Edge Isoperimetry and Rapid Mixing on Matroids and Geometric Markov Chains

Ravi Montenegro *   Jung-Bae Son†

ABSTRACT

We show how to bound the mixing time and log-Sobolev constants of Markov chains by bounding the edge-isoperimetry of their underlying graphs. To do this we use two recent techniques, one involving Average Conductance and the other log-Sobolev constants. We show a sort of strong conductance bound on a family of geometric Markov chains, give improved bounds for the mixing time of a Markov chain on balanced matroids, and in both cases find lower bounds on the log-Sobolev constants of these chains.

1. INTRODUCTION

Given a Markov chain $\mathcal{M}$ such as a random walk on a graph we are are interested in showing *rapid mixing*, that the chain approaches the steady state distribution after a polynomial number of steps. An early result in this area was by Sinclair and Jerrum [16] who used the notion of conductance to show rapid mixing on several combinatorial Markov chains.

Dyer, Frieze, and Kannan [6] adapted this technique to give the first provably polynomial time algorithm to approximate the volume of a convex body. Developments related to the volume problem were sometimes applied back to the original problem of rapid mixing on combinatorial Markov chains. In one case Karzanov and Khachiyan [12] used geometric properties of the underlying graphs through *isoperimetric inequalities* to show rapid mixing on a Markov chain related to counting linear extensions.

In all these early techniques the key to bounding conductance was finding a lower bound on the *cutset expansion*, the infinum $\inf_{0<|S|\leq|V|/2} |Cut(S)|/|S|$ where $Cut(S)$ is the set of edges from S to S^c. A related concept is the *edge-isoperimetry*, a bound on $\inf_{0<|S|\leq x|V|} |Cut(S)|/|S|$ when $x \leq 1/2$. The extra information given by varying x allows us to find lower bounds on the *conductance function*, which measures how well one step redistributes probabilities conditional on the set being a certain size. Significantly, this is all that is needed to use two recent techniques for bounding mixing time, *Average Conductance* [14] and isoperimetric bounds on log-Sobolev constants [9].

In our first result we re-visit Karzanov and Khachiyan's work on geometric Markov chains and show a "strong version" of Sinclair and Jerrum's theorem, that the mixing time $\tau = O(1/\Phi^2)$ when Φ is bounded geometrically. In particular this strengthens previous conductance based results on rapid mixing for a Markov chain on linear extensions to $\tau = O(n^4)$, which even outperforms a path coupling / comparison technique [2] and is close to the correct bound [17] of $\tau = \Theta(n^3 \log n)$.

In our second result we extend work of Feder and Mihail [7] to bound the edge-isoperimetry of a random walk on balanced matroids. In the case of the *natural bases-exchange walk*, this strengthens their upper bound for the mixing time from $\tau = O(m\, n^3 \log m)$ to $\tau = O(m^{3/2}\, n^2 \log n)$, an improvement for all regular matroids with a constant number of parallel elements (eg. graphic matroids with few multiple edges). The most interesting point in this proof is that our improved bound is found by lower bounding the log-Sobolev constant of the Markov chain, making this one of the more complicated problems solved with log-Sobolev constants.

These two results partially answer a question of [14] as to how Average Conductance can be applied to combinatorial problems. We have shown how to extend two methods of bounding conductance, geometry and induction, to bound Average Conductance. The third and most common method of bounding conductance involves canonical paths, however when canonical paths are available they can be used to bound the mixing time more directly through eigenvalues [15] and so conductance is generally not used in this case. Moreover, eigenvalue bounds on the mixing time are usually better than $1/\Phi^2$ while Average Conductance gives $1/\Phi^2$ at best, so it seems unlikely Average Conductance will give useful results for canonical paths problems.

In Section 2 we introduce notation and theorems used throughout this paper. In Section 3 we show how to use geometry to bound the edge-isoperimetry of a family of Markov Chains and in Section 4 we use induction to do the same on balanced matroids. Section 5 poses some problems for future study and the Appendix contains proofs of some theorems we use.

*Department of Mathematics, Yale University, New Haven, CT 06520-8283, monteneg@math.yale.edu; supported in part by a VIGRE grant

†Division of Informatics, King's Buildings, University of Edinburgh, Edinburgh EH9 3JZ, UK, jun@dcs.ed.ac.uk.; supported in part by a divisional grant

STOC'01, July 6-8, 2001, Hersonissos, Crete, Greece.

2. PRELIMINARIES

In this paper we consider finite state Markov chains $\mathcal{M}$ with state space V and transition probabilities p_{ij}. We assume that our chains are aperiodic and connected and so have a steady state distribution π. Moreover all chains will be assumed to be time reversible, that is if $q_{ij} = \pi_i p_{ij}$ is the flow from i to j then $q_{ij} = q_{ji}$ for every pair of states (i, j).

For probability distributions σ and τ on V the *variation distance* is

$$\|\sigma - \tau\|_{TV} = \frac{1}{2} \sum_{v \in V} |\sigma(v) - \tau(v)|$$

This is a measure of how far σ and τ are from equality.

The *mixing time* $\tau(\epsilon)$ measures how many steps it takes a Markov chain to come close to the uniform distribution.

$$\tau(\epsilon) = \max_{i \in V} \min \left\{ t : \|P_i^t - \pi\| \le \epsilon \right\}$$

where P_i is the probability distribution with mass 1 on i and 0 elsewhere, and P_i^t is the distribution after t steps of the Markov chain. By convention we will define the *mixing time* τ to be $\tau(e^{-1})$.

A common technique for bounding the mixing time on complicated Markov chains involves conductance. For disjoint sets $A, B \subseteq V$ we define the flow from A to B by $Q(A, B) = \sum_{i \in A, j \in B} \pi_i p_{ij}$. The *conductance function* of $\mathcal{M}$ is defined in terms of the flow,

$$\Phi(x) = \min_{\pi_0 \le \pi(S) \le x} \frac{Q(S, \overline{S})}{\pi(S)}$$

where $\pi_0 = \min_{v \in V} \pi_v$ and $\pi_0 \le x \le 1/2$. Also, the *conductance* Φ is defined by $\Phi = \Phi(1/2)$.

If a finite state Markov chain has a uniform (ie. constant) stationary distribution and all transition probabilities are either a constant p or 0 then bounding the conductance function is equivalent to bounding the edge-isoperimetry, as the following reduction makes clear

$$\Phi(x) = \min_{\pi_0 \le \pi(S) \le x} \frac{Q(S, \overline{S})}{\pi(S)} = p \min_{0 < |S| \le x|V|} \frac{|Cut(S)|}{|S|} \qquad (1)$$

Likewise bounding the conductance is equivalent to bounding cutset expansion.

The following two theorems can be used to bound the mixing time τ

THEOREM 2.1 (SINCLAIR & JERRUM). *The mixing time τ of any Markov chain is bounded by*

$$\tau \le \frac{2}{\Phi^2}(1 + \log(1/\pi_0))$$

THEOREM 2.2 (AVERAGE CONDUCTANCE). (LOVÁSZ & KANNAN) *The mixing time τ of any Markov chain is bounded by*

$$\tau \le K \left(14 \int_{\pi_0}^{1/2} \frac{dx}{x\Phi(x)^2} + \frac{4}{\Phi} \right)$$

where K is a constant independent of the Markov Chain.

Theorem 2.2 is essentially that given in [14], however we have corrected a minor mistake in their theorem (the $4/\Phi$ term was omitted) and have adjusted the constants to take into account the fact that our conductance function differs from theirs by roughly a factor of 2.

Many proofs of rapid mixing use the *spectral gap* λ or the *log-Sobolev constant* ρ.

$$\lambda = \inf_{Var(\phi) \ne 0} \frac{\mathcal{E}(\phi, \phi)}{Var(\phi)} \qquad \textit{and} \qquad \rho = \inf_{\mathcal{L}(\phi) \ne 0} \frac{\mathcal{E}(\phi, \phi)}{\mathcal{L}(\phi)}$$

where $\mathrm{Var}(\phi)$ is the variance of ϕ, $\mathcal{E}$ is the Dirichlet form, and $\mathcal{L}$ is the entropy

$$\mathcal{E}(\phi, \phi) = \frac{1}{2} \sum_{x,y} (\phi(y) - \phi(x))^2 \pi(x)\, P(x, y)$$

$$Var(\phi) = \frac{1}{2} \sum_{x,y} (\phi(y) - \phi(x))^2 \pi(y)\pi(x)$$

$$\mathcal{L}(\phi) = \sum_x |\phi(x)|^2 \log(|\phi(x)|^2 / \|\phi\|_2)\pi(x)$$

Some examples of log-Sobolev constants can be found in [4]. In particular it is shown that

THEOREM 2.3. *The mixing time τ of any Markov chain can be bounded by*

$$\tau \le \frac{1}{2\rho}(2 + \log\log(1/\pi_0))$$

where ρ is the log-Sobolev constant of the chain.

Until now it was extremely difficult to lower bound the log-Sobolev constants of complicated chains, but recent work [9] makes it no harder than bounding the conductance function. We rewrite the definitions from [9] in a form which is equivalent but where the relation to our current techniques is clearer.

$$g_1^+ = \inf_{\pi_0 \le x \le 1/2} \frac{\Phi(x)}{\sqrt{\log(1/x)}} \qquad \ell_1^+ = \inf_{\pi_0 \le x \le 1/2} \frac{\Phi(x)}{\log(1/x)} \quad (2)$$

THEOREM 2.4 (HOUDRÉ). *Let $\mathcal{M}$ be a Markov chain and let g_1^+ and ℓ_1^+ be as in (2). Then the log-Sobolev constant ρ is bounded by*

$$(i)\ \rho \ge \frac{1}{100}(g_1^+)^2 \qquad (ii)\ \rho \ge \frac{\lambda \ell_1^+}{2(\sqrt{\lambda} + 2\ell_1^+)} \ge \frac{\sqrt{\lambda}\ell_1^+}{12}$$

where we simplified the final inequality by using the fact $2\lambda \ge \Phi^2 \ge (\ell_1^+ \log 2)^2$.

Observe that Theorem's 2.3 and 2.4 (i) improve pure conductance bounds when $g_1^+ \approx \Phi$ and from our simplification of part (ii) we see the second log-Sobolev bound will improve on the spectral gap bound when $\lambda \approx \Phi^2$ and $\ell_1^+ \approx \Phi$.

3. GEOMETRY AND EDGE-ISOPERIMETRY ON LINEAR EXTENSIONS

The first case in which we use edge-isoperimetry to show faster mixing are Markov chains whose underlying graphs (G, V) have a natural geometric structure. The key to exploiting this geometric structure will be a type of isoperimetric inequality first used in [12], strengthened in [5] and to be further strengthened in this paper.

3.1 Isoperimetry

The key to bounding the edge isoperimetry is an *isoperimetric inequality* relating the surface area of a cut to the volume it encloses. First a few definitions.

DEFINITION 3.1. *Let F be a real-valued function on a convex set K. We say that F is* log-concave *if $\log F$ is a concave function on K. In particular concave functions are log-concave.*

DEFINITION 3.2. *Given a norm $\|\cdot\|$ define the* dual norm $\|\cdot\|^*$ *by :* $\quad \|x\|^* = \max\{ax : \|a\| = 1\}$

Previous uses of isoperimetry to bound conductance used the following Theorem of [5].

THEOREM 3.1 (DYER & FRIEZE). *Let $K \subseteq \mathbb{R}^n$ be a convex body and F a log-concave function on int K. Let $S \subseteq K$, with $\mu(S) \le \frac{1}{2}\mu(K)$, be such that $\partial S \setminus \partial K$ is a piecewise smooth surface σ, with $u(x)$ the Euclidean unit normal to σ at $x \in \sigma$. If $\mu'(S) = \int_\sigma F(x)\|u(x)\|^* dx$ and $\mu(S) = \int_S F(x)dx$ then*

$$\frac{\mu(S)}{\mu'(S)} \le \frac{1}{2}\, diam\, K$$

where the diameter diam K is measured with respect to $\|\cdot\|$.

When $F = 1$ and $\|u(x)\|^*$ isn't too big then the theorem says roughly that when K is cut by a surface S then the ratio of the $Volume\ of\ S$ to the $Surface\ Area\ of\ S$ is bounded above by $\frac{1}{2} diam\ K$. For the problems we consider the surface area will be related to $|Cut(S)|$ and the volume will be related to $|S|$ so this theorem will give a method of bounding Φ by (1).

To bound $\Phi(x)$ it will be necessary to bound $|Cut(S)|$ conditioned on the $|S|$ (as $|S| \le x|V|$), i.e. we need to find an edge-isoperimetric inequality. Lovász & Kannan [14] gave a result that amounts to a one dimensional generalization of Theorem 3.1 and used it to bound the conductance function of a random walk on convex sets. We require the n-dimensional form of their theorem.

THEOREM 3.2. *Suppose the conditions of Theorem 3.1 are satisfied and moreover $\mu(S) \le x\mu(K)$ where $x \le 1/2$. Then*

$$\frac{\mu(S)}{\mu'(S)} \le \frac{1}{1 + \log(1/x)}\, diam\, K$$

Proof This can be proven by a method similar to the proof in [5] or by using the Localization Lemma of [13]. We omit the proof as the details would distract us from our main goal of applying geometry to bound the conductance function.

The interested reader will find that the slightly weaker $\frac{\mu(S)}{\mu'(S)} \le \frac{3}{2}\frac{1}{log(1/x)}\, diam\, K$ can be proven by slightly modifying the proof of [5]. The modified proof will require Lemma 3.3 in [14] and the initial "trivial" case is when $\mu(B) \ge \frac{2}{3e}\mu(K)$ which follows from Theorem 3.1. □

This is tight to within a constant factor, for as Example 3.3 will show, edge-isoperimetric inequalities on $[k]^n$ found in [3] show that any inequality of the form $C/\log(1/x)$ cannot do better than $e^{-1}/\log(1/x)$.

The connection between Theorem 3.2 and edge-isoperimetry can be seen clearly by an example. This is done by associating vertices of graphs with simplexes such that two vertices are adjacent exactly when their associated simplexes share a face. The technique is similar to that developed in [5, 12] to bound the conductance of a graph, our contribution is in extending these inequalities to edge-isoperimetry.

Example 3.3 : Let G be the grid $[k]^n$, the n-dimensional cube of side length k, and write the vertices of G in Cartesian product form so that

$$G = \{v = (v_1, v_2, \ldots, v_n) : v_i \in [1, \ldots, k]\}$$

To each vertex $v \in G$ associate the polytope

$$P(v) = \{x \in \mathbb{R}^n : v_i - 1 \le x_i \le v_i \text{ for all } i \in [1, \ldots n]\}$$

and denote the image of G by $\Omega = \bigcup_{v\in G} P(v) = [0, k]^n \subset \mathbb{R}^n$.

Properties such as adjacency and cut size in G carry over well to Ω. Cuts S of G with $|S|/|G| \le x$ map to cuts $P(S)$ of Ω with

$$vol_n P(S)/vol_n\Omega = |S|/|G| \le x$$

Two vertices $v^1, v^2 \in G$ are adjacent if and only if $P(v^1)$ and $P(v^2)$ intersect at a face, and

$$|Cut(S)| = vol_{n-1}\left(\partial P(S) \cap int(\Omega)\right) \tag{3}$$

The right hand side of (3) is just $\mu'(P(S))$ when $F = 1$ (observe $\|u\|^* = 1$ for $F = 1$), while $\mu(P(S)) = vol_n P(S)$. This suggests the use of Theorem 3.2 to bound $|Cut(S)|$.

$$\frac{\mu(P(S))}{\mu'(P(S))} = \frac{vol_n P(S)}{vol_{n-1}\left(\partial P(S) \cap int(\Omega)\right)} \le \frac{diam_\infty \Omega}{1 + \log(1/x)} \tag{4}$$

Algebraic manipulation of (3) and (4), along with $diam_\infty\Omega = k$ and $vol_n P(S) = |S|$ give

$$\frac{|Cut(S)|}{|S|} \ge \frac{1}{k}(1 + \log(1/x)) \text{ when } |S| \le x|G| \tag{5}$$

This is within a factor e of the correct inequality [3].

The properties important for use of the isoperimetric inequality in the previous example are captured by the following definition. The interested reader can make the definition more precise but we simplify it for clarity.

DEFINITION 3.3. *We say a graph (G, V) is* geometrically constructible *if there is a 1-1 mapping*

$$\phi : V \to \textit{Simplexes in } \mathbb{R}^n$$

such that adjacency is preserved, all simplexes have the same volume, the image of V is a convex body, and the ℓ_2-unit-normal along the boundary of the polytopes is a constant.

These conditions can be weakened somewhat without significantly affecting the results of this section. With this we can generalize the result of Example 3.3.

THEOREM 3.4. *Let (G, V) be a geometrically constructible graph. Suppose the method of Example 3.3 with the isoperimetric inequality of Dyer & Frieze gives cutset expansion*

$$\frac{|Cut(S)|}{|S|} \ge \gamma \qquad \textit{when } |S| \le |V|/2$$

Then the graph has edge-isoperimetric inequality

$$\frac{|Cut(S)|}{|S|} \ge \frac{\gamma}{2}(1 + \log(1/x)) \qquad \textit{when } |S| \le x|V| \le |V|/2$$

Proof The only difference in the two approaches is the Theorem used at (4). Theorem 3.2 differs from the theorem of Dyer & Frieze by $(1 + \log(1/x))/2$ and so the cutset expansion and edge-isoperimetric inequalities differ by this same amount. □

3.2 Rapid Mixing

We are now in a position to apply the theorems from the preliminaries to obtain bound on the mixing time and log-Sobolev constants of these geometric Markov chains. The family of geometric Markov chains we will consider is :

Definition 3.4. *A Markov chain $\mathcal{M}$ is called* geometrically constructible *if the underlying graph is geometrically constructible, it has uniform stationary distribution, and all transition probabilities are equal (ie. 0 or p for some constant p).*

As before, all three conditions can be weakened somewhat without effecting the results significantly.

Example 3.5 : Consider the grid $[k]^n$ of Example 3.3. Define a random walk on this graph with equal transition probability $1/(4n)$ to any neighbor. This Markov chain is geometrically constructible with $p = 1/(4n)$, so by (1) and (5) $\Phi(x) \geq (1 + \log(1/x))/(4nk)$. Applying the Average Conductance Theorem with $\Phi = \Phi(1/2)$ we get

$$\tau \leq K\left(14\int_{\pi_0}^{1/2}\frac{dx}{x\Phi(x)^2} + \frac{4}{\Phi}\right) \leq 260\,K\,k^2n^2 = O(k^2n^2)$$

This is not far from the correct bound of $O(k^2 n \log n)$.

As before, we can extend this technique to general geometric Markov chains. Note that from here on we will use the notation Φ_g to denote the lower bound on conductance found by the approach of Examples 3.3, 3.5.

Theorem 3.6. *Let $\mathcal{M}$ be a geometrically constructible Markov chain. Suppose the method of Examples 3.3, 3.5 with the isoperimetric inequality of Theorem 3.1 shows $\Phi \geq \Phi_g$. Then*

$$\Phi(x) \geq \frac{1}{2}\,\Phi_g(1 + \log(1/x)); \qquad \tau \leq 38\,K/\Phi_g^2$$

and

$$(i)\ \rho \geq \Phi_g^2/36 \qquad (ii)\ \rho \geq \sqrt{\lambda}\ \Phi_g/24$$

Proof Let γ be as in Theorem 3.4. Then (1) and Theorem 3.4 give

$$\begin{aligned}\Phi(x) &= p\min_{0<|S|\leq x|V|}\frac{|Cut(S)|}{|S|}\\ &\geq p\frac{\gamma}{2}(1+\log(1/x)) = \frac{1}{2}\Phi_g(1+\log(1/x))\end{aligned}$$

The bound on τ follows by substituting this expression and $\Phi \geq \Phi_g$ into the Average Conductance Theorem. The second bound on ρ follows by using the lower bound on $\Phi(x)$ to bound ℓ_1^+ and substituting this into Theorem 2.4. The first bound on ρ comes from substituting $\lambda \geq \Phi^2/2$ into the second bound (or with a weaker constant if Theorem 2.4 (i) is used). □

Observe that Theorem 3.6 can do better than spectral gap, whereas conductance alone in Theorem 2.1 cannot. To see this notice that in Example 3.5 the mixing time was found to be $\tau = O(k^2\,n^2)$, on the other hand the spectral gap is $\lambda = \Omega(1/k^2\,n)$ so $\tau = O(k^2\,n^2\ \log k)$ which is a weaker result.

Example 3.7 : One Markov chain where geometry has been used to find upper bounds on the mixing time is a random walk on Linear Extensions [10, 12]. Given a partially ordered set $(V, \prec)$, $V = [n]$ the set of linear extensions of $\prec$ is defined by

$$\Omega = \{g \in Sym\,V : g(i) \prec g(j) \Rightarrow i \leq j, \text{ for all } i, j \in V\}$$

ie. the set of permutations on V that preserve the partial ordering.

Sample from Ω u.a.r. as follows. If X_t is the current state choose a transposition $(i, i+1)$ u.a.r., if $X_t \circ (i, i+1) \in \Omega$ then with probability 1/2 set $X_{t+1} = X_t \circ (i, i+1)$, otherwise $X_{t+1} = X_t$. This Markov chain is time reversible and symmetric so it has the uniform stationary distribution and all edges (i, j) have identical weight $1/(n|\Omega|)$.

It was found in [10] that $\Phi_g = 1/(2n(n-1))$. Thus

$$\tau \leq 152\,K\,n^2(n-1)^2 = O(n^4)$$

$$\textit{and} \qquad \rho \geq 1/(144n^2(n-1)^2) = \Omega(1/n^4)$$

This is a large improvement over the previous conductance bound of $\tau = O(n^5 \log n)$ and even beats the path-coupling and comparison bound [2] of $\tau = O(n^4 \log^2 n)$. It is also quite close to the correct bound [17] of $\Theta(n^3 \log n)$.

4. INDUCTION AND EDGE-ISOPERIMETRY ON BALANCED MATROIDS

A second means for bounding the cutset expansion is by induction, in particular we will use an inductive argument similar to [7] to bound the cutset expansion of balanced matroids.

There are many equivalent definitions of matroids. Here we follow a description given in [11]. A *matroid* M on a ground set $E(M)$ is entirely defined by its set of *bases* $\mathcal{B}(M) \subseteq 2^{E(M)}$ and the following two conditions: 1) all bases have the same size, namely the *rank* of M, 2) for every pair of bases $X, Y \in \mathcal{B}(M)$ and every element $e \in X$, there exists an element $f \in Y$ s.t. $(X \cup \{f\}) \setminus \{e\} \in \mathcal{B}(M)$. The *bases-exchange graph* $G(M)$ of M has as vertex set $\mathcal{B}(M)$ and two bases are connected by an edge if they differ in exactly one element (see condition 2). Two basic operations on matroids are contraction and deletion. If $e \in E(M)$, then the matroid $M \setminus e$ obtained by *deleting* e has ground set $E(M \setminus e) = E(M) \setminus \{e\}$ and bases $\mathcal{B}(M \setminus e) = \{X \subseteq E(M \setminus e) \mid X \in \mathcal{B}(M)\}$. The matroid M/e obtained by *contracting* e has ground set $E(M/e) = E(M) \setminus \{e\}$ and bases $B(M/e) = \{X \subseteq E(M/e) \mid X \cup \{e\} \in B(M)\}$. Any matroid obtained from a series of contractions and deletions is a *minor* of M.

If X is a basis uniformly chosen at random from $\mathcal{B}(M)$ and e is an element of $E(M)$, let by abuse of notation e denote the event $e \in X$, i.e. e is in the chosen basis. A matroid M is *negatively correlated* if for all pairs of distinct elements $e, f \in E(M)$ the inequality $\Pr[ef] \leq \Pr[e]\Pr[f]$ holds. A matroid M is said to be *balanced* if itself and all its minors are negatively correlated.

We define a Markov chain on the bases exchange graph as follows. Suppose the current state is $X \in \mathcal{B}(M)$, then choose a basis element $b \in X$ and an edge $e \in E(M)$ uniformly at random. If $X' = (X \setminus b \cup e) \in \mathcal{B}(M)$ then move to X' with probability 1/2, otherwise stay at X.

This Markov chain has been shown to mix rapidly by several authors, the strongest bounds on the mixing time were shown in [7]. We will apply Average Conductance and log-Sobolev techniques to these problems to obtain new mixing time bounds.

4.1 Edge-Isoperimetry

As in the geometric case, we first need an edge-isoperimetric inequality for cuts in balanced matroids. Similar to the proof of lemma 3.2 in [7], the proof of this inequality is done in an inductive fashion.

THEOREM 4.1 (MATROID EDGE-ISOPERIMETRY). *Let $G(M)$ be the bases-exchange graph of any balanced matroid M with bases $\mathcal{B}$. For all subsets $\mathcal{S} \subset V(G(M))$ such that $0 < |S| \leq |V(G(M))|/2$*

$$\frac{Cut(\mathcal{S})}{|\mathcal{S}|} \geq \log_2\left(\frac{|\mathcal{B}|}{|\mathcal{S}|}\right).$$

Proof We proceed by induction on the size of the ground set of M. For the base-case, $|M| = 1, 2$, the hypothesis is trivially true. Induction step, $|M| > 2$: Let $\mathcal{S} \subset \mathcal{B}$ be a collection of bases, with $|\mathcal{S}| \leq |\mathcal{B}|/2$, defining a cut in the bases-exchange graph of M. Let $\mathcal{S}_e = \mathcal{S} \cap \mathcal{B}_e$ and $\mathcal{S}_{\overline{e}} = \mathcal{S} \cap \mathcal{B}_{\overline{e}}$, where $\alpha|\mathcal{B}|$ and $(1-\alpha)|\mathcal{B}|$, $\alpha \in [0,1]$, are the sizes of $\mathcal{B}_e$ and $\mathcal{B}_{\overline{e}}$ respectively, and define x, y by $|\mathcal{S}_e| = x|\mathcal{B}_e|$ and $|\mathcal{S}_{\overline{e}}| = y|\mathcal{B}_{\overline{e}}|$, $x, y \in [0,1]$. The edges forming the cut are of three kinds: (i) those whose endpoints are both within $\mathcal{B}_e$, (ii) those whose endpoints are both within $\mathcal{B}_{\overline{e}}$ and (iii) those which span $\mathcal{B}_e$ and $\mathcal{B}_{\overline{e}}$. Since, as mentioned above, $\mathcal{B}_e$ and $\mathcal{B}_{\overline{e}}$ are isomorphic to $\mathcal{B}(M/e)$ and $\mathcal{B}(M \setminus e)$, they give rise to minors of M and the induction hypothesis is applicable. By induction hypothesis, the numbers of edges of kinds (i) and (ii) are $-\min\{x, 1-x\}\log_2(\min\{x, 1-x\})|\mathcal{B}_e|$ and $-\min\{y, 1-y\}\log_2(\min\{y, 1-y\})|\mathcal{B}_{\overline{e}}|$ respectively. To lower bound the number of edges of kind (iii), assume first that $x \geq y$. By [7](lemma 3.1), there are at least $x|\mathcal{B}_{\overline{e}}|$ bases in $\mathcal{B}_{\overline{e}}$ adjacent to some bases in $\mathcal{S}_e$; of these, at least $(x-y)|\mathcal{B}_{\overline{e}}|$ must lie outside $|\mathcal{S}_{\overline{e}}|$. Thus there are at least $(x-y)|\mathcal{B}_{\overline{e}}|$ edges of type (iii). This argument can equally well be applied in the opposite direction, starting at the set $\mathcal{B}_{\overline{e}} \setminus \mathcal{S}_{\overline{e}}$, yielding a second lower bound of $(x-y)|\mathcal{B}_e|$. Thus the number of edges of kind (iii) is at least $(x-y)\max\{|\mathcal{B}_e|, |\mathcal{B}_{\overline{e}}|\}$. Since the case $x < y$ is entirely symmetric, we obtain, summing the contributions from edges of kinds (i)-(iii):

$$\begin{aligned}|Cut(\mathcal{S})| \geq\ & -\min\{x, 1-x\}\log_2(\min\{x, 1-x\})|\mathcal{B}_e| \\ & -\min\{y, 1-y\}\log_2(\min\{y, 1-y\})|\mathcal{B}_{\overline{e}}| \\ & +|x-y|\max\{|\mathcal{B}_e|, |\mathcal{B}_{\overline{e}}|\}.\end{aligned}$$

To complete the proof, we must show that $|Cut(\mathcal{S})|$ is always at least $-(x\alpha + y(1-\alpha))\log_2(x\alpha + y(1-\alpha))|\mathcal{B}|$, whenever $|\mathcal{S}| \leq |\mathcal{B}|/2$. Note that this last condition may be expressed as

$$(\frac{1}{2} - x)|\mathcal{B}_e| + (\frac{1}{2} - y)|\mathcal{B}_{\overline{e}}| \geq 0. \tag{6}$$

This inequality shows that only one of x or y can be greater than $1/2$. It remains to establish that

$$\begin{aligned}& -\min\{x, 1-x\}\log_2(\min\{x, 1-x\})|\mathcal{B}_e| \\ & \quad -\min\{y, 1-y\}\log_2(\min\{y, 1-y\})|\mathcal{B}_{\overline{e}}| \\ & \quad +|x-y|\max\{|\mathcal{B}_e|, |\mathcal{B}_{\overline{e}}|\} \\ \geq\ & -(x\alpha + y(1-\alpha))\log_2(x\alpha + y(1-\alpha))|\mathcal{B}|.\end{aligned}$$

As the cases $0 \leq \alpha \leq 1/2$ and $1/2 < \alpha \leq 1$ are entirely symmetrical, we have to scrutinize only four of the eight cases. But before, notice that in the degenerate case, $\alpha = 1$ (or equally $\alpha = 0$), the induction hypothesis becomes immediately applicable. Therefore we can restrict ourselves to $1/2 \leq \alpha < 1$ and obtain the following four cases: 1.) $1/2 \geq x \geq y \geq 0$, 2.) $1/2 \geq y \geq x \geq 0$, 3.) $x > 1/2,\ y < 1/2$ and 4.) $y > 1/2,\ x < 1/2$. Each of these corresponds to one of the four lemmas below:

LEMMA 4.1. *For $1/2 \leq \alpha < 1$ and $1/2 \geq x \geq y \geq 0$*

$$\begin{aligned}f_1(x, y, \alpha) &= -\alpha x \log_2 x - (1-\alpha)y\log_2 y + (x-y)\alpha \\ &\quad +(x\alpha + y(1-\alpha))\log_2(x\alpha + y(1-\alpha)) \\ &\geq 0\end{aligned}$$

LEMMA 4.2. *For $1/2 \leq \alpha < 1$ and $1/2 \geq y \geq x \geq 0$*

$$\begin{aligned}f_2(x, y, \alpha) &= -\alpha x \log_2 x - (1-\alpha)y\log_2 y + (y-x)\alpha \\ &\quad +(x\alpha + y(1-\alpha))\log_2(x\alpha + y(1-\alpha)) \\ &\geq 0\end{aligned}$$

LEMMA 4.3. *For $1/2 \leq \alpha < 1$ and $x > 1/2,\ y < 1/2$*

$$\begin{aligned}& f_3(x, y, \alpha) = -\alpha(1-x)\log_2(1-x) - (1-\alpha)y\log_2 y \\ & \quad +(x-y)\alpha + (x\alpha + y(1-\alpha))\log_2(x\alpha + y(1-\alpha)) \\ & \geq 0\end{aligned}$$

LEMMA 4.4. *For $1/2 \leq \alpha < 1$ and $y > 1/2,\ x < 1/2$*

$$\begin{aligned}& f_4(x, y, \alpha) = -\alpha x\log_2 x - (1-\alpha)(1-y)\log_2(1-y) \\ & \quad +(y-x)\alpha + (x\alpha + y(1-\alpha))\log_2(x\alpha + y(1-\alpha)) \\ & \geq 0\end{aligned}$$

Proving these lemmas shows that the induction step is valid and thus concludes the proof of Theorem 4.1. As the proofs of the lemmas are rather technical, we will not give them here and refer the interested reader to the appendix. □

4.2 Rapid Mixing

As in the geometric case we now have all that is needed to show rapid mixing.

THEOREM 4.2. *The mixing time of the bases-exchange walk on any balanced matroid of rank n on a ground set of size m is at most $\tau \leq C\, m^2 n^2$ for some constant C independent of the matroid.*

Proof By Theorem 4.1 $|Cut(\mathcal{S})|/|\mathcal{S}| \geq \log_2(1/x)$. The Markov chain has $p = 1/(2mn)$, so by the Preliminaries

$$\Phi(x) \geq \frac{1}{2mn} \inf_{\pi_0 \leq \pi(\mathcal{S}) \leq x} \frac{|Cut(\mathcal{S})|}{|\mathcal{S}|} \geq \frac{\log_2(1/x)}{2mn} \tag{7}$$

Substituting (7) into the Average Conductance theorem gives the result. □

This Theorem is stronger than [7] Theorem 5.1 ($\tau = O(n^3 m \log m)$) when $n \log m = \Omega(m)$, eg. when $m = O(n \log n)$. In the case of graphic matroids this would be the case when the average degree of vertices is $O(\log n)$. However we can get a stronger result with log-Sobolev constants.

THEOREM 4.3. *The log-Sobolev constant and mixing time of the bases-exchange walk on any balanced matroid of rank n on a ground set of size m are bounded by*

$$\rho \geq \frac{1}{24\, m^{3/2} n^2} \qquad \tau \leq 24\, m^{3/2} n^2 (\log n + \log\log m)$$

Proof By (7) we see $\ell_1^+ = 1/[(2\log 2)m\,n]$. It was shown in [7] that $\lambda \geq 1/m\,n^2$. Thus

$$\rho \geq \sqrt{\lambda}\,\ell_1^+/12 \geq 1/(24m^{3/2}\,n^2)$$

and

$$\tau \leq 12m^{3/2}n^2(2+\log\log(m^n)) \leq 24m^{3/2}n^2(\log n+\log\log m)$$

□

This is stronger than [7] Theorem 5.1 ($\tau = O(n^3 m\log m)$) when $n\log m = \Omega(\sqrt{m}\log n)$, eg. when $m = O(n^2)$. According to a result by Heller [8] this is true for *simple* regular matroids (i.e. matroids without loops and parallel elements): $m \leq n(n+1)$, which is smaller than $2n^2$ for $n \geq 1$,and implies that $m = O(n^2)$ if the size of all parallel classes is bounded by a constant. In particular this includes all graphic matroids with few multiple edges.

5. CONCLUDING REMARKS

5.1 Observations

This paper bounded the edge-isoperimetry of two problems of combinatorial interest. A significant amount of work has been done on studying the edge-isoperimetry of various graphs, some information on this topic and a list of references can be found in [1].

From the matroid example we found that Theorem 2.4 (ii) can do better than Average Conductance when we have knowledge of the spectral gap. This isn't the case for Theorem 2.4 (i). From the definitions of g_1^+ and ℓ_1^+ we have $\Phi(x) \geq g_1^+\sqrt{\log(1/x)}$ and $\Phi(x) \geq \ell_1^+\log(1/x)$. Substituting these bounds into the Average Conductance Theorem and integrating gives

$$\tau \leq C \begin{cases} \frac{1}{\Phi^2}\log \pi_0^{-1} \\ \frac{1}{g_1^{+2}}\log\log \pi_0^{-1} \\ \frac{1}{\ell_1^{+2}} \end{cases} \tag{8}$$

with the constant C a constant independent of the Markov chain. The second term is just the mixing time bound given by Theorem 2.4 (i) and Theorem 2.3 so Average Conductance is always at least as good.

The quantities in (8) show g_1^+ and ℓ_1^+ are in a sense natural analogs of conductance. The different bounds are best under different circumstances. Both examples in this paper had $\ell_1^+ = \Omega(\Phi_g)$, in this case the third bound is the best. The first bound is best for the random walk on the complete graph K_n, we don't know of a situation when the second bound is best.

However the Average Conductance Theorem says more than this simple generalization. Consider the random walk on the barbell given by two complete graphs K_n connected by a single edge, but with transition probability along the central edge of ϵ/n. Then $\Phi(x) = \Omega(1)$ for $x < 1/2$, $\Phi(1/2) = \Theta(\epsilon/n^2)$ and Theorem 2.2 gives $\tau = O(n^2/\epsilon) = O(1/\Phi)$, the correct bound. This shows that in special cases Average Conductance can even hit the lower bound in $1/\Phi \leq \tau \leq (2/\Phi^2)(1+\log(1/\pi_0))$.

5.2 Problems

The two geometric examples show that the results in the paper are very close to the correct bounds. We ask whether our bounds in this and the matroid problem can be strengthened further.

Problem 1 : In both Examples 3.5 and 3.7 using geometry with transition probability p to bound mixing gave results $\Theta(p\log(1/p))$ from the correct bounds. We ask if Theorem 3.6 can be strengthened further, for example is it true that

$$\tau = O\left(\frac{p\log(1/p)}{\Phi_g^2}\right) \qquad \textit{and} \qquad \rho = \Omega\left(\Phi_g^2/p\right)$$

This may be true because we worked in a d-dimensional space, so adding dimensions should make the mixing time increase by $O(d\log d)$, but $1/\Phi^2$ overstates this by increasing the mixing time by $O(d^2)$ (Notice that $p = \Omega(1/d)$ in our examples so $p\log(1/p)$ would correct this overstatement).

Problem 2 : In Theorem 4.3 we have $m^{3/2}$ while the bound in [7] has only m^1. We ask if the fractional power can be removed and the bounds improved to

$$\rho = \Omega\left(\frac{1}{m\,n^2}\right) \qquad \textit{and} \qquad \tau = O(m\,n^2\,\log m)$$

Problem 2 has recently been solved by the second author who used an inductive argument on the second eigenvalue to show

$$\lambda \geq 1/m\,n \qquad \textit{and} \qquad \tau = O(m\,n^2\,\log m)$$

This is tight as the following example shows. Take the line of length n and copy every edge to make it a double edge (for a total of $2n$ edges). The graphic matroid on this graph has rank n, $m = 2n$ and the spectral gap of the Markov chain can easily be shown to be $\lambda = \Theta(1/n^2) = \Theta(1/m\,n)$ (it's just the $n-dim$ cube with delay $\Theta(n)$ between moves).

Observe that with this new spectral gap we can improve Theorem 4.3 a bit to

$$\rho \geq 1/(24m^{3/2}\,n^{3/2})$$

This is slightly weaker than the bound on ρ in Problem 2, since $m \geq n$. It is also unclear if either bound on ρ is tight since for the cube matroid both give $\rho = \Omega(1/n^3)$, when the correct bound is $\rho = \Theta(1/n^2)$. This raises another problem.

Problem 3 : In all examples where we know the log-Sobolev constant ρ exactly we have found the fractional powers left by $\sqrt{\lambda}$ in Theorem 2.4(ii) to be unnecessary. Is it true that

$$\rho = \Omega\left(\frac{1}{m\,n}\right) \qquad \textit{and} \qquad \tau = O(m\,n\,\log m)$$

This would put ρ at the upper bound of $\lambda/\log(1/\pi_0) \leq \rho \leq \lambda/2$ [4] and would give a tight bound on both the mixing time and log-Sobolev constants of the cube matroid.

6. ACKNOWLEDGEMENTS

The authors thank Ravi Kannan and Mark Jerrum for the many helpful suggestions that led to this work. Part of this work was done while the first author was a guest at the University of Edinburgh.

7. REFERENCES

[1] S.L. Bezrukov, *Edge Isoperimetric Problems on Graphs, in: Graph Theory and Combinatorial Biology*, Bolyai Soc. Math. Stud., 7:157–197, 1999.

[2] R. Bubley and M. Dyer, *Faster Random Generation of Linear Extensions*, Technical Report 97.41, School of Computer Studies, University of

Leeds, August 1997. Available at
ftp://agora.leeds.ac.uk/scs/doc/reports/1997/97_41.ps.Z

[3] B. Bollobás and I. Leader, *Edge-Isoperimetric Inequalities in the Grid*, Combinatorica, 11(4):299–314, 1991.

[4] P. Diaconis and L. Saloff-Coste, *Logarithmic Sobolev inequalities for finite Markov chains*, Ann. Appl. Probab., 6(3):695–750, 1996.

[5] M. Dyer and A. Frieze, *Computing the Volume of Convex Bodies : A Case Where Randomness Provably Helps*, Proc. Sympos. Appl. Math., 44:123–169, 1991.

[6] M. Dyer, A. Frieze and R. Kannan, *A Random Polynomial-time Algorithm for Approximating the Volume of Convex Bodies*, J. Assoc. Comput. Mach., 38(1):1–17, 1991.

[7] T. Feder and M. Mihail, *Balanced Matroids*, STOC '92, pp. 26–38, 1992.

[8] I. Heller, *On linear systems with integral valued solutions*, Pacific Journal of Mathematics, 7:1351–1364, 1957.

[9] C. Houdré, *Mixed and Isoperimetric Estimates on the Log-Sobolev Constants of Graphs and Markov Chains*, To appear in Combinatorica.

[10] M. Jerrum, *Mathematical foundations of the Markov chain Monte Carlo method*, Probabilistic Methods for Algorithmic Discrete Mathematics, Algorithms and Combinatorics, 16:116–165, 1998.

[11] M. Jerrum, *Matroids*, Preprint.

[12] A. Karzanov and L. Khachiyan, *On the conductance of order Markov chains*, Order, 8(1):7–15, 1991.

[13] R. Kannan, L. Lovász and M. Simonovitz, *Isoperimetric problems for convex bodies and a localization lemma*, J. Discr. Comput. Geom., 13:282–287, 1999.

[14] L. Lovász and R. Kannan, *Faster Mixing via average Conductance*, STOC '99, 282–287, 1999.

[15] A. Sinclair, *Improved bounds for mixing rates of Markov chains and multicommodity flow*, Combin. Probab. Comput., 1(4):351–370, 1992.

[16] A. Sinclair and M. Jerrum, *Approximate Counting, Uniform Generation, and Rapidly Mixing Markov Chains*, Information and Computation, 82:93–133, 1989.

[17] Wilson, *Mixing times of lozenge tiling and card shuffling Markov chains*, Preprint.

APPENDIX

Proof of lemma 4.1. First extend f_1 to the boundary by continuity. Treating α like a constant, the hessian of $f_1(x,y,\alpha)$ is

$$h_1(x,y,\alpha) =$$

$$\begin{pmatrix} -\frac{\alpha}{x\ln 2} + \frac{\alpha^2}{(x\alpha+y(1-\alpha))\ln 2} & \frac{(1-\alpha)\alpha}{(x\alpha+y(1-\alpha))\ln 2} \\ \frac{(1-\alpha)\alpha}{(x\alpha+y(1-\alpha))\ln 2} & -\frac{1-\alpha}{y\ln 2} + \frac{(1-\alpha)^2}{(x\alpha+y(1-\alpha))\ln 2} \end{pmatrix}.$$

As $\frac{\partial^2}{\partial x^2} f_1(x,y,\alpha)$ and $\frac{\partial^2}{\partial y^2} f_1(x,y,\alpha)$ are less than zero, the hessian cannot be positive definite, i.e. f_1 has no local minimum in the interior. Thus, we merely have to check the boundaries: $x = y$, $(y = 0 \wedge 1/2 \geq x \geq 0)$ and $(x = 1/2 \wedge 1/2 \geq y \geq 0)$. But, since $\frac{\partial^2}{\partial x^2} f_1(x,y,\alpha) \leq 0$ and also $\frac{\partial^2}{\partial y^2} f_1(x,y,\alpha) \leq 0$, the claim holds if $f_1(x,y,\alpha) \geq 0$ for $(x = 1/2, y = 0)$ and $x = y$. It is easy to check that $f_1(1/2,0,\alpha) \geq 0$ since $-1 \leq \log_2 \alpha \leq 0$ and $f_1(x,y,\alpha) = 0$ for $x = y$. □

Proof of lemma 4.2. Similarly to the first case, extend f_2 to the boundary by continuity first and then form the hessian, which turns out to be the same as for f_1. Since the hessian of $f_2(x,y,\alpha)$ is the same as in the first case, the same arguments apply and it suffices to look at certain points of the boundary. This time these are: $x = y$ and $(x = 0, y = 1/2)$. For $x = y$ the value of f_2 equals zero, again. For $x = 0, y = 1/2$ the function $f_2(x,y,\alpha)$ turns into a function in α only, which we will call $g(\alpha)$. The first derivative of g shows that the only local extremum of g lies outside the allowed range for α: $\frac{d\ g}{d\alpha} = \frac{1}{2} - \frac{1}{2\ln 2}(\ln(1-\alpha)+1) = 0 \Leftrightarrow \alpha = 1 - e^{\ln 2 - 1} < \frac{1}{2}$. Checking the boundaries of g, $\alpha = 1/2$ and $\alpha \to 1$, reveals that the claim holds. □

For the remaining two cases recollect that inequality (6) on page ensured that only one of x or y is greater than 1/2.

Proof of lemma 4.3. Extend f_3 to the boundary by continuity. As $|\mathcal{S}| \leq |\mathcal{B}|/2$, the valid range for x and y is given by $x\alpha + y(1-\alpha) \leq 1/2$, which in turn yields $x \leq 1/(2\alpha)$. To start with, we again form the hessian:

$$h_3(x,y,\alpha) =$$

$$\begin{pmatrix} -\frac{\alpha}{(1-x)\ln 2} + \frac{\alpha^2}{(x\alpha+y(1-\alpha))\ln 2} & \frac{(1-\alpha)\alpha}{(x\alpha+y(1-\alpha))\ln 2} \\ \frac{(1-\alpha)\alpha}{(x\alpha+y(1-\alpha))\ln 2} & -\frac{1-\alpha}{y\ln 2} + \frac{(1-\alpha)^2}{(x\alpha+y(1-\alpha))\ln 2} \end{pmatrix}$$

Noticing that $\frac{\partial^2}{\partial y^2} f_3(x,y,\alpha) \leq 0$ and $\frac{\partial^2}{\partial x^2} f_3(x,y,\alpha) = \frac{\alpha(\alpha - 2\alpha x + \alpha y - y)}{(1-x)(x\alpha+y(1-\alpha))\ln 2} \leq 0$, since $2x \geq 1$ and $1/2 \leq \alpha < 1$, the hessian $h_3(x,y,\alpha)$ cannot be positive definite. So, the minima of f_3 can only be found on points of the boundary: $(y = 0, x = 1/2)$ and $x\alpha + y(1-\alpha) = 1/2$ $(\Leftrightarrow y = \frac{1-2x\alpha}{2-2\alpha})$. Since $1/2 \leq \alpha < 1$, for $y = 0, x = 1/2$, it is easy to verify that $f_3(1/2,0,\alpha) \geq 0$. The case $y = \frac{1-2x\alpha}{2-2\alpha}$ needs more consideration. Still treating α like a constant, f_3 becomes a function in x, henceforth called $g(x) = -\alpha(1-x)\log_2(1-x) - (1-\alpha)(\frac{1-2x\alpha}{2-2\alpha})\log_2(\frac{1-2x\alpha}{2-2\alpha}) + \frac{2x\alpha-1}{2-2\alpha}$. The second derivative of $g(x)$, $\frac{d^2\ g(x)}{dx^2} = -\frac{\alpha}{(1-x)\ln 2} - \frac{2\alpha^2}{(1-2x\alpha)\ln 2}$, is less than zero for $\alpha \geq 1/2 \wedge x \leq 1/(2\alpha)$. Checking the boundaries $x = 1/2$ and $x = 1/(2\alpha)$ shows that $g(\frac{1}{2}) = 0$ and $g(\frac{1}{2\alpha}) = -\alpha(1-\frac{1}{2\alpha})\log_2(1-\frac{1}{2\alpha}) \geq 0$, concluding the proof of lemma 4.3. □

Proof of lemma 4.4. Again, extend f_4 to the boundary by continuity. Due to inequality (6) on page , only one of x or y can be greater 1/2. The legal range for x and y is again $x\alpha + y(1-\alpha) \leq 1/2$, which yields the constraint $y \leq 1/(2(1-\alpha))$. Notice that $1/(2(1-\alpha)) \geq 1$ for $\alpha \geq 1/2$ so that in fact, y is upper bounded by 1. Thus for $\alpha > 1/2$ the boundary is a trapezoid. Looking at the hessian of $f_4(x,y,\alpha)$,

$$h_4(x,y,\alpha) =$$

$$\begin{pmatrix} -\frac{\alpha}{x\ln 2}+\frac{\alpha^2}{(x\alpha+y(1-\alpha))\ln 2} & \frac{(1-\alpha)\alpha}{(x\alpha+y(1-\alpha))\ln 2} \\ \frac{(1-\alpha)\alpha}{(x\alpha+y(1-\alpha))\ln 2} & -\frac{1-\alpha}{(1-y)\ln 2}+\frac{(1-\alpha)^2}{(x\alpha+y(1-\alpha))\ln 2} \end{pmatrix},$$

discloses that $\frac{\partial^2}{\partial x^2}f_4(x,y,\alpha) \leq 0$ and $\frac{\partial^2}{\partial y^2}f_4(x,y,\alpha) = \frac{(1-\alpha)[(1-2y)(1-\alpha)-x\alpha]}{(1-y)(x\alpha+y(1-\alpha))\ln 2} \leq 0$. Consequently, the hessian is not positive definite, i.e. f_4 has no local minima in the interior. The minima must therefore be on the following points of the boundary (trapezoid): $(x=0, y=1/2)$, $(x=0, y=1)$ and $x\alpha + y(1-\alpha) = 1/2$ ($\Leftrightarrow x = \frac{1-2y(1-\alpha)}{2\alpha}$). Checking that $f_4(0,1/2,\alpha)$ and $f_4(0,1,\alpha)$ are greater than 0 is straightforward. Again, $x = \frac{1-2y(1-\alpha)}{2\alpha}$ needs more effort. As before, regard f_4 now as a function $g(y) = -\alpha(\frac{1-2y(1-\alpha)}{2\alpha})\log_2(\frac{1-2y(1-\alpha)}{2\alpha}) - (1-\alpha)(1-y)\log_2(1-y)+y-1$ in y only. Forming the second derivative, $\frac{d^2\ g(y)}{dy^2} = -\frac{2(1-\alpha)^2}{(1-2y(1-\alpha))\ln 2} - \frac{1-\alpha}{(1-y)\ln 2}$, we see that $g(y)$ has no local minima for $1/2 \leq y < 1$. The strict inequality $y < 1$ is necessary because $\frac{d^2\ g(y)}{dy^2}$ is not defined for $y = 1$ and $\alpha = 1/2$. For the boundaries $y = 1/2$ and $y = 1$ we obtain $g(\frac{1}{2}) = 0$ and by continuity $g(1) = -\alpha\frac{2\alpha-1}{2\alpha}\log_2\frac{2\alpha-1}{2\alpha} \geq 0.\square$