ISRAEL JOURNAL OF MATHEMATICS **124** (2001), 189–201

RANGE OF CUBE-INDEXED RANDOM WALK

BY

J. KAHN*

Department of Mathematics and RUTCOR, Rutgers University
New Brunswick, NJ 08903, USA
e-mail: jkahn@math.rutgers.edu

ABSTRACT

For given finite, connected, bipartite graph $G = (V, E)$ with distinguished $v_0 \in V$, set

$$\mathcal{F} = \{f: V \to \mathbf{Z} | f(v_0) = 0, \{x, y\} \in E \Rightarrow |f(x) - f(y)| = 1\}.$$

Our main result says there is a fixed b so that when G is a Hamming cube ($\{0,1\}^n$ with the usual adjacency), and $\mathbf{f}$ is chosen uniformly from $\mathcal{F}$, the probability that $\mathbf{f}$ takes more than b values is at most $e^{-\Omega(n)}$. This settles in a very strong way a conjecture of I. Benjamini, O. Häggström and E. Mossel [2].

The proof is based on entropy considerations and a new log-concavity result.

1. Introduction

For a (finite) connected, bipartite graph $G = (V, E)$ with distinguished $v_0 \in V$, set

$$\mathcal{F} = \mathcal{F}(G, v_0) = \{f: V \to \mathbf{Z} | f(v_0) = 0, \{x, y\} \in E \Rightarrow |f(x) - f(y)| = 1\}.$$

(That is, $\mathcal{F}$ is the set of *graph homomorphisms* from G to $\mathbf{Z}$, normalized to vanish at v_0. Some terminology is given at the end of this section. For graph theory background see, e.g., [4].)

* Supported by NSF.
Received February 17, 2000

In [2] such functions are studied from a probabilistic point of view, a motivating idea being that, for suitable G, a typical member of $\mathcal{F}$ should exhibit stronger concentration behavior than can be guaranteed for an arbitrary member.

Of particular interest is what happens when G is a Hamming cube ($\{0,1\}^n$ with the usual adjacency), since this is the starting point for most discussions of concentration, at least in discrete settings. Cubes are also the main concern of the present paper, and for the remainder of this section (actually for all of the paper except Section 5) we specialize to this case, taking $v_0 = \underline{0}$.

Here for a general $f \in \mathcal{F}$ we have the usual large deviation inequality: for v uniform from V,

$$\Pr(|f(v) - \mathbf{E}f| > \lambda\sqrt{n}) < 2e^{-\lambda^2/2}$$

for each $\lambda > 0$. These bounds are, of course, quite accurate; but a little reflection suggests that a *typical* f should be nearly constant on either even or odd vertices (and mainly just take two values on vertices of the other parity). Nonetheless, even the following conjecture from [2] does not seem easy to prove.

For $f \in \mathcal{F}$, set

$$R(f) = |\{f(x) : x \in V\}|.$$

Write $\mathbf{f}$ for a function chosen uniformly at random from $\mathcal{F}$. (Our title thinks of $\mathbf{f}$ as a "cube-indexed random walk," the name again taken from [2].)

CONJECTURE 1.1 ([2]): *For each* $t > 0$, $\Pr(R(\mathbf{f}) > tn) \to 0$ $(n \to \infty)$.

See [2] for further discussion. As suggested there (and above), one expects something much stronger than Conjecture 1.1 to be true. The purpose of the present paper is to prove such a statement:

THEOREM 1.2: *There is a constant b for which* $\Pr(R(\mathbf{f}) > b) < e^{-\Omega(n)}$.

CONJECTURE 1.3: $\Pr(R(\mathbf{f}) > 5) < e^{-\Omega(n)}$ *and* $\Pr(R(\mathbf{f}) = 5) > \Omega(1)$.

The proof of Theorem 1.2 is mainly based on entropy considerations. The basic approach was introduced in [8] (see also [7]) to deal with the following problem of C. Athanasiadis [1]. Define a *rank function* on $2^{[n]}$ to be an f: $2^{[n]} \to \mathbf{N}$ for which $f(\emptyset) = 0$ and $f(A) \leq f(A \cup x) \leq f(A) + 1$. Athanasiadis conjectured that the number of such functions is

$$\exp_2[2^{n-1}(1 + o(1))]. \tag{1}$$

This was proved in [8], where it was further conjectured that the number is in fact $O(2^{2^{n-1}})$.

As pointed out to the author by E. Mossel, mapping $f \to g$ with $g(A) = 2f(A) - |A|$ gives a bijection between rank functions and cube-indexed walks. So the number of cube-indexed walks is also at most (1), though this by itself does not seem to help with Conjecture 1.1. It follows from the present arguments that the $o(1)$ term in (1), which in [8] was $O(n^{-1/2})$, could be improved to C^{-n} for some constant $C > 1$.

The rest of the paper is organized as follows. It has seemed best to first give (in Section 2) the proof of Theorem 1.2 modulo proofs of the main entropy-based inequality (2) and one simple combinatorial fact (Proposition 2.1) which may be of independent interest. Entropy notions are then reviewed in Section 3 and applied in Section 4 to prove (2). Finally, the proof of Proposition 2.1 is given in Section 5.

USAGE. We use $\sim$ for adjacency and N_v for the set of neighbors of (vertices adjacent to) v. A vertex is *even* or *odd* depending on its distance from v_0. As usual, $[n] = \{1, \ldots, n\}$ and 2^S is the set of subsets of S. We use boldface for expectation ($\mathbf{E}$) and for random variables ($\mathbf{X}, \mathbf{f}, \ldots$).

2. Proof of Theorem

Here again G is the n-dimensional Hamming cube and $\mathbf{f}$ is chosen uniformly from $\mathcal{F} = \mathcal{F}(G, v_0 = \underline{0})$.

For $v \in V$ set $N_v = \{w \in V : w \sim v\}$ and define the event

$$Q_v = \{\mathbf{f} \text{ is constant on } N_v\}.$$

We also write $Q_{\overline{v}}$ for the complementary event and set (for $v \neq z \in V$) $Q_{vz} = Q_v Q_z$, $Q_{v\overline{z}} = Q_v Q_{\overline{z}}$ etc.

Most of our work is devoted to showing that for $v \sim z$, Q_{vz} and $Q_{\overline{v}\overline{z}}$ are unlikely; precisely,

$$\Pr(Q_{vz}) + \Pr(Q_{\overline{v}\overline{z}}) < e^{-\Omega(n)}. \tag{2}$$

Proof of this is deferred to Section 4.

Now (2) easily implies

$$\forall v \in V \quad \Pr(|\mathbf{f}(v)| > 2) < e^{-\Omega(n)}. \tag{3}$$

To see this, consider a shortest path $v_0 \sim v_1 \sim \cdots \sim v_k = v$. By (2) we have with probability $1 - e^{-\Omega(n)}$

$$Q_{v_{i-1}\overline{v}_i} \cup Q_{\overline{v}_{i-1}v_i}, \ 1 \leq i \leq k,$$

implying that $Q_{v_{i-1}v_{i+1}}$ holds for either all even i or all odd i in $\{1,\ldots,k-1\}$, and consequently that $\mathbf{f}$ is constant on either $\{v_i : 0 \le i \le k \text{ even}\}$ or $\{v_i : 1 \le i \le k \text{ odd}\}$. This gives (3).

One would think that Theorem 1.2 would now follow immediately; for (3) implies

$$|\{v : |\mathbf{f}(v)| > 2\}| < e^{-\Omega(n)}2^n \quad \text{a.s.},$$

so that failure of the theorem would require an extremely unnatural distribution of values for a typical f. But, perhaps stupidly, we were not able to see a very direct way to complete a proof, even of Conjecture 1.1. Nonetheless, we do finish fairly easily using the following observation, valid for general bipartite graphs.

Let $H = (V, E)$ be an arbitrary (connected, finite) bipartite graph with $v_0 \in V$, and let $\mathbf{f}$ be chosen uniformly at random from $\mathcal{F}(H, v_0)$. Fix $v \in V$ and, for $i \in \mathbf{Z}$, set

$$a_i(v) = \begin{cases} \Pr(\mathbf{f}(v) = 2i) & \text{if } v \text{ is even,} \\ \Pr(\mathbf{f}(v) = 2i+1) & \text{if } v \text{ is odd.} \end{cases}$$

PROPOSITION 2.1: *For each $v \in V$ the sequence $\{a_i(v)\}$ is log-concave.*

This is proved in Section 5.

To complete the proof of Theorem 1.2, consider any even v and set $a_i(v) = a_i$. Proposition 2.1 implies that a_i is maximized at $i = 0$ (this is also given by Proposition 5.1—taken from [2]—which we will use in the proof of Proposition 2.1), so by (3) we have $a_0 > \Omega(1)$. On the other hand, (3) gives $a_2, a_{-2} < e^{-\Omega(n)}$. But then for $|i| > 2$ we have, using log-concavity,

$$a_i^2 \le a_2^{|i|}/a_0^{|i|-2},$$

implying $a_i < C^{-|i|n}$ for some constant $C > 1$. This gives Theorem 1.2. ∎

3. Entropy

Here we briefly review relevant entropy background. For more thorough discussions see [9], [6].

In what follows $\mathbf{X}$, $\mathbf{Y}$ etc. are discrete random variables (r.v.'s), which in our usage are allowed to take values in any countable (here always finite) set. Throughout the paper we take $\log = \log_2$.

As usual, H is the (binary) entropy function, $H(\alpha) = \alpha\log(1/\alpha) + (1-\alpha)\log(1/(1-\alpha))$. The *entropy* of the r.v. $\mathbf{X}$ is

$$H(\mathbf{X}) = \sum_x p(x)\log\frac{1}{p(x)},$$

where we write $p(x)$ for $\Pr(\mathbf{X} = x)$ (and extend this convention in natural ways below). The *conditional entropy* of $\mathbf{X}$ given $\mathbf{Y}$ is

$$H(\mathbf{X}|\mathbf{Y}) = \mathbf{E}H(\mathbf{X}|\mathbf{Y}=y) = \sum_y p(y) \sum_x p(x|y) \log \frac{1}{p(x|y)}.$$

Notice that we are also writing $H(\mathbf{X}|Q)$ with Q an *event* (in this case $Q = \{\mathbf{Y} = y\}$):

$$H(\mathbf{X}|Q) = \sum p(x|Q) \log \frac{1}{p(x|Q)}.$$

In what follows we will often have these two types of conditioning simultaneously, and will use ";" to separate them, writing, for instance, $H(\mathbf{X}|Q, R; \mathbf{Y}, \mathbf{Z})$.

For a random vector $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_n)$ (note this is also a r.v.), we have

$$H(\mathbf{X}) = H(\mathbf{X}_1) + H(\mathbf{X}_2|\mathbf{X}_1) + \cdots + H(\mathbf{X}_n|\mathbf{X}_1, \ldots, \mathbf{X}_{n-1}). \tag{4}$$

We will make repeated, and usually unremarked, use of the inequalities

$$\begin{aligned} H(\mathbf{X}) &\le \log|\mathrm{range}(\mathbf{X})|, \\ H(\mathbf{X}|\mathbf{Y}) &\le H(\mathbf{X}), \end{aligned} \tag{5}$$

and, more generally,

$$\text{if } \mathbf{Y} \text{ determines } \mathbf{Z} \text{ then } H(\mathbf{X}|\mathbf{Y}) \le H(\mathbf{X}|\mathbf{Z}). \tag{6}$$

(It may be worth noting that while (6) implies—is in fact equivalent to—$H(\mathbf{X}|\mathbf{Y}, \mathbf{Z}) \le H(\mathbf{X}|\mathbf{Y})$, the formally similar $H(\mathbf{X}|Q, R) \le H(\mathbf{X}|Q)$ is nonsense.)

We will also need the following fact which slightly generalizes a lemma of J. Shearer (see [5, p. 33]). For random vector $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_m)$ and $A \subseteq [m]$, set $\mathbf{X}_A = (\mathbf{X}_i \colon i \in A)$.

LEMMA 3.1: *Let $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_m)$ be a random vector and $\mathcal{A}$ a collection of subsets (possibly with repeats) of $[m]$, with each element of $[m]$ contained in at least t members of $\mathcal{A}$. Then for any partial order $\prec$ on $[m]$,*

$$H(\mathbf{X}) \le \frac{1}{t} \sum_{A \in \mathcal{A}} H(\mathbf{X}_A | (X_i \colon i \prec A)).$$

(Of course $i \prec A$ means $i \prec a\ \forall a \in A$.) When the partial order is vacuous, Lemma 3.1 becomes Shearer's Lemma, and the further specialization $\mathcal{A} = \{\{1\}, \ldots, \{m\}\}$ gives the basic inequality

$$H(\mathbf{X}_1, \ldots, \mathbf{X}_n) \le \sum H(\mathbf{X}_i).$$

Proof of Lemma 3.1: . (This repeats a proof of Shearer's Lemma due, as far as I know ([10]), to Jaikumar Radhakrishnan.)

By (6) the statement is strongest when $\prec$ is a total order, w.l.o.g. just the usual order $<$. In this case for $A = \{j_1 < \cdots < j_k\} \in \mathcal{A}$ and $\mathbf{Y} = (\mathbf{X}_i : i < A)$, we have

$$\begin{aligned} H(\mathbf{X}_A|\mathbf{Y}) =& H(\mathbf{X}_{j_1}|\mathbf{Y}) + H(\mathbf{X}_{j_2}|\mathbf{X}_{j_1}, \mathbf{Y}) + \cdots + H(\mathbf{X}_{j_k}|\mathbf{X}_{j_1}, \ldots, \mathbf{X}_{j_{k-1}}, \mathbf{Y}) \\ \geq& H(\mathbf{X}_{j_1}|(\mathbf{X}_i : i < j_1)) + \cdots + H(\mathbf{X}_{j_k}|(\mathbf{X}_i : i < j_k)) \end{aligned}$$

(using (4) in the first line and (6) in the second).

Thus

$$\begin{aligned} \sum_{A\in\mathcal{A}} H(\mathbf{X}_A|(\mathbf{X}_i : i < A)) \geq& \sum_{A\in\mathcal{A}} \sum_{j\in A} H(\mathbf{X}_j|(\mathbf{X}_i : i < j)) \\ \geq& t\sum_j H(\mathbf{X}_j|(\mathbf{X}_i : i < j)) \\ =& tH(\mathbf{X}) \end{aligned}$$

(by our assumption on $\mathcal{A}$ and again using (4)). ∎

4. Main inequality

In this section we prove (2). We again take $G = (V, E)$ to be the n-dimensional cube, with $v_0 = \underline{0}$, and write $|v|$ for the size of v, regarded as a subset of $[n]$ in the usual way.

Fix some $i \in [n]$, and for $x \in V$ define $x' \in V$ by $x'_j = x_j$ iff $j \neq i$. Set

$$\begin{aligned} M_x &= N_x \setminus \{x'\}, \\ R_x &= \{\mathbf{f} \text{ is constant on } M_x\} \ (\supseteq Q_x), \\ R_{\overline{x}} &= \overline{R}_x, \ R_{xy} = R_x R_y \quad \text{etc.}, \\ r_x &= \Pr(R_x), \ \ r_{xy} = \Pr(R_{xy}) \quad \text{etc.}, \end{aligned}$$

and similarly $q_x = \Pr(Q_x)$ etc.

In [8] and [7] we worked directly with events analogous to the present Q_x's. A key to the present work was the discovery that—for reasons we still find a bit mysterious—one can do considerably better by working with pairs of events R_x, $R_{x'}$.

Throughout the following discussion we let v range over the set

(7) $$\{x \colon x_i = 0, \ x \text{ even}\}$$

and always take $z = v'$.

We will focus on the quantity

$$\varepsilon = r_{vz} + r_{\overline{vz}},$$

where v, z are an arbitrary pair as above (of course all such pairs give the same ε). It is easy to see that, although Q_x is a proper subset of R_x, we have

$$Q_{\overline{vz}} = R_{\overline{vz}},$$

and consequently, using symmetry, $\varepsilon \geq q_{xy} + q_{\overline{xy}}$ for any $x \sim y$ in V. So a slight strengthening of (2) is

$$\varepsilon < e^{-\Omega(n)}. \tag{8}$$

This is the statement we prove.

To begin we note a trivial lower bound on $H(\mathbf{f}) = \log|\mathcal{F}|$:

$$H(\mathbf{f}) \geq 2^{n-1}. \tag{9}$$

We will prove (8) by producing a nearly matching upper bound which actually falls *below* the lower bound unless ε is very small.

From this point we specialize to $i = n$. For $0 \leq k \leq n-1$ set

$$L_k = \{x \in \{0,1\}^n \colon \sum_{i=1}^{n-1} x_i = k\},$$

and define the partial order "$\prec$" on $\{0,1\}^n$ by taking $L_1 \prec L_0 \prec L_3 \prec L_2 \prec L_5 \prec \cdots$ (i.e., $L_{2i+1} \prec L_{2i} \prec L_{2i+3}$).

Recall that v ranges over the set (7) and $z = v'$, and set $\mathbf{X}_v = \mathbf{f}|_{M_v \cup M_z}$.

Our point of departure, an instance of Lemma 3.1, is

$$H(\mathbf{f}) \leq \sum_v H(\mathbf{f}(v), \mathbf{f}(z) | (\mathbf{f}(x) \colon x \prec v)) + \frac{1}{n-1} \sum_v H(\mathbf{X}_v | (\mathbf{f}(x) \colon x \prec M_v \cup M_z)). \tag{10}$$

(To put this in the framework of Lemma 3.1, let $\mathcal{A}$ consist of the sets $M_v \cup M_z$ together with $n-1$ copies of $\{v, z\}$ for each v.)

For $|v| \geq 4$ we associate with v two additional vertices: $w = w(v)$ some vertex satisfying $w < v$ and $|w| = |v| - 4$ ("$<$" referring to the lattice of subsets of $[n]$), and $t = w'$. We will use w, t to represent the information we need from

$(\mathbf{f}(x)\colon x \prec M_v \cup M_z)$ when we come to investigate the entropy of $\mathbf{X}_v$. (Note $w, t \prec M_v \cup M_z$.)

For terms in the first sum in (10) we have the easy bound

$$\begin{aligned} H(\mathbf{f}(v),\mathbf{f}(z)|(\mathbf{f}(x):x \prec v)) \leq & H(\mathbf{f}(v),\mathbf{f}(z)|(\mathbf{f}(x):x \in M_v \cup M_z)) \\ \leq & r_{vz}\log 3 + r_{v\bar{z}} + r_{\bar{v}z}. \end{aligned} \tag{11}$$

(Note that on $R_{\bar{v}\bar{z}}$, $\mathbf{f}|_{M_v\cup M_z}$ determines $\mathbf{f}(v)$ and $\mathbf{f}(z)$, whence the absence of an $r_{\bar{v}\bar{z}}$ term in (11). For the terms that do appear we have, for instance,

$$H(\mathbf{f}(v),\mathbf{f}(z)|R_{v\bar{z}};(\mathbf{f}(x)\colon x \in M_v \cup M_z)) \leq 1,$$

since on $R_{v\bar{z}}$, $\mathbf{f}|_{M_v\cup M_z}$ determines $\mathbf{f}(z)$ and leaves just two possibilities for $f(v)$. Substitution of $\log 3$ for the naive upper bound 2, here and later in the corresponding bound (21), may be regarded as driving our analysis.)

The second sum in (10) requires more care. Set $\mathbf{T}_v = (\mathbf{1}_{R_v}, \mathbf{1}_{R_z})$ and assume from now on that $|v| \geq 4$. (We will eventually give away a negligible $O(n^2)$ bit to allow for smaller v's.) We will find it convenient to set $\varphi(x) = x\log(x^{-1}+1)$.

To begin, we have

$$\begin{aligned} H(\mathbf{X}_v|(\mathbf{f}(x):x \prec M_v \cup M_z)) \leq & H(\mathbf{X}_v|\mathbf{1}_{Q_w},\mathbf{f}(w),\mathbf{f}(t)) \\ \leq & H(\mathbf{T}_v|\mathbf{1}_{Q_w}) + H(\mathbf{X}_v|\mathbf{1}_{Q_w},\mathbf{f}(w),\mathbf{f}(t),\mathbf{T}_v). \end{aligned} \tag{12}$$

The first term here will turn out to be small because $\mathbf{1}_{Q_w}$ nearly determines $\mathbf{T}_v$; precisely we have

$$\Pr(R_{v\bar{z}}|Q_w) > 1 - O(\varepsilon), \tag{13}$$

$$\Pr(R_{\bar{v}z}|Q_{\bar{w}}) > 1 - O(\varepsilon). \tag{14}$$

Proof: We just prove (13), (14) being similar. We have

$$\overline{R}_{v\bar{z}}Q_w \subseteq \overline{Q}_{v\bar{z}}Q_w \cup Q_{v\bar{z}}\overline{R}_{v\bar{z}}.$$

But with $w \lessdot u \lessdot x \lessdot y \lessdot v \lessdot z$ we have

$$\overline{Q}_{v\bar{z}}Q_w \subseteq Q_{wu} \cup Q_{\bar{u}\bar{x}} \cup Q_{xy} \cup Q_{\bar{y}\bar{v}} \cup Q_{vz}. \tag{15}$$

Moreover

$$Q_{v\bar{z}}\overline{R}_{v\bar{z}} \subseteq R_v\overline{R}_{v\bar{z}} = R_{vz}. \tag{16}$$

This gives (13), since each of the events on the right hand sides of (15) and (16) has probability less than ε (and since $q_w \geq (1-\varepsilon)/2$). ∎

Notice for future reference that this argument also shows

$$\Pr(Q_{\overline{u}x\overline{y}v\overline{z}}|Q_w), \Pr(Q_{u\overline{x}y\overline{v}z}|Q_{\overline{w}}) > 1 - O(\varepsilon). \tag{17}$$

From (13) and (14) it follows that

$$H(\mathbf{T}_v|\mathbf{1}_{Q_w}) \le H(O(\varepsilon), O(\varepsilon), O(\varepsilon), 1 - O(\varepsilon)) = O(\varphi(\varepsilon)). \tag{18}$$

For the second term on the right hand side of (12) we have

$$\begin{aligned} H(\mathbf{X}_v|\mathbf{1}_{Q_w}, \mathbf{f}(w), \mathbf{f}(t), \mathbf{T}_v) \le & \Pr(Q_{\overline{w}})H(\mathbf{X}_v|Q_{\overline{w}}; \mathbf{f}(w), \mathbf{T}_v) \\ & + \Pr(Q_w)H(\mathbf{X}_v|Q_w; \mathbf{f}(t), \mathbf{T}_v). \end{aligned} \tag{19}$$

To bound the first of the entropy terms on the right hand side, we will make use of the fact that conditioning on $Q_{\overline{w}}$ and $\mathbf{f}(w)$ leaves little information in $\mathbf{f}(v)$. We have

$$H(\mathbf{X}_v|Q_{\overline{w}}; \mathbf{f}(w), \mathbf{T}_v) \le H(\mathbf{f}(v)|Q_{\overline{w}}; \mathbf{f}(w)) + H(\mathbf{X}_v|Q_{\overline{w}}; \mathbf{T}_v, \mathbf{f}(v)). \tag{20}$$

From (17) it follows that $\Pr(\mathbf{f}(v) \neq \mathbf{f}(w)|Q_{\overline{w}}) = O(\varepsilon)$ (since $Q_{uy} \Rightarrow \mathbf{f}(v) = \mathbf{f}(w)$). So, since the number of possibilities for $\mathbf{f}(v)$ given $\mathbf{f}(w)$ is $O(1)$ in any case, we have

$$H(\mathbf{f}(v)|Q_{\overline{w}}; \mathbf{f}(w)) \le O(\varphi(\varepsilon)).$$

For the main term in (20) we have

$$\begin{aligned} H(\mathbf{X}_v|Q_{\overline{w}}; \mathbf{T}_v, \mathbf{f}(v)) = & \Pr(R_{vz}|Q_{\overline{w}})H(\mathbf{X}_v|Q_{\overline{w}}, R_{vz}; \mathbf{f}(v)) \\ & + \Pr(R_{v\overline{z}}|Q_{\overline{w}})H(\mathbf{X}_v|Q_{\overline{w}}, R_{v\overline{z}}; \mathbf{f}(v)) \\ & + \Pr(R_{\overline{v}z}|Q_{\overline{w}})H(\mathbf{X}_v|Q_{\overline{w}}, R_{\overline{v}z}; \mathbf{f}(v)) \\ & + \Pr(R_{\overline{v}\overline{z}}|Q_{\overline{w}})H(\mathbf{X}_v|Q_{\overline{w}}, R_{\overline{v}\overline{z}}; \mathbf{f}(v)) \end{aligned}$$

Here we may bound the entropies by

$$H(\mathbf{X}_v|Q_{\overline{w}}, R_{vz}; \mathbf{f}(v)) \le 2$$

(on R_{vz} specification of $\mathbf{f}(v)$ leaves just 4 possibilities for the restriction of $\mathbf{f}$ to $M_v \cup M_z$, though "$O(1)$" in place of "2" would also suffice here);

$$H(\mathbf{X}_v|Q_{\overline{w}}, R_{v\overline{z}}; \mathbf{f}(v)) \le n$$

(one bit for $\mathbf{f}(z)$, then $n-1$ bits for $\mathbf{f}|_{M_z}$, which also determines the value of $\mathbf{f}$ on M_v);

$$H(\mathbf{X}_v|Q_{\overline{w}}, R_{\overline{v}z}; \mathbf{f}(v)) \le n - 1;$$

and

$$H(\mathbf{X}_v|Q_{\overline{w}}, R_{\overline{v}\overline{z}}; \mathbf{f}(v)) \leq 1 + (n-1)\log 3$$

(allowing one bit for $\mathbf{f}(z)$ and then $\log 3$ bits for $(\mathbf{f}(x), \mathbf{f}(x'))$ for each $x \in M_v$).

Thus

$$\begin{aligned} H(\mathbf{X}_v|Q_{\overline{w}}; \mathbf{T}_v, \mathbf{f}(v)) \leq & \Pr(R_{vz}|Q_{\overline{w}}) \cdot 2 + \Pr(R_{v\overline{z}}|Q_{\overline{w}})n + \Pr(R_{\overline{v}z}|Q_{\overline{w}})(n-1) \\ & + \Pr(R_{\overline{v}\overline{z}}|Q_{\overline{w}})(1 + (n-1)\log 3). \end{aligned}$$

A similar analysis bounds the second entropy term on the right hand side of (19):

$$\begin{aligned} H(\mathbf{X}_v|Q_w; \mathbf{f}(t), \mathbf{T}_v) \leq & H(\mathbf{f}(z)|Q_w; \mathbf{f}(t)) + H(\mathbf{X}_v|Q_w; \mathbf{T}_v, \mathbf{f}(z)) \\ = & O(\varphi(\varepsilon)) + H(\mathbf{X}_v|Q_w; \mathbf{T}_v, \mathbf{f}(z)), \end{aligned}$$

and

$$\begin{aligned} H(\mathbf{X}_v|Q_w; \mathbf{T}_v, \mathbf{f}(z)) = & \Pr(R_{vz}|Q_w)H(\mathbf{X}_v|Q_w, R_{vz}; \mathbf{f}(z)) \\ & + \Pr(R_{v\overline{z}}|Q_w)H(\mathbf{X}_v|Q_w, R_{v\overline{z}}; \mathbf{f}(z)) \\ & + \Pr(R_{\overline{v}z}|Q_w)H(\mathbf{X}_v|Q_w, R_{\overline{v}z}; \mathbf{f}(z)) \\ & + \Pr(R_{\overline{v}\overline{z}}|Q_w)H(\mathbf{X}_v|Q_w, R_{\overline{v}\overline{z}}; \mathbf{f}(z)) \\ \leq & \Pr(R_{vz}|Q_w) \cdot 2 + \Pr(R_{v\overline{z}}|Q_w)(n-1) \\ & + \Pr(R_{\overline{v}z}|Q_w)n + \Pr(R_{\overline{v}\overline{z}}|Q_w)(1 + (n-1)\log 3). \end{aligned}$$

Combining the preceding bounds we find that the right hand side of (19) is at most

$$\begin{aligned} & O(\varphi(\varepsilon)) + 2r_{vz} + r_{\overline{v}\overline{z}}(1 + (n-1)\log 3) \\ (21) \qquad & \qquad + \Pr(Q_{\overline{w}}R_{v\overline{z}}) + \Pr(Q_w R_{\overline{v}z}) + (r_{v\overline{z}} + r_{\overline{v}z})(n-1) \\ & = (n-1)(r_{v\overline{z}} + r_{\overline{v}z}) + 2r_{vz} + r_{\overline{v}\overline{z}}(1 + (n-1)\log 3) + O(\varphi(\varepsilon)) \end{aligned}$$

(again using (13) and (14)).

Finally, reviewing (10), (11), (12), (18), and the preceding bound for (19), and allowing $O(n^2)$ bits for the terms in (10) corresponding to v's with $|v| \leq 2$, we have

$$\begin{aligned} H(\mathbf{f}) \leq & O(n^2) + \sum_v \{r_{vz}\log 3 + r_{v\overline{z}} + r_{\overline{v}z} + \frac{1}{n-1}[(n-1)(r_{v\overline{z}} + r_{\overline{v}z}) \\ & + 2r_{vz} + r_{\overline{v}\overline{z}}(1 + (n-1)\log 3) + O(\varphi(\varepsilon))]\} \\ = & O(n^2) + \sum_v \{2(r_{v\overline{z}} + r_{\overline{v}z}) + (r_{vz} + r_{\overline{v}\overline{z}})(\log 3 + O(1/n)) + O(\varphi(\varepsilon)/n)\}. \end{aligned}$$

Combining this with (9) and recalling that $r_{vz}, \ldots, r_{\overline{vz}}$ do not depend on the choice of v, we have, for any particular v,

$$\begin{aligned} 2 - O(n^2 2^{-n}) \leq & 2(r_{v\bar{z}} + r_{\bar{v}z}) + (r_{vz} + r_{\overline{vz}})(\log 3 + O(1/n)) + O(\varphi(\varepsilon)/n) \\ = & 2(1-\varepsilon) + \varepsilon(\log 3 + O(1/n)) + O(\varphi(\varepsilon)/n), \end{aligned}$$

or

$$(2 - \log 3 - O(1/n))\varepsilon \leq O(\varphi(\varepsilon)/n) + O(n^2 2^{-n}).$$

This gives (8). ∎

5. Log-concavity

Here we only assume that $G = (V, E)$ is connected and bipartite with $v_0 \in V$, and again take $\mathbf{f}$ to be uniform from $\mathcal{F} = \mathcal{F}(G, v_0)$. We need the following unimodality statement, which is Proposition 2.1 of [2].

Proposition 5.1: *For every $v \in V$ and $0 \leq |s| < |t|$ with $s, t \in \mathbf{Z}$ and $s \equiv t \pmod 2$,*

$$\Pr(\mathbf{f}(v) = t) \leq \Pr(\mathbf{f}(v) = s).$$

Write $\Gamma(G, v_0)$ for the set of Lipschitz functions on G that vanish at v_0; that is,

$$\Gamma(G, v_0) = \{\gamma \in \mathbf{Z}^V \colon \gamma(v_0) = 0;\ x \sim y \Rightarrow |\gamma(x) - \gamma(y)| \leq 1\}.$$

For $v \in V$ set

$$\mathcal{F}_i(G, v_0, v) = \{f \in \mathcal{F} \colon f(v) = i\}, \quad \Gamma_i(G, v_0, v) = \{\gamma \in \Gamma(G, v_0) \colon \gamma(v) = i\}.$$

For the rest of this discussion we fix $v \in V$ and abbreviate $\Gamma(G, v_0) = \Gamma$, $\mathcal{F}_i(G, v_0, v) = \mathcal{F}_i$ and $\Gamma_i(G, v_0, v) = \Gamma_i$.

For $f, g \in \mathcal{F}$ let $\gamma_{fg} = (f+g)/2$. Then γ_{fg} clearly belongs to Γ, and, more precisely,

$$f \in \mathcal{F}_i,\ g \in \mathcal{F}_j \Rightarrow \gamma_{fg} \in \Gamma_{(i+j)/2}.$$

(In fact one easily sees that each $\gamma \in \Gamma$ is of this form; for instance, we may define f, g by setting $f(v) = g(v) = \gamma(v)$ if $\gamma(v)$ has the same parity as v, and $f(v) - 1 = \gamma(v) = g(v) + 1$ otherwise.)

For $\gamma \in \Gamma_i$ and $j \in \mathbf{Z}$ set

$$A_j(\gamma) = \{(f, g) \in \mathcal{F}_{i+j} \times \mathcal{F}_{i-j} : \gamma_{fg} = \gamma\}.$$

We prove a somewhat stronger version of Proposition 2.1:

CLAIM 5.2: *For all i and $\gamma \in \Gamma_i$, $|A_2(\gamma)| \le |A_0(\gamma)|$.*

We will prove this by reducing it to an instance of Proposition 5.1. The reduction is given by several assertions whose straightforward verifications are mainly left to the reader. We assume throughout that $\gamma \in \Gamma$ and $f, g \in \mathcal{F}$. Define $f \sim \gamma$ (and $\gamma \sim f$) to mean

$$x \sim y, \gamma(x) \neq \gamma(y) \Rightarrow f(x) - f(y) = \gamma(x) - \gamma(y).$$

For $\gamma \in \Gamma_i$ set

$$B_j(\gamma) = \{f \in \mathcal{F}_{i+j}: f \sim \gamma\}.$$

One easily checks that

$$\gamma_{fg} = \gamma \Rightarrow f, g \sim \gamma$$

and

$$\text{if } \gamma \in \Gamma_i,\ f \in \mathcal{F}_j \text{ and } f \sim \gamma, \text{ then } \gamma \sim 2\gamma - f \in \mathcal{F}_{2i-j}.$$

Thus $f \mapsto (f, 2\gamma - f)$ gives a bijection between $B_j(\gamma)$ and $A_j(\gamma)$ (for each j), and Claim 5.2 is equivalent to

$$|B_2(\gamma)| \le |B_0(\gamma)|. \tag{22}$$

Now let $E(\gamma) = \{\{x, y\} \in E : \gamma(x) \neq \gamma(y)\}$ and $G/\gamma = G/E(\gamma)$ (the graph obtained from G by contracting the edges of $E(\gamma)$; see, for instance, [4]). Then, crucially, we have

$$G/\gamma \text{ is bipartite.} \tag{23}$$

(To see this, note that any cycle of G/γ is the contraction of some cycle $C = (x_0, x_1, \ldots, x_{2k} = x_0)$ of G, and that the number of parity changes in the sequence $(\gamma(x_0), \ldots, \gamma(x_{2k}))$ is $|E(C) \cap E(\gamma)|$.)

The vertices of G/γ are naturally identified with the components of $(V, E(\gamma))$ and we write x/γ for the vertex of G/γ containing $x \in V$.

For $\gamma \in \Gamma_i$ and $f \in B_j(\gamma)$, $f' := f - \gamma$ is easily seen to lie in Γ_j and to satisfy

$$x \sim_G y \Rightarrow [f'(x) = f'(y) \Leftrightarrow \gamma(x) \neq \gamma(y)].$$

In particular, f' is constant on each component of $E(\gamma)$, so maps naturally to a function $f'' \in \mathcal{F}_j(G/\gamma, v_0/\gamma, v/\gamma)$.

This construction is reversible: given such an f'', define $f' \in \Gamma_j$ by $f'(x) = f''(x/\gamma)$ and take $f = f' + \gamma \in B_j(\gamma)$.

So we have $|B_j(\gamma)| = |\mathcal{F}_j(G/\gamma, v_0/\gamma, v/\gamma)|$, and for (22) can substitute

$$|\mathcal{F}_2(G/\gamma, v_0/\gamma, v/\gamma)| \leq |\mathcal{F}_0(G/\gamma, v_0/\gamma, v/\gamma)|.$$

But in view of (23), this is contained in Proposition 5.1. ∎

ACKNOWLEDGEMENT: I would like to thank Itai Benjamini and Elchanan Mossel for telling me about these questions, and Elchanan also for helpful comments on the manuscript.

References

[1] C. A. Athanasiadis, *Algebraic combinatorics of graph spectra, subspace arrangements and Tutte polynomials*, PhD thesis, Massachusetts Institute of Technology, 1996.

[2] I. Benjamini, O. Häggström and E. Mossel, *On random graph homomorphisms into* **Z**, Journal of Combinatorial Theory. Series B **78** (2000), 86–114.

[3] I. Benjamini and G. Schechtman, *Upper bounds on the height difference of the Gaussian random field and the range of random graph homomorphisms into* **Z**, Random Structures & Algorithms **17** (2000), 20–25.

[4] B. Bollobás, *Modern Graph Theory*, Springer, New York, 1998.

[5] F. R. K. Chung, P. Frankl, R. Graham and J. B. Shearer, *Some intersection theorems for ordered sets and graphs*, Journal of Combinatorial Theory. Series A **48** (1986), 23–37.

[6] I. Csiszár and J. Körner, *Information Theory. Coding Theorems for Discrete Memoryless Systems*, Akadémiai Kiadó, Budapest, 1981.

[7] J. Kahn, *An entropy approach to the hard-core model on bipartite graphs*, submitted for publication.

[8] J. Kahn and A. Lawrenz, *Generalized rank functions and an entropy argument*, Journal of Combinatorial Theory. Series A **87** (1999), 398–403.

[9] R. J. McEliece, *The Theory of Information and Coding*, Addison-Wesley, London, 1977.

[10] A. Wigderson, *personal communication.*