An inequality for functions on the Hamming cube

Alex Samorodnitsky *

Abstract

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.

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 a graph with 2^n vertices indexed by boolean strings of length n. Two vertices are connected by an edge if they differ in one coordinate. The *edge-isoperimetric* inequality [8] for $\{0,1\}^n$ provides well-known example for 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 compares between the cardinality of the set and of its boundary:

$$|\partial A| \geq |A| \cdot \log_2 \left(\frac{2^n}{|A|}\right) \tag{1}$$

One of its implications is that a simple random walk in the cube doesn't 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 $g : \{0,1\}^n \rightarrow \mathbb{R}$, let the *Dirichlet quadratic form* of g be given by

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

*School of Engineering and Computer Science, The Hebrew University of Jerusalem, Jerusalem 91904, Israel.
Research partially supported by ISF grant 1241/11 and by 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) can be rewritten, for $g = 1_A$ as

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

It is natural to look for inequalities for real-valued functions g on the cube generalizing (1). One such inequality is the logarithmic Sobolev inequality [7] :

$$\mathcal{E}(g,g) \geq Ent\left(g^2\right) = \mathbb{E}g^2 \ln g^2 - \mathbb{E}g^2 \ln \mathbb{E}g^2$$

For $g = 1_A$ this becomes $|\partial A| \geq |A| \cdot \ln\left(\frac{2^n}{|A|}\right)$, recovering (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

$$\mathcal{E}(g,g) \geq 2\mathbb{E}g^2 \cdot \ln\left(\frac{\mathbb{E}g^2}{\mathbb{E}^2|g|}\right) \tag{3}$$

This extends (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[1], 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 (3) is in fact tight for general real-valued functions. Therefore, to recover correct constants, we need to look for different extensions of (1).

This paper gives one example of such a 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*

$$\mathcal{E}(g,g) \geq 2 \cdot \frac{1}{2^n \cdot |A|} \log_2\left(\frac{2^n}{|A|}\right) \cdot \left(\sum_{x \in A} |g(x)|\right)^2 \tag{4}$$

The dependence on g on the right hand side of this inequality is weaker than that in the logarithmic Sobolev inequality, or that in (3). However, it does give the right constant. In fact, substituting $g = 1_A$ recovers (1).

[1]Recall that a set is isoperimetric if it satisfies an isoperimetric inequality with equality.

It turns out that (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 $\{0,1\}^n$,

$$\mathbb{E}Y \le \frac{n}{\log_2\left(\frac{2^n}{|A|}\right)} \tag{5}$$

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. Some remarks on the structure of almost isoperimetric sets are given in Section 3. Theorem 1.1 is proved in section 4.

2 A random walk interpretation of Theorem 1.1

Inequality (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 $a \ne b$, $L(a,b) = -1$ if a, b are connected, and 0 if not.

Let $J := J_A$ be the $|A| \times |A|$ all-1 matrix. Then, (4) is equivalent to:

$$L \succeq \frac{1}{|A|}\log_2\left(\frac{2^n}{|A|}\right) \cdot J \tag{6}$$

This is an inequality 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), the matrix L is non-singular. Therefore

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

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 (6) amounts to

$$\left\langle L^{-1}\mathbf{1}, \mathbf{1}\right\rangle \le \frac{1}{r} \tag{7}$$

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 a subgraph of $\{0,1\}^n$ induced by the vertices in A.[2] The matrix $\frac{1}{n} \cdot E$ has eigenvalues smaller than 1, and therefore we can write

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

The inequality (7) can be rewritten as

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

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 (7)

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

proving (5).

Next, we verify that (5) holds with equality if A is a subcube, completing the proof of Theorem 1.2.

Let A be a d-dimensional subcube. Then $\Pr\{Y > k\} = \frac{d^k}{n^k}$, and therefore

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

∎

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 2.1: The number of length-2 walks inside the set A is

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

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 $2^d - 1$, this number is $\left(2^d - 1\right)^2 + \left(2^d - 1\right) = 2^d \cdot \left(2^d - 1\right)$, which is much larger. ∎

[2] Thus L is the "external" Laplacian of the subgraph induced by A.

3 Near-isoperimetric sets and their eigenvalues

Fix a small parameter $\epsilon > 0$. A set A is *nearly isoperimetric* if it satisfies the isoperimetric inequality (1) with nearly an equality, that is

$$|A| \log_2 \left(\frac{2^n}{|A|} \right) \le |\partial A| \le (1+\epsilon) \cdot |A| \log_2 \left(\frac{2^n}{|A|} \right) \tag{8}$$

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 (8) with $\epsilon = \frac{\delta}{\log_2(2^n/|A|)}$. Then there is a subcube C such that $|A \Delta C| \le O\left(\frac{\delta}{\log(1/\delta)} \cdot |A|\right)$.

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 $\log_2 \left(\frac{2^n}{|A|}\right)$ and the corresponding eigenvector is the all-1 vector $\mathbf{1}$.

We show in Corollary 3.2 below that if ϵ' is at most a small constant and $\epsilon = \frac{\epsilon'}{2n} \cdot \log_2 (2^n/|A|)$, then the subgraph induced by a set A satisfying (8) is nearly regular, will 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 $\log_2 \left(\frac{2^n}{|A|}\right)$ (we don't go into details). On the other end of the scale, for a very small $\epsilon \ll \frac{1}{n \cdot \log_2 \left(\frac{2^n}{|A|}\right)}$ 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 $\lambda_1 \ge \lambda_2 \ge \ldots \ge \lambda_m$ be the eigenvalues of E. The eigenvalues of L are $n - \lambda_1 \le n - \lambda_2 \le \ldots \le n - \lambda_m$. Let $v_1, \ldots, v_m$ 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 (7) translates to

$$\sum_{i=1}^m \frac{\alpha_i^2}{n - \lambda_i} \le \frac{1}{r} = \frac{|A|}{\log_2 \left(\frac{2^n}{|A|}\right)} \tag{9}$$

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 near-isoperimetric property (8) is equivalent to

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

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

Lemma 3.1: *Let g be a strictly positive-valued function on a finite domain satisfying*

$$\mathbb{E}\frac{1}{g}\cdot\mathbb{E}g \le 1+\epsilon$$

Then

$$\mathbb{E}\,(g-\mathbb{E}g)^2 \le \epsilon\cdot\mathbb{E}g\cdot\|g\|_\infty \tag{10}$$

Proof: We have

$$\mathbb{E}\left(\frac{(g-\mathbb{E}g)^2}{g}\right) = \mathbb{E}^2 g\cdot\mathbb{E}\left(\frac{1}{g}\right) - \mathbb{E}g = \mathbb{E}g\cdot\left(\mathbb{E}g\cdot\mathbb{E}\left(\frac{1}{g}\right)-1\right) \le \epsilon\cdot\mathbb{E}g$$

Therefore

$$\mathbb{E}\,(g-\mathbb{E}g)^2 \le \mathbb{E}\left(\frac{(g-\mathbb{E}g)^2}{g}\right)\cdot\|g\|_\infty \le \epsilon\cdot\mathbb{E}g\cdot\|g\|_\infty$$

∎

Corollary 3.2: *Let A satisfy (8) with $\epsilon = \frac{\epsilon'}{2n}\cdot\log_2\left(\frac{2^n}{|A|}\right))$, where $\epsilon' \le 1$. Fix a parameter $0 \le \delta \le 1$.*

Choose uniformly at random an element $x \in A$ and consider its outdegree[3] *$d_{out}(x)$. Then*

$$Pr\left\{(1-\delta)\cdot\log_2\left(\frac{2^n}{|A|}\right) \le d_{out}(x) \le (1+\epsilon)(1+\delta)\cdot\log_2\left(\frac{2^n}{|A|}\right)\right\} \ge 1-\frac{\epsilon'}{\delta^2}$$

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

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

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

[3] the number of neighbors of x outside A

Similarly, $\mathbb{E}d_{out}^2 = \mathbb{E}f^2$. Therefore, by Chebyshev's inequality and Lemma 10,

$$Pr\left\{\left|d_{out}(x) - \frac{|\partial A|}{|A|}\right| \geq \delta \cdot \frac{|\partial A|}{|A|}\right\} = Pr\left\{\left|d_{out}(x) - \mathbb{E}d_{out}\right| \geq \delta \cdot \mathbb{E}d_{out}\right\} \leq$$

$$\frac{Var\left(d_{out}\right)}{\delta^2 \cdot \mathbb{E}^2 d_{out}} = \frac{Var\left(f\right)}{\delta^2 \cdot \mathbb{E}^2 f} \leq \frac{\epsilon \cdot \|f\|_\infty}{\delta^2 \cdot \mathbb{E}f} \leq \frac{\epsilon'}{\delta^2}$$

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

4 Proof of Theorem 1.1

There are several simple assumptions we may and will make on the structure of the function g in (4).

First, we may assume $g \geq 0$, since replacing g with its absolute value preserves RHS of (4) and can only decrease its LHS.

Second, we may assume the support of g is the whole set A, otherwise we may replace A with the support of g in (4), increasing RHS.

Next, consider the partial order on $\{0,1\}^n$ in which $x \preceq y$ iff $x_i \leq y_i$, $i = 1, ..., n$. A function g on the cube is *downwards monotone* if $g(x) \geq g(y)$ when $x \preceq y$.

We may assume the function g in (4) to be monotone. This follows from two simple lemmas.

Lemma 4.1: *Fix a direction $1 \leq i \leq n$, 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

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

Then

$$\mathcal{E}(f, f) \leq \mathcal{E}(g, g)$$

Proof: This is a standard "shifting" argument [3], more commonly applied in the special case of g being a characteristic function. The claim of the lemma is easily seen to follow from its validity for 2-dimensional cubes. The two-dimensional case is verifiable by a direct calculation. ∎

Lemma 4.2: *Applying consecutive shifts in directions $i = 1, ..., n$ to a function on the cube produces a monotone function.*

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

The proof 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 (4) is the usual edge-isoperimetry. If $|A| = 2$, RHS in (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 all vectors with last coordinate 1. The function g and the set A decompose according to their restrictions to the subcubes.

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

Induction step amounts to proving

$$\mathcal{E}(g,g) = \frac{1}{2} \cdot \left(\mathcal{E}\left(g_0, g_0\right) + \mathcal{E}\left(g_1, g_1\right)\right) + \|g_0 - g_1\|_2^2 \geq_{\text{ind}}$$

$$\frac{1}{2} \cdot 2 \cdot \frac{1}{2^{n-1}|A_0|} \log\left(\frac{2^{n-1}}{|A_0|}\right) \left(\sum_{x \in A_0} g_0(x)\right)^2 +$$

$$\frac{1}{2} \cdot 2 \cdot \frac{1}{2^{n-1}|A_1|} \log\left(\frac{2^{n-1}}{|A_1|}\right) \left(\sum_{x \in A_1} g_1(x)\right)^2 + \|g_0 - g_1\|_2^2 \geq^{??}$$

$$2 \cdot \frac{1}{2^n|A|} \log\left(\frac{2^n}{|A|}\right) \left(\sum_{x \in A} g(x)\right)^2$$

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 analyze allows an additional simplifying assumption on g. We may assume both g_0, g_1 to be constant on A_1 and on $A_0 \setminus A_1$ (and of course g_i vanish 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 LHS and does not change RHS in the second inequality above.

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

Note that

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

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

$$f(t_0)\, s_0^2 + f(t_1)\, s_1^2 + \left(t_1(\alpha-\beta)^2 + (t_0-t_1)\gamma^2\right) \geq \frac{1}{2}\cdot f\left(\frac{t_0+t_1}{2}\right)(s_0+s_1)^2 \tag{11}$$

Expressing β and γ as functions of s_i, t_i and of α,[4] LHS of (11) is a quadratic in α with coefficients depending on s_i and t_i. Minimizing LHS in α, we arrive, after some simple calculations, to the following inequality we need to verify:

$$f(t_0)\, s_0^2 + f(t_1)\, s_1^2 + \frac{(s_0-s_1)^2}{t_0} \geq \frac{1}{2}\cdot f\left(\frac{t_0+t_1}{2}\right)(s_0+s_1)^2 \tag{12}$$

Next, let $R = \frac{s_0}{s_1}$.

Inequality (12) transforms to a quadratic inequality in R

$$f(t_0)\, R^2 + f(t_1) + \frac{(R-1)^2}{t_0} \geq \frac{1}{2}\cdot f\left(\frac{t_0+t_1}{2}\right)(R+1)^2 \tag{13}$$

We need to check $P(R) := aR^2 + bR + c \geq 0$ with the coefficients a, b, c coming from (13). We will, in fact, verify $a \geq 0$ and $D = b^2 - 4ac \leq 0$, 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 4.3: *The function $f(t)$ is decreasing and convex for $0 < t < 2^{n-1}$. It satisfies the identity*

$$f(\beta \cdot t) = \frac{1}{\beta}\cdot f(t) + \frac{1}{\beta}\log\frac{1}{\beta}\cdot\frac{1}{t} \tag{14}$$

for any $t, \beta > 0$.

[4] Dealing with the simple case $t_0 = t_1$ separately

Proof: Directly verifiable. ∎

Corollary 4.4: *Viewing inequality (13) in the form $aR^2 + bR + c \geq 0$, we have*

$$a \geq 0$$

Proof: It is easy to verify

$$a = \frac{2t_0 f(t_0) + 2 - t_0 f\left(\frac{t_0+t_1}{2}\right)}{2t_0}$$

By Lemma 4.3

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

completing the proof. ∎

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

$$2t_0 f(t_0) f(t_1) + 2(f(t_0) + f(t_1)) \geq t_0 f\left(\frac{t_0+t_1}{2}\right)(f(t_0) + f(t_1)) + 4f\left(\frac{t_0+t_1}{2}\right) \tag{15}$$

5 Proof of inequality (15)

Renaming the variables $x = t_0$ and $y = t_1$, and recalling the constraints on t_0 and t_1, we need to prove (15) for $1 \leq y < x \leq 2^{n-1}$.

Rearranging, this is easily seen to be equivalent to

$$\Delta(x,y) \geq \frac{x \cdot (f(x) - f(y))^2}{2x \cdot (f(x) + f(y)) + 8} \tag{16}$$

Here $\Delta(x,y) := \frac{f(x)+f(y)}{2} - f\left(\frac{x+y}{2}\right)$. Note that $\Delta \geq 0$ since f is convex.

We now substitute $y = \beta x$ in (16), witn $0 < \beta < 1$ and expand using (14). We have

$$\Delta(x,y) = \Delta(x, \beta \cdot x) = \frac{f(x) + f(\beta \cdot x)}{2} - f\left(\frac{1+\beta}{2} \cdot x\right) =$$

$$\frac{1}{2} \cdot \left(f(x) + \frac{1}{\beta} \cdot f(x) + \frac{1}{\beta} \log \frac{1}{\beta} \cdot \frac{1}{x}\right) - \left(\frac{2}{1+\beta} \cdot f(x) + \frac{2}{1+\beta} \log \frac{2}{1+\beta} \cdot \frac{1}{x}\right) =$$

$$\frac{(1-\beta)^2}{2\beta(1+\beta)} \cdot f(x) + \left(\frac{1}{2\beta} \log \frac{1}{\beta} - \frac{2}{1+\beta} \log \frac{2}{1+\beta}\right) \cdot \frac{1}{x}$$

As to RHS of (16), we have

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

Taking $z := xf(x)$,

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

where $z = xf(x) = \log\left(\frac{2^{n-1}}{x}\right) \geq 0$ and $A, B, ..., F$ depend only on β.

Specifically,

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

So, we need to show

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

Observe that $AE = C^2 = \frac{(1-\beta)^2}{\beta^2}$. Therefore, this reduces to a linear inequality in z:

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

This holds for all nonnegative z if and only if

$$\begin{cases} AF + BE &\geq 2CD \\ BF &\geq D^2 \end{cases}$$

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

Lemma 5.1:

$$AF + BE \geq 2CD$$

For $0 < \beta < 1$.

Proof: Simplifying and rearranging, this inequality reduces to

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

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

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

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 nonnegative.

∎

Lemma 5.2:

$$BF \geq D^2$$

for $0 < \beta < 1$.

Proof: We need to prove

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

Simplifying and rearranging, this reduces to

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

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

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, nonnegative on the interval.

As before, it will suffice to show that g' is convex, is positive in 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

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

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

To show this is nonnegative, we multiply by $\beta^3(1+\beta)^3$ and verify

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

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

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

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

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

Since $\frac{3}{\ln 2} \geq 4$ and $\beta^2 \geq \beta^3$ for $0 \leq \beta \leq 1$, it suffices to prove the quadratic inequality

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

Simplifying, this reduces to the trivial statement

$$7\beta + 1 \geq 8\beta^2$$

∎

References

[1] F. Barthe, *Log-concave and spherical models in isoperimetry*. GAFA, Geom. funct. anal., Vol. 12: 32-55, 2002.

[2] S. Bezrukov, *Isoperimetric problems in discrete spaces*. In Extremal Problems for Finite Sets, Vol. 3 of Bolyai Soc. Math. Stud. (P. Frankl, Z. Furedi, G. Katona and D. Miklos, eds), pp. 59-91, 1994.

[3] B. Bollobas, **Combinatorics: set systems, hypergraphs, families of vectors, and combinatorial probability**. Cambridge University Press New York, NY, USA 1986.

[4] D. Ellis, *Almost Isoperimetric Subsets of the Discrete Cube*. Combinatorics, Probability & Computing 20(3): 363-380, 2011.

[5] D. Falik and A. Samorodnitsky, *Edge-Isoperimetric Inequalities and Influences*. Combinatorics, Probability & Computing 16(5): 693-712, 2007.

[6] Oded Goldreich, Shafi Goldwasser, Eric Lehman, Dana Ron, Alex Samorodnitsky, *Testing Monotonicity*. Combinatorica 20(3): 301-337, 2000.

[7] L. Gross, *Logarithmic Sobolev inequalities*, Amer. J. of Math., 97, pp. 1061-1083, 1975.

[8] L. H. Harper, *Optimal numberings and isoperimetric problems on graphs*, J. Combin. Theory, 1, pp. 385-393, 1966.

[9] M. Jerrum, **Counting, Sampling and Integrating: Algorithms and Complexity**, Lectures in Mathematics, ETH Zurich. Birkhauser, Basel, 2003.

[10] P. Keevash, *Shadows and intersections: stability and new proofs*. Adv. Math. 218, pp. 1685–1703, 2008.

[11] R. Montenegro and P. Tetali, **Mathematical Aspects of Mixing Times in Markov Chains**. Foundations and Trends in Theoretical Computer Science 1(3), 2005.

[12] A. Ros. The isoperimetric problem. http://www.ugr.es/∼aros/isoper.htm, 2001.