An Inequality for Moments of Log-Concave Functions on Gaussian Random Vectors

Nikos Dafnis and Grigoris Paouris

Abstract We prove sharp moment inequalities for log-concave and log-convex functions, on Gaussian random vectors. As an application we take a reverse form of the classical logarithmic Sobolev inequality, in the case where the function is log-concave.

1 Introduction and Main Results

A function $f : \mathbb{R}^k \to [0, +\infty)$ is called *log-concave* (*on its support*), if and only if

$$f\big((1-\lambda)x + \lambda y\big) \geq f(x)^{(1-\lambda)} f(y)^{\lambda},$$

for every $\lambda \in [0, 1]$ and $x, y \in \mathrm{supp}(f)$. Respectively, f is called log-*convex* (*on its support*), if and only if

$$f\big((1-\lambda)x + \lambda y\big) \leq f(x)^{(1-\lambda)} f(y)^{\lambda},$$

for every $\lambda \in [0, 1]$ and $x, y \in \mathrm{supp}(f)$. The aim of this note is to present a sharp inequality for Gaussian moments of log-concave and log-convex functions, stated below as Theorem 1.1.

We work on $\mathbb{R}^k$, equipped with the standard scalar product $\langle \cdot, \cdot \rangle$. We denote by $|\cdot|$ the corresponding Euclidean norm and the absolute value of a real number. We use the notation $X \sim N(\xi, T)$, if X is a Gaussian random vector in $\mathbb{R}^k$, with expectation $\xi \in \mathbb{R}^k$ and covariance the $k \times k$ positive semi-definite matrix T. We say that X is a *standard Gaussian* random vector if it is centered (i.e. $\mathbb{E}X = 0$) with covariance matrix the identity in $\mathbb{R}^k$, where in that case γ_k stands for its distribution law. Finally,

N. Dafnis (✉)
Department of Mathematics, Technion - Israel Institute of Technology, Haifa 32000, Israel
e-mail: nikdafnis@gmail.com

G. Paouris
Department of Mathematics, Texas A&M University, College Station, TX 77843, USA
e-mail: grigorios.paouris@gmail.com

B. Klartag, E. Milman (eds.), *Geometric Aspects of Functional Analysis*,
Lecture Notes in Mathematics 2169, DOI 10.1007/978-3-319-45282-1_7

$\mathcal{L}^{p,s}(\gamma_k)$ stand for the class of all functions $f \in L^p(\gamma_k)$ whose partial derivatives up to order s, are also in $L^p(\gamma_k)$.

Theorem 1.1 *Let $k \in \mathbb{N}$ and X be a Gaussian random vector in $\mathbb{R}^k$. Let $f : \mathbb{R}^k \to [0, +\infty)$ be a* log-*concave and $g : \mathbb{R}^k \to [0, +\infty)$ be a* log-*convex function. Then,*

(i) for every $r \in [0, 1]$

$$\mathbb{E}f\big(\sqrt{r}X\big) \geq (\mathbb{E}f(X)^r)^{\frac{1}{r}} \quad \text{and} \quad \mathbb{E}g\big(\sqrt{r}X\big) \leq (\mathbb{E}g(X)^r)^{\frac{1}{r}}, \tag{1}$$

(ii) for every $q \in [1, +\infty)$

$$\mathbb{E}f\big(\sqrt{q}X\big) \leq (\mathbb{E}f(X)^q)^{\frac{1}{q}} \quad \text{and} \quad \mathbb{E}g\big(\sqrt{q}X\big) \geq (\mathbb{E}g(X)^q)^{\frac{1}{q}}. \tag{2}$$

In any case, equality holds if $r = 1 = q$ or if $f(x) = g(x) = e^{-\langle \mathrm{a},x\rangle + c}$, where $\mathrm{a} \in \mathbb{R}^k$ *and $c \in \mathbb{R}$.*

We prove Theorem 1.1 in Sect. 2, where we combine techniques from [7] along with Barthe's inequality [2].

The *entropy* of a function $f : \mathbb{R}^k \to \mathbb{R}$, with respect to a random vector X in $\mathbb{R}^k$, is defined to be

$$\mathrm{Ent}_X(f) := \mathbb{E}|f(X)| \log |f(X)| - \mathbb{E}|f(X)| \log \mathbb{E}|f(X)|,$$

provided all the expectations exist. Note that (for $f \geq 0$)

$$\mathrm{Ent}_X(f) = \frac{d}{dq}\Big[\big(\mathbb{E}f(X)^q\big)^{\frac{1}{q}}\Big]_{q=1}$$

and so, Theorem 1.1 implies the following entropy inequality:

Corollary 1.2 *Let $f : \mathbb{R}^k \to [0, +\infty)$ and X be a Gaussian random vector in $\mathbb{R}^k$.*

(i) If f is log-concave, then

$$\mathrm{Ent}_X(f) \geq \frac{1}{2}\mathbb{E}\langle X, \nabla f(X)\rangle. \tag{3}$$

(ii) If f is log-convex, then

$$\mathrm{Ent}_X(f) \leq \frac{1}{2}\mathbb{E}\langle X, \nabla f(X)\rangle. \tag{4}$$

In any case, equality holds if $f(x) = \exp\big(\langle \mathrm{a}, x\rangle + c\big)$, $\mathrm{a} \in \mathbb{R}^k$, $c \in \mathbb{R}$.

Proof Let $m(q) := \big(\mathbb{E}f(X)^q\big)^{\frac{1}{q}}$ and $h(q) := \mathbb{E}f(\sqrt{q}X)$. Then we have

$$m(1) = \mathbb{E}f(X) = h(1), \quad m'(1) = \mathrm{Ent}_X(f) \quad \text{and} \quad h'(1) = \frac{1}{2}\mathbb{E}\langle X, \nabla f(X)\rangle,$$

and Theorem 1.1 implies the desired result. □

The logarithmic Sobolev inequality, proved by Gross in [10], states that if $X \sim N(0, I_k)$, then

$$\mathrm{Ent}_X(f^2) \le 2\,\mathbb{E}|\nabla f(X)|^2, \tag{5}$$

for every function $f \in L^2(\gamma_k)$. Moreover, Carlen showed in [6], that equality holds if and only if f is an exponential function. For more details about the logarithmic Sobolev inequality we refer the reader to [4, 14, 19, 20] and to the references therein.

In Sect. 3, we show that Corollary 1.2, after an application of the Gaussian integration by parts formula (see Lemma 3.1), leads to the following reverse form of Gross' inequality, when the function is log concave:

Theorem 1.3 *Let X be a standard Gaussian random vector in $\mathbb{R}^k$ and $f = e^{-v} \in \mathcal{L}^{2,1}(\gamma_k)$, be a positive log-concave function (on its support). Then*

$$2\,\mathbb{E}|\nabla f(X)|^2 - \mathbb{E}f(X)^2 \Delta v(X) \le \mathrm{Ent}_X(f^2). \tag{6}$$

Theorem 1.3, ensures that if a log-concave function $f = e^{-v}$ is close to be an exponential, in the sense that $\mathbb{E}f(X)^2\Delta v(X)$ is small, then the logarithmic Sobolev inequality for f is close to be sharp.

For more properties and stability results on the logarithmic-Sobolev inequalities we refer to the papers [8, 9, 11] and the references therein.

2 Proof of the Main Result

The first ingredient of the proof of Theorem 1.1, is the following inequality for Gaussian random vectors, proved in [7]. We recall that for two square matrices A and B, we say that $A \le B$ if and only if $B - A$ is positive semi-definite.

Theorem 2.1 *Let $m, n_1, \dots, n_m \in \mathbb{N}$ and set $N = \sum_{i=1}^m n_i$. For every $i = 1, \dots, m$, let X_i be a Gaussian random vector in $\mathbb{R}^{n_i}$, such that $\mathbf{X} := (X_1, \dots, X_m)$ is a Gaussian random vector in $\mathbb{R}^N$ with covariance the $N \times N$ matrix $T = (T_{ij})_{1 \le i,j \le m}$, where T_{ij} is the covariance $n_i \times n_j$ matrix between X_i and X_j, $1 \le i, j \le m$. Let $p_1, \dots, p_m \in \mathbb{R}$ and consider the $N \times N$ block diagonal matrix $P = \mathrm{diag}(p_1 T_{11}, \dots, p_m T_{mm})$. Then, for any set of nonnegative measurable functions f_i on $\mathbb{R}^{n_i}$, $i = 1, \dots, m$,*

(*i*) *if $T \le P$, then*

$$\mathbb{E}\prod_{i=1}^m f_i(X_i) \le \prod_{i=1}^m \left(\mathbb{E}f_i(X_i)^{p_i}\right)^{\frac{1}{p_i}}, \tag{7}$$

(ii) if $T \geq P$, then

$$\mathbb{E}\prod_{i=1}^{m} f_i(X_i) \geq \prod_{i=1}^{m}\left(\mathbb{E}f_i(X_i)^{p_i}\right)^{\frac{1}{p_i}}. \tag{8}$$

Theorem 2.1 generalizes many fundamental results in analysis, such as Hölder inequality and its reverse, Young inequality with the best constant and its reverse [3] and [5], and Nelson's Gaussian Hypercontractivity and its reverse [17] and [15]. Actually, the first part of Theorem 2.1 is another formulation of the Brascamp-Lieb inequality [5, 13], while the second part provides a reverse form.

Moreover, (8) implies (see [7]) F. Barthe's reverse Brascamp-Lieb inequality [2], which the second main tool in our the proof of Theorem 1.1. For more extensions of Brascamp-Lieb inequality and similar results see [12] and [16].

For our purposes, we need the so-called *geometric* form (see [1]) of Barthe's theorem.

Theorem 2.2 *Let $n, m, n_1, \ldots, n_m \in \mathbb{N}$ with $n_i \leq n$ for every $i = 1, \ldots, m$. Let U_i be a $n_i \times n$ matrix with $U_i U_i^* = I_{n_i}$ for $i = 1, \ldots, m$ and $c_1, \ldots, c_m$ be positive real numbers such that*

$$\sum_{i=1}^{m} c_i\, U_i^* U_i = I_n.$$

Let $h : \mathbb{R}^n \to [0, +\infty)$ and $f_i : \mathbb{R}^{n_i} \to [0, +\infty)$, $i = 1, \ldots, m$, be measurable functions such that

$$h\left(\sum_{i=1}^{N} c_i U_i^* \xi_i\right) \geq \prod_{i=1}^{m} f_i(\xi_i)^{c_i} \quad \forall\, \xi_i \in \mathbb{R}^{n_i}, \tag{9}$$

$i = 1, \ldots, m$. Then

$$\int_{\mathbb{R}^n} h(x)\, d\gamma_n(x) \geq \prod_{i=1}^{m}\left(\int_{\mathbb{R}^{n_i}} f_i(x)\, d\gamma_{n_i}(x)\right)^{c_i}. \tag{10}$$

2.1 Decomposing the Identity

We will apply Theorem 2.1 in the special case where the covariance is the $kn \times kn$ matrix $T = \left([T_{ij}]\right)_{i,j\leq n}$, with $T_{ii} = I_k$ and $T_{ij} = tI_k$ if $i \neq j$, for some $t \in [-\frac{1}{n-1}, 1]$. Equivalently, in that case $\mathbf{X} := (X_1, \ldots, X_n) \sim N(0, T)$, where $X_1, \cdots, X_n$

are standard Gaussian random vectors in $\mathbb{R}^k$, such that

$$\mathbb{E}(X_i X_j^*) = \begin{cases} I_k \,, \; i = j \\ tI_k \,, \; i \neq j \end{cases} . \tag{11}$$

For any $t \in [0, 1]$, a natural way to construct such random vectors is to consider n independent copies $Z_1, \dots, Z_n$, of a $Z \sim N(0, I_k)$ and set

$$X_i := \sqrt{t} Z + \sqrt{1-t} Z_i \,, \qquad i = 1, \dots, n.$$

However, we are going to use a more geometric approach. First we will deal with the 1-dimensional case and then, by using a tensorization argument, we will pass to the general k-dimensional case, for any $k \in \mathbb{N}$. We begin with the definition of the SR-simplex.

Definition 2.3 We say that $S = \mathrm{conv}\{\mathrm{v}_1, \dots, \mathrm{v}_n\} \subseteq \mathbb{R}^{n-1}$ is the *spherico-regular simplex* (in short SR-simplex) in $\mathbb{R}^{n-1}$, if $\mathrm{v}_1, \dots, \mathrm{v}_n$ are unit vectors in $\mathbb{R}^{n-1}$ with the following two properties:

(SR1) $\langle \mathrm{v}_i, \mathrm{v}_j \rangle = -\frac{1}{n-1}$, for any $i \neq j$,
(SR2) $\sum_{i=1}^n \mathrm{v}_i = 0$.

Using the vertices of the SR-simplex in $\mathbb{R}^{n-1}$, we create n vectors in $\mathbb{R}^n$ with the same angle between them. This is done in the next lemma.

Lemma 2.4 *Let $n \geq 2$ and $\mathrm{v}_1, \dots, \mathrm{v}_n$ be the vertices of any RS-Simplex in $\mathbb{R}^{n-1}$. For every $t \in [-\frac{1}{n-1}\,, 1]$, let $u_1, \dots, u_n$ be the unit vectors in $\mathbb{R}^n$ with*

$$u_i = u_i(t) = \sqrt{\frac{t(n-1)+1}{n}} \; e_n \; + \; \sqrt{\frac{n-1}{n}(1-t)} \; \mathrm{v}_i \,, \tag{12}$$

$i = 1, \dots, n$. Then we have that

$$\langle u_i, u_j \rangle = t \,, \qquad \forall \, i \neq j. \tag{13}$$

Moreover,

(i) if $t \in [0, 1]$, then

$$\frac{1}{t(n-1)+1} \sum_{i=1}^n u_i u_i^* + \frac{nt}{t(n-1)+1} \sum_{j=1}^{n-1} e_j e_j^* = I_n, \tag{14}$$

(ii) if $t \in [-\frac{1}{n-1}, 0]$, then

$$\frac{1}{1-t} \sum_{i=1}^n u_i u_i^* + \frac{-nt}{1-t} e_n e_n^* = I_n. \tag{15}$$

Proof A direct computation, using the properties ($SR1$), ($SR2$) and the fact that

$$\frac{n-1}{n}\sum_{i=1}^{n} \mathrm{v}_i \mathrm{v}_i^* = I_{n-1},$$

shows that (13)–(15) holds true. □

Remark 2.5 If $Z \sim N(0, I_n)$, then $X_i := \langle u_i, Z\rangle$, $i = 1, \ldots, n$, are standard Gaussian random variables, satisfying the condition (11) in the 1-dimensional case.

For the general case we first recall the definition of the *tensor product* of two matrices:

Definition 2.6 For any matrices $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{k\times\ell}$, their tensor product is defined to be the $km \times \ell n$ matrix

$$A \otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1n}B \\ \vdots & \ddots & \vdots \\ a_{m1}B & \cdots & a_{mn}B \end{pmatrix}.$$

Every vector $a \in \mathbb{R}^n$ is considered to be a $n \times 1$ column matrix and with this notation, we state some basic properties for the tensor product, that we will use.

Lemma 2.7 *1. Let $a = (a_1, \ldots, a_m)^* \in \mathbb{R}^m$ and $b = (b_1, \ldots, b_n)^* \in \mathbb{R}^n$. Then*

$$a \otimes b^* = ab^* = \begin{pmatrix} a_1b_1 & \cdots & a_1b_n \\ \vdots & \ddots & \vdots \\ a_mb_1 & \cdots & a_mb_n \end{pmatrix} \in \mathbb{R}^{m\times n},$$

and as a linear transformation, $a \otimes b^* = ab^* : \mathbb{R}^n \to \mathbb{R}^m$ with

$$(a \otimes b^*)(x) = (ab^*)(x) = \langle x, b\rangle\, a, \quad x \in \mathbb{R}^n.$$

2. Let $A_i \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{k\times\ell}$. Then $\left(\sum_i A_i\right) \otimes B = \sum_i A_i \otimes B$.
3. Let $A_1 \in \mathbb{R}^{m\times n}$, $B_1 \in \mathbb{R}^{k\times\ell}$, and $A_2 \in \mathbb{R}^{n\times r}$, $B_2 \in \mathbb{R}^{\ell\times s}$. Then

$$(A_1 \otimes B_1)(A_2 \otimes B_2) = (A_1A_2) \otimes (B_1B_2) \in \mathbb{R}^{km\times rs}.$$

4. For any matrices A and B,

$$(A \otimes B)^* = A^* \otimes B^*.$$

For our k-dimensional construction, we consider the $k \times kn$ matrices

$$U_i := u_i^* \otimes I_k = \Big(\big[u_{i1} I_k\big] \cdots \big[u_{in} I_k\big] \Big), \tag{16}$$

$$E_j := e_j^* \otimes I_k = \Big(\big[e_{j1} I_k\big] \cdots \big[e_{jn} I_k\big] \Big), \tag{17}$$

for $i = 1 \ldots, n$. Note that

$$U_i^* U_i = (u_i^* \otimes I_k)^* (u_i^* \otimes I_k) = u_i u_i^* \otimes I_k$$

and

$$E_j^* E_j = (e_j^* \otimes I_k)^* (e_j^* \otimes I_k) = e_j e_j^* \otimes I_k,$$

for every $i, j = 1, \ldots, n$. Thus by taking the tensor product with I_k, in both sides of (14), we get that

$$\frac{1}{p} \sum_{i=1}^{n} U_i^* U_i + \frac{nt}{p} \sum_{j=1}^{n-1} E_j^* E_j = I_{kn}, \tag{18}$$

for every $t \in [0, 1]$, where $p := (n-1)t + 1$. Moreover, we can now construct the general case describing in (11). We summarize in the next lemma.

Lemma 2.8 *Suppose that $Z_1, \ldots, Z_n$ are iid standard Gaussian random vectors in $\mathbb{R}^k$ and set $\mathbf{Z} := (Z_1, \ldots, Z_n) \sim N(0, I_{kn})$. Consider the random vectors*

$$X_i := U_i \mathbf{Z} = \sum_{a=1}^{n} u_{ia} Z_a, \qquad i = 1, \ldots, n, \tag{19}$$

where U_i, $i = 1, \ldots, n$, are the matrices defined in (16). Then X_i is a standard Gaussian random vector in $\mathbb{R}^k$, for every $i = 1, \ldots, n$ and

$$\mathbb{E}\big[X_i \otimes X_j^*\big] = \Big(\mathbb{E}[X_{ir} X_{j\ell}]\Big)_{r,\ell \le k} = \Big(t\delta_{r\ell}\Big)_{r,\ell \le k} = tI_k, \tag{20}$$

for every $i \neq j$.

Proof Clearly, $\mathbb{E}X_i = 0$, for every $i, j = 1, \ldots, n$, and since

$$\mathbb{E}\big[Z_a \otimes Z_b^*\big] = \Big(\mathbb{E}[Z_{ar} Z_{b\ell}]\Big)_{r,\ell \le k} = \delta_{\alpha\beta} I_k$$

we have that

$$\begin{aligned}\mathbb{E}\left[X_{ir}X_{j\ell}\right] &= \mathbb{E}\left[\left(\sum_{a=1}^{n} u_{ia}Z_{ar}\right)\left(\sum_{b=1}^{n} u_{jb}Z_{b\ell}\right)\right] \\ &= \sum_{a=1}^{n}\sum_{b=1}^{n} u_{ia}u_{jb}\,\mathbb{E}\left[Z_{ar}Z_{b\ell}\right] \\ &= \sum_{a=1}^{n} u_{ia}u_{ja}\,\mathbb{E}\left[Z_{ar}Z_{a\ell}\right] \\ &= \sum_{a=1}^{n} u_{ia}u_{ja}\,\delta_{r\ell} \\ &= \langle u_i, u_j\rangle\,\delta_{r\ell}.\end{aligned}$$

The proof is complete, since $|u_i| = 1$ for all i's and by (13) $\langle u_i, u_j\rangle = t$ for all $i \neq j$. □

2.2 *Proof of Theorem 1.1*

The next proposition is the main ingredient for the proof of Theorem 1.1.

Proposition 2.9 *Let $t \in [0, 1]$, $k, n \in \mathbb{N}$, $p = t(n-1)+1$, X be a standard Gaussian random vector in $\mathbb{R}^k$ and $X_1, \cdots, X_n$ be copies of X such that*

$$\mathbb{E}\left[X_i \otimes X_j^*\right] = \left(\mathbb{E}[X_{ir}X_{j\ell}]\right)_{r,\ell\le k} = tI_k, \quad \forall\, i \neq j.$$

Then, for any log-*concave (on its support) function $f : \mathbb{R}^k \to [0, +\infty)$, we have that*

$$\mathbb{E}\left(\prod_{i=1}^{n} f(X_i)\right)^{\frac{1}{n}} \le \left(\mathbb{E}f(X)^{\frac{p}{n}}\right)^{\frac{n}{p}} \le \mathbb{E}f\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \tag{21}$$

Note that, the log-concavity of f implies that

$$\left(\prod_{i=1}^{n} f(X_i)\right)^{\frac{1}{n}} \le f\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right),$$

where equality is achieved for the exponential function $f(x) = e^{\langle \mathrm{a}, x\rangle + c}$, $\mathrm{a} \in \mathbb{R}^k$ and $c \in \mathbb{R}$.

Proof of Proposition 2.9 In order to prove the left-hand side inequality in (21), we will apply Theorem 2.1. Note that the assumption of log-concavity will not be used. The left-hand side inequality in (21) holds true for any non-negative measurable function f.

To be more precise, let $X_1, \ldots, X_n$ be standard Gaussian random vectors in $\mathbb{R}^k$ satisfying condition (20) and $t \in [-\frac{1}{n-1}, 1]$. Then, $\mathbf{X} := (X_1, \ldots, X_n)$, is a centered Gaussian vector in $\mathbb{R}^{kn}$ with covariance the $kn \times kn$ matrix $T = (T_{ij})_{i,j \leq n}$, with block entries the $k \times k$ matrices $T_{ii} = I_k$ and $T_{ij} = tI_k$, for $i \neq j$. Setting

$$p := (n-1)t + 1 \quad \text{and} \quad q := 1 - t,$$

it's not hard to check that, for any $t \in [0.1]$, p is the biggest and q is the smallest singular value of T, while for any $t \in [-\frac{1}{n-1}, 0]$, q is the biggest and p is the smallest singular value of T. Thus,

(i) if $t \geq 0$, then

$$qI_{kn} \leq T \leq pI_{kn},$$

(ii) if $t \leq 0$, then

$$pI_{kn} \leq T \leq qI_{kn}$$

In the above situation, Theorem 2.1 reads as follows:

Theorem 2.10 *Let $k, n \in \mathbb{N}$, $t \in [-\frac{1}{n-1}, 1]$ and let $X_1, \ldots, X_n$ be standard Gaussian random vectors in $\mathbb{R}^k$, with $\mathbb{E}\big[X_i \otimes X_j^*\big] = tI_k$, for all $i \neq j$. Set $p := (n-1)t+1$, $q := 1 - t$, and then for every measurable functions $f_i : \mathbb{R}^k \to [0, +\infty)$, $i = 1, \ldots, n$,*

(i) if $t \in [0, 1]$, then

$$\prod_{i=1}^{n} \Big(\mathbb{E} f_i(X_i)^q\Big)^{1/q} \leq \mathbb{E} \prod_{i=1}^{n} f_i(X_i) \leq \prod_{i=1}^{n} \Big(\mathbb{E} f_i(X_i)^p\Big)^{1/p}, \tag{22}$$

(ii) if $t \in [-\frac{1}{n-1}, 0]$, then

$$\prod_{i=1}^{n} \Big(\mathbb{E} f_i(X_i)^p\Big)^{1/p} \leq \mathbb{E} \prod_{i=1}^{n} f_i(X_i) \leq \prod_{i=1}^{n} \Big(\mathbb{E} f_i(X_i)^q\Big)^{1/q}. \tag{23}$$

Now, the left-hand side inequality of (21) follows immediately from (22), by taking $f_i = f^{1/n}$ for every $i = 1, \ldots, n$.

In order to prove the right-hand side inequality of (21) we apply Barthe's theorem, using the decomposition of the identity in (18). In the following lemma we gather some technical facts.

Lemma 2.11 *Let U_i and E_i, $i = 1, \dots, n$ the matrices defined in* (16) *and* (17)*, and set $p = (n-1)t + 1$, $q = 1 - t$. Then*

$$U_i^* = \sqrt{\frac{p}{n}}\, e_n \otimes I_k + \sqrt{\frac{n-1}{n} q}\, \mathrm{v}_i \otimes I_k \ \in \mathbb{R}^{kn \times k}.$$

$$U_i U_j^* = \langle u_i, u_j \rangle I_k$$

$$U_i E_j^* = \sqrt{\frac{n-1}{n} q}\, \langle \mathrm{v}_i, e_j \rangle I_k$$

for every $i \le n$ and $j \le n-1$.

Proof The first and the second assertion can be verified, just by using the definitions. For the third one, we have

$$\begin{aligned}
U_i E_j^* &= (u_i^* \otimes I_k)(e_j^* \otimes I_k)^* \\
&= \left(\sqrt{\frac{p}{n}}\, e_n^* \otimes I_k + \sqrt{\frac{n-1}{n} q}\, \mathrm{v}_i^* \otimes I_k \right) (e_j \otimes I_k) \\
&= \sqrt{\frac{p}{n}}\, (e_n^* \otimes I_k)(e_j \otimes I_k) + \sqrt{\frac{n-1}{n} q}\, (\mathrm{v}_i^* \otimes I_k)(e_j \otimes I_k) \\
&= \sqrt{\frac{p}{n}}\, e_n^* e_j \otimes I_k + \sqrt{\frac{n-1}{n} q}\, \mathrm{v}_i^* e_j \otimes I_k \\
&= \sqrt{\frac{p}{n}}\, \langle e_n, e_j \rangle I_k + \sqrt{\frac{n-1}{n} q}\, \langle \mathrm{v}_i, e_j \rangle I_k \\
&= \mathbb{O} + \sqrt{\frac{n-1}{n} q}\, \langle \mathrm{v}_i, e_j \rangle I_k.
\end{aligned}$$

□

To finish the proof of Proposition 2.9, we apply Barthe's Theorem 2.2, using the decomposition of the identity appearing in (18). We choose the parameters: $n \leftrightarrow kn$, $m := 2n - 1$, $n_i := k$ for all $i = 1, \dots, 2n - 1$, and

$$c_i := \begin{cases} \frac{1}{p} & , i = 1, \dots, n \\ \frac{nt}{p} & , i = n+1, \dots, 2n-1 \end{cases} .$$

Then, we apply Theorem 2.2 to the functions

$$\tilde{f}_i(x) := \begin{cases} f(x)^{\frac{p}{n}} & , i = 1, \dots, n \\ 1 & , i = n+1, \dots, 2n-1 \end{cases} , \quad x \in \mathbb{R}^k$$

and

$$h(x) := f\left(\frac{1}{n}\sum_{i=1}^{n} U_i x\right), \quad x \in \mathbb{R}^{kn}.$$

For any $\xi_1, \ldots, \xi_n \in \mathbb{R}^k$, by Lemma 2.11, we get that

$$\begin{aligned}
&h\left(\sum_{j=1}^{n}\frac{1}{p}U_j^*\xi_j + \sum_{a=1}^{n-1}\frac{nt}{p}E_a^*\xi_{n+a}\right)\\
&= f\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{1}{p}U_iU_j^*\xi_j + \frac{1}{n}\sum_{i=1}^{n}\sum_{a=1}^{n-1}\frac{nt}{p}U_iE_a^*\xi_{n+a}\right)\\
&= f\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{1}{p}U_iU_j^*\xi_j + \frac{1}{n}\sum_{i=1}^{n}\sum_{a=1}^{n-1}\frac{nt}{p}\sqrt{\frac{n-1}{n}}q\langle \mathrm{v}_i, e_a\rangle\xi_{n+a}\right)\\
&= f\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{1}{p}U_iU_j^*\xi_j\right) \qquad \left(\text{since } \sum \mathrm{v}_i = 0\right)\\
&= f\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\frac{1}{p}\langle u_i, u_j\rangle\xi_j\right)\\
&= f\left(\frac{1}{n}\sum_{i=1}^{n}\Big(\frac{1}{p}\xi_i + \sum_{j\neq i}\frac{t}{p}\xi_j\Big)\right)\\
&= f\left(\frac{1}{n}\sum_{i=1}^{n}\Big(\frac{1}{p} + (n-1)\frac{t}{p}\Big)\xi_i\right)\\
&= f\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i\right)\\
&\geq \prod_{i=1}^{n} f(\xi_i)^{\frac{1}{n}} = \prod_{i=1}^{n}\left(f(\xi_i)^{\frac{p}{n}}\right)^{\frac{1}{p}} = \prod_{i=1}^{n}\tilde{f}(\xi_i)^{c_i}.
\end{aligned}$$

Thus, Theorem 2.2 implies

$$\mathbb{E}f\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right) = \mathbb{E}f\left(\frac{1}{n}\sum_{i=1}^{n}U_iZ\right) \geq \prod_{i=1}^{n}\left(\mathbb{E}f(X_i)^{\frac{p}{n}}\right)^{\frac{1}{p}} = \left(\mathbb{E}f(X)^{\frac{p}{n}}\right)^{\frac{n}{p}}$$

and the proof is complete. □

We close this section with the proof of our primary result.

Proof of Theorem 1.1 Suppose first that $X \sim N(0, I_k)$. Then, under the notation of Lemma 2.8 we have that

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n} U_i\mathbf{Z} &= \frac{1}{n}\sum_{i=1}^{n}\sqrt{\frac{p}{n}}\,(e_n^* \otimes I_k)\mathbf{Z} \;+\; \frac{1}{n}\sum_{i=1}^{n}\sqrt{\frac{n-1}{n}q}\,(\mathrm{v}_i^* \otimes I_k)\mathbf{Z} \\ &= \sqrt{\frac{p}{n}}\,(e_n^* \otimes I_k)\mathbf{Z} \;+\; \frac{1}{n}\sqrt{\frac{n-1}{n}q}\,\left(\sum_{i=1}^{n}\mathrm{v}_i^*\right)\otimes I_k\,\mathbf{Z} \\ &= \sqrt{\frac{p}{n}}\,E_n\mathbf{Z} \;+\; \frac{1}{n}\sqrt{\frac{n-1}{n}q}\,\left(\sum_{i=1}^{n}\mathrm{v}_i\right)^{*}\otimes I_k\,\mathbf{Z} \\ &= \sqrt{\frac{p}{n}}\,Z_n.\end{aligned}$$

Thus, the right hand side of (21) can be written as

$$\mathbb{E}f\left(\sqrt{\frac{p}{n}}X\right) \geq \left(f(X)^{\frac{p}{n}}\right)^{\frac{n}{p}}. \tag{24}$$

where $p = (n-1)t + 1$, $n \in \mathbb{N}$, and $t \in [0, 1]$.

Consequently, if $f : \mathbb{R}^k \to [0, +\infty)$ is a log-concave function and $r \in (0, 1]$, then there exist $t \in [0, 1]$ and $n \in \mathbb{N}$, such that $r = \frac{p}{n} = \frac{(n-1)t+1}{n}$ and so by (24) we get that

$$\mathbb{E}f\big(\sqrt{r}X\big) \geq (\mathbb{E}f(X)^r)^{\frac{1}{r}} \tag{25}$$

for every $r \in (0, 1]$. We consider now the case where $r = 0$. Since f is *log*-concave, there exists a convex function $v : \mathbb{R}^k \to \mathbb{R}$ such that $f = e^{-v}$. Then, for $r = 0$, inequality (1) is equivalent to Jensen's inequality

$$v(0) = v(\mathbb{E}X) \leq \mathbb{E}v(X), \tag{26}$$

and the proof of (1) is complete.

For every $q \geq 1$ consider $r = \frac{1}{q} \in (0, 1]$. Let $F(x) = f(x/\sqrt{r})^{1/r}$ which is also log-concave and so (25) for F and r implies

$$\mathbb{E}f(X)^q \geq \big(\mathbb{E}f(\sqrt{q}X)\big)^q, \tag{27}$$

and (2) follows.

Assume now that $g : \mathbb{R}^n \to [0, +\infty)$ is log-convex and $r \in (0, 1]$. By the log-convexity of g and Theorem 2.10(i), we have that

$$\mathbb{E}g\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \leq \mathbb{E}\prod_{i=1}^{n} g(X_i)^{\frac{1}{n}} \leq \left(\mathbb{E}g(X)^{\frac{p}{n}}\right)^{\frac{n}{p}}. \tag{28}$$

As we have seen at the beginning of the proof $\frac{1}{n}\sum_{i=1}^{n} X_i \overset{d}{=} \sqrt{\frac{p}{n}}X$. So, using (28) for $t \in [0, 1]$ and $n \in \mathbb{N}$ such that $\frac{p}{n} = \frac{(n-1)t+1}{n} = r$, we derive that

$$\mathbb{E}g\left(\sqrt{r}X\right) \leq \left(\mathbb{E}g(X)^r\right)^{\frac{1}{r}},$$

for every $r \in (0, 1]$. The rest of the proof for a log-convex function g is identical to the log-concave case.

For the equality case, a straightforward computation shows that for $f(x) = e^{\langle \mathrm{a}, x\rangle + c}$, we have that

$$\mathbb{E}f(\sqrt{q}X) = C\exp\left(\frac{q}{2}|\mathrm{a}|^2\right) = \left(\mathbb{E}f(X)^q\right)^{\frac{1}{q}}.$$

for every $q \geq 0$.

Finally, suppose that X is a general Gaussian random vector in $\mathbb{R}^k$ with expectation $\xi \in \mathbb{R}^k$ and covariance matrix $T = UU^*$ where $U \in \mathbb{R}^{k\times k}$. Note, that if f is log-concave (or log-convex) and positive function on $\mathbb{R}^k$, then so is $F(x) := f(Ux - \xi)$. Moreover, if $Z \sim N(0, I_k)$ then $UZ - \xi \overset{d}{=} X \sim N(0, T)$. The general case follows then, by applying the previous case on function F. □

3 Reverse Logarithmic Sobolev Inequality

In the next lemma, we state the *Gaussian Integration by Parts* formula (see [18, Appendix 4] for a simple proof).

Lemma 3.1 *Let $X, Y_1, \ldots, Y_n$ be centered jointly Gaussian random variables, and F be a real valued function on $\mathbb{R}^n$, that satisfy the growth condition*

$$\lim_{|x|\to\infty} |F(x)| \exp\left(-a|x|^2\right) = 0 \qquad \forall\, a > 0. \tag{29}$$

Then

$$\mathbb{E}\big[XF(Y_1, \ldots, Y_n)\big] = \sum_{i=1}^{n} \mathbb{E}\big[XY_i\big]\, \mathbb{E}\big[\partial_i F(Y_1, \ldots, Y_n)\big]. \tag{30}$$

Involving this formula, we can further elaborate Corollary 1.2.

Let $\mathcal{G}_k$, be the class of all positive functions in $\mathbb{R}^k$, such that their first derivatives satisfy the growth condition (29). Then for any $f \in \mathcal{G}_k$, by Lemma 3.1, we get that

$$\begin{aligned}\mathbb{E}\big[\langle X, \nabla f(X)\rangle\big] &= \sum_{i=1}^{k} \mathbb{E}\big[X_i \partial_i f(X)\big] \\ &= \sum_{i=1}^{k}\sum_{j=1}^{k} \mathbb{E}\big[X_i X_j\big]\,\mathbb{E}\big[\partial_{ij} f(X)\big] = \mathbb{E}\big[\mathrm{tr}\big(T\,H_f(X)\big)\big],\end{aligned}$$

where T is the covariance matrix of X and $H_f(x)$ stands for the Hessian matrix of f at $x \in \mathbb{R}^k$. In the special case where $X \sim N(0, I_k)$, Corollary 1.2 implies the following:

Corollary 3.2 *Let $k \in \mathbb{N}$, and X be a standard Gaussian vector in $\mathbb{R}^k$. Then*

(i) for every log-concave function $f \in \mathcal{G}_k$, we have

$$\mathrm{Ent}_X(f) \geq \frac{1}{2}\mathbb{E}\Delta f(X), \tag{31}$$

(ii) for every log-convex function $f \in \mathcal{G}_k$, we have

$$\mathrm{Ent}_X(f) \leq \frac{1}{2}\mathbb{E}\Delta f(X). \tag{32}$$

Proof of Theorem 1.3 Let $f \in \mathcal{L}^{2,1}(\gamma_k)$. Without loss of generality we may also assume that $\mathbb{E}f^2(X) = 1$. Suppose first that f has a bounded support. Then $f^2 \in \mathcal{G}_k$ and Corollary 3.2, after an application of the chain rule $\frac{1}{2}\Delta f^2 = |\nabla f|^2 + f\Delta f$, gives that

$$\mathbb{E}|\nabla f(X)|^2 + \mathbb{E}f(X)\Delta f(X) \leq \mathrm{Ent}_X(f^2) \leq 2\,\mathbb{E}|\nabla f(X)|^2. \tag{33}$$

Let $f = e^{-v}$, where $v : supp(f) \to \mathbb{R}$ is a convex function. Again by the chain rule we have $f\Delta f = |\nabla f|^2 - f^2\Delta v$, and so

$$\mathbb{E}f(X)\Delta f(X) = \mathbb{E}|\nabla f(X)|^2 - \mathbb{E}f(X)^2\Delta v(X). \tag{34}$$

Equations (33) and (34), prove Theorem 1.3 in this case.

To drop the assumption of the bounded support, we consider the functions $f_n := f\,\mathbf{1}_{nB_2^k}$, where $\mathbf{1}_{nB_2^k}$ is the indicator function of the Euclidean Ball in $\mathbb{R}^k$ with radius $n \in \mathbb{N}$. Every f_n has bounded support and so by the previous case,

$$2\,\mathbb{E}|\nabla f_n(X)|^2 - \mathbb{E}f_n(X)^2\Delta v_n(X) \leq \mathrm{Ent}_X(f_n^2). \tag{35}$$

In order to avoid any possible problem of infiniteness of the derivatives of f_n, $n \in \mathbb{N}$, we define the functions

$$F_n = |\nabla f|^2 \cdot \mathbf{1}_{nB_2^k}, \qquad H_n = f^2 \Delta v \cdot \mathbf{1}_{nB_2^k}.$$

Notice that $F_n = |\nabla f_n|^2$ and $H_n = f_n^2 \Delta v_n$ almost everywhere, since they could only differ on the zero-measure set $\{x \in \mathbb{R}^k : \ |x| = n\}$. Thus,

$$0 \le f_n \nearrow f, \qquad 0 \le F_n \nearrow |\nabla f|^2, \qquad 0 \le H_n \nearrow f^2 \Delta v,$$

and by the monotone convergence theorem

$$\mathbb{E}|\nabla f_n(X)|^2 = \mathbb{E}F_n(X) \longrightarrow \mathbb{E}|\nabla f(X)|^2 \tag{36}$$

and

$$\mathbb{E}f_n(X)^2 \Delta v_n(X) = \mathbb{E}H_n(X) \longrightarrow \mathbb{E}f(X)^2 \Delta v(X). \tag{37}$$

Moreover, $f_n^2 \log f_n^2 \to f^2 \log f^2$ and $|f_n^2 \log f_n^2| \le |f^2 \log f^2|$, for every $n \in \mathbb{N}$ (where we have taken that $0 \log 0 = 0$). Since, by Gross' inequality, $f^2 \log f^2 \in L^1(\gamma_k)$, the Lebesgue's dominated convergence theorem implies that

$$\mathrm{Ent}_X(f_n^2) \longrightarrow \mathrm{Ent}_X(f^2). \tag{38}$$

Under the light of (36)–(38), the desired result follows by taking the limit in (35), as $n \to \infty$. □

Acknowledgements The research leading to these results is part of a project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 637851).

Part of this work was done while the first named author was a postdoctoral research fellow at the University of Crete, and he was supported by the Action Supporting Postdoctoral Researchers of the Operational Program Education and Lifelong Learning (Actions Beneficiary: General Secretariat for Research and Technology), co-financed by the European Social Fund (ESF) and the Greece State.

The second named author is supported by the US NSF grant CAREER-1151711 and BSF grant 2010288.

Finally, we would like to thank the anonymous referee whose valuable remarks improved the presentation of the paper.

References

1. K.M. Ball, *Volumes of Sections of Cubes and Related Problems*. Lecture Notes in Mathematics, vol. 1376 (Springer, Berlin, 1989), pp. 251–260
2. F. Barthe. On a reverse form of the Brascamp-Lieb inequality. Invent. Math. **134**, 335–361 (1998)

3. W. Beckner, Inequalities in Fourier analysis. Ann. Math. **102**, 159–182 (1975)
4. V. Bogachev, *Gaussian Measures*. Mathematical Surveys and Monographs, vol. 62 (American Mathematical Society, Providence, 1998)
5. H.J. Brascamp, E.H. Lieb, Best constants in Young's inequality, its converse, and its generalization to more than three functions. Adv. Math. **20**, 151–173 (1976)
6. E.A. Carlen, Superadditivity of Fisher's information and logarithmic Sobolev inequalities. J. Funct. Anal. **101**, 194–211 (1991)
7. W.-K. Chen, N. Dafnis, G. Paouris, Improved Hölder and reverse Hölder inequalities for Gaussian random vectors. Adv. Math. **280**, 643–689 (2015)
8. M. Fathi, E. Indrei, M. Ledoux, Quantitative logarithmic Sobolev inequalities and stability estimates. Discrete Contin. Dyn. Syst. 36(12), 6835–6853 (2016)
9. A. Figalli, F. Maggi, A. Pratelli. Sharp stability theorems for the anisotropic Sobolev and log-Sobolev inequalities on functions of bounded variation. Adv. Math. **242**, 80–101 (2013)
10. L. Gross. Logarithmic Sobolev inequalities. Am. J. Math. **97**, 1061–1083 (1975)
11. E. Indrei, D. Marcon, Quantitative log-Sobolev inequality for a two parameter family of functions. Int. Math. Res. Not. **20**, 5563–5580 (2014)
12. M. Ledoux, Remarks on Gaussian noise stability, Brascamp-Lieb and Slepian inequalities, in *Geometric Aspects of Functional Analysis*. Lecture Notes in Mathematics, vol. 2116 (Springer, Cham, 2014), pp. 309–333
13. E.H. Lieb, Gaussian kernels have only Gaussian maximizers. Inv. Math. **102**, 179–208 (1990)
14. E.H. Lieb, M. Loss, *Analysis*, 2nd edn. Graduate Studies in Mathematics, vol. 14 (American Mathematical Society, Providence, 2001)
15. E. Mossel, K. Oleszkiewicz, A. Sen, On reverse hypercontractivity. Geom. Funct. Anal. **23**(3), 1062–1097 (2013)
16. J. Neeman, A multi-dimensional version of noise stability. Electron. Commun. Probab. **19**(72), 1–10 (2014)
17. E. Nelson, The free Markov field. J. Funct. Anal. **12**, 211–227 (1973)
18. M. Talagrand, *Mean Field Models for Spin Glasses. Volume I* (Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge). A Series of Modern Surveys in Mathematics, vol. **54** (Springer, Berlin, 2010)
19. C. Villani, *Topics in Optimal Transportation*. Graduate Studies in Mathematics, vol. 58 (American Mathematical Society, Providence, 2003)
20. C. Villani, *Optimal Transport. Old and New* (Grundlehren der mathematischen Wissenschaften), vol. 338 (Springer, Berlin, 2009)