

THE BRASCAMP–LIEB POLYHEDRON

STEFÁN INGI VALDIMARSSON

ABSTRACT. A set of necessary and sufficient conditions for the Brascamp–Lieb inequality to hold has recently been found by Bennett, Carbery, Christ and Tao. We present an analysis of these conditions. This analysis allows us to give a concise description of the set where the inequality holds in the case where each of the linear maps involved has co-rank 1. This complements the result of Barthe concerning the case where the linear maps all have rank 1. Pushing our analysis further, we describe the case where the maps have either rank 1 or rank 2.

A separate but related question is to give a list of the finite number of conditions necessary and sufficient for the Brascamp–Lieb inequality to hold. We present an algorithm which generates such a list.

1. INTRODUCTION

The Brascamp–Lieb inequality unifies and generalises several of the most central inequalities in analysis, among others the inequalities of Hölder, Young and Loomis–Whitney. It has the form

$$(1) \quad \int_H \prod_{j=1}^m f_j^{p_j}(B_j x) \, dx \leq C \prod_{j=1}^m \left(\int_{H_j} f_j \right)^{p_j}$$

where H and H_j are finite dimensional Hilbert spaces of dimensions n and n_j respectively, $B_j : H \rightarrow H_j$ are linear maps, p_j are non-negative numbers, C is a finite constant and f_j are non-negative functions. We shall refer to $((B_j), (p_j))$ as the Brascamp–Lieb datum for this inequality.

The inequality was first written down by Brascamp and Lieb in [5] where they pose two questions. The first one is to find the necessary and sufficient conditions on the datum $((B_j), (p_j))$ for (1) to hold and the second one is to determine when the best constant for (1) is attained by a tuple of centred gaussian functions, $f_j(x) = e^{-\langle x, A_j x \rangle}$ with each A_j a symmetric and positive semi-definite linear transformation.

In [7] Lieb showed that gaussians exhaust the inequality in the following sense.

Theorem 1 (Lieb’s Theorem). *Let $C((B_j), (p_j))$ be the smallest constant we can take in (1) so that it holds for all tuples (f_j) of integrable functions*

2000 *Mathematics Subject Classification.* Primary 44A35; Secondary 14M15, 26D20.

and let $C_g((B_j), (p_j))$ be the smallest constant we can take so that it holds for tuples of centred gaussians. Then

$$(2) \quad C((B_j), (p_j)) = C_g((B_j), (p_j)).$$

Brascamp and Lieb proved this theorem in the case when each B_j has rank one already in [5]. With this theorem, the fundamental question of when is $C((B_j), (p_j))$ finite has been reduced to the question of when is $C_g((B_j), (p_j))$ finite. In [3] and [4] the question is further reduced by showing that the Brascamp–Lieb inequality (1) holds for the datum $((B_j), (p_j))$ if and only if we have

$$(3) \quad \dim V \leq \sum_j p_j \dim(B_j V)$$

for all subspaces V of H , the scaling condition

$$(4) \quad \dim H = \sum_j p_j \dim(H_j)$$

holds and

$$(5) \quad p_j \geq 0$$

for all j .

Let us fix the maps B_j . Then for which tuples (p_j) does the Brascamp–Lieb inequality hold, that is which tuples satisfy (3), (4) and (5)?

Since each of the conditions is a linear inequality or equality in the variables (p_j) and since the coefficients in (3) are dimensions of spaces which can only range through a finite set, it is clear that the set of tuples (p_j) such that these conditions hold is a convex set in \mathbb{R}^m whose boundary consists of a finite number of hyperplanes. It is thus a polyhedron and we shall refer to it as the *Brascamp–Lieb polyhedron* for the m -transformation (B_j) .

The scaling and positivity conditions (4) and (5) imply that this polyhedron lies in the intersection of a hyperplane and the first 2^m -tant in \mathbb{R}^m . What portion of this intersection the polyhedron occupies can vary greatly. In particular, for Hölder’s inequality the conditions in (3) do not give any restrictions and the polyhedron is this whole intersection. On the other hand, (3) for the Loomis–Whitney inequality restricts the polyhedron to the one point set $(p_j)_{1 \leq j \leq n} = (\frac{1}{n-1})_{1 \leq j \leq n}$.

The conditions (3), (4) and (5) give a description of the Brascamp–Lieb polyhedron, \mathcal{S} , in the sense that if we want to check whether a particular point (p_j) belongs to \mathcal{S} then we can do so by checking (p_j) against each one of these conditions and if it satisfies them all then the point belongs to the polyhedron. However, for two reasons it might be considered of benefit to give an alternative description. Firstly, the shape of the polyhedron can still seem quite unclear, in particular we do not have a result which says that the point (p_j) lies in the polyhedron if and only if it is of some prescribed form. Secondly, there is the question how many conditions are included in (3). Although, as we said above, it is only a finite number because the dimension

of the spaces involved can only range through a finite set, it remains unclear how to get an exhaustive list of the conditions as it would seem to require examining each subspace V of H . In this note, we will address both of these problems.

For the first problem, it is known by the Weyl–Minkowski theorem that a bounded polyhedron is a polytope, that is the convex hull of a finite set of points. Furthermore, it is a consequence of Carathéodory’s theorem that each point in a bounded polyhedron can be written as a convex combination of the vertices of the polyhedron. Here we say that a point (q_j) is a vertex of a polyhedron if there exists a hyperplane such that the intersection of the hyperplane and \mathcal{S} is the singleton $\{(q_j)\}$ and by writing (p_j) as a convex combination of the vertices we mean that (p_j) lies in the polyhedron if and only if we can write

$$p_j = \sum_{s=0}^{s_0} \lambda_s q_{s,j}$$

for all j , where $\lambda_s \geq 0$, $\sum_s \lambda_s = 1$ and q_s for $s = 1, \dots, s_0$ is an enumeration of the vertices. For these standard results in convexity see for example [2].

The problem of determining the vertices of \mathcal{S} has until now only been resolved in the rank-one case. There we have the following result.

Theorem 2 (Rank one case, Barthe [1]). *Let $B_j x = \langle v_j, x \rangle$ for vectors v_j in H . Then (q_j) is a vertex of \mathcal{S} if and only if $q_j = \chi_I(j)$ where χ_I denotes a characteristic function of an index set I such that $(v_j)_{j \in I}$ forms a basis for H .*

This result is reproved in [6] and [4].

In Section 2 we present a new analysis of the properties of the vertices which has the benefit that aside from yielding a new proof of the result of Barthe it makes it possible to determine the form of the vertices in several other cases.

Theorem 3 (Rank $n - 1$ case). *Assume B_j all have rank $n - 1$ and for each j let $\{v_j\}$ be a nonzero element in the kernel of B_j . Then (q_j) is a vertex of \mathcal{S} if and only if $q_j = \frac{1}{n-1} \chi_I(j)$ where I is an index set such that $(v_j)_{j \in I}$ forms a basis for H .*

The main lemma for our treatment of these results is the following.

Lemma 4. *Let (q_j) be a vertex of \mathcal{S} . Then the support of q , $\{j | q_j \neq 0\}$, can have at most n elements where n is the dimension of H .*

Finally, we will also push the analysis further to give a description of the vertices in the case when each B_j has rank either 1 or 2.

In Section 3 we address the second problem mentioned above, how can we know which conditions are included in (3). To state the result we make the following definition.

Definition 5. Let $(V_k)_{k \in K}$ be a family of subspaces of a common space. Then the *lattice* of (V_k) , denoted $\mathcal{L}_{(V_k)}$ is defined as follows

- (1) $V_k \in \mathcal{L}_{(V_k)}$ for each $k \in K$;
- (2) $V_1 \cap V_2, V_1 + V_2 \in \mathcal{L}_{(V_k)}$ for any $V_1, V_2 \in \mathcal{L}_{(V_k)}$.

We neither require $\{0\}$ nor the whole space to be elements of the lattice.

Remark 6. The lattice of a given family of spaces is the smallest set of spaces which contains each member of the family and is closed under the operations of set intersection and vector space addition; we say that the lattice is generated by the family.

Definition 7. For the m -transformation (B_j) we let $\mathcal{L}_{(B_j)}$ denote $\mathcal{L}_{(\ker(B_j))}$, the lattice generated by the kernels of B_j .

In Section 3 we prove the following theorem:

Theorem 8. *Let $((B_j), (p_j))$ be a Brascamp–Lieb datum. Then a necessary and sufficient condition for the the Brascamp–Lieb constant $C((B_j), (p_j))$ to be finite is that (4) and (5) hold and (3) holds for each subspace in $\mathcal{L}_{(B_j)}$.*

However, even with Theorem 8 there remain some questions. Firstly, do we know that the number of elements in $\mathcal{L}_{(B_j)}$ is finite? The answer to this seems to be no in general, see [8] for an overview discussion on lattice theory, to which this question belongs. However, it is clear that the number of elements is countable and it is straightforward to generate a list of elements which we can check (3) on in sequence. So for computational purposes, a more important variant of this question is: how do we know when to stop, that is, when can we be sure that we have got a list of all the conditions included in (3)? We will address this question towards the end of Section 3.

Remark 9. It is a comment of Michael Christ that by working through the induction proof of the Brascamp–Lieb inequality in [3] an algorithm which gives necessary and sufficient conditions for $C((B_j), (p_j))$ to be finite can be found. The proof we give of Theorem 8 is along these lines. The proof also establishes that the lattice $\mathcal{L}_{(B_j)}$ is sufficient.

This note forms part of my PhD thesis from the University of Edinburgh. I would like to thank my supervisor Tony Carbery for his support and in particular for discussions relating to the material in Section 3.

Part of this work was completed during a stay at the University of Athens. I would like to thank Apostolos Giannopoulos for his hospitality and for helpful comments.

2. THE VERTICES OF \mathcal{S}

Proof of Lemma 4. Assume U and W are two subspaces of H such that inequality (3) holds with equality for the point (q_j) of \mathcal{S} and U and W .

Then we get that

$$\begin{aligned}
& \sum_j q_j \dim(B_j U) + \sum_j q_j \dim(B_j W) \\
&= \sum_j q_j (\dim(B_j U) + \dim(B_j W)) \\
(6) \quad &= \sum_j q_j (\dim(B_j U \cap B_j W) + \dim(B_j U + B_j W)) \\
&\geq \sum_j q_j (\dim(B_j(U \cap W)) + \dim(B_j(U + W))) \\
&\geq (\dim(U \cap W) + \dim(U + W)) \\
&= (\dim U + \dim W)
\end{aligned}$$

where we have used twice the fact that $\dim U + \dim W = \dim(U + W) + \dim(U \cap W)$ for any subspaces U and W . Also for the first inequality we have used that $\dim(B_j U + B_j W) = \dim(B_j(U + W))$ and $\dim(B_j U \cap B_j W) \geq \dim(B_j(U \cap W))$. The second inequality follows since (q_j) belongs to the polyhedron and therefore the condition (3) holds with (q_j) and both $U \cap W$ and $U + W$.

Since we are assuming that the beginning and end of this chain are equal, we must in fact have equality all the way. This tells us that we have equality in inequality (3) for $U \cap W$ and $U + W$ and that for all j such that $q_j > 0$ we have

$$(7) \quad \dim(B_j U) + \dim(B_j W) = \dim(B_j(U \cap W)) + \dim(B_j(U + W)).$$

We note that so far we have proved the following.

Lemma 10. *Let U and W be critical subspaces of H for a Brascamp–Lieb datum $((B_j), (p_j))$. Then $U \cap W$ and $U + W$ are also critical and for all j such that $p_j > 0$ we have that (7) holds.*

Now, if (q_j) is a vertex of \mathcal{S} then we will have a set of indices, J , such that

$$(8) \quad q_j = 0 \quad \text{for } j \notin J$$

and a collection of subspaces, \mathcal{V} , such that

$$(9) \quad \dim V = \sum_j q_j \dim(B_j V) \quad \text{if } V \in \mathcal{V}.$$

A vertex of a polyhedron is the unique solution of the set of linear equations which the facets adjacent to the vertex satisfy. Thus, the system (8), (9) of linear equations determines the vertex (q_j) uniquely.

Let us now apply row operations to this system to simplify it. By subtracting the appropriate multiples of (8) from (9) we can substitute (9) with

$$(10) \quad \dim V = \sum_{j \in J} q_j \dim(B_j V) \quad \text{for } V \in \mathcal{V}.$$

Now, take $U, W \in \mathcal{V}$. By the above discussion, we have $U \cap W, U + W \in \mathcal{V}$ as well and furthermore, the equality for W can be deduced from the equality for $U \cap W$, U and $U + W$ as follows.

$$\begin{array}{r}
\left(\begin{array}{c} \dim(U \cap W) = \sum_{j \in J} q_j \dim(B_j(U \cap W)) \end{array} \right) \\
+ \left(\begin{array}{c} \dim(U + W) = \sum_{j \in J} q_j \dim(B_j(U + W)) \end{array} \right) \\
- \left(\begin{array}{c} \dim U = \sum_{j \in J} q_j \dim(B_j U) \end{array} \right) \\
\hline
= \left(\begin{array}{c} \dim W = \sum_{j \in J} q_j \dim(B_j W) \end{array} \right)
\end{array}$$

where we have used (7) to simplify the right hand side. This shows that we may remove the equation coming from W from (10) by row operations and thus without affecting the solution set.

Let us try and remove as many equations from (10) as we can. First of all, we may assume that $\{0\}$ is not in \mathcal{V} as (10) is content free for that space. Let us then take a $U_1 \in \mathcal{V}$ such that no proper subspace of U_1 is in \mathcal{V} . Clearly such a space exists as we cannot have an infinite chain of nested subspaces in H . Define $\mathcal{V}_{U_1} := \{W \in \mathcal{V} : U_1 \subset W\}$. Then all the equalities for the subspaces in \mathcal{V} can be deduced from the equalities for the subspaces in \mathcal{V}_{U_1} . To see this we note that if $W \in \mathcal{V} \setminus \mathcal{V}_{U_1}$ then $W \cap U_1 = \{0\}$ so the equality for W can be deduced from the equalities for U_1 and $U_1 + W$ which are elements of \mathcal{V}_{U_1} .

Next, let $U_2 \in \mathcal{V}_{U_1}$, $U_2 \neq U_1$ be such that no subspace $W \in \mathcal{V}_{U_1}$ lies properly between U_1 and U_2 . Then as in the last paragraph we see that all equalities for subspaces in \mathcal{V}_{U_1} can be deduced from the equalities for the subspaces in \mathcal{V}_{U_2} and the equality for U_1 . Continuing this process, we get a flag $U_1 \subsetneq U_2 \subsetneq \dots \subsetneq U_s$ such that all the equalities for the subspaces in \mathcal{V} can be deduced from the equalities for the spaces in this chain.

Thus we have seen that by using row operations we can remove all the equations from (10) except the ones coming from this flag, which we shall refer to as \mathcal{U} , and still have left the linear system

$$(11) \quad q_j = 0 \quad \text{for } j \notin J;$$

$$(12) \quad \dim U = \sum_j q_j \dim(B_j U) \quad \text{if } U \in \mathcal{U}$$

which is equivalent to the original one. Since H is n -dimensional, \mathcal{U} can have at most n elements so the number of equations in (12) is at most n . However, since the system (11), (12) is a linear system which has a unique solution in \mathbb{R}^m , there must be at least m equations in the system. Therefore, there must be at least $m - n$ elements not in the set J and so the solution to the system (q_j) can have at most n non-zero elements.

This completes the proof of the lemma. \square

The next lemma partly addresses the question how one can check that a particular point is a vertex.

Lemma 11. *Let a Brascamp–Lieb datum $((B_j), (p_j))$ be given and assume that $\mathcal{U} = (U_1, \dots, U_s)$ is a flag in H , that is $U_1 \subsetneq U_2 \subsetneq \dots \subsetneq U_s = H$, such that (12) holds. Assume also that the inequality (3) holds for any space \tilde{W} which can be added into the flag.*

Then inequality (3) holds for any subspace W of H so the Brascamp–Lieb inequality holds for this datum.

Remark 12. If \mathcal{U} is a maximal flag we cannot add any subspace to the flag so if we have a vector (q_j) for which (12) holds for a maximal flag \mathcal{U} then (q_j) is a vertex of the Brascamp–Lieb polyhedron. This is however not a necessary condition for (q_j) to be a vertex, see Remark 15.

Proof of Lemma 11. If we re-examine the calculations in (6) we see that if we assume that (3) holds for $U \cap W$ and $U + W$ and it holds with equality for U then we get that (3) holds for W .

Let us now define $t_0 \in \{0, \dots, s\}$ such that $U_{t_0} \subset W$ but $U_{t_0+1} \not\subset W$. To ensure that t_0 is well-defined we allow it to take the value 0 in which case we define $U_0 = \{0\}$. We see that if (3) holds for $W \cap U_{t_0+1}$ and $W + U_{t_0+1}$ then it holds for W . Since $U_{t_0} \subset W \cap U_{t_0+1} \subset U_{t_0+1}$ we see that (3) holds for $W \cap U_{t_0+1}$ by assumption. For $W + U_{t_0+1}$ we argue inductively. We note that $W + U_{t_0+1} \supset U_{t_0+1}$ so we can repeat this process for that space, that is find a $t_1 > t_0$ such that $U_{t_1} \subset W + U_{t_0+1}$ but $U_{t_1+1} \not\subset W + U_{t_0+1}$ and then (3) for $W + U_{t_0+1}$ will follow from the condition for $(W + U_{t_0+1}) \cap U_{t_1+1}$ which lies between U_{t_1} and U_{t_1+1} and the condition for $W + U_{t_1+1}$. This process will give us a flag $U_{t_0} \subset \dots \subset U_{t_r}$ which is a subflag of the flag \mathcal{U} and can therefore not contain more than s elements. Furthermore, this flag has the property that to confirm that (3) holds for W we need only to check that (3) holds for spaces V such that $U_t \subset V \subset U_{t+1}$ with $t \in \{t_0, \dots, t_r\}$. Since W was arbitrary we have proved the lemma. \square

Let us now list all the possible vertices in several cases. First let us assume that all the maps B_j have the same rank.

Proof of Theorem 2. As before, we let (q_j) be a vertex of the polyhedron and J be the set of indices j such that $q_j > 0$. If v_j for $j \in J$ do not span H then we do not have a solution to the system (3), (4) and (5). To see this, let V be a subspace of codimension 1 which contains v_j for all $j \in J$. Then V^\perp lies in the kernel of all the relevant B_j . Therefore, testing (3) on V^\perp gives $1 = \dim V^\perp \leq \sum_j q_j \dim(B_j V^\perp) = 0$ which is impossible.

This and Lemma 4 shows that $|J| = n$ and for each $j \in J$ there is a vector v_j such that $v_j \in \ker B_{j'}$ for all $j' \in J \setminus \{j\}$ but $v_j \notin \ker B_j$. Define

$$U_j = \sum_{\substack{j' \in J \\ j' \leq j}} \text{span}(v_{j'}).$$

Then $\mathcal{U} = (U_j)_{j \in J}$ is a maximal flag in H . With these definitions of J and \mathcal{U} we can see that (11) and (12) have the unique solution $q_j = 1$ for $j \in J$ and $q_j = 0$ otherwise. The note following Lemma 11 therefore gives that each vector of this form is a vertex of the polyhedron. \square

Proof of Theorem 3. With (q_j) and J as before, we first note that if $\ker B_j$ for $j \in J$ do not span H then we do not have a solution to the system (3), (4) and (5) as can be seen from testing (3) on a space V such that $\sum_{j \in J} \ker B_j \subset V$ and $\dim V = n - 1$. This gives $n - 1 = \dim V \leq \sum_j q_j \dim(B_j V) = (n - 2) \sum_j q_j$ whereas the scaling condition (4) gives $n = \sum_j q_j (n - 1)$.

From this and Lemma 4 we then see that $|J| = n$. Also, if we define

$$U_j = \sum_{\substack{j' \in J \\ j' \leq j}} \ker B_{j'}$$

then $\mathcal{U} := (U_j)_{j \in J}$ is a maximal flag in H . The set of equations (12) for this flag becomes

$$s_j = \sum_{\substack{j' \in J \\ j' \leq j}} q_{j'} (s_j - 1) + \sum_{\substack{j' \in J \\ j' > j}} q_{j'} s_j \quad j \in J$$

where $s_j := |\{j' \in J | j' \leq j\}|$. Since the number of terms in the first sum is s_j and the number of terms in the last sum is $n - s_j$ it is evident that $q_j = \frac{1}{n-1}$ for $j \in J$ is a solution. Since the system has rank n this is the only solution and since the flag is maximal we get a vertex for the polyhedron. \square

2.1. Mixed rank one and two. We can push this analysis further and examine the mixed rank case when each B_j has rank 1 or 2. Again, we assume (q_j) is a vertex of \mathcal{S} and J and $\mathcal{U} = (U_1 \subsetneq U_2 \subsetneq \dots \subsetneq U_s)$ are such that (11) and (12) hold.

By subtracting the equation for U_{k-1} from the equation for U_k we see that we can replace (12) with

$$(13) \quad \dim(U_k/U_{k-1}) = \sum_j q_j (\dim(B_j U_k) - \dim(B_j U_{k-1}))$$

for $U_k \in \mathcal{U}$, $k \geq 1$ and with $U_0 = \{0\}$. In this set of equations we note that the coefficients multiplying q_j sum up to the rank of B_j and the constant coefficients sum up to $\dim H$. Therefore, if we let J_1 and J_2 be the set of indices from J for the rank 1 and 2 transformations in the set $\{B_j | j \in J\}$ respectively and let m_1 and m_2 be the number of elements in these sets then the sum of the elements in the coefficient matrix of (13) equals $m_1 + 2m_2$.

Furthermore, since the set of equations (13) uniquely determines $(q_j)_{j \in J}$ and $|J| = m_1 + m_2$ we get that $s \geq m_1 + m_2$.

Now, for each $j \in J_1$ the coefficients of q_j in (13) must all be 0 except one which must be 1. Therefore, at most m_1 of the equations can contain a non-zero coefficient for an element q_j with $j \in J_1$ and these equalities must contain at least m_1 of the non-zero coefficients in the matrix.

There are now two cases, either there is equality in each step of this calculation, that is there are exactly m_1 of the equations which have a non-zero coefficient for an element q_j with $j \in J_1$ and these equations have only these non-zero coefficients and there are exactly m_2 equations left which have all of the non-zero coefficients for the q_j with $j \in J_2$ which sum up to $2m_2$. Moreover, each of these m_2 equations must have either one coefficient equal to 2 and all other 0 or two coefficients equal to 1 and all other 0. Otherwise, if any of this does not hold, then, by the pigeonhole principle, there must be an equation among these, all of whose coefficients are zero except one, for q_j with $j \in J_2$, which must be one.

Note that we have made heavy use of the fact that the coefficients in (13) must all be non-negative integers and we may assume that no equation has zero coefficients in front of all the q_j as such an equation can be removed from the linear system and the corresponding subspace can be removed from the flag.

In the first case, we get that for each $j \in J_1$, the relevant equation from (13) takes the form $1 = q_j$. The left hand side must be 1 as we know that $0 < q_j \leq 1$ for each $j \in J$. Let us say that this is the equation coming from the quotient U_{k_j}/U_{k_j-1} . From the fact that $\dim(B_{j'}U_{k_j}) = \dim(B_{j'}U_{k_j-1})$ for all $j' \neq j$ we get that the intersection of $\ker B_{j'}$ with $U_{k_j} \setminus U_{k_j-1}$ is non-empty. Now, $U_{k_j-1} \subset \ker B_j$ whereas $U_{k_j} \setminus U_{k_j-1}$ contains no vectors in $\ker B_j$ so we see that $\ker B_{j'} \cap (H \setminus \ker B_j)$ is non-empty for any $j' \neq j$. Since $\dim \ker B_j = n - 1$ we get by testing (3) on $\ker B_j$ that

$$(14) \quad \dim(H) - 1 \leq \sum_{j' \neq j} q_{j'} \dim(B_{j'}H).$$

Since we know that we have equality in (3) for H and since we have $q_j = 1$ we get by subtraction that (14) must in fact be an equality.

All in all, we get by repeating this process, rearranging and carrying out the reductions in the proof of Lemma 4 again that there exists a subspace H_1 of H such that $\dim B_j H_1 = 2$ for all $j \in J_2$ but H_1 lies in the kernel of all B_j for $j \in J_1$. Furthermore, H_1 is $n - m_1$ dimensional, the cosets $v_j + H_1$, $v_j \in \text{im } B_j$ for $j \in J_1$, form a basis for H/H_1 and $q_j = 1$ for all $j \in J_1$.

So we are left with a flag in H_1 and m_2 equations associated with it, all of whose non-zero coefficients are for q_j with $j \in J_2$. If we have that one of these equations has only one non-zero coefficient, which must then be 2, then that equation must take the form $2 = 2q_j$. This we see since the left hand side cannot be larger than 2 as q_j is at most 1 and since we must

always have

$$\dim(U_k/U_{k-1}) \geq \dim(B_j U_k) - \dim(B_j U_{k-1})$$

so the coefficient on the left hand side must be as large as any coefficient on the right hand side. Now, in the same way as before with the rank one spaces we get that there exists a subspace H_2 of H_1 and a flag in H_2 such that all of the equations associated to this flag have the form

$$(15) \quad t_{j,j'} = q_j + q_{j'}$$

for some $j, j' \in J_{2,1} \subset J_2$ and $t_{j,j'} \in \{1, 2\}$. Then if we let $J_{2,2} := J_2 \setminus J_{2,1}$ and for each $j \in J_{2,2}$ we let $\{v_{j,1}, v_{j,2}\}$ be a basis for $\text{im } B_j$ then the set of cosets $\{v_{j,l} + H_2 | j \in J_{2,2} \text{ and } l = 1, 2\}$ forms a basis for H_1/H_2 and $q_j = 1$ for all $j \in J_{2,2}$. We also note that the flag we get by adding the span of the vectors from this basis one by one to the subspace H_2 is a maximal flag between H_2 and H_1 and we have equality in (13) for each step.

Now, if we have an equation in the set (15) with $t_{j,j'} = 2$ then we must have $q_j = q_{j'} = 1$ as neither can be greater than 1. Then we can insert a space into the flag which splits the single equation into the two equations and those equations are of the form we originally split off from the main argument. We will deal with these equations in the next paragraph but one and see that the index set of those should properly be considered as part of $J_{2,2}$.

When that rearrangement has been done we can thus get that all the equations concerning q_j with $j \in J_{2,1}$ have the form $1 = q_j + q_{j'}$. Let us define a relation on $J_{2,1}$ such that j is related to j' if there is an equation of the form $1 = q_j + q_{j'}$ with these j, j' . If we draw the graph of this relation then each vertex j will have exactly two edges connected to it. Therefore we can see that the graph will be a collection of disjoint circles. Let us examine one of these circles. We can write all of the equations relating to the vertices in this circle in the form

$$(16) \quad \begin{array}{rcl} q_{j_1} + q_{j_2} & & = 1 \\ & q_{j_2} + q_{j_3} & = 1 \\ & & q_{j_{l-1}} + q_{j_l} = 1 \\ q_{j_1} & & + q_{j_l} = 1. \end{array}$$

The number of equations in this list is the same as the number of variables. However, if there is an even number of equations then the sum we get by adding the even numbered equations is the same as the sum we get by adding the odd numbered equations and so this system does not have a unique solution, contrary to our assumptions. Therefore, the number of equations in each circle is odd and in that case the system has a unique solution, which clearly is $q_j = \frac{1}{2}$ for all $j \in J_{2,1}$. We note that as the left hand side of these equations is always 1, the flag they come from must be maximal.

Finally, let us look at the other case, where one of the equations in (13) is of the form $1 = q_j$ with $j \in J_2$. Since the sum of the coefficients in front of q_j equals 2 there must be another equation with the term q_j . Either, it also takes the form $1 = q_j$ or the form $t = q_j + Q$ where $t > 1$ is an integer and Q stands for terms with $q_{j'}$, $j' \in J_2 \setminus \{j\}$. Let us first examine the second case. Assume that it comes from (13) with U_{k_j}/U_{k_j-1} where the codimension of U_{k_j-1} in U_{k_j} is t . Since the coefficient multiplying q_j is 1 we get that there are $t - 1$ independent vectors in the intersection of $\ker B_j/U_{k_j-1}$ and U_{k_j}/U_{k_j-1} . Let \tilde{U} denote the vector sum of the span of these and U_{k_j-1} . By testing (3) on \tilde{U} and subtracting (3) on U_{k_j-1} which we know gives an equality we get that $t - 1 \leq Q'$ where Q' denotes the contribution to this sum from terms $q_{j'}$, $j' \in J_2 \setminus \{j\}$. Now we get the chain of inequalities

$$t = 1 + (t - 1) \leq q_j + Q' \leq q_j + Q = t$$

and so we must have equality all the way and in particular this shows that we may add \tilde{U} to the flag which gives equalities and assume that both equalities involving q_j take the form $1 = q_j$.

For the purpose of determining the vector (q_j) uniquely these two identical equations do the same as the single equation $2 = 2q_j$. We can therefore merge them and remove one space U_k from the flag \mathcal{U} . This shows that we may assume that the only equation involving this q_j has the form $2 = 2q_j$ and this we had already analysed above.

All in all we have proved the following.

Theorem 13 (Mixed rank 1 and 2). *Let B_j for $j \in J_1$ be rank 1 linear transformations from H and let B_j for $j \in J_2$ be rank 2 linear transformations. Then (q_j) is a vertex of \mathcal{S} if and only if the following holds*

- (1) $q_j = 1$ for all $j \in J_1$;
- (2) the set J_2 can be divided into two sets $J_{2,1}$ and $J_{2,2}$ such that
 - $q_j = \frac{1}{2}$ for all $j \in J_{2,1}$ and
 - $q_j = 1$ for all $j \in J_{2,2}$ and
- (3) the indices in $J_{2,1}$ can be split into classes such that the equations for each class take the form (16) and the number of indices in each class is odd.
- (4) There exists a maximal flag $U_1 \subsetneq \cdots \subsetneq U_n$ in H and numbers $0 \leq s_1 \leq s_2 \leq n$ such that
 - $\dim B_j U_{s_1} = 2$ for all $j \in J_{2,1}$ but $U_{s_1} \subset \ker B_j$ for all $j \in J_{2,2}$ and $j \in J_1$; and
 - $\dim B_j U_{s_2} = 2$ for all $j \in J_{2,2}$ but $U_{s_2} \subset \ker B_j$ for all $j \in J_1$.

Remark 14. From the proof of the theorem it is clear that we may rearrange the flag so that the equations for q_j with $j \in J_{2,2} \cap J_1$ come in any order. However, this is not the case for U_{s_1} . In fact there might be only one way of choosing this maximal flag for U_{s_1} . An example of such a configuration is with $\dim H = 5$ and B_j for $j = 1, \dots, 5$ are the rank two projections onto

$\langle e_1, e_2 + e_3 \rangle$, $\langle e_1, e_4 \rangle$, $\langle e_2 + e_1, e_4 + e_3 \rangle$, $\langle e_2, e_5 \rangle$ and $\langle e_3, e_5 + e_4 \rangle$ respectively, where $\{e_i\}_{i=1,\dots,5}$ is an orthonormal basis for H and the angled brackets denote the span of the relevant vectors. Then the only maximal flag for which we have equality is

$$\langle e_5 \rangle \subset \langle e_4, e_5 \rangle \subset \langle e_3, e_4, e_5 \rangle \subset \langle e_2, e_3, e_4, e_5 \rangle \subset \langle e_1, e_2, e_3, e_4, e_5 \rangle.$$

Remark 15. In the cases we have looked at, all of the vertices have had associated with them flags of maximal length. However, this is not the case in general as can be seen from the following example. We take H of dimension 8 with an orthonormal basis $(e_i)_{i=1,\dots,8}$. For $j = 1, \dots, 4$ we take B_j to be the orthogonal projections onto the spaces $\langle e_1, e_2, e_5 \rangle$, $\langle e_2, e_4, e_7 \rangle$, $\langle e_1 + e_2, e_6, e_8 \rangle$ and $\langle e_3 + e_4, e_5 + e_6, e_7 + e_8 \rangle$ respectively. Then we have the flag

$$\langle e_1, e_2 \rangle \subset \langle e_1, e_2, e_3, e_4 \rangle \subset \langle e_1, e_2, e_3, e_4, e_5, e_6 \rangle \subset \langle e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8 \rangle$$

for which (13) becomes

$$\begin{aligned} p_1 + p_2 + p_3 &= 2 \\ p_1 + p_2 + p_4 &= 2 \\ p_1 + p_3 + p_4 &= 2 \\ p_2 + p_3 + p_4 &= 2 \end{aligned}$$

which has the solution $p_1 = p_2 = p_3 = p_4 = \frac{2}{3}$. It is straightforward to confirm that the inequality (3) is satisfied for any subspace V of H as from Lemma 11 we know that we need only to check it for subspaces which can be placed into the flag. However, no linear combination of the p_j with non-negative integer coefficients can equal 1 so there can be no one-dimensional subspace of H which has equality in (3).

Remark 16. If all the maps B_j have rank k then (4) gives that

$$(17) \quad \sum_j p_j = n/k$$

and we can rewrite (3) as

$$\dim V \leq \sum_j p_j \dim(B_j V) = \sum_j p_j (\dim V - \dim(\ker B_j \cap V))$$

which says

$$(18) \quad \sum_j p_j \dim(\ker B_j \cap V) \leq \frac{n-k}{k} \dim V.$$

We can carry out the analysis of this section with the conditions (5), (17) and (18) and in particular we can recover a theorem similar to Theorem 13 for the case when all B_j have rank $n - 2$.

3. THE FACETS OF \mathcal{S}

We begin this section with a proof of Theorem 8.

Proof. The necessity of the conditions follows immediately from [4] as they are a subset of the necessary conditions established there.

To show that the conditions are sufficient we use induction on $n+m$, where $n = \dim H$ and m is the degree of multilinearity of the form. For the base case we consider $m = 1$. Then testing (3) on $\ker B_1$ gives that $\dim \ker B_1 = 0$ so B_1 is surjective and then the scaling condition gives $\dim H_1 = \dim H$ and $p_1 = 1$. We see then the inequality evidently holds with equality if we take $C(B_1, p_1) = (\det B_1)^{-1}$.

For the inductive step we take a datum $((B_j), (p_j))$ and assume that the result holds for each datum for which the quantity $m + n$ is smaller.

As before, the conditions (4), (5) along with (3) for $V \in \mathcal{L}_{(B_j)}$ define a bounded convex polyhedron in \mathbb{R}^m and by multilinear interpolation, to show that the result holds everywhere in this polyhedron is enough to establish it at each vertex of it. As we have already dealt with the case $m = 1$ we may assume $m > 2$ and then we get that, at a vertex, more than one of the linear inequalities defining the polyhedron must be satisfied with equality.

There are now two cases. Either we have $p_{j_0} = 0$ for some j_0 or there is a space $U \in \mathcal{L}_{(B_j)} \setminus \{\{0\}, H\}$ such that

$$(19) \quad \dim U = \sum_j p_j \dim(B_j U).$$

In the first case we see that we may write the Brascamp–Lieb inequality without referring to j_0 and the result thus follows from the induction hypothesis since the degree of multilinearity has been reduced.

In the second case we can factor the Brascamp–Lieb form in the following way: Define

$$\begin{aligned} \tilde{B}_j &: U \rightarrow B_j U : x \mapsto B_j x \\ \tilde{\tilde{B}}_j &: U^\perp \rightarrow (B_j U)^\perp : x \mapsto \Pi_{(B_j U)^\perp} B_j x \\ \Gamma_j &: U^\perp \rightarrow B_j U : x \mapsto \Pi_{B_j U} B_j x \end{aligned}$$

where $\Pi_{(B_j U)^\perp}$ and $\Pi_{B_j U}$ denote the orthogonal projections onto the relevant spaces. Then we can calculate

$$\begin{aligned}
\int_H \prod_{j=1}^m f_j^{p_j}(B_j x) dx &= \int_{U^\perp} \int_U \prod_{j=1}^m f_j^{p_j}(\tilde{B}_j \tilde{x} + B_j \tilde{x}) d\tilde{x} d\tilde{x} \\
&\leq C((\tilde{B}_j), (p_j)) \int_{U^\perp} \prod_{j=1}^m \left(\int_{B_j U} f_j(\tilde{y} + B_j \tilde{x}) d\tilde{y} \right)^{p_j} d\tilde{x} \\
&= C((\tilde{B}_j), (p_j)) \int_{U^\perp} \prod_{j=1}^m \left(\int_{B_j U} f_j(\tilde{y} + \Gamma_j \tilde{x} + \tilde{B}_j \tilde{x}) d\tilde{y} \right)^{p_j} d\tilde{x} \\
&= C((\tilde{B}_j), (p_j)) \int_{U^\perp} \prod_{j=1}^m \left(\int_{B_j U} f_j(\tilde{y} + \tilde{B}_j \tilde{x}) d\tilde{y} \right)^{p_j} d\tilde{x} \\
&\leq C((\tilde{B}_j), (p_j)) C((\tilde{\tilde{B}}_j), (p_j)) \\
&\quad \prod_{j=1}^m \left(\int_{B_j U^\perp} \int_{B_j U} f_j(\tilde{y} + \tilde{\tilde{y}}) d\tilde{y} d\tilde{\tilde{y}} \right)^{p_j} \\
&= C((\tilde{B}_j), (p_j)) C((\tilde{\tilde{B}}_j), (p_j)) \prod_{j=1}^m \left(\int_{H_j} f_j(y) dy \right)^{p_j}.
\end{aligned}$$

Here we have used for the first inequality that for almost any $\tilde{x} \in U^\perp$ the tuple $(f_j(\cdot + B_j \tilde{x}))$ consists of non-negative integrable functions defined on $B_j U$ and we can therefore use the Brascamp–Lieb inequality for the datum $((\tilde{B}_j), (p_j))$. For the next equality we use the definitions of Γ_j and \tilde{B}_j and for the one below that we use the translation invariance of the inner integral and the fact that $\Gamma_j \tilde{x} \in B_j U$ for any $\tilde{x} \in U^\perp$. For the second inequality we use the fact that for any j the inner integral defines a non-negative function of $\tilde{B}_j \tilde{x}$ with domain $(B_j U)^\perp$ and we can therefore use the Brascamp–Lieb inequality for the datum $((\tilde{\tilde{B}}_j), (p_j))$.

Since we can perform this calculation for any tuple of non-negative integrable functions (f_j) defined on H_j , we have established the inequality

$$(20) \quad C((B_j), (p_j)) \leq C((\tilde{B}_j), (p_j)) C((\tilde{\tilde{B}}_j), (p_j)).$$

In particular this shows that if both $C((\tilde{B}_j), (p_j))$ and $C((\tilde{\tilde{B}}_j), (p_j))$ are finite then $C((B_j), (p_j))$ is finite. Since $\dim U < \dim H$ and $\dim U^\perp < H$ we may use the induction hypothesis to establish that this is the case. The positivity condition (5) clearly holds since the tuple (p_j) is inherited unchanged from the original datum. The scaling condition (4) for \tilde{B} holds by the assumption that U is critical and by subtracting that condition from the scaling condition for H we see that (4) holds for $\tilde{\tilde{B}}_j$.

So the only conditions that remain to be checked are (3) for any space in $\mathcal{L}_{(\tilde{B}_j)}$ and $\mathcal{L}_{(\tilde{\tilde{B}}_j)}$. First of all, we note that the first of these sets is a subset of $\mathcal{L}_{(B_j)}$. To see this we note that it is enough to show that the building blocks of $\mathcal{L}_{(\tilde{B}_j)}$, the sets $\ker \tilde{B}_j$, lie in $\mathcal{L}_{(B_j)}$. Since $\tilde{B}_j = B_j|_U$ we get that $\ker \tilde{B}_j = \ker B_j \cap U$ and the inclusion follows as both the sets on the right hand side are elements of $\mathcal{L}_{(B_j)}$. Now, for any $W \in \mathcal{L}_{(\tilde{B}_j)}$ we have that $W \subset U$ and therefore $\dim \tilde{B}_j W = \dim B_j W$. Therefore, the inequality

$$\dim W \leq \sum_j p_j \dim \tilde{B}_j W$$

is in the list in inequalities coming from $\mathcal{L}_{(B_j)}$.

Secondly, we study $\mathcal{L}_{(\tilde{\tilde{B}}_j)}$. Let us take an element $\tilde{\tilde{W}}$ from this set. Our aim is to establish that the inequality

$$\dim \tilde{\tilde{W}} \leq \sum_j p_j \dim \tilde{\tilde{B}}_j \tilde{\tilde{W}}$$

is in the list from $\mathcal{L}_{(B_j)}$. Since U is critical and the elements in the pairs U, W and $B_j U, \tilde{\tilde{B}}_j \tilde{\tilde{W}}$ are orthogonal to each other we see that we may equivalently establish the inequality

$$(21) \quad \dim(\tilde{\tilde{W}} + U) \leq \sum_j p_j \dim(\tilde{\tilde{B}}_j \tilde{\tilde{W}} + B_j U).$$

We note that the sets $\tilde{\tilde{B}}_j \tilde{\tilde{W}} + B_j U$ and $B_j(\tilde{\tilde{W}} + U)$ are the same. To see this take an element x from the former set. Then x has the form $\Pi_{(B_j U)^\perp} B_j y + B_j z$ with $y \in \tilde{\tilde{W}}$ and $z \in U$. Now there is an element $y' \in U$ such that $\Pi_{(B_j U)^\perp} B_j y = B_j y + B_j y'$. Then $x = B_j(y + (y' + z))$ with $y \in \tilde{\tilde{W}}$ and $y' + z \in U$. For the other direction we take $x \in B_j(\tilde{\tilde{W}} + U)$. Then we can write $x = B_j(y + z)$ with $y \in \tilde{\tilde{W}}$ and $z \in U$. We take y' as before and then $x = \tilde{\tilde{B}}_j y + B_j(z - y')$ with $y \in \tilde{\tilde{W}}$ and $z - y' \in U$.

Therefore, it is enough to show that $\tilde{\tilde{W}} + U \in \mathcal{L}_{(B_j)}$. To establish this we note first of all that if $\tilde{\tilde{W}} = \ker \tilde{\tilde{B}}_j$ then $\tilde{\tilde{W}} + U = \ker B_j + U$. To see this take $x \in \tilde{\tilde{W}}$. This means by definition that $B_j x \in B_j U$ so $x \in \ker B_j + U$. On the other hand, if we take $x \in \ker B_j$ and write $x = y + z$ with $y \in U$ and $z \in U^\perp$ then $B_j z = B_j x - B_j y = -B_j y \in B_j U$ so $\tilde{\tilde{B}}_j z = 0$ so $z \in \ker \tilde{\tilde{B}}_j$. We also note that for any $\tilde{\tilde{W}}_1, \tilde{\tilde{W}}_2 \in \mathcal{L}_{(\tilde{\tilde{B}}_j)}$ we have that $(\tilde{\tilde{W}}_1 + U) \cap (\tilde{\tilde{W}}_2 + U) = (\tilde{\tilde{W}}_1 \cap \tilde{\tilde{W}}_2) + U$ and $(\tilde{\tilde{W}}_1 + U) + (\tilde{\tilde{W}}_2 + U) = (\tilde{\tilde{W}}_1 + \tilde{\tilde{W}}_2) + U$. The first of those follows from the fact that both $\tilde{\tilde{W}}_1$ and $\tilde{\tilde{W}}_2$ lie in U^\perp and the second is self-evident. It is now clear that by using induction on the number of

operations needed to get to \tilde{W} that we can show that $\tilde{W} + U \in \mathcal{L}_{(B_j)}$ and we thus complete the proof of the theorem. \square

By examining the above proof we can give a procedure which tells us when we have found all the conditions included in (3).

We start by looking for necessary conditions by going through an enumeration of the elements of $\mathcal{L}_{(B_j)}$ and we decide (arbitrarily) to pause when we have found the necessary conditions (3) for $V \in \mathcal{V}$ where $\mathcal{V} \subset \mathcal{L}_{(B_j)}$. At this stage we wish to determine whether we have found all the necessary conditions for the Brascamp–Lieb inequality to hold. The conditions (3) for $V \in \mathcal{V}$, together with the conditions (4) and (5) restrict the set of tuples (p_j) for which the Brascamp–Lieb inequality holds to a polyhedron $\tilde{\mathcal{S}}_{(B_j)}$ and we wish to determine whether $\tilde{\mathcal{S}}_{(B_j)} = \mathcal{S}_{(B_j)}$ where $\mathcal{S}_{(B_j)}$ is the Brascamp–Lieb polyhedron for (B_j) . This will be the case if and only if each vertex of $\tilde{\mathcal{S}}_{(B_j)}$ is in $\mathcal{S}_{(B_j)}$. There exists an algorithm which lists all of the vertices of $\tilde{\mathcal{S}}_{(B_j)}$. For each vertex (q_j) in this list we know that m of the conditions (3) for $V \in \mathcal{V}$, (4) and (5) are satisfied with equality. If none of these equalities comes from (3) then the support of (q_j) can only contain one element q_{j_0} and we know from above that the Brascamp–Lieb inequality holds at this vertex if and only if $q_{j_0} = 1$ and $\ker B_{j_0} = \{0\}$. Otherwise there is a space $U \in \mathcal{V}$ which lies strictly between $\{0\}$ and H such that (3) holds with equality for U . By the proof above we see that the Brascamp–Lieb inequality holds at (q_j) if and only if it holds for the data $((\tilde{B}_j), (q_j))$ and $((\tilde{\tilde{B}}_j), (q_j))$, that is if $(q_j) \in \mathcal{S}_{(\tilde{B}_j)}$ and $(q_j) \in \mathcal{S}_{(\tilde{\tilde{B}}_j)}$.

To determine whether this is the case we run through the above algorithm for both $\mathcal{S}_{(\tilde{B}_j)}$ and $\mathcal{S}_{(\tilde{\tilde{B}}_j)}$. This recursion can only have n levels of depth and will therefore be completed in a finite number of steps and when it is completed we know whether (q_j) is in $\mathcal{S}_{(B_j)}$ in which case we move on to the next vertex, or whether (q_j) is not in $\mathcal{S}_{(B_j)}$ in which case we break the pause and continue looking for necessary conditions in the list of $\mathcal{L}_{(B_j)}$ until we decide again (arbitrarily) to pause and check whether we have now found all of the necessary conditions.

REFERENCES

- [1] Franck Barthe. On a reverse form of the Brascamp–Lieb inequality. *Invent. Math.*, 134(2):335–361, 1998.
- [2] Alexander Barvinok. *A course in convexity*, volume 54 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2002.
- [3] Jonathan Bennett, Anthony Carbery, Michael Christ, and Terence Tao. Finite bounds for Hölder–Brascamp–Lieb multilinear inequalities. *Math. Res. Lett.* to appear.
- [4] Jonathan Bennett, Anthony Carbery, Michael Christ, and Terence Tao. The Brascamp–Lieb inequalities: finiteness, structure, and extremals. *Geom. Funct. Anal.* to appear.

- [5] Herm J. Brascamp and Elliott H. Lieb. Best constants in Young's inequality, its converse, and its generalization to more than three functions. *Advances in Math.*, 20(2):151–173, 1976.
- [6] Eric A. Carlen, Elliott H. Lieb, and Michael Loss. A sharp analog of Young's inequality on S^N and related entropy inequalities. *J. Geom. Anal.*, 14(3):487–520, 2004.
- [7] Elliott H. Lieb. Gaussian kernels have only Gaussian maximizers. *Invent. Math.*, 102(1):179–208, 1990.
- [8] Gian-Carlo Rota. The many lives of lattice theory. *Notices Amer. Math. Soc.*, 44(11):1440–1445, 1997.

UCLA MATHEMATICS DEPARTMENT, BOX 951555, LOS ANGELES, CA 90095-1555
E-mail address: valdimarsson@math.ucla.edu