On the Volume of the Minkowski Sum of Line Sets and the Entropy-Power Inequality *

Ram Zamir and Meir Feder

Dept. of Electrical Engineering - Systems,
Tel-Aviv University, Tel-Aviv 69978 Israel.
e-mail: zamir@eng.tau.ac.il, meir@eng.tau.ac.il

October 24, 1997

Abstract

We derive a Brunn-Minkowski-type inequality regarding the volume of the Minkowski sum of degenerate sets, namely, line sets. Let $\mathcal{A}_1 \dots \mathcal{A}_n$ be one dimensional sets of unit length, and $\underline{v}_1 \dots \underline{v}_n$ vectors in $\mathcal{R}^d$, $d \le n$. Consider the Minkowski sum of the line sets $\mathcal{A}_i \underline{v}_i \triangleq \{x\underline{v}_i : x \in \mathcal{A}_i\}$ $i = 1, \dots, n$. We show that the volume of this set sum satisfies

$$V_d\Big(\sum_{i=1}^{n} \mathcal{A}_i \underline{v}_i\Big) \ge \sum_{j=1}^{\binom{n}{d}} |\det_j(T)| \, ,$$

where T is the $d \times n$ matrix $(\underline{v}_1; \underline{v}_2; \dots; \underline{v}_n)$, and $\det_j(T), j = 1 \dots \binom{n}{d}$ denote the determinants of all $d \times d$ submatrices of T. Equality holds if $d = n$ or if the $\mathcal{A}_i$'s are intervals, where in the latter case $\sum_i \mathcal{A}_i \underline{v}_i$ is a zonotope. This inequality parallels a recently obtained matrix generalization of the entropy-power inequality.

Key Words: Brunn-Minkowski Inequality, Entropy Power Inequality, Zonotope, Matrix Generalization of Information Theoretic Inequalities.

*This research was supported in part by the Wolfson Research Awards administered by the Israel Academy of Science and Humanities. This work was presented in the IEEE-IMS Workshop on Information Theory and Statistics, Alexandria, Virginia, Oct. 1994, and in the International Symposium on Information Theory, Whistler, Canada, Sept. 1995.

I. Motivation and Statement of Result

The Brunn-Minkowski Inequality (BMI) [5], [7], [3, Sec. 16.7] states that for any measurable sets $\mathcal{S}_1, \ldots, \mathcal{S}_n$ in $\mathcal{R}^d$,

$$V_d(\mathcal{S}_1 + \ldots + \mathcal{S}_n)^{1/d} \geq V_d(\mathcal{S}_1)^{1/d} + \ldots + V_d(\mathcal{S}_n)^{1/d}\ , \tag{1}$$

where $V_d(\cdot)$ denotes d-dimensional volume, and $\mathcal{S}_1 + \mathcal{S}_2 \stackrel{\Delta}{=} \{x + y : \ x \in \mathcal{S}_1, y \in \mathcal{S}_2\}$, i.e., the + sign denotes the Minkowski set sum. Since the BMI holds with equality for proportional convex sets, inequality (1) can be rewritten as

$$V_d\Big(\mathcal{S}_1 + \ldots + \mathcal{S}_n\Big) \geq V_d\Big(\widetilde{\mathcal{S}}_1 + \ldots + \widetilde{\mathcal{S}}_n\Big) \tag{2}$$

where $\widetilde{\mathcal{S}}_1 \ldots \widetilde{\mathcal{S}}_n$ are proportional convex sets, e.g., spheres or parallel cubes, with $V_d(\widetilde{\mathcal{S}}_i) = V_d(\mathcal{S}_i), i = 1 \ldots n$. For $d = 1$ the sets $\widetilde{\mathcal{S}}_i$ are of course intervals.

When the $\mathcal{S}_i$'s are degenerate sets in $\mathcal{R}^d$ (i.e., sets whose d-dimensional volume is zero), the BMI gives a trivial lower bound, zero, even though the Minkowski sum $\sum_i \mathcal{S}_i$ may have a positive volume. One such case of interest is when the $\mathcal{S}_i$'s are line sets, i.e., $\mathcal{S}_i = \mathcal{A}_i \underline{v}_i \stackrel{\Delta}{=} \{x\underline{v}_i : \ x \in \mathcal{A}_i\}$, where the $\mathcal{A}_i$'s are sets in $\mathcal{R}^1$ and the $\underline{v}_i$'s are vectors in $\mathcal{R}^d$. If $n \geq d > 1$ and $\underline{v}_1, \ldots, \underline{v}_n$ span $\mathcal{R}^d$, then $V_d(\mathcal{S}_i) = 0 \ \ \forall i$ but the Minkowski set sum of the $\mathcal{S}_i$'s is *not* degenerate in $\mathcal{R}^d$. For instance, if the $\mathcal{A}_i$'s are unit intervals, then $\sum_{i=1}^n \mathcal{S}_i$ is a parallelepiped for $n = d$, and a *zonotope* for $n > d$. (This is in fact the definition of a zonotope ; see e.g. [6, 8]).

The Minkowski sum of line sets above may be written as a linear transformation of a product set, a format which turns out to be both useful and inspiring. Let $T = (\underline{v}_1; \underline{v}_2; \ldots; \underline{v}_n)$ be the $d \times n$ matrix whose columns are the vectors $\underline{v}_1, \ldots, \underline{v}_n$, and let $\mathcal{A} = (\mathcal{A}_1 \times \ldots \times \mathcal{A}_n)$. Then

$$\sum_i \mathcal{S}_i = \sum_i \mathcal{A}_i \underline{v}_i = T \cdot \mathcal{A} \stackrel{\Delta}{=} \Big\{T\underline{x} : \ x_i \in \mathcal{A}_i\Big\}. \tag{3}$$

Our purpose is to estimate the volume of this set. If the $\mathcal{A}_i$'s are *unit intervals*, then this set is a zonotope, and we know its volume exactly for any d, [8]:

$$V_d\Big(T\mathcal{A}\Big) = \sum_{j=1}^{\binom{n}{d}} |\det_j(T)| \ , \tag{4}$$

where $\det_j(T), j = 1 \ldots \binom{n}{d}$ denote the determinants of all $d \times d$ submatrices of T, each composed of d out of the n columns on T. On the other hand, for $d = 1$ and $\mathcal{A}_i$ *arbitrary sets* of unit length the BMI gives

$$V_1\Big(\sum_i \mathcal{A}_i v_i\Big) \geq \sum_i V_1(\mathcal{A}_i v_i) = \sum_i |v_i|. \tag{5}$$

What can we say in the general ($d > 1$, $\mathcal{A}_i$ arbitrary set) case?

The problem of estimating the volume of $T(\mathcal{A}_1 \times \ldots \times \mathcal{A}_n)$ strongly resembles the question of estimating the differential entropy of the random vector $T\underline{X}$, where $\underline{X} = (X_1, \ldots, X_n)^t$ is a vector whose components are statistically independent. Suppose the X_i's have unit entropy powers, i.e., $P(X_i) \triangleq 2^{2h(X_i)}/2\pi e = 1, \ i = 1 \ldots n$, where $h(X) = -\int f(x) \log f(x) dx$ denotes the differential entropy of X. If the X_i's are Gaussian, then $T\underline{X}$ is a Gaussian vector whose entropy is given by [3]

$$h(T\underline{X}) = \frac{d}{2} \log\Big(2\pi e \cdot \det(T \cdot T^t)^{1/d}\Big) \ . \tag{6}$$

On the other hand, for $d = 1$ and X_i arbitrary, the Entropy Power Inequality (EPI) implies, [3],

$$P\Big(\sum_i X_i v_i\Big) \geq \sum_i P(X_i v_i) = \sum_i v_i^2 \ . \tag{7}$$

However, like in the case of the Minkowski sum of line sets, if one tries to apply the EPI to $T\underline{X} = \sum_i X_i \underline{v}_i$ for $d > 1$, then one gets a trivial lower bound

$$P\left(\sum_i X_i \underline{v}_i\right) \geq \sum_i P(X_i \underline{v}_i) = 0,$$

since $X_i \underline{v}_i$ is a degenerate random vector in $\mathcal{R}^d$, i.e., a vector whose entropy-power

$$P(X_i \underline{v}_i) \triangleq \frac{2^{\frac{2}{d} h(X_i \underline{v}_i)}}{2\pi e}$$

is zero.

The formal resemblance between cases (6)-(7) and cases (4)-(5), as well as the duality of the degenerate cases, is striking. This resemblance goes back to Costa and Cover [2], who pointed out the analogy between the BMI (1) and the EPI for statistically independent random vectors

$$P\left(\sum_i \underline{X}_i\right) \geq \sum_i P(\underline{X}_i). \tag{8}$$

Costa and Cover also proposed a form of the EPI which resembles (2)

$$h\left(\sum_i \underline{X}_i\right) \geq h\left(\sum_i \underline{\widetilde{X}}_i\right), \tag{9}$$

where $\{\underline{\widetilde{X}}_i\}$ are statistically independent Gaussian vectors with proportional covariances (e.g., white Gaussians), having the same entropies as $\{\underline{X}_i\}$, respectively. Dembo *et al* [5] showed that indeed the BMI and the EPI are two extreme points of the Young inequality.

In [9] the authors derived a matrix version of the EPI, which when applied to the random vector $T\underline{X}$ above yields the lower bound

$$h(T\underline{X}) \geq h(T\underline{\widetilde{X}}) = \frac{d}{2}\log\left(2\pi e \cdot \det(TT^t)^{1/d}\right), \tag{10}$$

where $\underline{\widetilde{X}} = (\widetilde{X}_1, \ldots, \widetilde{X}_n)$ is a vector of independent Gaussian variables with unit variances (= entropy powers). This inequality and the analogy discussed above motivate us to prove the following lower bound for the volume of the Minkowski sum of (degenerate) line sets.

Theorem 1 *Let $\mathcal{A}_1, \ldots, \mathcal{A}_n$ be measurable sets in $\mathcal{R}^1$ having unit length, and let $\underline{v}_1, \ldots, \underline{v}_n$ be vectors in $\mathcal{R}^d$. Then, the volume of the Minkowski sum $\sum_i \mathcal{A}_i \underline{v}_i$ is minimized when the $\mathcal{A}_i$'s are unit intervals, i.e., when $\sum_i \mathcal{A}_i \underline{v}_i$ is a* zonotope. *The lower bound for the volume is given by*

$$V_d\Big(\sum_i \mathcal{A}_i \underline{v}_i\Big) \geq V_d\Big(\sum_i \widetilde{A}_i \underline{v}_i\Big) = \sum_{j=1}^{\binom{n}{d}} |det_j(T)|, \tag{11}$$

where $\widetilde{\mathcal{A}}_1, \ldots, \widetilde{\mathcal{A}}_n$ are unit intervals, T is the $d \times n$ matrix $(\underline{v}_1; \underline{v}_2; \ldots; \underline{v}_n)$, and $\{det_j(T)\}$ are defined in (4).

Remarks:

1. In analogy with (10), Theorem 1 can be viewed as a *matrix generalization of the BMI (2)* for one dimensional sets. Namely

$$V_d\Big(T(\mathcal{A}_1 \times \ldots \times \mathcal{A}_n)\Big) \geq V_d\Big(T(\widetilde{\mathcal{A}_1} \times \ldots \times \widetilde{\mathcal{A}_n})\Big) \tag{12}$$

where the definition of a linear transformation of a product set, $T(\mathcal{A}_1 \times \ldots \times \mathcal{A}_n)$, appears in (3).

2. If $\mathcal{A}_1, \ldots, \mathcal{A}_n$ have arbitrary lengths $\ell_1, \ldots, \ell_n > 0$, then we may write $\sum_i \mathcal{A}_i \underline{v}_i = \sum_i \mathcal{A}'_i \underline{v}'_i$ where $\mathcal{A}'_i \stackrel{\Delta}{=} \frac{1}{\ell_i}\mathcal{A}_i$ is a scaled version of $\mathcal{A}_i$ having *unit* length, and $\underline{v}'_i \stackrel{\Delta}{=} \ell_i \underline{v}_i$. Substituting in Theorem 1 we thus obtain

$$V_d\Big(\sum_i \mathcal{A}_i \underline{v}_i\Big) \geq \sum_{j=1}^{\binom{n}{d}} |\det_j(T \cdot \mathrm{diag}[\ell_1 \ldots \ell_n])| \;, \tag{13}$$

where diag $[\ell_1 \ldots \ell_n]$ denotes the $n \times n$ diagonal matrix whose diagonal elements are $\ell_1, \ldots, \ell_n$.

3. For $d = 1$ and the row matrix $T = [1, \ldots, 1]$, (11) and (13) reduce to

$$V_1\left(\mathcal{A}_1 + \ldots + \mathcal{A}_n\right) \geq V_1(\mathcal{A}_1) + \ldots + V_1(\mathcal{A}_n) \;, \tag{14}$$

i.e., to the regular BMI (1) for one dimensional sets.

4. We have the following sufficient condition for equality in (11). Suppose the vectors $\underline{v}_1, \ldots, \underline{v}_n$ span $\mathcal{R}^d$ (otherwise $\sum \mathcal{A}_i \underline{v}_i$ is a degenerate set whose volume is zero), and that non of these vectors equals zero (otherwise omit the zero vectors and their associated sets without affecting the volume). Let $\mathcal{I}_R(T) \subseteq \{1 \ldots n\}$ denote the subset of indices j such that x_j is uniquely determined by $T\underline{x}$. We show in the Appendix that

Equality in (11) holds if the set $\mathcal{A}_j$ is an interval for all $j \notin \mathcal{I}_R(T)$.

In particular, equality holds if all sets $\mathcal{A}_1 \ldots \mathcal{A}_n$ are intervals, or if $n = d$. In the latter case we have $\mathcal{I}_R(T) = \{1, \ldots, n\}$, and the volume of $\sum_i \mathcal{A}_i \underline{v}_i$ is calculated via simple laws of integration

$$V_n\Big(\sum_i \mathcal{A}_i \underline{v}_i\Big) = V_n\Big(T(\mathcal{A}_1 \times \ldots \times \mathcal{A}_n)\Big) = |\det(T)|, \tag{15}$$

independently of the specific shape of the sets $\mathcal{A}_1, \ldots, \mathcal{A}_n$.

The proof of Theorem 1 is given in Section II. In Section III we summarize the various analogies between the BMI and the EPI in view of this paper. We also point out an interesting inequality for determinants that follow from Theorem 1, and consider a possible extension of Theorem 1 to linear transformation of d'-dimensional sets, $d' > 1$.

II. Proof of Theorem

We first state and prove the following lemma which plays a key role in proving the Theorem.

Lemma 1 (BMI for set + line-set) *Let $\mathcal{A}$ be a d-dimensional set whose projection on the first $d-1$ coordinates is denoted $\mathcal{A}_p$, i.e.,*

$$\mathcal{A}_p = \Big\{\underline{x} \in \mathcal{R}^{d-1} : (\underline{x}, y) \in \mathcal{A} \textit{ for some } y \in \mathcal{R}^1\Big\} . \tag{16}$$

Let $\mathcal{B}$ be a set in $\mathcal{R}^1$, and define the (degenerate) d-dimensional expansion of $\mathcal{B}$ as

$$\mathcal{B}^e = \left\{ (\overbrace{0, \ldots, 0}^{d-1}, y) : \ y \in \mathcal{B} \right\} . \tag{17}$$

Then,

$$V_d(\mathcal{A} + \mathcal{B}^e) \geq V_d(\mathcal{A}) + V_1(\mathcal{B}) \cdot V_{d-1}(\mathcal{A}_p) . \tag{18}$$

Equality in (18) holds if $\mathcal{B}$ is an interval, and for every $\underline{x} \in \mathcal{A}_p$ the intersection

$$\{y : \ (\underline{x}, y) \in \mathcal{A}\} \ , \tag{19}$$

is an interval.

Proof: By the definition of the expanded set (17), we have

$$\mathcal{A} + \mathcal{B}^e = \{\underline{u} + \underline{v} : \ \underline{u} \in \mathcal{A}, \underline{v} \in \mathcal{B}^e\} = \{(\underline{x}, y + z) : \ (\underline{x}, y) \in \mathcal{A}, z \in \mathcal{B}\} \tag{20}$$

(where vector dimensions can be understood unambiguously from their associated sets). Regard $\{y+z:\ (\underline{x},y)\in\mathcal{A}, z\in\mathcal{B}\}$ as a one-dimensional set parameterized by $\underline{x}\in\mathcal{A}_p$. Using the chain rule for a multiple integral,

$$V_d(\mathcal{A}+\mathcal{B}^e)=\int_{\mathcal{A}_p} V_1\Big(\{y+z:\ (\underline{x},y)\in\mathcal{A}, z\in\mathcal{B}\}\Big)d\underline{x}\ . \tag{21}$$

Applying the Brunn-Minkowski inequality (1) (with $d=1$) to the integrand of (21), we have

$$V_1\Big(\{y+z:\ (\underline{x},y)\in\mathcal{A}, z\in\mathcal{B}\}\Big)\geq V_1\Big(\{y:\ (\underline{x},y)\in\mathcal{A}\}\Big)+V_1(\mathcal{B})\ ,\quad \forall\underline{x}\in\mathcal{A}_p\ . \tag{22}$$

Substituting (22) in (21), and using

$$V_d(\mathcal{A})=\int_{\mathcal{A}_p} V_1\Big(\{y:\ (\underline{x},y)\in\mathcal{A}\}\Big)d\underline{x}$$

and

$$V_{d-1}(\mathcal{A}_p)=\int_{\mathcal{A}_p} d\underline{x}$$

(which are special cases of (21)), results in the desired inequality.

The condition for equality follows since if $\mathcal{B}$ is an interval, and the intersection $\{y:\ (\underline{x},y)\in\mathcal{A}\}$ is an interval, then (22) holds with equality. $\Box$

We are now in a position to prove the theorem. We want to prove that for any $m\times n$ matrix T,

$$V_m\Big(T(\mathcal{A}_1\times\ldots\times\mathcal{A}_n)\Big)\geq\sum_{j=1}^{\binom{n}{m}}|\det{}_j(T)| \tag{23}$$

for sets $\mathcal{A}_1\ldots\mathcal{A}_n$ having unit length, and

$$V_m\Big(T(\widetilde{\mathcal{A}}_1\times\ldots\times\widetilde{\mathcal{A}}_n)\Big)=\sum_{j=1}^{\binom{n}{m}}|\det{}_j(T)| \tag{24}$$

for $\widetilde{\mathcal{A}}_1,\ldots,\widetilde{\mathcal{A}}_n$ unit intervals. (Identity (24) follows from formula (4) for the volume of a zonotope but we include it here for completeness.) Note that in the proof we use the letter m (instead of d used in the Theorem) for the number of rows of T (= the dimension of the vectors $\underline{v}_i$). To make the proof more transparent, we divide it into a few steps.

1) The induction structure: The proof of (23) and (24) is by *double induction* over m and n. The *induction boundary conditions* are the line $m = 1$ (any n) and the line $m = n$ in the plain $(m, n) \in \mathcal{N}^2$. In the case $m = 1$ (23) and (24) hold by the regular Brunn-Minkowski inequality and its equality conditions since T is a row matrix; see (14). In the case $m = n$ the matrix T is invertible so both (23) and (24) hold with equality; see (15). We show below that if (23) holds for any $(m-1) \times (n-1)$ and $m \times (n-1)$ matrices, then it also holds for any $m \times n$ matrix. We do the same with respect to (24). This is the *induction step*. Figure 1 shows a path in the plain $\mathcal{N}^2$ from the boundary lines to an arbitrary point (m, n), which is followed by the induction steps to prove the theorem for any $m \times n$ matrix. Since m and n are arbitrary, Theorem 1 holds for any matrix provided that the induction step is proved.

2) Preliminaries: To prove the induction step, we need to make some matrix manipulations used in Gaussian elimination on the matrix $T = \{t_{i,j}\}$. Suppose the last column of T is not zero. Otherwise, i.e. if $t_{i,n} = 0$ for all i, then $T(\mathcal{A}_1 \times \ldots \times \mathcal{A}_n)$ does not depend on $\mathcal{A}_n$ and T is actually an $m \times (n-1)$ matrix for which the inequality holds by the induction assumption. Now if $t_{m,n} = 0$, permute a pair of rows of T so that after permutation $t_{m,n} \neq 0$. This permutation, if needed, does not affect the volume of $T(\mathcal{A}_1 \times \ldots \times \mathcal{A}_n)$.

The next step is to use row operations and make the first $(m-1)$ elements of the last column be zero. This is possible since we assured above that $t_{m,n} \neq 0$. Denote by $T' = \Omega T$ the matrix after the row operations, where the $m \times m$ matrix Ω has the form

$$\Omega = \begin{pmatrix} 1 & & & \omega_1 \\ & 1 & 0 & \omega_2 \\ & & . & . \\ & 0 & & . \\ & & & 1 \end{pmatrix} , \qquad \det(\Omega) = 1 \,. \tag{25}$$

Since $\det(\Omega) = 1$, the row operations do not change the volume, i.e., for any set $\mathcal{A}$ in $\mathcal{R}^n$

$$V_m(T\mathcal{A}) = V_m(\Omega^{-1}T'\mathcal{A}) = V_m(T'\mathcal{A})/|\det(\Omega)| = V_m(T'\mathcal{A}) \ , \tag{26}$$

where we used the formula for integration after a linear transformation as in (15).

Define some sub-matrices of T', as follows

$$T' = \left(\begin{array}{c|c} & 0 \\ & \cdot \\ S & \cdot \\ & 0 \\ \hline P & t_{m,n} \end{array}\right) = \left(\begin{array}{c|c} & \\ & \\ U & Q \\ & \\ & \\ & \end{array}\right)$$

where $P = (t_{m,1}, \ldots, t_{m,n-1})$ is the last row of T' (= the last row of T) without the last term, $Q = [0, \ldots, 0, t_{mn}]^t$ is the last column of T', U is obtained by dropping Q from T' ($\dim U = m \times (n-1)$), and S is obtained by dropping P from U ($\dim S = (m-1) \times (n-1)$).

3) Proof of induction step for (23): The induction assumption applies to the sub-matrices U and S, so we have

$$V_m\Big(U(\mathcal{A}_1 \times \ldots \times \mathcal{A}_{n-1})\Big) \geq \sum_{j=1}^{\binom{n-1}{m}} |\det_j(U)| \ , \tag{27}$$

and

$$V_{(m-1)}\Big(S(\mathcal{A}_1 \times \ldots \times \mathcal{A}_{n-1})\Big) \geq \sum_{j=1}^{\binom{n-1}{m-1}} |\det_j(S)| \ . \tag{28}$$

Since $T' = (U; Q)$ and $Q = [0, \ldots, 0, t_{m,n}]^t$, we have in analogy with (17) and (20)

$$T'(\mathcal{A}_1 \times \ldots \times \mathcal{A}_n) \quad = \quad U(\mathcal{A}_1 \times \ldots \times \mathcal{A}_{n-1}) + t_{m,n}\mathcal{A}_n^e \ , \tag{29}$$

where $\mathcal{A}_n^e$ denotes the m-dimensional expansion of the set $\mathcal{A}_n$ as in (17). Also, since

$$U = \begin{pmatrix} S \\ - \\ P \end{pmatrix} , \tag{30}$$

the projection of $U(\mathcal{A}_1 \times \ldots \times \mathcal{A}_{n-1})$ on the first $(m-1)$ coordinates is the set $S(\mathcal{A}_1 \times \ldots \times \mathcal{A}_{n-1})$. Using this fact, we apply Lemma 1 to the set sum in (29), and obtain

$$\begin{aligned} & V_m\Big(U(\mathcal{A}_1 \times \ldots \times \mathcal{A}_{n-1}) \quad + \quad t_{m,n}\mathcal{A}_n^e\Big) \geq \\ & V_m\Big(U(\mathcal{A}_1 \times \ldots \times \mathcal{A}_{n-1})\Big) \quad + \quad V_1\Big(t_{m,n}\mathcal{A}_n\Big) \cdot V_{m-1}\Big(S(\mathcal{A}_1 \times \ldots \times \mathcal{A}_{n-1})\Big)\,. \end{aligned} \tag{31}$$

Substituting the induction assumptions (27) and (28), and $V_1(t_{m,n}\mathcal{A}_n) = |t_{m,n}|$ in (31), we have (see also (26) and (29))

$$V_m\Big(T'(\mathcal{A}_1 \times \ldots \times \mathcal{A}_n)\Big) \geq \sum_{j=1}^{\binom{n-1}{m}} |\det_j(U)| + |t_{m,n}| \cdot \sum_{j=1}^{\binom{n-1}{m-1}} |\det_j(S)|\,. \tag{32}$$

In view of the mutual structure of the matrices T', U and S, it is easy to verify that their sub-determinant sets satisfy the relation

$$\begin{aligned} \left\{\det_j(U),\ j = 1 \ldots \binom{n-1}{m}\right\} \ \cup \ & \left\{t_{m,n} \cdot \det_j(S),\ j = 1 \ldots \binom{n-1}{m-1}\right\} \\ = \ & \left\{\det_j(T'),\ j = 1 \ldots \binom{n}{m}\right\}\,. \end{aligned} \tag{33}$$

(Note that $\binom{n-1}{m} + \binom{n-1}{m-1} = \binom{n}{m}$.) This implies that the lower bound in (32) equals

$$\sum_{j=1}^{\binom{n}{m}} |\det_j(T')| = \sum_{j=1}^{\binom{n}{m}} |\det_j(T)|\,, \tag{34}$$

where the last equality follows since by the definition of a sub-determinant $\det_j(T') = \det_j(\Omega T) = \det(\Omega) \cdot \det_j(T) = \det_j(T)$. This proves (23).

4) Proof of induction step for (24): By the induction assumption, Ineq. (27) and (28) hold with equality for the unit intervals $\widetilde{\mathcal{A}}_1, \ldots, \widetilde{\mathcal{A}}_{n-1}$. Note that the only other inequality in the induction step (Step 3 above) occurs in (31), where we used Lemma 1. So in order to assert (24), it is enough to show that intervals satisfy the condition for equality in Lemma 1 when applied to (31). Now, since T is a linear transformation and $\widetilde{\mathcal{A}}_1, \ldots, \widetilde{\mathcal{A}}_n$ are convex sets, $T(\widetilde{\mathcal{A}}_1 \times \ldots \times \widetilde{\mathcal{A}}_n)$ is also a convex set. Thus, the intersection of $T(\widetilde{\mathcal{A}}_1 \times \ldots \times \widetilde{\mathcal{A}}_n)$ with any line is an interval. In particular, for every $\underline{x} \in S(\widetilde{\mathcal{A}}_1 \times \ldots \times \widetilde{\mathcal{A}}_{n-1})$, the intersection

$$\Big\{y:\ (\underline{x}, y) \in T(\widetilde{\mathcal{A}}_1 \times \ldots \times \widetilde{\mathcal{A}}_n)\Big\} \tag{35}$$

is an interval. Combining with the fact that $t_{m,n}\widetilde{\mathcal{A}}_n$ is also an interval, we see from Lemma 1 that the condition for equality in (31) is satisfied. This shows (24), and the theorem is proved. □

III. Discussion

A. Analogies Between the BMI and the EPI

The analogy between the forms (2) and (9) of the BMI and the EPI, and the fact that *spheres* satisfy the BMI with equality, led Costa and Cover [2] to draw the connection between the role of the Gaussian distribution in the EPI, and the role of the sphere in the BMI. These inequalities and the analogies discussed above were the motivation for the work of this paper. The fact that intervals, i.e., one-dimensional spheres, satisfy the new inequality (11) with equality follows and further supports the line suggested by Costa and Cover. The possible extension of this matrix form of the BMI to d'-dimensional sets is discussed in Section III-C.

In the table below we summarize the various analogies between the BMI and the EPI:

BMI	EPI
measurable sets	random variables
Cartesian product	statistical independence
Minkowski sum	real sum
length V_1 (effective radius $V_d^{1/d}$)	entropy power P
interval (sphere / cube)	Gaussian
Lemma 1	conditional-EPI [1]

As Dembo *et al* show [5], both the EPI and the BMI can be derived as the two extreme points of a single inequality, namely, Young's inequality. Indeed, it remains an open question weather the matrix form of the EPI (10) and the matrix form of the BMI (11)-(12) are in fact the two extreme points of an (unknown) matrix generalization of the Young inequality.

B. An Inequality for Determinants

As [5] and [3] show, some interesting inequalities for determinants follow straightforwardly from the entropy power inequality. In this spirit, we give below an example for an interesting implication of the new matrix form of the Brunn-Minkowski inequality .

Let $\mathcal{A}_1 \ldots \mathcal{A}_n$ and $\mathcal{B}_1 \ldots \mathcal{B}_n$ be sets with unit length in $\mathcal{R}^1$. Let T_1 and T_2 be $n \times n$ matrices, and consider the volume of the set sum $T_1(\mathcal{A}_1 \times \ldots \times \mathcal{A}_n) + T_2(\mathcal{B}_1 \times \ldots \times \mathcal{B}_n)$. A direct application of the Brunn-Minkowski inequality (1) to this set sum gives

$$\begin{aligned} V_n\Big[T_1(\mathcal{A}_1 \times \ldots \times \mathcal{A}_n) &+ T_2(\mathcal{B}_1 \times \ldots \times \mathcal{B}_n)\Big]^{1/n} \\ &\geq V_n\Big[T_1(\mathcal{A}_1 \times \ldots \times \mathcal{A}_n)\Big]^{1/n} + V_n\Big[T_2(\mathcal{B}_1 \times \ldots \times \mathcal{B}_n)\Big]^{1/n} \\ &= |\det(T_1)|^{1/n} + |\det(T_2)|^{1/n} \ . \end{aligned} \tag{36}$$

On the other hand, viewing the set sum as the transformation of the set $(\mathcal{A}_1 \times \ldots \times \mathcal{A}_n \times \mathcal{B}_1 \times \ldots \times \mathcal{B}_n)$ by the $n \times 2n$ matrix $T = (T_1; T_2)$, Theorem 1 gives

$$V_n\Big[(T_1;T_2)(\mathcal{A}_1 \times \ldots \times \mathcal{A}_n \times \mathcal{B}_1 \times \ldots \times \mathcal{B}_n)\Big] \geq \sum_{j=1}^{\binom{2n}{n}} |\mathrm{det}_j(T_1;T_2)| \ . \tag{37}$$

But, by Theorem 1, for $\mathcal{A}_1 \ldots \mathcal{B}_n$ unit intervals (37) becomes an equality, while (36) remains an inequality unless T_1 and T_2 are proportional. Thus, we conclude that (37) is a tighter lower bound than (36), and obtain the inequality

$$\Big(|\det(T_1)|^{1/n} + |\det(T_2)|^{1/n}\Big)^n \leq \sum_{j=1}^{\binom{2n}{n}} |\mathrm{det}_j(T_1;T_2)| \ , \tag{38}$$

with equality if T_1 and T_2 are proportional. Note that for $T_2 = \lambda T_1$, both sides of (38) are equal to $(1+|\lambda|)^n |T_1|$.

C. A Matrix Generalization of the d-dimensional BMI

In an earlier version of this paper we made the following more general statement of Theorem 1. Let $T = \{t_{i,j}\}$ be an $m \times n$ matrix, and d an integer. Define the $(md) \times (nd)$ expanded matrix

$$T_e = \left[\begin{array}{c|c|c} t_{1,1} I_d & \ldots & t_{1,n} I_d \\ - & - & - \\ \vdots & \vdots & \vdots \\ - & - & - \\ t_{m,1} I_d & \ldots & t_{m,n} I_d \end{array} \right], \tag{39}$$

where I_d denotes the $d \times d$ identity matrix. Let $\mathcal{A}_1, \ldots, \mathcal{A}_n$ be measurable sets in $\mathcal{R}^d$, and let $\mathcal{A} = \mathcal{A}_1 \times \ldots \times \mathcal{A}_n$. Define the term "canonical cubes" as cubes whose edges parallel the axes. Then, for fixed volumes of $\mathcal{A}_1, \ldots, \mathcal{A}_n$, the volume of the md-dimensional set $T_e \mathcal{A}$ is minimized when $\mathcal{A}_1, \ldots, \mathcal{A}_n$ are canonical cubes, i.e.,

$$V_{md}\left(T_e \mathcal{A}\right) \geq V_{md}\left(T_e \widetilde{\mathcal{A}}\right) = \left(\sum_{j=1}^{\binom{n}{m}} \left| \det_j (T \cdot \mathrm{diag}[r_1 \ldots r_n]) \right| \right)^d, \tag{40}$$

where $\widetilde{\mathcal{A}} = \widetilde{\mathcal{A}}_1 \times \ldots \times \widetilde{\mathcal{A}}_n$ is the product set composed of d-dimensional canonical cubes $\widetilde{\mathcal{A}}_1, \ldots, \widetilde{\mathcal{A}}_n$ with volumes $V_d(\mathcal{A}_1), \ldots, V_d(\mathcal{A}_n)$, respectively, and $r_i \triangleq V_d(\mathcal{A}_i)^{1/d}$ denotes the "effective radius" of $\mathcal{A}_i$.

Note that for $m = 1$ and $T = [1, \ldots, 1]$, this inequality reduces to the regular d-dimensional form of the BMI (2), while for $d = 1$ and any m this inequality reduces to Theorem 1. Thus, (40) could have been a full matrix generalization of the BMI. Unfortunately the proof we gave for this more general inequality was wrong, and was based on an un-correct d-dimensional extension of Lemma 1. The general inequality (40) together with the false lemma appeared in the ISIT-95 proceedings [10], and the error in the proof was detected later by Dembo [4]. At the current moment it remains an open question weather the d-dimensional form of the matrix BMI in (40) is true or false. If true, it suggests that the canonical cube and not the sphere is the true analog in the BMI of the Gaussian distribution in the EPI.

Acknowledgments

The authors are indebted to Amir Dembo for pointing out a critical error in the early version of this paper, and for many helpful suggestions and comments. Thanks are due also to Noga Alon for pointing out the relation to zonotopes, and to Tamas Linder, Simon Litsyn and Nadav Shulman for helpful discussions.

Appendix: Sufficient Condition for Equality (Remark 4 after Theorem 1)

Let r be the size of the index set $\mathcal{I}_R(T)$. We can re-order the columns of T and use row operations, to get the $(d \times n)$ matrix

$$T'' = \begin{pmatrix} R & | & 0 \\ - & - & - \\ 0 & | & W \end{pmatrix},$$

where R is a non-singular $r \times r$ matrix, and W is a $(d-r) \times (n-r)$ matrix. We then correspondingly re-order the sets $\mathcal{A}_1, \ldots, \mathcal{A}_n$, so now the first r sets are the ones whose indices belonged to $\mathcal{I}_R(T)$, and the next $n-r$ sets are the ones whose indices belonged to $\{1 \ldots n\} \setminus \mathcal{I}_R(T)$. Denote the re-ordered sets $\mathcal{A}''_1, \ldots, \mathcal{A}''_n$, and let $\mathcal{A}'' = \mathcal{A}''_1 \times \ldots \times \mathcal{A}''_n$. Since re-ordering and row operations on T do not change the volume of $T(\mathcal{A}_1 \times \ldots \times \mathcal{A}_n)$ (see the discussion in Section II, Step 2), we have

$$V_d(T\mathcal{A}) = V_d(T''\mathcal{A}'').$$

By the structure of the matrix T'', the set $T''\mathcal{A}''$ has a product form

$$T''(\mathcal{A}''_1 \times \ldots \times \mathcal{A}''_n) = \left(R(\mathcal{A}''_1 \times \ldots \times \mathcal{A}''_r)\right) \;\times\; \left(W(\mathcal{A}''_{r+1} \times \ldots \times \mathcal{A}''_n)\right).$$

Thus $V_d(T\mathcal{A}) = V_r\left(R(\mathcal{A}''_1 \times \ldots \times \mathcal{A}''_r)\right) \cdot V_{d-r}\left(W(\mathcal{A}''_{r+1} \times \ldots \times \mathcal{A}''_n)\right)$. Since R is a square matrix, the value of the first volume is independent of the specific shape of the sets $\mathcal{A}''_1 \ldots \mathcal{A}''_r$; see (15). We conclude that it is sufficient for equality in Theorem 1 that the sets in the second multiplicand, $\mathcal{A}''_{r+1}, \ldots, \mathcal{A}''_n$, corresponding to the index set $\{1 \ldots n\} \setminus \mathcal{I}_R(T)$, are intervals.

References

[1] N. M. Blachman. The convolution inequality for entropy powers. *IEEE Trans. Information Theory*, IT-11:267–271, 1965.

[2] M.H.M. Costa and T. M. Cover. On the similarity of the entropy power inequality and the Brunn-Minkowski inequality. *IEEE Trans. Information Theory*, IT-30:837–839, Nov. 1984.

[3] T. M. Cover and J. A. Thomas. *Elements of Information Theory.* Wiley, New York, 1991.

[4] A. Dembo. Private communication. 1997.

[5] A. Dembo, T.M.Cover, and J.A.Thomas. Information theoretic inequalities. *IEEE Trans. Information Theory*, IT-37:1501–1518, Nov. 1991.

[6] David Eppstein. Zonohedra and zonotopes. *http://www.ics.uci.edu/ eppstein/junkyard/ ukraine/ukraine.html.*

[7] S. D. Gupta. Brunn-Minkowski inequality and its aftermath. *Journal of multivariate analysis analysis*, 10:296–318, 1980.

[8] G.C. Shephard. Combinatorial properties of associated zonotopes. *Canad. J. Math.*, XXVI No. 2:302–321, 1974.

[9] R. Zamir and M. Feder. A generalization of the Entropy Power Inequality with applications. *IEEE Trans. Information Theory*, IT-39:1723–1727, Sept. 1993.

[10] R. Zamir and M. Feder. A matrix form of the Brunn-Minkowski inequality. In *Proceedings of ISIT-95, Vancouver, Canada*, page 71, Sept. 1995.

Figure 1: Path from boundary lines followed by induction steps.