Combinatorics, Probability and Computing (2017) 26, 468–480. © Cambridge University Press 2017
doi:10.1017/S0963548316000432

An Inequality for Functions on the Hamming Cube

ALEX SAMORODNITSKY†

School of Engineering and Computer Science, The Hebrew University of Jerusalem, Jerusalem 91904, Israel
(e-mail: salex@cs.huji.ac.il)

Received 5 July 2012; revised 25 October 2016

We prove an inequality for functions on the discrete cube $\{0,1\}^n$ extending the edge-isoperimetric inequality for sets. This inequality turns out to be equivalent to the following claim about random walks on the cube: subcubes maximize 'mean first exit time' among all subsets of the cube of the same cardinality.

2010 *Mathematics subject classification*: Primary 05D05
Secondary 60C05

1. Introduction

Isoperimetric inequalities play an important role in describing the geometry of ambient spaces [2, 12]. This paper deals with one such space, the *discrete cube* $\{0,1\}^n$. This is the graph with 2^n vertices indexed by Boolean strings of length n, in which two vertices are connected by an edge if they differ in one coordinate. The *edge-isoperimetric* inequality [8] for $\{0,1\}^n$ provides a well-known example of a discrete isoperimetric inequality.

The *edge boundary* ∂A of a subset $A \subseteq \{0,1\}^n$ is the set of edges between A and its complement. The edge-isoperimetric inequality relates the cardinality of a set and that of its boundary:

$$|\partial A| \geqslant |A| \cdot \log_2\left(\frac{2^n}{|A|}\right). \tag{1.1}$$

One of its implications is that a simple random walk in the cube does not stay for too long in any given subset. This can be used to prove upper bounds on the mixing time of the walk [9].

This inequality can also be viewed as an inequality for characteristic functions on $\{0,1\}^n$. For a function $g : \{0,1\}^n \to \mathbb{R}$, let the *Dirichlet quadratic form* of g be given by

$$\mathcal{E}(g,g) = \mathbb{E}_x \sum_{y \sim x} (g(x) - g(y))^2.$$

† Research partially supported by ISF grant 1241/11 and BSF grant 2010451. Part of this work was done while the author was a fellow at the Radcliffe Institute in 2009–2010.

Here the expectation is taken with respect to the uniform probability measure on the cube. The notation $x \sim y$ means that x and y are connected by an edge. Then (1.1) can be rewritten for $g = 1_A$ as

$$\mathcal{E}(g,g) \geqslant 2\mathbb{E}(g^2) \cdot \log_2\left(\frac{\mathbb{E}(g^2)}{(\mathbb{E}|g|)^2}\right). \tag{1.2}$$

It is natural to look for inequalities for real-valued functions g on the cube generalizing (1.1). One such inequality is the *logarithmic Sobolev inequality* [7]:

$$\mathcal{E}(g,g) \geqslant Ent(g^2) = \mathbb{E}(g^2 \ln g^2) - \mathbb{E}(g^2) \ln \mathbb{E}(g^2).$$

For $g = 1_A$ this becomes

$$|\partial A| \geqslant |A| \cdot \ln\left(\frac{2^n}{|A|}\right),$$

recovering (1.1) up to a multiplicative factor of $1/\ln 2$. For a general real-valued function g, the logarithmic Sobolev inequality has been observed [5, 11] to imply

$$\mathcal{E}(g,g) \geqslant 2\mathbb{E}(g^2) \cdot \ln\left(\frac{\mathbb{E}(g^2)}{(\mathbb{E}|g|)^2}\right). \tag{1.3}$$

This extends (1.2), again up to a multiplicative factor of $1/\ln 2$.

It is useful to look for inequalities for general functions reducing to an isoperimetric inequality with the *correct* constant in the special case of characteristic functions. Such an inequality would, in particular, mean that the characteristic function of an *isoperimetric* set (that is, a set satisfying an isoperimetric inequality with equality), or an 'almost-isoperimetric' set, is an optimal (or nearly optimal) solution of a continuous extremal problem, and as such, might be expected to have an interesting structure. We refer to [1] for an example of relevant work in continuous analysis.

As observed in [5], the inequality (1.3) is in fact tight for general real-valued functions. Therefore, to recover correct constants, we need to look for different extensions of (1.1). This paper gives one example of such an inequality.

Theorem 1.1. *Let A be a subset of $\{0,1\}^n$ and let g be a real-valued function on $\{0,1\}^n$ supported on A. Then*

$$\mathcal{E}(g,g) \geqslant 2 \cdot \frac{1}{2^n \cdot |A|} \log_2\left(\frac{2^n}{|A|}\right) \cdot \left(\sum_{x \in A} |g(x)|\right)^2. \tag{1.4}$$

The dependence on g on the right hand side of this inequality is weaker than that in the logarithmic Sobolev inequality, or that in (1.3). However, it does give the right constant. In fact, substituting $g = 1_A$ recovers (1.1).

It turns out that (1.4) is equivalent to a statement about random walks in the cube. Let A be a subset of $\{0,1\}^n$. Let Y be a random variable defined as follows. Choose a uniformly random point $a \in A$ and consider the simple random walk in $\{0,1\}^n$ starting from a. Then Y measures the time it takes the walk to exit A for the first time. We refer to $\mathbb{E}Y$ as the *mean first exit time* of A. This is a parameter of a subset A of the cube.

The following claim is equivalent to Theorem 1.1.

Theorem 1.2. *Subcubes maximize mean first exit time among all subsets of the cube of the same cardinality. More precisely, for any subset A of $\{0,1\}^n$,*

$$\mathbb{E}Y \leqslant \frac{n}{\log_2(2^n/|A|)}. \tag{1.5}$$

If A is a subcube, this is an equality.

This paper is organized as follows. We show equivalence of Theorems 1.1 and 1.2 in Section 2. Theorem 1.1 is proved in Sections 3 and 4. Some remarks on the structure of almost isoperimetric sets are given in Section 5.

2. A random walk interpretation of Theorem 1.1

Inequality (1.4) is an inequality between two quadratic forms, which can be interpreted as a matrix inequality. Let $L = L_A$ be the $|A| \times |A|$ matrix indexed by the vertices of A, with the following entries: $L(a,a) = n$; and for $a \neq b$, $L(a,b) = -1$ if a,b are connected, and 0 if not. Let $J := J_A$ be the $|A| \times |A|$ all-1 matrix. Then, (1.4) is equivalent to

$$L \succeq \frac{1}{|A|} \log_2\left(\frac{2^n}{|A|}\right) \cdot J \tag{2.1}$$

Here $L \succeq M$ means that $L - M$ is a positive semidefinite matrix, that is, $\langle Lu, u\rangle \geqslant \langle Mu, u\rangle$ for any vector u.

The inequality (2.1) is of the form $L \succeq vv^t$ for a vector $v \in \mathbb{R}^A$. Note that if A is not the complete cube (which we may assume, since otherwise the claims of both Theorems 1.1 and 1.2 are trivially true), the matrix L is non-singular. Therefore

$$L \succeq vv^t \quad \Leftrightarrow \quad I \succeq (L^{-1/2}v)(L^{-1/2}v)^t \quad \Leftrightarrow \quad \langle L^{-1/2}v, L^{-1/2}v\rangle \leqslant 1 \quad \Leftrightarrow \quad \langle L^{-1}v, v\rangle \leqslant 1.$$

Let

$$r = \frac{1}{|A|} \log_2\left(\frac{2^n}{|A|}\right)$$

and let $\mathbf{1}$ be the all-1 vector in $\mathbb{R}^A$. Then (2.1) amounts to

$$\langle L^{-1}\mathbf{1}, \mathbf{1}\rangle \leqslant \frac{1}{r}. \tag{2.2}$$

This inequality allows a random walk interpretation. Write $L = n \cdot I - E$, where I is the identity matrix and E is the adjacency matrix of the subgraph of $\{0,1\}^n$ induced by the vertices in A. (Thus L is the 'external' Laplacian of the subgraph induced by A.) The matrix $\frac{1}{n} \cdot E$ has eigenvalues smaller than 1, and therefore we can write

$$L^{-1} = \frac{1}{n} \cdot \sum_{k=0}^{\infty} \frac{E^k}{n^k}.$$

https://doi.org/10.1017/S0963548316000432 Published online by Cambridge University Press

The inequality (2.2) can be rewritten as

$$\frac{n}{r} \geqslant n \cdot \langle L^{-1}\mathbf{1},\mathbf{1}\rangle = \sum_{k=0}^{\infty} \frac{\langle E^k\mathbf{1},\mathbf{1}\rangle}{n^k}.$$

Let Y be a random variable defined as follows. Choose a uniform random point $a \in A$ and consider a simple random walk in $\{0,1\}^n$ starting from a. Then Y measures the first time the walk exits A. Note that $E^k(a,b)$ counts the number of paths of length k in A between a and b. Hence

$$\frac{1}{n^k} \cdot \sum_{b\in A} E^k(a,b)$$

is the probability that the random walk starting from a remains in A for the first k steps, and

$$\frac{\langle E^k\mathbf{1},\mathbf{1}\rangle}{|A| \cdot n^k}$$

is the probability $Y > k$. Therefore, by (2.2),

$$\mathbb{E}Y = \sum_{k=0}^{\infty} \mathbb{P}\{Y > k\} = \frac{1}{|A|} \cdot \sum_{k=0}^{\infty} \frac{\langle E^k\mathbf{1},\mathbf{1}\rangle}{n^k} \leqslant \frac{n}{|A| \cdot r} = \frac{n}{\log_2(2^n/|A|)},$$

proving (1.5).

Next, we verify that (1.5) holds with equality if A is a subcube, completing the proof of Theorem 1.2. Let A be a d-dimensional subcube. Then

$$\mathbb{P}\{Y > k\} = \frac{d^k}{n^k},$$

and therefore

$$\mathbb{E}Y = \sum_{k=0}^{\infty} \mathbb{P}\{Y > k\} = \sum_{k=0}^{\infty} \frac{d^k}{n^k} = \frac{n}{n-d} = \frac{n}{\log_2(2^n/|A|)}.$$

One might consider the possibility that subcubes have a stronger property, namely that for a walk of *any* length the probability to remain in a subcube is maximal among all sets of the same size. This is true for walks of length 1, since subcubes have the smallest edge boundaries. However, the following example shows this to be false already for walks of length 2.

Example. The number of length-2 walks inside the set A is

$$\sum_{a,b\in A} E^2(a,b) = \langle E^2\mathbf{1},\mathbf{1}\rangle = \langle E\mathbf{1},E\mathbf{1}\rangle = \sum_{x\in A} d_x^2,$$

where d_x is the degree of x in the subgraph induced by A. Therefore, for a d-dimensional cube, the number of such walks is $2^d \cdot d^2$. But, for a radius-1 ball of dimension $2^d - 1$ (that is, a star with 2^d vertices), this number is $(2^d-1)^2 + (2^d-1) = 2^d \cdot (2^d-1)$, which is much larger.

3. Proof of Theorem 1.1

There are several simple assumptions we may and will make on the structure of the function g in (1.4). First, we may assume $g \geqslant 0$, since replacing g with its absolute value preserves the right-hand side of (1.4) and can only decrease its left-hand side. Second, we may assume the support

https://doi.org/10.1017/S0963548316000432 Published online by Cambridge University Press

of g is the whole set A, otherwise we may replace A with the support of g in (1.4), increasing the right-hand side.

Next, consider the partial order on $\{0,1\}^n$ in which $x \preceq y$ if and only if $x_i \leqslant y_i$, $i = 1,\ldots,n$. A function g on the cube is *downwards monotone* if $g(x) \geqslant g(y)$ when $x \preceq y$. We may assume the function g in (1.4) to be monotone. This follows from two simple lemmas.

Lemma 3.1. *Fix a direction $1 \leqslant i \leqslant n$, and let f be a function obtained from g by a downward shift in direction i. That is, for any pair of adjacent points x, y in the cube, with $x_i = 0$ and $y_i = 1$, set*

$$f(x) = \max\{g(x), g(y)\} \quad \textit{and} \quad f(y) = \min\{g(x), g(y)\}.$$

Then

$$\mathcal{E}(f,f) \leqslant \mathcal{E}(g,g).$$

Proof. This is a standard 'shifting' argument [3], more commonly applied in the special case of g being a characteristic function. We will reduce the claim of the lemma to the two-dimensional case, verifiable by a direct calculation.

For a point $x \in \{0,1\}^n$ and $1 \leqslant j \leqslant n$, let $x^{(j)}$ denote the point adjacent to x in direction j. That is, $x^{(j)}_k = x_k$ for any $k \neq j$, but $x^{(j)}_j \neq x_j$. For a function h on the cube, and x uniformly distributed in $\{0,1\}^n$, set $\Delta_j(h) = \mathbb{E}_x(f(x) - f(x^{(j)}))^2$ and note that $\mathcal{E}(h,h) = \sum_{j=1}^n \Delta_j(h)$.

Let f be obtained from g by a downward shift in direction i. Clearly $\Delta_i(f) = \Delta_i(g)$. We will show that $\Delta_j(f) \leqslant \Delta_j(g)$ for all $j \neq i$. The claim of the lemma will follow. Fix $j \neq i$ and assume, for ease of notation, that $j = n-1$ and $i = n$.

For a function h on $\{0,1\}^n$ and for $z \in \{0,1\}^{n-2}$, let h_z be the restriction of h to the 2-dimensional cube $\{x \in \{0,1\}^n : x_k = z_k \text{ for } 1 \leqslant k \leqslant n-2\}$. Observe that if z is uniformly distributed in $\{0,1\}^{n-2}$, then $\Delta_k(h) = \mathbb{E}_z \Delta_k(h_z)$, for $k = n-1, n$. For any z in the $(n-2)$-dimensional cube, f_z is a downward shift of g_z in direction n. We will verify that $\Delta_{n-1}(f_z) \leqslant \Delta_{n-1}(g_z)$. This will imply, by averaging over z, that $\Delta_{n-1}(f) \leqslant \Delta_{n-1}(g)$.

Fix $z \in \{0,1\}^{n-2}$, let $g = g_z$, and $f = f_z$. The only interesting case to consider, up to the symmetries of the cube, is

$$g = \begin{pmatrix} A & b \\ a & B \end{pmatrix} \quad \text{and hence } f = \begin{pmatrix} a & b \\ A & B \end{pmatrix}.$$

Here direction n is vertical, $A > a$ and $B > b$. Direct calculation gives

$$\Delta_{n-1}(g) - \Delta_{n-1}(f) = \frac{1}{2} \cdot (B-b)(A-a) > 0,$$

completing the proof. □

Lemma 3.2. *Applying consecutive shifts in directions $i = 1,\ldots,n$ to a function on the cube produces a monotone function.*

https://doi.org/10.1017/S0963548316000432 Published online by Cambridge University Press

Proof. Again, it suffices to verify this in the two-dimensional case. See [6], for example, where this argument is applied in the special case of characteristic functions. □

The proof of Theorem 1.1 proceeds by induction on the dimension. First, consider the base case $n = 1$. There are two choices for $|A|$. If $|A| = 1$, we are in the Boolean case, in which (1.4) is the usual edge-isoperimetry. If $|A| = 2$, the right-hand side in (1.4) is 0, and we are done.

Now we go to the induction step. The cube $\{0,1\}^n$ decomposes into two $(n-1)$-dimensional subcubes. The first subcube contains all vectors with last coordinate 0, and the second contains all vectors with last coordinate 1. The function g and the set A decompose according to their restrictions to the subcubes:

$$g \hookrightarrow (g_0, g_1), \quad A \hookrightarrow (A_0, A_1).$$

The induction step amounts to proving

$$\begin{aligned}\mathcal{E}(g,g) &= \frac{1}{2}\cdot(\mathcal{E}(g_0,g_0)+\mathcal{E}(g_1,g_1)) + \|g_0 - g_1\|_2^2 \\ &\geqslant_{\text{ind}} \frac{1}{2}\cdot 2\cdot \frac{1}{2^{n-1}|A_0|}\log\left(\frac{2^{n-1}}{|A_0|}\right)\left(\sum_{x\in A_0} g_0(x)\right)^2 \\ &\quad + \frac{1}{2}\cdot 2\cdot \frac{1}{2^{n-1}|A_1|}\log\left(\frac{2^{n-1}}{|A_1|}\right)\left(\sum_{x\in A_1} g_1(x)\right)^2 + \|g_0-g_1\|_2^2 \\ &\geqslant^{??} 2\cdot\frac{1}{2^n|A|}\log\left(\frac{2^n}{|A|}\right)\left(\sum_{x\in A} g(x)\right)^2.\end{aligned}$$

In the expressions above, the Dirichlet forms and the ℓ_2 distance for functions g_i on $(n-1)$-dimensional cubes are computed with respect to the uniform probability measure on these subcubes.

Note that, by our assumptions on g, the set A is downwards monotone, since it is the support of a monotone function g. This implies $A_1 \subseteq A_0$ (identifying the two subcubes in the natural way). The expression we need to analyse allows an additional simplifying assumption on g. We may assume g_0, g_1 to be constant on A_1 and on $A_0 \setminus A_1$ (and of course g_i vanishes on A_i^c; in particular g_1 is zero on $A_0 \setminus A_1$). In fact, replacing g_i with their averages on the corresponding subsets can only decrease the left-hand side and does not change the right-hand side in the second inequality above.

We proceed with the analysis, introducing some notation.

Notation

- Let $s_0 := \sum_{x\in A_0} g_0(x)$, $s_1 := \sum_{x\in A_1} g_1(x)$. Let $t_0 := |A_0|$, $t_1 := |A_1|$. We may and will assume $t_1 > 0$ and $s_1 > 0$; otherwise the problem reduces to a lower-dimensional case.
- Let α be the value of g_0 on A_1 and let γ be the value of g_0 on $A_0 \setminus A_1$. Let β be the value of g_1 on A_1.
- The 'f'-notation. Let

$$f(t) = f_{n-1}(t) := \frac{1}{t}\log_2\left(\frac{2^{n-1}}{t}\right).$$

https://doi.org/10.1017/S0963548316000432 Published online by Cambridge University Press

Note that

(1) $t_1\beta = s_1$,
(2) $t_1\alpha + (t_0 - t_1)\gamma = s_0$,
(3) $\|g_0 - g_1\|_2^2 = \frac{1}{2^{n-1}} \cdot (t_1(\alpha-\beta)^2 + (t_0-t_1)\gamma^2)$

With the new notation, the inequality to be verified for the induction step is

$$f(t_0)s_0^2 + f(t_1)s_1^2 + (t_1(\alpha-\beta)^2 + (t_0-t_1)\gamma^2) \geqslant \frac{1}{2} \cdot f\left(\frac{t_0+t_1}{2}\right)(s_0+s_1)^2. \tag{3.1}$$

Expressing β and γ as functions of s_i, t_i and of α (dealing with the simple case $t_0 = t_1$ separately), the left-hand side of (3.1) is a quadratic in α with coefficients depending on s_i and t_i. Minimizing the left-hand side in α, we arrive, after some simple calculations, at the following inequality we need to verify:

$$f(t_0)s_0^2 + f(t_1)s_1^2 + \frac{(s_0-s_1)^2}{t_0} \geqslant \frac{1}{2} \cdot f\left(\frac{t_0+t_1}{2}\right)(s_0+s_1)^2. \tag{3.2}$$

Next, let $R = s_0/s_1$. Inequality (3.2) transforms to a quadratic inequality in R:

$$f(t_0)R^2 + f(t_1) + \frac{(R-1)^2}{t_0} \geqslant \frac{1}{2} \cdot f\left(\frac{t_0+t_1}{2}\right)(R+1)^2. \tag{3.3}$$

We need to check $P(R) := aR^2 + bR + c \geqslant 0$ with the coefficients a, b, c coming from (3.3). We will, in fact, verify $a \geqslant 0$ and $D = b^2 - 4ac \leqslant 0$, which will conclude the proof. We start with some simple properties of the function

$$f(t) = \frac{1}{t}\log_2\left(\frac{2^{n-1}}{t}\right).$$

Lemma 3.3. *The function $f(t)$ is decreasing and convex for $0 < t < 2^{n-1}$. It satisfies the identity*

$$f(\beta \cdot t) = \frac{1}{\beta} \cdot f(t) + \frac{1}{\beta}\log\frac{1}{\beta} \cdot \frac{1}{t}, \tag{3.4}$$

for any $t, \beta > 0$.

Proof. Directly verifiable. □

Corollary 3.4. *Viewing inequality (3.3) in the form $aR^2 + bR + c \geqslant 0$, we have*

$$a \geqslant 0.$$

Proof. It is easy to verify that

$$a = \frac{2t_0 f(t_0) + 2 - t_0 f((t_0+t_1)/2)}{2t_0}.$$

https://doi.org/10.1017/S0963548316000432 Published online by Cambridge University Press

By Lemma 3.3,

$$t_0 f\left(\frac{t_0+t_1}{2}\right) \leqslant t_0 f\left(\frac{t_0}{2}\right) = 2t_0 f(t_0)+2,$$

completing the proof. □

It remains to verify the inequality $4ac \geqslant b^2$, which after some simplification reduces to

$$2t_0 f(t_0) f(t_1) + 2(f(t_0)+f(t_1)) \geqslant t_0 f\left(\frac{t_0+t_1}{2}\right)(f(t_0)+f(t_1)) + 4f\left(\frac{t_0+t_1}{2}\right). \tag{3.5}$$

4. Proof of inequality (3.5)

Renaming the variables $x = t_0$ and $y = t_1$, and recalling the constraints on t_0 and t_1, we need to prove (3.5) for $1 \leqslant y < x \leqslant 2^{n-1}$. Rearranging, this is easily seen to be equivalent to

$$\Delta(x,y) \geqslant \frac{x\cdot(f(x)-f(y))^2}{2x\cdot(f(x)+f(y))+8}. \tag{4.1}$$

Here

$$\Delta(x,y) := \frac{f(x)+f(y)}{2} - f\left(\frac{x+y}{2}\right).$$

Note that $\Delta \geqslant 0$ since f is convex.

We now substitute $y = \beta x$ in (4.1), with $0 < \beta < 1$, and expand using (3.4). We have

$$\begin{aligned}
\Delta(x,y) = \Delta(x,\beta\cdot x) &= \frac{f(x)+f(\beta\cdot x)}{2} - f\left(\frac{1+\beta}{2}\cdot x\right) \\
&= \frac{1}{2}\cdot\left(f(x)+\frac{1}{\beta}\cdot f(x)+\frac{1}{\beta}\log\frac{1}{\beta}\cdot\frac{1}{x}\right) - \left(\frac{2}{1+\beta}\cdot f(x)+\frac{2}{1+\beta}\log\frac{2}{1+\beta}\cdot\frac{1}{x}\right) \\
&= \frac{(1-\beta)^2}{2\beta(1+\beta)}\cdot f(x) + \left(\frac{1}{2\beta}\log\frac{1}{\beta} - \frac{2}{1+\beta}\log\frac{2}{1+\beta}\right)\cdot\frac{1}{x}.
\end{aligned}$$

As to the right-hand side of (4.1), we have

$$\mathrm{RHS}(x,y) = \mathrm{RHS}(x,\beta\cdot x) = \frac{x\cdot\left(\frac{1}{\beta}\cdot f(x)+\frac{1}{\beta}\log\frac{1}{\beta}\cdot\frac{1}{x}-f(x)\right)^2}{2x\cdot\left(f(x)+\frac{1}{\beta}\cdot f(x)+\frac{1}{\beta}\log\frac{1}{\beta}\cdot\frac{1}{x}\right)+8}.$$

Taking $z := xf(x)$,

$$\Delta \geqslant \mathrm{RHS} \quad\Longleftrightarrow\quad x\cdot\Delta \geqslant x\cdot\mathrm{RHS} \quad\Longleftrightarrow\quad Az+B \geqslant \frac{(Cz+D)^2}{Ez+F},$$

where

$$z = xf(x) = \log\left(\frac{2^{n-1}}{x}\right) \geqslant 0$$

and $A, B, \ldots, F$ depend only on β.

https://doi.org/10.1017/S0963548316000432 Published online by Cambridge University Press

Specifically,

$$\begin{aligned} A &= \frac{(1-\beta)^2}{2\beta(1+\beta)}, \\ B &= \frac{1}{2\beta}\log\frac{1}{\beta} - \frac{2}{1+\beta}\log\frac{2}{1+\beta}, \\ C &= \frac{1-\beta}{\beta}, \\ D &= \frac{1}{\beta}\log\frac{1}{\beta}, \\ E &= \frac{2+2\beta}{\beta}, \\ F &= \frac{2}{\beta}\log\frac{1}{\beta} + 8. \end{aligned}$$

So, we need to show that

$$(Az+B)(Ez+F) \geqslant (Cz+D)^2.$$

Observe that

$$AE = C^2 = \frac{(1-\beta)^2}{\beta^2}.$$

Therefore, this reduces to a linear inequality in z:

$$(AF+BE-2CD)\cdot z \geqslant D^2 - BF.$$

This holds for all non-negative z if and only if

$$\begin{aligned} AF+BE &\geqslant 2CD, \\ BF &\geqslant D^2. \end{aligned}$$

Hence, the problem is reduced to two univariate inequalities in β. We will prove them in the next two lemmas.

Lemma 4.1. *For $0 < \beta < 1$ we have $AF + BE \geqslant 2CD$.*

Proof. Simplifying and rearranging, this inequality reduces to

$$\frac{(1-\beta)^2}{\beta(1+\beta)}\cdot\log\frac{1}{\beta} + \frac{1+\beta}{\beta}\cdot\log\frac{1}{\beta} + 4\frac{(1-\beta)^2}{1+\beta} \geqslant 4\log\frac{2}{1+\beta} + 2\frac{1-\beta}{\beta}\cdot\log\frac{1}{\beta}$$

and hence to

$$\frac{\beta}{1+\beta}\cdot\log\frac{1}{\beta} + \frac{(1-\beta)^2}{1+\beta} \geqslant \log\frac{2}{1+\beta},$$

which is the same as

$$\beta\log\frac{1}{\beta} + (1-\beta)^2 \geqslant (1+\beta)\log\frac{2}{1+\beta}.$$

The derivative of

$$g(\beta)=\beta\log\frac{1}{\beta}+(1-\beta)^2-(1+\beta)\log\frac{2}{1+\beta}$$

is

$$\log\frac{1+\beta}{2\beta}-2(1-\beta).$$

This is a convex function, which means it can vanish in at most two points in the interval $(0,1]$. In addition, g' is positive close to 0 and it vanishes at 1. Taking into account the boundary conditions $g(0)=g(1)=0$, this means that g first increases from 0 at zero and then decreases to 0 at one, that is, it is non-negative. □

Lemma 4.2. *For $0<\beta<1$ we have $BF\geqslant D^2$.*

Proof. We need to prove

$$\left(\frac{1}{2\beta}\log\frac{1}{\beta}-\frac{2}{1+\beta}\log\frac{2}{1+\beta}\right)\cdot\left(\frac{2}{\beta}\log\frac{1}{\beta}+8\right)\geqslant\frac{1}{\beta^2}\log^2\frac{1}{\beta}.$$

Simplifying and rearranging, this reduces to

$$(1+\beta)\cdot\log\frac{1}{\beta}\geqslant\log\frac{2}{1+\beta}\cdot\log\frac{1}{\beta}+4\beta\cdot\log\frac{2}{1+\beta},$$

which is the same as

$$(\beta+\log(1+\beta))\cdot\log\frac{1}{\beta}\geqslant 4\beta\cdot\log\frac{2}{1+\beta}.$$

As in the preceding lemma, the function

$$g(\beta)=(\beta+\log(1+\beta))\cdot\log\frac{1}{\beta}-4\beta\cdot\log\frac{2}{1+\beta}$$

vanishes at the endpoints. We will (again) claim it increases from 0 at zero and then decreases from the maximum point to 0 at one, and is therefore non-negative on the interval.

As before, it will suffice to show that g' is convex, is positive at the beginning of the interval, and vanishes at 1. We have

$$\ln 2\cdot g'(\beta)=\left(1+\frac{1}{\ln 2\cdot(1+\beta)}\right)\cdot\ln\frac{1}{\beta}-\frac{\ln 2\cdot\beta+\ln(1+\beta)}{\ln 2\cdot\beta}-4\ln\frac{2}{1+\beta}+\frac{4\beta}{1+\beta}.$$

It is easy to verify that g' is positive for small positive β and that $g'(1)=0$. It remains to check g' is convex. Taking another two derivatives, we have

$$\ln 2\cdot g'''(\beta)=\left(\frac{1}{\beta^2}+\frac{3}{\ln 2}\cdot\left(\frac{1}{\beta^2(1+\beta)}+\frac{1}{\beta(1+\beta)^2}\right)+\frac{2}{\ln 2}\cdot\frac{\ln(1/\beta)}{(1+\beta)^3}\right)$$
$$-\left(\frac{4}{(1+\beta)^2}+\frac{8}{(1+\beta)^3}+\frac{2}{\ln 2}\cdot\frac{\ln(1+\beta)}{\beta^3}\right).$$

To show that this is non-negative, we multiply by $\beta^3(1+\beta)^3$ and verify

$$\beta(1+\beta)^3+\frac{3}{\ln 2}\cdot\beta(1+\beta)(1+2\beta)+\frac{2}{\ln 2}\cdot\beta^3\ln\frac{1}{\beta}\geqslant 4\beta^3(1+\beta)+8\beta^3+\frac{2}{\ln 2}\cdot(1+\beta)^3\ln(1+\beta).$$

https://doi.org/10.1017/S0963548316000432 Published online by Cambridge University Press

We show a stronger inequality (removing the third summand on the left):

$$\beta(1+\beta)^3 + \frac{3}{\ln 2}\cdot\beta(1+\beta)(1+2\beta) \geqslant 4\beta^3(1+\beta) + 8\beta^3 + \frac{2}{\ln 2}\cdot(1+\beta)^3\ln(1+\beta).$$

Note that $(1+\beta)\ln(1+\beta) \leqslant 2\ln 2\cdot\beta$, by convexity of $(1+\beta)\ln(1+\beta)$ on $[0,1]$. Substituting and simplifying, it suffices to show that

$$(1+\beta)^3 + \frac{3}{\ln 2}\cdot(1+\beta)(1+2\beta) \geqslant 4\beta^2(1+\beta) + 8\beta^2 + 4\cdot(1+\beta)^2.$$

Since $3/\ln 2 \geqslant 4$ and $\beta^2 \geqslant \beta^3$ for $0 \leqslant \beta \leqslant 1$, it suffices to prove the quadratic inequality

$$1 + 3\beta + 3\beta^2 + 4(1+\beta)(1+2\beta) \geqslant 15\beta^2 + 4\cdot(1+\beta)^2.$$

Simplifying, this reduces to the trivial statement

$$7\beta + 1 \geqslant 8\beta^2.$$

5. Nearly isoperimetric sets and their eigenvalues

Fix a small parameter $\varepsilon > 0$. A set A is *nearly isoperimetric* if it satisfies the isoperimetric inequality (1.1) almost as an equality, that is,

$$|A|\log_2\left(\frac{2^n}{|A|}\right) \leqslant |\partial A| \leqslant (1+\varepsilon)\cdot|A|\log_2\left(\frac{2^n}{|A|}\right). \tag{5.1}$$

We would like to understand the structure of nearly isoperimetric sets and, in particular, their possible similarity to subcubes.

This discussion is closely related to *stability* of isoperimetric inequalities. A stability-type result shows that a nearly isoperimetric set is close (in an appropriate metric) to a genuinely isoperimetric set. Such a result is proved in [4]. Let δ be at most a small constant, and let A be a set satisfying (5.1) with

$$\varepsilon = \frac{\delta}{\log_2(2^n/|A|)}.$$

Then there is a subcube C such that

$$|A\Delta C| \leqslant O\left(\frac{\delta}{\log(1/\delta)}\cdot|A|\right).$$

In this section we look at eigenvalues and eigenvectors of the Laplacian L (equivalently, of the adjacency matrix E) of a subgraph induced by an almost isoperimetric subset A of the cube. If A is a subcube, the induced subgraph is regular, of degree $\log_2|A|$. This means that the minimal eigenvalue of the Laplacian L is $\log_2(2^n/|A|)$ and the corresponding eigenvector is the all-1 vector $\mathbf{1}$. We show in Corollary 5.2 below that if ε' is at most a small constant and

$$\varepsilon = \frac{\varepsilon'}{2n}\cdot\log_2(2^n/|A|),$$

then the subgraph induced by a set A satisfying (5.1) is nearly regular, with the degrees of almost all the vertices close to $\log_2|A|$.

Similar arguments can be used to show that even for ε as large as a small constant, most of the spectral mass in the expansion of $\mathbf{1}$ in an eigenbasis of L is concentrated around the eigenvalue $\log_2(2^n/|A|)$ (we will not go into details). At the other end of the scale, for a very small

$$\varepsilon \ll \frac{1}{n \cdot \log_2(2^n/|A|)},$$

we can derive stability-type results in the sense of [4] (via a result of Keevash [10] on stability of the Kruskal–Katona inequality). Since this is weaker than the results in [4], we omit the details here as well.

We start with some notation. Let $|A| = m$, and let $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_m$ be the eigenvalues of E. The eigenvalues of L are $n-\lambda_1 \leqslant n-\lambda_2 \leqslant \cdots \leqslant n-\lambda_m$. Let $v_1,\dots,v_m$ be an orthonormal basis of eigenvectors, and let $\mathbf{1} = \sum_{i=1}^m \alpha_i v_i$ be the expansion of the constant-1 function $\mathbf{1}$ in this basis. Note, for future use, that $\sum_{i=1}^m \alpha_i^2 = \langle \mathbf{1},\mathbf{1}\rangle = |A|$.

The inequality (2.2) translates to

$$\sum_{i=1}^m \frac{\alpha_i^2}{n-\lambda_i} \leqslant \frac{1}{r} = \frac{|A|}{\log_2(2^n/|A|)}. \tag{5.2}$$

Note that the edge boundary of A is given by

$$|\partial A| = \langle L\mathbf{1},\mathbf{1}\rangle = \sum_{i=1}^m \alpha_i^2 (n-\lambda_i).$$

Therefore the nearly isoperimetric property (5.1) is equivalent to

$$|A|\log_2\left(\frac{2^n}{|A|}\right) \leqslant \sum_{i=1}^m \alpha_i^2(n-\lambda_i) \leqslant (1+\varepsilon)\cdot |A|\log_2\left(\frac{2^n}{|A|}\right).$$

Consider the probability distribution on $[m]$ given by $p_i = \alpha_i^2/|A|$, and let $f : i \mapsto n-\lambda_i$ be a positive function on $[m]$. Computing expectations according to p, we have $\mathbb{E}\frac{1}{f}\cdot\mathbb{E}f \leqslant 1+\varepsilon$. Intuitively, this should mean f is concentrated with respect to p. In the next lemma we state this formally.

Lemma 5.1. *Let g be a strictly positive-valued function on a finite domain satisfying*

$$\mathbb{E}\frac{1}{g}\cdot \mathbb{E}g \leqslant 1+\varepsilon.$$

Then

$$\mathbb{E}(g-\mathbb{E}g)^2 \leqslant \varepsilon\cdot\mathbb{E}g\cdot\|g\|_\infty. \tag{5.3}$$

Proof. We have

$$\mathbb{E}\left(\frac{(g-\mathbb{E}g)^2}{g}\right) = (\mathbb{E}g)^2\cdot\mathbb{E}\left(\frac{1}{g}\right) - \mathbb{E}g = \mathbb{E}g\cdot\left(\mathbb{E}g\cdot\mathbb{E}\left(\frac{1}{g}\right)-1\right) \leqslant \varepsilon\cdot\mathbb{E}g.$$

Therefore

$$\mathbb{E}(g-\mathbb{E}g)^2 \leqslant \mathbb{E}\left(\frac{(g-\mathbb{E}g)^2}{g}\right)\cdot\|g\|_\infty \leqslant \varepsilon\cdot\mathbb{E}g\cdot\|g\|_\infty.$$

□

https://doi.org/10.1017/S0963548316000432 Published online by Cambridge University Press

Corollary 5.2. *Let A satisfy* (5.1) *with*

$$\varepsilon = \frac{\varepsilon'}{2n} \cdot \log_2\left(\frac{2^n}{|A|}\right),$$

where $\varepsilon' \leqslant 1$. *Fix a parameter* $0 \leqslant \delta \leqslant 1$. *Choose uniformly at random an element* $x \in A$ *and let* $d_{\mathrm{out}}(x)$ *be the number of neighbours of x outside A. Then*

$$\mathbb{P}\left\{(1-\delta)\cdot\log_2\left(\frac{2^n}{|A|}\right) \leqslant d_{\mathrm{out}}(x) \leqslant (1+\varepsilon)(1+\delta)\cdot\log_2\left(\frac{2^n}{|A|}\right)\right\} \geqslant 1-\frac{\varepsilon'}{\delta^2}.$$

In particular, the subgraph induced by A is almost regular, similar to the isoperimetric case.

Proof. We use the notation above. Consider the random variable $d_{\mathrm{out}}(x)$, for x uniformly distributed in A. We have

$$\mathbb{E}d_{\mathrm{out}}(x) = \frac{|\partial A|}{|A|} = \frac{1}{|A|}\langle L\mathbf{1},\mathbf{1}\rangle = \sum_{i=1}^{m}\frac{\alpha_i^2}{|A|}(n-\lambda_i) = \mathbb{E}f.$$

Similarly, $\mathbb{E}(d_{\mathrm{out}}^2) = \mathbb{E}(f^2)$. Therefore, by Chebyshev's inequality and Lemma 5.3,

$$\mathbb{P}\left\{\left|d_{\mathrm{out}}(x) - \frac{|\partial A|}{|A|}\right| \geqslant \delta\cdot\frac{|\partial A|}{|A|}\right\} = \mathbb{P}\{|d_{\mathrm{out}}(x) - \mathbb{E}d_{\mathrm{out}}| \geqslant \delta\cdot\mathbb{E}d_{\mathrm{out}}\}$$
$$\leqslant \frac{\mathrm{Var}(d_{\mathrm{out}})}{\delta^2\cdot(\mathbb{E}d_{\mathrm{out}})^2} = \frac{\mathrm{Var}(f)}{\delta^2\cdot(\mathbb{E}f)^2} \leqslant \frac{\varepsilon\cdot\|f\|_\infty}{\delta^2\cdot\mathbb{E}f} \leqslant \frac{\varepsilon'}{\delta^2}.$$

In the last inequality we used the easy fact $\|f\|_\infty \leqslant 2n$. The claim follows. □

References

[1] Barthe, F. (2002) Log-concave and spherical models in isoperimetry. *Geom. Funct. Anal.* **12** 32–55.
[2] Bezrukov, S. (1994) Isoperimetric problems in discrete spaces. In *Extremal Problems for Finite Sets* (P. Frankl, Z. Füredi, G. Katona and D. Miklos, eds), Vol. 3 of *Bolyai Society Mathematical Studies*, pp. 59–91.
[3] Bollobas, B. (1986) *Combinatorics: Set Systems, Hypergraphs, Families of Vectors, and Combinatorial Probability*, Cambridge University Press.
[4] Ellis, D. (2011) Almost isoperimetric subsets of the discrete cube. *Combin. Probab. Comput.* **20** 363–380.
[5] Falik, D. and Samorodnitsky, A. (2007) Edge-isoperimetric inequalities and influences. *Combin. Probab. Comput.* **16** 693–712.
[6] Goldreich, O., Goldwasser, S., Lehman, E., Ron, D. and Samorodnitsky, A. (2000) Testing monotonicity. *Combinatorica* **20** 301–337.
[7] Gross, L. (1975) Logarithmic Sobolev inequalities. *Amer. J. Math.* **97** 1061–1083.
[8] Harper, L. H. (1966) Optimal numberings and isoperimetric problems on graphs. *J. Combin. Theory* **1** 385–393.
[9] Jerrum, M. (2003) *Counting, Sampling and Integrating: Algorithms and Complexity*, *Lectures in Mathematics, ETH Zürich*, Birkhäuser.
[10] Keevash, P. (2008) Shadows and intersections: Stability and new proofs. *Adv. Math.* **218** 1685–1703.
[11] Montenegro, R. and Tetali, P. (2005) *Mathematical Aspects of Mixing Times in Markov Chains*, Vol. 1:3 of *Foundations and Trends in Theoretical Computer Science*, Now Publishers Inc.
[12] Ros, A. (2001) The isoperimetric problem. http://www.ugr.es/~aros/isoper.htm