# Eigenvectors of random graphs: Nodal domains

Yael Dekel [*] Hebrew University
James R. Lee [†] University of Washington
Nathan Linial [‡] Hebrew University

**Abstract**

We initiate a systematic study of eigenvectors of random graphs. Whereas much is known about eigenvalues of graphs and how they reflect properties of the underlying graph, relatively little is known about the corresponding eigenvectors. Our main focus in this paper is on the *nodal domains* associated with the different eigenfunctions. In the analogous realm of Laplacians of Riemannian manifolds, nodal domains have been the subject of intensive research for well over a hundred years. Graphical nodal domains turn out to have interesting and unexpected properties. Our main theorem asserts that there is a constant $c$ such that for almost every graph $G$, each eigenfunction of $G$ has at most two large nodal domains, and in addition at most $c$ exceptional vertices outside these primary domains. We also discuss variations of these questions and briefly report on some numerical experiments which, in particular, suggest that almost surely there are just two nodal domains and no exceptional vertices.

## 1 Introduction

Let $G$ be a graph and let $A$ be its adjacency matrix. The eigenvalues of $A$ turn out to encode a good deal of interesting information about the graph $G$. Such phenomena have been intensively investigated for over half a century. We refer the reader to the book [14, Ch. 11] for a general discussion of this subject and to the survey article [10] for the connection between eigenvalues and expansion. Strangely, perhaps, not much is known about the *eigenvectors* of $A$ and how they are related to the properties of $G$. However, in many application areas such as machine learning and computer vision, eigenvectors of graphs are being used with great success in various computational tasks such as partitioning and clustering. For example, see the work of Shi and Malik [19], Coifman, et. al. [5, 6], Pothen, Simon and Liou [16] and others. In particular, a basic technique for spectral partitioning (e.g. Weiss [24]) involves splitting a graph according to its nodal domains. As far as we know, the success of these methods has not yet been given a satisfactory theoretical explanation and we hope that our investigations will help in shedding some light on these issues as well. We also mention that nodal domain counts in graphs are relevant to various studies in statistical physics; see the survey [2].

There is, on the other hand, a rich mathematical theory dealing with the spectrum and eigenfunctions of Laplacians on manifolds. We only mention this important and highly relevant background material and refer the reader who wants to know more about this theory to Chapter 8 in

[*]Work supported in part by a grant from the binational science foundation US-Israel.
[†]Much of this work was done during a visit of the author to the Hebrew University. Research partially supported by NSF CCF-0644037.
[‡]Work supported in part by a grant from the binational science foundation US-Israel.

Marcel Berger's monumental panorama of Riemannian Geometry [3]. Suffices it to say that the adjacency matrix of a graph is a discrete analogue of the Laplacian (we say a bit more about other analogs below). The geometric perspective of graphs is, in our opinion, among the most promising and most exciting aspects of present-day graph theory. Among the recent successes of this point of view is the metric theory of graphs and its computational applications. The success of the metric theory notwithstanding, we believe that other areas of geometry can be incorporated into this line of development.

The only necessary facts we require are that the (geometric) Laplacian has a discrete spectrum, that its smallest eigenvalue is zero and that the corresponding eigenfunction is the constant function. This is analogous to the fact that the first eigenvector of a finite connected graph is a positive vector and in particular, if the graph at hand is $d$-regular, then its first eigenvalue is $d$ and that in the corresponding eigenvector all coordinates are equal.

*Nodal domains* of eigenfunctions of the Laplacian have been studied in depth for more than a century. We will only discuss this concept in the realm of graphs and refer the interested reader to [3] for further information about the geometric setting. So what are nodal domains? Let $G$ be a finite connected graph. It is well-known that every eigenfunction $f$ but the first takes both positive and negative values. These values induce a partition of the vertex set $V(G)$ into maximal connected components on which $f$ does not change its sign; these are the nodal domains of $f$ (see below for the precise definition).

We maintain the following convention: If $G$ is an $n$-vertex graph, we denote the eigenvalues of its adjacency matrix by $\lambda_1 \geq \lambda_2 \geq \ldots$. We simply refer to the $\lambda_i$ as the *eigenvalues of* $G$ and let $f_1, f_2, \ldots$ be the corresponding eigenfunctions, normalized in $\ell_2$. A slight adaptation of classical theorem due to Courant (see, for example, [4]) shows that for every $k$, the eigenfunction $f_k$ has at most $k$ nodal domains. This statement is a bit inaccurate and we refer the reader to [8] for a full account of Courant's theorem for graphs.

We impose throughout some fixed (but arbitrary) ordering on the vertex set $V = \{v_1, v_2, \ldots, v_n\}$, andthe coordinates in eigenvectors of $G$ are arranged in this order. We freely interchange between the vector $(f(v_1), \ldots, f(v_n))$ and the corresponding function $f : V \to \mathbb{R}$. In the graph setting, one has to be careful in defining nodal domains properly.

**Definition 1.1.** *Let $G = (V, E)$ be a graph and let $f : V \to \mathbb{R}$ be any real function. A subset $D \subseteq V$ is a* weak nodal domain *of $f$ if it is a maximal subset of $V$ subject to the two conditions*

1. *$D$ is connected, and*
2. *if $x, y \in D$ then $f(x)f(y) \geq 0$.*

*We say that $D$ is a* strong nodal domain *if (ii) is replaced by*

2'. *if $x, y \in D$, then $f(x)f(y) > 0$.*

The main focus of our research is the following problem.

**Question 1.** *How many nodal domains (strong or weak) do the eigenvectors of $G$ tend to have for $G$ that comes from various random graph models?*

To get some initial idea, we started our research with a numerical experiment whose outcomes were quite unexpected. We sampled numerous graphs from the random graph space $G(n, \frac{1}{2})$. It

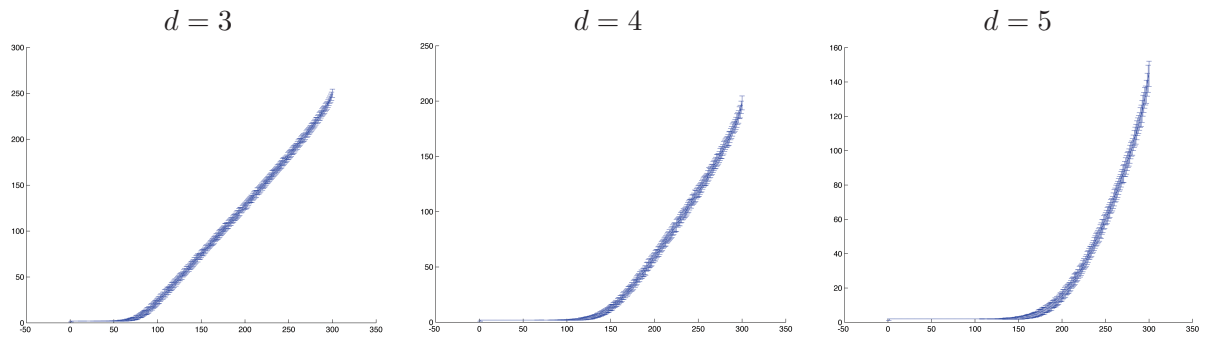


Figure 1: The number of nodal domains in a random $d$-regular 300 vertex graph. There are $y$ nodal domains corresponding to the $x$'th eigenvector (eigenvalues are sorted). For each $d$ we show the average and standard deviation of 100 such random graphs.

turned out that in each and every one of these cases, *all* eigenvectors of the adjacency matrix had *exactly two nodal domains*. (In the experiments, no eigenvector ever had a 0 coordinate, so the notions of strong and weak domains are equivalent.) The same experimental phenomenon was observed for several smaller values of $p > 0$ in the random graph model $G(n, p)$, provided that $n$ is large enough. Even more unexpected were the results obtained for *random regular graphs*. Some of these results are shown in Fig. 1. We found that quite a few of the first eigenvectors have just two nodal domains, and only then the number of nodal domains starts to grow.

As mentioned above, there are other discrete analogs to the geometric Laplacian. One often considers the so-called *combinatorial Laplacian* of a graph $G$. This is the matrix $D - A$ where $A$ is $G$'s adjacency matrix and $D$ is a diagonal matrix with $D_{ii}$ being the degree of the vertex $v_i$. It is well known that this matrix is positive semidefinite, and it is also of interest to investigate similar questions for the combinatorial Laplacian of random graphs. The convention here is that the eigenvalues are sorted as $0 = \mu_1 \leq \ldots \leq \mu_n$. For regular graphs, the question for the adjacency matrix and for the discrete Laplacian are equivalent, since $\lambda_i + \mu_{n-i+1} = d$ for every $i$ and the eigenfunction corresponding to $\lambda_i$ and to $\mu_{n-i+1}$ in the two matrices are identical. However, for graphs in $G(n, p)$ it turns out that the two spectra behave slightly differently. Computer simulations suggest that all eigenvectors of the combinatorial Laplacian that correspond to $\mu_2 \ldots \mu_{n-\Delta}$ for some small $\Delta$ have exactly two nodal domains. However, among the last $\Delta$ eigenfunctions, some have three nodal domains. Fig. 2 suggests that in a constant fraction of the graphs in $G\left(n, \frac{1}{2}\right)$ the eigenvector corresponding to $\mu_n$ has three nodal domains.

Our main theorem for $G(n, p)$ partly establishes these observed phenomena.

**Theorem 1.2.** *For every $p \in (0, 1)$, if $G \sim G(n, p)$, then asymptotically almost surely the following holds for every eigenvector of $G$. The two largest weak nodal domains cover all vertices in $G$ with the exception of at most $O_p(1)$ vertices, where $O_p(1)$ represents a constant depending only on $p$. The two largest strong nodal domains cover all the vertices $\{v \in V : f(v) \neq 0\}$, with the exception of at most $O_p(1)$ vertices. In particular, in both cases every eigenvector of $G$ has at most $O_p(1)$ (strong or weak) nodal domains.*

We remark that our bounds are quite reasonable (see Appendix A). For instance, for $p = 1/2$, we show that there are at most 46 exceptional vertices almost surely.

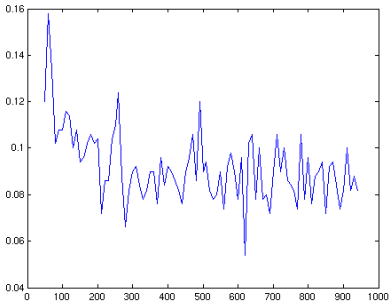


Figure 2: The probability in $G(n, \frac{1}{2})$ that the last eigenvector of the Laplacian has three nodal domains. For each $n$ there were 500 experiments carried out.

### 1.1 Overview of our approach

Given $G \sim G(n,p)$ and an eigenvector $f$ of $G$, we first partition $V(G) = \mathcal{P}_f \cup \mathcal{N}_f \cup \mathcal{E}_f$ where $\mathcal{P}_f, \mathcal{N}_f$ are the largest positive and negative nodal domains of $f$, respectively, and $\mathcal{E}_f$ is a set of exceptional vertices. All eigenvectors are normalized to be unit vectors in $\ell_2$.

We show that if $\mathcal{E}_f$ is large, then we can use the eigenvalue condition, combined with upper estimates on the eigenvalues of $G(n,p)$ to find a large subset $S$ of coordinates for which $\|f|_S\|_2$ is smaller than one would expect for a random unit vector $f \in S^{n-1}$. The last step is to show that the probability that some eigenvector has small 2-norm on a large set of coordinates is exponentially small, and then apply a union bound over all such subsets.

The problem is that the final step seems to require very strong upper bounds on $\|A^{-1}\|$ (i.e. lower bounds on the smallest singular value of $A$) for a random discrete matrix $A$. Although there has been a great deal of progress in this direction [13, 12, 17, 21, 7, 18] (see also the survey [23]), the known bounds for symmetric matrices are far too weak for our purpose. Thus it is crucial that we reduce to proving upper bounds on $\|A^{-1}\|$ when $A$ is a random *rectangular* $\pm 1$ Bernoulli matrix. In this case, we can employ the optimal bounds of [12, 13] which also yield the exponential failure probability that we require. This reduction is carried out in Theorem 3.1.

In Section 3.3, we show that it is possible to get significantly better control on the nodal domains of $G(n,p)$ if we could get slightly non-trivial bounds on the $\ell_\infty$ norms of eigenvectors. This is a possible avenue for future research.

### 1.2 Preliminaries

**Notation 1.1.** *We denote by $G(n,p)$ a random graph with $n$ vertices, where each edge is chosen independently with probability $p$. The set of neighbors of a vertex $x \in V(G)$ is denoted by $\Gamma(x)$.*

**Notation 1.2.** *For a graph $G = (V,E)$ a function $f : V \to \mathbb{R}$, and any subset $S \subseteq V$, we denote $f(S) = \sum_{y \in S} f(y)$. In particular, for every $x \in V$ we denote $f(\Gamma(x)) = \sum_{y \sim x} f(y) = \sum_{y \in V} A_{xy} f(y)$, where $A$ is the adjacency matrix of $G$.*

As usual, the inner product of $f, g : V \to \mathbb{R}$ is denoted $\langle f, g \rangle = \sum_{x \in V} f(x) g(x)$. For $p \geq 1$, we define $\|f\|_p = \left( \sum_{x \in V} |f(x)|^p \right)^{1/p}$, and $\|f\|_\infty = \max_{x \in V} |f(x)|$.

**Definition 1.3.** *For $p \in [0,1]$, we define the random variable $X_p$ by*

$$X_p = \begin{cases} p-1 & \text{with probability } p \\ p & \text{with probability } 1-p. \end{cases}$$

*In particular, $\mathbb{E}X_p = 0$.*

**Definition 1.4.** *Let $M_{m\times k}(p)$ be the $m \times k$ matrix whose entries are independent copies of $X_p$, and let $M_k^{\mathrm{sym}}(p)$ be the symmetric $k \times k$ matrix whose entries above the diagonal are independent copies of $X_p$, and whose diagonal entries are $p$.*

Unless otherwise stated, throughout the manuscript, *all eigenvectors are assumed to be normalized in $\ell_2$.* When $A$ is the adjacency matrix of a graph, we arrange its eigenvalues as $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$. The eigenvector corresponding to $\lambda_1$ is called the *first eigenvector,* and any other is a *non-first eigenvector.*

As is customary in this area, we occasionally say "almost surely" as shorthand for "asymptotically almost surely." All asymptotic statements are made for large $n$ and it is always implicitly assumed that $n$ is large enough.

## 2 Spectral properties of random matrices

We now review some relevant properties of random matrices.

**Theorem 2.1** (Tail bound for symmetric matrices)**.** *If $A \sim M_k^{\mathrm{sym}}(p)$, then for every $\xi > 0$*

$$\Pr\left(\|A\| \geq (2\sqrt{p(1-p)} + \xi)\sqrt{k}\right) \leq 4e^{-(1-o(1))\xi^2 k/8}$$

*Here $\|A\|$ stands for the $\ell_2$ operator norm of $A$.*

*Proof.* Füredi and Komlós prove in [9, §3.3] that the expected value of the largest magnitude eigenvalue of $A$ is at most $(2+o(1))\sqrt{p(1-p)}\sqrt{k}$. Alon, Krivelevich and Vu prove in [1] (see also [11]) that the probability that the largest eigenvalue of $A$ exceeds its median by $\xi\sqrt{k}$ is at most $2e^{-\xi^2 k/8}$, and so is the probability that the smallest eigenvalue of $A$ is smaller than its median by $\xi\sqrt{k}$. As usual in the context of sharp concentration, the expected value of the first and last eigenvalues differs from the median by at most $O(1)$. The conclusion follows. □

We also need a similar bound for $A \sim M_{m\times k}(p)$ whose proof is standard; we repeat it in Appendix A in order to record the exact dependence of the constants on $p$.

**Theorem 2.2** (Tail bound for non-symmetric matrices)**.** *For any $m \geq k$, if $A \sim M_{m\times k}(p)$ then*

$$\Pr\left(\|A\| \geq a_1\sqrt{p(1-p)}\sqrt{m}\right) \leq e^{-a_2 m}$$

*where $a_1, a_2$ are constants that depend only on $p$.*

**Theorem 2.3** (Tail bound for eigenvalues of $G(n,p)$)**.** *Let $G$ be a graph from $G(n,p)$ with $p \in (0,1)$ and let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of $G$'s adjacency matrix. Then for every $i \geq 2$ and every $\xi > 0$*

$$\Pr\left(|\lambda_i| \geq (2\sqrt{p(1-p)} + \xi)\sqrt{n}\right) \leq \exp(-\xi^2 n/32)$$

*Proof.* Füredi and Komlós prove in [9, §3.3] that $\mathbb{E}\left(\max_{i\geq 2}|\lambda_i|\right) \leq 2\sqrt{p(1-p)}\sqrt{n}(1+o(1))$. As before, the theorem is derived by using a tail bound from [1]. □

We now state the following theorem from [13, Thm 3.3] (a slight generalization of [12]), specialized to the case of the sub-gaussian random variables $X_p$.

**Theorem 2.4.** *[13] For any $p \in (0,1)$, $\delta > 0$ there exist constants $\alpha = \alpha(p,\delta) > 0, \beta = \beta(p) > 0$ such that the following holds for all sufficiently large $k$ and every $w \in \mathbb{R}^m$:*

$$\Pr\left[\exists v \in S^{k-1} \text{ s.t. } \|Qv - w\|_2 \leq \alpha\sqrt{m}\right] \leq \exp(-\beta m)$$

*where $m = (1+\delta)k$ and the probability is taken over $Q \sim M_{m\times k}(p)$.*

Strictly speaking the results in [13] only apply for symmetric random variables (and thus only to $X_{1/2}$ in our setting). Note, however, that the proof of the stated bound in [13] is based on three ingredients: (i) The exponential tail bound on the operator norm from Theorem 2.2; (ii) a small-ball probability estimate that is based on the Paley-Zygmund inequality; and (iii) a small-ball probability estimate, based on a Berry-Esséen-type inequality. We note that step (iii) does not require a symmetry assumption (see, e.g. [20, §2.1]).

In order to derive explicit bounds on $\alpha$ and $\beta$, in Appendix A.1, we give a simpler proof of [13, Thm 3.3] (which proof can essentially be read off from [13]) that works when $p$ and $\delta$ are large enough, and uses only (ii). For instance, this proof yields non-trivial results for random graphs in $G(n,p)$ when $p \in [0.18, 0.78]$. In general, this does not require $\delta$ to be small; for instance, the simpler proof yields positive values for $\alpha$ and $\beta$ when $p = 1/2$ and $1+\delta \geq 1000$, which is sufficient for the applications in Section 3.

These estimates are only used in the appendix for calculating explicit upper bounds on the number of exceptional vertices, whereas in the rest of the paper we use the statement of Theorem 2.4 above. Nevertheless, the proof is fairly straightforward, and carries pedagogical value for readers not familiar with such techniques. We refer to Appendix A.1.

The next result follows from taking a union bound in Theorem 2.4.

**Corollary 2.5.** *Maintaining the notations of Theorem 2.4, it is moreover true that*

$$\Pr\left[\exists c \in \mathbb{R}, \exists v \in S^{k-1} \text{ s.t. } \|Qv - cw\|_2 \leq \alpha\sqrt{m}\right] \leq \exp(-\beta m)$$

*for all sufficiently large $k$, and every $w \in \mathbb{R}^m$.*

*Proof.* We may assume that $\|w\|_2 = 1$. Since clearly $\|Qv\|_2 \leq m$, it suffices to prove the bound for $c \in [-2m, 2m]$. Apply Theorem 2.4 for every integer $j \in [-2m, 2m]$ and a union bound to conclude that

$$\Pr\left(\exists \text{ an integer } j \in [-2m,2m] \text{ and } v \in S^{k-1} \text{ s.t. } \|Qv - jw\|_2 \leq \alpha\sqrt{m}\right) \leq \exp\left(-\beta m + \ln(4m+1)\right)$$

But now for any $c \in [-2m, 2m]$, let $j \in [-2m, 2m]$ be the nearest integer. For any matrix $Q$ and $v \in S^{k-1}$

$$\|Qv - cw\|_2 \leq \|Qv - jw\|_2 + \|(j-c)w\|_2 \leq \|Qv - jw\|_2 + 1 \ .$$

□

# 3 Nodal domains

## 3.1 Eigenvectors are not too localized

We show first that the restriction of an eigenvector to a large set of vertices must have a substantial $\ell_2$ norm.

**Theorem 3.1** (Large mass on large subsets)**.** *For every $p \in (0,1)$ and every $\varepsilon > 0$, there exist $\alpha = \alpha(\varepsilon, p) > 0$ and $\beta = \beta(p) > 0$ such that for $n$ large enough, and for every* fixed *subset $S \subseteq [n]$ of size $|S| \geq (\frac{1}{2} + \varepsilon)n$,*

$$\Pr\left[\exists \textit{a non-first eigenvector } f \textit{ of } G \textit{ satisfying } \|f|_S\|_2 < \alpha\right] \leq \exp(-\beta n)$$

*where the probability is over the choice of $G \sim G(n,p)$.*

*Proof.* Let $A$ be the adjacency matrix of $G = (V, E)$, and let $f : V \to \mathbb{R}$ be a non-first eigenvector of $G$ with eigenvalue $\lambda$. Assume w.l.o.g. that $\alpha \leq \frac{1}{2}$, and let $S \subseteq V$ be as in the theorem.

For every $x \in S$, the eigenvector condition $\lambda f(x) = f(\Gamma(x))$ implies that

$$\left|\sum_{y \in V} A_{xy} f(y)\right| = |\lambda f(x)|.$$

Or equivalently,

$$\left|\sum_{y \in V} (p - A_{xy}) f(y) - p \sum_{y \in V} f(y)\right| = |\lambda f(x)|.$$

Squaring and summing over all $x \in S$ this yields

$$\sum_{x \in S} \left|\sum_{y \in V} (p - A_{xy}) f(y) - p \sum_{y \in V} f(y)\right|^2 = |\lambda|^2 \cdot \|f|_S\|_2^2. \tag{1}$$

Let us define $M = pJ - A$, where $J$ is the $n \times n$ all ones matrix, and let $B$ be the $|S| \times n$ sub-matrix of $M$ consisting of rows corresponding to vertices in $S$, then (1) implies that $\|Bf - p\bar{f}\mathbf{1}\|_2 = |\lambda| \cdot \|f|_S\|_2$, where $\bar{f} = \langle f, \mathbf{1} \rangle$.

Furthermore, if we decompose $B = [P \quad Q]$ where $P$ contains the columns corresponding to vertices in $S$, and $Q$ the others, then clearly $P \sim M^{\mathrm{sym}}_{|S|}(p)$ and $Q \sim M_{|S| \times (n - |S|)}(p)$. (It is useful to think of $P$ as corresponding to the subgraph of $G$ induced by $S$, whereas $Q$ corresponds to the bipartite graph corresponding to $S$ and its complement). Since $Bf = P(f|_S) + Q(f|_{\overline{S}})$, we can write

$$\|Q(f|_{\overline{S}}) - p\bar{f}\mathbf{1}\|_2 \leq \|Bf - p\bar{f}\mathbf{1}\|_2 + \|Pf|_S\|_2 = |\lambda| \cdot \|f|_S\|_2 + \|P(f|_S)\|_2 \tag{2}$$

The main point now is that, by Corollary 2.5, $Q$ (a random rectangular matrix with *all* entries being i.i.d. copies of $X_p$) is extremely unlikely to map $f|_{\overline{S}}$ near the 1-dimensional subspace spanned by $\mathbf{1}$. But we know from concentration results that if $f|_S$ has small 2-norm, then the RHS will be small. This leads to a contradiction.

Now, since $f$ is a non-first eigenvector, by Theorems 2.3 and 2.1, there exist constants $C = C(p), \beta' = \beta'(p) > 0$ such that

$$\Pr\left[|\lambda| + \|P\| \geq C\sqrt{n}\right] \leq \exp(-\beta' n). \tag{3}$$

If we assume that $|\lambda| + \|P\| \le C\sqrt{n}$ and also $\|f|_S\|_2 < \alpha \le \frac{1}{2}$, then $a = \|f|_{\overline{S}}\|_2 = \sqrt{1-\alpha^2} > \frac{1}{2}$. In this case, (2) implies that

$$\|Q(\tfrac{1}{a} f|_{\overline{S}}) - \tfrac{p}{a}\bar{f}\mathbf{1}\|_2 \le 2C\sqrt{n}\alpha. \tag{4}$$

Now let $k = n - |S|$. By Corollary 2.5, for any $\delta \ge \frac{4\varepsilon}{1-2\varepsilon} > 0$, there exists $\alpha, \beta > 0$ (with $\beta$ depending only on $p$) such that for $n$ large enough,

$$\Pr_{Q \sim M_{k(1+\delta)\times k}(p)} \left[ \exists v \in S^{k-1}, \exists c \in \mathbb{R} \text{ s.t. } \|Qv - c\mathbf{1}\|_2 \le 2C\alpha\sqrt{n} \right] \le \exp(-\beta n), \tag{5}$$

but this implies that (4) and $|\lambda| + \|P\| \le C\sqrt{n}$ occur with probability at most $\exp(-\beta n)$. Taking a union bound over (5) and (3), we see that for some $\alpha > 0$,

$$\Pr[\|f|_S\|_2 < \alpha] \le \exp(-\beta n) + \exp(-\beta' n),$$

where both $\beta, \beta' > 0$ depend only on $p$. This completes the proof. □

We record the following simple corollary.

**Corollary 3.2.** *For every $p \in (0,1)$ there exist $r = r(p) > 0$ and $\varepsilon = \varepsilon(p)$ with $0 < \varepsilon < \frac{1}{2}$ such that for almost all $G \sim G(n,p)$ and* every *subset $S \subseteq V(G)$ with $|S| \ge (\frac{1}{2} + \varepsilon)n$ we have $\|f|_S\|_2 \ge r$.*

*Proof.* We need to take a union bound over all subsets $S \subseteq [n]$ with $|S| \ge (\frac{1}{2} + \varepsilon)n$. But for every value $\beta = \beta(p)$ from Theorem 3.1, there exists an $\varepsilon < \frac{1}{2}$ such that the number of such subsets is $o(\exp(\beta n))$, hence the union bound applies. □

### 3.2 Bounding the number of exceptional vertices

We now turn to prove Theorem 1.2. Fix some $p \in (0,1)$ and consider $G(V,E) \sim G(n,p)$. We concentrate on the part of the theorem that concerns weak nodal domains; exactly the same proof works for strong domains after we delete the vertices at which an eigenvector vanishes (which do not contribute to the count of strong nodal domains).

Let $r = r(p)$ and $\varepsilon = \varepsilon(p)$ be chosen as in Corollary 3.2. In everything that follows, we assume that $n$ is sufficiently large. By Theorem 2.3, it almost surely holds that every non-first eigenvalue $\lambda$ of $G$ satisfies $|\lambda| = O(\sqrt{p(1-p)n})$. We thus can and will assume that this bound holds for the remainder of this section.

**Lemma 3.3.** *Almost surely every non-first eigenvector $f : V \to \mathbb{R}$ of $G$ satisfies $|\langle f, \mathbf{1}\rangle| \le O(1/\sqrt{p})$.*

*Proof.* As usual, $A$ is the adjacency matrix of $G$, $\lambda$ is the eigenvalue of $f$ and $J$ is the all-ones matrix. Let $M = pJ - A$. Then,

$$\lambda\langle f, \mathbf{1}\rangle = \langle f, A\mathbf{1}\rangle = \langle f, pJ\mathbf{1}\rangle - \langle f, M\mathbf{1}\rangle = pn\langle f, \mathbf{1}\rangle - \langle f, M\mathbf{1}\rangle.$$

It follows that

$$|\langle f, \mathbf{1}\rangle| = \frac{|\langle f, M\mathbf{1}\rangle|}{pn - \lambda} \le \frac{\|M\mathbf{1}\|_2}{pn - \lambda} \le \frac{\sqrt{n}\|M\|}{pn - \lambda}.$$

But since $M \sim M_n^{\mathrm{sym}}(p)$, we know that $\|M\| = O(\sqrt{p(1-p)n})$ almost surely. Therefore almost surely we have

$$|\langle f, \mathbf{1}\rangle| \le \frac{O(n\sqrt{(1-p)})}{pn - O(\sqrt{(1-p)n})} \le O\left(\frac{1}{\sqrt{p}}\right).$$

□

We now derive an almost sure bound on the number of weak nodal domains. To this end, for any function $f : V \to \mathbb{R}$, we let $\mathcal{P}_f$ and $\mathcal{N}_f$ be the largest non-negative and non-positive domains in $f$ (with respect to the random graph $G$). Define $\mathcal{E}_f = V \setminus (\mathcal{P}_f \cup \mathcal{N}_f)$, and let $\mathcal{Z}_f = \{v \in V : f(v) = 0\}$. In particular, observe that $\mathcal{E}_f \cap \mathcal{Z}_f = \emptyset$ since we are discussing weak domains. Although the next lemma is stated in terms of nodal domains, it should be clear that it is, in fact, a simple combinatorial observation about random graphs.

**Lemma 3.4.** *For any $f : V \to \mathbb{R}$, if $D_1, \ldots, D_m$ are the weak nodal domains in $\mathcal{E}_f$, then almost surely $m = O(p^{-1} \log n)$ and $|D_i| = O(p^{-1} \log n)$ for every $i \in [m]$.*

*Proof.* If $P_1, P_2, \ldots, P_s$ are non-negative nodal domains, then selecting one element from every $P_i$ yields an independent set of size $s$. By a standard fact (which follows immediately from a union bound) about graphs in $G(n,p)$, almost surely $s = O(p^{-1} \log n)$. Furthermore, since each $|P_i| \leq |\mathcal{P}_f|$ and there are no edges between the two sets, almost surely $|P_i| \leq O(p^{-1} \log n)$. The same holds for the non-positive nodal domains in $\mathcal{E}_f$. □

We are now ready to complete the proof of Theorem 1.2. We will use the following straightforward fact about $G(n,p)$.

**Fact 3.5.** *For any fixed $k \in \mathbb{N}$, the following holds almost surely over the choice $G \sim G(n,p)$. For any $x_1, x_2, \ldots, x_k \in V$, $|\Gamma(x_1) \cup \cdots \cup \Gamma(x_k)| = (1 - (1-p)^k \pm o(1))n$ and $|\Gamma(x_1) \cap \cdots \cap \Gamma(x_k)| = (p^k \pm o(1))n$.*

**Lemma 3.6.** *It almost surely holds that every non-first eigenvector satisfies $|\mathcal{E}_f| = O_p(1)$.*

*Proof.* Since we seek a result that holds asymptotically almost surely, we take the following liberty: Any property that holds for almost every graph in $G(n,p)$ and is needed in the proof is assumed to hold for $G$.

We can assume that $f$ has a constant sign on $\mathcal{E}_f$, for if $x, y \in \mathcal{E}_f$ satisfy $f(x) > 0 > f(y)$, then also $\Gamma(x) \cap \Gamma(y) \subseteq \mathcal{E}_f$. But from Lemma 3.4, $|\mathcal{E}_f| = O(\log^2 n)$, whereas $|\Gamma(x) \cap \Gamma(y)| = (p^2 + o(1))n$ by properties of random graphs.

Let $k$ be an integer such that $\frac{\frac{1}{2}+\varepsilon}{1-(1-p)^k} < 1$. Let $\{x_1, x_2, \ldots, x_k\} \subseteq \mathcal{E}_f$ and assume without loss of generality that $f(x_i) < 0$ for $i = 1, \ldots, k$. This implies that $|\mathcal{P}_f \cup \mathcal{E}_f| \geq (1 - (1-p)^k - o(1))n$ (because this set contains the union of the neighborhoods of $x_1, x_2, \ldots, x_k$). Therefore $|\mathcal{N}_f \setminus \mathcal{Z}_f| \leq ((1-p)^k + o(1))n$. By Lemma 3.3, we have

$$\sum_{x \in \mathcal{P}_f \cup \mathcal{E}_f} f(x) \leq O(1/\sqrt{p}) + \sum_{x \in \mathcal{N}_f \setminus \mathcal{Z}_f} |f(x)| \leq O(1/\sqrt{p}) + \sqrt{|\mathcal{N}_f \setminus \mathcal{Z}_f|} \sqrt{\sum_{x \in \mathcal{N}_f} f(x)^2} \leq \sqrt{((1-p)^k + o(1))n}.$$

By Markov's inequality, we know that there exists a subset $S \subseteq \mathcal{P}_f \cup \mathcal{E}_f$ such that $|S| \geq \frac{\frac{1}{2}+\varepsilon}{1-(1-p)^k} |\mathcal{P}_f \cup \mathcal{E}_f|$, and for every $y \in S$,

$$\begin{aligned} f(y) &\leq \frac{\sqrt{((1-p)^k + o(1))n}}{\left(1 - \frac{\frac{1}{2}+\varepsilon}{1-(1-p)^k}\right) |\mathcal{P}_f \cup \mathcal{E}_f|} \\ &\leq \frac{\left(1-(1-p)^k\right) \sqrt{((1-p)^k + o(1))n}}{(\frac{1}{2} - \varepsilon - (1-p)^k) |\mathcal{P}_f \cup \mathcal{E}_f|} \\ &\leq \frac{\sqrt{((1-p)^k + o(1))}}{(\frac{1}{2} - \varepsilon - (1-p)^k)\sqrt{n}}. \end{aligned}$$

Consequently,

$$\|f|_S\|_2 \leq \sqrt{|\mathcal{P}_f \cup \mathcal{E}_f|}\frac{\sqrt{((1-p)^k+o(1))}}{(\frac{1}{2}-\varepsilon-(1-p)^k)\sqrt{n}} \leq \frac{\sqrt{((1-p)^k+o(1))}}{\frac{1}{2}-\varepsilon-(1-p)^k}.$$

Corollary 3.2 yields that $\|f|_S\|_2 \geq r = r(p)$. It follows that

$$\frac{\sqrt{((1-p)^k+o(1))}}{\frac{1}{2}-\varepsilon-(1-p)^k} \geq r. \tag{6}$$

We conclude that $k = O\left(\frac{1}{p}\log\left(\frac{1}{r(\frac{1}{2}-\varepsilon)}\right)\right)$, which finishes the proof since $r, \varepsilon$ can be chosen depending only on $p$. □

### 3.3 Future directions

First, we suspect that the following question should not be too difficult to resolve.

**Conjecture 1.** *For every fixed $p \in (0,1)$, for $G \sim G(n,p)$, almost surely every eigenfunction $f$ of $G$ satisfies*

$$\{v \in V(G) : f(v) = 0\} = \emptyset.$$

For simplicity in what follows, we only discuss weak nodal domains. As before, $\mathcal{P}_f$ and $\mathcal{N}_f$ are the largest non-negative and non-positive weak domains, and $\mathcal{E}_f = V \setminus (\mathcal{P}_f \cup \mathcal{N}_f)$ is the set of exceptional vertices. We observe that if sufficiently good lower bounds on $|\mathcal{N}_f|$ and $|\mathcal{P}_f|$ hold, then the number of exceptional vertices is at most one. We illustrate this for $p = \frac{1}{2}$, but the extension to general $p$ is straightforward.

**Lemma 3.7.** *Suppose that there exists an $\epsilon_0 > 0$ such that almost surely, every non-first eigenvector $f$ of $G \sim G(n, \frac{1}{2})$, satisfies $|\mathcal{P}_f|, |\mathcal{N}_f| \geq (\frac{1}{4}+\epsilon_0)n$. Then almost surely every eigenvector has at most one exceptional vertex.*

*Proof.* Almost surely, every pair of vertices $x, y \in V(G)$, satisfies $|\Gamma(x) \cup \Gamma(y)| \geq (\frac{3}{4} - o(1))n$. Assuming the stated lower bound on $|\mathcal{P}_f|$ and $|\mathcal{N}_f|$, there must be an edge from $\{x, y\}$ to both $\mathcal{P}_f$ and $\mathcal{N}_f$. But if $x$ and $y$ are both exceptional, then $f$ must have the same sign on $x$ and $y$ (see, e.g. the proof of Lemma 3.6), implying that there cannot be both types of edges. □

Next, we show that if one can obtain a slightly non-trivial upper bound on $\|f\|_\infty$ for eigenvectors $f$ of $G(n,p)$, then almost surely all eigenvectors of such graphs have at most $O(\frac{1}{p})$ exceptional vertices, and e.g. at most one exceptional vertex for $p \in [0.21, 0.5]$. First, we pose the following natural problem.

**Question 2.** *Is it true that, almost surely, every eigenvector $f$ of $G(n,p)$ has $\|f\|_\infty = o(1)$? The natural guess would be that for almost all graphs, every eigenfunction satisfies $\|f\|_\infty = n^{-\frac{1}{2}+o(1)}$.*

Very recently we learned that Tao and Vu [22, Prop. 58] have made significant progress on this question by showing that it holds for all but $o(n)$ of the eigenvectors. The preceding question, though, is still open. A positive answer yields more precise control on the nodal domains of $G(n,p)$.

**Theorem 3.8.** *Suppose that for almost every $G \sim G(n,p)$ it holds that all eigenvectors $f$ of $G$ satisfy $\|f\|_\infty = o(1)$. Then almost surely every eigenvector has at most $k_p = \lfloor \frac{1}{\log_2(1/(1-p))} \rfloor$ exceptional vertices.*

In order to prove this statement, we need the following strengthening of Theorem 3.1. In particular, we require that the subset of vertices $S$ be allowed to (weakly) depend on the random choice of $G \sim G(n,p)$.

**Theorem 3.9.** *For every $p \in (0,1)$, and $\varepsilon > 0$, there exist values $\alpha = \alpha(\varepsilon, p) > 0$ and $\beta = \beta(p) > 0$ such that the following holds. Suppose $G \sim G(n,p)$ and $A$ is the adjacency matrix of $G$. Suppose further that $S \subseteq [n]$ is a (possibly) random subset which is allowed to depend on the rows of $A$ indexed by a set $T \subseteq [n]$ with $|T| = o(n)$. Then for all sufficiently large $n$, we have*

$$\Pr\left[\exists \text{a non-first eigenvector } f \text{ of } G \text{ with } \|f|_S\|_2 < \alpha \|f|_{V\setminus T}\|_2 \text{ and } |S| \geq (\tfrac{1}{2}+\varepsilon)n\right] \leq \exp(-\beta n).$$

*Proof.* Let $A$ be the adjacency matrix of $G = (V, E)$, and let $f : V \to \mathbb{R}$ be a non-first eigenvector of $G$ with eigenvalue $\lambda$. Let $S \subseteq V$ be a possibly random subset with $|S| \geq (\frac{1}{2}+\varepsilon)n$. Let $T \subseteq [n]$ be such that $|T| \leq o(n)$ and $S$ is determined after conditioning on the values of $A$ in the rows indexed by $T$. Assume, furthermore, that $\|f|_S\|_2 < \alpha \|f|_{V\setminus T}\|_2$.

Again, for every $x \in S$, the eigenvalue condition $\lambda f(x) = f(\Gamma(x))$ implies that

$$\left| \sum_{y \in V\setminus T} A_{xy} f(y) + \sum_{y \in T} A_{xy} f(y) \right| = |\lambda f(x)|.$$

Or equivalently,

$$\left| \sum_{y \in V\setminus T} (p - A_{xy}) f(y) - p \sum_{y \in V\setminus T} f(y) + \sum_{y \in T} A_{xy} f(y) \right| = |\lambda f(x)|.$$

Squaring and summing over all $x \in S \setminus T$ yields

$$\sum_{x \in S\setminus T} \left( \sum_{y \in V\setminus T} (p - A_{xy}) f(y) - p \sum_{y \in V\setminus T} f(y) + \sum_{y \in T} A_{xy} f(y) \right)^2 = |\lambda|^2 \cdot \|f|_{S\setminus T}\|_2^2 \tag{7}$$

As above we define $M = pJ - A$, but now we let $B$ be the $|S\setminus T| \times |V\setminus T|$ sub-matrix of $M$ consisting of rows corresponding to vertices in $S \setminus T$ and columns corresponding to vertices in $V \setminus T$. Now let $g = \frac{f|_{V\setminus T}}{\|f|_{V\setminus T}\|_2}$, then (7) implies that

$$\|Bg - c_f w_T\|_2 \leq 2|\lambda| \cdot \alpha$$

where $w_T \in \mathbb{R}^{|S\setminus T|}$ is a vector which depends only on the rows of $A$ indexed by $T$ and $c_f \in \mathbb{R}$ is some constant depending on $f$. Note that we have used the fact that $\|f|_S\|_2 \leq \alpha \|f|_{V\setminus T}\|_2$. Furthermore, $B$ and $S$ are independent random variables conditioned on $T$. From this point, the proof proceeds just as in Theorem 3.1. □

*Proof of Theorem 3.8.* Our goal is to show that almost surely $|\mathcal{E}_f| \leq k_p$ for every non-first eigenvector $f$ with associated eigenvalue $\lambda$. Suppose, to the contrary, that $|\mathcal{E}_f| > k_p$, and let $U \subseteq \mathcal{E}_f$ have $|U| = k_p + 1$. Consider $\Gamma = \bigcup_{u \in U} \Gamma(u)$. By properties of random graphs, it holds that

$$|\Gamma| \geq [1 - (1-p)^{k_p+1} - o(1)]n \geq \left(\frac{1}{2} + \varepsilon_p - o(1)\right) n,$$

for some positive $\varepsilon_p$. Thus for $n$ large enough, we may assume that indeed $|\Gamma| \geq \left(\frac{1}{2} + \varepsilon_p\right) n$.

Again, by Theorem 2.3, we have $|\lambda| = O_p(\sqrt{n})$ almost surely, and thus we assume this bound holds for the remainder of the proof. Now for each $u \in U$, let $D_u$ be the nodal domain of $u$ with respect to $f$. Using the eigenvalue condition, for every $u \in U$, we have $|f(\Gamma(u))| = |\lambda f(u)| = O_p(\sqrt{n})|f(u)|$. Every neighborhood $\Gamma(u)$ has non-trivial intersection with only one of $\mathcal{P}_f$ or $\mathcal{N}_f$, hence grouping terms by sign, we have

$$\begin{aligned}
\sum_{x \in \Gamma(u)} |f(x)| &\leq \left| f\left(\Gamma(u) \cap (\mathcal{P}_f \cup \mathcal{N}_f)\right) + f\left(\Gamma(u) \cap \mathcal{E}_f \setminus D_u\right)\right| + |f(D_u)| \\
&\leq |f(\Gamma(u))| + |D_u| \\
&\leq O_p(\sqrt{n})|f(u)| + O(\log n),
\end{aligned}$$

where we have used the estimate on $|D_u|$ from Lemma 3.4 which holds almost surely.

In particular,

$$\sum_{u \in U} \sum_{x \in \Gamma(u)} |f(x)| \leq (k_p + 1)\|f\|_\infty \cdot O_p(\sqrt{n}) + O(k_p \log n). \tag{8}$$

Now let $\Gamma' = \{x \in \Gamma : |f(x)| \leq \frac{c}{\sqrt{n}}\}$ for some $c = c(n) > 0$ to be chosen momentarily. Using (8), we see that $|\Gamma \setminus \Gamma'| \leq \frac{O(k_p\sqrt{n})|\lambda| \cdot \|f\|_\infty}{c}$. Under the assumption $|\lambda| \cdot \|f\|_\infty = o(\sqrt{n})$, we can choose $c = o(1)$ so that $|\Gamma \setminus \Gamma'| = o(n)$, in which case we may assume that for $n$ large enough, $|\Gamma'| \geq \left(\frac{1}{2} + \varepsilon_p'\right) n$ for some $\varepsilon_p' > 0$. We have $\|f|_{\Gamma'}\|_2 = o(1)$, hence also

$$\|f|_{\Gamma'}\|_2 = o(1) \cdot \|f|_{V \setminus U}\|_2, \tag{9}$$

since $\|f|_U\|_2 \leq \sqrt{|U|} \cdot \|f\|_\infty = o(1)$.

Since $\Gamma'$ depends on $f$ and not just on the rows of the adjacency matrix corresponding to $U$, we cannot directly apply Theorem 3.9. Instead, we will apply the theorem to a collection of sets $\mathcal{U}$ which are determined by $\Gamma$ (which *is* determined by the sets $\{\Gamma(u)\}_{u \in U}$), with the guarantee that $\Gamma' \in \mathcal{U}$.

To this end, let $y(n) = o(n)$ be an upper bound on the size of $|\Gamma \setminus \Gamma'|$, and consider

$$\mathcal{U} = \{W \subseteq \Gamma : |W| \geq |\Gamma| - y(n)\}.$$

Then, we have $|\mathcal{U}| \leq \binom{n}{y(n)}$, $\Gamma' \in \mathcal{U}$, and the collection $\mathcal{U}$ is completely determined by the rows of the adjacency matrix of $G$ corresponding to the vertices in $U$ (as it is determined by $\Gamma = \bigcup_{u \in U} \Gamma(u)$). We may enumerate $\mathcal{U} = \{U_1, U_2, \ldots\}$ in such a way that each $U_i$ is determined after conditioning on $\Gamma$ (simply by canonically ordering all subsets of the vertices, and taking the induced ordering on $\mathcal{U}$).

We can thus apply Theorem 3.9 to each of the $\binom{n}{y(n)}$ sets $U_i \in \mathcal{U}$ (one of which will always be the set $\Gamma'$) and take a union bound to obtain, for some $\beta = \beta(p) > 0$,

$$\Pr[\exists U \text{ s.t. } \|f|_{\Gamma'}\|_2 = o(1) \cdot \|f|_{V\setminus U}\|_2 \text{ and } |\Gamma'| \geq (\tfrac{1}{2} + \varepsilon'_p)n] \leq \binom{n}{k_p+1}\binom{n}{y(n)} \exp(-\beta n),$$

and the latter quantity is $o(1)$ since $y(n) = o(n)$, but this contradicts (9). □

**Remark 3.1.** *Observe that even under the preceding assumptions, we are not able to rule out the case of one exceptional vertex $v$ with, say $\Gamma(v) = \mathcal{N}_f$ and $\mathcal{P}_f = V \setminus (\mathcal{N}_f \cup \{v\})$.*

## APPENDIX

# A Properties of random matrices

We first recall that the random variables $X_p$ (cf. Definition 1.3) have a subgaussian tail.

**Lemma A.1.** *For every $0 \leq p \leq \frac{1}{2}$,*

$$\mathbb{E}e^{tX_p} \leq e^{(1-p)^2t^2/2} .$$

*For every $\frac{1}{2} < p \leq 1$,*

$$\mathbb{E}e^{tX_p} \leq e^{p^2t^2/2} .$$

*Proof.* Note that

$$\mathbb{E}e^{tX_p} = pe^{-t(1-p)} + (1-p)e^{tp}.$$

In both cases, the claim follows by comparing the Taylor series on both sides. □

*Proof of Theorem 2.2.* Let $N_1 \subseteq S^{k-1}$ and $N_2 \subseteq S^{m-1}$ be $\frac{1}{3}$-nets. By standard estimates, $|N_1| \leq 9^k, |N_2| \leq 9^m$.

Fix some $x \in N_1$ and $y \in N_2$, and estimate $\Pr\left(\langle y, Ax\rangle > t\sqrt{m}\right)$ as follows

$$\Pr\left(\langle y, Ax\rangle > t\sqrt{m}\right) = \Pr\left(e^{\lambda\langle y,Ax\rangle} > e^{\lambda t\sqrt{m}}\right) \leq \frac{\mathbb{E}e^{\lambda\langle y,Ax\rangle}}{e^{\lambda t\sqrt{m}}}$$

Define $q = \max\{p, 1-p\}$. By Lemma A.1 we have

$$\begin{aligned}\mathbb{E}e^{\lambda\langle y,Ax\rangle} &= \prod_{i=1}^{k}\prod_{j=1}^{m}\mathbb{E}e^{\lambda A_{ij}x_iy_j} \\ &\leq \prod_{i=1}^{k}\prod_{j=1}^{m}e^{q^2\lambda^2x_i^2y_j^2/2} = e^{q^2\lambda^2/2} .\end{aligned}$$

The optimal choice is $\lambda = \frac{t\sqrt{m}}{q^2}$ which yields

$$\Pr\left(\langle y, Ax\rangle > t\sqrt{m}\right) \leq e^{-\frac{t^2m}{2q^2}}$$

and therefore

$$\Pr\left(\exists x \in N_1, y \in N_2 : |\langle y, Ax\rangle| > t\sqrt{m}\right) \leq 2e^{-\frac{t^2 m}{2q^2}}|N_1||N_2| \leq 2e^{-\frac{t^2 m}{2q^2}}9^{k+m}.$$

By successive approximation, express $x \in S^{k-1}$ as $x = \sum_{i\geq 0}\alpha_i x_i$ where for all $i$, $x_i \in N_1$ and $|\alpha_i| \leq 3^{-i}$. Likewise for $y \in S^{m-1}$ and $y = \sum_{i\geq 0}\beta_i y_i$ with $y_i \in N_2$ and $|\beta_i| \leq 3^{-i}$. Therefore, if $|\langle y_i, Ax_j\rangle| \leq t\sqrt{m}$ for all $x_j \in N_1$ and $y_i \in N_2$, then

$$|\langle y, Ax\rangle| \leq \sum_{i,j\geq 0} 3^{-i-j}|\langle y_i, Ax_i\rangle| \leq 3t\sqrt{m},$$

which means that

$$\Pr\left(\|A\| \geq 3t\sqrt{m}\right) \leq 9^{m+k}e^{-\frac{t^2 m}{2q^2}}.$$

Now select $t$ to satisfy $\frac{2+\delta}{1+\delta}\ln 9 < \frac{t^2}{2q^2}$, where $m = (1+\delta)k$. Concretely, let

$$t = 2q\sqrt{2\cdot\frac{2+\delta}{1+\delta}\ln 9}.$$

With this choice, we can take

$$a_1 = \frac{18q}{4}\sqrt{2\cdot\frac{2+\delta}{1+\delta}\ln 9}$$
$$a_2 = 3\frac{2+\delta}{1+\delta}\ln 9.$$

□

## A.1 Quantitative bounds

First, we recall the Payley-Zygmund inequality [15].

**Lemma A.2.** *For any positive random variable $Z$ and any $0 \leq \lambda \leq 1$*

$$\Pr\left(Z \geq \lambda\mathbb{E}(Z)\right) \geq (1-\lambda)^2\frac{(\mathbb{E}Z)^2}{\mathbb{E}(Z^2)}.$$

We can now prove (ii), specialized to the random variables $X_p$.

**Lemma A.3.** *Let $Y_1, \ldots, Y_n$ be independent copies of $X_p$ and let $a_1, \ldots, a_n \in \mathbb{R}$ satisfy $\sum a_i^2 = 1$, then for every $s \in \mathbb{R}$ and every $0 \leq \eta \leq 1$*

$$\Pr\left(\left|\sum_{i=1}^n a_i Y_i - s\right| > \eta\sqrt{p(1-p)}\right) > C(1-\eta^2)^2$$

*where $C = \frac{p^2(1-p)^2}{128q^4}$ and $q = \max\{p, 1-p\}$.*

*Proof.* Let $X = \sum_{i=1}^n a_i Y_i - s$. By independence, one has

$$\mathbb{E}[X^2] = \sum_{i=1}^n a_i^2 \mathbb{E}(Y_i^2) + s^2 = p(1-p) + s^2.$$

Also, setting $Y = \sum_{i=1}^n a_i Y_i = X + s$, we have

$$\begin{aligned} \mathbb{E}[Y^4] &= 4\int_0^\infty t^3 \Pr\left(\left|\sum_{i=1}^n a_i Y_i\right| > t\right) dt \\ &\leq 8\int_0^\infty t^3 e^{-t^2/2q^2}\, dt = 16q^4. \end{aligned}$$

This inequality follows from the subgaussian tail bound in Lemma A.1. For $\Pr\left(\sum a_i Y_i > t\right)$ we use the result for $X_p$ and for $\Pr\left(-\sum a_i Y_i > t\right)$ we use the result for $-X_p = X_{1-p}$. Here it is convenient to reformulate Lemma A.1 by saying that $\mathbb{E}e^{tX_p} \leq e^{q^2t^2/2}$ for all $0 \leq p \leq 1$.

Using $(a-b)^4 \leq 8(a^4+b^4)$, we have

$$\mathbb{E}[X^4] \leq 8(\mathbb{E}[Y^4] + s^4) \leq 128q^4 + 8s^4.$$

We can now apply Lemma A.2 with $Z = X^2$ and conclude that

$$\Pr\left(X^2 \geq \lambda(p(1-p)+s^2)\right) \geq (1-\lambda)^2 \frac{(p(1-p)+s^2)^2}{128q^4 + 8s^4}$$

This last expression has a single local minimum at $s = 0$ and tends to $1/8$ as $s \to \infty$. Its value when $s = 0$ is $\frac{p^2(1-p)^2}{128q^4}$ $(= C)$ which is at most $\frac{1}{128} < \frac{1}{8}$.

Thus we have

$$\Pr\left(X^2 \geq \lambda p(1-p)\right) \geq \Pr\left(X^2 \geq \lambda(p(1-p)+s^2)\right) \geq C(1-\lambda)^2$$

Letting $\eta = \sqrt{\lambda}$ gives us the desired result. □

We now present a proof of Theorem 2.4 which only works for certain values of $p, \delta$ in order to calculate explicit upper bounds in the appendix.

*Proof of Theorem 2.4 holding only for certain values of $p, \delta$.* We are seeking an upper bound for

$$\Pr\left(\exists x \in S^{k-1} \text{ s.t. } \|Ax - w\|_2 \leq \alpha\sqrt{m}\right)$$

for $A \sim M_{m\times k}(p)$. We separately estimate this probability depending on whether $\|A\| > a_1\sqrt{m}$ or not, where $a_1$ is the constant from Theorem 2.2. An upper bound on the probability that $\|A\| > a_1\sqrt{m}$ is given in Thm. 2.2. For the complementary case we use an $\gamma$-net $N$ on $S^{k-1}$. Let $x \in N$ and $w \in \mathbb{R}^m$. Let $f_i = |\sum_{j=1}^k A_{ij}x_j - w_i|$ for every $i \leq m$. By Lemma A.3 we have

$$\Pr\left(f_i > \eta\right) > C\left(1 - \tfrac{\eta^2}{p(1-p)}\right)^2 .$$

Define $b = C\left(1 - \frac{\eta^2}{p(1-p)}\right)^2$.

Let us bound the probability that $\|Ax - w\|^2 \leq c^2 m$ for some constant $c$ to be chosen later.

$$\begin{aligned}
\Pr\left(\|Ax-w\|^2 \leq c^2 m\right) &= \Pr\left(\sum_{i=1}^{m} f_i^2 \leq c^2 m\right) \\
&= \Pr\left(m - \frac{1}{c^2}\sum_{i=1}^{m} f_i^2 \geq 0\right) \\
&= \Pr\left(\exp\left(\tau m - \frac{\tau}{c^2}\sum_{i=1}^{m} f_i^2\right) \geq 1\right) \\
&\leq \mathbb{E}\left(\exp\left(\tau m - \frac{\tau}{c^2}\sum_{i=1}^{m} f_i^2\right)\right) = e^{\tau m}\prod_{i=1}^{m}\mathbb{E}\left(\exp\left(-\frac{\tau f_i^2}{c^2}\right)\right),
\end{aligned}$$

for every $\tau > 0$.

For every $i \leq m$

$$\begin{aligned}
\mathbb{E}\left(\exp\left(-\frac{\tau f_i^2}{c^2}\right)\right) &= \int_0^1 \Pr\left(\exp\left(-\frac{\tau f_i^2}{c^2}\right) > t\right) dt \\
&\leq \int_0^{e^{-\tau\eta^2/c^2}} dt + \int_{e^{-\tau\eta^2/c^2}}^1 (1-b)dt \\
&= e^{-\tau\eta^2/c^2} + (1-b)\left(1 - e^{-\tau\eta^2/c^2}\right) = 1 - b\left(1 - e^{-\tau\eta^2/c^2}\right)
\end{aligned}$$

so we have

$$\Pr\left(\|Ax-w\|^2 \leq c^2 m\right) \leq e^{\tau m}\left(1 - b\left(1 - e^{-\tau\eta^2/c^2}\right)\right)^m \leq e^{\tau m - b\left(1-e^{-\tau\eta^2/c^2}\right)m}$$

This expression is minimized for

$$\tau = \frac{c^2}{\eta^2}\ln\frac{b\eta^2}{c^2}\ .$$

For $\tau$ to be positive (as it should), we must have $c < \sqrt{b}\eta = \sqrt{C}\left(1 - \frac{\eta^2}{p(1-p)}\right)\eta$. Letting $\eta = \sqrt{\frac{p(1-p)}{3}}$ gives

$$b = \frac{4C}{9}\ .$$

Let

$$c = (1-\theta)\frac{2}{3}\sqrt{\frac{p(1-p)C}{3}}$$

for some $0 < \theta \leq 1$, which gives

$$\tau = \frac{4}{9}(1-\theta)^2 C \ln\frac{1}{(1-\theta)^2}\ .$$

Using these values for $\eta, b, \tau$ and $c$ gives

$$\Pr\left(\|Ax-w\|^2 \leq c^2 m\right) \leq \exp\left(-\frac{4C}{9}\left[1 + (1-\theta)^2\left(2\ln(1-\theta) - 1\right)\right]m\right)$$

The size of an $\gamma$-net on $S^{k-1}$ is at most $(1+2/\gamma)^k \leq \left(\frac{3}{\gamma}\right)^k$ assuming $\gamma \leq 1$. Repeating this argument for every $x \in N$ and using the union bound we get that

$$\begin{aligned}\Pr\left(\exists x \in N \text{ s.t. } \|Ax-w\|_2 \leq c\sqrt{m}\right) &\leq \exp\left(-\frac{4C}{9}\left[1+(1-\theta)^2\left(2\ln(1-\theta)-1\right)\right]m + k\ln\frac{3}{\gamma}\right) \\ &= \exp\left(-\left\{\frac{4C}{9}\left[1+(1-\theta)^2\left(2\ln(1-\theta)-1\right)\right] - \frac{\ln\frac{3}{\gamma}}{1+\delta}\right\}m\right)\end{aligned}$$

Now we seek a constant $\alpha$ such that if (i) $\|A\| \leq a_1\sqrt{m}$, and (ii) there exists $x \in S^{k-1}$ with $\|Ax-w\|_2 \leq \alpha\sqrt{m}$, then $\exists x' \in N$ with $\|Ax'-w\|_2 \leq c\sqrt{m}$. If $x' \in N$ is chosen so that $\|x-x'\|_2 \leq \gamma$, then

$$\|Ax'-w\|_2 = \|Ax-w+A(x'-x)\|_2 \leq \|Ax-w\|_2 + \|A(x'-x)\|_2 \leq \alpha\sqrt{m} + \gamma a_1\sqrt{m}\ .$$

By Theorem 2.2 we have $\Pr(\|A\| > a_1\sqrt{m}) \leq e^{-a_2 m}$, so taking

$$\alpha = (1-\theta)\frac{2}{3}\sqrt{\frac{p(1-p)C}{3}} - \gamma a_1 \tag{10}$$

and

$$\beta = \min\left(a_2, \frac{4C}{9}\left[1+(1-\theta)^2\left(2\ln(1-\theta)-1\right)\right] - \frac{\ln\frac{3}{\gamma}}{1+\delta}\right) \tag{11}$$

gives us the desired result. □

This preceding argument yields the conclusion of Theorem 2.4 as long as there exists a choice of $\gamma, \theta \in (0,1)$ such that both $\alpha$ and $\beta$ are positive. These bounds are used in Appendix A to show that our techniques, if analyzed in gory detail, yield reasonable bounds on the number of nodal domains for various values of $p$.

# B A few examples

We used a simple MATLAB program to compute an upper bound on the constant behind the $O_p(1)$ term in Lemma 3.6 for various $p$'s. It calculates the largest $k$ such that inequality (6) holds, using $r = \frac{\alpha\sqrt{\frac{1}{2}+\varepsilon}}{2D}$, and the explicit bounds computed in Theorem 2.2 and (10) and (11). Our bounds for various $p$'s are given in the following table.

| $p$ | $k$ |
|---|---|
| 0.78 | 29 |
| 0.74 | 30 |
| 0.7 | 32 |
| 0.66 | 34 |
| 0.62 | 37 |
| 0.58 | 39 |
| 0.54 | 43 |
| 0.5 | 46 |
| 0.46 | 54 |
| 0.42 | 63 |
| 0.38 | 75 |
| 0.34 | 90 |
| 0.3 | 109 |
| 0.26 | 137 |
| 0.22 | 181 |
| 0.18 | 277 |

## References


[1] Noga Alon, Michael Krivelevich, and Van H. Vu. On the concetration of eigenvalues of random symmetric matrices. *Israel J. Math.*, 131:259–267, 2002.

[2] Ram Band, Idan Oren, and Uzy Smilansky. Nodal domains on graphs—how to count them and why? In *Analysis on graphs and its applications*, volume 77 of *Proc. Sympos. Pure Math.*, pages 5–27. Amer. Math. Soc., Providence, RI, 2008.

[3] Marcel Berger. *A Panoramic View of Riemannian Geometry*. Springer-Verlag, 2003.

[4] Isaac Chavel. *Eigenvalues in Riemannian Geometry*. Academic Press, 1984.

[5] R. R. Coifman, S. Lafon, A. B. Lee, M. Maggioni, B. Nadler, F. Warner, and S. Zucker. Geometric Diffusion as a tool for harmonic analysis and structure definition of data, part i: Diffusion maps. *Proc. Natl. Acad. Sci.*, 102(21):7426–7431, 2005.

[6] R. R. Coifman, S. Lafon, A. B. Lee, M. Maggioni, B. Nadler, F. Warner, and S. Zucker. Geometric Diffusion as a tool for harmonic analysis and structure definition of data, part i: Multiscale methods. *Proc. Natl. Acad. Sci.*, 102(21):7432–7437, 2005.

[7] Kevin P. Costello, Terence Tao, and Van Vu. Random symmetric matrices are almost surely nonsingular. *Duke Math. J.*, 135(2):395–413, 2006.

[8] E. Brian Davies, Graham M. L. Gladwell, Josef Leydold, and Peter F. Stadler. Discrete nodal domain theorems. *Linear Algebra and its Applications*, 336(1–3):51–60, 2001.

[9] Zoltán Füredi and János Komlós. The eigenvalues of random symmetric matrices. *Combinatorica*, 1(3):233–241, 1981.

[10] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bull. AMS*, 43(4):439–561, October 2006.

[11] Michael Krivelevich and Van H. Vu. Approximating the independence number and the chromatic number in expected polynomial time. *J. Comb. Optim.*, 6(2):143–155, 2002.

[12] A. E. Litvak, A. Pajor, M. Rudelson, and N. Tomczak-Jaegermann. Smallest singular value of random matrices and geometry of random polytopes. *Adv. Math.*, 195(2):491–523, 2005.

[13] A. E. Litvak, A. Pajor, M. Rudelson, N. Tomczak-Jaegermann, and R. Vershynin. Euclidean embeddings in spaces of finite volume ratio via random matrices. *J. Reine Angew. Math.*, 589:1–19, 2005.

[14] L. Lovász. *Combinatorial Problems and Exercises*. Elsevier, 1979.

[15] R. Paley and A. Zygmund. A note on analytic functions in the unit circle. *Proc. Camb. Phil. Soc.*, 28:266–272, 1932.

[16] Alex Pothen, Horst D. Simon, and Kan-Pu Liou. Partitioning sparse matrices with eigenvectors of graphs. *SIAM J. Matrix Anal. Appl.*, 11(3):430–452, 1990.

[17] M. Rudelson. Invertibility of random matrices: norm of the inverse. To appear, *Ann. Math.*, 2007.

[18] M. Rudelson and R. Vershynin. The Littlewood-Offord problem and invertibility of random matrices. *Adv. Math.*, 218:600–633, 2008.

[19] Jianbo Shi and Jitendra Malik. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):888–905, 2000.

[20] Daniel W. Stroock. *Probability theory, an analytic view*. Cambridge University Press, Cambridge, 1993.

[21] T. Tao and V. Vu. Inverse Littlewood-Offord theorems and the condition number of random discrete matrices. To appear, *Ann. Math.*, 2007.

[22] T. Tao and V. Vu. Random matrices: The distribution of the smallest singular values. *ArXiv e-prints*, 2009.

[23] V. Vu. Random discrete matrices. arXiv:math/0611321, 2007.

[24] Y. Weiss. Segmentation using eigenvectors: A unifying view. In *Proc. of the IEEE conf. on Computer Vision and Pattern Recognition*, pages 520–527, 1999.