THE BRASCAMP-LIEB & BARTHE INEQUALITIES: A SURVEY

Albert Zhang

May 13, 2022

1. The Brascamp-Lieb Inequality

1.1. Introduction. Let H and $(H_j)_{1\le j\le m}$ be Euclidean spaces, $\mathbf{B} = (B_j)_{1\le j\le m}$ be linear surjective maps $B_j : H \to H_j$, and $\mathbf{c} = (c_j)_{1\le j\le m}$ be nonnegative real numbers. Then we denote by $(\mathbf{B}, \mathbf{c})$ a *Brascamp-Lieb datum.* For a fixed Brascamp-Lieb datum, we define an *input* to be an m-tuple $\mathbf{f} = (f_j)_{1\le j\le m}$ of nonnegative measurable functions $f_j : H_j \to \mathbb{R}^+$ satisfying $0 < \int_{H_j} f_j < \infty$. We define

$$\mathrm{BL}(\mathbf{B}, \mathbf{c}; \mathbf{f}) := \frac{\int_H \prod_{j=1}^m (f_j \circ B_j)^{c_j}}{\prod_{j=1}^m (\int_{H_j} f_j)^{c_j}}.$$

We then define the *Brascamp-Lieb constant* $\mathrm{BL}(\mathbf{B}, \mathbf{c})$ to be the supremum of $\mathrm{BL}(\mathbf{B}, \mathbf{c}; \mathbf{f})$ over all inputs $\mathbf{f}$. That is, $\mathrm{BL}(\mathbf{B}, \mathbf{c})$ is the smallest constant for which the functional inequality

$$\int_H \prod_{j=1}^m (f_j \circ B_j)^{c_j} \le \mathrm{BL}(\mathbf{B}, \mathbf{c}) \prod_{j=1}^m \left(\int_{H_j} f_j \right)^{c_j}$$

holds for any choice of nonnegative measureable functions $f_j : H_j \to \mathbb{R}^+$.

We are interested in two questions. The first is regarding finiteness of the constant, i.e. when is $\mathrm{BL}(\mathbf{B}, \mathbf{c}) < \infty$? The second concerns *extremizability.* We say that a Brascamp-Lieb datum is *extremizable* if $\mathrm{BL}(\mathbf{B}, \mathbf{c}) < \infty$ and there exists an input $\mathbf{f}$ such that $\mathrm{BL}(\mathbf{B}, \mathbf{c}; \mathbf{f}) = \mathrm{BL}(\mathbf{B}, \mathbf{c})$. Before diving into the answers to these questions, we provide some special cases of the Brascamp-Lieb inequality and their optimal constants.

Example 1.1 (Hölder's inequality)**.** Let $H_j = H$ and $B_j = \mathrm{Id}_H$ for $1 \le j \le m$. Then under the constraint $\sum_{j=1}^m c_j = 1$, we have Hölder's inequality:

$$\left\| \prod_{j=1}^m f_j^{c_j} \right\|_{L^1(H)} \le \prod_{j=1}^m \|f_j^{c_j}\|_{L^{1/c_j}(H)}$$

which asserts that $\mathrm{BL}(\mathbf{B}, \mathbf{c}) = 1$. For any other choice of $\mathbf{c}$, we have $\mathrm{BL}(\mathbf{B}, \mathbf{c}) = \infty$.

Example 1.2 (Loomis-Whitney inequality)**.** Let $H = \mathbb{R}^n$ and $H_j = e_j^\perp$, where $\{e_1, \ldots, e_n\}$ is the standard basis of $\mathbb{R}^n$. Set $B_j = P_j : \mathbb{R}^n \to e_j^\perp$ to be the orthogonal projection and $c_j = \frac{1}{n-1}$ for each $1 \le j \le n$. Then the Loomis-Whitney inequality asserts that $\mathrm{BL}(\mathbf{B}, \mathbf{c}) = 1$. For any other choice of c_j, we have $\mathrm{BL}(\mathbf{B}, \mathbf{c}) = \infty$. In the case of $n = 3$, the inequality reads

$$\int\int\int f(y,z)^{1/2} g(x,z)^{1/2} h(x,y)^{1/2} dxdydz \le \|f\|_{L^1(\mathbb{R}^2)}^{1/2} \|g\|_{L^1(\mathbb{R}^2)}^{1/2} \|h\|_{L^1(\mathbb{R}^2)}^{1/2}.$$

Example 1.3 (Sharp Young's inequality)**.** Let $H = \mathbb{R}^d \times \mathbb{R}^d$ and $H_j = \mathbb{R}^d$ for $1 \le j \le 3$. The maps $B_j : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ are given by

$$B_1(x,y) = x; \quad B_2(x,y) = y; \quad B_3(x,y) = x - y.$$

So we are looking at the inequality

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x)^{c_1} g(y)^{c_2} h(x-y)^{c_3} \le \mathrm{BL}(\mathbf{B}, \mathbf{c}) \left(\int_{\mathbb{R}^d} f \right)^{c_1} \left(\int_{\mathbb{R}^d} g \right)^{c_2} \left(\int_{\mathbb{R}^d} h \right)^{c_3}$$

It was shown in [Bec] that

$$\mathrm{BL}(\mathbf{B}, \mathbf{c}) = \left(\prod_{j=1}^3 \frac{(1-c_j)^{1-c_j}}{c_j^{c_j}} \right)^{d/2}$$

if $c_1 + c_2 + c_3 = 2$ and $0 \le c_1, c_2, c_3 \le 1$, and $\mathrm{BL}(\mathbf{B}, \mathbf{c}) = \infty$ for any other values of (c_1, c_2, c_3).

For $A_j \in \mathcal{S}^+(H_j)$, the space of positive definite (symmetric) operators on H_j, we may consider the (centered) *Gaussian input* $(g_{A_j}(x))_{1\le j\le m} := (\exp(-\pi\langle A_j x, x\rangle_H))_{1\le j\le m}$. Recall the formula

$$\int_{H_j} g_{A_j}(x)dx = (\det_{H_j} A_j)^{-1/2}. \tag{1}$$

Clearly we have

$$\mathrm{BL}_g(\mathbf{B},\mathbf{c};\mathbf{A}) \le \mathrm{BL}(\mathbf{B},\mathbf{c}),$$

where $\mathrm{BL}_g(\mathbf{B},\mathbf{c};\mathbf{A})$ is defined as

$$\mathrm{BL}_g(\mathbf{B},\mathbf{c};\mathbf{A}) := \left(\frac{\prod_{j=1}^m (\det_{H_j} A_j)^{c_j}}{\det_H(\sum_{j=1}^m c_j B_j^* A_j B_j)}\right)^{1/2}.$$

Now, if we set

$$\mathrm{BL}_g(\mathbf{B},\mathbf{c}) := \sup\{\mathrm{BL}_g(\mathbf{B},\mathbf{c};\mathbf{A}) : \mathbf{A} \text{ is a gaussian input for } (\mathbf{B},\mathbf{c})\}$$

then we have

$$\mathrm{BL}_g(\mathbf{B},\mathbf{c}) \le \mathrm{BL}(\mathbf{B},\mathbf{c}). \tag{2}$$

The following theorem of Lieb [Lieb] says that (2) is actually an equality.

Theorem 1.1 (Lieb)**.** *For any Brascamp-Lieb datum $(\boldsymbol{B},\boldsymbol{c})$, we have $\mathrm{BL}(\boldsymbol{B},\boldsymbol{c}) = \mathrm{BL}_g(\boldsymbol{B},\boldsymbol{c})$.*

Note that Lieb's theorem says nothing about whether the constants are finite, or whether extremals exist and what they are. In [BCCT], necessary and sufficient conditions for $\mathrm{BL}(\mathbf{B},\mathbf{c})$ and $\mathrm{BL}_g(\mathbf{B},\mathbf{c})$ to be finite were determined:

Theorem 1.2 (Finiteness of the Brascamp-Lieb constant)**.** *The constants $\mathrm{BL}(\boldsymbol{B},\boldsymbol{c})$ and $\mathrm{BL}_g(\boldsymbol{B},\boldsymbol{c})$ are finite if and only if the scaling condition*

$$\dim(H) = \sum_{j=1}^m c_j \dim(H_j) \tag{3}$$

and the dimension condition

$$\dim(V) \le \sum_{j=1}^m c_j \dim(B_j V) \text{ for all subspaces } V \subseteq H \tag{4}$$

are satisfied.

Before we show finiteness in Section 1.3, we first prove the Geometric Brascamp-Lieb inequality and the equivalence between geometricity and Gaussian-extremizability in Section 1.2. We say that a Brascamp-Lieb datum $(\mathbf{B},\mathbf{c})$ is *Gaussian-extremizable* if there exists a Gaussian input $\mathbf{A}$ such that $\mathrm{BL}_g(\mathbf{B},\mathbf{c};\mathbf{A}) = \mathrm{BL}_g(\mathbf{B},\mathbf{c})$. In particular, if $(\mathbf{B},\mathbf{c})$ is Gaussian-extremizable, then $\mathrm{BL}_g(\mathbf{B},\mathbf{c}) < \infty$. It turns out that Gaussian-extremizability is equivalent to extremizability. Observe that Lieb's theorem asserts that Gaussian-extremizability implies extremizability. The converse will be shown in Section 1.4.

1.2. The geometric Brascamp-Lieb inequality. We first answer the question of finiteness for the geometric case. Then, as it turns out, every Gaussian-extremizable datum is equivalent, in a sense to be clarified later, to a geometric datum. In particular, we say that a Brascamp-Lieb datum $(\mathbf{B},\mathbf{c})$ is *geometric* if for every $1 \le j \le m$ we have $B_j B_j^* = \mathrm{Id}_{H_j}$, i.e. B_j^* is an isometry, and

$$\sum_{j=1}^m c_j B_j^* B_j = \mathrm{Id}_H . \tag{5}$$

Since B_j^* is an isometry, we may identify H_j with the $E_j := \operatorname{ran}(B_j^*)$, and B_j with the orthogonal projection from H to $E_j \subset H$. Therefore the datum is geometric in the sense of Hilbert space geometry.

Now, taking traces of (5), we obtain (3). Furthermore, if we let $\Pi_V : H \to H$ be the orthogonal projection onto a subspace $V \subset H$, then multiplying (5) by Π_V and taking traces we get

$$\sum_{j=1}^{m} c_j \operatorname{tr}(B_j^* B_j \Pi_V) = \dim(V).$$

Since $B_j^* B_j \Pi_V$ is a contraction with range $B_j V$, we obtain (4) also.

The proof of the geometric Brascamp-Lieb inequality relies on heat flow techniques. In particular, let $f_j : H_j \to \mathbb{R}^+$ be nonnegative Schwartz functions and let $g_j : H \times \mathbb{R}^+ \to \mathbb{R}^+$ be a solution to the initial value problem:

$$\begin{cases} g_j(x,0) = f_j(B_j x) & x \in H \\ \frac{\partial}{\partial t} g_j(x,t) = \Delta g_j(x,t) & x \in H,\ t > 0. \end{cases} \tag{6}$$

Then define the product

$$F(x,t) = \prod_{j=1}^{m} g_j(x,t)^{c_j}.$$

The idea is to show that $\int F(x,t)dx$ is monotone nondecreasing in time, and that $\lim_{t\to 0} \int F(x,t)dx$ and $\lim_{t\to\infty} \int F(x,t)dx$ give us the l.h.s. and r.h.s. of the geometric Brascamp-Lieb inequality.

Lemma 1.3 (Monotonicity)**.** *For the product function $F(x,t)$ defined as above, we have*

$$\frac{\partial}{\partial t} \int F(x,t)dx \geq 0 \quad \forall t > 0.$$

Proof. From (6) we see that

$$\begin{aligned} \frac{\partial}{\partial t} F &= F \sum_{j=1}^{m} c_j \frac{\Delta g_j}{g_j} \\ &= F \sum_{j=1}^{m} c_j \frac{\operatorname{div}(v_j g_j)}{g_j} \\ &= \underbrace{F \sum_{j=1}^{m} c_j \langle v_j, v_j \rangle}_{I} + \underbrace{F \sum_{j=1}^{m} c_j \operatorname{div}(v_j)}_{J} \end{aligned}$$

where we set $v_j = \frac{\nabla g_j}{g_j}$. Applying integration by parts and the divergence theorem we get

$$\begin{aligned} \int J &= \int F \sum_{j=1}^{m} c_j \operatorname{div}(v_j) \\ &= \sum_{j=1}^{m} c_j \int F \operatorname{div}(v_j) \\ &= \sum_{j=1}^{m} c_j \left(\int \operatorname{div}(F v_j) - \int \langle \nabla F, v_j \rangle \right) \\ &= -\int F \left\langle \sum_{i=1}^{m} c_i v_i, \sum_{j=1}^{m} c_j v_j \right\rangle. \end{aligned}$$

The use of the divergence theorem is justified in eliminating $\int \operatorname{div}(Fv_j)$ since the integrand is sufficiently smooth and decays rapidly at infinity. For details, see [Val]. Now, note that

$$v_j(x) = \frac{\nabla g_j(x)}{g_j(x)} = \frac{\nabla(f_j \circ B_j)(x)}{(f_j \circ B_j)(x)} = \frac{B_j^* \nabla f_j(B_j x)}{f_j(B_j x)} = B_j^* \nabla h_j(B_j x)$$

where $h_j = \log f_j$. Then we have

$$\begin{aligned} \frac{\partial}{\partial t} \int F(x,t)dx &= \int F(x,t) \sum_{j=1}^m c_j \langle B_j^* \nabla h_j(B_j x), B_j^* \nabla h_j(B_j x)\rangle dx \\ &\quad - \int F(x,t) \left\langle \sum_{i=1}^m c_i B_i^* \nabla h_i(B_i x), \sum_{j=1}^m c_j B_j^* \nabla h_j(B_j x) \right\rangle dx. \end{aligned}$$

Let $T : H \to \oplus_j H_j$ be the map

$$T = \begin{bmatrix} \dots & \sqrt{c_1} B_1 & \dots \\ & \vdots & \\ \dots & \sqrt{c_m} B_m & \dots \end{bmatrix} \tag{7}$$

and $A(x) \in \oplus_j H_j$ be the vector

$$A(x) = \begin{bmatrix} \sqrt{c_1} \nabla h_1(B_1 x) \\ \vdots \\ \sqrt{c_m} \nabla h_m(B_m x) \end{bmatrix} \tag{8}$$

Note that $T^*T = \sum_j c_j B_j^* B_j = \mathrm{Id}_H$ by geometricity, so $P = TT^* = T(T^*T)^{-1}T^*$ is an orthogonal projection. Then the difference above can be rewritten more succinctly as

$$\begin{aligned} \frac{\partial}{\partial t} \int F(x,t)dx &= \int F(x,t) \left[\langle A(x), A(x)\rangle - \langle T^*A(x), T^*A(x)\rangle\right] dx \\ &= \int F(x,t) \langle (I-P)A(x), A(x)\rangle. \end{aligned}$$

Since P is an orthogonal projection, $I - P$ is positive semidefinite, and so we deduce that $\int F(x,t)dx$ is monotonic in time. $\square$

Lemma 1.4. *The Gaussian integral*

$$g_j(x,t) = \int_H \frac{1}{(4\pi t)^{n/2}} e^{-\frac{\|x-y\|^2}{4t}} f_j(B_j y) dy$$

solves (6).

Proof. Verifying the relationship between the derivatives is a straightforward exercise in differentiation, so it remains to check that $g_j(x,0) = f_j(B_j x)$. Using the orthogonal decomposition $H = H_j \oplus H_j^\perp$ and $B_j^\perp$ to denote the orthogonal prrojection onto H_j, write

$$g_j(x,t) = \int_{H_j} \int_{H_j^\perp} \frac{1}{(4\pi t)^{n/2}} e^{-\frac{\|B_j x - y_1\|^2 + \|B_j^\perp - y_2\|^2}{4t}} f_j(y_1) dy_2 dy_1.$$

Using the change of variables $y_2 \mapsto y_2 - B_j^\perp x$ this becomes

$$\int_{H_j} \left(\int_{H_j^\perp} \frac{1}{(4\pi t)^{n/2}} e^{-\frac{\|y_2\|^2}{4t}} dy_2 \right) e^{-\frac{\|B_j x - y_1\|^2}{4t}} f_j(y_1) dy_1 = \int_{H_j} \frac{1}{(4\pi t)^{n/2}} e^{-\frac{\|B_j x - y_1\|^2}{4t}} f_j(y_1) dy_1.$$

Now, this is just the standard convolution with a Gaussian kernel in the variable $B_j x$, and since Gaussian kernels are approximations to the identity, we conclude that $\lim_{t\to 0} g_j(x,t) = f_j(B_j x)$. $\square$

Lemma 1.5. *We have*

$$\lim_{t\to 0^+}\int F(x,t)dx \geq \int \prod_{j=1}^m f_j(B_jx)^{c_j}\,dx$$
$$\lim_{t\to\infty}\int F(x,t)dx = \prod_{j=1}^m \left(\int_{H_j} f_j\right)^{c_j}$$

Remark. The inequality above is actually an equality, but we will not need this fact.

Proof. By definition we have

$$\lim_{t\to 0} F(x,t) = \prod_{j=1}^m f_j(B_jx)^{c_j}.$$

Then by Fatou's lemma, we get

$$\lim_{t\to 0}\int F(x,t)dx \geq \int \lim F(x,t)dx = \int \prod_{j=1}^m f_j(B_jx)^{c_j}\,dx.$$

For the limit at infinity, let g_j be as in Lemma 1.4. Using the scaling condition (3) and the form of g_j in the proof of Lemma 1.4, we have

$$\begin{aligned}\int_H \prod_{j=1}^m g_j(x,t)^{c_j}\,dx &= \int_H \prod_{j=1}^m \left(t^{n_j/2} g_j(t^{1/2}w,t)\right)^{c_j} dw\\ &= \int_H \prod_{j=1}^m \left(\int_{H_j} \frac{1}{(4\pi)^{n_j/2}} e^{-\frac{\|B_jw - t^{-1/2}y_1\|^2}{4}} f_j(y_1)dy_1\right)^{c_j} dw.\end{aligned}$$

Now, (5) implies that $\cap_{j=1}^m \ker(B_j) = \{0\}$, so we may apply dominated convergence to get

$$\begin{aligned}\lim_{t\to\infty}\int_H \prod_{j=1}^m g_j(x,t)^{c_j}\,dx &= \int_H \prod_{j=1}^m \left(\frac{1}{(4\pi)^{n_j/2}} e^{-\frac{\|B_jw\|^2}{4}}\right)^{c_j} dw \prod_{j=1}^m \left(\int_{H_j} f_j\right)^{c_j}\\ &= \int_H \frac{1}{(4\pi)^{n/2}} e^{-\frac{1}{4}\langle \sum_{j=1}^m c_jB_j^*B_jw,w\rangle} dw \prod_{j=1}^m \left(\int_{H_j} f_j\right)^{c_j}\\ &= \prod_{j=1}^m \left(\int_{H_j} f_j\right)^{c_j},\end{aligned}$$

where we have used the scaling condition (3) and the geometric condition (5) to simplify the Gaussian term. $\square$

Theorem 1.6 (Geometric Brascamp-Lieb inequality)**.** *Let $(\boldsymbol{B},\boldsymbol{c})$ be a geometric Brascamp-Lieb datum. Then*

$$\mathrm{BL}(\boldsymbol{B},\boldsymbol{c}) = \mathrm{BL}_g(\boldsymbol{B},\boldsymbol{c}) = 1.$$

Furthermore, $(\boldsymbol{B},\boldsymbol{c})$ is Gaussian-extremizable and therefore also extremizable.

Proof. Putting together Lemmas 1.3 and 1.5, we get

$$\int_H \prod_{j=1}^m f_j(B_jx)^{c_j}\,dx \leq \lim_{t\to 0}\int F(x,t)dx \leq \lim_{t\to\infty}\int F(x,t)dx = \prod_{j=1}^m \left(\int_{H_j} f_j\right)^{c_j}.$$

It follows that $\mathrm{BL}(\mathbf{B},\mathbf{c}) = 1$. To see that $\mathrm{BL}_g(\mathbf{B},\mathbf{c}) \geq 1$, simply consider the Gaussian input with covariances $(\mathrm{Id}_{H_j})_{1\leq j\leq m}$. $\square$

Every geometric datum is Gaussian extremizable; it turns out that the converse is true, up to a notion of equivalence. We say that two data $(H, (H_j), (B_j), (c_j))$ and $(H', (H_j'), (B_j'), (c_j'))$ are *equivalent* if there exist invertible linear maps $C : H' \to H$ and $C_j : H_j' \to H_j$ such that $B_j' = C_j^{-1} B_j C$ and $c_j' = c_j$ for all $1 \le j \le m$. We call C and C_j the *intertwining operators* of the two data. Note that from a change of variables $x = C^{-1}y$, we have the following relations:

$$\mathrm{BL}(\mathbf{B}', \mathbf{c}') = \frac{\prod_{j=1}^m |\det_{H_j' \to H_j} C_j|^{c_j}}{|\det_{H' \to H} C|} \mathrm{BL}(\mathbf{B}, \mathbf{c}) \tag{9}$$

$$\mathrm{BL}_g(\mathbf{B}', \mathbf{c}') = \frac{\prod_{j=1}^m |\det_{H_j' \to H_j} C_j|^{c_j}}{|\det_{H' \to H} C|} \mathrm{BL}_g(\mathbf{B}, \mathbf{c}) \tag{10}$$

Proposition 1.7 (Gaussian extremizability is equivalent to geometricity)**.** *Let $\boldsymbol{A} = (A_j)_{1 \le j \le m}$ be a Gaussian input for $(\boldsymbol{B}, \boldsymbol{c})$. Define $M : H \to H$ to be the positive semi-definite transformation $M := \sum_{j=1}^m c_j B_j^* A_j B_j$. Then the following are equivalent:*

(a) $(\boldsymbol{B}, \boldsymbol{c})$ is Gaussian extremizable with $\mathrm{BL}_g(\boldsymbol{B}, \boldsymbol{c}) = \mathrm{BL}_g(\boldsymbol{B}, \boldsymbol{c}; \boldsymbol{A})$.

(b) $(\boldsymbol{B}, \boldsymbol{c})$ is equivalent to a geometric datum $(\boldsymbol{B}', \boldsymbol{c}')$ with intertwining operators $C = M^{-1/2}$ and $C_j = A_j^{-1/2}$, and

$$\mathrm{BL}(\boldsymbol{B}, \boldsymbol{c}) = \mathrm{BL}_g(\boldsymbol{B}, \boldsymbol{c}) = \mathrm{BL}_g(\boldsymbol{B}, \boldsymbol{c}; \boldsymbol{A}).$$

Proof. Suppose (a) holds with a Gaussian extremizer $\mathbf{A}$. Then $\mathbf{A}$ is a local maximizer of the quantity

$$\sum_{j=1}^m c_j \log \det_{H_j} A_j - \log \det_H \sum_{j=1}^m c_j B_j^* A_j B_j.$$

Now, fix a $j \in \{1, \dots, m\}$, and consider perturbing A_j by some ϵQ_j, where $Q_j : H_j \to H_j$ is an arbitrary self-adjoint map and $\epsilon > 0$ is arbitrary. Then we have

$$\frac{d}{d\epsilon} \left(c_j \log \det_{H_j}(A_j + \epsilon Q_j) - \log \det_H (M + \epsilon c_j B_j^* Q_j B_j) \right)\Big|_{\epsilon=0} = 0.$$

Subtracting off the constant terms $c_j \log \det_{H_j} A_j - \log \det_H M$ we obtain

$$\frac{d}{d\epsilon} \left(c_j \log \det_{H_j}(\mathrm{Id}_{H_j} + \epsilon A_j^{-1} Q_j) - \log \det_H (\mathrm{Id}_H + \epsilon c_j M^{-1} B_j^* Q_j B_j) \right)\Big|_{\epsilon=0} = 0.$$

From Taylor expansion we have

$$\frac{d}{d\epsilon} \log \det_H (\mathrm{Id}_H + A)|_{\epsilon=0} = \mathrm{tr}_H(A).$$

Therefore our first order condition comes down to

$$\mathrm{tr}_{H_j}((A_j^{-1} - B_j M^{-1} B_j^*) Q_j) = 0.$$

Since Q_j was arbitrary, we must have $A_j^{-1} = B_j M^{-1} B_j^*$ for each $1 \le j \le m$. With $B_j' := A_j^{1/2} B_j M^{-1/2}$, we get that $B_j' B_j'^* = \mathrm{Id}_{H_j}$. Furthermore, we have

$$\sum_{j=1}^m c_j B_j'^* B_j' = \sum_{j=1}^m M^{-1/2} c_j B_j^* A_j B_j M^{-1/2} = M^{-1/2} M M^{-1/2} = \mathrm{Id}_H,$$

so $(\mathbf{B}', \mathbf{c}')$ is geometric. Thus by Theorem 1.6, (9) and (10), we conclude that

$$\mathrm{BL}(\mathbf{B}, \mathbf{c}) = \mathrm{BL}_g(\mathbf{B}, \mathbf{c}) = \mathrm{BL}_g(\mathbf{B}, \mathbf{c}; \mathbf{A}).$$

The converse (b) $\implies$ (a) is trivial. $\square$

1.3. Finiteness. In this section we prove Theorems 1.1 and 1.2. The proof of necessity of (4) and (3) is easy and follows from a scaling argument. To show sufficiency, we will develop some structural theory for Brascamp-Lieb data, from which the proof follows by induction. In particular, the idea is to repeatedly break down the datum until it is "indecomposable", at which point the claim follows trivially.

Proof of necessity of (3) and (4). Assume that $\mathrm{BL}(\mathbf{B},\mathbf{c})<\infty$. Then $\mathrm{BL}_g(\mathbf{B},\mathbf{c})<\infty$. If we consider the Gaussian input $(t\,\mathrm{Id}_{H_j})_{1\le j\le m}$, then we have

$$\mathrm{BL}_g(\mathbf{B},\mathbf{c}) \ge \frac{t^{\frac{1}{2}\sum_{j=1}^m c_j \dim(H_j)-\dim(H)}}{\det\left(\sum_{j=1}^m B_j^* B_j\right)^{1/2}}.$$

Letting $t\to\infty$, we see that (3) must hold, otherwise the right hand side blows up. Next, consider the Gaussian input $(A_j)_{1\le j\le m}$ given by $A_j=(\epsilon\,\mathrm{Id}_{B_jV})\oplus\mathrm{Id}_{(B_jV)^\perp}$. Then $\det_{H_j}(A_j)$ decays as $\epsilon^{\dim(B_jV)}$ as $\epsilon\to0$. Also note that $\sum_{j=1}^m B_j^*A_jB_j$ is bounded uniformly in ϵ, and decays linearly in ϵ when restricted to V. Thus $\det_H(\sum_{j=1}^m B_j^*A_jB_j)=O(\epsilon^{\dim(V)})$. Therefore, we have

$$\mathrm{BL}_g(\mathbf{B},\mathbf{c})=\Omega\left(\epsilon^{\frac{1}{2}\sum_{j=1}^m c_j\dim(B_jV)-\dim(V)}\right),$$

which implies the necessity of (4). □

Now, define a *critical subspace* V for $(\mathbf{B},\mathbf{c})$ to be a non-zero proper subspace of H satisfying

$$\dim(V)=\sum_{j=1}^m c_j\dim(B_jV).$$

We say that the datum $(\mathbf{B},\mathbf{c})$ is *simple* if it has no critical subspaces. Next, for any subspace $V\subset H$ define the restriction $\mathbf{B}_V$ of $\mathbf{B}$ to be

$$\mathbf{B}_V=(V,(B_jV)_{1\le j\le m},(B_j|_V)_{1\le j\le m}).$$

Also define the quotient $\mathbf{B}_{H/V}$ of $\mathbf{B}$ to be

$$\mathbf{B}_{H/V}=(H/V,(H_j/(B_jV))_{1\le j\le m},(B_j|_{H/V})_{1\le j\le m}).$$

The following lemmas are shown in [BCCT].

Lemma 1.8 (Necessary conditions split)**.** *Let V be a critical subspace. Then $(\boldsymbol{B},\boldsymbol{c})$ satisfies conditions (3) and (4) if and only if both $(\boldsymbol{B}_V,\boldsymbol{c})$ and $(\boldsymbol{B}_{H/V},\boldsymbol{c})$ satisfy the conditions (3) and (4) for all subspaces of U of V and H/V, respectively.*

Lemma 1.9 (Brascamp-Lieb constants split)**.** *If V is a critical subspace, then*

$$\mathrm{BL}(\boldsymbol{B},\boldsymbol{c})=\mathrm{BL}(\boldsymbol{B}_V,\boldsymbol{c})\mathrm{BL}(\boldsymbol{B}_{H/V},\boldsymbol{c})$$
$$\mathrm{BL}_g(\boldsymbol{B},\boldsymbol{c})=\mathrm{BL}_g(\boldsymbol{B}_V,\boldsymbol{c})\mathrm{BL}_g(\boldsymbol{B}_{H/V},\boldsymbol{c})$$

Lemma 1.10. *Let $(\boldsymbol{B},\boldsymbol{c})$ be a Brascamp-Lieb datum for which (3) and (4) hold. Then $\mathrm{BL}_g(\boldsymbol{B},\boldsymbol{c})<\infty$. Furthermore, if $(\boldsymbol{B},\boldsymbol{c})$ is simple, then $(\boldsymbol{B},\boldsymbol{c})$ is Gaussian-extremizable.*

Proof of sufficiency of (3) and (4). Assume (3) and (4) hold. We induct on the dimension $\dim(H)$. When $\dim(H)=0$ the claim is vacuously true. Now suppose $\dim(H)>0$, and that the claim has been proven for smaller values of $\dim(H)$.

By Lemma 1.10, we know that $\mathrm{BL}_g(\mathbf{B},\mathbf{c})<\infty$, so it suffices to show that $\mathrm{BL}(\mathbf{B},\mathbf{c})=\mathrm{BL}_g(\mathbf{B},\mathbf{c})$. First suppose that $(\mathbf{B},\mathbf{c})$ is simple. Then Lemma 1.10 tells us $(\mathbf{B},\mathbf{c})$ is Gaussian-extremizable, and so the claim follows from Proposition 1.7. Next consider the case where $(\mathbf{B}, \mathbf{c})$ is not simple. Then there is a criticial subspace, and we can split the datum $(\mathbf{B}, \mathbf{c})$ into $(\mathbf{B}_V, \mathbf{c})$ and $(\mathbf{B}_{H/V}, \mathbf{c})$. By Lemma 1.8 and the induction hypothesis, we have

$$\mathrm{BL}(\mathbf{B}_V, \mathbf{c}) = \mathrm{BL}_g(\mathbf{B}_V, \mathbf{c}) < \infty$$
$$\mathrm{BL}(\mathbf{B}_{H/V}, \mathbf{c}) = \mathrm{BL}_g(\mathbf{B}_{H/V}, \mathbf{c}) < \infty.$$

Therefore, by Lemma 1.9, we conclude that $\mathrm{BL}(\mathbf{B}, \mathbf{c}) = \mathrm{BL}_g(\mathbf{B}, \mathbf{c}) < \infty$. □

1.4. Extremizability. The goal of this section is to discuss extremizers and prove the equivalence of extremizability and Gaussian extremizability. We start with some closure properties of extremizers:

Lemma 1.11. *Suppose* $\mathrm{BL}(\boldsymbol{B}, \boldsymbol{c}) < \infty$, *with extremizers* $(f_j)_{1\leq j\leq m}$, $(f_j')_{1\leq j\leq m}$.

- *(Scale invariance) For any nonzero real* λ, $(f_j(\lambda\cdot))_{1\leq j\leq m}$ *is also an extremizer.*
- *(Homogeneity) For any nonzero real numbers* $c_1, \ldots, c_m$, $(c_j f_j)_{1\leq j\leq m}$ *is also an extremizer.*
- *(Translation invariance) For any* $x_0 \in H$, $(f_j(\cdot - B_j x_0))_{1\leq j\leq m}$ *is also an extremizer.*
- *(Limit in* L^1*) If* $(f_j^{(n)})_{1\leq j\leq m}$ *is a sequence of extremizers which converge in* L^1 *to another input* $(f_j)_{1\leq j\leq m}$ *for each* j, *then* $(f_j)_{1\leq j\leq m}$ *is also an extremizer.*
- *(Convolution)* $(f_j * f_j')_{1\leq j\leq m}$ *is also an extremizer.*
- *(Multiplication) The input* $(f_j(\cdot - B_j x_0) f_j'(\cdot))_{1\leq j\leq m}$ *is an extremizer for almost every* $x_0 \in H$ *for which* $\int_H f_j(x - B_j x_0) f_j'(x)dx > 0$ *for all* j.

Theorem 1.12. *A Brascamp-Lieb datum* $(\boldsymbol{B}, \boldsymbol{c})$ *is extremizable if and only if it is Gaussian extremizable.*

Proof. The idea is to start with an extremizer $\mathbf{f}$ and repeatedly convolve, rescale, and multiply so that we may apply the central limit theorem along with the closure properties of extremizers listed above to obtain a Gaussian extremizer. The technical part of the proof is to verify enough moment conditions to apply the central limit theorem. For details on this see [BCCT, Proposition 6.5].

Now, set $\mathbf{f}^{(n)} = (f_j^{(n)})_{1\leq j\leq m}$ to be the rescaled n-fold convolution

$$f_j^{(n)}(x) := n^{(\dim H_j)/2} f_j * \cdots * f_j(\sqrt{n}x).$$

The central limit theorem says that each $f_j^{(n)}$ converges weakly to a centered Gaussian. The additional moment conditions allow us to apply Berry-Esseen's quantitative central limit theorem to upgrade this to L^1 convergence. By Lemma 1.11 we have obtained a centered Gaussian extremizer. □

2. The Reverse Brascamp-Lieb Inequality

2.1. The Brunn-Minkowski and Prékopa-Leindler inequalities. As a simple instance of the reverse Brascamp-Lieb inequality, or sometimes referred to as Barthe's inequality, we start with the Prékopa-Leindler inequality. This is closely related to the Brunn-Minkowski inequality for measurable sets, see [Gar]. For two sets X and Y, we use $X+Y$ to denote their *Minkowski sum*:

$$X + Y = \{x + y : x \in X, y \in Y\}.$$

For a $c \in \mathbb{R}$, we use cX to denote the *dilation* of X:

$$cX = \{cx : x \in X\}.$$

If X is a compact measurable set in $\mathbb{R}^n$, we use $V_n(X)$ to denote its *volume*, i.e. V_n is the Lebesgue measure on $\mathbb{R}^n$.

Theorem 2.1 (Brunn-Minkowski in $\mathbb{R}$)**.** *Let $0 < \lambda < 1$ and let X and Y be nonempty measurable sets in $\mathbb{R}$ such that the convex combination $(1-\lambda)X + \lambda Y$ is also measurable. Then*

$$V_1((1-\lambda)X + \lambda Y) \geq (1-\lambda)V_1(X) + \lambda V_1(Y). \tag{11}$$

Proof. Note that by approximation from within, it suffices to show the inequality for compact sets X and Y. If X and Y are compact, then so is $X+Y$. Without loss of generality, we may translate X and Y so that $X \cap Y = \{0\}$, $X \subset \mathbb{R}_{\leq 0}$, and $Y \subset \mathbb{R}_{\geq 0}$. Then clearly $X \cup Y \subset X + Y$, so

$$V_1(X+Y) \geq V_1(X \cup Y) = V_1(X) + V_1(Y).$$

Replacing X and Y with $(1-\lambda)X$ and λY, respectively, gives us (11). □

Theorem 2.2 (PL)**.** *Let $0 < \lambda < 1$ and $f, g, h \geq 0$ be integrable functions on $\mathbb{R}$ satisfying*

$$h((1-\lambda)x + \lambda y) \geq f(x)^{1-\lambda} g(y)^{\lambda}$$

for all $x, y \in \mathbb{R}$. Then

$$\int_{\mathbb{R}} h(x)dx \geq \left(\int_{\mathbb{R}} f(x)dx\right)^{1-\lambda} \left(\int_{\mathbb{R}} g(x)dx\right)^{\lambda}.$$

We write down two proofs of this inequality: the first following from an application of (11), and the second a precursor to the transportation methods which we will see later in the proofs of the more general Barthe's inequality.

First proof. Let $L(f,t) = \{x : f(x) \geq t\}$ denote the super-level set of f at level t. By Fubini's theorem, we may write

$$\int f(x)dx = \int \int_0^{f(x)} 1dtdx = \int_0^{\infty} \int_{L(f,t)} 1dxdt = \int_0^{\infty} V_n(L(f,t))dt.$$

By a straightforward application of the dominated convergence theorem, we may assume that $f, g \in L^{\infty}$. Then, by homogeneity we may assume that $\sup_{x\in\mathbb{R}} f(x) = \sup_{x\in\mathbb{R}} g(x) = 1$. If $0 \leq t < 1$ and $f(x), g(y) \geq t$, then we have $h((1-\lambda)x + \lambda y) \geq t$. That is,

$$L(h,t) \supset (1-\lambda)L(f,t) + \lambda L(g,t).$$

Therefore, we may apply (11) and the AM-GM inequality to obtain

$$\begin{aligned}
\int h(x)dx &\geq \int_0^1 V_1(L(h,t))dt \\
&\geq \int_0^1 V_1((1-\lambda)L(f,t)+\lambda L(g,t))dt \\
&\geq (1-\lambda)\int_0^1 V_1(L(f,t))dt + \lambda\int_0^\infty V_1(L(g,t))dt \\
&= (1-\lambda)\int f(x)dx + \lambda\int g(x)dx \\
&\geq \left(\int f(x)dx\right)^{1-\lambda}\left(\int g(x)dx\right)^{\lambda},
\end{aligned}$$

which is what we wanted. □

Second proof. By homogeneity, we can assume that

$$\int_{\mathbb{R}} f(x)dx = \int_{\mathbb{R}} g(x)dx = 1.$$

Define $u, v : [0,1] \to \mathbb{R}$ such that $u(t)$ and $v(t)$ are the smallest numbers satisfying

$$\int_{-\infty}^{u(t)} f(x)dx = \int_{-\infty}^{v(t)} g(x)dx = t.$$

Note that $u(t)$ and $v(t)$ are strictly increasing functions, so they are differentiable almost everywhere. Differentiate the above with respect to t to get

$$f(u(t))u'(t) = g(v(t))v'(t) = 1$$

almost everywhere. Now set

$$w(t) = (1-\lambda)u(t) + \lambda v(t),$$

so that by AM-GM we obtain

$$\begin{aligned}
w'(t) &= (1-\lambda)u'(t) + \lambda v'(t) \\
&\geq u'(t)^{1-\lambda}v'(t)^{\lambda} \\
&= f(u(t))^{-(1-\lambda)}g(v(t))^{-\lambda}
\end{aligned}$$

whenever $f(u(t)) \neq 0$ and $g(v(t)) \neq 0$. Thus, we have

$$\begin{aligned}
\int_{\mathbb{R}} h(x)dx &\geq \int_0^1 h(w(t))w'(t)dt \\
&\geq \int_0^1 f(u(t))^{1-\lambda}g(v(t))^{\lambda}f(u(t))^{-(1-\lambda)}g(v(t))^{-\lambda}dt = 1,
\end{aligned}$$

proving our claim. □

Remark. The Prékopa-Leindler inequality can also be written in one line in the form

$$\int \sup\{f(x)^{1-\lambda}g(y)^{\lambda} : (1-\lambda)x + \lambda y = z\}dz \geq \left(\int f(x)dx\right)^{1-\lambda}\left(\int g(x)dx\right)^{\lambda}.$$

2.2. Barthe's inequality & transportation argument. Let $m \geq n$. Fix $(c_i)_{i=1}^m$ to be positive reals and $(n_i)_{i=1}^m$ integers such that

$$\sum_{i=1}^m c_i n_i = n.$$

Let $B_i : \mathbb{R}^n \to \mathbb{R}^{n_i}$ be a linear surjective map, such that $\cap_{i=1}^m \ker B_i = \{0\}$. For f_i nonnegative, integrable functions define

$$I((f_i)_{i=1}^m) := \int_{\mathbb{R}^n} \sup \left\{ \prod_{i=1}^n f_i^{c_i}(y_i) : \sum_{i=1}^m c_i B_i^* y_i = x, y_i \in \mathbb{R}^{n_i} \right\} dx$$

and

$$J((f_i)_{i=1}^m) := \int_{\mathbb{R}^n} \prod_{i=1}^m f_i^{c_i}(B_i x) dx.$$

Define $E = \mathrm{RBL}(\mathbf{B}, \mathbf{c})$ to be the Reverse Brascamp-Lieb constant:

$$\mathrm{RBL}(\mathbf{B}, \mathbf{c}) := \inf \left\{ \frac{I((f_i)_{i=1}^m}{\prod_{i=1}^m \left(\int_{\mathbb{R}^{n_i}} f_i\right)^{c_i}} : \ f_i \text{ nonnegative, integrable on } \mathbb{R}^{n_i} \text{ with } \int f_i \neq 0,\ 1 \leq i \leq m \right\}$$

In other words, E is the largest constant such that for all such $(f_i)_{i=1}^m$,

$$I((f_i)_{i=1}^m) \geq E \prod_{i=1}^m \left(\int_{\mathbb{R}^{n_i}} f_i \right)^{c_i}. \tag{12}$$

As with $F_g = \mathrm{BL}_g(\mathbf{B}, \mathbf{c})$, define $E_g = \mathrm{RBL}_g(\mathbf{B}, \mathbf{c})$ to be the Reverse Brascamp-Lieb constant when computed under only centered Gaussian inputs. Let $F = \mathrm{BL}(\mathbf{B}, \mathbf{c})$ be the Brascamp-Lieb constant, which is given by:

$$\mathrm{BL}(\mathbf{B}, \mathbf{c}) := \sup \left\{ \frac{J((f_i))_{i=1}^m}{\prod_{i=1}^m \left(\int_{\mathbb{R}^{n_i}} f_i\right)^{c_i}} : \ f_i \text{ nonnegative, integrable on } \mathbb{R}^{n_i} \text{ with } \int f_i \neq 0, 1 \leq i \leq m \right\}.$$

The Reverse Brascamp-Lieb inequality (12), due to [Bar], generalizes the Prékopa-Leindler inequality. In the following proof, based on mass transportation from [Bar], we will see that Barthe's inequality is dual to Brascamp-Lieb's inequality, in the sense that their extremizing constants are reciprocals of each other.

Theorem 2.3 (Barthe)**.** *The constants E and F can be computed with centered Gaussians:*

$$\mathrm{RBL}(\boldsymbol{B}, \boldsymbol{c}) = \mathrm{RBL}_g(\boldsymbol{B}, \boldsymbol{c}),$$
$$\mathrm{BL}(\boldsymbol{B}, \boldsymbol{c}) = \mathrm{BL}_g(\boldsymbol{B}, \boldsymbol{c}).$$

Furthermore, if we define

$$D = \sup \left\{ \frac{\det\left(\sum_{i=1}^m c_i B_i^* A_i B_i\right)}{\prod_{i=1}^m (\det A_i)^{c_i}} : A_i \in \mathcal{S}^+(\mathbb{R}^{n_i}) \right\},$$

then $E = D^{1/2}$ and $F = D^{-1/2}$.

We divide the proof of Theorem 2.3 into two parts. The first is to show that $E_g \cdot F_g = 1$. Clearly by (1), we have that $F_g = D^{1/2}$. Therefore, for the second half it would suffice to show that $I((f_i)_{i=1}^m) \geq D \cdot J((h_i)_{i=1}^m)$ for all $f_i, g_i \in L_1^+(\mathbb{R}^{n_i})$ satisfying $\int_{\mathbb{R}^{n_i}} f_i = \int_{\mathbb{R}^{n_i}} h_i$.

Lemma 2.4. *We have $E_g = 0$ if and only if $F_g = +\infty$. Otherwise, we have*

$$E_g \cdot F_g = 1.$$

Proof. For $1 \le i \le m$ let $A_i \in \mathcal{S}^+(\mathbb{R}^{n_i})$ and define Q to be the quadratic form

$$Q(y) = \left\langle \sum_{i=1}^m c_i B_i^* A_i B_i y, y \right\rangle.$$

Also define Q^* to be the dual quadratic form of Q, given by

$$Q^*(x) = \sup\{|\langle x, y\rangle|^2 : Q(y) \le 1\}.$$

Now, we claim that $Q^* = R$ where R is defined by

$$R(x) = \inf\left\{ \sum_{i=1}^m c_i \langle A_i^{-1} x_i, x_i\rangle : \sum_{i=1}^m c_i B_i^* x_i = x,\ x_i \in \mathbb{R}^{n_i} \text{ for all } 1 \le i \le m \right\}.$$

Fix $x = \sum_{i=1}^m c_i B_i^* x_i$, so that

$$|\langle x, y\rangle|^2 = \left| \left\langle \sum_{i=1}^m c_i B_i^* x_i, y \right\rangle \right|^2 = \left| \sum_{i=1}^m \langle \sqrt{c_i} A_i^{-1/2} x_i, \sqrt{c_i} A_i^{1/2} B_i y\rangle \right|^2.$$

Then by Cauchy-Schwarz, applied to the quadratic form ϕ on $\mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_m}$ defined by

$$\phi(x_1, \ldots, x_m) = \sum_{i=1}^m \langle x_i, x_i\rangle,$$

we get

$$\begin{aligned} |\langle x, y\rangle|^2 &\le \left(\sum_{i=1}^m \left| \sqrt{c_i} A_i^{-1/2} x_i \right|^2 \right) \left(\sum_{i=1}^m \left| \sqrt{c_i} A_i^{1/2} B_i y \right|^2 \right) \\ &= \left(\sum_{i=1}^m c_i \langle x_i, A_i^{-1} x_i\rangle \right) \left(\left\langle \sum_{i=1}^m c_i B_i^* A_i B_i y, y \right\rangle \right). \end{aligned}$$

Furthermore, we have equality above when[1]

$$\begin{aligned} y &= \left(\sum_{i=1}^m c_i B_i^* A_i B_i \right)^{-1} x \\ x_i &= A_i B_i y \quad i = 1, \ldots, m. \end{aligned}$$

Therefore $R = Q^*$. Now, for centered Gaussians $g_{A_1}, \ldots, g_{A_m}$, we have

$$\begin{aligned} \frac{J(g_{A_1}, \ldots, g_{A_m})}{\prod_{i=1}^m \left(\int g_{A_i}\right)^{c_i}} &= \sqrt{\frac{\prod_{i=1}^m (\det A_i)^{c_i}}{\det Q}} \\ \frac{I(g_{A_1^{-1}}, \ldots, g_{A_m^{-1}})}{\prod_{i=1}^m \left(\int g_{A_i^{-1}}\right)^{c_i}} &= \sqrt{\frac{\prod_{i=1}^m (\det A_i)^{-c_i}}{\det R}}. \end{aligned}$$

[1]Here, the definition of y assumes $\cap_{i=1}^m \ker(B_i) = \{0\}$, which occurs if and only if $E_g > 0$.

By duality, we have $\det Q \cdot \det R = 1$, so we get

$$\frac{J(g_{A_1}, \dots, g_{A_m})}{\prod_{i=1}^m \left(\int g_{A_i}\right)^{c_i}} \cdot \frac{I(g_{A_1^{-1}}, \dots, g_{A_m^{-1}})}{\prod_{i=1}^m \left(\int g_{A_i^{-1}}\right)^{c_i}} = 1.$$

It follows that $E_g \cdot F_g = 1$. $\square$

Before we move on to the second part of the proof, we need to introduce the machinery of optimal transport. Recall in the second proof of the Prékopa-Leindler inequality, where we made use of mappings $u(t)$ and $v(t)$ such that

$$\int_{-\infty}^{u(t)} f(x)dx = \int_{-\infty}^{v(t)} g(x)dx.$$

One can think of u and v as transporting the mass between densities f and g. In higher dimensions, we will use the Brenier mapping, see [Bre].

Theorem 2.5. *Let f, h be non-negative measurable functions on $\mathbb{R}^n$ with $\int f = \int h = 1$. Then there exists a convex function ϕ on $\mathbb{R}^n$ such that the map $T = \nabla\phi$ transports mass from h to f, i.e. for every non-negative Borel function b on $\mathbb{R}^n$, we have*

$$\int_{\mathbb{R}^n} b(T(x))h(x)dx = \int_{\mathbb{R}^n} b(x)f(x)dx. \tag{13}$$

Furthermore, when $\phi \in C^2$, the map T is a solution to the equation:

$$\det(\nabla T(x))h(T(x)) = f(x). \tag{14}$$

Remark. In the proof of Theorem 2.5, we actually need the transport maps T to be differentiable, so that we can make use of the relation (14). It turns out that it suffices to reduce to the case of Lipschitz functions f, h, from which a regularity theorem of Caffarelli [Caf] gives us differentiability. For details on the reduction, see [Bar].

Lemma 2.6. *For $1 \leq i \leq m$, let $f_i, h_i \in L_1^+(\mathbb{R}^{n_i})$ satisfy $\int_{\mathbb{R}^{n_i}} f_i = \int_{\mathbb{R}^{n_i}} h_i$. Then*

$$I((f_i)_{i=1}^m) \geq D \cdot J((h_i)_{i=1}^m).$$

Proof. By homogeneity we may suppose that $\int f_i = \int h_i = 1$ for $1 \leq i \leq m$. Define Ω_i be the subset of the domain where h_i is positive. Let T_i be the differentiable Brenier map between f_i and h_i, so that we have

$$\det(dT_i(x)) \cdot f_i(T_i(x)) = h_i(x).$$

Furthermore, we know that $dT_i(x) \in \mathcal{S}^+(\mathbb{R}^{n_i})$ on Ω_i for each i. Define Θ on $S := \cap_{i=1}^m B_i^{-1}(\Omega_{h_i})$ by

$$\Theta(y) = \sum_{i=1}^m c_i B_i^*(T_i(B_i y)).$$

Its derivative is symmetric positive definite,

$$d\Theta(y) = \sum_{i=1}^m c_i B_i^* dT_i(B_i y) B_i,$$

since

$$\det\left(\sum_{i=1}^m c_i B_i^* dT_i(B_i y) B_i\right) \geq D \prod_{i=1}^m (\det dT_i(B_i y))^{c_i} > 0.$$

Now, write

$$\begin{aligned}
\int_{\mathbb{R}^n} \prod_{i=1}^m h_i^{c_i}(B_i y) dy &= \int_S \prod_{i=1}^m h_i^{c_i}(B_i y) dy \\
&= \int_S \prod_{i=1}^m (f_i(T_i(B_i y)) \det(dT_i(B_i y)))^{c_i} dy \\
&\leq \frac{1}{D} \int_S \prod_{i=1}^m f_i(T_i(B_i y))^{c_i} \det\left(\sum_{i=1}^m c_i B_i^* dT_i(B_i y) B_i\right) dy \\
&\leq \int_S \sup_{\Theta(y)=\sum_{i=1}^m c_i B_i^* x_i} \left\{ \left(\prod_{i=1}^m f_i(x_i)^{c_i}\right) \det(d\Theta(y)) \right\} dy \\
&= \frac{1}{D} \int_{\mathbb{R}^n} \sup_{x=\sum_{i=1}^m c_i B_i^* x_i} \left(\prod_{i=1}^m f_i(x_i)^{c_i}\right) dx,
\end{aligned}$$

which shows that $I((f_i)_{i=1}^m) \geq D \cdot J((h_i)_{i=1}^m)$. □

As a corollary of Theorem 2.3, we obtain the necessary and sufficient conditions for strict positivity of E. In particular, the relation $E = F^{-1}$ implies that the conditions are the same for the Reverse Brascamp-Lieb inequality as they are for the Forward Brascamp-Lieb inequality, i.e. (3) and (4).

The equivalence of geometricity and Gaussian extremizability still holds for the reverse inequality. A quick way to see this is by combining Proposition 1.7 and Theorem 2.3. Alternatively, one could replicate the proof in Proposition 1.7.

The question of whether extremizability is equivalent to Gaussian extremizability is still unknown, however it is known that extremizers of the reverse inequality are preserved under convolution, so one could potentially use an argument similar to that used in the proof of Theorem 1.12, as suggested by [BKX].

Remark. The Reverse Brascamp-Lieb inequality in the rank one case has also been proved using heat semigroup techniques by Barthe, Cordero-Erausquin [BarCor]. In the general case, both the forward and reverse inequality have been proved using heat semigroup techniques by Barthe, Huet [BarHu]. In the next section, we introduce probabilistic arguments for showing both the Forward and Reverse Brascamp-Lieb inequalities, due to Lehec [Leh2].

3. The Stochastic Perspective

3.1. Forward-Reverse Brascamp-Lieb inequalities. Let $(E_i)_{i=1}^n$ and $(E^j)_{j=1}^m$ be Euclidean spaces and define $E_0 = \oplus_{i=1}^n E_i$. Let $\pi_{E_i} : E_0 \to E_i$ be the orthogonal projection of E_0 onto E_i, and $\boldsymbol{B} := (B_{ij})_{1\le i\le n, 1\le j\le m}$, where each $B_{ij} : E_i \to E^j$ is a linear map. Also define $B_j : E_0 \to E^j$ by

$$B_j x := \sum_{i=1}^n B_{ij}\pi_{E_i}(x).$$

In particular, the collection $(B_j)_{j=1}^m$ gives an equivalent representation of $\boldsymbol{B}$.

Suppose $\boldsymbol{c} := (c_i)_{i=1}^n$ and $\boldsymbol{d} := (d_j)_{j=1}^m$ are collections of positive real numbers satisfying the *dimension condition*

$$\sum_{i=1}^n c_i \dim(E_i) = \sum_{j=1}^m d_j \dim(E^j). \tag{15}$$

Then we refer to $(\boldsymbol{c}, \boldsymbol{d}, \boldsymbol{B})$ as a *datum*.

Given a datum $(\boldsymbol{c}, \boldsymbol{d}, \boldsymbol{B})$, we are interested in the best constant D in the following statement: If measurable $f_i : E_i \to \mathbb{R}^+$, $1 \le i \le k$ and $g_j : E^j \to \mathbb{R}^+$, $1 \le j \le m$ satisfy

$$\prod_{i=1}^n f_i^{c_i}(x_i) \le \prod_{j=1}^m g_j^{d_j}\left(\sum_{i=1}^n c_i B_{ij} x_i\right) \qquad \forall x_i \in E_i,\ 1 \le i \le n, \tag{16}$$

then

$$\prod_{i=1}^n \left(\int_{E_i} f_i\right)^{c_i} \le e^D \prod_{j=1}^m \left(\int_{E^j} g_j\right)^{d_j}. \tag{17}$$

Definition 3.1. We define $D(\boldsymbol{c}, \boldsymbol{d}, \boldsymbol{B})$ to be the smallest constant D such that (17) holds for all nonnegative measurable functions satisfying the constraints (16). Together, we may refer to (16) and (17) as the *Forward-Reverse Brascamp-Lieb inequality*, or forward-reverse inequality, or FRBL for short.

Theorem 3.1 (Geometric FRBL)**.** *Let $(\boldsymbol{c}, \boldsymbol{d}, \boldsymbol{B})$ be a datum, satisfying the dimension condition (15), as well as the geometric condition: for some $\Sigma \in \Pi(\mathrm{Id}_{E_1}, \dots, \mathrm{Id}_{E_n})$, i.e. operators $\Sigma : E_0 \to E_0$ whose restriction onto each E_i is equal to Id_{E_i} for $1 \le i \le n$,*

$$B_j \Sigma B_j^* = \mathrm{Id}_{E^j} \qquad 1 \le j \le m$$
$$\sum_{j=1}^m d_j B_j^* B_j \le \Lambda_c := \bigoplus_{i=1}^n c_i \,\mathrm{Id}_{E_i}$$

The first constraint says that B_j is an orthogonal projection under the inner product defined with respect to Σ. The second constraint is also known as the frame condition. If measurable maps $f_i : E_i \to \mathbb{R}^+$, $1 \le i \le n$ and $g_j : E^j \to \mathbb{R}^+$, $1 \le j \le m$ satisfy

$$\prod_{i=1}^n f_i^{c_i}(\pi_{E_i}(x)) \le \prod_{j=1}^m g_j^{d_j}(Q_j x) \qquad \forall x \in E_0, \tag{18}$$

then

$$\prod_{i=1}^n \left(\int_{E_i} f_i\right)^{c_i} \le \prod_{j=1}^m \left(\int_{E^j} g_j\right)^{d_j}. \tag{19}$$

That is, the optimal constant satisfies $e^{D(\boldsymbol{c}, \boldsymbol{d}, \boldsymbol{B})} = 1$.

Remark. By taking $n = 1$, $c_1 = 1$ we get the forward inequality. On the other hand, taking $m = 1$ and $d_1 = 1$ gives us the reverse inequality.

3.2. Entropic formulations. In what follows, it will be convenient to recast our problems as entropic ones. This was done by [CaCo] for Brascamp-Lieb inequalities. We shall extend their work to the more general Forward-Reverse Brascamp-Lieb inequalities. Let f be the probability density of a random variable $X \sim \mu$ on $\mathbb{R}^n$, and let m be a reference measure with $\mu \ll m$. We define the *entropy* of X (or sometimes we say of f or μ) to be

$$h(\mu) := -\int f(x) \ln f(x) dx$$

and the *relative entropy* of μ with respect to m by

$$D(\mu \| m) := \int \frac{d\mu}{dm} \ln \left(\frac{d\mu}{dm} \right) dm,$$

where $\frac{d\mu}{dm}$ is the Radon-Nikodym derivative. Note that if $m = \lambda$ is the Lebesgue measure, $D(\mu\|\lambda) = -h(\mu)$. For a map B onto $\mathbb{R}^n$, we use $f_{(B)}$ to denote the density of the push-forward measure $B\#\mu := \mu \circ B^{-1}$.

Theorem 3.2 (Carlen & Cordero)**.** *Suppose the datum $(\boldsymbol{c}, \boldsymbol{B})$ is given. Then for any $D \in \mathbb{R}$, the following two assertions are equivalent:*

(i) For all probability densities f with $|h(f)| < \infty$, we have

$$-\sum_{j=1}^{m} d_j h(f_{(B_j)}) \leq -h(f) + D. \tag{20}$$

(ii) For any m nonnegative functions $(f_j)_{j=1}^m$ on H_i, we have

$$\int_{\mathbb{R}^n} \prod_{j=1}^{m} f_j^{d_j}(B_i x) dx \leq e^D \prod_{j=1}^{m} \left(\int_{\mathbb{R}} f_j(y_j) dy_j \right)^{d_j}. \tag{21}$$

Remark 3.2. Assuming finite second moments and a geometric datum, the entropic form (20) can be rewritten in terms of relative entropies as

$$\sum_{j=1}^{m} d_j D(B_j\#\mu \| \gamma_{E^j}) \leq D(\mu\|\gamma), \tag{22}$$

where γ_{E^j} is the standard Gaussian measure on $\operatorname{ran} B_j = E^j$, given by density

$$f_{E^j}(x) = (2\pi)^{-\dim(E^j)/2} e^{-|x|^2/2}, \quad x \in E^j.$$

To see this, let $m = \gamma$ be a standard Gaussian measure on $\mathbb{R}^n$. Then we have

$$\begin{aligned} D(\mu\|\gamma) &= \int f(x) \ln f(x) dx - \int f(x) \ln \left(\frac{1}{(2\pi)^{n/2}} e^{-\|x\|_2^2/2} \right) dx \\ &= -h(\mu) + n \ln \sqrt{2\pi} + \frac{1}{2} \mathbb{E}_\mu \|X\|_2^2 \end{aligned}$$

Hence (20) becomes

$$\sum_{j=1}^{m} d_j \left[D(B_j\#\mu\|\gamma_{E^j}) - \dim(H_j) \ln \sqrt{2\pi} - \frac{1}{2} \mathbb{E}_\mu \|B_j X\|_2^2 \right] \leq D(\mu\|\gamma) - n \ln \sqrt{2\pi} - \frac{1}{2} \mathbb{E}_\mu \|X\|_2^2.$$

The second terms on both sides cancel out by the dimension condition, and the third terms on both sides cancel out by the frame condition, and so we are left with (22). It turns out that a generalization of Proposition 1.7 holds for the forward-reverse inequality, see [CourLiu, Theorem 1.23]. Henceforth, we will only consider geometric instances for sake of convenience, and without loss of generality.

The proof of equivalence between the functional and entropic forms relies on an expression for the (relative) entropy via Legendre duality. In the probability literature, this formula is also known as the Donsker-Varadhan representation.

Lemma 3.3 (Donsker-Varadhan)**.** *We have*

$$D(\mu\|\nu) = \sup_{g} \left\{ \mathbb{E}_\mu[g(X)] - \ln \mathbb{E}_\nu[e^{g(X)}] \right\} \tag{23}$$

where the supremum is over all g such that e^g is integrable.

Proof. Let g be any function such that e^g is integrable. Let $f = \frac{d\mu}{d\nu}$ be the Radon-Nikodym derivative. By Jensen's inequality,

$$\begin{aligned} \ln \int e^g d\nu &\geq \int f \ln\left(\frac{e^g}{f}\right) d\nu \\ &= \int fg d\nu - D(\mu\|\nu). \end{aligned}$$

Rearranging, we have

$$D(\mu\|\nu) \geq \int g(x) d\mu(x) - \ln \int e^{g(x)} d\nu(x) = \mathbb{E}_\mu[g(X)] - \ln \mathbb{E}_\nu[e^{g(X)}].$$

From the equality cases of Jensen's inequality, we see that equality above is achieved if and only if e^g is a constant multiple of f on the support of f. □

Instead of showing Theorem 3.2 directly, we derive the entropic dual of FRBL, from which the entropic forms of BL and RBL fall out immediately.

Theorem 3.4 (Entropic FRBL)**.** *Given a geometric FRBL datum $(\boldsymbol{c}, \boldsymbol{d}, \boldsymbol{B})$, the forward-reverse inequality is equivalent to the following entropic inequality:*

$$\sum_{i=1}^n c_i D(\mu_i\|\gamma_{E_i}) \geq \inf_{\mu \in \Pi(\mu_1,\ldots,\mu_n)} \left\{ \sum_{j=1}^m d_j D(B_j \# \mu \| \gamma_{E^j}) \right\}. \tag{24}$$

Here $\Pi(\mu_1, \ldots, \mu_n)$ refers to the space of couplings, i.e. measures μ whose marginals are equal to $\mu_1, \ldots, \mu_n$.

Proof. Suppose the entropic inequality holds, and let (e^{f_i}), (e^{g_j}) be functions satisfying (18). Then, as in Lemma 3.3, let μ_i be measures such that $f_i = \frac{d\mu_i}{d\gamma_{E_i}}$ and

$$D(\mu_i\|\gamma_{E_i}) = \mathbb{E}_{\mu_i}[f_i(X_i)] - \log \mathbb{E}_{\gamma_{E_i}}[e^{f_i(X_i)}].$$

Let $\epsilon > 0$ be arbitrary. Choose μ to be ϵ-close to attaining the infimum in (24). Then, as in Lemma 3.3, we have

$$D(B_j \# \mu \| \gamma_{E^j}) \geq \mathbb{E}_\mu[g_j(B_j X)] - \log \mathbb{E}_{\gamma_{E^j}}[e^{g_j(B_j X)}].$$

In particular, we may write

$$\begin{aligned} \epsilon + \sum_{i=1}^n c_i \left(\mathbb{E}_{\mu_i}[f_i(X_i)] - \log \mathbb{E}_{\gamma_{E_i}}[e^{f_i(X_i)}] \right) &= \epsilon + \sum_{i=1}^n c_i D(\mu_i\|\gamma_{E_i}) \\ &\geq \sum_{j=1}^m d_j D(B_j \# \mu \| \gamma_{E^j}) \\ &\geq \sum_{j=1}^m d_j \left(\mathbb{E}_\mu[g_j(B_j X)] - \log \mathbb{E}_\gamma[e^{g_j(B_j X)}] \right). \end{aligned}$$

We can take $\epsilon \to 0$, rearrange, and apply (18) to get

$$\begin{aligned} 0 &\geq \sum_{i=1}^{n} c_i \mathbb{E}_{\mu_i}[f_i(X_i)] - \sum_{j=1}^{m} \mathbb{E}_{\mu}[g_j(B_j X)] \\ &\geq \log \left[\frac{\prod_{i=1}^{n} (\mathbb{E}_{\gamma_{E_i}}[e^{f_i(X)}])^{c_i}}{\prod_{j=1}^{m} (\mathbb{E}_{\gamma}[e^{g_j(B_j X)}])^{d_j}} \right] \end{aligned}$$

Exponentiating and rearranging gives us the functional form.

For the other direction, we will use a minimax principle. In particular, let $\Pi = \Pi(\mu_1, \ldots, \mu_n)$ and write

$$\begin{aligned} &\inf_{\mu \in \Pi} \left\{ \sum_{j=1}^{m} d_j D(B_j \# \mu \| \gamma_{E^j}) \right\} \\ &= \inf_{\mu \in \Pi} \sup_{v_i} \left\{ \sum_{j=1}^{m} d_j D(B_j \# \mu \| \gamma_{E^j}) + \sum_{i=1}^{n} c_i \left(\mathbb{E}\left[v_i(X_i)\right] - \mathbb{E}\left[v_i(X_i)\right] \right) \right\} \\ &= \inf_{\mu \in \Pi} \sup_{u_j, v_i} \left\{ \sum_{j=1}^{m} d_j \left(\mathbb{E}_{\mu}[u_j(B_j X)] - \ln \mathbb{E}_{\gamma_{E^j}} \left[e^{u_j(X)} \right] \right) + \sum_{i=1}^{n} c_i \left(\mathbb{E}\left[v_i(X_i)\right] - \mathbb{E}\left[v_i(X_i)\right] \right) \right\} \\ &= \sup_{u_j, v_i} \left\{ \sum_{i=1}^{n} c_i \mathbb{E}[v_i(X_i)] - \sum_{j=1}^{m} d_j \ln \mathbb{E}_{\gamma_{E^j}} \left[e^{u_j(X)} \right] + \inf_{\mu \in \Pi} \left\{ \sum_{j=1}^{m} d_j \mathbb{E}_{\mu}[u_j(B_j X)] - \sum_{i=1}^{n} c_i \mathbb{E}[v_i(X_i)] \right\} \right\} \\ &= \sup_{u_j, v_i : \sum_{j=1}^{m} d_j u_j \circ B_j \geq \sum_{i=1}^{n} c_i v_i \circ \pi_{E_i}} \left\{ \sum_{i=1}^{n} c_i \mathbb{E}[v_i(X_i)] - \sum_{j=1}^{m} d_j \ln \mathbb{E}_{\gamma_{E^j}} \left[e^{u_j(X)} \right] \right\} \\ &\leq \sup_{v_i} \sum_{i=1}^{n} c_i \left(\mathbb{E}_{\mu_i}[v_i(X)] - \ln \mathbb{E}_{\gamma_{E_i}} \left[e^{v_i(X)} \right] \right) \\ &= \sum_{i=1}^{n} c_i D(\mu_i \| \gamma_{E_i}) \end{aligned}$$

The third line is due to (23). The fourth line we exchanged the order of infimum and supremum, which can be justified by Sion's minimax theorem. The fifth line holds since if $\sum_{j=1}^{m} d_j u_j \circ B_j \geq \sum_{i=1}^{n} c_i v_i \circ \pi_{E_i}$ does not hold almost everywhere, then the inner infimum can be chosen to blow up towards $-\infty$. The sixth line follows from the functional form of FRBL. Finally, the last line is another application of (23). $\square$

Corollary 3.5 (Entropic BL & RBL)**.** *Assuming geometricity, the forward inequality is equivalent to the following entropic inequality:*

$$D(\mu \| \gamma) \geq \sum_{j=1}^{m} d_j D(B_j \# \mu \| \gamma_{E^j}), \tag{25}$$

and the reverse inequality is equivalent to the following entropic inequality:

$$\sum_{i=1}^{n} c_i D(\mu_i \| \gamma_{E_i}) \geq \inf_{\mu \in \Pi(\mu_1, \ldots, \mu_n)} D(B \# \mu \| \gamma_{E_0}). \tag{26}$$

Remark. To see the connection between the map B and those in Barthe's inequality, write P_i for the B_i found in Section 2.2. Then

$$B = \begin{bmatrix} \vdots & & \vdots \\ c_1 P_1^* & \cdots & c_m P_m^* \\ \vdots & & \vdots \end{bmatrix}$$

Proof. In (24), let $n = 1$ and $c_1 = 1$ to obtain (25). On the other hand, let $m = 1$ and $d_1 = 1$ to obtain (26). □

3.3. Stochastic proof of BL and RBL.

We begin with a representation theorem for the relative entropy, as seen in [Leh].

Proposition 3.6. *Let B be a Brownian motion defined on some filtered probability space and let U be a drift with respect to the same filtration. Let μ be the law of $B + U$. Then we have*

$$D(\mu\|\gamma) \leq \frac{1}{2}\mathbb{E}\|U\|^2.$$

Now, let $(\mathbb{W}, \mathcal{B}, \gamma)$ be a Wiener space, with a natural filtration $(\mathcal{G}_t)_{t\geq 0}$. Let $\mathbb{H}$ be the Cameron-Martin space. Let μ be a measure on $\mathbb{W}$, absolutely continuous with respect to γ. Then there exists a process $u_t : \mathbb{W} \to \mathbb{R}^d$, known as the Föllmer drift, such that the following statements hold:

1. *The process $U_t := \int_0^t u_s(X)ds$ belongs to $\mathbb{H}$ almost surely.*
2. *$D(\mu\|\gamma) = \frac{1}{2}\mathbb{E}\|U\|^2$.*

We present now the stochastic proofs of BL and RBL via Föllmer drifts, due to Lehec [Leh2].

Theorem 3.7 (Lehec's proof of BL)**.** *Suppose our datum satisfies the geometricity conditions. For every probability measure μ on $\mathbb{W}$ we have*

$$D(\mu\|\gamma) \geq \sum_{j=1}^{m} d_i D(\mu_i\|\gamma_i),$$

where $\mu_i = P_i\#\mu$ and $\gamma_i = P_i\#\gamma$.

Proof. By Proposition 3.6, we have a Brownian motion B on H and a drift U such that $\mu \sim B + U$ and

$$D(\mu\|\gamma) = \frac{1}{2}\mathbb{E}\|U\|^2.$$

Since P_i is an orthogonal projection, the process $P_i B$ is the standard Brownian motion on H_i. Now, $P_i B + P_i U$ has law μ_i and so we have

$$D(\mu_i\|\gamma_i) \leq \frac{1}{2}\mathbb{E}\|P_i U\|^2, \quad i = 1, \ldots, m.$$

Finally, the geometric condition implies that

$$\|U\|^2 = \sum_{i=1}^{m} d_i \|P_i U\|^2,$$

which gives the result. □

Theorem 3.8 (Lehec's proof of RBL)**.** *Suppose our datum satisfies the geometricity conditions. Let γ be the Wiener measure and γ_i its projections onto each H_i. Then there exist coupled random processes $X_1, \ldots, X_m$ with laws $\mu_1, \ldots, \mu_m$. Letting μ^* be the law of $X := \sum_{i=1}^m c_i X_i$, we have*

$$D(\mu^*\|\gamma) \leq \sum_{i=1}^{n} c_i D(\mu_i\|\gamma_i).$$

Proof. Let B be a standard Brownian motion on H. Then P_iB is a standard Brownian motion on H_i. There is a drift u_i such that $P_iB + u_i \sim \mu_i$ and

$$D(\mu_i \| \gamma_i) = \frac{1}{2}\mathbb{E}\|u_i\|^2.$$

On the other hand, from geometricity we have $X = \sum_{i=1}^m c_i(P_iB + u_i) = B + \sum_{i=1}^m c_iu_i$, and so

$$D(\mu\|\gamma) \le \frac{1}{2}\mathbb{E}\left\|\sum_{i=1}^m c_iu_i\right\|^2$$

Now, if $x_1, \dots, x_m$ belong to $H_1, \dots, H_m$ respectively, then for any $y \in H$, the Cauchy-Schwarz inequality and frame condition yields

$$\begin{aligned}\left\langle \sum_{i=1}^m c_ix_i, y \right\rangle &= \sum_{i=1}^m c_i\langle x_i, P_iy\rangle \\ &\le \left(\sum_{i=1}^m c_i|x_i|^2\right)^{1/2}\left(\sum_{i=1}^m c_i|P_iy|^2\right)^{1/2} \\ &= \left(\sum_{i=1}^m c_i|x_i|^2\right)^{1/2}|y|.\end{aligned}$$

Substituting $y = \sum_{i=1}^m c_ix_i$, we get

$$\left|\sum_{i=1}^m c_ix_i\right|^2 \le \sum_{i=1}^m c_i|x_i|^2.$$

Thus,

$$\left\|\sum_{i=1}^m c_iu_i\right\|^2 \le \sum_{i=1}^m c_i\|u_i\|^2,$$

and so the claim follows. $\square$

4. Characterization of Extremizers

4.1. Structure. We start with some definitions. Recall that a subspace V of H is *critical* if $V \neq \{0\}, H$ and

$$\dim(V) = \sum_j c_j \dim(B_j V).$$

A subspace K of H is said to be *independent* if $K \neq \{0\}$ and

$$K = \cap_{j=1}^m H_j^a,$$

where $H_j^a = H_j$ or $H_j^\perp$. It turns out that any independent subspace is also critical. Note that any two independent subspaces are orthogonal to each other. Thus we have the *independent decomposition* of H, given by

$$H = K_{\text{ind}} \bigoplus K_{\text{dep}} = \left(\bigoplus_{k=1}^{k_0} K_k \right) \bigoplus K_{\text{dep}},$$

where $\{K_k : k = 1, \ldots, k_0\}$ is an enumeration of the independent subspaces of H and K_{dep} is the orthogonal complement of K_{ind}.

4.2. Characterization in BL. Without loss of generality, we may assume (B, c) is a geometric datum, and let $\oplus_{k=0}^{k_0} K_k \oplus K_{\text{dep}}$ be the independent decomposition of H. Then Valdimarsson [Val] states the following result:

Theorem 4.1 (Characterization for BL)**.** *Suppose (f_j) are extremizers for the datum (B, c). Then there exists a critical decomposition*

$$H = \underbrace{\left(\oplus_{k=0}^{k_0} K_k\right)}_{K_{ind}} \oplus \underbrace{\left(\oplus_{k=k_0+1}^{k_1} K_k\right)}_{K_{dep}}$$

with integrable $u_k : H_k \to \mathbb{R}$, $k = 1, \ldots, k_0$, positive constants C_j, d_k, and an element $b \in K_{dep}$ such that

$$f_j(x) = C_j \prod_{k=1}^{k_0} u_k(P_{j,k} B_j^* x) \prod_{k=k_0+1}^{k_1} e^{-d_k \langle P_{j,k} B_j^* x, P_{j,k}(B_j^* x + b)\rangle}, \tag{27}$$

where $P_{j,k}$ is the orthogonal projection from H onto $H_j \cap K_k$. Conversely, all functions of this form are extremizers for (B, c).

The method of proof involves analyzing the equality conditions in the heat equation proof. As such, we rely on notation used in Section 1.2. In particular, we have

$$\frac{\partial}{\partial t} \int F(x, t) dx = 0, \quad t > 0.$$

Henceforth we will fix a time $t > 0$ which will be implicit. Recalling the definitions of T and A as defined in (7) and (8), with $P = TT^*$, we have $\langle (I - P)A, A \rangle = 0$, and in particular $PA = A$. Recall that $P = TT^* = T(T^*T)^{-1}T^*$ is the projection onto the image of T. Therefore there must exist a map $\beta : H \to H$ so that

$$A(x) = T\beta(x).$$

By reading off the rows we get

$$\nabla h_j(B_j x) = B_j \beta(x), \tag{28}$$

and hence

$$\nabla \log F(x) = \sum_j p_j B_j^* \nabla h_j(B_j x) = \sum_j p_j B_j^* B_j \beta(x) = \beta(x)$$

where the last equality follows from geometricity.

Let $b_j = B_j^* e_j$, where $e_j \in H_j$. Then we have

$$\langle \nabla \log F(x), b_j\rangle = \langle \beta(x), b_j\rangle = \langle B_j\beta(x), e_j\rangle = \langle \nabla h_j(B_j x), e_j\rangle.$$

Now if we let $b_j^\perp = B_j^{\perp *} e_j^\perp$, where we take $B_j^\perp$ to be the orthogonal projection of H onto H_j and $e_j^\perp \in H_j^\perp$, and we differentiate this with respect to $b_j^\perp$, we get

$$D^2(\log F)(b_j, b_j^\perp) = 0. \tag{29}$$

This restricts $\log F$ to be of the form $u_j(B_j x) + u_j^\perp(B_j^\perp x)$. In particular for any j, j' we can write

$$\log F = u_j(B_j x) + u_j^\perp(B_j^\perp x) = u_{j'}(B_{j'} x) + u_{j'}^\perp(B_{j'}^\perp x) \tag{30}$$

At this point, we may use (30) to prove two lemmas which will be useful in determining the extremizers.

Lemma 4.2. *There exists functions u_{K_k} and $U_{K_{dep}}$ such that*

$$\log F = \left(\sum_{k=1}^{k_0} u_{K_k}(P_{K_k} x)\right) + u_{K_{dep}}(P_{K_{dep}} x).$$

Proof. It suffices to show that the second derivative of $\log F$ with respect to any pair of vectors from different components of the independent decomposition is zero.

If the two vectors come from two distinct independent subspaces K_k and $K_{k'}$, then there must be a j such that $K_k \subset H_j$ and $K_{k'} \subset H_j^\perp$. Then the result follows immediately from (29).

Now suppose one vector $b_1 \in K_k$ and another vector $b_2 \in K_{\text{dep}}$. Since

$$K_{\text{dep}} \subset K_k^\perp = \sum_j H_j^{a\perp},$$

b_2 can be written as some linear combination of vectors in $H_j^{a\perp}$. Then we have $D^2(\log F)(b_1, b_j^{a\perp})$ for any $b_j^{a\perp} \in H_j^{a\perp}$. The result then follows by linearity. □

Lemma 4.3. *If a space H has no independent subspaces, then any extremizer must be Gaussian.*

Proof. In [Val, Lemma 5], it is shown that $\nabla \log F$ has at most linear growth in each x_j. In particular, if $\nabla \log F$ is a linear polynomial, then (28) together with $\nabla \log F = \beta$ implies that each f_j is Gaussian.

To show that $\nabla \log F$ is a polynomial, it suffices to show that the Fourier transform is a distribution supported only at the origin. Considering (30), they compute the Fourier transform of functions of the form $u(B_j x)$ and show that it is supported in H_j. For details see [Val, Lemma 14]. It then follows that the transform of $\nabla \log F$ is supported in $H_j \cup H_j^\perp$, and in particular supported on

$$\cap_j (H_j \cup H_j^\perp).$$

Since H has no independent subspaces, we know this intersection is just $\{0\}$. □

Remark. By Lemma 4.3 we see that the $u_{K_{\text{dep}}}$ appearing in Lemma 4.2 is Gaussian. After some more massaging we may obtain the form in (27). Finally, the proofs above relied on the functions (f_j) being Schwartz. The technical details for the reduction to Schwartz functions are at the end of [Val].

4.3. Equality cases via the transportation method. Recall the optimal transport proof of Barthe's inequality, which involved the following relationship between constants:

$$I(f_1, \ldots, f_m) \geq D \cdot J(h_1, \ldots, h_m). \tag{31}$$

We may use this to glean information about the extremizers of both BL and RBL. In particular, suppose (h_i) and (f_i) are extremizers for BL and RBL, respectively. Then (31) must hold with equality. Following the optimal transport proof, we must have

$$\det\left(\sum_{i=1}^m c_i B_i^* dT_i(B_i y) B_i\right) = D \prod_{i=1}^m (\det(dT_i(B_i y)))^{c_i}. \tag{32}$$

This is equivalent to saying that the Gaussians corresponding to covariances $(dT_i(B_i y))_{i=1}^m$ are maximizers for BL, and that $(dT_i(B_i y)^{-1})_{i=1}^m$ are minimizers for RBL.

Furthermore, we must satisfy

$$\prod_{i=1}^m f_i(T_i(B_i y))^{c_i} \leq \sup_{\Theta(y)=\sum_{i=1}^m c_i B_i^* x_i} \prod_{i=1}^m f_i(x_i)^{c_i} \tag{33}$$

with equality. That is, for each y, we must have

$$\prod_{i=1}^m f_i(T_i(B_i y))^{c_i} \geq \prod_{i=1}^m f_i(x_i)^{c_i}$$

for all (x_i) satisfying

$$\sum_{i=1}^m c_i B_i^* T_i(B_i y) = \sum_{i=1}^m c_i B_i^* x_i.$$

Let's start with some special cases.

Example 4.1 (Prékopa Leindler)**.** For $0 < \lambda < 1$ and nonnegative f, g, h, suppose

$$h((1-\lambda)x + \lambda y) \geq f(x)^{1-\lambda} g(y)^\lambda$$

for all $x, y \in \mathbb{R}$. Then

$$\int h(s)ds \geq \left(\int f\right)^{1-\lambda} \left(\int g\right)^\lambda.$$

The datum is given by

$$B_1 = B_2 = I_1; \quad c_1 = 1-\lambda; \quad c_2 = \lambda,$$

so the forward inequality can be written as

$$\int f(x)^{1-\lambda} g(x)^\lambda dx \leq \left(\int f\right)^{1-\lambda} \left(\int g\right)^\lambda,$$

which is just Hölder's inequality. Suppose (h_1, h_2) are extremizers in Hölder and (f_1, f_2) extremizers in PL, so that we have equality in (31). Then (32) becomes

$$(1-\lambda)T_1'(y) + \lambda T_2'(y) = (T_1'(y))^{1-\lambda}(T_2'(y))^\lambda.$$

The equality conditions for AM-GM tell us that $T_1'(y) = T_2'(y)$ for all y. That is, we may write $T_2(y) = T_1(y) + c$ for some constant $c \in \mathbb{R}$. Now, equation (33) becomes

$$f_1(T_1(y))^{1-\lambda} f_2(T_2(y))^\lambda = \sup\left[f_1(x_1)^{1-\lambda} f_2(x_2)^\lambda\right],$$

where the supremum is over all x_1, x_2 satisfying

$$(1-\lambda)T_1(y) + \lambda T_2(y) = (1-\lambda)x_1 + \lambda x_2.$$

Applying logarithms and writing $\varphi_i = \log f_i$ for $i = 1, 2$, we get

$$(1-\lambda)\varphi_1(T_1(y)) + \lambda\varphi_2(T_1(y) + c) = \sup\left[(1-\lambda)\varphi_1(x_1) + \lambda\varphi_2(x_2)\right]$$

where the supremum is over x_1, x_2 satisfying

$$T_1(y) + \lambda c = (1-\lambda)x_1 + \lambda x_2.$$

Optimizing the Lagrangian[2], we get that $\varphi_1'(T_1(y)) = \varphi_2'(T_1(y) + c)$ for all y. Since T_1 is monotonic increasing and bijective, we see that $\varphi_2(x_2) = \varphi_1(x_2 - c) + d$, for some $d \in \mathbb{R}$. So our equation becomes

$$\varphi_1(T_1(y)) = \sup\left[(1-\lambda)\varphi_1(x_1) + \lambda\varphi_1(x_2 - c)\right]$$

where x_1, x_2 satisfy $T_1(y) = (1-\lambda)x_1 + \lambda(x_2 - c)$. But this is just the definition of concavity. That is, f_1, f_2 are log-concave, satisfying the relation

$$f_2(x) = e^d f_1(x - c).$$

Now, to recover the extremizers for Hölder's inequality, we use the relation

$$T_i'(y) f_i(T_i(y)) = h_i(y).$$

From this we see that $h_2(y) = e^d h_1(y)$ for all y.

Example 4.2 (Young)**.** To simplify our analysis, we will assume that $p = q = r = 3/2$. Then our inequality becomes

$$\int\int f(x)^{2/3} g(x-y)^{2/3} h(y)^{2/3} dxdy \leq F\left(\int f\right)^{2/3}\left(\int g\right)^{2/3}\left(\int h\right)^{2/3}$$

Then, suppose that $(h_i)_{i=1}^3$ are extremizers for the above inequality, and that $(f_i)_{i=1}^3$ are extremizers for the reverse inequality. We have

$$B_1 = \begin{bmatrix}1 & 0\end{bmatrix}; \quad B_2 = \begin{bmatrix}1 & -1\end{bmatrix}; \quad B_3 = \begin{bmatrix}0 & 1\end{bmatrix}; \quad c_1 = c_2 = c_3 = 1/p = 2/3.$$

So the reverse inequality can be stated as follows: if $\varphi : \mathbb{R}^2 \to \mathbb{R}_{\geq 0}$ satisfies

$$\varphi\left(\frac{2}{3}\begin{bmatrix} y_1 + y_2 \\ -y_2 + y_3 \end{bmatrix}\right) \geq f(y_1)^{2/3} g(y_2)^{2/3} h(y_3)^{2/3},$$

then

$$\int\int \varphi(x,y) dxdy \geq E\left(\int f\right)^{2/3}\left(\int g\right)^{2/3}\left(\int h\right)^{2/3}.$$

The condition (32) becomes

$$\det\left(\frac{2}{3}\begin{bmatrix} T_1'(y_1) + T_2'(y_1 - y_2) & -T_2'(y_1 - y_2) \\ -T_2'(y_1 - y_2) & T_3'(y_2) + T_2'(y_1 - y_2)\end{bmatrix}\right) = D[T_1'(y_1) T_2'(y_1 - y_2) T_3'(y_2)]^{2/3},$$

[2]Here we are assuming differentiability of the extremizers. For the general case, we can use the following reduction. Let g_n be a Gaussian density with variance $1/n$, and (f_j) are extremizers of the inequality. Then each $f_j * g_n$ is smooth, and also an extremizer of the inequality; hence they are log-concave. It is well-known that log-concavity is preserved under weak limits, and since $f_j * g_n \to f_j$ weakly, it follows that f_j itself is log-concave.

where we may explicitly compute the constant to be $D = 4/3$. Therefore we have

$$\frac{1}{3}[T_1'(y_1)T_3'(y_2) + T_1'(y_1)T_2'(y_1 - y_2) + T_3'(y_2)T_2'(y_1 - y_2)] = [T_1'(y_1)T_2'(y_1 - y_2)T_3'(y_2)]^{2/3}.$$

Once again equality conditions for AM-GM imply

$$T_1'(y_1) = T_2'(y_1 - y_2) = T_3'(y_2)$$

for all y_1, y_2. Thus $T_i(x) = cx + d_i$ for some constants $d_i \in \mathbb{R}$. Now, we know that the extremizers (h_1, h_2, h_3) are Gaussians of the same shape. So by the change of variables relation

$$h_i(x) = g_i(T_i(x))T_i'(x),$$

we see that the extremizers (g_1, g_2, g_3) must also be Gaussians of the same shape.

Theorem 4.4. *Suppose our datum $(\boldsymbol{c}, \boldsymbol{B})$ is geometric, dependent, and our space $H = \mathbb{R}^n$ is simple with respect to this datum. Then any extremizers of Barthe's inequality take the form:*

$$f_j(x) = \alpha_j e^{\alpha \langle x, x + \beta_j \rangle}$$

where $a_j, \alpha \in \mathbb{R}$ and $\beta_j \in H = \mathbb{R}^{n_j}$ for $1 \leq j \leq m$. Conversely, all functions of this form are extremizers for this datum.

Proof. Let (h_j) be extremizers of the forward inequality. Then by Valdimarsson's characterization, we have for each $1 \leq j \leq m$,

$$h_j(x) = a_j e^{-a \langle P_j B_j^* x, P_j(B_j^* x + b) \rangle},$$

where $a_j, a \in \mathbb{R}$, $b \in H$ and P_j are the orthogonal projections from H onto H_j. In particular, $P_j = B_j$ and by geometricity we get

$$h_j(x) = a_j e^{-a(\|x\|^2 + x^* B_j b)}. \tag{34}$$

Assuming (f_j) extremize the reverse inequality, we must satisfy (32). So consider any collection of transport maps (T_j). Fix any $y \in H$. We know that $(dT_j(B_j y))$ corresponds to the covariances of Gaussian maximizers of BL. From Valdimarsson's characterization and (34), we see that for each $1 \leq j \leq m$,

$$dT_j(B_j y) = a' \operatorname{Id}_{H_j},$$

for some $a' \in \mathbb{R}$. Since B_j is onto, we may write $dT_j(z) = a'(z) \operatorname{Id}_{H_j}$ for all $z \in H_j$. Thus for all $y \in H$ we have

$$dT_{j_1}(B_1 y)_{i_1 i_1} = dT_{j_2}(B_2 y)_{i_2 i_2}$$

for all $1 \leq j_k \leq m$, $1 \leq i_k \leq \dim(H_{j_k})$ for $k \in \{1, 2\}$. Now, note that dependence implies $\sum_{j=1}^m \ker(B_j) = H$. Therefore each $dT_j(B_j y)_{ii}$ is constant, and in particular we may write

$$T_j(x) = a'x + d_j,$$

for some $d_j \in H_j$. By the change of variables formula between h_j and f_j, we get

$$\begin{aligned} f_j(x) &= h_j(T_j^{-1}(x)) \det(d(T_i^{-1})(x)) \\ &= \frac{a_j}{a'} \exp\left(-\frac{a}{a'^2}\langle x - d_j, x - d_j + B_j b\rangle\right), \end{aligned}$$

which is of the desired form. ∎

Corollary 4.5. *Suppose the datum $(\boldsymbol{c}, \boldsymbol{B})$ is geometric and dependent. Let $H = \oplus_{k=1}^{k_1} K_k$ be a maximal critical decomposition. Then any extremizers of Barthe's inequality take the form*

$$f_j(x) = \alpha_j \prod_{k=1}^{k_1} e^{-\alpha_k \langle P_{j,k} B_j^* x, P_{j,k}(B_j^* x + \beta) \rangle}$$

where $P_{j,k}$ is the orthogonal projection from H onto $H_j \cap K_k$.

Remark. Our result strictly generalizes theorem 4 in [Bar], which considers the case where each $n_j = 1$, and the space $\mathbb{R}^n$ is irreducible with respect to the datum $(v_j)_{j=1}^m$. Indeed, their irreducibility condition implies that the whole space is dependent. For suppose otherwise, i.e. there exists a nonzero independent subspace U. Then we must have $\dim U = 1$; otherwise the subspace U has dimension at least 2 and is orthogonal to every v_j, which prevents any subset of the $\{v_j\}$ from spanning H. Now, suppose $U = \mathrm{span}(v_i)$ for some $i \in [m]$. There exists $n-1$ vectors v_j, with $j \neq i$, that span $U^\perp$. Note that v_i is orthogonal to each of these v_j, and so cannot belong to the same equivalence class, contradicting irreducibility.

We now present the full statement of characterization for RBL, due recently to [BKX].

Theorem 4.6 (Characterization for RBL)**.** *Suppose our datum $(\boldsymbol{B}, \boldsymbol{c})$ is geometric. Let $F_1, \dots, F_l$ be the independent subspaces, and F_{dep} be the dependent subspace. Suppose (f_i) are extremizers for $(\boldsymbol{B}, \boldsymbol{c})$. Then there exist $\theta_i > 0$, $b_i \in E_i \cap F_{dep}$, and $w_i \in E_i \cap F_{dep}^\perp$ for $i = 1, \dots, n$, as well as arbitrary $g_j : F_j \mapsto \mathbb{R}^+$, log-concave $h_j : F_j \to \mathbb{R}^+$, and a matrix $A \in \mathcal{S}^+(F_{dep})$ such that the eigenspace of A are critical subspaces and*

$$f_i(x) = \theta_i e^{-\langle A P_{F_{dep}} x, P_{F_{dep}} x - b_i \rangle} \prod_{\substack{F_j \subset E_i \\ \dim(F_j)=1}} g_j(P_{F_j}(x - w_i)) \prod_{\substack{F_j \subset E_i \\ \dim(F_j)>1}} h_j(P_{F_j}(x - w_i)).$$

Conversely, any function of the form above is an extremizer of RBL.

Remark. The proof in [BKX] relies on the Brenier map utilized in [Bar]. The rough idea is to restrict to the independent subspaces and reduce the problem to the case of the Prékopa-Leindler inequality, for which we know the extremizers are log-concave. The dimension one arbitrary g_j's in the above form simply come from the degenerate case of $m = 1$ in Prékopa-Leindler. An alternate method of proof for these characterization results is to go through the stochastic analysis done in Section 3.3.

Acknowledgements

This survey was done as part of my senior thesis at UC Berkeley. I would like to thank Professor Thomas Courtade, who introduced me to the topic, and Professor Michael Christ, who advised this thesis. I am grateful for both of their guidance throughout this process.

References

[Bar] F. Barthe, *On a reverse form of the Brascamp-Lieb inequality*, Invent. Math. 134(1998), 335-361.

[Bar2] F. Barthe, *Optimal Young's inequality and its converse: a simple proof*, Geom. Func. Anal. 8(1998), 234-242.

[BarCor] F. Barthe, D. Cordero-Erausquin, *Inverse Brascamp-Lieb inequalities along the heat equation* Geom. Func. Anal. 65-71.

[BarHu] F. Barthe, N. Huet, *On Gaussian Brunn-Minkowski inequalities*, Studia Math. 191(2009), 283-304.

[Bec] W. Beckner, *Inequalities in Fourier analysis*, Annals of Mathematics, 102(1975), 159-182.

[BCCT] J. Bennett, A. Carbery, M. Christ, and T. Tao, *The Brascamp-Lieb inequalities: finiteness, structure and extremals*, Geom. Func. Anal. 17(5): 1343-1415, 2008.

[BKX] K. J. Boroczky, P. Kalantzopoulos, Dongmeng Xi, *About the case of equality in the Reverse Brascamp-Lieb inequality*, arXiv:2203.01428v2, 2022.

[Bre] Y. Brenier, *Décomposition polaire et réarrangement monotone des champs de vecteurs*, C. R. Acad. Sci. Paris Sér. I Math., 305:805-808, 1987.

[CaCo] E. Carlen and D. Cordero-Erausquin, *Subadditivity of the entropy and its relation to Brascamp-Lieb type inequalities*, Geom. Func. Anal. 19 (2009), 373-405.

[Caf] L. Caffarelli, *The regularity of mappings with a convex potential*, J. Amer. math. Soc., 4:99-104, 1992.

[CourLiu] T.A. Courtade, J. Liu, *Euclidean forward-reverse Brascamp-Lieb inequalities: finiteness, structure, and extremals*, J. Geom. Anal. 31(2021), 3300-3350.

[Gar] R. Gardner, *The Brunn-Minkowski inequality*, Bulletin of the American Mathematical Society, 39(3): 355-405, 2002.

[Leh] J. Lehec, *Representation formula for the entropy and functional inequalities*, Annales de l'IHP Probabilités et statistiques, 49(2013), 885-899.

[Leh2] J. Lehec, *Short probabilistic proof of the Brascamp-Lieb and Barthe theorems*, Canad. math. Bull. 57(2014), 585-597.

[Lieb] E. H. Lieb, *Gaussian kernels have only Gaussian maximizers*, Inven. Math. 102(1990), 179-208.

[Val] S. I. Valdimarsson, *Optimisers for the Brascamp-Lieb inequality*, Israel J. Math. 168(2008), 253-274.