Relating singular values and discrepancy of weighted directed graphs

Steven Butler*

Abstract

Various parameters have been discovered which give a measurement of the "randomness" of a graph. We consider two such parameters for directed graphs: the singular values of the (normalized) adjacency matrix and discrepancy (a measurement of how randomly edges have been placed). We will show that these two are equivalent by bounding one by the other so that if one is small then both are small. We will also give a related result for discrepancy of walks when the in-degree and out-degree at each vertex is equal. Both of these results follow from a more general discrepancy property of nonnegative matrices which we will state and prove.

1 Introduction.

A recent area of interest in graph theory is the study of the "quasirandom" properties of graphs [5, 6]. These are measurable properties which we would expect a random graph to have which are equivalent in the sense that if a graph satisfies any of the quasirandom properties it satisfies all of them. Two such quasirandom properties for simple undirected graphs are the eigenvalues of the (normalized) adjacency matrix and discrepancy (a measurement of how randomly the edges are placed).

For undirected graphs it has been known for some time that one could bound the discrepancy by the size of the largest nontrivial eigenvalue [4]. Only recently has progress been made in the other direction. Bollobás and Nikiforov [2] showed that the largest nontrivial eigenvalue can be greater than any constant multiple of discrepancy, while Bilu and Linial [1] were able to get a bound for the largest nontrivial eigenvalue in terms of discrepancy in the case of an undirected regular simple graph. This latter result was generalized by the author to hold for all undirected weighted graphs without isolated vertices [3]. In this note we generalize these results to directed graphs.

We will proceed as follows. In Section 2 we develop discrepancy for directed graphs and state how it is equivalent to the largest nontrivial singular value of the normalized adjacency matrix, we will also note a discrepancy property for walks in the case when the in-degree and out-degree of every vertex is equal. These results follow from a matrix discrepancy result which we will state in Section 3 and prove in Section 4.

1.1 In- and out-degrees for weighted directed graphs.

Our object of consideration will be weighted directed graphs $G = (V, E)$ where we have a weight function, w, on the edges. This weight function is nonnegative and has the property that $w(u{\to}v) > 0$ if and only if $u{\to}v$ is a directed edge of G. The weight function is used to define the adjacency matrix A of G, namely $(A)_{uv} = w(u{\to}v)$.

We use the edge weights to define the in- and out-degrees, respectively d_{in} and d_{out}, of vertices in G,

$$d_{in}(u) = \sum_{v\in V} w(v{\to}u) \text{ and } d_{out}(u) = \sum_{v\in V} w(u{\to}v).$$

When our weight function is 1 on the edges in G, we get the usual definition of in- and out-degree. We will also use D_{in} and D_{out} to denote the in- and out-degree diagonal matrices respectively.

2 Discrepancy for directed graphs.

Discrepancy is used to measure how randomly the edges are placed in the graph. This is done by the bound between the actual edges and expected edges between two subsets of vertices of the graph G. Thus the discrepancy for a directed graph G, denoted $\operatorname{disc} G$, will be the minimal β such that for all $X, Y \subseteq V$

$$\Big|\underbrace{e(X{\to}Y)}_{\text{actual edges}} - \underbrace{\frac{\operatorname{vol}_{out} X \ \operatorname{vol}_{in} Y}{\operatorname{vol} G}}_{\text{expected edges}}\Big| \le \beta \underbrace{\sqrt{\operatorname{vol}_{out} X \ \operatorname{vol}_{in} Y}}_{\text{error normalizing term}}.$$

We now define the terms used above. The actual edges is the sum of the weights of all edges that start in X and end in Y, i.e.,

$$e(X{\to}Y) = \sum_{\substack{u\in X\\ v\in Y}} w(u{\to}v).$$

The in- and out-volume of a subset of the vertices is the sum of the in- and out-degrees of that subset respectively, i.e.,

$$\operatorname{vol}_{in} X = \sum_{u\in X} d_{in}(u) \text{ and } \operatorname{vol}_{out} X = \sum_{u\in X} d_{out}(u),$$

*University of California, San Diego, Dept. of Mathematics. email: sbutler@math.ucsd.edu

and $\operatorname{vol} G := \operatorname{vol}_{in} G = \operatorname{vol}_{out} G$.

As in the undirected case there is a close connection between discrepancy and the (normalized) adjacency matrix.

THEOREM 2.1. *Let G be a weighted directed graph where every vertex has positive in- and out-degree. Then*

$$\operatorname{disc} G \le \sigma_2(D_{out}^{-1/2} A D_{in}^{-1/2}) \le 150 \operatorname{disc} G(1 - 8 \log \operatorname{disc} G).$$

Throughout we will use σ_i to refer to the ith singular value of a matrix. This shows that the discrepancy of a directed graph is equivalent to the nontrivial singular value of the normalized adjacency matrix in the sense that if for some family of graphs when one of the two parameters goes to zero then both go to zero.

2.1 When in- and out-degree are equal. When the in- and out-degree are equal for each vertex then we can generalize the notion of discrepancy. Discrepancy as defined above deals with weighted edges between two subsets, this can also be viewed as weighted walks of length 1. To generalize we will consider weighted walks of length t.

Let $P = x_0 \to x_1 \to x_2 \to \cdots \to x_t$ be a walk of length t from x_0 to x_t. We define the weight of the walk P by

$$w(P) = \frac{w(x_0 \to x_1) w(x_1 \to x_2) \cdots w(x_{t-1} \to x_t)}{d_{x_1} d_{x_2} \cdots d_{x_{t-1}}}.$$

Note that since the in- and out-degrees are equal we have that $d_u := d_{in}(u) = d_{out}(u)$, $\operatorname{vol} X := \operatorname{vol}_{in} X = \operatorname{vol}_{out} X$, and $D := D_{in} = D_{out}$. If we let $\mathcal{P}_t(X, Y)$ denote the set of all paths of length t joining a vertex in X to a vertex in Y, then the actual sum of weighted paths is

$$e_t(X \to Y) = \sum_{P \in \mathcal{P}_t(X \to Y)} w(P).$$

Note that $e_t(X \to Y)/\operatorname{vol} G$ is the probability that a randomly generated walk of length t starts in X and ends in Y. With $e_t(X \to Y)$ we now define $\operatorname{disc}_t G$ to be the minimal β such that for all $X, Y \subseteq V$

$$\left| e_t(X \to Y) - \frac{\operatorname{vol} X \ \operatorname{vol} Y}{\operatorname{vol} G} \right| \le \beta \sqrt{\operatorname{vol} X \ \operatorname{vol} Y}.$$

With this definition we get the following theorem.

THEOREM 2.2. *Let G be a directed graph where in-degree equals out-degree at each vertex. Then*

$$\operatorname{disc}_t G \le \left(\sigma_2(D^{-1/2} A D^{-1/2}) \right)^t.$$

Further, if the in-degree and out-degree are positive for each vertex in G then

$$|\lambda_2|^t \le 150 \operatorname{disc}_t G(1 - 8 \log \operatorname{disc}_t G),$$

where $1 = \lambda_1 \ge |\lambda_2| \ge \cdots$ are the eigenvalues of $D^{-1/2} A D^{-1/2}$.

For undirected graphs $|\lambda_2| = \sigma_2(D^{-1/2} A D^{-1/2})$, showing that the distribution of t-walks and the first non-trivial eigenvalue are equivalent in that case.

3 Matrix theoretic result.

Theorems 2.1 and 2.2 are consequences of the following matrix theoretic results which will be proved in the next section. We use $\langle x, y \rangle$ to denote the usual inner product of x and y, and ψ_S to denote 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 3.1. *Let $B \in M_{m,n}$ with nonnegative entries, and let $R \in M_{m,m}$ and $C \in M_{n,n}$ 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 $T \subseteq [n]$*

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

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

THEOREM 3.2. *Let B, R, C be as above with the further assumption that there are no 0 rows or columns of A (equivalently all diagonal entries of R and C are positive). Then if for all $S \subseteq [m]$ and $T \subseteq [n]$*

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

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

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

These theorems are closely related to discrepancy, note that $\langle \psi_S, B\psi_T \rangle$ is the sum of the entries of the principal submatrix generated by S and T. Similarly, the other term, i.e., $(\langle \psi_S, A\mathbf{1} \rangle \langle \mathbf{1}, B\psi_T \rangle)/\langle \mathbf{1}, A\mathbf{1} \rangle$, is the expected sum of the entries of the principal submatrix if the entries were distributed "randomly" as dictated by the row and column sums. So these theorems show that the largest nontrivial singular value is bounded by the randomness of the entries.

We now show how Theorems 2.1 and 2.2 follow from the above theorems.

Proof. [of Theorem 2.1] Using $B = A$ we have that $R^{-1/2} = D_{out}^{-1/2}$ and $C^{-1/2} = D_{in}^{-1/2}$. It is also easy to see that $e(X \to Y) = \langle \psi_X, A\psi_Y \rangle$, $\mathrm{vol}_{out}\, X = \langle \psi_X, A\mathbf{1} \rangle$ and $\mathrm{vol}_{in}\, Y = \langle \mathbf{1}, A\psi_Y \rangle$. The result now follows from Theorems 3.1, 3.2 and the definition of discrepancy. □

Proof. [of Theorem 2.2] Using

$$B = AD^{-1}AD^{-1}\cdots D^{-1}A = D^{1/2}(D^{-1/2}AD^{-1/2})^t D^{1/2}$$

we have that $R^{-1/2} = C^{-1/2} = D^{-1/2}$. A calculation shows that $e_t(X \to Y) = \langle \psi_X, B\psi_Y \rangle$, $\mathrm{vol}\, X = \langle \psi_X, B\mathbf{1} \rangle$ and $\mathrm{vol}\, Y = \langle \mathbf{1}, B\psi_Y \rangle$. From Theorems 3.1, 3.2 and the definition of discrepancy we have

$$\begin{aligned}\mathrm{disc}_t\, G \;&\le\; \sigma_2\big((D^{-1/2}AD^{-1/2})^t\big)\\ &\le\; 150\,\mathrm{disc}_t\, G(1 - 8\log \mathrm{disc}_t\, G).\end{aligned}$$

As we will see below $D^{1/2}\mathbf{1}$ is the left and the right vector corresponding to the largest singular value of $(D^{-1/2}AD^{-1/2})^t$ and in this case it is also the left and right eigenvector associated with the largest eigenvalue. From this it follows that

$$\begin{aligned}&\sigma_2\big((D^{-1/2}AD^{-1/2})^t\big)\\ &\quad= \sigma_1\Big((D^{-1/2}AD^{-1/2})^t - \frac{1}{\mathrm{vol}\, G}D^{1/2}JD^{1/2}\Big)\\ &\quad= \sigma_1\Big((D^{-1/2}AD^{-1/2} - \frac{1}{\mathrm{vol}\, G}D^{1/2}JD^{1/2})^t\Big)\\ &\quad\le \Big(\sigma_1(D^{-1/2}AD^{-1/2} - \frac{1}{\mathrm{vol}\, G}D^{1/2}JD^{1/2})\Big)^t\\ &\quad= \big(\sigma_2(D^{-1/2}AD^{-1/2})\big)^t.\end{aligned}$$

On the other hand, by biorthogonality we have that the right eigenvector of $(D^{-1/2}AD^{-1/2})^t$ corresponding to λ_2, which we will denote by y, is orthogonal to the left eigenvector of 1, i.e., $D^{1/2}\mathbf{1}$. So we have that

$$\begin{aligned}&\sigma_2\big((D^{-1/2}AD^{-1/2})^t\big)\\ &\quad= \sup_{x:\langle x, D^{1/2}\mathbf{1}\rangle = 0} \frac{\|(D^{-1/2}AD^{-1/2})^t x\|}{\|x\|}\\ &\quad\ge \frac{\|(D^{-1/2}AD^{-1/2})^t y\|}{\|y\|} \;=\; |\lambda_2|^t.\end{aligned}$$

The result now follows. □

4 Proofs of the main result.

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

Items (i) and (ii) are immediate. Item (iii) follows by noting that $(R^{-1/2}BC^{-1/2})^*(R^{-1/2}BC^{-1/2})$ is a non-negative matrix with eigenvector $C^{1/2}\mathbf{1}$ associated with eigenvalue 1, then use the Peron-Frobenius Theorem. Finally, item (iv) follows by subtracting 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 3.1] This follows from the general fact that $|\langle x, My\rangle| \le \sigma_1(M)\|x\|\,\|y\|$, i.e.,

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

It is an easy calculation to verify 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 the other result we combine the approaches of [1, 2], the idea will be to take the actual vectors which are used to produce the singular value and approximate them by "step" vectors which we can use the assumptions of the theorem on. With this in mind we have the following two lemmas.

Lemma 4.1. *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 $\left(\frac{4}{5}\right)^j e^{2\pi i k/29}$ for j an integer and $0 \le k < 29$.*

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 $x_t = re^{i\theta}$. There is a unique integer j so that

$$\big(\frac{4}{5}\big)^j < \frac{r}{d_t} \le \big(\frac{4}{5}\big)^{j-1}.$$

We now set $y_t = \left(\frac{4}{5}\right)^j e^{2\pi i\lfloor 29\theta/2\pi\rfloor/29}$. By construction we have

$$\begin{aligned}0 < |x_t| - |d_t y_t| &\le \Big(\big(\frac{4}{5}\big)^{j-1} - \big(\frac{4}{5}\big)^j\Big)d_t\\ &= \big(\frac{5}{4} - 1\big)\big(\frac{4}{5}\big)^j d_t < \frac{1}{4}|x_t|,\end{aligned}$$

while the argument between x_t and y_t is bounded above by $2\pi/29$. Note, by construction, that $|d_t y_t| \le |x_t|$ from which it follows that $\|Dy\| \le \|x\| = 1$.

If we are given $z_1 = a_1 e^{i\phi_1}$ and $z_2 = a_2 e^{i\phi_2}$, then by the law of cosines we have

$$|z_1 - z_2|^2 = (a_1 - a_2)^2 + 2a_1 a_2\big(1 - \cos(\phi_1 - \phi_2)\big).$$

Applying this to our case, $z_1 = x_t$ and $z_2 = d_t y_t$, we have

$$|x_t - d_t y_t|^2 \le \Big(\frac{1}{4}|x_t|\Big)^2 + 2|x_t|^2\big(1 - \cos(2\pi/29)\big) \le \frac{1}{9}|x_t|^2.$$

From this it follows that

$$\|x - Dy\|^2 = \sum_t |x_t - d_t y_t|^2 \le \frac{1}{9}\sum_t |x_t|^2 = \frac{1}{9}. \qquad \square$$

LEMMA 4.2. *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 $|\langle x, My\rangle| \le \sigma_1(M)\|x\|\,\|y\|$.

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

rearranging then gives the result. $\square$

Proof. [of Theorem 3.2] We will 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 $\|y'\| = 1$ such that

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

Applying Lemma 4.1 twice, there exist (step) vectors x, y with $\|x\|, \|y\| \le 1$; $\|x' - R^{1/2}x\|$, $\|y' - C^{1/2}y\| \le \frac{1}{3}$. It now follows from Lemma 4.2 that

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

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 $[n]$ and get $y = \sum_s \big(\frac{4}{5}\big)^s y^{(s)}$. We now have by the triangle inequality that

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

Note by our assumption, we have that when w, z are 0-1 vectors that

$$\begin{aligned}
|\langle w, \mathcal{B}z\rangle| &= \Big|\langle w, Bz\rangle - \frac{\langle w, B\mathbf{1}\rangle\,\langle \mathbf{1}, Bz\rangle}{\langle \mathbf{1}, B\mathbf{1}\rangle}\Big| \\
&\le \alpha\sqrt{\langle w, B\mathbf{1}\rangle\,\langle \mathbf{1}, Bz\rangle}.
\end{aligned}$$

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

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

where $|w|, |z|$ are gotten from the vectors w, z by taking the absolute value of each entry.

From this it follows that

$$|\langle x^{(t)}, \mathcal{B}y^{(s)}\rangle| \le 29\alpha\sqrt{\langle |x^{(t)}|, B\mathbf{1}\rangle\,\langle \mathbf{1}, B|y^{(s)}|\rangle}. \tag{4.1}$$

We also have that

$$\sum_s |\langle x^{(t)}, \mathcal{B}y^{(s)}\rangle| \le 2\langle |x^{(t)}|, B\mathbf{1}\rangle \quad \text{and} \tag{4.2}$$

$$\sum_t |\langle x^{(t)}, \mathcal{B}y^{(s)}\rangle| \le 2\langle \mathbf{1}, B|y^{(s)}|\rangle.$$

To see this, let $|w|, |M|, |z|$ be as before then using the triangle inequality we have $|\langle w, Mz\rangle| \le \langle |w|, |M|\,|z|\rangle$.

So

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

the other result is proved similarly.

We now have all of the tools which we will need. Let $\gamma = \log_{4/5}\alpha$ and consider

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

To verify this we compare the coefficient of $|\langle x^{(t)}, \mathcal{B}y^{(s)}\rangle|$ on both sides. Clearly when $|s-t| \le \gamma$ 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 hand side. For the first term we note that

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

The inequalities follow from (respectively) (4.1), the geometric-arithmetic mean inequality, any term can show up at *most* $2\gamma+1$ times, and

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

For the second term we use (4.2) to get

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

and for the third term we again use (4.2) to get

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

Putting this together we have that

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

Acknowledgment: The author wishes to express his appreciation to Fan Chung for introducing him to this problem and for many stimulating discussions.

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, "On eigenvalues and discrepancy of graphs", preprint.

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

[5] F. Chung, R. L. Graham and R. M. Wilson, "Quasi-random graphs", Combinatorica 9 (1989), 345–362.

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

[7] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1990.

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

[9] D. West, *Introduction to Graph Theory*, Prentice Hall, Upper Saddle River, 1996.