that the exponential convergence of ε_k in (31) might hold in general. A proof of this conjecture is beyond the scope of this paper.

---

[3] All computations are on an i5-6300 CPU with 2.4 GHz dual core and 16 GB RAM.

Table 1 Moments of X from (24) and final bounds on partial moments of X from X^s from Algorithm 1 ($s \in \{+, -\}$)

i	μ_i	$\bar{\mu}_i^s - \underline{\mu}_i^s$	$\frac{\bar{\mu}_i^s - \underline{\mu}_i^s}{\mu_i}$ (%)	$\underline{\mu}_i^+$	$\bar{\mu}_i^+$	$\frac{\bar{\mu}_i^+ - \underline{\mu}_i^+}{\underline{\mu}_i^+}$ (%)	$\underline{\mu}_i^-$	$\bar{\mu}_i^-$	$\frac{\bar{\mu}_i^- - \underline{\mu}_i^-}{\underline{\mu}_i^-}$ (%)
0	1	0.546103489	54.61	0.226948256	0.773051744	240.6	0.226948256	0.773051744	240.6
1	0	0.282797632	NA	0.294552637	0.577350269	96.01	0.294552637	0.577350269	96.0
2	1.333333333	0.351808693	26.39	0.521500893	0.873309586	67.46	0.460023748	0.811832441	76.5
3	0.25	0.356825789	142.7	1.144408357	1.501234146	31.18	0.894408357	1.251234146	39.9
4	5.7	0.823473982	14.45	2.969661482	3.793135463	27.73	1.906864537	2.730338518	43.2
5	4.083333333	1.369867367	33.55	8.671382353	10.04124972	15.80	4.58804902	5.957916387	29.9
6	45.39285714	4.041940508	8.90	29.34109989	33.3830404	13.78	12.00981674	16.05175725	33.7
7	76.5	8.981713213	11.74	110.8509285	119.8326417	8.10	34.35092848	43.3326417	26.1
8	597.1111111	28.15618167	4.72	466.6217147	494.7778964	6.03	102.3332147	130.4893964	27.5
9	1831.9	85.29274135	4.66	2,167.547328	2,252.840069	3.93	335.6473277	420.9400691	25.4
10	12,404.59091	244.8679167	1.97	11,046.69865	11,291.56656	2.22	1,113.024347	1,357.892263	22.0
11	57,412.16667	824.5595536	1.44	61,434.33867	62,258.89822	1.34	4,022.172	4,846.731554	20.5
12	388,204.0769	2,801.793702	0.72	370,904.6141	373,706.4078	0.76	14,497.66914	17,299.46284	19.3
13	2,362,230.429	10,625.03282	0.45	2,419,009.945	2,429,634.978	0.44	56,779.51632	67,404.54913	18.7
14	17,225,582.07	46,838.45573	0.27	16,962,951.18	17,009,789.64	0.28	215,792.4297	262,630.8854	21.7
15	126,430,633.6	196,502.6027	0.16	127,335,953.2	127,532,455.8	0.15	905,319.6441	1,101,822.247	21.7
16	1,023,903,852	1,450,140.450	0.14	1,019,135,849	1,020,585,990	0.14	3,317,861.951	4,768,002.401	43.7

Table 2 Bounds on the 10-th moment of X^s for different iterations of Algorithm 1 ($s \in \{+, -\}$)

Iter.	$\bar{\mu}_i^s - \underline{\mu}_i^s$	$\underline{\mu}_{10}^+$	$\bar{\mu}_{10}^+$	$\frac{\bar{\mu}_{10}^+ - \underline{\mu}_{10}^+}{\underline{\mu}_{10}^+}$ (%)	$\underline{\mu}_{10}^-$	$\bar{\mu}_{10}^-$	$\frac{\bar{\mu}_{10}^- - \underline{\mu}_{10}^-}{\underline{\mu}_{10}^-}$ (%)
Init		68.33	12,404.59	18,054.28	12.17	12,404.59	101,802.21
1	467.8618868	10,917.87	11,385.73	4.29	1018.86	1486.72	45.92
2	467.8618868	10,917.87	11,385.73	4.29	1018.86	1486.72	45.92
5	419.4245296	10,966.31	11,385.73	3.82	1018.86	1438.28	41.17
10	283.9040018	11,031.32	11,315.23	2.57	1089.36	1373.27	26.06
15	255.1186226	11,042.66	11,297.78	2.31	1106.81	1361.93	23.05
20	247.0662494	11,045.83	11,292.90	2.24	1111.69	1358.76	22.22
25	245.3784058	11,046.50	11,291.88	2.22	1112.71	1358.09	22.05
50	244.8681644	11,046.70	11,291.57	2.22	1113.02	1357.89	22.00
83	244.8679167	11,046.70	11,291.57	2.22	1113.02	1357.89	22.00

Table 3 Bounds on the 16-th moment of X^s for different iterations of Algorithm 1 ($s \in \{+, -\}$)

| Iter.	$\bar{\mu}_i^s - \underline{\mu}_i^s$	$\underline{\mu}_{16}^+$	$\bar{\mu}_{16}^+$	$\frac{\bar{\mu}_{16}^+ - \underline{\mu}_{16}^+}{\underline{\mu}_{16}^+}$ (%)	$\underline{\mu}_{16}^-$	$\bar{\mu}_{16}^-$	$\frac{\bar{\mu}_{16}^- - \underline{\mu}_{16}^-}{\underline{\mu}_{16}^-}$ (%)
Init		477,337.99	1,023,903,851.56	214,402.90	2,357.07	1,023,903,851.56	43,439,619.84
1	467.8618868	1,013,938,219.28	1,022,509,778.80	0.85	1,394,072.75	9,965,632.28	614.86
2	467.8618868	1,016,240,773.36	1,022,509,778.80	0.62	1,394,072.75	7,663,078.20	449.69
5	419.4245296	1,017,903,197.71	1,022,128,400.82	0.42	1,775,450.74	6,000,653.85	237.98
10	283.9040018	1,018,871,498.78	1,020,927,857.30	0.20	2,975,994.26	5,032,352.78	69.10
15	255.1186226	1,019,074,462.63	1,020,679,642.95	0.16	3,224,208.61	4,829,388.93	49.79
20	247.0662494	1,019,124,022.75	1,020,605,983.69	0.15	3,297,867.87	4,779,828.81	44.94
25	245.3784058	1,019,132,743.60	1,020,590,701.45	0.14	3,313,150.11	4,771,107.96	44.01
50	244.8681644	1,019,135,847.48	1,020,585,991.93	0.14	3,317,859.63	4,768,004.08	43.71
83	244.8679167	1,019,135,849.16	1,020,585,989.61	0.14	3,317,861.95	4,768,002.40	43.71
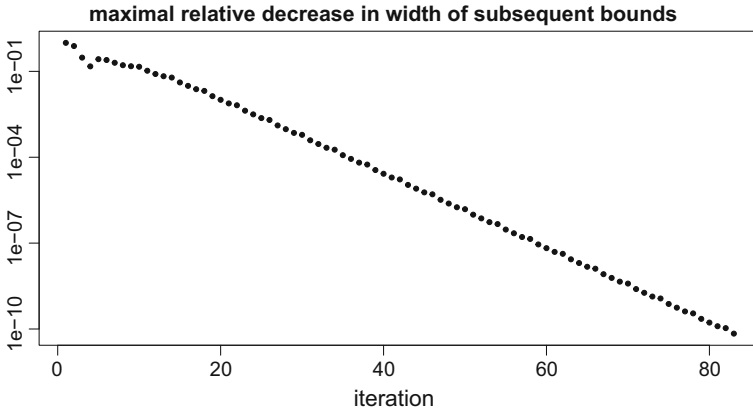
Fig. 1 Relative decrease ε_k in the width of the error bound $\bar{\mu}_i - \underline{\mu}_i$ of subsequent iterations (see (30)). Maximum is taken over moments $i = 0, 1, \ldots 16$

The product $\prod_{k=1}^{\infty} (1 - \varepsilon_k)$ is a measure for the cumulative decrease in the width of the error bound over all iterations. Provided (31) is the correct model and $\tilde{b} \leq 0$, the lower bound and upper bound do not converge to each other:

$$\log \prod_{k=1}^{\infty} (1 - \varepsilon_k) = \sum_{k=1}^{\infty} \log (1 - \varepsilon_k) \approx - \sum_{k=1}^{\infty} \varepsilon_k = - \sum_{k=1}^{\infty} e^{\tilde{a} + \tilde{b}k}$$

$$= -a \sum_{k=1}^{\infty} b^k = -\frac{ab}{1 - b} \neq 0, \tag{32}$$

where $a = e^{\tilde{a}}$ and $b = e^{\tilde{b}}$. An estimation of the linear regression (31) using $\varepsilon_{20}, \ldots, \varepsilon_{83}$ of this example gives $\tilde{a} = 1.45$ and $\tilde{b} = -0.299$. This predicts a small cumulative relative decrease in the width of the error bounds after iteration 83:

$$\log \prod_{k=84}^{\infty} (1 - \varepsilon_k) = -\frac{ab^{84}}{1 - b} \approx -2.05 \times 10^{-10} \tag{33}$$

Because ε_k is a maximum over different orders of moments, the value in (33) can be interpreted as an upper bound on the cumulative decrease of the partial moments $0, 1, \ldots, 16$.

In some cases, initial bounds on the moment ratios can be difficult to obtain. Instead of the initial bounds (27)–(28), suppose we were to initialize the upper bounds on the moment ratios by

$$m_i^+ \leq m_i \left(\frac{1}{2} T \right) + m_i(|Z|) \quad m_i^- \leq m_i(U) + m_i(|Z|) \quad i = 1, \ldots, 16. \tag{34}$$

Table 4 Initialization of the upper bounds $\bar{m}_i^+$ and $\bar{m}_i^-$ on the moment ratios using the bounds (27)–(28) and the weaker bounds in (34)

| i | $\bar{m}_i^+$ | | $\bar{m}_i^-$ | |
| | Table 1: (27) | Table 5: (34) | Table 1: (28) | Table 5: (34) |
	$m_i\left(\frac{1}{2}T + \lvert Z\rvert\right)$	$m_i\left(\frac{1}{2}T\right) + m_i\left(\lvert Z\rvert\right)$	$m_i\left(U + \lvert Z\rvert\right)$	$m_i\left(U\right) + m_i\left(\lvert Z\rvert\right)$
1	1.297884561	1.297884561	1.297884561	1.297884561
2	1.770484549	2.253314137	1.642070457	1.919980804
3	2.194451388	3.095769122	1.944265621	2.345769122
4	2.594931663	3.879971206	2.217710144	2.679971206
5	2.98780916	4.627692162	2.466624793	2.961025495
6	3.383670411	5.349964007	2.694189748	3.207106865
7	3.789507816	6.053230595	2.903662984	3.428230595
8	4.20945247	6.741624675	3.098069248	3.630513564
9	4.645191724	7.417977822	3.279945116	3.817977822
10	5.096404819	8.08432776	3.451307413	3.993418669
11	5.561330189	8.74219758	3.613732271	4.158864247
12	6.037415265	9.392760536	3.768455453	4.315837459
13	6.521919256	10.03694281	3.916458278	4.465514244
14	7.012352809	10.67549058	4.05853235	4.608823914
15	7.506709879	11.30901534	4.195326246	4.746515339
16	8.003520007	11.93802562	4.327378953	4.879202092

By Lemma 5(iii), this gives less strict initial bounds on the moment ratios than (27)–(28) (Table 4).

The effect of the initial bounds on the final bounds is substantial as can be seen by comparing Tables 1 and 5. Each difference $\bar{\mu}_i^s - \underline{\mu}_i^s$ is higher in Table 5 than in Table 1. This underlines the importance of providing initial bounds in Algorithm 1 that are as strict as possible. Particularly, the bounds on the moments of X^- are sensitive to the initial bounds. More specifically, the accuracy of the bounds on the highest order moment of X^- decreases by three orders of magnitude when the less strict initial bounds in (34) are imposed (rightmost column in Tables 1 and 5).

4.2 Example 2: the exponential function on a quadratic form

Suppose one is interested in $\mathbb{E}\big[\exp(X)\big]$ with the quadratic form $X = Z'\tilde{A}Z = \frac{1}{2}Z'\big[\tilde{A} + \tilde{A}'\big]Z$ where $Z \sim N(\mathbf{0}, I)$ has dimension d and $\tilde{A}$ is a $d \times d$ matrix with eigenvalues less than $\frac{1}{2}$. Diagonalize the symmetric matrix $A := \frac{1}{2}\big[\tilde{A} + \tilde{A}'\big]$ as $V\Lambda V'$ with Λ a diagonal matrix with diagonal entries the eigenvalues $\lambda_1 \leq \cdots \leq \lambda_d < \frac{1}{2}$. Since $V'Z \sim Z$, the random variable X is a weighted summation of d independent Chi-squared distributions with one degree of freedom,

Table 5 Moments of X from (24) and final bounds on partial moments of X from X^s from Algorithm 1 ($s \in \{+, -\}$)

i	μ_i	$\bar{\mu}_i^s - \underline{\mu}_i^s$	$\frac{\bar{\mu}_i^s - \underline{\mu}_i^s}{\mu_i}$ (%)	$\underline{\mu}_i^+$	$\bar{\mu}_i^+$	$\frac{\bar{\mu}_i^+ - \underline{\mu}_i^+}{\underline{\mu}_i^+}$ (%)	$\underline{\mu}_i^-$	$\bar{\mu}_i^-$	$\frac{\bar{\mu}_i^- - \underline{\mu}_i^-}{\underline{\mu}_i^-}$ (%)
0	1	0.605711267	60.57	0.209580046	0.815291313	289.0	0.184708687	0.790419954	327.9
1	0	0.363152443	NA	0.214197827	0.577350269	169.5	0.214197827	0.577350269	169.5
2	1.333333333	0.449005467	33.68	0.482654991	0.931660458	93.0	0.401672875	0.850678342	111.8
3	0.25	0.579079733	231.63	0.924325142	1.503404875	62.6	0.674325142	1.253404875	85.9
4	5.7	0.934374188	16.39	2.958453848	3.892828036	31.6	1.807171964	2.741546152	51.7
5	4.083333333	2.700086746	66.12	7.379772891	10.07985964	36.6	3.296439558	5.996526304	81.9
6	45.39285714	5.63222536	12.41	29.18859785	34.82082321	19.3	10.57203393	16.20425929	53.3
7	76.5	23.0277477	30.10	97.26060675	120.2883545	23.7	20.76060675	43.78835445	110.9
8	597.1111111	64.01829326	10.72	458.5608128	522.579106	14.0	74.53200508	138.5502983	85.9
9	1831.9	278.5404199	15.21	1,991.745206	2,270.285626	14.0	159.8452059	438.3856258	174.3
10	12,404.59091	1,180.922339	9.52	10,803.28546	11,984.2078	10.9	420.383113	1,601.305452	280.9
11	57,412.16667	4,947.565136	8.62	58,313.74211	63,261.30724	8.48	901.5754397	5,849.140575	548.8
12	388,204.0769	23,119.311	5.96	363,151.2004	386,270.5114	6.37	1,933.565475	25,052.87647	1,196
13	2,362,230.429	103,158.9599	4.37	2,366,377.253	2,469,536.213	4.36	4,146.824857	107,305.7848	2,488
14	17,225,582.07	485,659.9711	2.82	16,731,028.6	17,216,688.57	2.90	8,893.495781	494,553.4669	5,461
15	126,430,633.6	2,328,332.164	1.84	126,449,707	128,778,039.2	1.84	19,073.45253	2,347,405.617	12,207
16	1,023,903,852	11,412,560.48	1.11	1,012,450,385	1,023,862,946	1.13	40,905.91577	11453466.4	27,900

Instead of (27)–(28) as in Table 1, the less strict bounds from Lemma 5(iii) are used

$$X = Z'AZ \sim Z'\Lambda Z = \sum_{i=1}^{d} \lambda_i Z_i^2.$$

We can infer the exact outcome of $\mathbb{E}[\exp(X)]$ from the moment generating function of a Chi-squared distribution:

$$\mathbb{E}[\exp(X)] = \prod_{i=1}^{d} \mathbb{E}[\exp(\lambda_i Z_i^2)] = \frac{1}{\prod_{i=1}^{d} \sqrt{1 - 2\lambda_i}} \quad \text{if } \max_i(\lambda_i) < \frac{1}{2} \quad (35)$$

The expectation of $\exp(X)$ is infinite if $\max_i \lambda_i \geq \frac{1}{2}$. The outcome in (35) enables a direct comparison with the bounds that we obtain. It should be stressed that an exact representation of $\mathbb{E}[f(X)]$ is in general unavailable.

Consider the Taylor series expansion

$$\mathbb{E}[\exp(X)] = \sum_{i=0}^{\infty} \frac{1}{i!} \mathbb{E}[X^i]. \quad (36)$$

The moments of X are (Magnus 1986, Lemma 3)

$$\mathbb{E}[X^i] = \sum_{\nu} \gamma_i(\nu) \prod_{j=1}^{i} \left(\sum_k \lambda_k^j \right)^{n_j} \qquad \gamma_i(\nu) = \frac{i! 2^i}{\prod_{j=1}^{i} [n_j! (2j)^{n_j}]}. \quad (37)$$

where the summation is over all $\nu = (n_1, \ldots, n_i)$ with each $n_j \in \mathbb{N}_0$ and $\sum_{j=1}^{i} n_j j = i$. This procedure is computationally expensive for moments of a high order i. A procedure based on Algorithm 2 enables us to bound high moments of X, and thus $\mathbb{E}[\exp(X)]$ in (36). The bounds follow from a few additional steps we outline below.

Using a polar coordinate system, it can be verified that $X \sim Z'\Lambda Z \sim (\sqrt{R}U_d)'\Lambda(\sqrt{R}U_d)$ where the direction vector U_d follows a uniform distribution on the unit d-sphere S_d and R, which is the squared distance to the origin, follows a $\chi^2(d)$-distribution. The latter distribution is a Gamma distribution with shape parameter $d/2$ and scale parameter 2. This means that X is an infinite mixture of Gamma distributions $X_{\mathbf{u}}$ with each component characterized by some $\mathbf{u} \in S^d$. More specifically, each component $X_{\mathbf{u}}$ in the mixture is a Gamma distribution scaled by $\lambda(\mathbf{u}) := \mathbf{u}'\Lambda\mathbf{u}$. The scaling parameter $\lambda(\mathbf{u})$ varies between $\lambda_{\min} = \min_i \lambda_i = \lambda_1$ and $\lambda_{\max} = \max_i \lambda_i = \lambda_d$. For each component $X_{\mathbf{u}}$, a positive (negative) $\lambda(\mathbf{u})$ indicates that the corresponding Gamma distribution is added (subtracted). We loosely write $X_{\mathbf{u}} \sim \text{Gamma}(d/2, 2\lambda(\mathbf{u}))$ for all $\lambda(\mathbf{u})$, including negative $\lambda(\mathbf{u})$.

Define the scaled remainder $\xi_n(Y)$ of a random variable Y by the functional

$$\xi_n(Y) := \frac{1}{\mathbb{E}[Y^n]} \sum_{i=1}^{\infty} \frac{\mathbb{E}[Y^{n+i}]}{(n+i)!} \qquad n \in \mathbb{N}_0^+. \quad (38)$$

For a random variable Y that degenerates at x, i.e., $Y \equiv x$, the functional in (38) can be written as a function of x,

$$\xi_n(x) = \sum_{i=1}^{\infty} \frac{x^i}{(n+i)!} = \frac{{}_1F_1(1, n+2; x) - 1}{(n+1)!} \qquad n \in \mathbb{N}_0^+,$$

where ${}_1F_1(a, b; x)$ is the confluent hypergeometric function of the first kind with parameters a and b. By (36) and (38),

$$\mathbb{E}\big[\exp(X)\big] = 1 + \sum_{i=1}^{n} \frac{1}{i!}\mu_i + \mu_n^+\xi_n(X^+) + \mu_n^-\xi_n(-X^-) \qquad n \in \mathbb{N}^+. \tag{39}$$

The moments $\mu_1, \ldots, \mu_n$ are obtained from (37), while Algorithm 1 produces bounds on μ_n^+ and μ_n^-. The following lemma derives bounds on $\xi_n(X^+)$ and $\xi_n(-X^-)$ in (39).

Lemma 9 *Let $n \in \mathbb{N}^+$.*

(i) *The functionals $\xi_n(X^+)$ and $\xi_n(-X^-)$ are bounded by*

$$0 \le \xi_n(\underline{m}_{n+1}^+) \le \xi_n(X^+)$$
$$\xi_n(-\bar{m}_{n+1}^-) \le \xi_n(-X^-) \le \xi_n(-\underline{m}_n^-) \le 0.$$

(ii) *If the random variable X is a mixture of Gamma distributions $T_\theta \sim \text{Gamma}(k, \theta)$ with fixed shape parameter $k > 0$, scale parameter $\theta \in [\theta_{\min}, \theta_{\max}] \subseteq (-1, 1)$, and $T_\theta = -T_{-\theta}$ for $\theta < 0$, then*

$$\xi_n\left(\frac{(k+n)\underline{m}_n^+}{k+n-1}\right) \le \xi_n(X^+) \le \frac{1}{n!}\left[{}_2F_1\left(1, k+n; n+1; [\theta_{\max}]^+\right) - 1\right]$$
$$\frac{1}{n!}\left[{}_2F_1\left(1, k+n; n+1; -[\theta_{\min}]^-\right) - 1\right] \le \xi_n(-X^-),$$

where ${}_2F_1$ is the Gaussian hypergeometric function.

The bounds in Lemma 9(i) can deal with any distribution X, particularly any eigenvalue of A. In contrast, the two bounds in Lemma 9(ii) that are based on the Gaussian hypergeometric function ${}_2F_1$ require that each absolute eigenvalue of A is less than $\frac{1}{2}$, because $X_\mathbf{u}$ has scale parameter $\theta(\mathbf{u}) = 2\mathbf{u}^T A\mathbf{u} = 2\lambda(\mathbf{u})$ with $\mathbf{u} \in S^d$.

We present example cases where the minimal eigenvalue $\lambda_{\min}$ is -0.1, -0.2, -0.3, or -0.4, while the maximal eigenvalue $\lambda_{\max}$ is 0.1, 0.2, 0.3, or 0.4. The other eigenvalues are either equally spaced between $\lambda_{\min}$ and $\lambda_{\max}$, or equally split at the two extremes.

Define

$$Y := \exp(X) \quad r := \frac{\bar{\mu}_Y - \underline{\mu}_Y}{\mu_Y} \quad w := \frac{\mu_Y - \underline{\mu}_Y}{\bar{\mu}_Y - \underline{\mu}_Y}, \tag{40}$$

where $\underline{\mu}_Y$ and $\bar{\mu}_Y$ are a lower and an upper bound on the mean μ_Y of Y, respectively. The ratio r represents the size of the maximal error relative to the mean μ_Y. A small r indicates more accurate bounds. The weight $w \in [0, 1]$ is the normalized location of μ_Y on the interval $\left[\underline{\mu}_Y, \bar{\mu}_Y\right]$. A value of w close to zero (one) reflects that the lower

Table 6 $d = 2$, $n = 14$, and stopping criterion $\varepsilon_k < 10^{-6}$

Model	$\lambda_{\min}$	$\lambda_{\max}$	μ_Y	$\underline{\mu}_Y$	$\bar{\mu}_Y$	r	w	Time (ms)	Num. iter
1	−0.1	0.1	1.020620726	1.020620726	1.020620726	2.17E−12	0.52	16.3	6
2	−0.2	0.1	0.944911183	0.944911180	0.944911186	6.09E−09	0.39	24.1	11
3	−0.3	0.1	0.883883477	0.883882666	0.883884578	2.16E−06	0.42	27.5	13
4	−0.4	0.1	0.833333333	0.833283349	0.833394146	1.33E−04	0.45	27.4	100
5	−0.1	0.2	1.178511302	1.178511290	1.178511313	2.00E−08	0.52	24.4	11
6	−0.2	0.2	1.091089451	1.091089396	1.091089489	8.52E−08	0.59	16.2	6
7	−0.3	0.2	1.020620726	1.020619754	1.020622298	2.49E−06	0.38	21.0	9
8	−0.4	0.2	0.962250449	0.962204899	0.962307082	1.06E−04	0.45	24.5	11
9	−0.1	0.3	1.443375673	1.443355594	1.443386736	2.16E−05	0.64	27.3	13
10	−0.2	0.3	1.336306210	1.336285709	1.336316480	2.30E−05	0.67	20.8	9
11	−0.3	0.3	1.250000000	1.249955835	1.250023095	5.38E−05	0.66	16.5	6
12	−0.4	0.3	1.178511302	1.178330474	1.178721082	3.31E−04	0.46	16.2	6
13	−0.1	0.4	2.041241452	2.034127240	2.044169747	4.92E−03	0.71	27.2	100
14	−0.2	0.4	1.889822365	1.883325573	1.892480556	4.84E−03	0.71	24.5	11
15	−0.3	0.4	1.767766953	1.760843241	1.770546599	5.49E−03	0.71	16.2	6
16	−0.4	0.4	1.666666667	1.657677360	1.670628051	7.77E−03	0.69	16.3	6

Moment μ_Y from (35), bounds $\underline{\mu}_Y$ and $\bar{\mu}_Y$ from Algorithm 1, (39), and Lemma 9, ratios r and w from (40)

Table 7 $d = 20$, $n = 14$, stopping criterion $\varepsilon_k < 10^{-6}$, and λ equally spaced on $[\lambda_{\min}, \lambda_{\max}]$

Model	$\lambda_{\min}$	$\lambda_{\max}$	μ_Y	$\underline{\mu}_Y$	$\bar{\mu}_Y$	r	w	Time (ms)	Num. iter
1	−0.1	0.1	1.077538350	1.077538350	1.07753835	1.39E−10	0.39	13.6	4
2	−0.2	0.1	0.443711701	0.443707006	0.443712681	1.28E−05	0.83	18.6	7
3	−0.3	0.1	0.199777427	0.195641935	0.200118265	2.24E−02	0.92	32.9	16
4	−0.4	0.1	0.096647862	−0.302971521	0.126456813	4.44E+00	0.93	33.9	16
5	−0.1	0.2	3.547592932	3.547589205	3.547635372	1.30E−05	0.08	18.4	7
6	−0.2	0.2	1.365847679	1.365845671	1.365853055	5.41E−06	0.27	13.8	4
7	−0.3	0.2	0.580844471	0.580257471	0.581123237	1.49E−03	0.68	22.1	9
8	−0.4	0.2	0.267482930	0.180807024	0.294552335	4.25E−01	0.76	19.2	7
9	−0.1	0.3	15.42441363	15.41526999	15.63192003	1.41E−02	0.04	33.2	16
10	−0.2	0.3	5.429690103	5.427968854	5.457873354	5.51E−03	0.06	21.7	9
11	−0.3	0.3	2.141558899	2.140439149	2.149350941	4.16E−03	0.13	14.0	4
12	−0.4	0.3	0.924611333	0.902789338	0.959364519	6.12E−02	0.39	21.3	8
13	−0.1	0.4	113.2382513	107.889769	1034.580947	8.18E+00	0.01	33.4	16
14	−0.2	0.4	34.88184583	33.77084803	223.2829373	5.43E+00	0.01	19.1	7
15	−0.3	0.4	12.31834995	12.04519323	57.51232046	3.69E+00	0.01	21.1	8
16	−0.4	0.4	4.843952405	4.693885913	22.87053246	3.75E+00	0.01	15.2	4

Moment μ_Y from (35), bounds $\underline{\mu}_Y$ and $\bar{\mu}_Y$ from Algorithm 1, (39), and Lemma 9, ratios r and w from (40)

Table 8 $d = 20, n = 14$, stopping criterion $\varepsilon_k < 10^{-6}$, and multiplicity $= 10$ for both λ_{min} and λ_{max}

Model	λ_{min}	λ_{max}	μ_Y	$\underline{\mu}_Y$	$\bar{\mu}_Y$	r	w	Time (ms)	Num. iter
1	-0.1	0.1	1.226433020	1.226433013	1.226433030	1.42E−08	0.43	12.0	3
2	-0.2	0.1	0.567426856	0.567333752	0.567443729	1.94E−04	0.85	23.3	10
3	-0.3	0.1	0.291038305	0.243588920	0.297152834	1.84E−01	0.89	25.3	11
4	-0.4	0.1	0.161505583	−3.297354903	0.559174340	2.39E+01	0.90	33.5	16
5	-0.1	0.2	5.168178653	5.168021983	5.169214106	2.31E−04	0.13	23.5	10
6	-0.2	0.2	2.391132100	2.390814679	2.391710330	3.75E−04	0.35	12.2	3
7	-0.3	0.2	1.226433020	1.207276594	1.238770476	2.57E−02	0.61	23.3	10
8	-0.4	0.2	0.680583197	−0.978452311	0.984470818	2.88E+00	0.85	24.4	10
9	-0.1	0.3	39.24585664	38.74123015	42.66696115	1.00E−01	0.13	25.0	11
10	-0.2	0.3	18.15765938	17.92105434	19.49668755	8.68E−02	0.15	23.7	10
11	-0.3	0.3	9.313225746	9.055087015	10.22003614	1.25E−01	0.22	12.5	3
12	-0.4	0.3	5.168178653	4.095785072	7.408452526	6.41E−01	0.32	22.5	9
13	-0.1	0.4	1255.867413	730.6466560	15211.39638	1.15E+01	0.04	33.6	16
14	-0.2	0.4	581.0451003	347.7114576	7084.213495	1.16E+01	0.03	24.0	10
15	-0.3	0.4	298.0232239	175.9155469	3655.652693	1.17E+01	0.04	22.3	9
16	-0.4	0.4	165.3817169	79.97024849	2598.933043	1.52E+01	0.03	14.0	3

Moment μ_Y from (35), bounds $\underline{\mu}_Y$ and $\bar{\mu}_Y$ from Algorithm 1, (39), and Lemma 9, ratios r and w from (40)

(upper) bound is the most accurate bound on μ_Y. The bounds are equally accurate if $w = \frac{1}{2}$.

Algorithm 1 stops after iteration k if $\varepsilon_k < 10^{-6}$ with ε_k as in (30), or if $k = 100$. Subsequently, the expression in (39) can be bounded. Table 6 reports several statistics for the case with two eigenvalues ($d = 2$). The mean μ_Y is in all cases of the same order of magnitude. The maximal relative error r is minimal when both extreme eigenvalues $\lambda_{\min}$ and $\lambda_{\max}$ are close to zero; a similar observation holds with $d = 20$ in Table 7 (eigenvalues equally spaced on the interval $[\lambda_{\min}, \lambda_{\max}]$) and in Table 8 (10 eigenvalues at both $\lambda_{\min}$ and $\lambda_{\max}$). Indeed, we can perfectly estimate $\mu_Y = 1$ for the case where each eigenvalue equals zero.

It follows from w in Table 6 that with $d = 2$, the upper bound tends to be more accurate than the lower bound if $\lambda_{\max}$ is high, thus when X^+ tends to be large. This observation is reversed for $d = 20$ (Tables 7 and 8). Compared to $\lambda_{\max}$, the value of $\lambda_{\min}$ has a smaller impact on w. The accuracy decreases with the dimension d and the magnitude of the eigenvalues of A. More specifically, we observe a lower accuracy r if $d = 20$ and $\max(|\lambda_{\min}|, |\lambda_{\max}|) \geq 0.3$.

The computation time in Tables 6, 7, 8 is the mean CPU time of 5,000 computations of each model. The standard error is at most 0.12 ms. The computation time and the number of iterations are lower in cases where $|\lambda_{\min}| = |\lambda_{\max}|$.

5 Discussion and conclusions

This paper has presented an iterative algorithm that bounds the lower and upper partial moment at $a \in \mathbb{R}$ of the random variable X with a known finite sequence of moments of X. In a numerical example, particularly the higher order partial moments can have narrow bounds. The obtained bounds imply bounds on unobserved higher order moments of X which is useful for bounding moments of the transformation $f(X)$. In another application, a transformation $f(x) = e^x$ is considered for the quadratic form $X = Z'AZ$ where Z is a multivariate normal distribution. Numerical experiments suggest that the obtained bounds on $\mathbb{E}\big[\exp(X)\big]$ are most accurate if X^+ is not too large. The accuracy depends on the dimension as well as the eigenvalues of A.

Appendix

Proof of Lemma 1 Since the equality cases are straightforward, we prove strict bounds for X with both $X^+ \not\equiv 0$ and $X^- \not\equiv 0$. The constraints in (2) and the first inequality

in (3) follow from the identity $\mu_i = \mu_i^+ + (-1)^i \mu_i^-$ ($i \in \mathbb{N}_0^+$). The second and third inequality in (3) follow from ($n \in \mathbb{N}^+$)

$$m_{2n-1}^+ = \frac{\mu_{2n-1}^+}{\mu_{2n-2}^+} > \frac{\mu_{2n-1}^+ - \mu_{2n-1}^-}{\mu_{2n-2}^+ + \mu_{2n-2}^-} = m_{2n-1}$$

$$\frac{1}{m_{2n}^+} = \frac{\mu_{2n-1}^+}{\mu_{2i}^+} > \frac{\mu_{2n-1}^+ - \mu_{2n-1}^-}{\mu_{2i}^+ + \mu_{2i}^-} = \frac{1}{m_{2n}}$$

$$m_{2n-1}^- = \frac{\mu_{2n-1}^-}{\mu_{2n-2}^-} > \frac{\mu_{2n-1}^- - \mu_{2n-1}^+}{\mu_{2n-2}^+ + \mu_{2n-2}^-} = -m_{2n-1}$$

$$\frac{1}{m_{2n}^-} = \frac{\mu_{2n-1}^-}{\mu_{2n}^-} > \frac{\mu_{2n-1}^- - \mu_{2n-1}^+}{\mu_{2n}^+ + \mu_{2n}^-} = -\frac{1}{m_{2n}}.$$

$\square$

Proof of Lemma 2 See Theorem 3.1 in Gavriliadis and Athanassoulis (2009). $\square$

Proof of Lemma 3 The j zeros $x_j^{(1)} < x_j^{(2)} < \cdots < x_j^{(j)}$ of each polynomial P_j are real and distinct (Akhiezer 1965; Szegö 1975). Lemma 2 implies that the functions L and U are a lower bound and an upper bound of the cdf F of X, respectively:

$$L(x) \leq F(x) \leq U(x)$$

Therefore,

$$\mu_i^+ = \int_0^\infty x^i \, dF(x) \geq \int_0^\infty x^i \, dU(x) = \sum_{j=j_0^+}^N \left(\tilde{x}^{(j)}\right)^i dU_j$$

$$\mu_i^- = \int_{-\infty}^0 (-x)^i \, dF(x) \geq \int_{-\infty}^0 (-x)^i \, dL(x) = \sum_{j=1}^{j_0^+ - 1} \left(-\tilde{x}^{(j)}\right)^i dL_j,$$

where j_0^+, dL_j, and dU_j ($j = 1, \ldots, N$) are defined in Lemma 3. $\square$

Proof of Lemma 4 (i) The Cauchy–Schwarz inequality with the inner product $\langle Y, Z \rangle = \mathbb{E}[YZ]$ is $|\mathbb{E}[YZ]|^2 \leq \mathbb{E}[Y^2]\mathbb{E}[Z^2]$. Substituting $Y = X^{n/2-1}$ and $Z = X^{n/2}$ leads to the desired result for $n > 2$, and for $n = 2$ provided $\mu_0 := \mathbb{P}(X \neq 0) = 1$. An additional step is needed for the case where $n = 2$ and $\mu_0 < 1$. Consider the Cauchy–Schwarz inequality $\tilde{\mu}_1^2 = \mathbb{E}[\tilde{X}]^2 \leq \mathbb{E}[\tilde{X}^0]\mathbb{E}[\tilde{X}^2] = \tilde{\mu}_0 \tilde{\mu}_2$ with $\tilde{X} \sim X_{X \neq 0}$. Since $\tilde{\mu}_i = \mu_i / \mathbb{P}(X \neq 0)$, multiplying both hand sides of this inequality by $[\mathbb{P}(X \neq 0)]^2$ gives the result with $n = 2$ and $\mu_0 < 1$.

The Cauchy–Schwarz inequality holds with equality if and only if (iff) $Y = \lambda Z$ for some constant λ, $Y \equiv 0$, or $Z \equiv 0$. None of these three conditions holds if X is nqd. It follows that $\mu_{n-1}^2 = \mu_n \mu_{n-2}$ holds iff X is quasi-degenerate.

(ii) Follows from $\mu_1 > 0$ and (i).

$\hfill\square$

Proof of Lemma 5 Without loss of generality, assume $J = \{1, \ldots, |J|\}$ and $0 < h_1 \le h_2 \le \cdots$. When summing over all $j \in J$, the set J is omitted for convenience.

(i) Notice that

$$0 < m_i(Y_1) = g_i h_1 \le m_i(Y_2) = g_i h_2 \le \cdots .$$

A convex combination of larger $m_i(Y_j)$, $j \ge j_0$, is larger than a convex combination of this combination and another convex combination of smaller $m_i(Y_j)$ with $j < j_0$:

$$\sum_{j \ge j_0} m_i(Y_j) \frac{\mathbb{E}\left[Y_j^{i-1}\right]}{\sum_{r \ge j_0} \mathbb{E}\left[Y_r^{i-1}\right]} \ge \sum_j m_i(Y_j) \frac{\mathbb{E}\left[Y_j^{i-1}\right]}{\sum_r \mathbb{E}\left[Y_r^{i-1}\right]}$$

Rearranging terms,

$$\frac{\sum_{j \ge j_0} \mathbb{E}\left[Y_j^i\right]}{\sum_j \mathbb{E}\left[Y_j^i\right]} = \frac{\sum_{j \ge j_0} m_i(Y_j)\mathbb{E}\left[Y_j^{i-1}\right]}{\sum_j m_i(Y_j)\mathbb{E}\left[Y_j^{i-1}\right]} \ge \frac{\sum_{r \ge j_0} \mathbb{E}\left[Y_r^{i-1}\right]}{\sum_r \mathbb{E}\left[Y_r^{i-1}\right]}. \tag{41}$$

Let $w_j^{(i)} := \mathbb{E}\left[Y_j^{i-1}\right]/\sum_{r \in J} \mathbb{E}\left[Y_r^{i-1}\right]$. By (41), the measure $w^{(i+1)}$ stochastically dominates the measure $w^{(i)}$:

$$\sum_{j \ge j_0} w_j^{(i+1)} = \sum_{j \ge j_0} \frac{\mathbb{E}\left[Y_j^i\right]}{\sum_r \mathbb{E}[Y_r^i]} \ge \frac{\sum_{j \ge j_0} \mathbb{E}\left[Y_j^{i-1}\right]}{\sum_r \mathbb{E}\left[Y_r^{i-1}\right]} = \sum_{j \ge j_0} w_j^{(i)}$$

Write

$$m_i(Y) = \frac{\mathbb{E}\left[\left(\sum_j Y_j\right)^i\right]}{\mathbb{E}\left[\left(\sum_r Y_r\right)^{i-1}\right]} = \sum_j \frac{\mathbb{E}\left[Y_j^i\right]}{\sum_r \mathbb{E}\left[Y_r^{i-1}\right]} = \sum_j \frac{\mathbb{E}\left[Y_j^{i-1}\right]}{\sum_r \mathbb{E}\left[Y_r^{i-1}\right]} \frac{\mathbb{E}\left[Y_j^i\right]}{\mathbb{E}\left[Y_j^{i-1}\right]}$$
$$= \sum_j w_j^{(i)} m_i(Y_j) = g_i \sum_j w_j^{(i)} h_j.$$

Because $\{h_j\}_{j \in J}$ is a positive nondecreasing sequence and $w^{(i+1)}$ stochastically dominates $w^{(i)}$, for each $r \in J$,

$$\frac{m_{i+1}(Y)}{m_{i+1}(Y_r)} = \frac{g_{i+1} \sum_j w_j^{(i+1)} h_j}{g_{i+1} h_r} = \frac{1}{h_r} \sum_j w_j^{(i+1)} h_j \geq \frac{1}{h_r} \sum_j w_j^{(i)} h_j = \frac{m_i(Y)}{m_i(Y_r)}.$$

(42)

Inequality (5) follows from rearranging (42) and $m_i, m_{i+1} > 0$.

(ii) Under the assumption that Y is a mixture, i.e., $\mathbb{P}(Y_j Y_r = 0) = 1 \ (j \neq r)$, the moment ratio $m_i(Y)$ is a weighted average of $m_i(Y_j)$ with positive weights $\mathbb{E}[Y_j^{i-1}] / \sum_r \mathbb{E}[Y_r^{i-1}]$:

$$m_i(Y) = \frac{\mathbb{E}[(\sum_j Y_j)^i]}{\mathbb{E}[(\sum_r Y_r)^{i-1}]} = \frac{\sum_j \mathbb{E}[Y_j^i]}{\sum_r \mathbb{E}[Y_r^{i-1}]} = \sum_j \frac{\mathbb{E}[Y_j^i]}{\mathbb{E}[Y_j^{i-1}]} \frac{\mathbb{E}[Y_j^{i-1}]}{\sum_r \mathbb{E}[Y_r^{i-1}]}$$

$$\leq \max_j \left(\frac{\mathbb{E}[Y_j^i]}{\mathbb{E}[Y_j^{i-1}]} \right) = \max_j m_i(Y_j)$$

(iii) The highest moment in the expansion of $\left(\sum_{r \in J} Y_r\right)^{i-1}$ equals $i - 1$. By Lemma 4(ii),

$$\mathbb{E}[Y_j^{r_j+1}] \leq \mathbb{E}[Y_j^{r_j}] \frac{\mathbb{E}[Y_j^i]}{\mathbb{E}[Y_j^{i-1}]} \qquad r_j + 1 \leq i. \qquad (43)$$

Hence, by independence of each $\{Y_r\}_{r \in J}$ and Lemma 4(ii),

$$\mathbb{E}\left[\left(\sum_r Y_r\right)^{i-1} Y_j\right] = \sum_{\sum_j r_j = i-1} \binom{i-1}{r_1, \dots, r_{|J|}} \mathbb{E}[Y_1^{r_1}] \dots \mathbb{E}[Y_j^{r_j+1}] \dots \mathbb{E}[Y_{|J|}^{r_{|J|}}]$$

$$= \sum_{\sum_j r_j = i-1} \binom{i-1}{r_1, \dots, r_{|J|}} \mathbb{E}[Y_1^{r_1}] \dots \mathbb{E}[Y_{|J|}^{r_{|J|}}] \frac{\mathbb{E}[Y_j^{r_j+1}]}{\mathbb{E}[Y_j^{r_j}]}$$

$$\leq \sum_{\sum_j r_j = i-1} \binom{i-1}{r_1, \dots, r_{|J|}} \mathbb{E}[Y_1^{r_1}] \dots \mathbb{E}[Y_{|J|}^{r_{|J|}}] \frac{\mathbb{E}[Y_j^i]}{\mathbb{E}[Y_j^{i-1}]}$$

$$= \mathbb{E}\left[\left(\sum_r Y_r\right)^{i-1}\right] \frac{\mathbb{E}[Y_j^i]}{\mathbb{E}[Y_j^{i-1}]}. \qquad (44)$$

Substituting (44),

$$m_i(Y) = \frac{\mathbb{E}[Y^i]}{\mathbb{E}[Y^{i-1}]} = \frac{\sum_j \mathbb{E}[(\sum_r Y_r)^{i-1} Y_j]}{\mathbb{E}[(\sum_r Y_r)^{i-1}]} \leq \sum_j \frac{\mathbb{E}[(\sum_r Y_r)^{i-1}]}{\mathbb{E}[(\sum_r Y_r)^{i-1}]} \frac{\mathbb{E}[Y_j^i]}{\mathbb{E}[Y_j^{i-1}]}$$

$$= \sum_j \frac{\mathbb{E}[Y_j^i]}{\mathbb{E}[Y_j^{i-1}]} = \sum_j m_i(Y_j).$$

(iv) As the case $I = J$ is straightforward, assume $I \subset J$. Define the complement of I as $I^c := J \backslash I$. Since each Y_j is independent and log-concave, the convolutions $Y_I := \sum_{j \in I} Y_j$, $Y_{I^c} := \sum_{j \in I^c} Y_j$, and $Y = Y_I + Y_{I^c}$ have log-concave distributions (Ibragimov 1956). Denote the probability density function (pdf) of the nonnegative random variables Y_I, Y_{I^c}, and Y by $f_{Y_I}(x)$, $f_{Y_{I^c}}(x)$, and $f_Y(x) := \int_0^\infty f_{Y_I}(x - z) \, dF_{Y_{I^c}}(z)$, respectively. For any $i \geq 1$, consider $\tilde{Y}_I$ and $\tilde{Y}$ with pdf

$$f_{\tilde{Y}_I}(x) = \frac{x^{i-1} f_{Y_I}(x)}{\int_0^\infty x^{i-1} f_{Y_I}(x) \, dx} \qquad f_{\tilde{Y}}(x) = \frac{x^{i-1} f_Y(x)}{\int_0^\infty x^{i-1} f_Y(x) \, dx}. \qquad (45)$$

Because $m_i(Y_I) = \frac{\mathbb{E}[[Y_I]^i]}{\mathbb{E}[[Y_I]^{i-1}]} = \mathbb{E}[\tilde{Y}_I]$ and $m_i(Y) = \frac{\mathbb{E}[Y^i]}{\mathbb{E}[Y^{i-1}]} = \mathbb{E}[\tilde{Y}]$, we need to show that $\mathbb{E}[\tilde{Y}_I] \leq \mathbb{E}[\tilde{Y}]$ holds. To do this, it suffices to show that the marginal likelihood ratio $f_{\tilde{Y}_j}/f_{\tilde{Y}} \propto f_{Y_j}/f_Y$ is nonincreasing on $(0, \infty)$ as this indicates that $\tilde{Y}$ stochastically dominates $\tilde{Y}_j$. Since Y_I is a log-concave distribution, we have that $\frac{d}{dx} \log f_{Y_I}(x)$ is nonincreasing on $[0, \infty)$. This gives for the density function of $Y = Y_I + Y_{I^c}$,

$$\frac{d}{dx} \log f_{Y_I}(x) = \frac{d}{dx} \log \mathbb{E}[f_Y(x + Y_{I^c})] \leq \frac{d}{dx} \log f_Y(x) \qquad x \geq 0.$$

This proves that $\log \left(f_{Y_I}(x)/f_Y(x) \right)$, and thus the ratio $f_{\tilde{Y}_I}/f_{\tilde{Y}}$, is nonincreasing on $[0, \infty)$. Therefore, $\mathbb{E}[\tilde{Y}_I] \leq \mathbb{E}[\tilde{Y}]$ and thus $m_i(Y_I) \leq m_i(Y)$ for any $i \geq 1$.

□

Proof of Lemma 6 We prove the bounds on $m_i^+ := m(X^+)$ and $m_i(Y)$, the statements on m_i^- and $m_i(Z)$ follow analogously.

(i) Since $X = \sum_{j \in I} X_j$ is a mixture, the positive part of X is $X^+ = \sum_{j \in I} X_j^+$. Apply Lemma 5(ii).

(ii) Since each Y_j is independent and log-concave, the convolution $Y = \sum_{j \in J} Y_j$ is log-concave (Ibragimov 1956). Denote the probability density function (pdf) of X and Y by $f_X(x) := \int_0^\infty f_Y(x + z) \, dF_Z(z)$ and $f_Y(x)$, respectively. For any

$i \geq 1$, consider $\tilde{X}$ and $\tilde{Y}$ with pdf

$$f_{\tilde{X}}(x) = \frac{x^{i-1} f_{\tilde{X}}(x)}{\int_0^\infty x^{i-1} f_{\tilde{X}}(x)\, \mathrm{d}x} \qquad f_{\tilde{Y}}(x) = \frac{x^{i-1} f_{\tilde{Y}}(x)}{\int_0^\infty x^{i-1} f_{\tilde{Y}}(x)\, \mathrm{d}x}. \qquad (46)$$

Because $m_i^+ = m_i(X^+) = \frac{\mathbb{E}\left[[X^+]^i\right]}{\mathbb{E}\left[[X^+]^{i-1}\right]} = \mathbb{E}\left[\tilde{X}\right]$ and $m_i(Y) = \frac{\mathbb{E}[Y^i]}{\mathbb{E}[Y^{i-1}]} = \mathbb{E}\left[\tilde{Y}\right]$,

we need to show that $\mathbb{E}\left[\tilde{X}\right] \leq \mathbb{E}\left[\tilde{Y}\right]$ holds. To do this, it suffices to show that the marginal likelihood ratio $f_{\tilde{X}}/f_{\tilde{Y}} \propto f_X/f_Y$ is nonincreasing on $(0, \infty)$ as this indicates that $\tilde{Y}$ stochastically dominates $\tilde{X}$. Since Y is a log-concave distribution, $\frac{\mathrm{d}}{\mathrm{d}x} \log f_Y(x)$ is nonincreasing on $[0, \infty)$. This gives for the density function of $X = Y - Z$ with Z any nonnegative and independent distribution,

$$\frac{\mathrm{d}}{\mathrm{d}x} \log f_X(x) = \frac{\mathrm{d}}{\mathrm{d}x} \log \mathbb{E}[f_Y(x + Z)] \leq \frac{\mathrm{d}}{\mathrm{d}x} \log f_Y(x) \qquad x \geq 0.$$

This proves that $\log(f_X(x)/f_Y(x))$, and thus the ratio $f_{\tilde{X}}/f_{\tilde{Y}}$, is nonincreasing on $[0, \infty)$. Therefore, $\mathbb{E}\left[\tilde{X}\right] \leq \mathbb{E}\left[\tilde{Y}\right]$ and thus $m_i(X^+) \leq m_i(Y)$ for any $i \in \mathbb{N}^+$.

(iii) Follows immediately from $X^+ = \max(X, 0) \equiv Y$.

□

Proof of Theorem 1 The moments μ_{n-2}, μ_{n-1}, and μ_n are known and satisfy $\mu_i = \mu_i^+ + (-1)^i \mu_i^-$ ($i \in \{n-2, n-1, n\}$ and n even). Combining this with $0 < m_{n-1}^+ \leq m_n^+$ from Lemma 4(ii) gives for $t \in \{0, 1\}$

$$0 \leq \mu_{n-2}^- = \mu_{n-2} - \frac{\mu_{n-1}^+}{m_{n-1}^+} \leq \mu_{n-2} - \frac{\mu_{n-1}^+}{m_{n-t}^+} \qquad (47)$$

$$\mu_{n-1}^- = \mu_{n-1}^+ - \mu_{n-1} \qquad (48)$$

$$0 \leq \mu_n^- = \mu_n - m_n^+ \mu_{n-1}^+ \leq \mu_n - m_{n-t}^+ \mu_{n-1}^+. \qquad (49)$$

Define for $s \in \{+, -\}$

$$g^s(x) := \frac{\mu_n}{x} - 2s\,\mu_{n-1} + \mu_{n-2}x \qquad\qquad x > 0.$$

The function g^s is minimal at $b := \sqrt{\mu_n/\mu_{n-2}}$, monotonically decreases on $(0, b)$, and monotonically increases on (b, ∞). Apply Lemma 4(i) to the nqd X,

$$g^s(x) \geq g^s(b) = 2\left(\sqrt{\mu_n \mu_{n-2}} - s\,\mu_{n-1}\right) > 0 \qquad\qquad x > 0. \qquad (50)$$

By (47)–(50) and (12), for $t \in \{0, 1\}$

$$(\mu_{n-1}^-)^2 \leq \mu_{n-2}^- \mu_n^- \Rightarrow (\mu_{n-1}^+ - \mu_{n-1})^2 \leq \left(\mu_{n-2} - \frac{\mu_{n-1}^+}{m_{n-t}^+}\right)(\mu_n - m_{n-t}^+ \mu_{n-1}^+)$$

$$\Leftrightarrow \mu_{n-1}^+ \leq \frac{\mu_n \mu_{n-2} - \mu_{n-1}^2}{\mu_n / m_{n-t}^+ - 2 \mu_{n-1} + \mu_{n-2} m_{n-t}^+}$$

$$\Leftrightarrow \mu_{n-1}^+ \leq u^+(m_{n-t}^+) := \frac{a}{g^+(m_{n-t}^+)}. \tag{51}$$

Notably, the derivation above started with bounds on the moments of X^- and ends with bounds on the moments of X^+.

Since $-X = X^- - X^+$, we can switch all superscript $+$ and $-$ signs and use $\mathbb{E}[(-X)^i] = (-1)^i \mu_{i-1}$ to generalize (51) to

$$\mu_{n-1}^s \leq u^s(m_{n-t}^s) := \frac{a}{g^s(m_{n-t}^s)} \qquad s \in \{+, -\} \qquad t \in \{0, 1\}. \tag{52}$$

Using (52), we bound μ_{n-1}^s by substituting appropriate values for the unobserved m_{n-t}^s. The function u^s increases on $(0, b)$, while it decreases on (b, ∞). Therefore, the most conservative upper bound from (52) equals $u^s(b) = a/g^s(b)$. Possibly tighter bounds on μ_{n-1}^s follow from the more general bound $\min(u^s(c), u^s(d))$ with $c = \operatorname{argmin}_x\{|x - b| : x \in [\underline{m}_{n-1}^s, \bar{m}_{n-1}^s]\}$ and $d = \operatorname{argmin}_x\{|x - b| : x \in [\underline{m}_n^s, \bar{m}_n^s]\}$. Thus, the upper bound on μ_{n-1}^s is $u^s(b)$ in cases where both $b \in [\underline{m}_{n-1}^s, \bar{m}_{n-1}^s]$ and $b \in [\underline{m}_n^s, \bar{m}_n^s]$ hold. By (15), this is equivalent to cases where $b \in [\underline{m}_n^s, \bar{m}_{n-1}^s]$.

If $\bar{m}_{n-1}^s \leq b \leq \underline{m}_n^s$, one can use $\min(u^s(\bar{m}_{n-1}^s), u^s(\underline{m}_n^s))$ as a bound, because (15) and the monotonicity of u^s imply that $\bar{m}_{n-1}^s$ or $\underline{m}_n^s$ must give the lowest possible upper bound on μ_{n-1}^s. The bottom two components of v^s in (16) follow from the monotonicity of the function u^s on $(0, b)$ and (b, ∞).

The obtained upper bound on μ_{n-1}^s implies an upper bound on μ_{n-2}^s from $\mu_{n-2}^s \leq \bar{\mu}_{n-1}^s / \underline{m}_{n-1}^s$. A tighter bound is available if the previous bound on μ_{n-1}^s is based on $\bar{m}_{n-1}^s$. By (52),

$$\mu_{n-2}^s \leq \frac{a}{h^s(m_{n-1}^s)} \qquad s \in \{+, -\}, \tag{53}$$

where

$$h^s(x) := x\, g^s(x) = \mu_n - 2s\, \mu_{n-1} x + \mu_{n-2} x^2 \qquad x > 0.$$

In (53), an upper bound on μ_{n-2}^s corresponds to a lower bound on h^s. The continuous function h^s monotonically decreases on $(0, s\, m_{n-1})$ and monotonically increases on $([s\, m_{n-1}]^+, \infty)$. By (14), $\underline{m}_{n-1}^s \geq s\, m_{n-1}$ such that $\underline{m}_{n-1}^s$ gives the maximal upper bound on μ_{n-2}^s in (53). As a result, using $\underline{m}_{n-1}^s$ (instead of $\bar{m}_{n-1}^s$) can tighten the upper bound on μ_{n-2}^s. This gives the bound on μ_{n-2}^s in (16).

Similarly, the upper bound on μ_{n-1}^s implies the upper bound $\mu_n^s \leq \bar{m}_n^s \bar{\mu}_{n-1}^s$. A tighter bound is available if the previous bound on μ_{n-1}^s is based on $\underline{m}_n^s$. By (52),

$$\mu_n^s \leq \frac{a}{k^s(m_n^s)} \qquad s \in \{+, -\}, \tag{54}$$

where

$$k^s(x) := \frac{g^s(x)}{x} = \frac{\mu_n}{x^2} - \frac{2s\,\mu_{n-1}}{x} + \mu_{n-2} \qquad x > 0.$$

In (54), an upper bound on μ_n^s corresponds to a lower bound on k^s. The function k^s is continuous on $(0, \infty)$ and monotonically increases on the interval $([s\,m_n]^+, \infty)$. Provided $s\,m_n > 0$, k^s monotonically decreases on $(0, s\,m_n)$. By (14), $1/\bar{m}_n^s \geq 1/(s\,m_n)$ such that $\bar{m}_n^s$ gives the maximal upper bound on μ_n^s in (54). Thus, using $\bar{m}_n^s$ (instead of $\underline{m}_n^s$) can tighten the upper the bound on μ_n^s. This proves the last inequality in (16). □

Proof of Lemma 7 The inequality in (17) follows from Lemma 4(i). Next, consider Hölder's inequality

$$\mathbb{E}[|YZ|] \leq \left(\mathbb{E}\big[|Y|^p\big]\right)^{1/p} \left(\mathbb{E}\big[|Z|^q\big]\right)^{1/q} \qquad p, q \in (1, \infty) \qquad \frac{1}{p} + \frac{1}{q} = 1.$$

The first inequality in (18) is obtained by substituting $Y = (X^s)^j$, $Z \equiv 1$, $p = i/j$ and $q = p/(p-1)$. The first inequality in (19) follows from $Y = (X^s)^i$, $p = j/i$, and $q = p/(p-1)$. The second inequality in both (18) and (19) follows from Lemma 4(i) and $\mu_i^s = m_i^s \mu_{i-1}^s$ $(i \in \mathbb{N}^+)$. The inequalities in (20)–(21) are based on $\mu_i = \mu_i^+ + (-1)^i \mu_i^-$ where $i \in \mathbb{N}_0^+$. □

Proof of Lemma 8 We prove the case with strict bounds, since the case with equality signs is trivial. Both inequalities in (22) are straightforward by $m_i^s = \mu_i^s / \mu_{i-1}^s$. The two inequalities in (23) follow from Lemma 4(ii):

$$m_{i-1}^s < \sqrt{m_{i-1}^s m_i^s} = \sqrt{\frac{\mu_i^s}{\mu_{i-2}^s}} < m_i^s$$

□

Proof of Lemma 9

(i) Define the function

$$r_{m,n}(x) = \sum_{i=m}^{\infty} \frac{x^i}{(n+i)!} \qquad m \in \{1, 2\} \qquad n = 0, 1, \dots$$

We prove the properties of $r_{m,n}$ in Tabel 9. Notice that

$$r_{1,n}(x) = \xi_n(x) = \frac{{}_1F_1(1, n+2; x) - 1}{(n+1)!}. \tag{55}$$

Apply Abramowitz and Stegun (1964, Eqs. (13.4.9) and (13.1.27)),

$$\frac{\mathrm{d}^k}{\mathrm{d}x^k} r_{1,n}(x) = \frac{\mathrm{d}^k}{\mathrm{d}x^k} \frac{{}_1F_1(1, n+2; x) - 1}{(n+1)!}$$

Table 9 Properties of $r_{1,n}(x)$, and $r_{2,n}(x)$

x	$r_{1,n}$	$r'_{1,n}$	$r''_{1,n}$	$r_{2,n}$	$r'_{2,n}$	$r''_{2,n}$
$(-\infty, 0)$	$-$	$+$	$+$	$+$	$-$	$+$
0	0	$+$	$+$	0	0	$+$
$(0, \infty)$	$+$	$+$	$+$	$+$	$+$	$+$

$$= \frac{k!}{(k+n+1)!} \left[{}_1F_1(k+1, k+n+2; x) - 1_{k=0} \right] \tag{56}$$

$$= \frac{k!}{(k+n+1)!} \left[e^x {}_1F_1(n+1, k+n+2; -x) - 1_{k=0} \right] \quad k, n \in \mathbb{N}_0^+ \tag{57}$$

with 1_A the indicator function which equals 1 if the event A is true, and 0 otherwise. Equations (56) and (57) indicate $r_{1,n}(0) = 0$ and $\frac{d^k}{dx^k} r_{1,n}(x) > 0$ for all $x \in \mathbb{R}$ and $k \in \{1, 2\}$. Thus, the function $r_{1,n}$ is increasing and convex with $r_{1,n}(x) < 0$ if $x < 0$, $r_{1,n}(0) = 0$, $r_{1,n}(x) > 0$ if $x > 0$, and $r'_{1,n}(0) = 1/(n+1)!$.

Since $r_{2,n}(x) = x r_{1,n+1}(x)$ and $r_{2,n}(x) = r_{1,n}(x) - x/(n+1)!$, the function $r_{2,n}$ is nonnegative and convex with $r_{2,n}(0) = r'_{2,n}(0) = 0$. Therefore, $r_{2,n}(x)$ decreases on $(-\infty, 0)$ and increases on $(0, \infty)$.

Next, we apply the properties in Table 9 on $\xi_n(x) = r_{1,n}(x)$. Consider $\tilde{X}_n$ with the cdf $\tilde{F}_n$ in terms of the cdf F of X,

$$d\tilde{F}_n(x) = \frac{x^n \, dF(x)}{\int x^n \, dF(x)} = \frac{x^n \, dF(x)}{\mathbb{E}[X^n]}.$$

It follows that $\mathbb{E}\left[r_{1,n}(\tilde{X}_n) \right] = \xi_n(X)$ and $\mathbb{E}\left[\tilde{X}_n \right] = m_{n+1}$. Using Lemma 4(ii), Jensen's inequality, and the convexity of $r_{1,n}$, we find the lower bounds on $\xi_n(X^+)$ and $\xi_n(-X^-)$:

$$0 \le \xi_n(m_{n+1}^+) = r_{1,n}(m_{n+1}^+) = r_{1,n}\left(\mathbb{E}\left[\tilde{X}_n^+ \right] \right) \le \mathbb{E}\left[r_{1,n}(\tilde{X}_n^+) \right] = \xi_n(X^+)$$

$$\xi_n(-m_{n+1}^-) = r_{1,n}(-m_{n+1}^-) = r_{1,n}\left(\mathbb{E}\left[-\tilde{X}_n^- \right] \right) \le \mathbb{E}\left[r_{1,n}(-\tilde{X}_n^-) \right] = \xi_n(-X^-).$$

The upper bound on $\xi_n(-X^-)$ follows by applying Jensen's inequality to the convex function $r_{2,n-1}$:

$$\xi_n(-X^-) = \frac{1}{m_n^- \mathbb{E}[(-X^-)^{n-1}]} \sum_{i=1}^{\infty} \frac{\mathbb{E}\left[(-X^-)^{n+i} \right]}{(n+i)!} = -\frac{1}{m_n^-} \sum_{i=2}^{\infty} \frac{\mathbb{E}\left[(-\tilde{X}_{n-1}^-)^i \right]}{(n-1+i)!}$$

$$= -\frac{1}{m_n^-} \mathbb{E}\left[r_{2,n-1}(-\tilde{X}_{n-1}^-) \right]$$

$$\leq -\frac{1}{m_n^-} r_{2,n-1}\left(\mathbb{E}\left[-\tilde{X}_{n-1}^-\right]\right) = -\frac{1}{m_n^-} \sum_{i=2}^{\infty} \frac{(-m_n^-)^i}{(n-1+i)!} = \sum_{i=1}^{\infty} \frac{(-m_n^-)^i}{(n+i)!}$$

$$= \xi_n(-m_n^-) = r_{1,n}(-m_n^-) \leq 0$$

(ii) The moments of $T_\theta \sim \Gamma(k,\theta)$ are $\mathbb{E}\left[T_\theta^n\right] = \theta^n \Gamma(k+n)/\Gamma(k)$, which gives $m_{n+1}(T_\theta) = \theta(k+n)$. By Lemma 5(i), we have for any mixture of Gamma distributions with fixed shape parameter k,

$$\frac{k+n}{k+n-1} \leq \frac{m_{n+1}}{m_n}. \tag{58}$$

Combining (58), $\xi_n'(x) = r_{1,n}'(x) > 0$ $(x \in \mathbb{R})$, and Lemma 9(i) produces the lower bound on $\xi_n(X^+)$ in Lemma 9(ii). The scaled remainder of T_θ is

$$\begin{aligned}
\xi_n(T_\theta) &:= \frac{1}{\mathbb{E}\left[T_\theta^n\right]} \sum_{i=1}^{\infty} \frac{\mathbb{E}\left[T_\theta^{n+i}\right]}{(n+i)!} \\
&= \frac{\Gamma(k)}{n!\,\theta^n \Gamma(k+n)} \sum_{i=1}^{\infty} \frac{\theta^{n+i} \Gamma(k+n+i)}{(n+1)\dots(n+i)\,0(k)} \\
&= \frac{1}{n!} \sum_{i=1}^{\infty} \theta^i \prod_{j=1}^{i} \left(\frac{k-1}{n+j}+1\right) \\
&= \frac{1}{n!} \left[{}_2F_1(1, k+n; n+1; \theta) - 1\right].
\end{aligned}$$

By well-known properties of the Gaussian hypergeometric function ${}_2F_1$, the function $h_n(\theta) := \xi_n(T_\theta)$ increases monotonically on $(-1,1)$ with $h_n(0) = 0$. Because $T_{[\theta_{\max}]^+}$ stochastically dominates the mixture X^+ and $\xi_n'(x) > 0$ for $x \in \mathbb{R}$, we must have $\xi_n(X^+) \leq \xi_n(T_{[\theta_{\max}]^+})$ if $[\theta_{\min}, \theta_{\max}] \subseteq (-1,1)$. Similarly, the mixture $-X^-$ stochastically dominates $-T_{[\theta_{\min}]^-}$, which leads to $\xi_n(T_{-[\theta_{\min}]^-}) \leq \xi_n(-X^-)$. This proves (ii).

$\square$

References

Abramowitz M, Stegun IA (1964) Handbook of mathematical functions: with formulas, graphs, and mathematical tables, vol 55. Courier Corporation, Chelmsford

Akhiezer NI (1965) The classical moment problem: and some related questions in analysis, vol 5. Oliver & Boyd, Edinburgh

Barnett NS, Dragomir SS, Agarwal R (2002) Some inequalities for probability, expectation, and variance of random variables defined over a finite interval. Comput Math Appl 43(10):1319–1357

Bawa VS, Lindenberg EB (1977) Capital market equilibrium in a mean-lower partial moment framework. J Financ Econ 5(2):189–200

Berg C (1995) Indeterminate moment problems and the theory of entire functions. J Comput Appl Math 65(1):27–55

Dokov SP, Morton DP (2005) Second-order lower bounds on the expectation of a convex function. Math Oper Res 30(3):662–677

Frauendorfer K (1988) Solving slp recourse problems with arbitrary multivariate distributions-the dependent case. Math Oper Res 13(3):377–394

Gavriliadis PN (2008) Moment information for probability distributions, without solving the moment problem, I: where is the mode? Commun Stat Theory Methods 37(5):671–681

Gavriliadis PN, Athanassoulis GA (2009) Moment information for probability distributions, without solving the moment problem, II: main-mass, tails and shape approximation. J Comput Appl Math 229(1):7–15

Goria MN, Tagliani A (2003) Bounds on the tail probability and absolute difference between two distributions. Communications in Statistics-Theory and Methods 32(3):519–532

Gut A (2002) On the moment problem. Bernoulli 3(8):407–421

Ibragimov IA (1956) On the composition of unimodal distributions. Theory Prob Appl 1(2):255–260

Kall P (1991) An upper bound for SLP using first and total second moments. Ann Oper Res 30(1):267–276

Kreĭn MG, Nudelman AA (1977) The Markov moment problem and extremal problems: ideas and problems of P.L. Čebyšev and A.A. Markov and their further development. American Mathematical Society, Providence

Lin GD (1997) On the moment problems. Stat Prob Lett 35(1):85–90

Lindsay BG, Basak P (2000) Moments determine the tail of a distribution (but not much else). Am Stat 54(4):248–251

Magnus JR (1986) The exact moments of a ratio of quadratic forms in normal variables. Annales d'Économie et de Statistique 4:95–109

Milev M, Novi Inverardi P, Tagliani A (2012) Moment information and entropy evaluation for probability densities. Appl Math Comput 218(9):5782–5795

Mnatsakanov RM (2008a) Hausdorff moment problem: reconstruction of distributions. Stat Prob Lett 78(12):1612–1618

Mnatsakanov RM (2008b) Hausdorff moment problem: reconstruction of probability density functions. Stat Prob Lett 78(13):1869–1877

Mnatsakanov RM, Hakobyan AS (2009) Recovery of distributions via moments. Lect Notes-monogr Ser 57:252–265

Pakes AG, Hung WL, Wu JW (2001) Criteria for the unique determination of probability distributions by moments. Aust N Z J Stat 43(1):101–111

Shohat JA, Tamarkin JD (1943) The problem of moments, vol 1. American Mathematical Society, Providence

Stoyanov J (2000) Krein condition in probabilistic moment problems. Bernoulli 5(6):939–949

Stoyanov JM (2013) Counterexamples in probability. Courier Corporation, Chelmsford

Szegö G (1975) Orthogonal polynomials, 4th edn. American Mathematical Society, Providence

Winkler RL, Roodman GM, Britney RR (1972) The determination of partial moments. Manage Sci 19(3):290–296