Eigenvectors of the discrete Laplacian on regular graphs—a statistical approach

2008 J. Phys. A: Math. Theor. 41 435203

(http://iopscience.iop.org/1751-8121/41/43/435203)

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **41** (2008) 435203 (17pp) doi:10.1088/1751-8113/41/43/435203

Eigenvectors of the discrete Laplacian on regular graphs—a statistical approach

Yehonatan Elon

Department of Physics of Complex Systems, The Weizmann Institute of Science, 76100 Rehovot, Israel

Received 17 April 2008, in final form 17 August 2008
Published 30 September 2008
Online at stacks.iop.org/JPhysA/41/435203

Abstract

In an attempt to characterize the structure of eigenvectors of random regular graphs, we investigate the correlations between the components of the eigenvectors associated with different vertices. In addition, we provide numerical observations, suggesting that the eigenvectors follow a Gaussian distribution. Following this assumption, we reconstruct some properties of the nodal structure which were observed in numerical simulations, but were not explained so far (Dakel *et al* 2007 *APPROX-RANDOM* pp 436–48).

PACS numbers: 02.10.Ox, 05.45.Mt

(Some figures in this article are in colour only in the electronic version)

1. Introduction

In the past few decades, the spectral properties of regular graphs had attracted considerable attention of researchers from diverse disciplines such as combinatorics, information theory, theoretical and applied computer science, quantum chaos and spectral theory, to list only a few (for a detailed review, consider [1]). In order to understand better the eigenvectors of the Laplacian on such graphs, we try to establish some analogies between those eigenvectors, to eigenfunctions of chaotic manifolds.

In the following, we consider the statistical properties of $G(n, d) = (V, E)$—a graph which is chosen uniformly at random from the set of d-regular graphs on n vertices. The graph can be uniquely described by its *adjacency matrix* A (also known as the connectivity matrix), where $A_{i,j} = 1$ if v_i and v_j are adjacent vertices in G, or zero otherwise. The action of the (discrete) Laplacian on a function $f : V \to \mathbb{R}$ is

$$Lf_i = \sum_{i \sim j} (f_i - f_j),$$

where $f_i \equiv f(v_i)$, and the summation is over all vertices v_j which are adjacent to v_i. For regular graphs, the Laplacian can be expressed in a matrix form as $L = dI - A$, therefore

an eigenvector of the adjacency matrix with an eigenvalue λ, is also an eigenvector of the Laplacian, with an eigenvalue $\mu = d - \lambda$.

The eigenvalues and eigenvectors of A contain valuable information about the structure of the graph. The relations between the spectrum of the adjacency matrix and the expansion properties of G (see section 2.2) have been thoroughly investigated and were found useful for coping with a variety of tasks. To mention some, the study of expanders is related to the evaluation of convergence rates for Markov chains [2] and the study of metric embeddings [3] in mathematics. In computer science, one uses expanders for the analysis of communication networks, construction of efficient error-correcting codes and in the theory of pseudorandomness [4]. The eigenvectors of A are being successfully used in various algorithms, such as partitioning and clustering (e.g. [5–7]).

As one can learn from spectral properties about the structure of the graph, we can go the other way around. In the study of quantum properties of (classically) chaotic systems, one is commonly interested in statistical properties of the spectrum and eigenstates of the corresponding Schrödinger operator. While quantum operators on graphs (such as the Laplacian) are well defined, the classical analog is not obvious. A plausible classical extension would be to consider a random walk on the graph. For a connected graph which is not bipartite, it is known that random walks are mixing exponentially fast (e.g. [8]). Since a fast mixing system is chaotic, one might expect that the quantum properties of a generic graph will be related in some manner to those of chaotic systems. This conjecture was supported by numerical simulations, relating the spectral statistics of regular graphs to the GOE ensemble [9]. It recently found an explicit formulation [10], relating spectral properties to cycles in G, through a trace formula.

The main goal of this paper is the characterization of statistical properties of eigenvectors of (n, d) graphs. As this work is inspired by analog findings for chaotic wavefunctions, and uses extensively several combinatorial properties of random regular graphs, we dedicate the following section to review some relevant results concerning chaotic billiards and (n, d) graphs.

In section 3, we examine the correlations between different components of the eigenvectors. We derive an explicit limiting expression for the (short distance) empirical covariance, which depends only on the eigenvalue λ.

In section 4, we provide a numerical evidence which suggests that the distribution of the eigenvectors can be approximated by a Gaussian measure.

In section 5, we calculate the Gaussian prediction for several properties of the nodal pattern of eigenvectors, such as the expected number of nodal domains and their expected structure.

2. A brief review of previous results

2.1. Eigenfunctions of chaotic billiards

A classical billiard system is defined as a point particle, which is confined to a domain $\mathcal{D} \subset \mathbb{R}^n$. The particle moves with a constant speed along geodesics and collides specularly with the boundary of $\mathcal{D}$. Depending on the shape of the boundary, the dynamics of the particle can be classified as chaotic or regular. A quantum analog would be to consider eigenstates of the Schrödinger operator for a particle confined to $\mathcal{D}$:

$$-\Delta_D \psi(\mathbf{r}) = k^2 \psi(\mathbf{r}),$$

where Δ_D is the Laplace–Beltrami operator,restricted to $\mathcal{D}$, with Dirichlet boundary condition.

The statistics of the wavefunctions $\psi(\mathbf{r})$ rely on the classical properties of $\mathcal{D}$. In [11], a limiting expression for the autocorrelation function

$$C(\mathbf{r}_0, \mathbf{r}, k) = \frac{\langle \psi(\mathbf{r}_0 + \mathbf{r}/2)\psi^*(\mathbf{r}_0 - \mathbf{r}/2)\rangle}{\langle |\psi(\mathbf{r})|^2\rangle} \tag{1}$$

was calculated, where $\langle \cdots \rangle$ denotes averaging over an appropriate spectral window $[k, k+\epsilon k]$, in the semi-classical limit $k \to \infty$.[1] It was shown that for an integrable domain, $C(\mathbf{r}_0, \mathbf{r}, k)$ is anisotropic and depends on the symmetries of the domain. For chaotic domains, the limit of the autocorrelation function is isotropic and universal (for points which are far enough from the boundary [12]), and can be written explicitly as

$$\lim_{k\to\infty} \mathcal{C}(\mathbf{r}, k) = \frac{\Gamma(n/2) J_{n/2-1}(|k\mathbf{r}|)}{(|k\mathbf{r}|)^{n/2-1}}, \tag{2}$$

where $J_\nu(x)$ is the νth Bessel function of the first kind, which decays asymptotically as $J_\nu(x) \sim \cos(x + \phi_\nu)/\sqrt{x}$. Moreover, it was suggested that in the semi-classical limit, the eigenfunction statistics reproduce a Gaussian measure (the *random wave model*), i.e. for $\{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_m\} \in \mathcal{D}$, the probability density of $\psi(\mathbf{R}) \equiv (\psi(\mathbf{r}_1), \ldots, \psi(\mathbf{r}_m))^T$ for a wavefunction chosen uniformly from the spectral window $[k, k + \epsilon k]$, is converging to

$$p(\psi(\mathbf{R})) = \frac{1}{(2\pi)^{m/2}\sqrt{|C_k|}} \exp\left[-\frac{1}{2}\psi(\mathbf{R})^T C_k^{-1} \psi(\mathbf{R})\right], \tag{3}$$

where the covariance matrix is given by $(C_k)_{ij} = \lim_{k\to\infty} \mathcal{C}(\mathbf{r}_i - \mathbf{r}_j, k)$. Although the random wave model is not supported by any rigorous derivation, it was found consistent with some numerical observations, such as [13, 14].

2.2. *Some properties of large regular graphs*

Throughout this paper, we will focus our attention on (n, d) graphs, where $d \geqslant 3$ is fixed, and $n \to \infty$. With a high probability, those graphs are highly connected, or *expanding*. An expander graph $G = (V, E)$ has the property that for every (small enough) subset $S \subset V$, the edge boundary ∂S, which is the set of edges connecting S to $G\backslash S$, is proportional in size to S itself.

A related property of (n, d) graphs, which will be used repeatedly in the following, is the *local tree property*. It is known [15] that for $k < \log_{d-1}(n)$, the numbers C_k of cycles of length k in an (n, d) graph, are distributed asymptotically as independent Poisson random variables with mean $E(C_k) = (d-1)^k/2k$. Therefore, for any $\epsilon > 0$ and as $n \to \infty$, almost all of the vertices of an (n, d) graph are not contained in a cycle of shorter length than $(1-\epsilon)\log_{d-1}(n)$, with a high probability. Equivalently, the ball of radius $\frac{1}{2+\epsilon}\log_{d-1}(n)$ around most of the vertices is a tree.

We will also be interested in the diameter of G, which grows asymptotically as: $\mathrm{diam}(G) = \log_{d-1}(n \log_{d-1} n) + O(1)$ [16].

In the following, we will express logarithms in the natural tree base: $\log(x) \equiv \log_{d-1}(x)$.

The adjacency matrix of a graph is real and symmetric, therefore it has a real spectrum, which is supported on $[-d, d]$. As $n \to \infty$, the spectral measure on A converges to the Kesten–McKay measure [17]:

$$p(\lambda) = \begin{cases} \dfrac{\mathrm{d}}{2\pi} \dfrac{\sqrt{4(d-1) - \lambda^2}}{d^2 - \lambda^2} & \text{for} \quad |\lambda| \leqslant 2\sqrt{d-1} \\ 0 & \text{for} \quad |\lambda| > 2\sqrt{d-1}. \end{cases} \tag{4}$$

[1] As the density of states is scaled as k^{n-1}, in order to obtain a spectral window which is narrow, but contains many states, we demand that $\epsilon k \to 0$, but $\epsilon k^n \to \infty$.

We will use the following notation throughout this paper: eigenvalues of the adjacency matrix are denoted by λ; superscript indices denote eigenvectors: $Af^{(i)} = \lambda_i f^{(i)}$; subscript indices will denote vertices: $f_j^{(i)} = f^{(i)}(v_j)$. We choose the normalization $\langle f, f\rangle = n$, so that $E\left(f_i^2\right) = 1$, irrespective of n. We enumerate the eigenvalues in the customary order: $d = \lambda_1 \geqslant \lambda_2 \cdots \geqslant \lambda_n$. The first eigenvector (or the ground state of L) is the constant vector $f^{(1)} = (1, 1, \ldots, 1)^T$. As the eigenvectors are orthogonal, we get for $i > 1$ that $\sum_j f_j^{(i)} = \langle f^{(1)}, f^{(i)}\rangle = 0$. We would like to emphasize again that for a regular graph, the eigenvectors of the adjacency matrix and the Laplacian are identical. Therefore, the results of this work are applicable (up to rescaling of the eigenvalues) to both of the operators.

For a graph $G = (V, E)$ and a function $f : V \to \mathbb{R}$, a positive (negative) nodal domain of f is a maximal connected component of G, so that $f(v) \geqslant 0$ ($f(v) \leqslant 0$) for all of the vertices in the component[2]. The nodal count of f, which will be denoted by ν, is the number of nodal domains of f.

In [18], Courant theorem is generalized to connected discrete graphs, showing that the jth eigenvector of the Laplacian contains no more than j nodal domains. A constraint on the allowed shapes of domains was derived in [19]: since an adjacency eigenvector satisfies $\lambda f_i = \sum_{i\sim j} f_j$, if $\lambda > k$ (for $k \in \mathbb{N}$), then for every positive (negative) nodal domain, the vertex which obtains the maximal (minimal) value of f on the domain must have at least $k + 1$ adjacent vertices of the same sign, therefore the minimal size of a domain is $k + 2$. Similarly, if $\lambda < 0$, every vertex has at least one adjacent vertex with an opposite sign, therefore for a negative eigenvalue, nodal domains cannot have inner vertices. In addition, by adding assumptions on the structure of the graph (for example, by considering trees only), it is possible to bound the minimal size of a domain for a given eigenvalue [20]. We refer the reader to [21] for a review on the nodal pattern of general graphs.

3. The covariance of an eigenvector

In this section, we would like to estimate the correlation between two distinct components of an adjacency eigenvector in an (n, d) graph. The distance in G between two vertices $v_i, v_j \in V$ is the length of the shortest walk in G from v_i to v_j—we denote the distance by $|i - j|$. Setting the *k-adjacency operator* to be

$$(\tilde{A}_k)_{ij} = \begin{cases} 1, & \text{for } |i - j| = k, \\ 0, & \text{otherwise.} \end{cases}$$

we evaluate the correlations between two components of an eigenvector f at distance k, by computing the *empirical k-covariance* of f and G, defined as

$$\mathrm{Cov}_k^{\mathrm{emp}}(f, G) = \frac{1}{\mathcal{M}_k} \sum_{|i-j|=k} f_i f_j = \frac{1}{\mathcal{M}_k} \langle f, \tilde{A}_k f\rangle, \tag{5}$$

where $\mathcal{M}_k = \sum_{i,j} (\tilde{A}_k)_{ij}$ is the number of (directed) k-neighbors in G.

For $k < \frac{1}{2}\log n$, we can take advantage of the local tree property, in order to find an explicit limiting expression for (5). Under the tree approximation, $\mathcal{M}_k = nd(d-1)^{k-1}$. Moreover, for a tree, $(\tilde{A}_k)_{ij} = 1$ if and only if there is a (unique) walk of length k from v_i to v_j which do not retrace itself at any step. Therefore, for a tree, the operator $\tilde{A}_k$ is identical to the 'non-retracing operator', introduced and calculated in [22]. Clearly, $\tilde{A}_0 = I$, $\tilde{A}_1 = A$, where I is the identity matrix. $\tilde{A}_2 = A^2 - dI$, as one has to eliminate from A^2 (which correspond to

[2] In the following, we will ignore the possibility that $f(v)$ vanishes for some vertex, as this event is of low measure for eigenvectors of (n, d) graphs.

all possible walks of length 2 in G) the walks which return back to their origin at the second step. In a similar manner, one gets for $k > 2$ that

$$\tilde{A}_k = A\tilde{A}_{k-1} - (d-1)\tilde{A}_{k-2}. \tag{6}$$

The first term is due to all paths of length k which do not retrace in the first $k-1$ steps, while the second term eliminates paths which have not retraced in the first $k-1$ steps, but do retrace in the kth step. Since $A^k f = \lambda^k f$, we get by substituting (6) into (5) that in the limit $n \to \infty$, the empirical covariance depends only on the eigenvalue λ and converges to

$$\mathrm{Cov}_k^{\mathrm{tree}}(\lambda) = \frac{1}{d(d-1)^{k-1}} P_k(\lambda), \tag{7}$$

where $P_k(\lambda)$ is given by the recursion relation:

$$\begin{cases} P_1(\lambda) = \lambda \\ P_2(\lambda) = \lambda^2 - d \\ P_{k+2}(\lambda) = \lambda P_{k+1}(\lambda) - (d-1)P_k(\lambda). \end{cases}$$

Introducing Chebyshev polynomials of the second kind [23]:

$$U_k(\cos\theta) = \frac{\sin((k+1)\theta)}{\sin\theta}. \tag{8}$$

The solution to this recursion relation, subject to the initial conditions, can be written as

$$\mathrm{Cov}_k^{\mathrm{tree}}(\lambda) = \frac{1}{d(d-1)^{k/2}} \left((d-1)U_k\left(\frac{\lambda}{2\sqrt{d-1}}\right) - U_{k-2}\left(\frac{\lambda}{2\sqrt{d-1}}\right)\right). \tag{9}$$

The functions $\mathrm{Cov}_k^{\mathrm{tree}}(\lambda)$ are orthogonal polynomials of degree k in λ with respect to the Kesten–McKay measure (4), satisfying

$$\int \mathrm{Cov}_k^{\mathrm{tree}}(\lambda)\,\mathrm{Cov}_{k'}^{\mathrm{tree}}(\lambda)\,p(\lambda)\,\mathrm{d}\lambda = \frac{1}{d(d-1)^{k-1}}\delta_{k,k'}. \tag{10}$$

This relation can be derived as well from the following combinatorial argument: following (5), the left-hand side of (10) is nothing but

$$\lim_{n\to\infty} \frac{n}{\mathcal{M}_k \mathcal{M}_{k'}} \mathrm{Tr}(\tilde{A}_k \tilde{A}_{k'}) = \frac{1}{d(d-1)^{k-1}} \lim_{n\to\infty} \frac{1}{\mathcal{M}_{k'}} \mathrm{Tr}(\tilde{A}_k \tilde{A}_{k'}).$$

Note that $\mathrm{Tr}(\tilde{A}_k \tilde{A}_{k'})$ is the number of closed walks in G, which are combined from a non-retracing walk of length k, followed by a non-retracing walk of length k'. The only way to perform such a walk on a tree is by going back and forth, therefore $\lim_{n\to\infty} \mathrm{Tr}(\tilde{A}_k \tilde{A}_{k'})/\mathcal{M}_k = \delta_{k,k'}$, resulting in (10).

As $|U_k(x)| \leqslant k+1$ for $|x| \leqslant 1$, $\mathrm{Cov}_k^{\mathrm{tree}}(\lambda)$ is an oscillatory function, which decays as $(d-1)^{-k/2}$ (for $|\lambda| \leqslant 2\sqrt{d-1}$). This behavior is analogous to the expected rate of decay for continuous chaotic manifolds (2). The surface of a ball of radius r in $\mathbb{R}^n$ grows as $S(r) \sim r^{n-1}$, while for regular trees, the surface of the ball grows as $S(r) \sim (d-1)^r$. Therefore, in both of the cases, the rate of decay of the covariance is proportional to the root of the area of the sphere.

While for short distances the empirical covariance converges to $\mathrm{Cov}_k^{\mathrm{tree}}(\lambda)$, the validity of the approximation is expected to deteriorate as k exceeds $\frac{1}{2}\log n$. As the computation presented above does not provide an error estimate, we have turned to numerical simulations. We have calculated numerically the empirical covariance (5) for several realizations of regular graphs, and compared them to the limiting expression (9). The graphs were generated following [24] and using *MATLAB*. As expected, for short distances the matching is very good, while for $k \sim \log n$ the deviations are evident. Figure 1 demonstrates this behavior for a realization of (4000, 3) graph, where $\log_2(4000) = 11.97$.

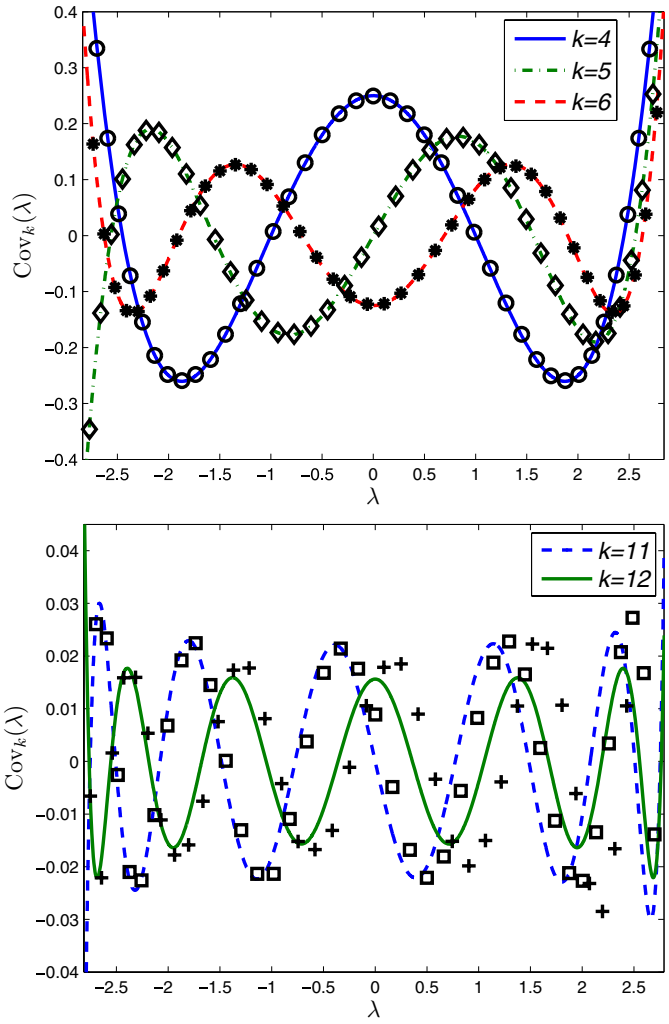

Figure 1. A comparison between $\mathrm{Cov}_k^{\mathrm{tree}}(\lambda)$ (marked by lines), and $\mathrm{Cov}_k^{\mathrm{emp}}(f, G)$, for a single realization of $G(4000, 3)$ (denoted by different markers), where $\log_{d-1}(n) = 11.97$. Upper figure: a comparison for $k = 4, 5, 6$. Lower figure: $k = 11, 12$.

Although for large distances, the empirical covariance deviates from $\mathrm{Cov}_k^{\mathrm{tree}}(\lambda)$, it seems that the expected rate of decay is reproduced quite well, so that asymptotically $\left|\mathrm{Cov}_k^{\mathrm{emp}}(f, G)\right| \sim (d-1)^{-k/2}$. In order to test this assumption, we introduce for a given realization $G \subset G(n, d)$ and k, the two norms

$$N_k^{\mathrm{emp}}(G) = \frac{1}{n}\sqrt{\sum_i \left(\mathrm{Cov}_k^{\mathrm{emp}}(f^{(i)}, G)\right)^2},$$

$$N_k^{\mathrm{tree}}(G) = \frac{1}{n}\sqrt{\sum_i \left(\mathrm{Cov}_k^{\mathrm{tree}}(\lambda_i)\right)^2},$$

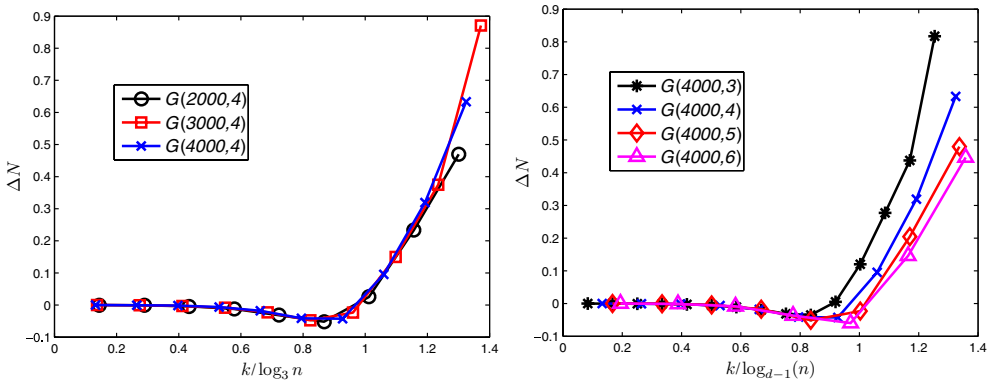

Figure 2. The scaled norm deviation (11) for several realizations of regular graphs. Left: ΔN, as a function of $k/\log_3(n)$, for three realizations of 4-regular graphs consisting of 2000, 3000 and 4000 vertices. The three curves seem to share a joint limit. Right: ΔN, as a function of $k/\log_{d-1}(n)$, for (4000, 3), (4000, 4), (4000, 5) and (4000, 6) graphs. The deviation seems to decrease as d increases.

and define the (scaled) norm deviation as

$$\Delta N(G,k) = \frac{N_k^{\text{emp}} - N_k^{\text{tree}}}{N_k^{\text{emp}} + N_k^{\text{tree}}}. \tag{11}$$

If the asymptotic rate of decay of $\text{Cov}_k^{\text{tree}}$ and $\text{Cov}_k^{\text{emp}}$ is similar[3], then $|\Delta N|$ will be bounded away from 1 for all k and n.

The expected norm deviation of a graph is a function of the parameters n, d and k. However, a comparison of the norm deviation for several realizations of (n, d) graphs with various values of n and d, suggest that it might be well approximated by a function of two parameters only—d, and the scaled parameter $k/\log_{d-1}(n)$ (see figure 2, left). For a fixed value of $k/\log_{d-1}(n)$, the deviation decreases as d increases (figure 2, right). In addition, for $k/\log_{d-1}(n) \leqslant 1$, the norm deviation is bounded away from one, for any d. Since the diameter of an (n, d) graph is $\log(n \log n) + O(1)$ [16], we get that $\max(k)/\log n$ approaches one as n approaches infinity. Therefore we expect $|\Delta N|$ to be bounded away from one for all k.

4. The limiting distribution of eigenvectors

In the previous section we have shown some of the similarities between the limiting expressions for the covariance of regular graphs and the autocorrelation of chaotic billiards (2). In this section and the following, we present extensive numerical evidence, which suggests that the distribution of eigenvectors of (n, d) graphs follows a Gaussian measure, as suggested in [11] for eigenfunctions of quantum billiards (see section 2.1).

For a given realization $G = (V, E)$ of an (n, d) graph, we consider the distribution of the components of its eigenvectors with respect to two ensembles. In *ensemble I*, we fix an eigenvector, satisfying $Af = \lambda f$, and ask for the empirical distribution of $f(v_i)$, where $v_i \in V$ is a uniformly chosen vertex. In *ensemble II*, we fix a vertex $v \in V$, and ask for the empirical distribution of $f^{(j)}(v)$, where $f^{(j)}$ is a uniformly chosen eigenvector.

[3] Meaning that for some positive $c_1(d), c_2(d)$ and for all n and k, we get with a high probability that $c_1 N_k^{\text{tree}}(G) \leqslant N_k^{\text{emp}}(G) \leqslant c_2 N_k^{\text{emp}}(G)$.

Denoting by $p(x)$ and $\Phi(x)$, the density and the cumulative distribution function (cdf) of the standard normal variable:

$$p(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2}, \qquad \Phi(x) = \int_{-\infty}^{x} p(y)\,\mathrm{d}y.$$

We suggest the following limits, as $n \to \infty$:

- *Hypothesis I.* For almost every (n, d) graph $G = (V, E)$ and almost every eigenvector f, the probability that $f(v_i) < x$, where $v_i \in V$ is a uniformly chosen vertex, is bounded by
$$|\mathbb{P}(f(v_i) < x) - \Phi(x)| < \Delta_1(n), \tag{12}$$
where $\Delta_1(n) = O(1/\sqrt{n})$.
- *Hypothesis II.* For almost every (n, d) graph $G = (V, E)$ and almost every $v \in V$, the probability that $f^{(j)}(v) < x$, where $f^{(j)}$ is a uniformly chosen adjacency eigenvector, is bounded by
$$|\mathbb{P}(f^{(j)}(v) < x) - \Phi(x)| < \Delta_2(n), \tag{13}$$
where $\Delta_2(n) = O(1/\sqrt{n})$.

Note that the assumed convergence does not provide information about the tails of the distribution, i.e.events which happens with probability smaller than $1/\sqrt{n}$.

In order to test these hypotheses, we have followed the next procedure. For a given $G \subset G(n, d)$ and an eigenvector f, we define, following hypothesis I, the Kolmogorov–Smirnov (KS) distance for G and f as

$$D_1(G, f) = \max_{x\in\mathbb{R}} |\mathbb{P}(f(v_i) < x) - \Phi(x)|,$$

where $\mathbb{P}(f(v_i) < x) = \frac{1}{n}\sharp\{i | f(v_i) < x\}$. Similarly, for $v \in V$, and following hypothesis II we define

$$D_2(G, v) = \max_{x\in\mathbb{R}} |\mathbb{P}(f^{(j)}(v) < x) - \Phi(x)|,$$

where $\mathbb{P}(f^{(j)}(v) < x) = \frac{1}{n}\sharp\{j | f^{(j)}(v) < x\}$.

Repeating the experiment for all the eigenvectors (vertices), we obtain the scaled cumulative distributions:

$$\mathcal{F}_1(G, x) = \mathbb{P}\left(D_1(G, f) \leqslant \frac{x}{\sqrt{n}}\right), \qquad \mathcal{F}_2(G, x) = \mathbb{P}\left(D_2(G, v) \leqslant \frac{x}{\sqrt{n}}\right). \tag{14}$$

We have executed the process for several values of n and d. In addition, we have also calculated the distribution of the KS distance for i.i.d. Gaussian variables, where the scaled distribution is known [25, 26] to converge to

$$\mathcal{F}_{KS}(x) = 1 - 2\sum_{q=1}^{\infty}(-1)^{q-1}\,\mathrm{e}^{-2q^2x^2}. \tag{15}$$

The numerical simulations (figures 3 and 4) seem to support the two hypotheses, as for both of the ensembles, the KS distribution scales as $1/\sqrt{n}$. In addition, they suggest that the convergence rate for ensemble II is slightly faster than (15), with only a minor dependence in d. The convergence rate for ensemble I is even faster and increases with d.

As a Gaussian distribution is determined by its covariance, and as a limiting expression for the (short distance) covariance of regular graphs was calculated in the previous section, it is not hard to generalize the above hypotheses for the multivariate distribution of an eigenvector. We begin by introducing some additional notations: for a graph $G = (V, E)$, $U = \{u_1, \ldots, u_m\} \subset V$ and a function $f : V \to \mathbb{R}$, we denote $f(U) \equiv (f(u_1), f(u_2), \ldots, f(u_m))$.

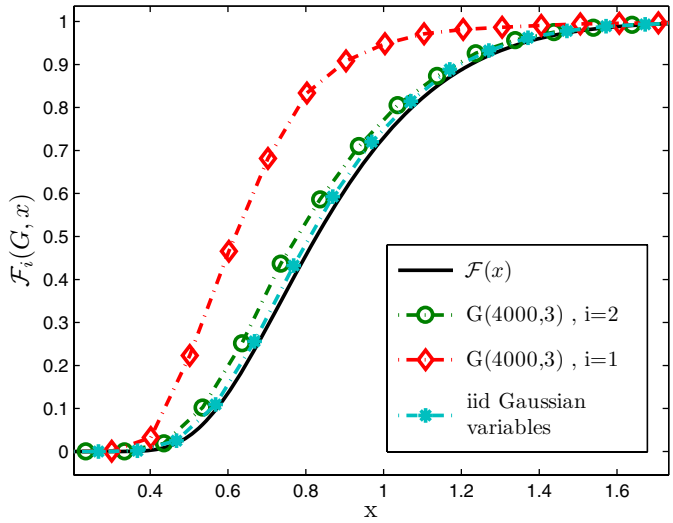

Figure 3. The distribution of the scaled KS distance (14) for a single realization of a (4000,3) graph, with respect to ensemble I (diamonds) and ensemble II (circles), compared to the KS distribution for 4000 vectors of 4000 i.i.d. random variables (asterisks), and to the limiting distribution (15) for i.i.d. variables (solid line).

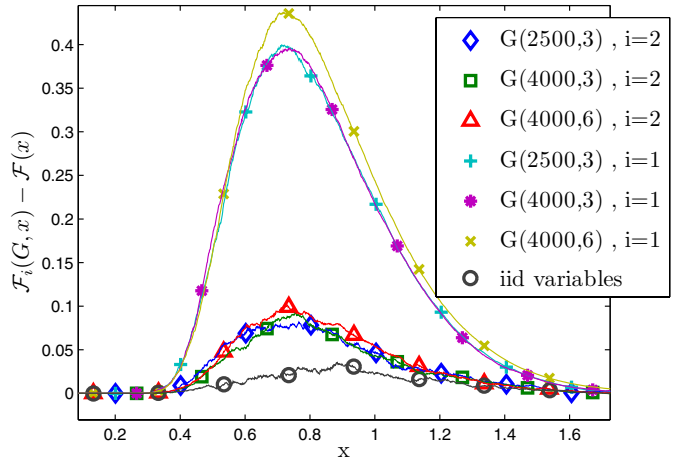

Figure 4. The observed deviations between the measured distribution of the scaled KS distance (14) for several realizations of (n, d) graphs and the limiting distribution for i.i.d. random variables (15). A positive (negative) deviation implies a faster (slower) convergence rate. One can note that the convergence rate with respect to ensemble II is slightly faster than (15), where the dependence in n and d is not evident. The convergence with respect to ensemble I is significantly faster and seems to increase as d increases.

The diameter of U is the maximal distance between two elements in U, while the *distance matrix* $D_G(U)$ is given by $(D_G(U))_{ij} = |i - j|$. A distance matrix is 'good', if it can be embedded in a d-regular tree. Similarly, $U \subset V$ is good, if its distance matrix is good.

The *limiting covariance matrix* $\mathcal{C}_\lambda(U)$, is given by $(\mathcal{C}_\lambda)_{ij} = \mathrm{Cov}^{\mathrm{tree}}_{D_{ij}}(\lambda)$, where $\mathrm{Cov}^{\mathrm{tree}}_k(\lambda)$ is defined in (9).

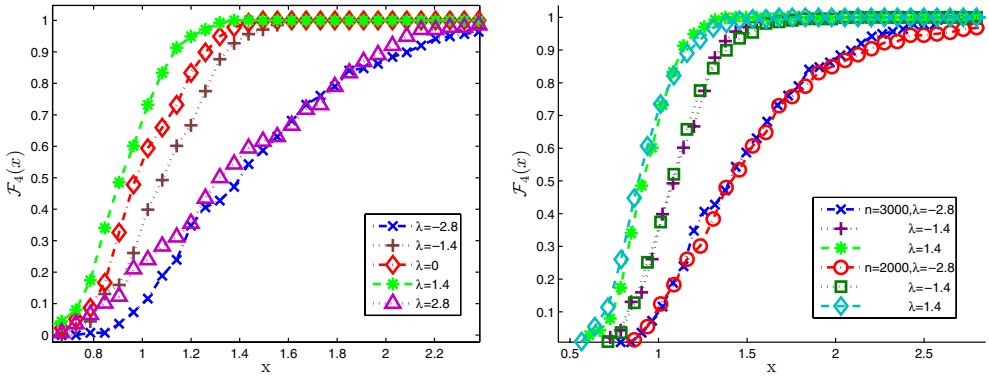

Figure 5. The distribution $\mathcal{F}_3(x)$ of the KS distance. In the left figure we compare the distribution for (3000, 3) graphs and several values of λ. In the right figure, we compare for three values of λ, the distribution of (3000, 3) graphs and (2000, 3). The results suggest that the limiting distribution of $\mathcal{F}_3(x)$ do not depend on n.

We will denote by $p(\mathbf{x}, \Sigma)$ and $\Phi(\mathbf{x}, \Sigma)$, the density and the cdf of the Gaussian vector $\mathbf{x} = (x_1, \ldots, x_m)^T$ with mean zero and covariance matrix Σ. For a positive-definite Σ, they are given by

$$p(\mathbf{x}, \Sigma) = \frac{1}{\sqrt{(2\pi)^m |\Sigma|}} \mathrm{e}^{-\langle \mathbf{x}, \Sigma^{-1}\mathbf{x}\rangle/2}, \qquad \Phi(\mathbf{x}, \Sigma) = \int_{-\infty}^{\mathbf{x}} p(\mathbf{y}, \Sigma)\, \mathrm{d}\mathbf{y}, \quad (16)$$

where the last integral is an m-dimensional integral over the different components.

Finally, for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$, we define $\mathbf{x} < \mathbf{y}$ if $x_i < y_i$ for all $1 \leqslant i \leqslant m$.

Equipped with this notation, we suggest the following, for $\lambda \in [-2\sqrt{d-1}, 2\sqrt{d-1}]$ and fixed $m_1, m_2 \in \mathbb{N}^+$:

- *Hypothesis I (multivariate).* for almost every (n, d) graph $G = (V, E)$ and eigenvector f (satisfying $Af = \lambda f$), and for a good distance matrix D, the probability that $f(U_i) < \mathbf{x}$, where U_i is chosen uniformly from the subsets $\{U \subset V | D_G(U) = D\}$, is bounded by

$$|\mathbb{P}(f(U_i) < \mathbf{x}) - \Phi(\mathbf{x}, \mathcal{C}_\lambda)| < \Delta_3(n), \quad (17)$$

  where $\Delta_3(n) = O(1/\sqrt{n})$.
- *Hypothesis II (multivariate).* for almost every (n, d) graph $G = (V, E)$ and almost every good $U \subset V$ of size m_1 and diameter m_2, the probability that $f^{(j)}(U) < \mathbf{x}$, where $f^{(j)}$ is an adjacency eigenvector, chosen uniformly from the spectral window $[\lambda, \lambda + \epsilon]$, is bounded by

$$|\mathbb{P}(f^{(j)}(U) < \mathbf{x}) - \Phi(\mathbf{x}, \mathcal{C}_\lambda)| < \Delta_4(n, \epsilon), \quad (18)$$

  where $\Delta_4(n, \epsilon) \to 0$ as $\epsilon \to 0$ but $\epsilon n \to \infty$.

A comprehensive numerical examination of the multivariate versions is a hard task. As a beginning, we had to make do with the comparison of the empirical cdf of two adjacent vertices to $\Phi(\mathbf{x}, \mathcal{C}_\lambda)$, where $(\mathcal{C}_\lambda)_{11} = (\mathcal{C}_\lambda)_{22} = 1$, $(\mathcal{C}_\lambda)_{12} = (\mathcal{C}_\lambda)_{21} = \lambda/d$.

In order to test hypothesis I, we fix n, d and λ, generate an (n, d) realization and consider the eigenvector with the closest eigenvalue to λ. We measure the joint empirical distribution of nearest neighbors and calculate its KS distance from the predicted Gaussian distribution. Repeating the process over many realizations with the same parameters, and scaling the distribution by a factor of $\sqrt{n}$, we obtain a distribution $\mathcal{F}_3(x)$ (see equation (14)). In figure 5,

we compare the distribution $\mathcal{F}_3(x)$ for several values of n, d and λ. The observations suggest that the distribution depends on λ (left figure) in a non-trivial manner. More important, for a fixed value of d and λ, $\mathcal{F}_3(x)$ scales similarly for different values of n (right figure), implying that the rate of convergence to the Gaussian measure is indeed $1/\sqrt{n}$.

The second hypothesis is harder to test directly, as the condition for convergence is the existence of a narrow spectral window, which contains a large number of eigenvalues. In order to obtain that, one has to generate and diagonalize an adjacency matrix, which is as large as possible. By using MATLAB's function eigs.m, we have explored relatively narrow spectral windows ($\lambda_{\max} - \lambda_{\min} \approx 0.1$) of (13000, 3) graphs, which contains between 200 and 400 eigenvectors (the exact number depends on the spectral density at λ). The KS distance between the empirical cdf for those windows and $\Phi(\mathbf{x}, \mathcal{C}_\lambda)$, indeed seems to decrease as the window size increases. However, the numerical findings are not sufficient to evaluate the rate of convergence.

5. The nodal structure of eigenvectors

For a graph $G = (V, E)$ and a function $f : V \to \mathbb{R}$, we define the *induced nodal graph* $\tilde{G}_f = (V, \tilde{E}_f)$, by the deletion of edges, which connect vertices of opposite signs in f: $\tilde{E}_f = \{(v_i, v_j) \in E | f_i f_j > 0\}$. We denote by ν_f the number of connected components (nodal domains) of f. In this section we analyze the predictions of the Gaussian model, as stated in hypothesis II, for the structure of the nodal pattern of eigenvectors.

5.1. Distribution of valency

We begin by calculating $p_e(\lambda)$—the probability of an edge $e \in E$ of a random graph, to belong to $\tilde{E}_f$ for an eigenvector f with eigenvalue λ. This is twice the probability of two adjacent vertices to be positive, which according to hypothesis I, should equal

$$p_e(\lambda) = 2\int_0^\infty\int_0^\infty \mathrm{d}\mathbf{f} \frac{1}{2\pi\sqrt{|\mathcal{C}_\lambda|}} \exp\left(-\frac{1}{2}\langle \mathbf{f}, \mathcal{C}_\lambda^{-1}\mathbf{f}\rangle\right),$$

where $(\mathcal{C}_\lambda)_{11} = (\mathcal{C}_\lambda)_{22} = 1$, $(\mathcal{C}_\lambda)_{12} = (\mathcal{C}_\lambda)_{21} = \lambda/d$. Integrating, we get that

$$p_e(\lambda) = \frac{1}{2} + \frac{1}{\pi}\arcsin\left(\frac{\lambda}{d}\right). \tag{19}$$

Note that $p_e(\lambda)$ is symmetric with respect to $\lambda = 0$. In addition, since $|\lambda| \leqslant 2\sqrt{d-1}$, for small values of d, $p_e(\lambda)$ varies considerably along the spectrum (thus, for $d = 3$, $p_e(\lambda)$ can take values in the interval $[0.1, 0.9]$), while as d increases, the allowed interval is being narrowed. As demonstrated in figure 6, equation (19) describes with a high accuracy the observed probability, for various values of d.

The Gaussian model predicts the distribution of the valency of vertices in $\tilde{G}_f$ as well. In order to evaluate $p_j(\lambda)$—the probability of a vertex $v_0 \in V$ to be of valency j in $\tilde{G}_f$, one should consider the joint distribution of f_0 and its d neighbors $\{f_1, \ldots, f_d\}$. Following hypothesis I, we obtain that

$$p_j(\lambda) = 2\binom{d}{j}\int p(\mathbf{f}, \mathcal{C}_\lambda)\prod_{i=0}^{j}\theta(f_i)\prod_{l=j+1}^{d}\theta(-f_l)\,\mathrm{d}\mathbf{f}, \tag{20}$$

where θ is the heaviside step function, $p(\mathbf{f}, \mathcal{C}_\lambda)$ is the Gaussian density and the prefactor is due to the sign symmetry and the different alternatives to choose j out of d adjacent

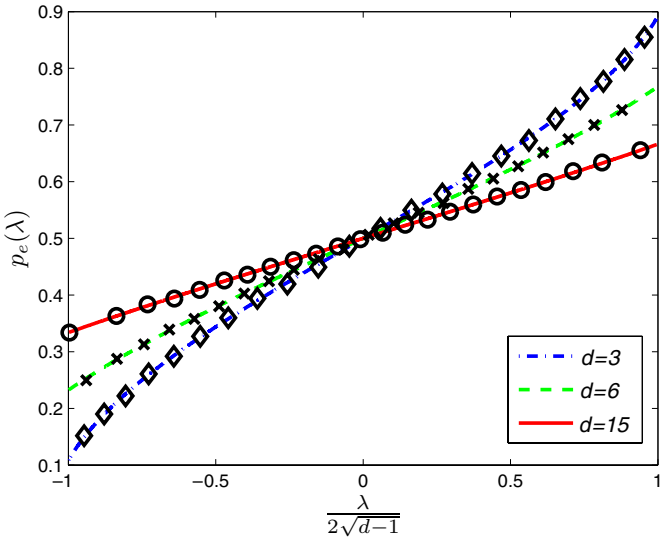

Figure 6. A comparison between the Gaussian prediction for $p_e(\lambda)$ (lines) and the empirical result (markers) for a single random realization of 3, 6 and 15 regular graphs on 4000 vertices.

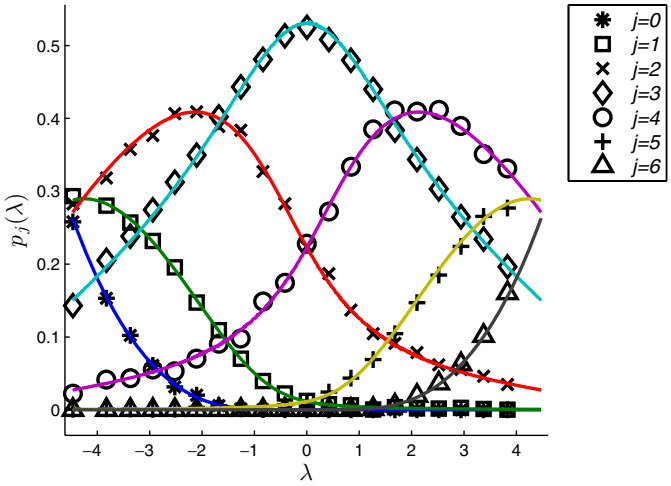

Figure 7. A comparison between the Gaussian prediction (solid lines) of $p_j(\lambda)$ and the empirical result (markers) for a single realization of $G(4000, 6)$.

vertices. The covariance entries are given by $(\mathcal{C}_\lambda)_{ii} = 1, (\mathcal{C}_\lambda)_{0j} = (\mathcal{C}_\lambda)_{j0} = \lambda/d$ and $(\mathcal{C}_\lambda)_{jk} = (\lambda^2 - d)/d(d-1)$, otherwise.

$\mathcal{C}_\lambda$ is singular, as $(\lambda, -1, -1, \ldots, -1)^T$ is an eigenvector of $\mathcal{C}_\lambda$ with eigenvalue zero. This singularity is due to the constraint on the distribution $\lambda f_0 = \sum_j f_j$. As a result $p(\mathbf{f}, \mathcal{C}_\lambda)$ cannot be written in the form (16). Nevertheless, the integral (20) can be evaluated numerically, e.g.by the method of [27]. As in the study of $p_e(\lambda)$, the Gaussian prediction is very close to the observed results. As an example, in figure 7 we compare the Gaussian prediction for $p_j(\lambda)$ and $d = 6$, evaluated by the function qscmvnv.m [28], to the measured result for a single realization of a (4000, 6) graph.

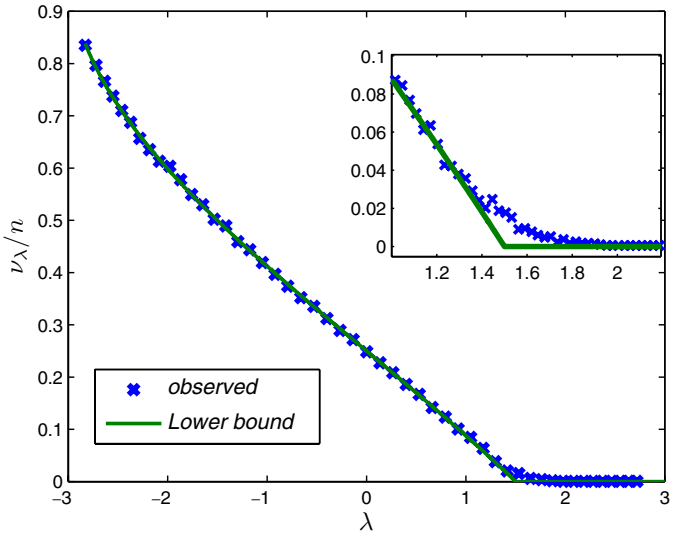

Figure 8. A comparison between the Gaussian bound on the expected nodal count (21) and the observed count for a single realization of $G(4000, 3)$. The inset is a magnification of the spectral window near the bound's flexion—the only part of the spectrum, in which the observed count deviates considerably from the bound.

5.2. The nodal count of an eigenvector

In [29], the following intriguing properties of the nodal count $\{\nu_j\}_{j=2}^n$ for the eigenvectors of (n, d) graphs were observed. First, for all $j < j_0(d, n)$, ν_j was found to be exactly 2, where the relative part j_0/n of eigenvectors with exactly two nodal domains is increasing with d. Second, for small values of d, and for $j > j_0$, the nodal count increases approximately linearly with j. While the known bounds on the nodal count (see section 2.2) are far from being satisfactory in explaining this behavior, we would like to demonstrate in this section, how these numerical findings emerge from the Gaussian model.

Following (19), it is possible to derive a lower bound on the expected nodal count of an eigenvector. The number N of connected components of a graph $G = (V, E)$ is given by $N = V - E + C$, where C is the number of independent cycles in G (cycles are independent, if each cycle contains at least one unique edge). Since on average, the induced nodal graph $\tilde{G}_f$ possesses $p_e(\lambda) \cdot |E| = \frac{nd}{2} p_e(\lambda)$ edges and n vertices, the expected nodal count should be bounded from below (for all of the eigenvectors but the first) by

$$E\left(\nu(n, \lambda)\right) \geqslant \max\left\{2, n\left(1 - \frac{d}{2} p_e(\lambda)\right)\right\}. \tag{21}$$

We should note that this bound is effective only for $d \leqslant 7$, as for larger values, $1 - dp_e(\lambda)/2$ is negative for $|\lambda| \leqslant 2\sqrt{d-1}$.

For low values of d, this crude bound matches surprisingly well the observed nodal count, as is demonstrated in figure 8 for a (4000, 3) graph. The good agreement can be understood if we consider the critical properties of the nodal pattern. Numerical observations [30] suggest that for $d \leqslant 5$, the induced nodal graph $\tilde{G}_f$ exhibits a phase transition at some λ_c, as $n \to \infty$. In the subcritical phase ($\lambda < \lambda_c$), the size of the largest nodal domains is proportional to $\log n$, while in the supercritical phase ($\lambda > \lambda_c$), two giant components of order n emerge.

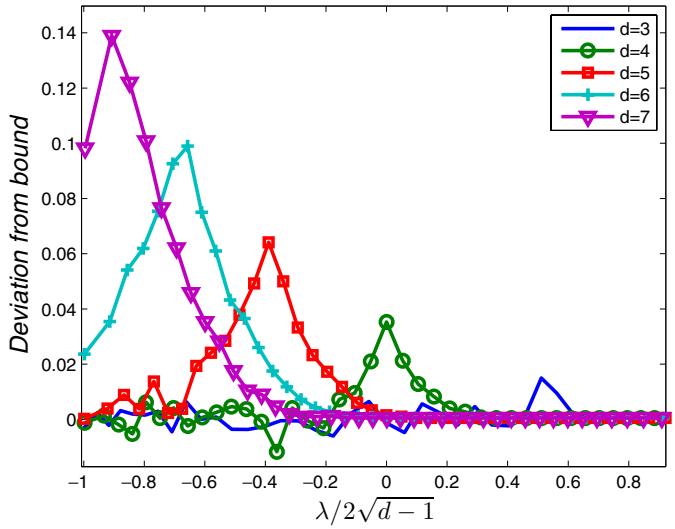

Figure 9. The deviation of ν_λ/n from the lower bound (21), for 3–7 regular graphs on 4000 vertices.

As the number of connected components of size $\log n$ in an (n, d) graph, which contain $\log(\log n)$ cycles is almost surely zero, the expected number of independent cycles in $\tilde{G}_f$ in the subcritical regime must be much smaller than $n \log(\log n)/\log n$ (as there cannot be more than $n/\log n$ domains comparable in size to $\log n$). As a result, for $\lambda < \lambda_c$ the deviation between (21) to the expected count is at most of order $n/\log(n)$. This result is reflected in figure 9, which demonstrates that for $d \leqslant 5$ (where a subcritical phase is observed), the measured count converges to (21), for low enough eigenvalues.

The fact that only two nodal domains are observed for a large number of eigenvectors is also consistent with the properties of a supercritical phase. A general property of supercritical systems is the scarcity of large but finite clusters. the expected number of clusters of size s decays asymptotically as $\exp(-s(p - p_c)^\gamma)$, for some (model-dependent) positive γ. Therefore, the supercritical phase consists of a giant component and 'dust'. When we consider the nodal pattern of supercritical eigenvector (i.e. those with $\lambda > \lambda_c$) two special phenomena occur. The first is the appearance of two giant components—a positive and a negative domain. The second is the rarity of small domains: as was mentioned in section 2.2, the distribution of the eigenvectors is constrained, preventing the existence of small domains for large enough values of λ. As a result, we expect to find only rarely more than two nodal domains, for a considerable amount of first eigenvectors (which are deep enough in the supercritical regime). Moreover, as d increases, the value of λ_c decreases, therefore the expected number of such eigenvectors is supposed to increase with d (as is indeed observed).

As the size distribution of clusters is expected to decay rapidly, we can tighten the bound on the expected nodal count considerably, by calculating $\nu_k(n, \lambda)$—the expected number of domains of size k. It is easy to see that $\nu_1(n, \lambda) = np_0(\lambda)$ where $p_0(\lambda)$ is given by (20). For $k > 1$ the calculation can be carried out in the same spirit: the probability for k given vertices to form a nodal domain of size k can be evaluated, through a $(k(d-1)+2)$-dimensional integral (over the k vertices and their $k(d-2)+2$ neighbors), in a similar manner to (20). Finally, $\nu_k(n, \lambda)$ is given (up to small corrections) by summing the probabilities over all trees of size k and maximal valency d, multiplied by the number of such trees in G. The

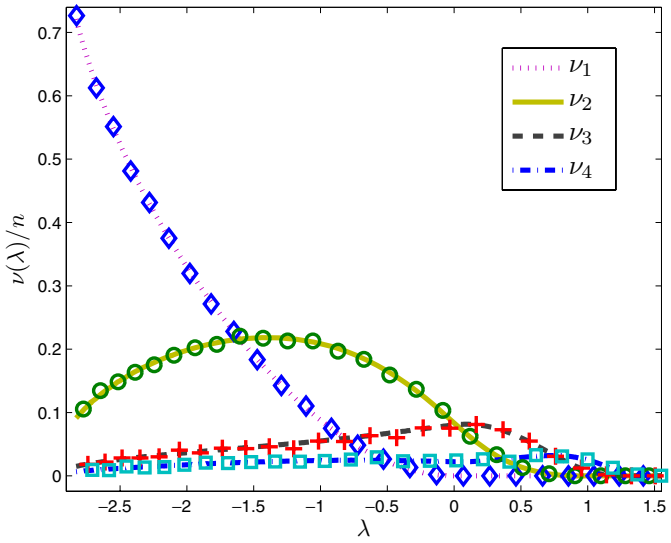

Figure 10. The Gaussian prediction for ν_k/n for $d = 3$ and $1 \leqslant k \leqslant 4$ (lines), compared to the observed count (markers) for a single realization of $G(4000, 3)$.

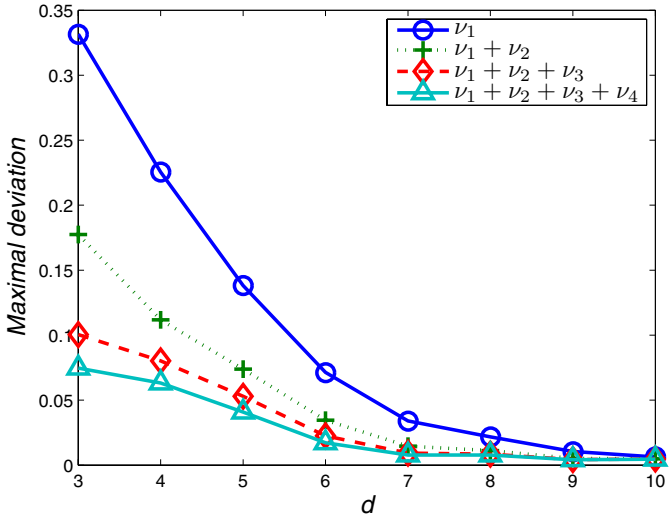

Figure 11. The maximal deviation between the measured nodal count of $3 \leqslant d \leqslant 10$ regular graphs on 4000 vertices to the expected number of domains, of size $1 \leqslant k \leqslant 4$.

agreement between the Gaussian prediction to the observed distribution of ν_k is demonstrated in figure 10.

As $E(\nu) = \sum_{k=1}^{\infty} \nu_k$, and since ν_k is expected to decay fast with k, we get that $\sum_{k=1}^{m} \nu_k(n, \lambda)$ should converge to $E(\nu(\lambda, n))$. This behavior is demonstrated in figure 11, where we plot the maximal deviation between the Gaussian prediction for $\sum \nu_k$ and the measured count of $(4000, d)$ graphs for $3 \leqslant d \leqslant 10$ and $1 \leqslant k \leqslant 4$. In fact, for $d = 3$

we obtain that $\nu_4(\lambda) > 0.9\nu(\lambda)$ for all values of λ, while the convergence rate improves as d increases.

6. Conclusions

As a summary, we collect the main new results of this work concerning the structure of adjacency (Laplacian) eigenvectors:

(i) In the limit $n \to \infty$, the empirical covariance (5) of an eigenvector of a uniformly chosen (n, d) graph, for a distance $k < \frac{1}{2}\log_{d-1}(n)$ is converging to

$$\mathrm{Cov}_k^{\mathrm{tree}}(\lambda) = \frac{1}{d(d-1)^{k/2}}\left((d-1)U_k\left(\frac{\lambda}{2\sqrt{d-1}}\right) - U_{k-2}\left(\frac{\lambda}{2\sqrt{d-1}}\right)\right), \qquad (22)$$

which decays exponentially with k. For $k > \frac{1}{2}\log_{d-1}(n)$, this approximation loses its accuracy; however, the observed rate of decay is still exponential in k.

(ii) We provide a numerical evidence in support of the hypothesis that the distribution of the eigenvectors follows a Gaussian measure, where the covariance is determined by (22).

(iii) We have calculated the Gaussian measure predictions for the expected number of nodal domains in an eigenvector and its dependence in the eigenvalue. We have shown the consistency of the Gaussian hypothesis with various nodal properties, such as valency and size distribution.

Acknowledgments

I wish to thank U Smilansky for guiding me patiently through this work—without his help and advices, this paper could have not been written. I am grateful to N Linial for exposing me to the fascinating world of expanders, and for raising some of the problems I came to investigate. I would like to acknowledge I Oren, G Kozma, O Zeitouni, Y Rinott, S Sodin, J Breuer and A Aronovitch for useful discussions and remarks. Thanks to A Genz for sharing his codes and adjusting it to our needs. The work was supported by the Minerva Center for non-linear Physics and the Einstein (Minerva) Center at the Weizmann Institute, and by grants from the BSF (grant 2006065), the GIF (grant I-808-228.14Q2003) and the ISF (grant 168/06).

References

[1] Hoory S, Linial N and Wigderson A 2006 Expander graphs and their applications *Bull. Am. Math. Soc. (NS)* **43** 439–561 (electronic)

[2] Lovász L and Winkler P 1998 Mixing times *Microsurveys in Discrete Probability (Princeton, NJ, 1997)* (*DIMACS Ser. Discrete Math. Theoret. Comput. Sci.* vol 41) (Providence, RI: American Mathematical Society) pp 85–133

[3] Linial N 2002 Finite metric spaces-combinatorics, geometry and algorithms *Proc. Int. Congress of Mathematicians, vol III (Beijing, 2002)* (Beijing: Higher Education Press) pp 573–86

[4] Spielman D A 1996 Linear-time encodable and decodable error-correcting codes *IEEE Trans. Inf. Theory* **42** 1723–31 (codes and complexity)

[5] Shi J and Malik J 2000 Normalized cuts and image segmentation *IEEE Trans. Pattern Anal. Mach. Intell.* **22** 888–905

[6] Coifman R R 2005 Perspectives and challenges to harmonic analysis and geometry in high dimensions: geometric diffusions as a tool for harmonic analysis and structure definition of data *Perspectives in Analysis* (*Math. Phys. Stud.* vol 27) (Berlin: Springer) pp 27–35

[7] Pothen A, Simon H D and Liou K P 1990 Partitioning sparse matrices with eigenvectors of graphs *SIAM J. Matrix Anal. Appl.* **11** 430–52

[8] Lovász L 1996 Random walks on graphs: a survey *Combinatorics, Paul Erdős is eighty, vol 2 (Keszthely, 1993)* (*Bolyai Soc. Math. Stud.* vol 2) (Budapest: János Bolyai Mathematical Society) pp 353–97
[9] Jakobson D, Miller S D, Rivin I and Rudnick Z 2003 Eigenvalue spacings for regular graphs arXiv:hep-th/0310002
[10] Smilansky U 2007 Quantum chaos on discrete graphs *J. Phys. A: Math. Theor.* **40** F621–30
[11] Berry M V 1977 Regular and irregular semiclassical wavefunctions *J. Phys. A: Math. Gen.* **10** 2083–91
[12] Urbina J D and Richter K 2004 Semiclassical construction of random wave functions for confined systems *Phys. Rev.* E **70** 015201
[13] Blum G, Gnutzmann S and Smilansky U 2002 Nodal domains statistics: a criterion for quantum chaos *Phys. Rev. Lett.* **88** 114101
[14] Elon Y, Gnutzmann S, Joas C and Smilansky U 2007 Geometric characterization of nodal domains: the area-to-perimeter ratio *J. Phys. A: Math. Theor.* **40** 2689–707
[15] Bollobás B 1980 A probabilistic proof of an asymptotic formula for the number of labelled regular graphs *Eur. J. Comb.* **1** 311–6
[16] Bollobás B and De la Vega W F 1982 The diameter of random regular graphs *Combinatorica* **2** 125–34
[17] McKay B D 1981 The expected eigenvalue distribution of a large regular graph *Linear Algebr. Appl.* **40** 203–16
[18] Davies E B, Gladwell G M L, Leydold J and Stadler P F 2001 Discrete nodal domain theorems *Linear Algebr. Appl.* **336** 51–60
[19] Band R, Oren I and Smilansky U 2008 Nodal domains on graphs—how to count them and why? *Analysis on Graphs and its Applications* Proc. Symp. Pure Math. (Providence, RI: American Mathematical Society) pp 5–28
[20] Bıyıkoğlu T and Leydold J 2007 Faber–{K}rahn type inequalities for trees *J. Comb. Theory Ser.* B **97** 159–74
[21] Bıyıkoğlu T, Leydold J and Stadler P F 2007 *Laplacian Eigenvectors of Graphs* (*Lecture Notes in Mathematics* vol1915) (Berlin: Springer) (Perron–Frobenius and Faber–Krahn-type theorems)
[22] Alon N, Benjamini I, Lubetzky E and Sodin S 2007 Non-backtracking random walks mix faster *Commun. Contemp. Math.* **9** 585–603
[23] Abramowitz M and Stegun I A 1964 *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (New York: Dover) (9th Dover printing, 10th GPO printing edition)
[24] Steger A and Wormald N C 1999 Generating random regular graphs quickly *Comb. Probab. Comput.* **8** 377–96 (Random Graphs and Combinatorial Structures (Oberwolfach, 1997))
[25] Kolmogorov A 1933 Sulla determinazione empirica di una legge di distribuzione *Giornale dell'Istituto Italiano degli Attuari* **4** 1–11
[26] Smirnov N 1939 Sur les écarts de la courbe de distribution empirique *Recueil Mathématique* **6** 3–26
[27] Genz A 1992 Numerical computation of multivariate normal probabilities *J. Comput. Graph Stat.* **1** 141–50
[28] Genz A http://www.math.wsu.edu/faculty/genz/homepage
[29] Dekel Y, Lee J and Linial N 2007 Eigenvectors of random graphs: nodal domains *APPROX-RANDOM* pp 436–48
[30] Elon Y and Smilansky U Level sets of eigenvectors on regular graphs (in preparation)