Journal of Theoretical Probability, Vol. 20, No. 2, June 2007 (© 2007)
DOI: 10.1007/s10959-006-0049-7

The Spectrum of a Random Geometric Graph is Concentrated

Sanatan Rai[1]

Received October 11, 2004; revised January 3, 2005

Consider n points, $x_1, \ldots, x_n$, distributed uniformly in $[0, 1]^d$. Form a graph by connecting two points x_i and x_j if $\|x_i - x_j\| \leq r(n)$. This gives a random geometric graph, $G(\mathcal{X}_n; r(n))$, which is connected for appropriate $r(n)$. We show that the spectral measure of the transition matrix of the simple random walk on $G(\mathcal{X}_n; r(n))$ is concentrated, and in fact converges to that of the graph on the deterministic grid.

KEY WORDS: Random geometric graphs; Spectral measure.

2000 MATHEMATICS SUBJECT CLASSIFICATION: Primary 60D05; Secondary 34L20.

1. INTRODUCTION

Let S be finite set contained in $[0, 1]^d$. Form a graph by connecting two points $u, v \in S$ if $\|u - v\| \leq r$, obtaining a graph $G(S; r)$.

Let $\mathcal{X}_n$ be a set on n points distributed iid $\mathsf{Unif}[0, 1]^d$, we call $G \sim G(\mathcal{X}_n; r(n))$ a random geometric graph. The function $r(n)$ is chosen to be such that $r(n) \to 0$ as $n \to \infty$, but such that G is a.s. and **whp** connected. Herein, **whp** denotes with high probability, i.e. with a probability greater than $1 - n^{-c}$ for some constant $c > 0$. We shall write $\mathcal{D}_n$ for the set of n grid points that are the intersections of axes parallel lines with separation $n^{-1/d}$. Hence, $G(\mathcal{D}_n; r(n))$ is a deterministic graph (see Figure 1).

[1]Department of Management Science and Engineering, Stanford University, Stanford, CA 94305, USA. E-mail: sanatan@gmail.com

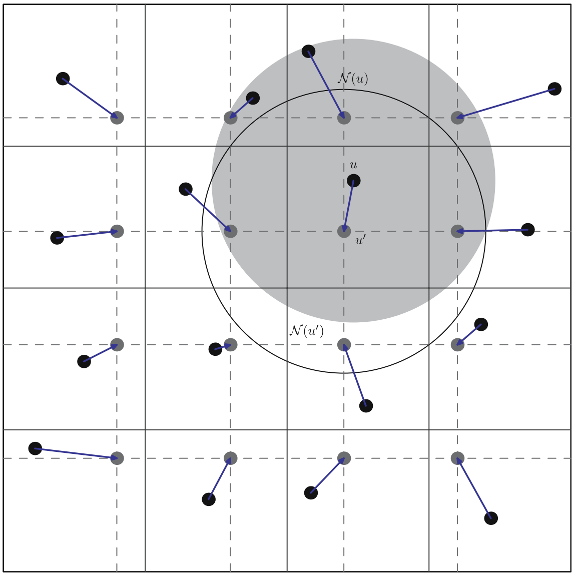

Fig. 1. The grey points are the set $\mathcal{D}_n$ for $n = 16$, the black points form an instance of the random set $\mathcal{X}_n (n = 16)$. The arrows show the optimal matching. The shaded circle represents the neighbourhood of $\mathcal{N}(u)$ of u, i.e. all the points of $\mathcal{X}_n$ that u is connected to. Likewise, $\mathcal{N}(u')$ is the unfilled circle, and is the neighbourhood of u' in $\mathcal{D}_n$.

For a graph G, $P(G)$ shall denote the transition probability matrix for the simple random walk (SRW) on G. That is:

$$P(G)_{uv} = \frac{\chi[u \sim v]}{|\mathcal{N}(u)|},$$

where $\mathcal{N}(u)$ is the set of neighbours of vertex u, and $\chi[u \sim v]$ is the indicator of the event $u \sim v$. Henceforth, we shall write $P(\mathcal{X}_n)$ and $P(\mathcal{D}_n)$ for $P(G(\mathcal{X}_n; r(n)))$ and $P(G(\mathcal{D}_n; r(n)))$, respectively, with **Spec**$(\mathcal{X}_n)$ and **Spec**$(\mathcal{D}_n)$ denoting their spectra. Furthermore, $\mu(\mathcal{X}_n)$ and $\mu(\mathcal{D}_n)$ shall stand for their spectral measures, respectively.

Let M_n denote the **minimum bottleneck matching distance** between X_n and $\mathcal{D}_n$ in d-dimensions (see Eq. (2.1)).

Our main result is:

Theorem 1.1. **Spec**$(\mathcal{X}_n)$ and **Spec**$(\mathcal{D}_n)$ are asymptotically equidistributed **whp**. Moreover, for $r(n)$ such that $M_n = o(r(n))$ a.s. and **whp**, the random spectral measure $\mu(\chi_n)$:

$$\mu(\mathcal{X}_n)(-\infty, x] := \frac{1}{n}|\{\lambda \in \mathbf{Spec}(\mathcal{X}_n) : \lambda \leq x\}|$$

is concentrated, i.e. there exists a sequence of deterministic measures $\{\mu(\mathcal{D}_n)\}$ and $c_d > 0$ such that $(a(n) := n\pi_d r(n)^d)$:

$$\mathbb{P}\left\{\|\mu(\mathcal{X}_n) - \mu(\mathcal{D}_n)\|_{\mathrm{ws}} > \frac{t}{a(n)^{1/4}}\right\} \leq \frac{32na(n)^{1/4}}{t} \times \left[2\exp\left(-\frac{1}{8}\left(\frac{t^4}{t^4+2048d}\right)^2 a(n)\right) + \exp\left(-c_d\frac{t^8}{512}a(n)\right)\right].$$

Here π_d is the volume of the unit ball in $\mathbb{R}^d$. The norm $\|\cdot\|_{\mathrm{WS}}$ is the **Wasserstein distance:**

$$\|\mu - v\|_{\mathrm{ws}} := \sup_{f \text{ Lipschitz}} \left|\int f d\mu - \int f dv\right|.$$

1.1. Related Work

The literature related to this paper may be classified into two groups: results on eigenvalues of random matrices and results on random geometric graphs.

1.1.1. Eigenvalues of Random Matrices

Eigenvalues of random matrices with iid entries have been studied extensively, and can be said to have begun with the work of Wigner.[17] He showed that the spectral measure of a random symmetric matrix with iid entries drawn from a (fixed) distribution with mean zero, variance $1/n$ and finite higher moments, converges to the semicircle law in distribution. The rate of convergence for this class of random matrices was computed in Bai.[1] The work closest in spirit to ours is the paper by Guionnet and Zeituni,[6] wherein they prove rates of convergence using concentration of measure ideas. There is a vast body of research in this area, and rather than provide a comprehensive review, we only note here that the results and methods from this literature are not applicable to our case. In many papers an exact expression of the limiting distributions of eigenvalues is used, and in some the Stiltjes transform is the main tool. Neither technique seems useful in this case. For instance, in the work of Guionnet and Zeituni,[6] the continuity of the trace as a function of the eigenvalues allows them to apply Talagrand's inequality directly. In our case, the randomness comes from the distribution of the points, and the entries of the matrix are not continuous in the position of the points. Hence one cannot apply Talagrand. Our proofs are elementary, and require only a knowledge of Chernoff-Höffding bounds (see Appendix A) and basic real analysis.

Bounds on the second largest eigenvalue modulus of $\mathrm{P}(\mathcal{X}_n)$ were obtained by Boyd *et al.*,[3] and used to establish the mixing times of the SRW and the fastest walk on $G(\mathcal{X}_n; r(n))$.

1.1.2. Random Geometric Graphs

The class of random graphs $G(\mathcal{X}_n; r(n))$ is extensively studied in Penrose.[13] The asymptotic behaviour of various graph properties such as connectivity, chromatic number and the degree distribution are presented therein. This class of graphs is used to model wireless sensor networks, and there is a vast literature on applications. We refer the reader to Booth *et al.*,[2] and Gupta and Kumar,[7] for some representative applications and issues.

1.1.3. Other Similar Results

We also note that a Wielandt–Hoffman theorem for the eigenvalues of discrete matrix approximations to learning kernels were proved in Koltchinskii and Giné[10] and Koltchinskii.[9]

2. PROOF OF THE MAIN THEOREM

Two sequences of real numbers $\{x_n\}$ and $\{y_n\}$ are said to be **asymptotically equidistributed** if for any L^1 function f (Pólya and Szëgo)[14]:

$$\lim_{n\to\infty}\left|\frac{1}{n}\sum_{i=1}^{n} f(x_i)-\frac{1}{n}\sum_{i=1}^{n} f(y_i)\right|=0.$$

Suppose $\{A_n\}$ and $\{B_n\}$ are two sequences of matrices such that $A_n,\ B_n\ \epsilon\ \mathbb{M}_{n\times n}(\mathbb{R})$. We say that the two sequences are asymptotically equivalent if

(1) $\|A_n\|$ and $B\|_n\|$ are uniformly bounded, and
(2) $\|A_n-B_n\|_{\mathrm{HS}}\to 0$ as $n\to\infty$.

Here $\|\cdot\|$ is the usual operator norm:

$$\|A\|:=\sup_{\|x\|=1}|Ax|$$

and $\|\cdot\|_{\text{HS}}$ is the **Hilbert–Schmidt norm:**

$$\|A\|_{\text{HS}} := \left(\frac{1}{n}\sum_{i,j} A_{ij}^2\right)^{1/2} = \left(\frac{1}{n}\text{Tr}(A^T A)\right)^{1/2}.$$

Hereafter, we shall use $\lambda_i(A)$ for the ith largest eigenvalue of matrix A.

With a few straightforward estimates, our main result shall follow easily from Gray[5]:

Theorem 2.1. For any two matrices A and B:

$$\left|\frac{1}{n}\sum_{i=1}^{n}\lambda_i(A) - \frac{1}{n}\sum_{i=1}^{n}\lambda_i(B)\right| \leq \|A-B\|_{\text{HS}}.$$

Therefore, if $\{A_n\}$ and $\{B_n\}$ are two asymptotically equivalent sequences of matrices, then their spectra are asymptotically equivalent.

To prove theorem 1.1 it suffices to show that $\{P(\mathcal{X}_n)\}$ and $\{P(\mathcal{D}_n)\}$ are asymptotically equivalent **whp**, and that $\|P(\mathcal{X}_n) - P(\mathcal{D}_n)\|_{\text{HS}}$ is concentrated.

Let M_n denote the length of the minimum bottleneck matching between $\mathcal{X}_n$ and $\mathcal{D}_n$:

$$M_n \equiv M_n := \min_{\substack{\phi:\mathcal{X}_n \to \mathcal{D}_n \\ \phi \text{ matching}}} \max_{u \in \mathcal{X}_n} \|u - \phi(u)\|. \tag{2.1}$$

Then it is well known that:

Theorem 2.2. **Whp** and a.s. as $n \to \infty$,

$$M_n = \begin{cases} O\left(\left(\frac{\log n}{n}\right)^{1/d}\right), & \text{when } d \geq 3,^{(15)} \\ O\left(\left(\frac{\log^{3/2} n}{n}\right)^{1/2}\right), & \text{when } d = 2,^{(12)} \\ O\left(\sqrt{\frac{\log \varepsilon^{-1}}{n}}\right), & \text{with prob.} \geq 1-\varepsilon, \text{ when } d = 1.^{(4)} \end{cases}$$

Henceforth, for $u \in \mathcal{X}_n$ we shall write u' for its matched point in $\mathcal{D}_n$ under a minimum bottleneck matching. We shall denote by $\mathcal{N}(u)$ the set of neighbours of u, and by $\mathcal{N}(u, u')$ the set of neighbours of u that are mapped to a neighbour of u'. We are now in a position to state a concentration result:

Lemma 2.3. If $r(n)$ is such that $M_n = o(r(n))$ a.s. and **whp,** then there is a constant $c_d > 0$ such that:

$$\mathbb{P}\left\{\|P(\mathcal{X}_n) - P(\mathcal{D}_n)\|_{\mathrm{HS}}^2 > \frac{t}{a(n)}\right\} \leq 2n \times \left[2\exp\left(-\frac{1}{8}\left(\frac{t}{t+8d}\right)^2 a(n)\right) + \exp\left(-c_d \frac{t^2}{2} a(n)\right)\right],$$

where $a(n) := n\pi_d r(n)^d$.

Proof.

$$\begin{aligned}\|P(\mathcal{X}_n) - P(\mathcal{D}_n)\|_{\mathrm{HS}}^2 &= \frac{1}{n}\sum_{u\in\mathcal{X}_n}\sum_{v\in\mathcal{X}_n}\left(\frac{[u\sim v]}{|\mathcal{N}(u)|} - \frac{[u'\sim v']}{|\mathcal{N}(\acute{u})|}\right)^2 \\ &= \frac{1}{n}\sum_{u\in\chi_n}\left(\frac{1}{|\mathcal{N}(u)|} + \frac{1}{|\mathcal{N}(u')|} - \frac{2|\mathcal{N}(u,u')|}{|\mathcal{N}(u)| \quad |\mathcal{N}(u')|}\right) \\ &\leq \max_u \left(\frac{1}{|\mathcal{N}(u)|} + \frac{1}{\mathcal{N}(u')|} - \frac{2|\mathcal{N}(u,u')|}{|\mathcal{N}(u)||\mathcal{N}(u')|}\right). \qquad (2.2)\end{aligned}$$

Therefore,

$$\mathbb{P}\left\{\|P(\mathcal{X}_n) - P(\mathcal{D}_n)\|_{\mathrm{HS}}^2 > \frac{t}{a(n)}\right\} \leq n\max_u \mathbb{P}\left\{\left(\frac{1}{|\mathcal{N}(u)|} + \frac{1}{|\mathcal{N}(u')|} - \frac{2|\mathcal{N}(u,u')|}{|\mathcal{N}(u)||\mathcal{N}(u')|}\right) > \frac{t}{a(n)}\right\}.$$

Now

$$\mathbb{P}\left\{\left(\frac{1}{|\mathcal{N}(u)|} + \frac{1}{|\mathcal{N}(u')|} - \frac{2|\mathcal{N}(u,u')|}{|\mathcal{N}(u)||\mathcal{N}(u')|}\right) > \frac{t}{a(n)}\right\} \leq \mathbb{P}\left\{\frac{2}{|\mathcal{N}(u')|} + \left|\frac{1}{|\mathcal{N}(u)|} + \frac{1}{|\mathcal{N}(u')|}\right| - \frac{2|\mathcal{N}(u,u')|}{|\mathcal{N}(u)||\mathcal{N}(u')|} > \frac{t}{a(n)}\right\}. \qquad (2.3)$$

To bound the last probability, note that for

$$\frac{2}{|\mathcal{N}(u')|} + \left|\frac{1}{|\mathcal{N}(u)|} - \frac{1}{|\mathcal{N}(u')|}\right| - \frac{2|\mathcal{N}(u,u')|}{|\mathcal{N}(u)||\mathcal{N}(u')|} > \frac{t}{a(n)}$$

to hold, it must be that either

$$\left|\frac{1}{|\mathcal{N}(u)|} + \frac{1}{|\mathcal{N}(u')|}\right| > \frac{t}{2a(n)}$$

or

$$\frac{|\mathcal{N}(u)| - |\mathcal{N}(u,u')|}{|\mathcal{N}(u)|} > t\frac{|\mathcal{N}(u')|}{4a(n)}$$

holds. Thus

$$\begin{aligned}
&\mathbb{P}\left\{\frac{2}{|\mathcal{N}(u')|} + \left|\frac{1}{|\mathcal{N}(u)|} - \frac{1}{|\mathcal{N}(u')|}\right| - \frac{2|\mathcal{N}(u,u')|}{|\mathcal{N}(u)| \quad |\mathcal{N}(u')|} > \frac{t}{a(n)}\right\}\\
&\leq \mathbb{P}\left\{\left|\frac{1}{|\mathcal{N}(u)|} - \frac{1}{|\mathcal{N}(u')|}\right| > \frac{t}{2a(n)}\right\} \qquad (2.4)\\
&+ \mathbb{P}\left\{\frac{|\mathcal{N}(u)| - |\mathcal{N}(u,u')|}{|\mathcal{N}(u)|} > t\frac{|\mathcal{N}(u')|}{4a(n)}\right\}. \qquad (2.5)
\end{aligned}$$

To bound the term in (2.5), observe that $a(n)/2d \leq |\mathcal{N}(u')| \leq a(n)$ This relation follows because $|\mathcal{N}(u')|$ is the number of neighbours of the (deterministic) point u' and $a(n) = n\pi_d r(n)^d$ is just the expected number of points of $\mathcal{X}_n$ in a ball of radius $r(n)$. The $1/2d$ factor is for the corner points. Hence

$$\begin{aligned}
&\mathbb{P}\left\{\frac{|\mathcal{N}(u)| - |\mathcal{N}(u,u')|}{|\mathcal{N}(u)|} > t\frac{|\mathcal{N}(u')|}{4a(n)}\right\} \leq \mathbb{P}\left\{\frac{|\mathcal{N}(u)| - |\mathcal{N}(u,u')|}{|\mathcal{N}(u)|} > \frac{t}{8d}\right\}\\
&\leq \mathbb{P}\left\{\frac{\big||\mathcal{N}(u)| - |\mathcal{N}(u')|\big| + \big||\mathcal{N}(u')| - |\mathcal{N}(u,u')|\big|}{|\mathcal{N}(u)|} > \frac{t}{8d}\right\}.
\end{aligned}$$

However,

$$\begin{aligned}
&\frac{\big||\mathcal{N}(u)| - |\mathcal{N}(u')|\big| + \big||\mathcal{N}(u')| - |\mathcal{N}(u,u')|\big|}{|\mathcal{N}(u)|} > \frac{t}{8t}\\
&\Longleftrightarrow \big||\mathcal{N}(u)| - |\mathcal{N}(u')|\big| + \big||\mathcal{N}(u')| - |\mathcal{N}(u,u')|\big| > \frac{t}{8d}|\mathcal{N}(u)|,
\end{aligned}$$

which imples that

$$\begin{aligned}
&\mathbb{P}\left\{\frac{\big||\mathcal{N}(u)| - |\mathcal{N}(u')|\big| + \big||\mathcal{N}(u')| - |\mathcal{N}(u,u')|\big|}{|\mathcal{N}(u)|} > \frac{t}{8d}\right\}\\
&= \mathbb{P}\left\{\big||\mathcal{N}(u)| - |\mathcal{N}(u')|\big| + \big||\mathcal{N}(u')| - |\mathcal{N}(u,u')|\big| > \frac{t}{8d}|\mathcal{N}(u)|\right\}\\
&\leq \mathbb{P}\left\{\big||\mathcal{N}(u)| - |\mathcal{N}(u')|\big| + \big||\mathcal{N}(u')| - |\mathcal{N}(u,u')|\big| > \frac{t}{8d}\left(|\mathcal{N}(u')| - \big||\mathcal{N}(u)| - |\mathcal{N}(u')|\big|\right)\right\}.
\end{aligned}$$

We can break the last probability further as:

$$\mathbb{P}\left\{\big||\mathcal{N}(u)|-|\mathcal{N}(u')|\big|+\big||\mathcal{N}(u')|-|\mathcal{N}(u,u')|\big|>\frac{t}{8d}\left(|\mathcal{N}(u')|-\big||\mathcal{N}(u)|-|\mathcal{N}(u')|\big|\right)\right\}$$
$$\leq \mathbb{P}\left\{\big||\mathcal{N}(u)|-|\mathcal{N}(u')|\big|>\frac{t}{2t+16d}|\mathcal{N}(u')|\right\}$$
$$+\mathbb{P}\left\{\big||\mathcal{N}(u')|-|\mathcal{N}(u,u')|\big|>\frac{t}{16d}|\mathcal{N}(u')|\right\}.$$

Before continuing further, we note that by the above reductions, we have shown that:

$$\mathbb{P}\left\{\|P(\mathcal{X}_n)-P(\mathcal{D}_n)\|_{\text{HS}}^2>\frac{t}{a(n)}\right\}\leq n\max_u\left[\mathbb{P}\left\{\left|\frac{1}{|\mathcal{N}(u)|}+\frac{1}{|\mathcal{N}(u')|}\right|>\frac{t}{2a(n)}\right\}\right.$$
$$+\mathbb{P}\left\{\big||\mathcal{N}(u)|-|\mathcal{N}(u')|\big|>\frac{t}{2t+16d}|\mathcal{N}(u')|\right\}$$
$$\left.+\mathbb{P}\left\{\big||\mathcal{N}(u')|-|\mathcal{N}(u,u')|\big|>\frac{t}{16d}|\mathcal{N}(u')|\right\}\right].$$

To bound:

$$\mathbb{P}\left\{\big||\mathcal{N}(u')|-|\mathcal{N}(u,u')|\big|>\frac{t}{16d}|\mathcal{N}(u')|\right\},$$

observe that if $\|u-v\|\leq r(n)-2M_n$, then $\|u'-v'\|\leq r(n)$. Thus, all points within a radius of $r(n)-2M_n$ of u map to neighbours of u'. Now $|\mathcal{N}(u,u')|$ is stochastically greater than $\mathsf{Bin}(n,\pi_d r(n)^d(1-2M_n/n)^d)$,so that:

$$\mathbb{P}\left\{\big||\mathcal{N}(u')|-|\mathcal{N}(u,u')|\big|>\frac{t}{16d}|\mathcal{N}(u')|\right\}\leq 2\exp\left(-c_d\frac{t^2}{2}n\pi_d r(n)^d\right),$$

where $c_d>0$ is such that

$$c_d\frac{t}{8}\leq\min\left(\frac{1+t/8}{(1-2M_n^+/r(n))^d}-1,\left|\frac{1-t/8}{(1-2M_n^+/r(n))^d}-1\right|\right)$$

for all large enough n. Here M_n^+ is a constant depending only on n such that $M_n\leq M_n^+$, a.s. and **whp**. Since, we have chosen $r(n)$ such that $M_n/r(n)\to 0$, such a constant c_d always exists.

Putting all our estimates together, and using Lemma A.2 to bound (2.4) after noting that $|\mathcal{N}(u')| = \mathbb{E}[|\mathcal{N}(u)|]$, we obtain:

$$\mathbb{P}\left\{\|P(\mathcal{X}_n) - P(\mathcal{D}_n)\|_{\mathrm{HS}}^2 > \frac{t}{a(n)}\right\} \leq 2n \times \left[\exp\left(-\frac{1}{2}\left(\frac{t}{t+8}\right)^2 a(n)\right) + \exp\left(-\frac{1}{8}\left(\frac{t}{t+8d}\right)^2 a(n)\right) + \exp\left(-c_d \frac{t^2}{2} a(n)\right)\right]$$

and the conclusion of the lemma follows. □

Proposition 2.4. When $r(n)$ is such that $M_n = r(n)$ a.s. and **whp,** for $a(n) \equiv n\pi_d r(n)^d \to \infty$ the two sequences $\{P(\mathcal{X}_n)\}$ and $\{P(\mathcal{D}_n)\}$ are asymptotically equivalent **whp.**

Proof. Since, the matrices $P(\mathcal{X}_n), P(\mathcal{D}_n)$ are stochastic, $\|P(\mathcal{X}_n)\| = \|P(\mathcal{D}_n)\| = 1$, and hence they are of uniformly bounded norm. Lemma 2.3 implies that, **whp**, $\|P(\mathcal{X}_n) - P(\mathcal{D}_n)\|_{\mathrm{HS}} \to 0$. Hence $\{P(\mathcal{X}_n)\}$ and $\{P(\mathcal{D}_n)\}$ are asymptotically equivalent **whp.**

We now prove that the concentration of the Hilbert–Schmidt norm of the matrices $P(\mathcal{X}_n)$ implies concentration of the spectrum. We first need a Lemma:

Lemma 2.5.

$$\mathbb{P}\left\{\sup_{f \text{ Lipschitz}} \left|\frac{1}{n}\sum_{i=1}^{n} f(\lambda_i(\chi_n)) - \frac{1}{n}\sum_{i=1}^{n} f(\lambda_i(\mathcal{D}_n))\right| > 4\varepsilon\right\} \leq \frac{2}{\varepsilon}\mathbb{P}\left\{\|P(\chi_n) - P(\mathcal{D}_n)\|_{\mathrm{HS}} > \varepsilon^2\right\}.$$

Proof. When f is linear, the result follows directly from theorem 2.1 and proposition 2.4. To obtain a uniform bound over *all* Lipschitz f, we use the trick in Guionnet and Zeituni[6] of approximating by a finite class of functions, for a given error bound.

To that end, let f be Lipschitz on $[-1, 1]$, and fix $\varepsilon > 0$. Define:

$$g_\varepsilon(x) := x\chi[0 \leq x \leq \varepsilon] + \varepsilon\chi[x > \varepsilon].$$

Then for f_ε recursively defined by, $f_\varepsilon(x)=0$ for $x<-1$ and for $x\in[-1,1]$:

$$f_\varepsilon(x)=\sum_{i=0}^{\lceil (x+1)/\varepsilon\rceil}\left(2\chi\left[f(-1+(i+1)\varepsilon)>f_\varepsilon(-1+i\varepsilon)\right]-1\right)g_\varepsilon(x+1-i\varepsilon),$$

we must have $\|f-f_\varepsilon\|\le\varepsilon$. Thus, we can approximate any Lipschitz f to within ε by a weighted sum of at most $4/\varepsilon$ functions, each of which has two constant portions joined by a linear part. Since the weights are ± 1, we have:

$$\begin{aligned}
&\mathbb{P}\left\{\sup_{f\text{ Lipschitz}}\left|\frac{1}{n}\sum_{i=1}^{n}f(\lambda_i(\mathcal{X}_n))-\frac{1}{n}\sum_{i=1}^{n}f(\lambda_i(\mathcal{D}_n))\right|>4\varepsilon\right\}\\
&\le\frac{4}{\varepsilon}\sup_k\mathbb{P}\left\{\left|\frac{1}{n}\sum_{i=1}^{n}g_k(\lambda_i(\mathcal{X}_n))-\frac{1}{n}\sum_{i=1}^{n}g_k(\lambda_i(\mathcal{D}_n))\right|>\varepsilon^2\right\}\\
&\le\frac{4}{\varepsilon}\mathbb{P}\left\{\|P(\mathcal{X}_n)-P(\mathcal{D}_n)\|_{\mathrm{HS}}>\varepsilon^2\right\},
\end{aligned}$$

where $g_k(x):=g_\varepsilon(x-k\varepsilon)$, and the last inequality follows from Theorem 2.1. □

Proof of theorem 1.1. Proposition 2.4 and Theorem 2.1 imply that the spectra asymptotically equidistributed **whp.** By Lemma 2.5, and the estimate in Lemma 2.3:

$$\begin{aligned}
&\mathbb{P}\left\{\|\mu(\mathcal{X}_n)-\mu(\mathcal{D}_n)\|_{\mathrm{WS}}>\frac{t}{a(n)^{1/4}}\right\}\\
&\le\frac{16a(n)^{1/4}}{t}\mathbb{P}\left\{\|P(\chi_n)-P(\mathcal{D}_n)\|_{\mathrm{HS}}>\frac{t^2}{16\sqrt{a(n)}}\right\}\\
&\le\frac{32na(n)^{1/4}}{t}\\
&\quad\times\left[2\exp\left(-\frac{1}{8}\left(\frac{t^4}{t^4+2048d}\right)^2a(n)\right)+\exp\left(-c_d\frac{t^8}{512}a(n)\right)\right].
\end{aligned}$$

3. DISCUSSION AND OPEN PROBLEMS

Note that if we define $\mu_n^{\{x\}}$ to be the empirical measure

$$\mu_n^{\{x\}}(\alpha)=\frac{1}{n}\sum_{i=1}^{n}\delta(\alpha-x_i),$$

then $\{x_n\}$ and $\{y_n\}$ are asymptotically equidistributed if and only if $\left\|\mu_n^{\{x\}} - \mu_n^{\{y\}}\right\|_{\mathrm{WS}} \to 0$, since we may approximate integrable functions by Lipschitz ones.

Our simulations suggest that in fact the eigenvalues of $P(\mathcal{X}_n)$ and $P(\mathcal{D}_n)$ are **asymptotically absolutely equally distributed**,[16] that is:

Conjecture 1. For any $f \in \mathcal{C}[-1,1]$:

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} |f(\lambda_i(\chi_n)) - f(\lambda_i(\mathcal{D}_n))|^2 = 0,$$

where $\lambda_i(\mathcal{X}_n)$ and $\lambda_i(\mathcal{D}_n)$ are the ith largest eigenvalues of $P(\mathcal{X}_n)$ and $P(\mathcal{D}_n)$, respectively.

Such a result would have followed immediately, if the matrices were symmetric. This is, of course, not true. However, the matrices are *almost* symmetric in the asymptotic limit, so one would expect the matrices to be *almost* diagonalisable via unitary matrices. Then the conjecture would follow immediately from the theory of asymptotically equivalent matrices. As it stands, one needs some other manner of bounding the differences

$$\frac{1}{n}\sum_{i=1}^{n} |\lambda_i(\chi_n) - \lambda_i(\mathcal{D}_n)|^2,$$

to obtain say, a Wielandt–Hoffman type theorem. Then the conjecture would follow by the Stone–Weierstrass theorem.[5]

It also seems reasonable to posit exponential tail bounds for the Wielandt–Hoffman type result.

Another natural question to ask is: what is the behaviour of the resolvent $R_n(x) := (I - xP(\mathcal{X}_n))^{-1}$, for $x \notin \mathbf{Spec}(\mathcal{X}_n)$. Whereas $(I - P(\mathcal{X}_n)) \to \Delta$, the asymptotic behaviour of the resolvent is not known. Convergence results for the resolvents of random walks in random environments (specifically, on $\mathbb{Z}^d$ with ergodic random bond percolation) were proved in Künnemann.[11]

Appendix A

Let $X \sim \mathsf{Bin}(n, p)$ then $\mathbb{E}X = np$. The following result is standard (see, e.g., Janson *et al.*:[8]):

Lemma A.1. (Chernoff–Höffding). For $t \geq 0$:

$$\mathbb{P}\{|X - \mathbb{E}X| > t\} \leq 2\exp\left(-\frac{t^2}{2}\mathbb{E}X\right).$$

We have used a modification in our proofs:

Lemma A.2. (Chernoff–Höffding bounds for reciprocals). For $t \geq 0$:

$$\mathbb{P}\left\{\left|\frac{1}{X} - \frac{1}{\mathbb{E}X}\right| > \frac{t}{\mathbb{E}X}\right\} \leq 2\exp\left(-\frac{1}{2}\left(\frac{t}{1+t}\right)^2 \mathbb{E}X\right).$$

Proof. Note that:

$$\begin{aligned}
\mathbb{P}\left\{\left|\frac{1}{X} - \frac{1}{\mathbb{E}X}\right| > \frac{t}{\mathbb{E}X}\right\} &= \mathbb{P}\{|X - \mathbb{E}X| > tX\} \\
&\leq \mathbb{P}\{|X - \mathbb{E}X| > t(\mathbb{E}X - |X - \mathbb{E}X|)\} \\
&= \mathbb{P}\left\{|X - \mathbb{E}X| > \frac{t}{1+t}\mathbb{E}X\right\} \\
&\leq 2\exp\left(-\frac{1}{2}\left(\frac{t}{1+t}\right)^2 \mathbb{E}X\right).
\end{aligned}$$

□

APPENDIX B. $\mu(\mathcal{D}_n)$ FOR $d = 1, 2$

Figures 2 and 3 show the spectra of $P(\mathcal{D}_n)$ for various values of $a(n)$, when $n = 529$ for $d = 1$ and $d = 2$, respectively. The similarities of the spectra follow from the fact that for large $n, a(n)$ is the major determiner of the behaviour of the system $P(\mathcal{D}_n)\mathbf{x} = \lambda\mathbf{x}$. To see this, consider the $d = 1$ case. The ith smallest point has min $(a(n), i)$ neighbours on the left, and min $(a(n), n - i)$ neighbours on the right. Hence, when the system $P(\mathcal{D}_n)\mathbf{x} = \lambda\mathbf{x}$ is written out in full, the ith equation is:

$$\sum_{k=i-\min(a(n),i)}^{i+\min(a(n),n-i)} x_k = \deg(i)\lambda x_i,$$

where $\deg(i) = \min(a(n), i) + \min(a(n), n - i)$ is the degree of point i. A similar structure is exhibited for higher dimensions. Thus, informally at least, it is not surprising that the spectra are similar for the same values of $a(n)$ irrespective of dimension.

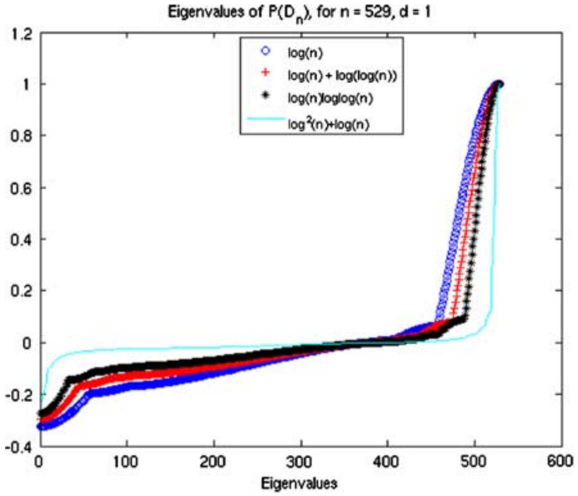

Fig. 2. Eigenvalues of $P(\mathcal{D}_n)$ for $n = 529$ and $d = 1$, for $a(n) = \log n, \log(n) + \log\log n, \log n \log\log n, \log^2 n + \log n$.

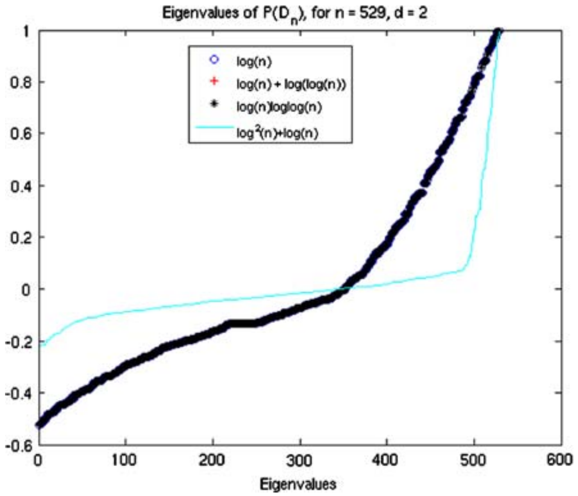

Fig. 3. Eigenvalues of $P(\mathcal{D}_n)$ for $n = 529$ and $d = 2$, for $a(n) = \log n, \log(n) + \log\log n, \log n \log\log n, \log^2 n + \log n$.

ACKNOWLEDGMENTS

The author gratefully acknowledges support from NSF Award CCR-0339262.

REFERENCES

1. Bai, Z. D. (1993). Convergence rate of expected spectral distributions of large random matrices, part i: Wigner matrices, part ii: sample covariance matrices. *Ann. Probab.*
2. Booth, L., Bruck, J., Franceschetti, M., and Meester, R. (2003). Covering algorithms, continuum percolation and the geometry of wireless networks. *Ann. Appl. Probab*. **13**(2), 722–741.
3. Boyd, S., Ghosh, A., Prabhakar, B., and Shah, D. (2004). Mixing times of random walk on geometric random graphs. *Submitted for publication*.
4. Goel, A., Rai, S., and Krishnamachari, B. (2004). Sharp thresholds for monotone properties in random geometric graphs. In *Proceedings of the Symposium on the Theory of Computing*, pages 13–23.
5. Gray, R. M. (1971). Toeplitz and circulant matrices: A review. *Revised 2002*, URL: http://ee.stanford.edu/ gray/toplitz.pdf.
6. Guionnet, A., and Zeituni, O. (2000). Concentration of the spectral measure for large matrices. *Electron. Commun. Probab.* **9**, 119–236.
7. Gupta, P., and Kumar, P. R. (1998). *Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W.H. Fleming*, chapter Critical power for asymptotic connectivity in wireless networks. Birkhaüser, Boston.
8. Janson, S., and Luczak, T., Ruciński, A. (2000). *Random Graphs*, Wiley Interscience series in discrete mathematics and optimization. Wiley-Interscience, New York.
9. Koltchinskii, V.I. (1998). Asymptotics of spectral projections of some random matrices approximating integral operators. In Eberlein, E., Hahn, M., Talagrand, M. (eds.). *High Dimensional Probability*, volume 43 of *Progress in Probability*, Birkhäuser, pages 191–227.
10. Koltchinskii, V.I., Giné, E. (2000). Random matrix approximation of spectra of integral operator. *Bernoulli*.
11. Künnemann, R. (1983). The diffusion limit for reversible jump processes on $\mathbb{Z}^d$ with ergodic random bond conductivities. *Commun. Math. Phys.* **90**, 27–68.
12. Leighton, F.T., Shor, P.W. (1989). Tight bounds for minimax grid matching, with applications to the average case analysis of algorithms. *Combinatorica* **9**, 161–187.
13. Penrose, M.D. (2003). *Random Geometric Graphs*, Volume 5 of *Oxford Studies in Probability*, Oxford University Press, oxford.
14. Pólya, G., Szẽgo, G. (1998). *Problems and Theorems in Analysis* I. Springer-Verlag, Berlin (reprint edition).
15. Shor, P.W., Yukich, J.E. (1991). Minmax grid matching and empirical measures. *Ann. Probab.* **19**(3), 1338–1348.
16. Trench, W.F. (2003). Absolute equal distribution of the spectra of hermitian matrices. *Linear Algebra and Its Applications*.
17. Wigner, E.P. (1985). On the distribution of the roots of certain symmetric matrices. *Ann. Math.* **67**, 325–327.