Moment-entropy inequalities for a random vector

Erwin Lutwak, Deane Yang, and Gaoyong Zhang

***Abstract*—The p-th moment matrix is defined for a real random vector, generalizing the classical covariance matrix. Sharp inequalities relating the p-th moment and Renyi entropy are established, generalizing the classical inequality relating the second moment and the Shannon entropy. The extremal distributions for these inequalities are completely characterized.**

***Index Terms*—random vector, entropy, Renyi entropy, covariance, covariance matrix, moment, moment matrix, information theory, information measure**

I. Introduction

In [1] the authors demonstrated how the classical information theoretic inequality for the Shannon entropy and second moment of a real random variable could be extended to inequalities for Renyi entropy and the p-th moment. The extremals of these inequalities were also completely characterized. Moment-entropy inequalities, using Renyi entropy, for discrete random variables have also been obtained by Arikan [2].

We describe how to extend the definition of the second moment matrix of a real random vector to that of the p-th moment matrix. Using this, we extend the moment-entropy inequalities and the characterization of the extremal distributions proved in [1] to higher dimensions.

The results in this paper extend earlier work of the authors (with O. Guleryuz) [3] and Costas-Hero-Vignat [4] (also, see recent work of Johnson-Vignat [5]). Variants and generalizations of the theorems presented can be found in work of the authors [6], [7], [8], [9] and Bastero-Romance [10].

The authors would like to thank Christoph Haberl for his careful reading of this paper and valuable suggestions for improving it.

II. The p-th moment matrix of a random vector

A. Basic notation

Throughout this paper we denote:

$$\begin{aligned} \mathbb{R}^n &= \text{ } n\text{-dimensional Euclidean space} \\ x \cdot y &= \text{ standard Euclidean inner product of } x, y \in \mathbb{R}^n \\ |x| &= \sqrt{x \cdot x} \\ S &= \text{ positive definite symmetric } n\text{-by-}n \text{ matrices} \\ |A| &= \text{ determinant of an } n\text{-by-}n \text{ matrix } A \end{aligned}$$

For each $A \in S$, define the norm $|\cdot|_A$ by

$$|x|_A = |Ax| = \sqrt{Ax \cdot Ax},$$

for each $x \in \mathbb{R}^n$.

Throughout this paper, we will denote the standard Lebesgue density on $\mathbb{R}^n$ by dx.

If X is a random vector in $\mathbb{R}^n$, then the associated probability measure on $\mathbb{R}^n$ will be denoted by m_X. If the measure m_X is absolutely continuous with respect to Lebesgue measure, then the corresponding Radon-Nikodym derivative is called the *density function of the random vector* X and denoted by f_X.

If A is an invertible n-by-n matrix, then

$$f_{AX}(y) = |A|^{-1} f_X(A^{-1}y), \tag{1}$$

for each $y \in \mathbb{R}^n$.

If Φ is a continuous scalar-, vector-, or matrix-valued function on $\mathbb{R}^n$, then the expected value of $\Phi(X)$ is given by

$$E[\Phi(X)] = \int_{\mathbb{R}^n} \Phi(x)\, dm_X(x).$$

If $v \in \mathbb{R}^n$, we denote by $v \otimes v$ the n-by-n matrix whose (i,j)-th component is $v_i v_j$. We call a random vector X *nondegenerate*, if the matrix $E[X \otimes X]$ is positive definite.

B. The p-th moment of a random vector

For $p \in (0, \infty)$, the *standard p-th moment* of a random vector X is given by

$$E[|X|^p] = \int_{\mathbb{R}^n} |x|^p \, dm_X(x). \tag{2}$$

More generally, the p-th moment with respect to the norm $|\cdot|_A$ is

$$E[|X|_A^p] = \int_{\mathbb{R}^n} |x|_A^p \, dm_X(x).$$

C. The p-th moment matrix

The second moment matrix of a random vector X is defined to be

$$M_2[X] = E[X \otimes X].$$

Recall that $M_2[X - E[X]]$ is the covariance matrix. An important observation is that the definition of the moment matrix does not use the inner product on $\mathbb{R}^n$.

A characterization of the second moment matrix is the following: *The matrix $M_2[X]^{-1/2}$ is the unique positive definite symmetric matrix with maximal determinant among all matrices $A \in S$ satisfying $E[|X|_A^2] = n$.*

We extend this characterization to a definition of the p-th moment matrix $M_p[X]$ for all $p \in (0, \infty)$.

Theorem 1: If $p \in (0, \infty)$ and X is a nondegenerate random vector in $\mathbb{R}^n$ with finite p-th moment, then there exists a unique matrix $A \in S$ such that

$$E[|X|_A^p] = n$$

E. Lutwak (elutwak@poly.edu), D. Yang (dyang@poly.edu), and G. Zhang (gzhang@poly.edu) are with the Department of Mathematics, Polytechnic University, Brooklyn, New York. and were supported in part by NSF Grant DMS-0405707.

and

$$|A| \geq |A'|,$$

for each $A' \in S$ such that $E[|X|^p_{A'}] = n$. Moreover, the matrix A is the unique matrix in S satisfying

$$I = E[|AX|^{p-2}(AX) \otimes (AX)]. \tag{3}$$

We define the *p-th moment matrix* of a random vector X to be $M_p[X] = A^{-p}$, where A is given by the theorem above.

The proof of the theorem is given in §IV

III. Moment-entropy inequalities

A. Entropy

The *Shannon entropy* of a random vector X is defined to be

$$h[X] = -\int_{\mathbb{R}^n} f_X \log f_X \, dx,$$

provided that the integral above exists. For $\lambda > 0$ the *λ-Renyi entropy power* of a density function is defined to be

$$N_\lambda[X] = \begin{cases} \left(\int_{\mathbb{R}^n} f_X^\lambda\right)^{\frac{1}{1-\lambda}} & \text{if } \lambda \neq 1, \\ e^{h[f]} & \text{if } \lambda = 1, \end{cases}$$

provided that the integral above exists. Observe that

$$\lim_{\lambda \to 1} N_\lambda[X] = N_1[X].$$

The *λ–Renyi entropy* of a random vector X is defined to be

$$h_\lambda[X] = \log N_\lambda[X].$$

The entropy $h_\lambda[X]$ is continuous in λ and, by the Hölder inequality, decreasing in λ. It is strictly decreasing, unless X is a uniform random vector.

It follows by (1) that

$$N_\lambda[AX] = |A| N_\lambda[X], \tag{4}$$

for each $A \in S$.

B. Relative entropy

Given two random vectors X, Y in $\mathbb{R}^n$, their *relative Shannon entropy* or *Kullback–Leibler distance* [11], [12], [13] (also, see page 231 in [14]) is defined by

$$h_1[X,Y] = \int_{\mathbb{R}^n} f_X \log\left(\frac{f_X}{f_Y}\right) dx, \tag{5}$$

provided that the integral above exists. Given $\lambda > 0$, we define the *relative λ–Renyi entropy power of X and Y* as follows. If $\lambda \neq 1$, then

$$N_\lambda[X,Y] = \frac{\left(\int_{\mathbb{R}^n} f_Y^{\lambda-1} f_X \, dx\right)^{\frac{1}{1-\lambda}} \left(\int_{\mathbb{R}^n} f_Y^\lambda \, dx\right)^{\frac{1}{\lambda}}}{\left(\int_{\mathbb{R}^n} f_X^\lambda \, dx\right)^{\frac{1}{\lambda(1-\lambda)}}}, \tag{6}$$

and, if $\lambda = 1$, then

$$N_1[X,Y] = e^{h_1[X,Y]},$$

provided in both cases that the righthand side exists. Define the *λ–Renyi relative entropy of random vectors X and Y* by

$$h_\lambda[X,Y] = \log N_\lambda[X,Y].$$

Observe that $h_\lambda[X,Y]$ is continuous in λ.

Lemma 2: If X and Y are random vectors such that $h_\lambda[X]$, $h_\lambda[Y]$, and $h_\lambda[X,Y]$ are finite, then

$$h_\lambda[X,Y] \geq 0.$$

Equality holds if and only if $X = Y$.

Proof: If $\lambda > 1$, then by the Hölder inequality,

$$\int_{\mathbb{R}^n} f_Y^{\lambda-1} f_X \, dx \leq \left(\int_{\mathbb{R}^n} f_Y^\lambda \, dx\right)^{\frac{\lambda-1}{\lambda}} \left(\int_{\mathbb{R}^n} f_X^\lambda \, dx\right)^{\frac{1}{\lambda}},$$

and if $\lambda < 1$, then we have

$$\begin{aligned} \int_{\mathbb{R}^n} f_X^\lambda &= \int_{\mathbb{R}^n} (f_Y^{\lambda-1} f_X)^\lambda f_Y^{\lambda(1-\lambda)} \\ &\leq \left(\int_{\mathbb{R}^n} f_Y^{\lambda-1} f_X\right)^\lambda \left(\int_{\mathbb{R}^n} f_Y^\lambda\right)^{1-\lambda}. \end{aligned}$$

The inequality for $\lambda = 1$ follows by taking the limit $\lambda \to 1$.

The equality conditions for $\lambda \neq 1$ follow from the equality conditions of the Hölder inequality. The inequality for $\lambda = 1$, including the equality condition, follows from the Jensen inequality (details may be found, for example, on page 234 in [14]). ■

C. Generalized Gaussians

We call the extremal random vectors for the moment-entropy inequalities *generalized Gaussians* and recall their definition here.

Given $t \in \mathbb{R}$, let

$$t_+ = \max(t, 0).$$

Let

$$\Gamma(t) = \int_0^\infty x^{t-1} e^{-x} \, dx$$

denote the Gamma function, and let

$$\beta(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

denote the Beta function.

For each $p \in (0,\infty)$ and $\lambda \in (n/(n+p), \infty)$, let Z be the random vector in $\mathbb{R}^n$ whose density function $f_Z : \mathbb{R}^n \to [0,\infty)$ is given by

$$f_Z(x) = \begin{cases} a_{p,\lambda}(1 + (1-\lambda)|x|^p)_+^{1/(\lambda-1)} & \text{if } \lambda \neq 1 \\ a_{p,1} e^{-|x|^p} & \text{if } \lambda = 1, \end{cases} \tag{7}$$

where

$$a_{p,\lambda} = \begin{cases} \dfrac{(1-\lambda)^{\frac{n}{p}+1}\Gamma(\frac{n}{2}+1)}{\pi^{\frac{n}{2}}\beta(\frac{n}{p}+1, \frac{1}{1-\lambda} - \frac{n}{p})} & \text{if } \lambda < 1, \\ \dfrac{\Gamma(\frac{n}{2}+1)}{\pi^{\frac{n}{2}}\Gamma(\frac{n}{p}+1)} & \text{if } \lambda = 1, \\ \dfrac{(\lambda-1)^{\frac{n}{p}+1}\Gamma(\frac{n}{2}+1)}{\pi^{\frac{n}{2}}\beta(\frac{n}{p}+1, \frac{1}{\lambda-1})} & \text{if } \lambda > 1, \end{cases}$$

Define the *standard generalized Gaussian* to be the random vector $\widehat{Z}$ defined by

$$\widehat{Z} = [\lambda(n+p)-n]^{1/p} Z. \tag{8}$$

Any random vector Y in $\mathbb{R}^n$ that can be written as $Y = AZ$, for some invertible n-by-n matrix A is called a *generalized Gaussian*.

D. Information measures of generalized Gaussians

If $0 < p < \infty$ and $\lambda > n/(n+p)$, then the p-th moment of the random vector Z is

$$E[|Z|^p] = \frac{n}{\lambda(n+p)-n},$$

and therefore the standard generalized Gaussian $\widehat{Z}$ is

$$E[|\widehat{Z}|^p] = n.$$

Its p-th moment matrix is $M_p[\widehat{Z}] = I$.

If $0 < p < \infty$ and $\lambda > n/(n+p)$, the λ-Renyi entropy power of the random vector Z is given by

$$N_\lambda[Z] = \begin{cases} \left(1 + \dfrac{n(\lambda-1)}{p\lambda}\right)^{\frac{1}{\lambda-1}} a_{p,\lambda}^{-1} & \text{if } \lambda \neq 1 \\ e^{\frac{n}{p}} a_{p,1}^{-1} & \text{if } \lambda = 1 \end{cases}$$

It follows by (4) and (8) that

$$N_\lambda[\widehat{Z}] = [\lambda(n+p)-n]^{\frac{n}{p}} N_\lambda[Z].$$

Define the constant

$$\begin{aligned} c(n,p,\lambda) &= \frac{E[|Z|^p]^{1/p}}{N_\lambda[Z]^{1/n}} \\ &= a_{p,\lambda}^{1/n} \left[\lambda\left(1+\frac{p}{n}\right)-1\right]^{-\frac{1}{p}} b(n,p,\lambda), \end{aligned} \tag{9}$$

where

$$b(n,p,\lambda) = \begin{cases} \left(1 - \frac{n(1-\lambda)}{p\lambda}\right)^{\frac{1}{n(1-\lambda)}} & \text{if } \lambda \neq 1 \\ e^{-1/p} & \text{if } \lambda = 1. \end{cases}$$

Observe that if $\lambda \neq 1$ and $0 < p < \infty$, then

$$\int_{\mathbb{R}^n} f_Z^\lambda = a_{p,\lambda}^{\lambda-1}(1 + (1-\lambda)E[|Z|^p]), \tag{10}$$

and if $\lambda = 1$, then

$$h[Z] = -\log a_{p,1} + E[|Z|^p]. \tag{11}$$

We will also need the following scaling identities:

$$f_{tZ}(x) = t^{-n} f_Z(t^{-1}x), \tag{12}$$

for each $x \in \mathbb{R}^n$. Therefore,

$$\int_{\mathbb{R}^n} f_{tZ}^\lambda \, dx = t^{n(1-\lambda)} \int_{\mathbb{R}^n} f_Z^\lambda \, dx. \tag{13}$$

E. Spherical moment-entropy inequalities

The proof of Theorem 2 in [1] extends easily to prove the following. A more general version can be found in [7].

Theorem 3: If $p \in (0,\infty)$, $\lambda > n/(n+p)$, and X is a random vector in $\mathbb{R}^n$ such that $N_\lambda[X], E[|X|^p] < \infty$, then

$$\frac{E[|X|^p]^{1/p}}{N_\lambda[X]^{1/n}} \geq c(n,p,\lambda),$$

where $c(n,p,\lambda)$ is given by (9). Equality holds if and only if $X = tZ$, for some $t \in (0,\infty)$.

Proof: For convenience let $a = a_{p,\lambda}$. Let

$$t = \left(\frac{E[|X|^p]}{E[|Z|^p]}\right)^{1/p} \tag{14}$$

and $Y = tZ$.

If $\lambda \neq 1$, then by (12) and (7), (2), (14), and (10),

$$\begin{aligned} &\int_{\mathbb{R}^n} f_Y^{\lambda-1} f_X \\ &= a^{\lambda-1} t^{n(1-\lambda)} \int_{\mathbb{R}^n} (1 + (1-\lambda)|t^{-1}x|^p)_+ f_X(x)\, dx \\ &\geq a^{\lambda-1} t^{n(1-\lambda)} \left(1 + (1-\lambda)t^{-p} \int_{\mathbb{R}^n} |x|^p f_X(x)\, dx\right) \\ &= a^{\lambda-1} t^{n(1-\lambda)} (1 + (1-\lambda)t^{-p} E[|X|^p]) \\ &= a^{\lambda-1} t^{n(1-\lambda)} (1 + (1-\lambda)E[|Z|^p]) \\ &= t^{n(1-\lambda)} \int_{\mathbb{R}^n} f_Z^\lambda, \end{aligned} \tag{15}$$

where equality holds if $\lambda < 1$. It follows that if $\lambda \neq 1$, then by Lemma 2, (6), (13) and (15), and (14), we have

$$\begin{aligned} 1 &\leq N_\lambda[X,Y]^\lambda \\ &= \left(\int_{\mathbb{R}^n} f_Y^\lambda\right) \left(\int_{\mathbb{R}^n} f_X^\lambda\right)^{-\frac{1}{1-\lambda}} \left(\int_{\mathbb{R}^n} f_Y^{\lambda-1} f_X\right)^{\frac{\lambda}{1-\lambda}} \\ &\leq t^n \frac{N_\lambda[Z]}{N_\lambda[X]} \\ &= \frac{E[|X|^p]^{n/p}}{N_\lambda[X]} \frac{N_\lambda[Z]}{E[|Z|^p]^{n/p}}. \end{aligned}$$

If $\lambda = 1$, then by Lemma 2, (5) and (7), and (11) and (14),

$$\begin{aligned} 0 &\leq h_1[X,Y] \\ &= -h[X] - \log a + n \log t + t^{-p} E[|X|^p] \\ &= -h[X] + h[Z] + \frac{n}{p} \log \frac{E[|X|^p]}{E[|Z|^p]}. \end{aligned}$$

Lemma 2 shows that equality holds in all cases if and only if $Y = X$. ■

F. Elliptic moment-entropy inequalities

Corollary 4: If $A \in S$, $p \in (0,\infty)$, $\lambda > n/(n+p)$, and X is a random vector in $\mathbb{R}^n$ satisfying $N_\lambda[X], E[|X|^p] < \infty$, then

$$\frac{E[|X|_A^p]^{1/p}}{|A|^{1/n} N_\lambda[X]^{1/n}} \geq c(n,p,\lambda), \tag{16}$$

where $c(n,p,\lambda)$ is given by (9). Equality holds if and only if $X = tA^{-1}Z$ for some $t \in (0,\infty)$.

Proof: By (4) and Theorem 3,

$$\frac{E[|X|_A^p]^{1/p}}{|A|^{1/n}N_\lambda[X]^{1/n}} = \frac{E[|AX|^p]^{1/p}}{N_\lambda[AX]^{1/n}} \geq \frac{E[|Z|^p]^{1/p}}{N_\lambda[Z]^{1/n}},$$

and equality holds if and only if $AX = tZ$ for some $t \in (0,\infty)$. ∎

G. Affine moment-entropy inequalities

Optimizing Corollary 4 over all $A \in S$ yields the following affine inequality.

Theorem 5: If $p \in (0,\infty)$, $\lambda > n/(n+p)$, and X is a random vector in $\mathbb{R}^n$ satisfying $N_\lambda[X], E[|X|^p] < \infty$, then

$$\frac{|M_p[X]|^{1/p}}{N_\lambda[X]} \geq n^{-n/p} c(n,p,\lambda)^n,$$

where $c(n,p,\lambda)$ is given by (9). Equality holds if and only if $X = A^{-1}Z$ for some $A \in S$.

Proof: Substitute $A = M_p[X]^{-1/p}$ into (16) ∎

Conversely, Corollary 4 follows from Theorem 5 by Theorem 1.

IV. Proof of Theorem 1

A. Isotropic position of a probability measure

A Borel measure μ on $\mathbb{R}^n$ is said to be *in isotropic position*, if

$$\int_{\mathbb{R}^n} \frac{x \otimes x}{|x|^2}\, d\mu(x) = \frac{1}{n} I, \tag{17}$$

where I is the identity matrix.

Lemma 6: If $p \geq 0$ and μ is a Borel probability measure in isotropic position, then for each $A \in S$,

$$|A|^{-1/n} \left(\int_{\mathbb{R}^n} \frac{|Ax|^p}{|x|^p}\, d\mu(x) \right)^{1/p} \geq 1,$$

with equality holding if and only if $A = aI$ for some $a > 0$.

Proof: By Hölder's inequality,

$$\left(\int_{\mathbb{R}^n} \frac{|Ax|^p}{|x|^p}\, d\mu(x) \right)^{1/p} \geq \exp\left(\int_{\mathbb{R}^n} \log \frac{|Ax|}{|x|}\, d\mu(x) \right),$$

so it suffices to prove the $p = 0$ case only.

By (17),

$$\int_{\mathbb{R}^n} \frac{(x \cdot e)^2}{|x|^2}\, d\mu(x) = \frac{1}{n}, \tag{18}$$

for any unit vector e.

Let $e_1, \ldots, e_n$ be an orthonormal basis of eigenvectors of A with corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$. By the concavity of log, and (18),

$$\begin{aligned}\int_{\mathbb{R}^n} \log \frac{|Ax|}{|x|}\, d\mu(x) &= \frac{1}{2} \int_{\mathbb{R}^n} \log \frac{|Ax|^2}{|x|^2}\, d\mu(x) \\ &= \frac{1}{2} \int_{\mathbb{R}^n} \log \sum_{i=1}^n \lambda_i^2 \frac{(x \cdot e_i)^2}{|x|^2}\, d\mu(x) \\ &\geq \frac{1}{2} \int_{\mathbb{R}^n} \sum_{i=1}^n \frac{(x \cdot e_i)^2}{|x|^2} \log \lambda_i^2\, d\mu(x) \\ &= \log |A|^{1/n}.\end{aligned}$$

The equality condition follows from the strict concavity of log. ∎

B. Proof of theorem

Lemma 7: If $p > 0$ and X is a nondegenerate random vector in $\mathbb{R}^n$ with finite p-th moment, then there exists $c > 0$ such that

$$E[|e \cdot X|^p] \geq c, \tag{19}$$

for every unit vector e.

Proof: The assumption that X is nondegenerate and has finite p-th moment implies that the left side of (19) is a positive continuous function of e in the unit sphere, which is compact. ∎

Theorem 8: If $p \geq 0$ and X is a nondegenerate random vector in $\mathbb{R}^n$ with finite p-th moment, then there exists $A \in S$, unique up to a scalar multiple, such that

$$|A|^{-1/n} E[|AX|^p]^{1/p} \leq |A'|^{-1/n} E[|A'X|^p]^{1/p} \tag{20}$$

for every $A' \in S$.

Proof: Let $S' \subset S$ be the subset of matrices whose maximum eigenvalue is exactly 1. This is a bounded set inside the set of all symmetric matrices, with its boundary $\partial S'$ equal to positive semidefinite matrices with maximum eigenvalue 1 and minimum eigenvalue 0. Given $A' \in S'$, let e be an eigenvector of A' with eigenvalue 1. By Lemma 7,

$$\begin{aligned}|A'|^{-1/n} E[|A'X|^p]^{1/p} &\geq |A'|^{-1/n} E[|X \cdot e|^p]^{1/p} \\ &\geq c^{1/p} |A'|^{-1/n}.\end{aligned} \tag{21}$$

Therefore, if A' approaches the boundary $\partial S'$, the left side of (21) grows without bound. Since the left side of (21) is a continuous function on S', the existence of a minimum follows.

Let $A \in S$ be such a minimum and $Y = AX$. For each $B \in S$, let $(BA)^s = [(BA)^t(BA)]^{1/2}$ and observe that $|(BA)x| = |(BA)^s x|$, for each $x \in \mathbb{R}^n$. Therefore,

$$\begin{aligned}|B|^{-1/n} E[|BY|^p]^{1/p} &= |A|^{1/n} |BA|^{-1/n} E[|(BA)X|^p]^{1/p} \\ &= |A|^{1/n} |(BA)^s|^{-1/n} E[|(BA)^s X|^p]^{1/p} \\ &\geq |A|^{1/n} |A|^{-1/n} E[|AX|^p]^{1/p} \\ &= E[|Y|^p]^{1/p},\end{aligned} \tag{22}$$

with equality holding if and only if equality holds for (20) with $A' = (BA)^s$. Setting $B = I + tB'$ for $B' \in S$, we get

$$|I + tB'|^{-1/n} E[|(I + tB')Y|^p]^{1/p} \geq E[|Y|^p]^{1/p},$$

for each t near 0. It follows that

$$\left.\frac{d}{dt}\right|_{t=0} |I + tB'|^{-1/n} E[|(I + tB')Y|^p]^{1/p} = 0,$$

for each $B' \in S$. A straightforward computation shows that this holds only if

$$\frac{1}{n} E[|Y|^p] I = E[Y \otimes Y |Y|^{p-2}]. \tag{23}$$

Applying Lemma 6 to

$$d\mu(x) = \frac{|x|^p\, dm_Y(x)}{E[|Y|^p]},$$

implies that equality holds for (22) only if $B = aI$ for some $a \in (0, \infty)$. This, in turn, implies that equality holds for (20) only if $A' = aA$. ■

Theorem 1 follows from Theorem 8 by rescaling A so that $E[|Y|^p] = n$. Equation (3) follows by substituting $Y = AX$ into (23).

REFERENCES

[1] E. Lutwak, D. Yang, and G. Zhang, "Cramer-Rao and moment-entropy inequalities for Renyi entropy and generalized Fisher information," *IEEE Trans. Inform. Theory*, vol. 51, pp. 473–478, 2005.

[2] E. Arikan, "An inequality on guessing and its application to sequential decoding," *IEEE Trans. Inform. Theory*, vol. 42, pp. 99–105, 1996.

[3] O. G. Guleryuz, E. Lutwak, D. Yang, and G. Zhang, "Information-theoretic inequalities for contoured probability distributions," *IEEE Trans. Inform. Theory*, vol. 48, pp. 2377–2383, 2002.

[4] J. A. Costa, A. O. Hero, and C. Vignat, "A characterization of the multivariate distributions maximizing Renyi entropy," in *Proceedings of 2002 IEEE International Symposium on Information Theory*, 2002, p. 263.

[5] O. Johnson and C. Vignat, "Some results concerning maximum Rényi entropy distributions," 2006, preprint.

[6] E. Lutwak, D. Yang, and G. Zhang, "The Cramer-Rao inequality for star bodies," *Duke Math. J.*, vol. 112, pp. 59–81, 2002.

[7] ——, "Moment–entropy inequalities," *Annals of Probability*, vol. 32, pp. 757–774, 2004.

[8] ——, "L^p John ellipsoids," *Proc. London Math. Soc.*, vol. 90, pp. 497–520, 2005.

[9] ——, "Optimal Sobolev norms and the L^p Minkowski problem," *Int. Math. Res. Not.*, pp. 62 987, 1–21, 2006.

[10] J. Bastero and M. Romance, "Positions of convex bodies associated to extremal problems and isotropic measures," *Adv. Math.*, vol. 184, no. 1, pp. 64–88, 2004.

[11] S. Kullback and R. A. Leibler, "On information and sufficiency," *Ann. Math. Statistics*, vol. 22, pp. 79–86, 1951.

[12] I. Csiszár, "Information-type measures of difference of probability distributions and indirect observations," *Studia Sci. Math. Hungar.*, vol. 2, pp. 299–318, 1967.

[13] S.-i. Amari, *Differential-geometrical methods in statistics*, ser. Lecture Notes in Statistics. New York: Springer-Verlag, 1985, vol. 28.

[14] T. M. Cover and J. A. Thomas, *Elements of information theory*. New York: John Wiley & Sons Inc., 1991, a Wiley-Interscience Publication.

Erwin Lutwak Erwin Lutwak is a professor at Polytechnic University, from which he received his B.S. in mathematics when it was the Polytechnic Institute of Brooklyn and his M.S., and Ph.D. degrees in mathematics when it was the Polytechnic Institute of New York.

Deane Yang Deane Yang is a professor at Polytechnic University. He received his B.A. in mathematics and physics from University of Pennsylvania and Ph.D. in mathematics from Harvard University.

Gaoyong Zhang Gaoyong Zhang is a professor at Polytechnic University. He received his B.S. degree in mathematics from Wuhan University of Science and Technology, M.S. degree in mathematics from Wuhan University, Wuhan, China, and Ph.D. degree in mathematics from Temple University, Philadelphia.