On a multidimensional spherically invariant extension of the Rademacher-Gaussian comparison

Iosif Pinelis
Michigan Technological University

Follow this and additional works at: https://digitalcommons.mtu.edu/math-fp
Part of the Mathematics Commons

Recommended Citation
Retrieved from: https://digitalcommons.mtu.edu/math-fp/4

Follow this and additional works at: https://digitalcommons.mtu.edu/math-fp
Part of the Mathematics Commons
On a multidimensional spherically invariant extension of the Rademacher–Gaussian comparison

Iosif Pinelis*

Abstract

It is shown that
\[ P(\|a_1U_1 + \cdots + a_nU_n\| > u) \leq c P(\|Z_d\| > u) \]
for all real \(a\), where \(U_1, \ldots, U_n\) are independent random vectors uniformly distributed on the unit sphere in \(\mathbb{R}^d\), \(a_1, \ldots, a_n\) are any real numbers, \(a := \sqrt{(a_1^2 + \cdots + a_n^2)/d}\).

Theorem 2.3 in [9] states that this constant \(C\) is about \(89\) times as small as the corresponding bound recently obtained in [8].

To provide a relevant context, let us begin by introducing the class \(C^2_{\text{conv}}\) of all even twice differentiable functions \(h: \mathbb{R} \rightarrow \mathbb{R}\) whose second derivative \(h''\) is convex. Let \(\varepsilon, \varepsilon_1, \ldots, \varepsilon_n\) be independent Rademacher random variables (r.v.’s), and let \(\xi_1, \ldots, \xi_n\) be any independent symmetric r.v.’s with \(E\xi_i^2 = 1\) for all \(i\).

Take any natural \(d\). For any vectors \(x\) and \(y\) in \(\mathbb{R}^d\), let, as usual, \(x \cdot y\) denote the standard inner product of \(x\) and \(y\), and then let \(\|x\| := \sqrt{x \cdot x}\).

Theorem 2.3 in [9] states that \(Eh(\sqrt{\varepsilon A\xi^T}) \leq Eh(\sqrt{\xi A\xi^T})\) for any \(h \in C^2_{\text{conv}}\) and any nonnegative definite \(n \times n\) matrix \(A \in \mathbb{R}^{n \times n}\), where \(\varepsilon := [\varepsilon, \ldots, \varepsilon_n]\) and \(\xi := [\xi_1, \ldots, \xi_n]\). This can be restated as the following generalized moment comparison:
\[ Eh(\|\varepsilon x_1 + \cdots + \varepsilon_n x_n\|) \leq Eh(\|\xi_1 x_1 + \cdots + \xi_n x_n\|) \]  
(1)

for any \(h \in C^2_{\text{conv}}\) and any (nonrandom) vectors \(x_1, \ldots, x_n\) in \(\mathbb{R}^d\); indeed, any nonnegative definite matrix \(A \in \mathbb{R}^{n \times n}\) is the Gram matrix of some vectors \(x_1, \ldots, x_n\) in \(\mathbb{R}^d\) for some natural \(d\), and then \(\|a_1 x_1 + \cdots + a_n x_n\| = \sqrt{\alpha A\alpha^T}\) for any \(\alpha := [\alpha_1, \ldots, \alpha_n] \in \mathbb{R}^{1 \times n}\). From the comparison (1) of generalized moments of the r.v.’s \(\|\varepsilon x_1 + \cdots + \varepsilon_n x_n\|\) and

---

*Michigan Technological University, United States of America. E-mail: ipinellis@mtu.edu
where $f_U$ (see e.g. [7]). A special case of (2) is the inequality
\[ P(\|\varepsilon_1 x_1 + \cdots + \varepsilon_n x_n\| > u) < c P(\|Z_r\| > u) \] (2)
for all real $u$, where $x_1, \ldots, x_n$ are any (nonrandom) vectors in $\mathbb{R}^d$ whose Gram matrix is an orthoprojector of rank $r$, $Z_r$ is a standard normal random vector in $\mathbb{R}^r$, and
\[ c = c_3 := 2e^3/9 = 4.46 \ldots \] (3)
A special case of (2) is the inequality
\[ P(\varepsilon_1 a_1 + \cdots + \varepsilon_n a_n) > u) \leq c P(\|Z_1\| > u) \] (4)
for all real $u$, where $a_1, \ldots, a_n$ are any real numbers such that
\[ a_1^2 + \cdots + a_n^2 = 1. \]
The quoted results generalize and refine results of [4, 5]. In turn, they were further developed in [10, 11].

A simple inductive argument, which was direct rather than based on a generalized moment comparison, was offered in [3], where (4) was proved with $c \approx 12.01$. Based in part on that inductive argument in [3], the constant $c$ in (4) was improved to $\approx 1.01c_n$ in [13] and then to $c_n$ in [2], where $c_n := P(|\varepsilon_1 + \varepsilon_2| \geq 2)/P(\|Z_1\| \geq \sqrt{2}) = 3.17 \ldots$, so that $c_n$ is the best possible value of $c$ in (4).

In [1], another kind of multidimensional generalized moment comparison was obtained. A continuous function $f: \mathbb{R}^d \to \mathbb{R}$ is called bisubharmonic if the (Sobolev–Schwartz) distribution $\Delta^2 f$ is a nonnegative Radon measure on $\mathbb{R}^d$, where $\Delta$ is the Laplace operator on $\mathbb{R}^d$. By [1, Theorem 3], for any continuous function $f: \mathbb{R}^d \to \mathbb{R}$ one has
\[ f \text{ is bisubharmonic if and only if } \mathbb{E} f(y + Ut) \text{ is convex in } t \in (0, \infty) \text{ for each } y \in \mathbb{R}^d. \] (5)
where $U$ is a random vector uniformly distributed on the unit sphere $S^{d-1}$ in $\mathbb{R}^d$.

Let $U_1, \ldots, U_n$ be independent copies of $U$. Theorem 1 in [1] states that
\[ \mathbb{E} f(a_1 U_1 + \cdots + a_n U_n) \leq \mathbb{E} f(b_1 U_1 + \cdots + b_n U_n), \] (6)
where $f$ is a bisubharmonic function and $a_1, \ldots, a_n, b_1, \ldots, b_n$ are real numbers such that the $n$-tuple $(b_1^2, \ldots, b_n^2)$ is majorized by $(a_1^2, \ldots, a_n^2)$ in the sense of the Schur majorization (see e.g. [7]).

One may note that, whereas in (1) each of the random summands $\varepsilon_1 x_1, \ldots, \varepsilon_n x_n$, $\xi_1 x_1, \ldots, \xi_n x_n$ is distributed on a straight line through the origin, each of the random summands $a_1 U_1, \ldots, a_n U_n, b_1 U_1, \ldots, b_n U_n$ in (6) is uniformly distributed on a sphere centered at the origin.

Since the distributions of the random vectors $a_1 U_1 + \cdots + a_n U_n$ and $b_1 U_1 + \cdots + b_n U_n$ are clearly spherically invariant, without loss of generality one may assume that the function $f$ in (6) is spherically invariant as well, that is, $f(x)$ depends on $x \in \mathbb{R}^d$ only through $\|x\|$. If $f$ is indeed a spherically invariant bisubharmonic function, it then follows from (6) and [1, formulas (1.2), (1.3)] that
\[ \mathbb{E} f(a_1 U_1 + \cdots + a_n U_n) \leq \mathbb{E} f(a Z_d), \] (7)
where
\[ a := \sqrt{(a_1^2 + \cdots + a_n^2)/d}; \] (8)
\[ cf. [1, Corollary 1]. \]
Let $C^2_{\text{conv}}(H)$ denote the class of all spherically invariant twice differentiable functions $f$ from a Hilbert space $H$ to $\mathbb{R}$ whose second derivative $f''$ is convex in the sense that the function $H \ni x \mapsto f''(x; y, y)$ is convex for each $y \in H$, where $f''(x; y, y)$ is the value of the second derivative of the function $\mathbb{R} \ni t \mapsto f(x + ty)$ at $t = 0$. The class $C^2_{\text{conv}}(H)$ was characterized in [12], with some applications. Clearly, $C^2_{\text{conv}}(\mathbb{R})$ coincides with the class $C^2_{\text{conv}}$ defined in the beginning of this note.

K. Oleszkiewicz conjectured [8] that

$$P(\|a_1 U_1 + \cdots + a_n U_n\| > u) \leq c P(\|Z_d\| > u) \quad (9)$$

for some universal constant $c$ and all real $u$, where $a_1, \ldots, a_n, U_1, \ldots, U_n, Z_d$ are as before; clearly, (9) is a generalization of (4). This conjecture was proved in [8] with $c = 397$ based, in part, on the idea from [3].

Using inequality (2.6) in [9], one can improve the lower bound $1/397$ in [8, Lemma 1] to $1/e^2$ and thus improve the constant $c$ in (9) from 397 to $c^2 = 7.38 \ldots$. Indeed, let, as usual, $\Phi$ denote the standard normal distribution function. Then, by inequality (2.6) in [9], $g(d) := P(\|Z_d\| \geq \sqrt{d+2}) > 1 - \Phi((\sqrt{d+2} - \sqrt{d-1})\sqrt{2}) =: q(d)$, which latter is clearly increasing in $d$, with $q(4) > 1/e^2$, whence $g(d) > 1/e^2$ for $d = 4, 5, \ldots$, whereas $g(2) = 1/e^2 < g(3)$. So, $P(\|Z_d\| \geq \sqrt{d+2}) = g(d) > 1/e^2$ for $d = 2, 3, \ldots$. Similarly, $P(\|Z_d\| \geq \sqrt{d}) > 1/e$ for $d = 2, 3, \ldots$ (but a lower bound on $P(\|Z_d\| > \sqrt{d})$ is not really needed in the proof of the main result in [8]).

The aim of this note is to point out that, based on the generalized moment comparison (7) and results in [9, 10], one can further improve the constant $c$ in (9):

**Theorem 1.** Inequality (9) holds (for all real $u$) with $c$ as in (3). The strict version of (9), again with $c$ as in (3), also holds.

Our method is quite different from that of [8]. In view of (7), Theorem 1 is an immediate corollary of the following two lemmas.

**Lemma 1.** For any function $h \in C^2_{\text{conv}}$, the function $f : \mathbb{R}^d \to \mathbb{R}$ defined by the formula $f(x) := h(\|x\|)$ for $x \in \mathbb{R}^d$ is a spherically invariant bisubharmonic function.

**Lemma 2.** Let $\xi$ be any nonnegative r.v. such that

$$E h(\xi) \leq E h(\|Z_d\|) \quad \text{for all } h \in C^2_{\text{conv}}. \quad (10)$$

Then

$$P(\xi > u) < c_3 P(\|Z_d\| > u) \quad (11)$$

for all real $u$, with $c_3$ defined in (3).

**Proof of Lemma 1.** Let $U$ be as in (5) and then let $\varepsilon$ be a Rademacher r.v. independent of $U$. For all $t \in (0, \infty)$ and $y \in \mathbb{R}^d$

$$E f(y + U\sqrt{t}) = E f(y + \varepsilon U\sqrt{t}) = E h(\|y + \varepsilon U\sqrt{t}\|) = E E_U g_{y, h}(\beta_U + \varepsilon \sqrt{t}), \quad (12)$$

where $E_U$ denotes the conditional expectation given $U$, $g_{y,h}(t) := h(\sqrt{t^2 + b})$ for $b \in [0, \infty)$ and $u \in \mathbb{R}$, $\beta_U := y \cdot U$, and $h_U := \|y\|^2 - (y \cdot U)^2 \geq 0$, so that the r.v. $\varepsilon$ is independent of the pair $(\beta_U, h_U)$, which latter is a function of $U$. By [9, Lemma 3.1], $g_{y,h} \in C^2_{\text{conv}}$, for each $b \in [0, \infty)$. Hence, by [14, Lemma 3.1] or [9, Proposition A.1], $E_U g_{y,h}(\beta_U + \varepsilon \sqrt{t})$ is convex in $t \in (0, \infty)$. So, in view of (12), $E f(y + U\sqrt{t})$ is convex in $t \in (0, \infty)$. Now it follows by (5) that the function $f$ is indeed bisubharmonic. That $f$ is spherically invariant is trivial.

**Proof of Lemma 2.** Taken almost verbatim, the proof of Theorem 2.4 in [9] (based on Theorem 2.3 in [9]) can also serve as a proof of Lemma 2. Indeed, no properties of the
Spherically invariant extension of the Rademacher–Gaussian comparison

r.v. $\epsilon \Pi e^T$ were used in the proof of [9, Theorem 2.4] except that this nonnegative r.v. satisfies the inequality in [9, Theorem 2.3] with $\epsilon = 1$ and $\xi = Z_n$, which can then be written as (10) with $\xi = \sqrt{\epsilon \Pi e^T}$ and $d$ equal the rank of $\Pi$. (Note here a typo in [9]: in place of “Theorem 2.3” in line 7- on page 36 there, it should be “Theorem 2.4”.)

Instead of following the entire proof of [9, Theorem 2.4], one can alternatively reason as follows. Let $\xi$ be any nonnegative r.v. such that (10) holds. Then [9, Lemma 3.5] holds with $\xi^2$ in place of $\epsilon \Pi e^T$. So, in view of [9, formula (3.11)] and [10, formula (22) in Theorem 3.11], inequality (11) holds for $u \geq \mu_r$, with $r := d$ and $\mu$ defined on page 362 in [9]. The cases $r^{1/2} \leq u \leq \mu_r$ and $0 \leq u \leq r^{1/2}$ are considered as was done at the end of the proof of [9, Lemma 3.6], starting at the middle of page 365 in [9]. The case $u < 0$ is trivial.

An immediate application of Theorem 1 is

**Corollary 1.** Let $X_1, \ldots, X_n$ be any independent spherically invariant random vectors in $\mathbb{R}^d$, which are also independent of the Gaussian random vector $Z_d$. Then

$$P(\|X_1 + \cdots + X_n\| > u) < \frac{2e^3}{9} P \left( \sqrt{\|X_1\|^2 + \cdots + \|X_n\|^2} \|Z_d\| > u \right)$$  \hspace{2cm} (13)

for all real $u$.

This corollary follows from Theorem 1 by the conditioning on $\|X_1\|, \ldots, \|X_n\|$, because for each $i = 1, \ldots, n$ the conditional distribution of the spherically invariant random vector $X_i$ given $\|X_i\| = a_i$ is the distribution of $a_i U_i$.

In the case when the independent spherically invariant random vectors $X_1, \ldots, X_n$ are bounded almost surely by positive real numbers $b_1, \ldots, b_n$, respectively, one can obviously replace $\sqrt{\|X_1\|^2 + \cdots + \|X_n\|^2}$ in the bound in (13) by $\sqrt{b_1^2 + \cdots + b_n^2}$. The resulting bound, but with the constant factor $397$ in place of $\frac{2e^3}{9} = 4.46 \ldots$, was obtained in [8].

Similarly to the extension (13) of inequality (9), one can extend (7) as follows:

$$E f(X_1 + \cdots + X_n) \leq E f \left( \sqrt{\|X_1\|^2 + \cdots + \|X_n\|^2} Z_d \right)$$  \hspace{2cm} (14)

for any spherically invariant bisubharmonic function $f$, where $X_1, \ldots, X_n$ are as in Corollary 1.

A related result was obtained in [6]: if $X_1, \ldots, X_n$ are independent identically distributed spherically invariant random vectors in $\mathbb{R}^d$ such that $E h(\|X_1\|^2) \leq E h(\|Z_d\|^2)$ for all nonnegative convex functions $h: \mathbb{R} \to \mathbb{R}$, then

$$E \|a_1 X_1 + \cdots + a_n X_n\|^p \leq E \|a Z_d \sqrt{d}\|^p$$  \hspace{2cm} (15)

for real $p \geq 3$, where $a_1, \ldots, a_n, a$ are as in (7)-(8).

**References**


Spherically invariant extension of the Rademacher–Gaussian comparison


Advantages of publishing in EJP-ECP

• Very high standards
• Free for authors, free for readers
• Quick publication (no backlog)
• Secure publication (LOCKSS\(^1\))
• Easy interface (EJMS\(^2\))

Economical model of EJP-ECP

• Non profit, sponsored by IMS\(^3\), BS\(^4\), ProjectEuclid\(^5\)
• Purely electronic

Help keep the journal free and vigorous

• Donate to the IMS open access fund\(^6\) (click here to donate!)
• Submit your best articles to EJP-ECP
• Choose EJP-ECP over for-profit journals

\(^1\)LOCKSS: Lots of Copies Keep Stuff Safe [http://www.lockss.org/](http://www.lockss.org/)
\(^5\)Project Euclid: [https://projecteuclid.org/](https://projecteuclid.org/)
\(^6\)IMS Open Access Fund: [http://www.imstat.org/publications/open.htm](http://www.imstat.org/publications/open.htm)