Using discrepancy to control singular values for nonnegative matrices

Steve Butler*

Abstract

We will consider two parameters which can be associated with a nonnegative matrix: the second largest singular value of the "normalized" matrix, and the discrepancy of the entries (which is a measurement between the sum of the actual entries in blocks versus the expected sum). Our main result is to show that these are related in that discrepancy can be bounded by the second largest singular value and vice versa. These matrix results are then used to derive some (edge/alternating walks) discrepancy properties of edge-weighted directed graphs.

Key words: singular values; discrepancy; directed graphs; alternating walks

AMS Mathematics Subject Classification: 15A18; 05C50

1 Introduction

Bollobás and Nikiforov [2] have shown that there is a constant C so that for any Hermitian matrix $A = (a_{ij})_{n\times n}$, $\sigma_2(A) \le C\,\mathrm{disc}(A)\log n$ where $\sigma_2(A)$ is the second singular value of A and $\mathrm{disc}(A)$ is the minimal α so that for all $S, T \subseteq [n]$

$$\left|\Big(\sum_{i\in S}\sum_{j\in T} a_{ij}\Big) - \rho|S|\,|T|\right| \le \alpha\sqrt{|S||T|} \quad \text{where} \quad \rho = \frac{1}{n^2}\sum_{i,j=1}^{n} a_{ij}. \qquad (1)$$

Their approach was to approximate the vector associated with $\sigma_2(A)$ as a linear combination of at most $C'\log n$ 0-1 vectors for a constant C' depending only on how close the approximation needs to be. They then used this approximation to collapse the matrix A and get the desired result.

Seperately, Bilu and Linial [1] showed (among other things) that for the special case when A is the adjacency matrix of a d-regular (undirected) graph that $\sigma_2(A) \le O\big(\alpha\big(1 + \log(d/\alpha)\big)\big)$. Their approach also involved an approximation of the vector, but this time the entries of the approximation were powers of 2,

*University of California, San Diego, Dept. of Mathematics, La Jolla, CA 92093-0112. email: sbutler@math.ucsd.edu

and instead of collapsing the matrix they used some clever manipulation of the sums.

We will combine the approximation ideas in [2] and the manipulation of the resulting sums in [1] to obtain a discrepancy result which we state below for nonnegative (not necessarily square) matrices. The proofs of these results will be given in Section 2 and an application to directed graphs will be given in Section 3.

In this paper we will use $\langle x, y\rangle$ to denote the usual inner product of vectors x and y, J to denote the matrix of all 1s, and ψ_S the characteristic vector of a subset S, i.e.,

$$\psi_S(x) = \begin{cases} 1 & \text{if } x \in S; \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 1. *Let $B \in M_{m\times n}$ be a matrix with nonnegative matrices and no zero rows/columns. Also, let $R \in M_{m\times m}$ and $C \in M_{n\times n}$ be the unique diagonal matrices such that $B\mathbf{1} = R\mathbf{1}$ and $\mathbf{1}B = \mathbf{1}C$ (here $\mathbf{1}$ denotes the all 1's vector of the appropriate size). Then for all $S \subseteq [m]$ and $T \subseteq [n]$*

$$\left|\langle \psi_S, B\psi_T\rangle - \frac{\langle \psi_S, B\mathbf{1}\rangle \langle \mathbf{1}, B\psi_T\rangle}{\langle \mathbf{1}, B\mathbf{1}\rangle}\right| \le \sigma_2(R^{-1/2}BC^{-1/2})\sqrt{\langle \psi_S, B\mathbf{1}\rangle \langle \mathbf{1}, B\psi_T\rangle}.$$

Note that the diagonal entries of R and C are the row sums and column sums (respectively) of B.

Theorem 2. *Let B, R, C be as above. If for all $S \subseteq [m]$ and $T \subseteq [n]$*

$$\left|\langle \psi_S, B\psi_T\rangle - \frac{\langle \psi_S, B\mathbf{1}\rangle \langle \mathbf{1}, B\psi_T\rangle}{\langle \mathbf{1}, B\mathbf{1}\rangle}\right| \le \alpha\sqrt{\langle \psi_S, B\mathbf{1}\rangle \langle \mathbf{1}, B\psi_T\rangle} \tag{2}$$

(we can and will assume that $\alpha \le 1$), then

$$\sigma_2(R^{-1/2}BC^{-1/2}) \le 150\alpha(1 - 8\log\alpha).$$

The minimal α satisfying equation (2) is a discrepancy of A which we denote by $\mathrm{Disc}(A)$. For nonnegative Hermitian matrices the difference between (1) and (2) can be viewed as how rows/columns are weighted. In (1) each row is given equal weight, and so the important measurement is the number of rows, while in (2) each row is weighted according to its row sum, and so the important measurement is the sum of the row sums (similarly for the columns). It is because of this different approach that we need to normalize the matrix A by multiplying on the left by $R^{-1/2}$ and on the right by $C^{-1/2}$.

2 Proofs of Theorem 1 and 2

Before beginning our proofs we note the following:

(i) $(R^{-1/2}BC^{-1/2})C^{1/2}\mathbf{1} = R^{1/2}\mathbf{1}$.

(ii) $\mathbf{1}R^{1/2}(R^{-1/2}BC^{-1/2}) = \mathbf{1}C^{1/2}$.

(iii) $\sigma_1(R^{-1/2}BC^{-1/2}) = 1$.

(iv) $\sigma_2(R^{-1/2}BC^{-1/2}) = \sigma_1\big(R^{-1/2}BC^{-1/2} - \frac{1}{\langle \mathbf{1}, B\mathbf{1}\rangle} R^{1/2}JC^{1/2}\big)$.

Equalities (i) and (ii) are an easy calculation, while (iii) follows by the Perron-Frobenius Theorem on $(R^{-1/2}BC^{-1/2})^*(R^{-1/2}BC^{-1/2})$ which has eigenvector $C^{1/2}\mathbf{1}$ associated with eigenvalue 1. For (iv) we subtract out the largest singular value which by (i)-(iii) has left and right vectors $\mathbf{1}R^{1/2}$ and $C^{1/2}\mathbf{1}$ respectively, and noting that $\|\mathbf{1}R^{1/2}\|^2 = \|C^{1/2}\mathbf{1}\|^2 = \langle \mathbf{1}, B\mathbf{1}\rangle$.

Proof of Theorem 1. This follows from $|\langle x, My\rangle| \le \sigma_1(M)\|x\|\,\|y\|$ (see [6]), i.e.,

$$\begin{aligned}
\big|\langle \psi_S, B\psi_T\rangle - \frac{\langle \psi_S, B\mathbf{1}\rangle\,\langle \mathbf{1}, B\psi_T\rangle}{\langle \mathbf{1}, B\mathbf{1}\rangle}\big| \;&=\; \big|\langle \psi_S, \big(B - \frac{RJC}{\langle \mathbf{1}, B\mathbf{1}\rangle}\big)\psi_T\rangle\big| \\
&= \big|\big\langle R^{1/2}\psi_S, \big(R^{-1/2}BC^{-1/2} - \frac{R^{1/2}JC^{1/2}}{\langle \mathbf{1}, B\mathbf{1}\rangle}\big)C^{1/2}\psi_T\big\rangle\big| \\
&\le \sigma_1\big(R^{-1/2}BC^{-1/2} - \frac{R^{1/2}JC^{1/2}}{\langle \mathbf{1}, B\mathbf{1}\rangle}\big)\|R^{1/2}\psi_S\|\,\|C^{1/2}\psi_T\|.
\end{aligned}$$

A calculation shows that $\|R^{1/2}\psi_S\|^2 = \langle \psi_S, B\mathbf{1}\rangle$ and $\|C^{1/2}\psi_T\|^2 = \langle \mathbf{1}, B\psi_T\rangle$, which with the above comments concludes the proof. □

For Theorem 2 we need the following approximation lemmas.

Lemma 3. *Let $x \in \mathbb{C}^n$ with $\|x\| = 1$, and D a diagonal matrix with positive entries, d_t, on the diagonal. Then there is a vector $y \in \mathbb{C}^n$ such that $\|Dy\| \le 1$, $\|x - Dy\| \le \frac{1}{3}$ and the nonzero entries of y are of the form $\big(\frac{4}{5}\big)^j e^{2\pi i k/29}$ for j, k integers with $0 \le k < 29$.*

Proof. Let $x = (x_t)_{1\le t\le n}$ then we define $y = (y_t)_{1\le t\le n}$ entrywise. If $x_t = 0$ then set $y_t = 0$. Otherwise for some $r > 0$ and $0 \le \theta < 2\pi$, we have $x_t = re^{i\theta}$. For the unique integer j so that $\big(4/5\big)^j < r/d_t \le \big(4/5\big)^{j-1}$, set $y_t = \big(\frac{4}{5}\big)^j e^{2\pi i\lfloor 29\theta/2\pi\rfloor/29}$. By construction we have

$$0 < |x_t| - |d_t y_t| \le \bigg(\big(\frac{4}{5}\big)^{j-1} - \big(\frac{4}{5}\big)^j\bigg)d_t = \big(\frac{5}{4} - 1\big)\big(\frac{4}{5}\big)^j d_t < \frac{1}{4}|x_t|,$$

while the argument between x_t and y_t is bounded above by $2\pi/29$.

By use of the law of cosines it follows that $|x_t - d_t y_t|^2 \le |x_t|^2/9$, which implies $\|x - Dy\|^2 = \sum_t |x_t - d_t y_t|^2 \le \frac{1}{9}\sum_t |x_t|^2 = \frac{1}{9}$. □

Lemma 4. *Let M be a matrix and x', y' vectors such that $\|x'\| = \|y'\| = 1$ and $\sigma_1(M) = |\langle x', My'\rangle|$. If x, y are vectors such that $\|x\|, \|y\| \le 1$ and $\|x' - x\|$, $\|y' - y\| \le \frac{1}{3}$, then $\sigma_1(M) \le \frac{9}{2}|\langle x, My\rangle|$.*

Proof. We again use $|\langle x, My\rangle| \le \sigma_1(M)\|x\|\,\|y\|$.

$$\begin{aligned}\sigma_1(M) &= |\langle x', My'\rangle| = |\langle x+(x'-x), M\big(y+(y'-y)\big)\rangle| \\ &\le |\langle x, My\rangle| + |\langle x, M(y'-y)\rangle| + |\langle (x'-x), My\rangle| + |\langle (x'-x), M(y'-y)\rangle| \\ &\le |\langle x, My\rangle| + \frac{1}{3}\sigma_1(M) + \frac{1}{3}\sigma_1(M) + \frac{1}{9}\sigma_1(M),\end{aligned}$$

rearranging then gives the result. □

Proof of Theorem 2. Let $\mathcal{B} = B - \frac{1}{\langle \mathbf{1}, B\mathbf{1}\rangle} RJC$, so that $\sigma_2(R^{-1/2}BC^{-1/2}) = \sigma_1(R^{-1/2}\mathcal{B}C^{-1/2})$. There exists vectors x', y' such that $\|x'\| = 1$ and $\|y'\| = 1$ where

$$\sigma_1(R^{-1/2}\mathcal{B}C^{-1/2}) = |\langle x', R^{-1/2}\mathcal{B}C^{-1/2}y'\rangle|.$$

Applying Lemma 3 twice, there exist (step) vectors x, y with $\|x\|, \|y\| \le 1$, and $\|x' - R^{1/2}x\|$, $\|y' - C^{1/2}y\| \le \frac{1}{3}$. It follows from Lemma 4 that

$$\sigma_1(R^{-1/2}\mathcal{B}C^{-1/2}) \le \frac{9}{2}|\langle R^{1/2}x, (R^{-1/2}\mathcal{B}C^{-1/2})C^{1/2}y\rangle| = \frac{9}{2}|\langle x, \mathcal{B}y\rangle|.$$

We now partition $[m]$ according to the vector x. Let $X^{(t)} = \{j : |x_j| = \big(\frac{4}{5}\big)^t\}$, and let $x = \sum_t \big(\frac{4}{5}\big)^t x^{(t)}$, where $x^{(t)}$ is the "signed" indicator function of $X^{(t)}$, i.e.,

$$x_j^{(t)} = \begin{cases} x_j/|x_j| & \text{if } |x_j| = \big(\frac{4}{5}\big)^t; \\ 0 & \text{otherwise.}\end{cases}$$

We similarly partition $[n]$ to get $y = \sum_s \big(\frac{4}{5}\big)^s y^{(s)}$. We now have

$$\sigma_2(R^{-1/2}BC^{-1/2}) \le \frac{9}{2}|\langle x, \mathcal{B}y\rangle| \le \frac{9}{2}\sum_t\sum_s \big(\frac{4}{5}\big)^{t+s}|\langle x^{(t)}, \mathcal{B}y^{(s)}\rangle|.$$

By assumption, we have for any 0-1 vectors w and z that

$$|\langle w, \mathcal{B}z\rangle| = \big|\langle w, Bz\rangle - \frac{\langle w, B\mathbf{1}\rangle\,\langle \mathbf{1}, Bz\rangle}{\langle \mathbf{1}, B\mathbf{1}\rangle}\big| \le \alpha\sqrt{\langle w, B\mathbf{1}\rangle\,\langle \mathbf{1}, Bz\rangle}.$$

More generally, if $w = \sum_{k=0}^{28} e^{2\pi ik/29} w^{\langle k\rangle}$ and $z = \sum_{\ell=0}^{28} e^{2\pi i\ell/29} z^{\langle \ell\rangle}$ where $w^{\langle k\rangle}, z^{\langle \ell\rangle}$ are 0-1 vectors and the $w^{\langle k\rangle}$ ($z^{\langle \ell\rangle}$) are mutually orthogonal, then by the triangle and Cauchy-Schwarz inequalities we have

$$\begin{aligned}|\langle w, \mathcal{B}z\rangle| &= |\langle \sum_{k=0}^{28} e^{2\pi ik/29} w^{\langle k\rangle}, \mathcal{B}\sum_{\ell=0}^{28} e^{2\pi i\ell/29} z^{\langle \ell\rangle}\rangle| \le \sum_{k=0}^{28}\sum_{\ell=0}^{28}|\langle w^{\langle k\rangle}, \mathcal{B}z^{\langle \ell\rangle}\rangle| \\ &\le \alpha\sum_{k=0}^{28}\sum_{\ell=0}^{28}\sqrt{\langle w^{\langle k\rangle}, B\mathbf{1}\rangle\,\langle \mathbf{1}, Bz^{\langle \ell\rangle}\rangle} \le 29\alpha\sqrt{\sum_{k=0}^{28}\sum_{\ell=0}^{28}\langle w^{\langle k\rangle}, B\mathbf{1}\rangle\,\langle \mathbf{1}, Bz^{\langle \ell\rangle}\rangle} \\ &= 29\alpha\sqrt{\langle \sum_{k=0}^{28} w^{\langle k\rangle}, B\mathbf{1}\rangle\,\langle \mathbf{1}, B\sum_{\ell=0}^{28} z^{\langle \ell\rangle}\rangle} = 29\alpha\sqrt{\langle |w|, B\mathbf{1}\rangle\,\langle \mathbf{1}, B|z|\rangle},\end{aligned}$$

where $|x|$ denotes the vector of the absolute value of the entries of x. Applying this to $x^{(t)}$ and $y^{(s)}$ we have

$$|\langle x^{(t)}, \mathcal{B}y^{(s)}\rangle| \leq 29\alpha\sqrt{\langle |x^{(t)}|, B\mathbf{1}\rangle\, \langle \mathbf{1}, B|y^{(s)}|\rangle}. \tag{3}$$

We also have that

$$\sum_s |\langle x^{(t)}, \mathcal{B}y^{(s)}\rangle| \leq 2\langle |x^{(t)}|, B\mathbf{1}\rangle \text{ and } \sum_t |\langle x^{(t)}, \mathcal{B}y^{(s)}\rangle| \leq 2\langle \mathbf{1}, B|y^{(s)}|\rangle. \tag{4}$$

To see this, by the triangle inequality we have $|\langle w, Mz\rangle| \leq \langle |w|, |M|\,|z|\rangle$, and so

$$\begin{aligned}\sum_s |\langle x^{(t)}, \mathcal{B}y^{(s)}\rangle| \ &\leq\ \langle |x^{(t)}|, |\mathcal{B}| \sum_s |y^{(s)}|\rangle \\ &\leq\ \langle |x^{(t)}|, \big(B + \frac{RJC}{\langle \mathbf{1}, B\mathbf{1}\rangle}\big)\mathbf{1}\rangle \ =\ 2\langle |x^{(t)}|, B\mathbf{1}\rangle.\end{aligned}$$

The other result is proved similarly.

We now let $\gamma = \log_{4/5}\alpha$ and consider

$$\begin{aligned}\sum_t\sum_s \big(\frac{4}{5}\big)^{t+s}|\langle x^{(t)}, \mathcal{B}y^{(s)}\rangle| \ \leq\ &\sum_{|s-t|\leq\gamma} \big(\frac{4}{5}\big)^{t+s}|\langle x^{(t)}, \mathcal{B}y^{(s)}\rangle| \\ &+ \sum_t \big(\frac{4}{5}\big)^{2t+\gamma}\sum_s |\langle x^{(t)}, \mathcal{B}y^{(s)}\rangle| + \sum_s \big(\frac{4}{5}\big)^{2s+\gamma}\sum_t |\langle x^{(t)}, \mathcal{B}y^{(s)}\rangle|. \end{aligned}\tag{5}$$

The inequality can be verified by comparing the coefficient of $|\langle x^{(t)}, \mathcal{B}y^{(s)}\rangle|$ on both sides. Clearly when $|s-t| \leq \gamma$ the result holds, when $t > s+\gamma$ then $s + t > 2s + \gamma$ so that $\big(\frac{4}{5}\big)^{s+t} < \big(\frac{4}{5}\big)^{2s+\gamma}$, similarly when $s > t + \gamma$ then $\big(\frac{4}{5}\big)^{s+t} < \big(\frac{4}{5}\big)^{2t+\gamma}$ and the inequality follows.

We now bound the three terms on the right side of (5). For the first term we have

$$\begin{aligned}&\sum_{|s-t|\leq\gamma} \big(\frac{4}{5}\big)^{s+t}|\langle x^{(t)}, \mathcal{B}y^{(s)}\rangle| \\ &\quad\leq\ \frac{29}{2}\alpha \sum_{|s-t|\leq\gamma} 2\sqrt{\big(\frac{4}{5}\big)^{2t}\langle |x^{(t)}|, B\mathbf{1}\rangle\, \big(\frac{4}{5}\big)^{2s}\langle \mathbf{1}, B|y^{(s)}|\rangle} \\ &\quad\leq\ \frac{29}{2}\alpha \sum_{|s-t|\leq\gamma} \Big(\big(\frac{4}{5}\big)^{2t}\langle |x^{(t)}|, B\mathbf{1}\rangle + \big(\frac{4}{5}\big)^{2s}\langle \mathbf{1}, B|y^{(s)}|\rangle\Big) \\ &\quad\leq\ \frac{29}{2}\alpha(2\gamma+1)\Big(\sum_t \big(\frac{4}{5}\big)^{2t}\langle |x^{(t)}|, B\mathbf{1}\rangle + \sum_s \big(\frac{4}{5}\big)^{2s}\langle \mathbf{1}, B|y^{(s)}|\rangle\Big) \\ &\quad\leq\ 29\alpha(2\gamma+1).\end{aligned}$$

The inequalities follow from (respectively) (3), the geometric-arithmetic mean inequality, the fact that any term can show up at *most* $2\gamma+1$ times, and

$$\begin{aligned}\sum_t \big(\frac{4}{5}\big)^{2t}\langle|x^{(t)}|, B\mathbf{1}\rangle &= \|R^{1/2}x\|^2 \ \le\ 1 \quad \text{and}\\ \sum_s \big(\frac{4}{5}\big)^{2s}\langle\mathbf{1}, B|y^{(s)}|\rangle &= \|C^{1/2}y\|^2 \ \le\ 1.\end{aligned}$$

For the second term we use (4) to get

$$\sum_t \big(\frac{4}{5}\big)^{2t+\gamma}\sum_s |\langle x^{(t)}, \mathcal{B}y^{(s)}\rangle| \ \le\ 2\big(\frac{4}{5}\big)^{\gamma}\sum_t \big(\frac{4}{5}\big)^{2t}\langle|x^{(t)}|, B\mathbf{1}\rangle \ \le\ 2\big(\frac{4}{5}\big)^{\gamma},$$

a similar statement holds for the third term.

Putting this together we have that

$$\sigma_2(R^{-1/2}BC^{-1/2}) \ \le\ \frac{9}{2}\big(29\alpha(2\gamma+1)+4\big(\frac{4}{5}\big)^{\gamma}\big) \le\ 150\alpha(1-8\log\alpha). \qquad \square$$

3 Discrepancy for directed graphs

In this section we consider directed graphs which have a weight function w which assigns $w_{uv} > 0$ to each edge $u\to v$. This weight function also is used to give the adjacency matrix $A = A(G)$ by $A_{uv} = w(u\to v)$ for all edges $u\to v$ and 0 otherwise. The in- and out-degrees are $d_{in} = \sum_v w(v\to u)$ and $d_{out} = \sum_v w(u\to v)$, respectively the column and row sums of A, and form the entries of the diagonal matrices D_{in} and D_{out}. While the in- and out-volume of subsets X of vertices are $\mathrm{Vol}_{in}(X) = \sum_{x\in X} d_{in}(x)$ and $\mathrm{Vol}_{out}(X) = \sum_{x\in X} d_{out}(x)$.

We also have a discrepancy for directed graphs, denoted $\mathrm{Disc}(G)$, which is the minimal α so that for any subsets X, Y of vertices

$$\Big|\big(\sum_{u\in X}\sum_{v\in Y} w(u\to v)\big) - \frac{\mathrm{Vol}_{out}(X)\,\mathrm{Vol}_{in}(Y)}{\mathrm{Vol}\,(G)}\Big| \le \alpha\sqrt{\mathrm{Vol}_{out}\,(X)\,\mathrm{Vol}_{in}(Y)}, \qquad (6)$$

where $\mathrm{Vol}\,(G) := \mathrm{Vol}_{in}\,(V) = \mathrm{Vol}_{out}(V)$. The discrepancy for a directed graph and of a matrix are related by $\mathrm{Disc}(G) = \mathrm{Disc}(A(G))$. Applying Theorems 1 and 2 we get the following result.

Theorem 5. *For G a weighted directed graph without sources or sinks,*

$$\mathrm{Disc}(G) \le \sigma_2(D_{out}^{-1/2}AD_{in}^{-1/2}) \le 150\,\mathrm{Disc}(G)\big(1-8\log\mathrm{Disc}(G)\big).$$

This shows that for a directed graph having a small second singular value gives control on discrepancy and vice versa.

Another type of discrepancy for graphs is based on $\mathrm{disc}(A(G))$, the difference between these two discrepancies can be viewed in how a set of vertices are weighted. While in $\mathrm{disc}(A(G))$ each vertex is given equal weight so that the measure is the number of vertices, in $\mathrm{Disc}(A(G))$ the vertices are weighted by their degree so that the measure is the sum of the degrees. This idea of normalizing the weights has been used with great success in spectral techniques by Chung [4].

3.1 Alternating walks

Chung and Graham [5] have generalized discrepancy for undirected graphs by considering the discrepancy of walks of length t (the case $t = 1$ gives the original form of discrepancy). There has been limited success in generalizing these results to directed graphs (see [3]). The difficulty seems to lie in that to count walks we look at a matrix such as $AA \cdots A$ (t terms) which works well with eigenvalues but not with singular values. However, if we consider alternating walks, a walk where at every step we reverse direction, which are counted by a matrix such as $AA^*AA^* \cdots$ (t terms), these do work well with singular values. Here we will consider a discrepancy for alternating walks.

For an alternating walk $P = x_0{\rightarrow}x_1{\leftarrow}x_2{\rightarrow}x_3{\leftarrow}x_4 \cdots x_t$ we associate a weight

$$w(P) = \begin{cases} \dfrac{w(x_0{\rightarrow}x_1)w(x_1{\leftarrow}x_2)\cdots w(x_{t-1}{\leftarrow}x_t)}{d_{in}(x_1)d_{out}(x_2)d_{in}(x_3)\cdots d_{in}(x_{t-1})} & t \text{ even;} \\ \dfrac{w(x_0{\rightarrow}x_1)w(x_1{\leftarrow}x_2)\cdots w(x_{t-1}{\rightarrow}x_t)}{d_{in}(x_1)d_{out}(x_2)d_{in}(x_3)\cdots d_{out}(x_{t-1})} & t \text{ odd.} \end{cases}$$

There is a slight difference between the case t odd and t even, which corresponds to the direction of the last edge.

Let $\mathcal{P}_t(x{\rightarrow}y)$ denote the set of all alternating walks of length t starting at x and ending at y. Then define $w_t(x{\rightarrow}y) = \sum_{P\in\mathcal{P}_t(x\rightarrow y)} w(P)$, equivalently, $w_t(x{\rightarrow}y)/\operatorname{Vol}(G)$ is the probability that a randomly generated alternating walk of length t starts at x and ends at y.

We now define the discrepancy of alternating t-walks, denoted $\mathrm{AltDisc}_t(G)$, to be the minimal β such that for all $X, Y \subseteq V$

$$\left|\sum_{x\in X}\sum_{y\in Y} w_t(x{\rightarrow}y) - \frac{\operatorname{Vol}_{out}(X)\ \operatorname{Vol}_{out}(Y)}{\operatorname{Vol}(G)}\right| \le \beta\sqrt{\operatorname{Vol}_{out}(X)\ \operatorname{Vol}_{out}(Y)} \qquad t \text{ even;}$$

$$\left|\sum_{x\in X}\sum_{y\in Y} w_t(x{\rightarrow}y) - \frac{\operatorname{Vol}_{out}(X)\ \operatorname{Vol}_{in}(Y)}{\operatorname{Vol}(G)}\right| \le \beta\sqrt{\operatorname{Vol}_{out}(X)\ \operatorname{Vol}_{in}(Y)} \qquad t \text{ odd.}$$

Theorem 6. *For G a weighted directed graph without sources or sinks,*

$$\mathrm{AltDisc}_t(G) \le \big(\sigma_2(D_{out}^{-1/2}AD_{in}^{-1/2})\big)^t \le 150\,\mathrm{AltDisc}_t(G)(1-8\log\mathrm{AltDisc}_t(G)).$$

Proof. We consider the case t odd (t even is handled similarly). Let

$$\begin{aligned} B &= AD_{in}^{-1}A^*D_{out}^{-1}AD_{in}^{-1}A^*D_{out}^{-1}\cdots D_{out}^{-1}A \\ &= D_{out}^{1/2}\underbrace{(D_{out}^{-1/2}AD_{in}^{-1/2})(D_{out}^{-1/2}AD_{in}^{-1/2})^*\cdots(D_{out}^{-1/2}AD_{in}^{-1/2})}_{t\text{ terms}}D_{in}^{1/2}. \end{aligned}$$

We have $D_{out} = R$, $D_{in} = C$, $\sum_{x\in X}\sum_{y\in Y} w_t(x\to y) = \langle \psi_X, B\psi_Y\rangle$, $\mathrm{Vol}_{out}(X) = \langle \psi_X, B\mathbf{1}\rangle$ and $\mathrm{Vol}_{in}(Y) = \langle \mathbf{1}, B\psi_Y\rangle$. From Theorems 1, 2 and the definition of discrepancy for alternating t-walks we have

$$\mathrm{AltDisc}_t(G) \;\le\; \sigma_2\big(D_{out}^{-1/2} B D_{in}^{-1/2}\big) \;\le\; 150\,\mathrm{AltDisc}_t(G)\big(1 - 8\log \mathrm{AltDisc}_t(G)\big).$$

It remains to show that $\sigma_2\big(D_{out}^{-1/2} B D_{in}^{-1/2}\big) = \big(\sigma_2(D_{out}^{-1/2} A D_{in}^{-1/2})\big)^t$. But this follows immediately from the definition of B and the fact that for a matrix F, $\sigma_2(\underbrace{FF^*FF^*F\cdots F}_{t\text{ terms}}) = \big(\sigma_2(F)\big)^t$. $\square$

Acknowledgment:

The author expresses his appreciation to Fan Chung for helpful discussions and guidance, and also to the anonymous referee whose comments helped improve the paper.

References

[1] Y. Bilu and N. Linial, "Constructing expander graphs by 2-lifts and discrepancy vs. spectral gap", FOCS 2004 Proceedings, 404–412.

[2] B. Bollobás and V. Nikiforov, "Hermitian matrices and graphs: singular values and discrepancy", Discrete Math. 285 (2004), 17–32.

[3] S. Butler, "Relating singular values and discrepancy of weighted directed graphs", SODA 2006 Proceedings, 1112-1116.

[4] F. Chung, *Spectral Graph Theory*, AMS Publications, 1997.

[5] F. Chung and R. L. Graham, Quasi-random graphs with given degree sequences, preprint.

[6] G. W. Stewart, *Matrix Algorithms*, SIAM, 1998.