ESAIM: PS
August 2007, Vol. 11, p. 412–426
DOI: 10.1051/ps:2007027

ESAIM: Probability and Statistics
www.edpsciences.org/ps

TOWARD THE BEST CONSTANT FACTOR FOR THE RADEMACHER-GAUSSIAN TAIL COMPARISON

Iosif Pinelis[1]

Abstract. It is proved that the best constant factor in the Rademacher-Gaussian tail comparison is between two explicitly defined absolute constants c_1 and c_2 such that $c_2 \approx 1.01\, c_1$. A discussion of relative merits of this result *versus* limit theorems is given.

Mathematics Subject Classification. 60E15, 62G10, 62G15, 60G50, 62G35.

Received June 7, 2006. Revised November 2, 2006.

Introduction and summary

Let $\varepsilon_1, \dots, \varepsilon_n$ be independent Rademacher random variables (r.v.'s), so that $\mathsf{P}(\varepsilon_i = 1) = \mathsf{P}(\varepsilon_i = -1) = \frac{1}{2}$ for all i. Let

$$S_n := a_1\varepsilon_1 + \cdots + a_n\varepsilon_n,$$

where $a_1, \dots, a_n$ are any real numbers such that

$$a_1^2 + \cdots + a_n^2 = 1.$$

The best upper exponential bound on the tail probability $\mathsf{P}(Z \geqslant x)$ for a standard normal random variable Z and a nonnegative number x is $\inf_{t\geqslant 0} \mathrm{e}^{-tx}\, \mathsf{E}\, \mathrm{e}^{tZ} = \mathrm{e}^{-x^2/2}$. Thus, a factor of the order of magnitude of $\frac{1}{x}$ is "missing" in this bound, compared with the asymptotics $\mathsf{P}(Z \geqslant x) \sim \frac{1}{x}\,\varphi(x)$ as $x \to \infty$, where $\varphi(x) := \mathrm{e}^{-x^2/2}/\sqrt{2\pi}$ is the density function of Z. Now it should be clear that any exponential upper bound on the tail probabilities for sums of independent random variables must be missing the $\frac{1}{x}$ factor.

Eaton [6] obtained an upper bound on $\mathsf{P}(S_n \geqslant x)$, which is asymptotic to $c_3\, \mathsf{P}(Z \geqslant x)$ as $x \to \infty$, where

$$c_3 := \frac{2\mathrm{e}^3}{9} \approx 4.46,$$

and he conjectured that $\mathsf{P}(S_n \geqslant x) \leqslant c_3 \frac{1}{x}\,\varphi(x)$ for $x > \sqrt{2}$. The stronger form of this conjecture,

$$\mathsf{P}(S_n \geqslant x) \leqslant c\, \mathsf{P}(Z \geqslant x) \tag{1}$$

Keywords and phrases. Probability inequalities, Rademacher random variables, sums of independent random variables, Student's test, self-normalized sums.

[1] Department of Mathematical Sciences, Michigan Technological University, Houghton, Michigan 49931 USA; ipinelis@mtu.edu

© EDP Sciences, SMAI 2007

Article published by EDP Sciences and available at http://www.esaim-ps.org or http://dx.doi.org/10.1051/ps:2007027

for all $x \in \mathbb{R}$ with $c = c_3$ was proved by Pinelis [11], along with a multidimensional extension.

Edelman [7] proposed inequality $\mathsf{P}(S_n \geqslant x) \leqslant \mathsf{P}(Z \geqslant x - 1.5/x)$ for all $x > 0$, but his proof appears to have a gap. A more precise upper bound, with $\ln c_3 = 1.495\ldots$ in place of 1.5, was recently shown [20] to be a rather easy corollary of the mentioned result of [11]. Various generalizations and improvements of inequality (1) as well as related results were given by Pinelis [12, 13, 15–17, 19, 20] and Bentkus [1–3].

Bobkov, Götze and Houdré (BGH) [4] gave a simple proof of (1) with a constant factor $c \approx 12.01$. Their method was based on the Chapman-Kolmogorov identity for the Markov chain (S_n). Such an identity was used, *e.g.*, in [14] to disprove a conjecture by Graversen and Peškir [9] on $\max_{k \leqslant n} |S_k|$.

In this paper, we shall show that a modification of the BGH method can be used to obtain inequality (1) with a constant factor $c \approx 1.01\, c_*$, where c_* is the best (that is, the smallest) possible constant factor c in (1).

Let $\overline{\Phi}$ and r denote the tail function of Z and the inverse Mills ratio:

$$\overline{\Phi}(x) := \mathsf{P}(Z \geqslant x) = \int_x^\infty \varphi(u)\,\mathrm{d}u \quad \text{and} \quad r := \frac{\varphi}{\overline{\Phi}}\cdot$$

Theorem 1 (Main)**.** *For the least possible absolute constant factor c_* in inequality* (1) *one has*

$$c_* \in [c_1, c_2] \approx [3.18, 3.22], \quad \textit{where}$$
$$c_1 := \frac{1}{4\overline{\Phi}(\sqrt{2})} \quad \textit{and} \quad c_2 := c_1 \cdot \Big(1 + \tfrac{1}{250}\big(1 + r(\sqrt{3}\,)\big)\Big) \approx c_1 \cdot 1.01.$$

Here we shall present just one application of Theorem 1, to self-normalized sums

$$V_n := \frac{X_1 + \cdots + X_n}{\sqrt{X_1^2 + \cdots + X_n^2}},$$

where, following Efron [8], we assume that the X_i's satisfy the orthant symmetry condition: the joint distribution of $\delta_1 X_1, \ldots, \delta_n X_n$ is the same for any choice of signs $\delta_1, \ldots, \delta_n \in \{1, -1\}$, so that, in particular, each X_i is symmetrically distributed. It suffices that the X_i's be independent and symmetrically (but not necessarily identically) distributed. In particular, $V_n = S_n$ if $X_i = a_i \varepsilon_i\ \forall i$. It was noted by Efron that (i) Student's statistic T_n is a monotonic function of the self-normalized sum: $T_n = \sqrt{\frac{n-1}{n}}\, V_n / \sqrt{1 - V_n^2/n}$ and (ii) the orthant symmetry implies in general that the distribution of V_n is a mixture of the distributions of normalized Rademacher sums S_n. Thus, one obtains

Corollary 1. *Theorem 1 holds with V_n in place of S_n.*

Various limit theorems for sums and self-normalized sums are available. In particular, the central limit theorem approximation for $\mathsf{P}(S_n \geqslant x)$ and $\mathsf{P}(V_n \geqslant x)$ is simply $\mathsf{P}(Z \geqslant x)$, without any extra factor c. However, (i) such asymptotic relations, without an upper bound on the rate of convergence, are impossible to use in statistical practice when one needs to be certain that the tail probability does not exceed a prescribed level; (ii) when an upper bound (say of the Berry-Esseen type) on the rate of convergence is available, usually it is relatively too large to be useful in statistics, especially in the tail area; (iii) usually, large deviation asymptotics are valid at best in the zone $x = o(n)$, and this zone is defined only qualitatively; (iv) the summands $X_1, \ldots, X_n$ are usually required to be identically, or nearly identically, distributed. If these conditions fail to hold then – as Theorem 1, Corollary 1, and the discussion below in the beginning of Section 1 show – the asymptotic approximations may be inadequate. Also, it was pointed out in Theorem 2.8 of [11], that, since the normal tail decreases fast, inequality (1) even with $c \approx 4.46$ implies that relevant quantiles of S_n and V_n may exceed the corresponding standard normal quantiles only by a relatively small amount; thus, one can use Corollary 1 rather efficiently to test symmetry even for non-i.i.d. observations.

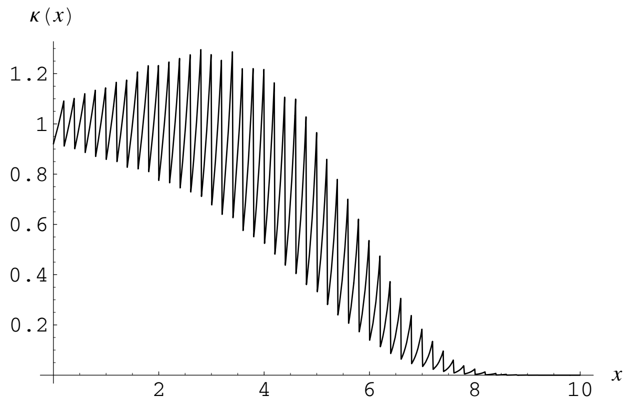

FIGURE 1. Ratio of the Rademacher and Gaussian tails for $n = 100$ and $a_1 = \cdots = a_{100} = \frac{1}{10}$.

1. Proof of Theorem 1

Theorem 1 follows immediately from Lemma 1, Theorem 2, and Lemma 3, stated in Section 1.1 below. In particular, Lemma 3 implies that the upper bound $h_1(x)$ on $\mathsf{P}(S_n \geqslant x)$ provided by Theorem 2 is somewhat better than the upper bound $c_2\,\mathsf{P}(Z \geqslant x)$, implied by Theorem 1.

While S_n represents a simplest case of the sum of independent non-identically distributed r.v.'s, it is still very difficult to control in a precise manner. Figure 1 shows the graph of the ratio $\kappa(x) := \mathsf{P}(S_n \geqslant x)/\,\mathsf{P}(Z \geqslant x)$ for $n = 100$ and $a_1 = \cdots = a_n$.

One can see that even for such a fairly large value of n and equal coefficients $a_1, \dots, a_n$, ratio $\kappa(x)$ oscillates rather wildly. In view of this, the existence of a high-precision inductive argument in the general setting with possibly unequal a_i's may seem very unlikely. However, such an argument will be presented in this paper.

The key idea in the proof of Theorem 1 is the construction of the upper bound h_1 and, in particular, the function g defined by (3) and (2), which allows an inductive argument to prove Theorem 2, a refined version of Theorem 1.

The proof of Theorem 2 is based on a number of lemmas. The proofs of all lemmas are deferred to Section 1.2.

1.1. Statements of lemmas and the proof of Theorem 2

Lemma 1. *One has $c_* \geqslant c_1$.*

For $a \in [0, 1)$ and $x \in \mathbb{R}$, introduce

$$g(x) := c_1 \cdot \Big(1 + \tfrac{1}{250}\big(1 + r(x)\big)\Big)\overline{\Phi}(x) = \tfrac{c_1}{250} \cdot \big(251\,\overline{\Phi}(x) + \varphi(x)\big); \tag{2}$$

$$h(x) := c_1 \cdot \Big(1 + \tfrac{1}{250}\big(1 + r(\sqrt{3}\,)\big)\Big)\overline{\Phi}(x) = c_2 \cdot \overline{\Phi}(x);$$

$$h_1(x) := \begin{cases} 1 & \text{if } x \leqslant 0, \\ \frac{1}{2} & \text{if } 0 < x \leqslant 1, \\ \frac{1}{2x^2} & \text{if } 1 \leqslant x < \sqrt{2}, \\ g(x) & \text{if } \sqrt{2} \leqslant x \leqslant \sqrt{3}, \\ h(x) & \text{if } x \geqslant \sqrt{3}; \end{cases} \tag{3}$$

$$K(a,x) := h_1(u) + h_1(v) - 2h_1(x), \quad \text{where} \tag{4}$$

$$u := u(a,x) := \frac{x-a}{\sqrt{1-a^2}} \quad \text{and} \tag{5}$$

$$v := v(a,x) := \frac{x+a}{\sqrt{1-a^2}}. \tag{6}$$

Theorem 2 (Refined)**.** *One has*

$$\mathsf{P}(S_n \geqslant x) \leqslant h_1(x) \tag{7}$$

for all $x \in \mathbb{R}$.

Lemma 2. *One has* $g \leqslant h$ *on* $(-\infty, \sqrt{3}\,]$ *and* $g \geqslant h$ *on* $[\sqrt{3}, \infty)$.

Lemma 3. *One has* $h_1 \leqslant h$ *on* $\mathbb{R}$.

Lemma 4. *One has* $K(a,x) \leqslant 0$ *for all* $(a,x) \in [0,1) \times [\sqrt{2}, \sqrt{3}\,]$.

Lemma 5. *One has* $K(a,x) \leqslant 0$ *for all* $(a,x) \in [0,1) \times [\sqrt{3}, \infty)$.

Now we can present

Proof of Theorem 2. Theorem 2 will be proved by induction in n. It is obvious for $n = 1$. Let now $n \in \{2, 3, \dots\}$ and assume that Theorem 2 holds with $n-1$ in place of n.

Note that for $x \leqslant 0$ inequality (7) is trivial. For $x \in (0, \sqrt{2})$, it follows by the symmetry of S_n and Chebyshev's inequality. Therefore, assume without loss of generality that $x \geqslant \sqrt{2}$ and $0 \leqslant a_n < 1$. By the Chapman-Kolmogorov identity and induction,

$$\begin{aligned} \mathsf{P}(S_n \geqslant x) &= \tfrac{1}{2}\, \mathsf{P}(S_{n-1} \geqslant x - a_n) + \tfrac{1}{2}\, \mathsf{P}(S_{n-1} \geqslant x + a_n) \\ &\leqslant \tfrac{1}{2}\, h_1\big(u(a_n, x)\big) + \tfrac{1}{2}\, h_1\big(v(a_n, x)\big) \\ &= h_1(x) + \tfrac{1}{2}\, K(a_n, x) \end{aligned}$$

for all $x \in \mathbb{R}$. It remains to refer to Lemmas 4 and 5. ☐

Lemmas 2, 3, and 5 are much easier to prove than Lemma 4. The least elementary of Lemmas 2, 3, and 5 is Lemma 3, whose proof uses the following l'Hospital-type rule for monotonicity.

Proposition 1. [18] *Let* $-\infty \leqslant a < b \leqslant \infty$. *Let* f *and* g *be real-valued differentiable functions defined on the interval* (a,b) *such that* $f(b-) = g(b-) = 0$. *It is assumed that* g *and* g' *do not take on the zero value on* (a,b). *Suppose finally that* $\frac{f'}{g'} \nearrow\searrow$ *on* (a,b); *that is,* $\frac{f'}{g'}$ *switches from* $\nearrow$ *(increase) to* $\searrow$ *(decrease) on* (a,b). *Then* $\frac{f}{g} \nearrow\searrow$ *or* $\searrow$ *on* (a,b).

This proposition follows immediately from [18, Proposition 4.3 and Remark 5.3].

A significant difficulty in the proof of Lemma 4 is that the "profiles" of the function K given by the cross-sections of the graph of K on rectangle

$$\begin{aligned} R &:= \{(a,x) \in \mathbb{R}^2 \colon 0 < a < 1, \sqrt{2} < x < \sqrt{3}\} \quad \text{(or on its closure} \\ \overline{R} &:= \{(a,x) \in \mathbb{R}^2 \colon 0 \leqslant a \leqslant 1, \sqrt{2} \leqslant x \leqslant \sqrt{3}\}\ ) \end{aligned}$$

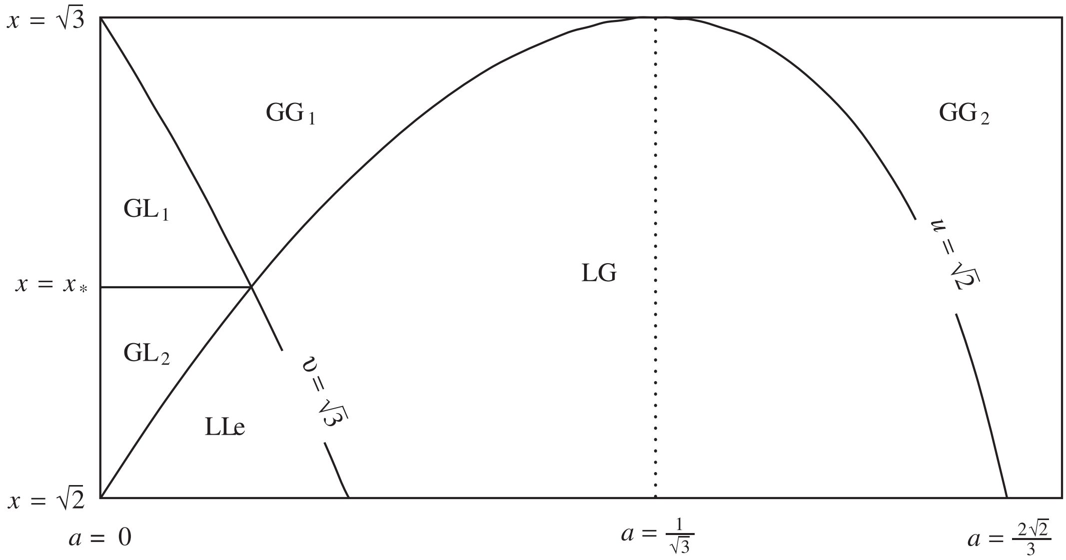

FIGURE 2. Partition of $R = (0,1) \times (\sqrt{2}, \sqrt{3})$.

with $a \in (0,1)$ or $x \in (\sqrt{2}, \sqrt{3})$ fixed are very complicated; in part, this is caused by the fact that the function h_1 is defined in (3) by three different expressions over the interval $[1, \infty)$. To overcome this difficulty, an idea is to partition R into "pieces" so that on each piece there is a direction in which the profiles of K are easier to deal with. This idea comes naturally from the following considerations.

Remark 1.

- Observe that $u = u(a,x) \in (1, \sqrt{3})$ and $v = v(a,x) \in (\sqrt{2}, \infty)$ for all $(a,x) \in R$. Therefore and in view of definitions (4) and (3), the form of expression of K on R depends on whether $u < \sqrt{2}$ and on whether $v < \sqrt{3}$. The curves $u = \sqrt{2}$ and $v = \sqrt{3}$ (which are ellipses) partition R into 5 connected "pieces" (as illustrated by Fig. 2), one of which may be naturally cut further into two pieces by the line $x = x_*$, where

$$x_* := \sqrt{\frac{5 + 2\sqrt{6}}{9 - 2\sqrt{6}}} = \sqrt{\frac{23 + 28\sqrt{\frac{2}{3}}}{19}} \approx 1.55 \tag{8}$$

may be also defined by the following condition: $\big(\, (a,x) \in R \;\;\&\;\; u = \sqrt{2} \;\;\&\;\; v = \sqrt{3} \, \big) \implies x = x_*$.

Thus, one comes to the following definitions of the "pieces":

$$
\begin{aligned}
\mathrm{LLe} &:= \{(a,x)\in R\colon u<\sqrt{2}, v\leqslant\sqrt{3}\};\\
\mathrm{LG} &:= \{(a,x)\in R\colon u<\sqrt{2}, v>\sqrt{3}\};\\
\mathrm{GL}_1 &:= \{(a,x)\in R\colon u>\sqrt{2}, v<\sqrt{3}, x\leqslant x_*\};\\
\mathrm{GL}_2 &:= \{(a,x)\in R\colon u>\sqrt{2}, v<\sqrt{3}, x> x_*\};\\
\mathrm{GG}_1 &:= \{(a,x)\in R\colon u>\sqrt{2}, v>\sqrt{3}, a<\tfrac{1}{\sqrt{3}}\};\\
\mathrm{GG}_2 &:= \{(a,x)\in R\colon u>\sqrt{2}, v>\sqrt{3}, a\geqslant\tfrac{1}{\sqrt{3}}\};\\
\mathrm{GE} &:= \{(a,x)\in R\colon u>\sqrt{2}, v=\sqrt{3}\};\\
\mathrm{ELe} &:= \{(a,x)\in R\colon u=\sqrt{2}, v\leqslant\sqrt{3}\};\\
\mathrm{EG}_1 &:= \{(a,x)\in R\colon u=\sqrt{2}, v>\sqrt{3}, a<\tfrac{1}{\sqrt{3}}\};
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{EG}_2 &:= \{(a,x)\in R\colon u=\sqrt{2}, v>\sqrt{3}, a\geqslant\tfrac{1}{\sqrt{3}}\};\\
\mathrm{A}_1 &:= \{(a,x)\in \overline{R}\colon a=0\};\\
\mathrm{A}_2 &:= \{(a,x)\in \overline{R}\colon a=1\};\\
\mathrm{X}_{1,1} &:= \{(a,x)\in \overline{R}\colon x=\sqrt{2}, 0<a<\tfrac{1}{\sqrt{2}}\};\\
\mathrm{X}_{1,2} &:= \{(a,x)\in \overline{R}\colon x=\sqrt{2}, \tfrac{1}{\sqrt{2}}\leqslant a<\tfrac{2\sqrt{2}}{3}\};\\
\mathrm{X}_{1,3} &:= \{(a,x)\in \overline{R}\colon x=\sqrt{2}, \tfrac{2\sqrt{2}}{3}\leqslant a<1\};\\
\mathrm{X}_2 &:= \{(a,x)\in \overline{R}\colon x=\sqrt{3}\},
\end{aligned}
$$

where u and v are defined by (5) and (6). Here, for example, the L in the first position in symbol LLe refers to "less than" in inequality $u<\sqrt{2}$, while the ligature Le in the second position refers to "less than or equal to" in inequality $v\leqslant\sqrt{3}$. Similarly, G and E in this notation refer to "greater than" and "equal to", respectively. Symbol A refers here to a fixed value of a, and X to a fixed value of x.

Observe that the set $\overline{R}$ is the union of the "pieces" $\mathrm{LLe},\dots,\mathrm{X}_2$. Indeed, let $\{C\}$ denote, for brevity, the set $\{(a,x)\in R\colon C\}$, where C stands for a condition. Then, basically following the just explained meaning of the notation for the "pieces", one has

$$
\underbrace{\mathrm{LLe}\cup\mathrm{LG}}_{\{u<\sqrt{2}\}}\cup\underbrace{\underbrace{\mathrm{GL}_1\cup\mathrm{GL}_2}_{\{u>\sqrt{2},v<\sqrt{3}\}}\cup\underbrace{\mathrm{GG}_1\cup\mathrm{GG}_2}_{\{u>\sqrt{2},v>\sqrt{3}\}}\cup\mathrm{GE}}_{\{u>\sqrt{2}\}}\cup\underbrace{\mathrm{ELe}\cup\underbrace{\mathrm{EG}_1\cup\mathrm{EG}_2}_{\{u=\sqrt{2},v>\sqrt{3}\}}}_{\{u=2\}}=R. \tag{9}
$$

(In fact, the sets on the left-hand side of (9) form a partition of R.) It is also clear that the union $\mathrm{A}_1\cup\mathrm{A}_2\cup\mathrm{X}_{1,1}\cup\mathrm{X}_{1,2}\cup\mathrm{X}_{1,3}\cup\mathrm{X}_2$ equals the boundary of $\overline{R}$. This verifies the entire "union-observation", which is illustrated in Figure 2, where only the labels of the two-dimensional members of the partition of R are shown.

It will be understood that the function K (defined by (4)) is extended to A_2 by continuity, so that

$$
K(a,x):=-2g(x)\quad\forall(a,x)\in\mathrm{A}_2\,. \tag{10}
$$

- Another difficulty to overcome in the proof of Lemma 4 is that the expressions for h_1 and hence for K contain the transcendental function $\overline{\Phi}$. Yet, fixing an arbitrary value of u, v, or x, one will thereby fix the value of one of the three terms in expression (4) – $h_1(u)$, $h_1(v)$, or $2h_1(x)$, respectively, so that the (partial) derivative of K in (say) a with the value of u or v or x fixed will be of the form $k_1 := A_1 e^{A_2} + A_3 e^{A_4}$, where $A_1, \dots, A_4$ are algebraic expressions in (a, x) or, equivalently, in (a, u) or (a, v). Assuming that $\operatorname{sign} A_3$ is constant and nonzero on a given "piece" (say P) of rectangle R, the sign of k_1 on P is the same as or opposite to that of $\tilde{k}_1 := A_1 e^{A_2 - A_4}/A_3 + 1$. Therefore, the sign (on P) of the derivative (say k_2) of $\tilde{k}_1$ in (say) a is the same as that of a certain algebraic expression. Since the equations of the boundaries of the "pieces" are algebraic as well, the sign pattern of k_2 on P can be determined in a completely algorithmic manner, according to a result by Tarski [5, 10, 21]. The implementation of this scheme will require a great amount of symbolic and numerical computation. We have done that with the help of Mathematica™ 5.2, which is rather effective and allows complete and easy control over the accuracy. The Tarski algorithm is implemented in Mathematica 5.2 via `Reduce` and related commands. In particular, command

```
Reduce[cond1 && cond2 && ..., {var1,var2,...,}, Reals]
```

returns a simplified form of the system of algebraic conditions (equations or inequalities) `cond1`, `cond2`, ... over real variables `var1`, `var2`, ... However, the execution of such a command may take a long time if the algebraic system is more than a little complicated; in such cases, Mathematica can use some human help.

On each of the "pieces" ELe, ..., X_2 we shall consider its own "coordinate system", the one that will be most convenient for us. In particular, we shall use the "coordinate" pair (a, v) on each of the pieces LLe and LG; the pair (a, x) on each of the pieces GL_1, GL_2, and GG_2; and the pair (a, u) on the piece GG_1. The choice of these coordinate systems was motivated by a consideration of contour plots of the function K on R and on particular "pieces". For instance, it will be shown in the proofs of Lemmas 6 and 7 below that the monotonicity pattern of the function K in a with v fixed over each of the "pieces" LLe and LG is $\searrow$ (decreasing), $\nearrow$ (increasing), or $\searrow\nearrow$ (switching from decrease to increase as a increases). Similarly over the other two-dimensional pieces: it will be shown that K is (i) $\searrow$ or $\searrow\nearrow$ in a with x fixed over each of the "pieces" GL_1 and GL_2; (ii) $\searrow$ in a with u fixed over GG_1; and (iii) $\searrow$ in a with x fixed over GG_2. The monotonicity patterns of K over the one-dimensional "pieces" of $\overline{R}$ are somewhat easier to establish.

Remark 2. Since the function h_1 defined by (3) is upper-semicontinuous on $\mathbb{R}$ and continuous on $[\sqrt{2}, \infty)$, the function K is upper-semicontinuous on the compact rectangle $\overline{R}$ and therefore attains its maximum on $\overline{R}$. We conclude that K attains its maximum over $\overline{R}$ on at least one of the sets LLe, ..., X_2.

In view of this remark, the main lemma – Lemma 4 – will be easily obtained from the following series of lemmas.

Lemma 6 (LLe)**.** *The function K does not attain a maximum on* LLe*.*

Lemma 7 (LG)**.** *The function K does not attain a maximum on* LG*.*

Lemma 8 (GL_1)**.** *The function K does not attain a maximum on* GL_1*.*

Lemma 9 (GL_2)**.** *The function K does not attain a maximum on* GL_2*.*

Lemma 10 (GG_1)**.** *The function K does not attain a maximum on* GG_1*.*

Lemma 11 (GG_2)**.** *The function K does not attain a maximum on* GG_2*.*

Lemma 12 (GE)**.** *One has $K \leqslant 0$ on* GE*.*

Lemma 13 (ELe)**.** *One has $K \leqslant 0$ on* ELe*.*

Lemma 14 (EG_1)**.** *One has* $K \leqslant 0$ *on* EG_1*.*

Lemma 15 (EG_2)**.** *One has* $K \leqslant 0$ *on* EG_2*.*

Lemma 16 (A_1)**.** *One has* $K = 0$ *on* A_1*.*

Lemma 17 (A_2)**.** *One has* $K \leqslant 0$ *on* A_2*.*

Lemma 18 ($\mathrm{X}_{1,1}$)**.** *One has* $K \leqslant 0$ *on* $\mathrm{X}_{1,1}$*.*

Lemma 19 ($\mathrm{X}_{1,2}$)**.** *One has* $K \leqslant 0$ *on* $\mathrm{X}_{1,2}$*.*

Lemma 20 ($\mathrm{X}_{1,3}$)**.** *One has* $K \leqslant 0$ *on* $\mathrm{X}_{1,3}$*.*

Lemma 21 (X_2)**.** *One has* $K \leqslant 0$ *on* X_2*.*

1.2. **Proofs of the lemmas**

Proof of Lemma 1. Let $n = 2$ and $a_1 = a_2 = \frac{1}{\sqrt{2}}$. Then $\mathsf{P}(S_n \geqslant \sqrt{2}\,) = \frac{1}{4} = c_1\,\mathsf{P}(Z \geqslant \sqrt{2}\,)$. □

Proof of Lemma 2. This follows from the well-known and easy-to-prove fact that the inverse Mills ratio r is increasing. □

Proof of Lemma 3. On interval $(-\infty, 0]$, one has $h_1 = 1 < 1.6 < h(0) \leqslant h$, since h is decreasing.

On interval $(0, 1]$, one similarly has $h_1 = \frac{1}{2} < 0.51 < h(1) \leqslant h$.

On interval $[\sqrt{2}, \sqrt{3}\,]$, one has $h_1 = g \leqslant h$, by Lemma 2.

On interval $[\sqrt{3}, \infty)$, one has $h_1 = h$.

It remains to consider the interval $(1, \sqrt{2})$. For $x \in (1, \sqrt{2})$, one has $h_1(x) = p(x) := \frac{1}{2x^2}$. One has $p(\infty-) = h(\infty-) = 0$ and, for some constant $C > 0$, $h'(x)/p'(x) = Cx^3\varphi(x)$, so that $\frac{h'}{p'} \nearrow\searrow$ on $(0, \infty)$. Hence, in view of the l'Hospital-type rule for monotonicity provided by Proposition 1, one has $\frac{h}{p} \searrow$ or $\nearrow\searrow$ on $(0, \infty)$, and so, $\inf_{(1,\sqrt{2})} \frac{h}{h_1} = \min_{[1,\sqrt{2}]} \frac{h}{p} = \min_{\{1,\sqrt{2}\}} \frac{h}{p} \approx 1.01 > 1$, whence $h_1 < h$ on $(1, \sqrt{2})$. □

Proof of Lemma 4. This lemma follows from Remark 2 and Lemmas 6–21, proved below. Indeed, by the conclusion of Remark 2, the function K attains its maximum over $\overline{R}$ on at least one of the sets $\mathrm{LLe}, \dots, \mathrm{X}_2$. By Lemmas 6–11, none of the sets $\mathrm{LLe}, \dots, \mathrm{GG}_2$ can be a set on which K attains its maximum over $\overline{R}$. Therefore, K attains its maximum over $\overline{R}$ on at least one of the sets $\mathrm{GE}, \dots, \mathrm{X}_2$. Finally, Lemmas 12–21 imply that this maximum is no greater than 0. □

Proof of Lemma 5. Let $0 \leqslant a < 1$ and $x \geqslant \sqrt{3}$. Then $u \geqslant \sqrt{2}$ and $v \geqslant \sqrt{3}$, where u and v are defined by (5) and (6), as before. Therefore and in view of Lemma 3 and (3), one has $h_1(u) \leqslant h(u)$, $h_1(v) = h(v)$, and $h_1(x) = h(x)$, so that

$$K(a, x) \leqslant 2c_2 \cdot \big(\tfrac{1}{2}\overline{\Phi}(u) + \tfrac{1}{2}\overline{\Phi}(v) - \overline{\Phi}(x)\big) \leqslant 0,$$

as shown in the mentioned proof in [4]. □

Proof of Lemma 6 (LLe). Expressing x and u in view of (5) and (6) in terms of a and v as $x(a, v) := \sqrt{1 - a^2}\,v - a$ and $\tilde{u}(a, v) := v - \frac{2a}{\sqrt{1-a^2}}$, respectively, one has

$$K(a, x) = k(a, v) := k_{\mathrm{LLe}}(a, v) := \frac{1}{2\tilde{u}(a, v)^2} + g(v) - 2g(x(a, v)) \quad \forall (a, x) \in \mathrm{LLe}, \tag{11}$$

since $u > 1$ and $v > \sqrt{2}$ on R. For $(a, x) \in R$, let

$$(D_a k)(a, v) := 4\overline{\Phi}(\sqrt{2})\big(x(a, v) - a\big)^3 \frac{\partial k}{\partial a}(a, v); \tag{12}$$

$$(D_{a,a} k)(a, x) := \frac{125(1 - a^2)^2}{(x - a)^2\varphi(x)} \frac{\partial(D_a k)}{\partial a}\big(a, v(a, x)\big). \tag{13}$$

The coefficients of the partial derivatives in (12) and (13) are chosen in order to make $(D_{a,a}k)(a,x)$ equal to an algebraic expression (and even a polynomial) in a and x, and at that $\operatorname{sign}(D_{a,a}k)(a,x) = \operatorname{sign}\frac{\partial(D_a k)}{\partial a}\big(a, v(a,x)\big)$. With Mathematica 5.2, one can therefore use the command

```
Reduce[daakx<=0 && 0<a<1 && Sqrt[2]<x<Sqrt[3]]
```

(where `daakx` stands for $(D_{a,a}k)(a,x)$), which outputs `False`, meaning that $D_{a,a}k > 0$ on R. By (13), this implies $\frac{\partial(D_a k)}{\partial a}\big(a, v(a,x)\big) > 0$ for $(a,x) \in R$, so that $(D_a k)(a,v)$ is increasing in a for every fixed value of v; more exactly, $(D_a k)(a,v)$ is increasing in $a \in \big(a_1(v), a_2(v)\big)$ for every fixed value of $v \in (\sqrt{2}, \sqrt{3}]$, where a_1 and a_2 are certain functions, such that

$$\big(a, x(a,v)\big) \in \mathrm{LLe} \iff \Big(v \in (\sqrt{2}, \sqrt{3}] \ \&\ a \in \big(a_1(v), a_2(v)\big)\Big).$$

Thus, for every fixed value of $v \in (\sqrt{2}, \sqrt{3}]$ the sign pattern of $(D_a k)(a,v)$ in $a \in \big(a_1(v), a_2(v)\big)$ is $-$ or $+$ or $-+$; that is, $(D_a k)(a,v)$ may change sign only from $-$ to $+$ as a increases. By (12), $\frac{\partial k}{\partial a}(a,v)$ has the same sign pattern. Hence, for every fixed value of $v \in (\sqrt{2}, \sqrt{3}]$ one has $k(a,v) \searrow$ or $\nearrow$ or $\searrow\nearrow$ in $a \in \big(a_1(v), a_2(v)\big)$. Now Lemma 6 follows. $\square$

Proof of Lemma 7 (LG). This proof is almost identical to that of Lemma 6, except that the term $g(v)$ in (11) is replaced here by $h(v)$, and the interval $(\sqrt{2}, \sqrt{3}]$ is replaced by $(\sqrt{3}, 5\sqrt{2}\,)$. However, both terms $g(v)$ and $h(v)$ are constant for any fixed value of v. $\square$

Proof of Lemma 8 (GL_1). One has

$$(a,x) \in \mathrm{GL}_1 \iff \Big(\sqrt{2} < x \leqslant x_* \ \&\ 0 < a < a_1(x)\Big), \tag{14}$$

where $a_1(x) := \frac{x}{3} - \frac{1}{3}\sqrt{6 - 2x^2}$. Since $\sqrt{2} < u < v < \sqrt{3}$ on GL_1 (where again u and v are defined by (5) and (6)), one has

$$K(a,x) = k(a,x) := k_{\mathrm{GL}}(a,x) := g(u) + g(v) - 2g(x) \quad \forall (a,x) \in \mathrm{GL}_1\,. \tag{15}$$

For $(a,x) \in R$, let

$$(D_a k)(a,x) := \frac{\partial k}{\partial a}(a,x) \cdot \frac{125\,\overline{\Phi}(\sqrt{2}\,)\,\big(1-a^2\big)^{3/2}}{(1+ax)(251+v)\varphi(v)}; \tag{16}$$

$$(D_{a,a}k)(s,x) := \frac{\partial(D_a k)}{\partial a}(a,x) \cdot (1-a^2)^3 (1+ax)^2 (251+v)^2 \mathrm{e}^{-\frac{2ax}{1-a^2}}, \tag{17}$$

where $\sqrt{1-s^2}$ is substituted for a in the right-hand side of (17). Then $(D_{a,a}k)(s,x)$ is a polynomial in s and x. Using again the Mathematica command `Reduce`, namely

$$\texttt{Reduce[daak[s,x]>0 \&\& Sqrt[2]<x<Sqrt[3] \&\& 0<s<1]}, \tag{18}$$

where `daak[s,x]` stands for $(D_{a,a}k)(s,x)$, one sees that for every $(s,x) \in R$

$$(D_{a,a}k)(s,x) > 0 \iff \Big(\sqrt{2} < x < x_{**} \ \&\ s_{*,1}(x) < s < s_{*,2}(x)\Big), \tag{19}$$

where x_{**} is a certain number between $\sqrt{2}$ and $\sqrt{3}$, and $s_{*,1}$ and $s_{*,2}$ are certain functions. $\big($In fact, $x_{**} \approx 1.678696$ is a root of a certain polynomial of degree 32 and, for each $x \in (\sqrt{2}, x_{**})$, the values $s_{*,1}(x)$ and $s_{*,2}(x)$ are two of the roots s of the polynomial $(D_{a,a}k)(s,x)$.$\big)$ Next,

```
Reduce[daak[95/100,x]<=0 && Sqrt[2]<x<=xx]
```

produces `False`; here `xx` stands for x_* – recall the definition of GL_1 and (8); that is, $(D_{a,a}k)(\frac{95}{100},x) > 0$ $\forall x \in (\sqrt{2}, x_*]$. Hence and in view of (19),

$$s_{*,1}(x) < \tfrac{95}{100} < s_{*,2}(x) \quad \forall x \in (\sqrt{2}, x_*]. \tag{20}$$

On the other hand, setting

$$s_1(x) := \sqrt{1 - a_1(x)^2} \tag{21}$$

with a_1 as in (14), one has $s_1 > \frac{95}{100}$ on $(\sqrt{2}, x_*]$. Hence, by (20), $s_1 > s_{*,1}$ on $(\sqrt{2}, x_*]$. So, in view of (19), the sign pattern of $(D_{a,a}k)(s,x)$ in $s \in (s_1(x), 1)$ is $-$ or $+-$, depending on whether $s_1(x) \geqslant s_{*,2}(x)$ or not, for each $x \in (\sqrt{2}, x_*]$. Now (17) and (21) imply that the sign pattern of $\frac{\partial(D_a k)}{\partial a}(a,x)$ in $a \in (0, a_1(x))$ is $-$ or $-+$, so that $(D_a k)(a,x) \searrow$ or $\searrow\nearrow$ in $a \in (0, a_1(x))$, for each $x \in (\sqrt{2}, x_*]$. Also, $(D_a k)(0,x) = 0$ for all $x \in \mathbb{R}$. So, the sign pattern of $(D_a k)(a,x)$ in $a \in (0, a_1(x))$ is $-$ or $-+$, for each $x \in (\sqrt{2}, x_*]$; in view of (16), $\frac{\partial k}{\partial a}(a,x)$ has the same sign pattern. Thus, $k(a,x) \searrow$ or $\searrow\nearrow$ in $a \in (0, a_1(x))$, for each $x \in (\sqrt{2}, x_*]$. Recalling (14), we complete the proof of Lemma 8. □

Proof of Lemma 9 (GL_2). This proof is similar to that of Lemma 8. Here one has

$$(a,x) \in \mathrm{GL}_2 \iff \Big(x_* < x < \sqrt{3}\ \&\ 0 < a < a_2(x)\Big), \tag{22}$$

where $a_2(x) := -\frac{x}{4} + \frac{1}{4}\sqrt{12 - 3x^2}$. Relation (15) holds here, and we retain definitions (16) and (17). Definition (21) is replaced here by

$$s_2(x) := \sqrt{1 - a_2(x)^2}. \tag{23}$$

Letting

$$s_* := s_2(x_*) = \frac{1}{12}\sqrt{\frac{1728 + 384\sqrt{6}}{19}} \approx 0.98761,$$

one can see that

$$s_2(x) \geqslant s_* \quad \forall x \in (x_*, \sqrt{3}\,). \tag{24}$$

On the other hand, using instead of (18) the Mathematica command

```
Reduce[daak[s,x]>0 && xx<x<Sqrt[3] && ss<s<1, Quartics->True],
```

where `xx` stands for x_* and `ss` stands for s_*, one sees that

$$\Big((D_{a,a}k)(s,x) > 0\ \&\ x_* < x < \sqrt{3}\ \&\ s_* < s < 1\Big) \iff \Big(x_* < x < x_{***}\ \&\ s_* < s < s_{*,2}(x)\Big), \tag{25}$$

where $x_{***} \approx 1.678694$ is a root of a certain polynomial of degree 20 and $s_{*,2}(x)$ is the same root in s of the polynomial $(D_{a,a}k)(s,x)$ as $s_{*,2}(x)$ in (19). Hence, in view of (25) and (24), the sign pattern of $(D_{a,a}k)(s,x)$ in $s \in (s_2(x), 1)$ is $-$ or $+-$, for each $x \in (x_*, \sqrt{3}\,)$.

Now (17) and (23) imply that the sign pattern of $\frac{\partial(D_a k)}{\partial a}(a,x)$ in $a \in (0, a_2(x))$ is $-$ or $-+$, so that $(D_a k)(a,x) \searrow$ or $\searrow\nearrow$ in $a \in (0, a_2(x))$, for each $x \in (x_*, \sqrt{3}\,)$. Also, $(D_a k)(0,x) = 0$ for all $x \in \mathbb{R}$. So, the sign pattern of $(D_a k)(a,x)$ in $a \in (0, a_2(x))$ is $-$ or $-+$, for each $x \in (x_*, \sqrt{3}\,)$; in view of (16), $\frac{\partial k}{\partial a}(a,x)$ has the same sign pattern. Thus, $k(a,x) \searrow$ or $\searrow\nearrow$ in $a \in (0, a_2(x))$, for each $x \in (x_*, \sqrt{3}\,)$. Recalling (22), we complete the proof of Lemma 9. □

Proof of Lemma 10 (GG_1). Expressing x and v in view of (5) and (6) in terms of a and u as $\tilde{x}(a,u) := \sqrt{1-a^2}\,u + a$ and $\tilde{v}(a,u) := u + \frac{2a}{\sqrt{1-a^2}}$, respectively, one has

$$K(a,x) = k(a,u) := k_{\mathrm{GG}_1}(a,u) := g(u) + h(\tilde{v}(a,u)) - 2g(\tilde{x}(a,u)) \quad \forall (a,x) \in \mathrm{GG}_1; \tag{26}$$

note that

$$(a,x)\in \mathrm{GG}_1 \implies \big(a,u(a,x)\big)\in(0,\tfrac{1}{\sqrt3})\times(\sqrt2,\sqrt3\,). \tag{27}$$

For $(a,u)\in R$, let

$$\begin{aligned} d(a,u) &:= \frac{\partial k}{\partial a}(a,u)\cdot 125\,\overline{\Phi}(\sqrt2)\,(1-a^2)^{3/2}/\varphi(\tilde v(a,u)); \\ d_a(a,u) &:= \frac{\partial d}{\partial a}(a,u)\cdot(1-a^2)\,\frac{\varphi(\tilde v(a,u))}{\varphi(\tilde x(a,u))}; \\ d_u(a,u) &:= \frac{\partial d}{\partial u}(a,u)\cdot\frac{\varphi(\tilde v(a,u))}{\varphi(\tilde x(a,u))\sqrt{1-a^2}}. \end{aligned} \tag{28}$$

The idea of the proof of Lemma 10 is to show that $d(a,u)<0$ and hence $\frac{\partial k}{\partial a}(a,u)<0$ for all $(a,u)\in R_0 := [0,\frac{1}{\sqrt3}]\times[\sqrt2,\sqrt3]$. This is done by considering separately the interior and the four sides of the rectangle R_0, at that using the fact that $d_a(a,u)$ and $d_u(a,u)$ are polynomials in a, u, and $\sqrt{1-a^2}$. Employing the command

```
Reduce[da[a,u]==0 && du[a,u]==0 && 0<a<1/Sqrt[3] &&Sqrt[2]<u<Sqrt[3], {a,u}, Reals],
```

where `da[a,u]` and `du[a,u]` stand for $d_a(a,u)$ and $d_u(a,u)$, one sees that the system of equations $d_a(a,u)=0=d_u(a,u)$ has a unique solution $(a_*,u_*)\approx(0.11918,1.57770)$ in $(a,u)\in(0,\frac{1}{\sqrt3})\times(\sqrt2,\sqrt3\,)$, and $d(a_*,u_*)\approx -0.44<0$.

Let us consider next the values of d on the boundary of the rectangle $(0,\frac{1}{\sqrt3})\times(\sqrt2,\sqrt3\,)$.

First, $d(0,u)$ is an increasing affine function of u, and $d(0,\sqrt3\,)\approx-0.4<0$. Hence, $d(0,u)<0\ \forall u\in(\sqrt2,\sqrt3\,)$.

Second,

$$\frac{27}{2\,\mathrm{e}^{(5+4\sqrt2u+u^2)/6}}\,\frac{\partial d}{\partial u}(\tfrac{1}{\sqrt3},u) = -\sqrt2u^3-\Big(3+251\sqrt3\Big)\,u^2-\sqrt2\,\Big(3+251\sqrt3\Big)\,u+251\sqrt3+7$$

is decreasing in u and takes on value $-9-753\sqrt3<0$ at $u=\sqrt2$, so that $\frac{\partial d}{\partial u}(\frac{1}{\sqrt3},u)<0$ for $u\in(\sqrt2,\sqrt3\,)$ and $d(\frac{1}{\sqrt3},u)\searrow$ in $u\in(\sqrt2,\sqrt3\,)$. Moreover, $d(\frac{1}{\sqrt3},\sqrt2)<0$. Thus, $d(\frac{1}{\sqrt3},u)<0\ \forall u\in(\sqrt2,\sqrt3\,)$.

Third,

$$\big(1-a^2\big)\,\frac{\varphi(\tilde v(a,\sqrt2\,))}{\varphi(\tilde x(a,\sqrt2\,))}\,\frac{\partial d}{\partial a}(a,\sqrt2)=p_1(a)+\sqrt{1-a^2}\,p_2(a),$$

where $p_1(a)$ and $p_2(a)$ are certain polynomials in a. Therefore, the roots a of $\frac{\partial d}{\partial a}(a,\sqrt2)$ are among the roots of the polynomial $p_{1,2}(a):=p_1(a)^2-(1-a^2)p_2(a)^2$, which has exactly two roots $a\in(0,\frac{1}{\sqrt3})$. Of the latter roots, one is not a root of $\frac{\partial d}{\partial a}(a,\sqrt2)$. Also, $\frac{\partial d}{\partial a}(0,\sqrt2)=1>0$ and $\frac{\partial d}{\partial a}(\frac{1}{\sqrt3},\sqrt2)<0$. Hence, $\frac{\partial d}{\partial a}(a,\sqrt2)$ has exactly one root, $a_*\approx0.2224$, in $a\in(0,\frac{1}{\sqrt3})$ and, moreover, $d(a,\sqrt2)\nearrow$ in $a\in(0,a_*]$ and $d(a,\sqrt2)\searrow$ in $a\in[a_*,\frac{1}{\sqrt3})$. Besides, $d(a_*,\sqrt2)\approx-0.088<0$. Thus, $d(a,\sqrt2)<0\ \forall a\in(0,\frac{1}{\sqrt3})$.

Fourth (very similar to third),

$$\big(1-a^2\big)\,\frac{\varphi(\tilde v(a,\sqrt3\,))}{\varphi(\tilde x(a,\sqrt3\,))}\,\frac{\partial d}{\partial a}(a,\sqrt3\,)=p_1(a)+\sqrt{1-a^2}\,p_2(a),$$

where $p_1(a)$ and $p_2(a)$ are certain polynomials in a, different from the polynomials $p_1(a)$ and $p_2(a)$ in the previous paragraph. Therefore, the roots a of $\frac{\partial d}{\partial a}(a,\sqrt3)$ are among the roots of the polynomial $p_{1,2}(a):=p_1(a)^2-(1-a^2)p_2(a)^2$, which has exactly two roots $a\in(0,\frac{1}{\sqrt3})$. Of the latter roots, one is not a root of $\frac{\partial d}{\partial a}(a,\sqrt3\,)$. Also, $\frac{\partial d}{\partial a}(0,\sqrt3\,)=1>0$ and $\frac{\partial d}{\partial a}(\frac{1}{\sqrt3},\sqrt3\,)<0$. Hence, $\frac{\partial d}{\partial a}(a,\sqrt3\,)$ has exactly one root, $a_*\approx0.06651$, in $a\in(0,\frac{1}{\sqrt3})$

and, moreover, $d(a,\sqrt{3}\,)\nearrow$ in $a\in(0,a_*]$ and $d(a,\sqrt{3}\,)\searrow$ in $a\in[a_*,\frac{1}{\sqrt{3}})$. Besides, $d(a_*,\sqrt{3}\,)\approx-0.358<0$. Thus, $d(a,\sqrt{3}\,)<0\ \forall a\in(0,\frac{1}{\sqrt{3}})$.

We conclude that $d(a,u)<0\ \forall(a,u)\in[0,\frac{1}{\sqrt{3}}]\times[\sqrt{2},\sqrt{3}\,]$. By (28), the same holds for $\frac{\partial k}{\partial a}(a,u)$. It remains to recall (26) and (27). □

Proof of Lemma 11 (GG$_2$). In view of Lemma 2,

$$K(a,x)=(g\wedge h)(u(a,x))+h(v(a,x))-2g(x)\quad\forall(a,x)\in\mathrm{GG}_2\,. \tag{29}$$

One has $\frac{\partial u}{\partial a}>0$ on GG$_2$ and $\frac{\partial v}{\partial a}>0$ on R. Since $g=\frac{c_1}{250}\left(251\overline{\Phi}+\varphi\right)$ and $h=c_2\,\overline{\Phi}$ are decreasing on $[0,\infty)$, we conclude that $K(a,x)$ is decreasing in a on GG$_2$. □

Proof of Lemma 12 (GE). One has

$$(a,x)\in\mathrm{GE}\iff\Big(x_*<x<\sqrt{3}\ \&\ a=a_2(x):=\tfrac{1}{4}\sqrt{12-3x^2}-\tfrac{x}{4}\,\Big),$$

where, as before, x_* is defined by (8). Therefore, for all $(a,x)\in\mathrm{GE}$

$$K(a,x)=k(x):=k_{\mathrm{GE}}(x):=K(a_2(x),x)=g\big(u(a_2(x),x)\big)+g(\sqrt{3})-2g(x),$$

and it suffices to show that $k\leqslant0$ on $[x_*,\sqrt{3}]$. For $x\in[x_*,\sqrt{3}]$, let

$$k_1(x):=\frac{k'(x)\,\overline{\Phi}(\sqrt{2}\,)}{\varphi(x)(251+x)};$$
$$k_2(x):=k_1'(x)\cdot\frac{125\,(x+251)^2\,\left(4-x^2\right)^{3/2}\rho(x)^{13/2}}{16\sqrt{2}\,\varphi\big(u(a_2(x),x)\big)/\varphi(x)},$$

where $\rho(x):=x^2+x\sqrt{3}\sqrt{4-x^2}+2$. Then $k_2(x)=p_1(x)+\sqrt{\rho(x)}\,p_2(x)$, where $p_1(x)$ and $p_2(x)$ are some polynomials in x and $\sqrt{4-x^2}$. Hence, the roots of $k_2(x)$ are among the roots of

$$p_{1,2}(x):=p_1(x)^2-\rho(x)p_2(x)^2=p_{1,2,1}(x)+\sqrt{4-x^2}\,p_{1,2,2}(x),$$

where $p_{1,2,1}(x)$ and $p_{1,2,2}(x)$ are some polynomials in x. Hence, the roots of $k_2(x)$ are among the roots of

$$\frac{p_{1,2,1}(x)^2-(4-x^2)p_{1,2,2}(x)^2}{1024\,(x^2-1)^{14}},$$

which is a polynomial of degree 32 and has exactly one root in $[x_*,\sqrt{3}]$. Also, $k_2(x_*)\approx1.39\times10^7>0$ and $k_2(\sqrt{3}\,)\approx-2.07\times10^6<0$. Therefore, k_2 and hence k_1' have the sign pattern $+-$ on $[x_*,\sqrt{3}]$. Next, $k_1(x_*)\approx-4.8494<0$ and $k_1(\sqrt{3}\,)=0$, so that k_1 and hence k' have the sign pattern $-+$ on $[x_*,\sqrt{3}]$. It follows that k does not have a local maximum on $(x_*,\sqrt{3})$. At that, $k(x_*)\approx-3.0133\times10^{-6}<0$ and $k(\sqrt{3}\,)=0$. Thus, $k\leqslant0$ on $[x_*,\sqrt{3}]$. □

Proof of Lemma 13 (ELe). This proof is very similar to that of Lemma 12 (GE). One has

$$(a,x)\in\mathrm{ELe}\iff\Big(\sqrt{2}<x\leqslant x_*\ \&\ a=a_1(x):=\tfrac{x}{3}-\tfrac{1}{3}\sqrt{6-2x^2}\,\Big), \tag{30}$$

where, as before, x_* is defined by (8). Therefore, for all $(a,x)\in\mathrm{LLe}$

$$K(a,x)=k(x):=k_{\mathrm{ELe}}(x):=K(a_1(x),x)=g(\sqrt{2})+g\big(v(a_1(x),x)\big)-2g(x),$$

and it suffices to show that $k \leqslant 0$ on $[\sqrt{2}, x_*]$. For $x \in [\sqrt{2}, x_*]$, let

$$k_1(x) := \frac{500\,\overline{\Phi}(\sqrt{2})\,k'(x)}{\varphi(x)(251+x)};$$
$$k_2(x) := k_1'(x) \cdot \frac{\sqrt{2}(x+251)^2\,\big(3-x^2\big)^{3/2}\,\rho(x)^{13/2}}{9\varphi\big(v(a_1(x),x)\big)/\varphi(x)},$$

where $\rho(x) := x^2 + 2x\sqrt{2}\sqrt{3-x^2} + 3$. Then $k_2(x) = p_1(x) + \sqrt{\rho(x)}\,p_2(x)$, where $p_1(x)$ and $p_2(x)$ are some polynomials in x and $\sqrt{3-x^2}$. Hence, the roots of $k_2(x)$ are among the roots of

$$p_{1,2}(x) := p_1(x)^2 - \rho(x)\,p_2(x)^2 = p_{1,2,1}(x) + \sqrt{3-x^2}\,p_{1,2,2}(x),$$

where $p_{1,2,1}(x)$ and $p_{1,2,2}(x)$ are some polynomials in x. Hence, the roots of $k_2(x)$ are among the roots of

$$\frac{p_{1,2,1}(x)^2 - (3-x^2)p_{1,2,2}(x)^2}{125524238436\,(x^2-1)^{14}},$$

which is a polynomial of degree 32 and has exactly one root in $[\sqrt{2}, x_*]$. Also, $k_2(\sqrt{2}) \approx -6.32 \times 10^7 < 0$ and $k_2(x_*) \approx 1.06 \times 10^8 > 0$. Therefore, k_2 and hence k_1' have the sign pattern $-+$ on $[\sqrt{2}, x_*]$. Next, $k_1(\sqrt{2}) = 0$ and $k_1(x_*) \approx 0.000426 > 0$, so that k_1 and hence k' have the sign pattern $-+$ on $[\sqrt{2}, x_*]$. It follows that k does not have a local maximum on $(\sqrt{2}, x_*)$. At that, $k(\sqrt{2}) = 0$ and $k(x_*) \approx -3.0133 \times 10^{-6} < 0$. Thus, $k \leqslant 0$ on $[\sqrt{2}, x_*]$. □

Proof of Lemma 14 (EG_1). This proof is similar to that of Lemma 13. One has

$$(a,x) \in \mathrm{EG}_1 \iff \Big(x_* < x < \sqrt{3}\ \&\ a = a_1(x) := \tfrac{x}{3} - \tfrac{1}{3}\sqrt{6-2x^2}\,\Big). \tag{31}$$

Therefore, for all $(a,x) \in \mathrm{EG}_1$

$$K(a,x) = k(x) := k_{\mathrm{EG}_1}(x) := K(a_1(x),x) = g(\sqrt{2}\,) + h\big(v(a_1(x),x)\big) - 2g(x),$$

and it suffices to show that $k \leqslant 0$ on $[x_*, \sqrt{3}]$. For $x \in [x_*, \sqrt{3}]$, let

$$k_1(x) := \frac{500\,\overline{\Phi}(\sqrt{2})\,k'(x)}{\varphi(x)(251+x)};$$
$$k_2(x) := k_1'(x) \cdot \frac{\sqrt{2}\,(x+251)^2\,\big(3-x^2\big)^{3/2}\,\rho(x)^{9/2}}{9\,r(\sqrt{3}\,)\varphi\big(v(a_1(x),x)\big)/\varphi(x)},$$

where $\rho(x) := x^2 + 2x\sqrt{2}\sqrt{3-x^2} + 3$.

Then $k_2(x) = p_1(x) + \sqrt{3-x^2}\,p_2(x)$, where $p_1(x)$ and $p_2(x)$ are some polynomials in x. Hence, the roots of $k_2(x)$ are among the roots of

$$p_{1,2}(x) := \frac{p_1(x)^2 - (3-x^2)p_2(x)^2}{4374\,\big(1 + 251/r(\sqrt{3}\,)\big)^2\,(x^2-1)^4},$$

which is a polynomial of degree 14 and has exactly one root in $[x_*, \sqrt{3}]$, $x_\# \approx 1.6012$. Also, $k_2(x_*) \approx 1.1722 \times 10^6 > 0$ and $k_2(\sqrt{3}) \approx -3.8778 \times 10^7 < 0$. Therefore, k_2 and hence k_1' have the sign pattern $+-$ on $[x_*, \sqrt{3}]$, so that $\max_{[x_*,\sqrt{3}]} k_1 = k_1(x_\#) \approx -0.00034907 < 0$. It follows that $k' < 0$ and hence $k \searrow$ on $[x_*, \sqrt{3}]$. At that, $k(x_*) \approx -3.0133 \times 10^{-6} < 0$. Thus, $k \leqslant 0$ on $[x_*, \sqrt{3}]$. □

Proof of Lemma 15 (EG_2). One has

$$(a,x)\in \mathrm{EG}_2 \iff \Big(\sqrt{2}<x<\sqrt{3}\ \&\ a=a_2(x):=\tfrac{x}{3}+\tfrac{1}{3}\sqrt{6-2x^2}\,\Big). \tag{32}$$

Therefore, for all $(a,x)\in \mathrm{EG}_2$

$$K(a,x)=k(x):=k_{\mathrm{EG}_2}(x):=K(a_2(x),x)=g(\sqrt{2})+h\big(v(a_2(x),x)\big)-2g(x),$$

and it suffices to show that $k\leqslant 0$ on $(\sqrt{2},\sqrt{3})$.

For the functions a_1 (defined in (30) and (31)) and a_2 (defined in (32)), and for $x\in(\sqrt{2},\sqrt{3})$, one has $a_2(x)\geqslant a_1(x)$; also, $h(z)$ is decreasing in z and $v(a,x)$ is increasing in a. Hence, $h\big(v(a_2(x),x)\big)\leqslant h\big(v(a_1(x),x)\big)$, so that $k_{\mathrm{EG}_2}\leqslant k_{\mathrm{EG}_1}$ on $[x_*,\sqrt{3})$.

Similarly, in view of Lemma 2 one has $h\big(v(a_2(x),x)\big)\leqslant g\big(v(a_2(x),x)\big)\leqslant g\big(v(a_1(x),x)\big)$ $\forall x\in(\sqrt{2},x_*]$, so that $k_{\mathrm{EG}_2}\leqslant k_{\mathrm{ELe}}$ on $(\sqrt{2},x_*]$.

Now Lemma 15 follows, because it was shown in the proofs of Lemmas 13 and 14, respectively, that $k_{\mathrm{ELe}}\leqslant 0$ on $[\sqrt{2},x_*]$ and $k_{\mathrm{EG}_1}\leqslant 0$ on $[x_*,\sqrt{3}\,]$. □

Proof of Lemma 16 (A_1). This is trivial. □

Proof of Lemma 17 (A_2). This is also trivial, in view of (10). □

Proof of Lemma 18 ($\mathrm{X}_{1,1}$). On $\mathrm{X}_{1,1}$, one has $u<\sqrt{2}\leqslant v$. Also, by Lemma 2, $g\geqslant h$ on $[\sqrt{3},\infty)$. Therefore, for all $(a,x)\in \mathrm{X}_{1,1}$

$$K(a,x)\leqslant k(a):=k_{\mathrm{X}_{1,1}}(a):=\frac{1}{2u(a,\sqrt{2}\,)^2}+g\big(v(a,\sqrt{2}\,)\big)-2g(\sqrt{2}\,), \tag{33}$$

and it suffices to show that $k\leqslant 0$ on $[0,\frac{1}{\sqrt{2}})$. For $a\in[0,\frac{1}{\sqrt{2}})$, let

$$k_1(a):=\frac{2000\,\overline{\Phi}(\sqrt{2})\,k'(a)}{\lambda(a)},\quad \lambda(a):=\frac{1-a\sqrt{2}}{\big(\sqrt{2}-a\big)^3}>0;$$
$$k_2(a):=k_1'(a)\cdot(\sqrt{2}-a)^4\left(1-a^2\right)^4\,\lambda(a)^2/\Big(\sqrt{2}\,\varphi\big(v(a,\sqrt{2}\,)\big)\Big).$$

Then $k_2(a)=\sqrt{1-a^2}\,p_1(a)+p_2(a)$, where $p_1(a)$ and $p_2(a)$ are some polynomials in a. Hence, the roots of $k_2(a)$ are among the roots of

$$p_{1,2}(a):=(1-a^2)p_1(a)^2-p_2(a)^2,$$

which is a polynomial of degree 12 and has exactly two roots in $[0,\frac{1}{\sqrt{2}})$. Of those two roots, one is not a root of $k_2(a)$, so that $k_2(a)$ has at most one root in $[0,\frac{1}{\sqrt{2}})$. Also, $k_2(0)=251\sqrt{2}>0$ and $k_2(\frac{1}{\sqrt{2}})=-127<0$. Therefore, k_2 and hence k_1' have the sign pattern $+-$ on $[0,\frac{1}{\sqrt{2}}]$, so that $k_1\nearrow\searrow$ on $[0,\frac{1}{\sqrt{2}}]$. At that the values of k_1 at points 0, $\frac{6}{10}$, and $\frac{7}{10}$ are approximately $-52<0$, $48>0$, and $-344<0$, respectively. Therefore, k_1 and hence k' have the sign pattern $-+-$ on $[0,\frac{1}{\sqrt{2}})$ and, moreover, the only local maximum of k on $(0,\frac{1}{\sqrt{2}}]$ occurs only between $\frac{6}{10}$ and $\frac{7}{10}$; in fact, it occurs at $a\approx 0.67433$ and equals $\approx -0.00013578<0$. It remains to note that $k(0)\approx -0.0028660<0$. □

Proof of Lemma 19 ($\mathrm{X}_{1,2}$). This proof is similar to that of Lemma 11. In place of (29), here one still has relation (33) for all $(a,x)\in \mathrm{X}_{1,2}$, since $u<\sqrt{2}\leqslant v$ on $\mathrm{X}_{1,2}$ as well. Since $u(a,\sqrt{2}\,)\nearrow$ and $v(a,\sqrt{2}\,)\nearrow$ in $a\in[\frac{1}{\sqrt{2}},\frac{2\sqrt{2}}{3})$, one has $k\searrow$ on $[\frac{1}{\sqrt{2}},\frac{2\sqrt{2}}{3})$, so that the maximum of k on $[\frac{1}{\sqrt{2}},\frac{2\sqrt{2}}{3})$ equals $k(\frac{1}{\sqrt{2}})$, which is negative, in view of Lemma 18 and the continuity of k. □

Proof of Lemma 20 ($X_{1,3}$). This proof is similar to that of Lemma 19. In place of (33), here one has

$$K(a,x) \leqslant k(a) := k_{X_{1,3}}(a) := g\big(u(a,\sqrt{2}\,)\big) + g\big(v(a,\sqrt{2}\,)\big) - 2g(\sqrt{2}\,),$$

for all $(a,x) \in X_{1,3}$, since $u \geqslant \sqrt{2}$ and $v \geqslant \sqrt{2}$ on $X_{1,3}$. Since $k \searrow$ in $a \in [\frac{2\sqrt{2}}{3}, 1)$, the maximum of k on $[\frac{2\sqrt{2}}{3}, 1)$ equals $k(\frac{2\sqrt{2}}{3}) \approx -0.25287 < 0$. □

Proof of Lemma 21 (X_2). This follows immediately from Lemma 5. □

References

[1] V. Bentkus, A remark on the inequalities of Bernstein, Prokhorov, Bennett, Hoeffding, and Talagrand. *Lithuanian Math. J.* **42** (2002) 262–269.

[2] V. Bentkus, An inequality for tail probabilities of martingales with differences bounded from one side. *J. Theoret. Probab.* **16** (2003) 161–173

[3] V. Bentkus, On Hoeffding's inequalities. *Ann. Probab.* **32** (2004) 1650–1673.

[4] S.G. Bobkov, F. Götze, C. Houdré, On Gaussian and Bernoulli covariance representations. *Bernoulli* **7** (2002) 439–451.

[5] G.E. Collins, Quantifier elimination for the elementary theory of real closed fields by cylindrical algebraic decomposition. *Lect. Notes Comput. Sci.* **33** (1975) 134–183.

[6] M.L. Eaton, A probability inequality for linear combinations of bounded random variables. *Ann. Statist.* **2** (1974) 609–614.

[7] D. Edelman, An inequality of optimal order for the tail probabilities of the T statistic under symmetry. *J. Amer. Statist. Assoc.* **85** (1990) 120–122.

[8] B. Efron, Student's t test under symmetry conditions. *J. Amer. Statist. Assoc.* **64** (1969) 1278–1302.

[9] S.E. Graversen, G. Peškir, Extremal problems in the maximal inequalities of Khintchine. *Math. Proc. Cambridge Philos. Soc.* **123** (1998) 169–177.

[10] S. Łojasiewicz, Sur les ensembles semi-analytiques. *Actes du Congrès International des Mathématiciens (Nice, 1970).* Tome 2, Gauthier-Villars, Paris (1970) 237–241.

[11] I. Pinelis, Extremal probabilistic problems and Hotelling's T^2 test under a symmetry condition. *Ann. Statist.* **22** (1994) 357–368.

[12] I. Pinelis, Optimal tail comparison based on comparison of moments. *High dimensional probability (Oberwolfach, 1996).* Birkhäuser, Basel *Progr. Probab.* . **43** (1998) 297–314.

[13] I. Pinelis, Fractional sums and integrals of r-concave tails and applications to comparison probability inequalities *Advances in stochastic inequalities (Atlanta, GA, 1997).* Amer. Math. Soc., Providence, RI. **234** *Contemp. Math.,* . (1999) 149–168.

[14] I. Pinelis, On exact maximal Khinchine inequalities. *High dimensional probability, II (Seattle, WA, 1999).* Birkhäuser Boston, Boston, MA *Progr. Probab..* **47** (2000) 49–63.

[15] I. Pinelis, Birkhäuser, Basel L'Hospital type rules for monotonicity: applications to probability inequalities for sums of bounded random variables. *J. Inequal. Pure Appl. Math.* **3** (2002) Article 7, 9 pp. (electronic).

[16] I. Pinelis, Binomial upper bounds on generalized moments and tail probabilities of (super)martingales with differences bounded from above. IMS Lecture Notes Monograph Series **51** (2006) 33-52.

[17] I. Pinelis, On normal domination of (super)martingales. Electronic Journal of Probality **11** (2006) 1049-1070.

[18] I. Pinelis, On l'Hospital-type rules for monotonicity. *J. Inequal. Pure Appl. Math.* **7** (2006) art40 (electronic).

[19] I. Pinelis, Exact inequalities for sums of asymmetric random variables, with applications. Probability Theory and Related Fields (2007) DOI 10.1007/s00440-007-0055-4.

[20] I. Pinelis, On inequalities for sums of bounded random variables. Preprint (2006) `http://arxiv.org/abs/math.PR/0603030`.

[21] A.A. Tarski, *A Decision Method for Elementary Algebra and Geometry.* RAND Corporation, Santa Monica, Calif. (1948).