

NODAL COUNT FOR ORTHOGONALLY INVARIANT ENSEMBLES

LIOR ALON, DAN MIKULINCER, AND JOHN URSCHER

ABSTRACT. We investigate the nodal count of eigenvectors of random matrices interpreted as operators on signed complete graphs. Our focus is on orthogonally invariant ensembles, with particular attention to the Gaussian Orthogonal Ensemble (GOE). We establish that, as the matrix size tends to infinity, the distribution of nodal counts converges to the same limiting law as the eigenvalue distribution. In the GOE case, this limit is the semicircle law. This result refutes a conjecture, motivated by quantum chaos and quantum graphs, which predicted Gaussian behavior of the nodal count.

1. INTRODUCTION

The purpose of this paper is to investigate the asymptotic nodal count sequence (size of the nodal sets of eigenvectors) of a random real symmetric $n \times n$ matrix as $n \rightarrow \infty$. The nodal count sequence of a matrix A is $\phi(A) = (\phi(A, k))_{k=1}^n$, with

$$\phi(A, k) := \left| \{i < j, : A_{ij} \varphi_i^{(k)} \varphi_j^{(k)} > 0\} \right|,$$

where $\varphi^{(k)}$ is the eigenvector of the k -th smallest eigenvalue of A . We consider ensembles of matrices that are invariant under orthogonal transformations, most notably, the Gaussian orthogonal ensemble (GOE). Our result for GOE matrices disproves a recent conjecture about the nodal count for randomly signed graphs [3]. While our focus is on the discrete setting—namely, nodal counts for matrices—our motivation stems from foundational results on the measure of nodal sets of eigenfunctions on manifolds, as well as the significant progress over the past two decades on the study of nodal sets of random waves.

The study of nodal patterns dates back to the 17th century when it was popularized by Ernst Chladni, who revealed intricate sand patterns on vibrating plates. Over time, mathematicians developed a rigorous theory to understand the interplay between order and chaos in these patterns, paving the way for spectral geometry. The first rigorous study of nodal sets was by Sturm in 1836 [28], who showed that the k -th eigenfunction of a Sturm–Liouville operator on an interval has exactly $k - 1$ zeros. In higher dimensions, nodal sets—the zero sets of eigenfunctions—exhibit much greater complexity. This naturally leads to questions about their geometric properties and how these depend on the underlying domain or manifold.

A fundamental question in spectral geometry regarding the size of nodal sets was asked by S.T. Yau [29]. For a smooth compact manifold M , let λ_k denote the k -th smallest eigenvalue of the Laplacian Δ_M and $\phi(\Delta_M, k)$ denote the measure of the nodal set of its eigenfunction. Yau’s conjecture states that $c_1 \sqrt{\lambda_k} < \phi(\Delta_M, k) < c_2 \sqrt{\lambda_k}$ for some positive constants c_1, c_2 . The proof for real analytic manifolds was given by Donnelly and Fefferman [14]. For smooth manifolds, Logunov and Malinnikova proved the lower bound and gave a polynomial upper bound [21, 20, 22].

DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MA, 02139 USA.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF WASHINGTON, SEATTLE, WA 98109 USA.

E-mail addresses: lioralon@mit.edu, danmiku@uw.edu, urschel@mit.edu.

2020 *Mathematics Subject Classification.* Primary 05C50, 15A18, 15B52.

Key words and phrases. nodal count, orthogonally invariant ensemble, spectral graph theory.

To go beyond asymptotic bounds, additional structure is required. For instance, much more can be computed explicitly in the presence of Gaussian randomness. Berry [11] proposed that if a compact planar domain $\Omega \subset \mathbb{R}^2$ is chaotic (in terms of billiard dynamics), then the local behavior of a Dirichlet eigenfunction with eigenvalue λ can be modeled by a Gaussian random field F_λ satisfying $-\Delta F_\lambda = \lambda F_\lambda$ almost surely. He showed that the nodal set of F_λ in Ω has average length $C \cdot \text{area}(\Omega) \sqrt{\lambda}$ and variance of order $\text{area}(\Omega) \log(\lambda)$ [11]. Since then, nodal sets of random wave models and random eigenfunctions have been extensively studied, with notable results for Laplace eigenfunctions on the two-dimensional sphere [24, 25] and torus [27, 19]. It was long believed that in all such cases, the properly normalized nodal length has a limiting Gaussian distribution. However, recent remarkable non-universality results identified sequences of eigenvalues for which the limiting nodal length distribution exists but is non-Gaussian [23, 26]. Our results reveal a comparable breakdown of universality.

We now turn to our main objects of interest, the nodal count of a matrix, or equivalently, of a finite graph. Given a real symmetric $n \times n$ matrix A , we associate a weighted signed graph G with n vertices and an edge (i, j) for each pair $i < j$ with non-zero off-diagonal entry $A_{ij} \neq 0$. Every edge has an edge weight $|A_{ij}|$, and the edge sign is $-\text{sgn}(A_{ij})$. Thus, for example, if A is a graph Laplacian, then G is unweighted and unsigned. Sort the eigenvalues of a A in increasing order, $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$, and denote the k -th eigenvector by $\varphi^{(k)}$.

Definition 1.1. *The nodal count $\phi(A, k)$ is the number of edges (i, j) of G for which $\varphi^{(k)}$ changes sign with respect to the sign of the edge, namely $A_{ij} \varphi_i^{(k)} \varphi_j^{(k)} > 0$. Under the generic assumption that $\varphi_j^{(k)} \neq 0$ for all k, j , we write*

$$\phi(A, k) := \left| \{i < j, : A_{ij} \varphi_i^{(k)} \varphi_j^{(k)} > 0\} \right| = \frac{1}{2} \sum_{i < j} (1 + \text{sgn}(A_{ij} \varphi_i^{(k)} \varphi_j^{(k)})).$$

Fiedler was the first to introduce and study this count [15], who showed that an analogue of Sturm's theorem holds: if G is a tree, then $\phi(A, k) = k - 1$ for all k . Conversely, Band [8] showed that if $\phi(A, k) = (k - 1)$ for all k , then G must be a tree. As in the continuous setting, there are no such uniform estimates for general graphs. Instead, Berkolaiko [9] generalized Fiedler's result and showed that for general graphs the following inequality always holds:

$$0 \leq \phi(A, k) - (k - 1) \leq \beta,$$

where $\beta = \beta(G) \geq 0$ is the first Betti number of the graph, and $\beta(G) = 0$ if and only if G is a tree. In the presence of randomness, if G is unweighted, and φ is replaced by a random Gaussian vector, then the nodal count is binomial and converges to a Gaussian as $n \rightarrow \infty$ [7]. These heuristics, together with numerical simulations and analogous results for quantum graphs [16, 1], led to the common belief that the nodal surplus sequence $\sigma(A, k) := \phi(A, k) - (k - 1)$ for $k = 1, \dots, n$ should behave like a Gaussian centered at $\beta/2$ as $\beta \rightarrow \infty$. We introduce some notation for this matter:

Notations 1.2. *For $\mathbf{x} \in \mathbb{R}^n$, we define $\text{avg}(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i$, $\text{std}(\mathbf{x}) = \left(\frac{1}{n} \sum_{i=1}^n (\mathbf{x}_i - \text{avg}(\mathbf{x}))^2 \right)^{1/2}$, the empirical distribution of its coordinates*

$$\text{emp}(\mathbf{x}) := \frac{1}{n} \sum_{k=1}^n \delta_{\mathbf{x}_k},$$

and the normalized empirical measure

$$\text{emp}_{\mathcal{N}}(\mathbf{x}) := \text{emp} \left(\frac{\mathbf{x} - \text{avg}(\mathbf{x}) \mathbf{1}}{\text{std}(\mathbf{x})} \right).$$

Let $\phi(A) = (\phi(A, k))_{k=1}^n$ and $\sigma(A) = (\sigma(A, k))_{k=1}^n$. The common belief was that $\text{emp}_{\mathcal{N}}(\sigma(A))$ is very close to a Gaussian centered at $\beta/2$ with variance of order β , for large n and β . This was disproved

by two of the authors [5], as they proved the following sharp bounds

$$\frac{\beta}{n} \leq \text{avg}(\sigma(A)) \leq \beta - \frac{\beta}{n}.$$

Examples of matrices saturating the upper and lower bounds were given for any possible choice of n and β [5]. These bounds were based on Berkolaiko's nodal-magnetic theorem [10, 12], which states that the nodal surplus is a Morse index of the eigenvalue. More precisely, by assigning to each edge (i, j) an angle $\theta_{ij} = -\theta_{ji}$, the eigenvalues of the Hermitian matrix $(A_\theta)_{rs} = e^{i\theta_{rs}} A_{rs}$ are functions of θ , that have critical points when $\theta_{ij} \in \{0, \pi\}$ for all edges. At such a critical point, the Morse index of λ_k turns out to be precisely equal to the nodal surplus $\sigma(A_\theta, k)$. Note that such a critical point corresponds to a signing $e^{-i\theta_{ij}} \in \{-1, 1\}$ of the edges of the graphs. Moreover, since these signs can be taken to be random and independent from one another, common wisdom about Gaussian universality suggests a possible refinement to the above-mentioned Gaussian conjecture. Encouraged by numerical evidence, as well as theoretical results which we detail below, this conjecture was formalized in [3]. Specifically, the first author and Goresky conjectured that when a graph is randomly signed, namely A_θ with random $\theta_{ij} \in \{0, \pi\}$, the nodal surplus does converge to a Gaussian. To be precise, given a fixed real symmetric matrix A , a random signing $(A_{\text{sgn}})_{ij} = \pm A_{ij}$ is a symmetric matrix with $|(A_{\text{sgn}})_{ij}| = |(A)_{ij}|$, and with random signs drawn independently and uniformly from $\{-1, 1\}$. It was conjectured that [3, p.3]:

For any graph G and a generic matrix A supported on G ,

$$\text{emp}_{\mathcal{N}}(\sigma(A_{\text{sgn}})) \rightarrow N(0, 1), \quad \text{as } \beta(G) \rightarrow \infty. \quad (1.1)$$

As mentioned, this conjecture was confirmed for some specific cases. In [3] it is shown that if G is the complete graph and the diagonal entries of A are distinct and sufficiently large, then $\sigma(A_{\text{sgn}}, k) \sim \text{Bin}(\beta, \frac{1}{2})$ for all k , and in particular $\text{emp}(\sigma(A_{\text{sgn}})) \sim \text{Bin}(\beta, \frac{1}{2})$.

For sparser graphs, according to [4], if the simple cycles of the graph are pairwise disjoint, then similarly, $\sigma(A_{\text{sgn}}, k) \sim \text{Bin}(\beta, \frac{1}{2})$ for all k , and so $\text{emp}(\sigma(A_{\text{sgn}})) \sim \text{Bin}(\beta, \frac{1}{2})$, for any generic choice of A supported on G . In all mentioned cases the Gaussianity conjecture 1.1 holds as $\beta \rightarrow \infty$, and follows from standard probabilistic considerations about the binomial distribution.

In this work, we derive a non-universality result, akin to the continuous results of [23, 26], which refutes the Gaussianity conjecture 1.1. Specifically, in Corollary 1.5 we show that if A is a random $n \times n$ GOE matrix,

$$\lim_{n \rightarrow \infty} \text{emp}_{\mathcal{N}}(\sigma(A_{\text{sgn}})) = \rho_{\text{s.c}} \neq N(0, 1) \quad \text{almost surely,}$$

where $\rho_{\text{s.c}}$ is the normalized semicircle distribution. To facilitate this result, we will investigate the nodal count of random real symmetric matrices of size $n \times n$ for large n . As mentioned, our goal is to study GOE random matrices. However, our results hold more generally and apply to any orthogonally invariant ensemble. An ensemble of random real symmetric matrices of size $n \times n$ is called an *orthogonally invariant ensemble* if it is invariant under conjugation by orthogonal matrices. We defer the formal definitions to Section 2.2 below, but mention for now that a key point is that these ensembles are uniquely determined by the distribution of their eigenvalues. So, if $p(\mathbf{\Lambda})$ is a distribution over eigenvalues $\mathbf{\Lambda} = (\lambda_1, \dots, \lambda_n)$, we will denote the corresponding orthogonally invariant ensemble by $\text{OE}_n(p(\mathbf{\Lambda}))$.

Our main result states that, for any orthogonally invariant ensemble satisfying some very mild regularity assumptions, the empirical nodal count distribution converges to a properly normalized version of its eigenvalue distribution. Below, we use $f(n) = \tilde{O}(n^\alpha)$ to denote

$$|f(n)| < C_1 n^\alpha \log^{C_2}(n) \quad \text{for all } n > N_0 \quad \text{for some positive constants } N_0, C_1, C_2.$$

Theorem 1.3. *Let $A \sim \text{OE}_n(p(\mathbf{\Lambda}))$ satisfy Spectral Growth Bound 2.1. Then*

$$\mathbb{E}[\phi(A, k)] = \binom{n}{2} \left(\frac{1}{2} + \frac{\sqrt{2}}{\pi^{3/2}} \mathbb{E} \left[\frac{\lambda_k - \text{avg}(\mathbf{\Lambda})}{\text{std}(\mathbf{\Lambda})} \right] n^{-1/2} + \tilde{O} \left(n^{-3/2} \right) \right) \quad (1.2)$$

and

$$\text{Var} [\phi(A, k)] = \tilde{O}(n^{5/2}) \quad (1.3)$$

uniformly over $k \in [n]$.

At a technical level, the orthogonal invariance allows us to represent $A_{ij}\varphi_i^{(k)}\varphi_j^{(k)}$ as a low-degree polynomial over the orthogonal group, equipped with its Haar measure. However, the existence of the highly non-linear sgn function in the definition of $\phi(A, k)$ precludes the use of classical tools, such as the Weingarten calculus, for the calculation of $\mathbb{E}[\phi(A, k)]$ and $\text{Var}[\phi(A, k)]$. Instead, our approach, which could be of potential independent interest, goes by a reduction to integration over Gaussian variables. While such a reduction would typically require considering a Gaussian space with dimension depending on n , we show that through careful conditioning, it is enough to consider a bounded number of Gaussians. In practice, this reduction allows us to estimate correlations between quadratic forms on the orthogonal group with very high precision, see Proposition 3.2. As we shall explain, $\phi(A, k)$ can be represented in this form which leads to Theorem 1.3.

To better understand Theorem 1.3 consider the following normalization

$$\phi_{\mathcal{N}}(A, k) := \frac{\pi^{3/2}}{\sqrt{2}} \left(\frac{\phi(A, k)}{\binom{n}{2}} - \frac{1}{2} \right) n^{1/2}.$$

Note that $\phi(A, k)$ is invariant to scaling and translating the matrix $A \mapsto c_1 A - c_2 \mathbf{I}$ for $c_1 > 0$ and $c_2 \in \mathbb{R}$. This invariance is also reflected in the quantitative estimate appearing in Theorem 1.3, where we consider the normalized eigenvalue. Therefore, we can and always will assume that $\text{avg}(\mathbf{\Lambda}) = 0$ and $\text{std}(\mathbf{\Lambda}) = 1$. Under this normalization, Theorem 1.3 can be rephrased as

$$\mathbb{E}[\phi_{\mathcal{N}}(A, k)] = \mathbb{E}[\lambda_k] + \tilde{O}(n^{-1}), \quad \text{and} \quad \text{Var}[\phi_{\mathcal{N}}(A, k)] = \tilde{O}(n^{-1/2}),$$

which immediately affords the following corollary concerning the asymptotics of the limiting empirical distribution $\text{emp}_{\mathcal{N}}(\phi(A))$, and as a result, also $\text{emp}_{\mathcal{N}}(\sigma(A))$. Recall that $\sigma(A) = \phi(A) + O(n)$ by definition, and that a sequence μ_n of random probability measures is said to *converge almost surely* to a deterministic probability measure ρ if, for any compactly supported continuous function φ , the random variables $\int \varphi d\mu_n$ converge to the number $\int \varphi d\rho$ almost surely. We denote such convergence by $\mu_n \xrightarrow{\text{a.s.}} \rho$.

Corollary 1.4 (General convergence). *Under the assumptions of Theorem 1.3, suppose further that $\text{avg}(\mathbf{\Lambda}) = 0$ and $\text{std}(\mathbf{\Lambda}) = 1$, and that there exists some probability measure ρ on \mathbb{R} such that*

$$\text{emp}(\mathbf{\Lambda}) \xrightarrow{\text{a.s.}} \rho, \quad \text{as } n \rightarrow \infty.$$

Then, the normalized nodal count $\phi_{\mathcal{N}}(A) = \left(\frac{\pi^{3/2}}{\sqrt{2}} \left(\frac{\phi(A, k)}{\binom{n}{2}} - \frac{1}{2} \right) n^{1/2} \right)_{k \in [n]}$ converges as well:

$$\text{emp}(\phi_{\mathcal{N}}(A)) \xrightarrow{\text{a.s.}} \rho.$$

The normalized nodal surplus $\sigma_{\mathcal{N}}(A) = \left(\frac{\pi^{3/2}}{\sqrt{2}} \left(\frac{\sigma(A, k)}{\binom{n}{2}} - \frac{1}{2} \right) n^{1/2} \right)_{k \in [n]}$ satisfies the same convergence

$$\text{emp}(\sigma_{\mathcal{N}}(A)) \xrightarrow{\text{a.s.}} \rho.$$

Corollary 1.4 suggests a general machinery of constructing different matrices with different limiting distributions for $\sigma_{\mathcal{N}}(A)$. Specifically, we can choose an appropriate orthogonal ensemble, sample a matrix A from this ensemble, and compute its normalized nodal surplus $\sigma_{\mathcal{N}}(A)$. We demonstrate this procedure in Figure 1, where we have chosen an orthogonal ensemble with a tri-modal eigenvalue distribution. The figure contains histograms showing $\text{emp}(\phi_{\mathcal{N}}(A))$ and $\text{emp}(\mathbf{\Lambda})$, in this case. These

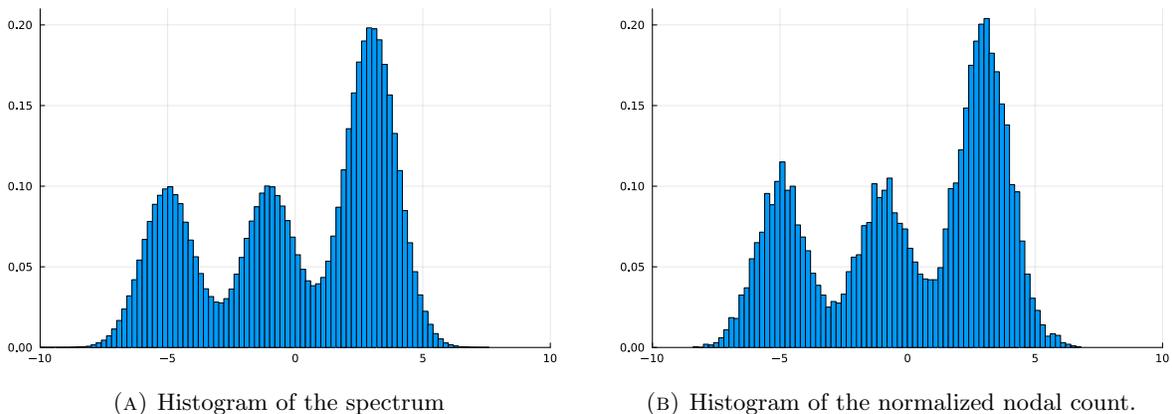


FIGURE 1. **Eigenvalue–nodal count agreement for a Gaussian mixture.** Numerical experiment for a *single* $n \times n$ random matrix $A = O^T \text{diag}(\Lambda)O$ with $n = 10^4$, where O is Haar orthogonal and the entries of Λ are sampled independently from the Gaussian mixture distribution $f(x) = \frac{1}{4\sqrt{2\pi}} \left(e^{-\frac{(x+5)^2}{2}} + e^{-\frac{(x+1)^2}{2}} + 2e^{-\frac{(x-3)^2}{2}} \right)$. We emphasize that this figure shows results for a single sampled matrix, not an average over many realizations.

histograms stand in sharp contrast to the usual unimodal histograms ubiquitous in the nodal count literature.

As mentioned, Corollary 1.4 is of particular interest when A is a GOE matrix. In that case, the symmetries of the Gaussian distribution imply that A_{sgn} has the same distribution as A , which leads to a refutation of (1.1).

Corollary 1.5 (GOE convergence). *Let A be a GOE_n matrix and let A_{sgn} be a random signing of A . Then A_{sgn} is also a GOE_n matrix, so*

$$\text{emp}_{\mathcal{N}}(\sigma(A_{\text{sgn}})) \xrightarrow{\text{a.s.}} \rho_{\text{s.c}} \neq N(0, 1),$$

where ρ_{sc} is the normalized semicircle distribution. In particular, Conjecture [3, pp. 1227] is false.

See Figure 3a, where the convergence $\text{emp}_{\mathcal{N}}(\sigma(A_{\text{sgn}})) \rightarrow \rho_{\text{s.c}}$ is shown in a numerical experiment.

CLT conjecture and further questions: Corollary 1.5 establishes that randomly signing the edges of the matrix is not enough to ensure universality of the nodal surplus count, and different randomly signed matrices will exhibit different behaviors.

To probe finer universality properties, we focus on the normalized fluctuations of the k -th nodal count. That is, let A be random $n \times n$ GOE matrix, and fix k . Consider the normalized nodal count

$$\overline{\phi(A, k)} := \frac{\phi(A, k) - \mathbb{E}[\phi(A, k)]}{\sqrt{\text{Var}(\phi(A, k))}}.$$

It is reasonable to expect that $\overline{\phi(A, k)}$ will be approximately Gaussian. As motivation, Theorem 3.2 (7) in [3] provides an explicit sufficient condition on a matrix A that ensures $\overline{\phi(A_{\text{sgn}}, k)}$ is approximately Gaussian: Given A a real symmetric $n \times n$ matrix, for any θ real antisymmetric matrix define the Hermitian matrix $(A_\theta)_{rs} = A_{rs}e^{i\theta_{rs}}$. If A is such that for every θ , all eigenvalues of A_θ are simple with nowhere-vanishing eigenvectors, then A has $\sigma(A_{\text{sgn}}, k) \sim \text{Bin}(\beta, \frac{1}{2})$ for all k .

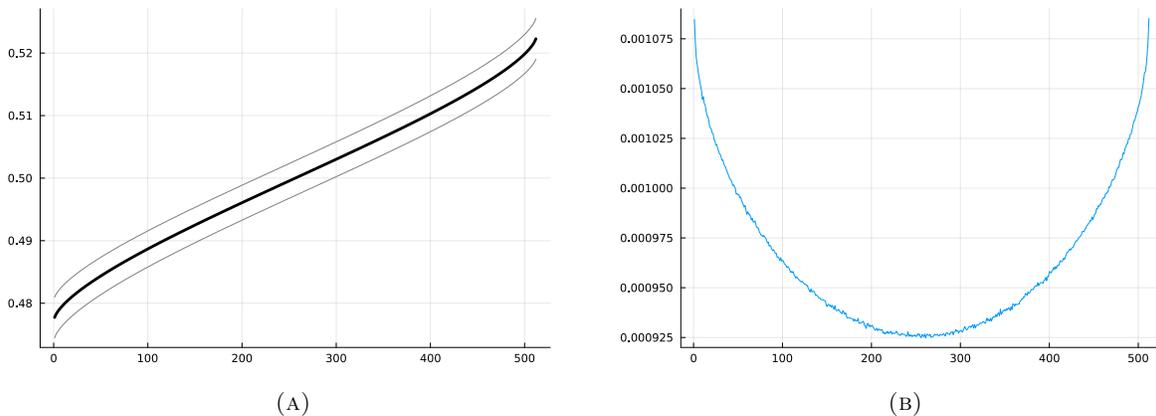


FIGURE 2. **Concentration of the nodal count for GOE matrices.** In this experiment we sampled 10^6 independent random GOE matrices of size $n = 2^9$. On the left, the thick curve shows the empirical mean $\mathbb{E}[\phi(A, k)/\binom{n}{2}]$, while the thin curves indicate ± 3 standard deviations. The mean takes values between 0.48 and 0.52, consistent with fluctuations of order $n^{-1/2}$ around $1/2$. On the right, the standard deviation $\text{std}[\phi(A, k)/\binom{n}{2}]$ is plotted as a function of k , taking values between 0.925×10^{-3} and 1.075×10^{-3} , of order n^{-1} , which means $\text{Var}[\phi(A, k)]$ is of order n^2 .

In particular, a standard argument on convergence of binomial to Gaussian allows us to conclude that any matrix A satisfying the assumption of Theorem 3.2 (7) in [3] has

$$d_{\text{Kol}}\left(\overline{\phi(A_{\text{sgn}}, k)}, N(0, 1)\right) = o(1),$$

as n grows, independently of k and the choice of A , and with d_{Kol} standing for Kolmogorov distance.

However, the k independence of $\sigma(A_{\text{sgn}}, k) \sim \text{Bin}(\beta, \frac{1}{2})$ implies $\sigma(A_{\text{sgn}}) \sim \text{Bin}(\beta, \frac{1}{2})$, which is not the case for a random GOE matrix, by Corollary 1.5. That is, a random GOE matrix A **does not** satisfy the assumption of Theorem 3.2 (7) in [3]. Nevertheless, recall that $\phi(A, k) = \frac{1}{2} \sum_{i < j} (1 + \text{sgn}(A_{ij}) \varphi_i^{(k)} \varphi_j^{(k)})$ is a sum of simple random variables. While these random variables are not independent, their joint law is invariant to certain permutations, and they form an exchangeable double array. That is, the law remains unchanged when either permuting the rows or the columns. We verify, for a random GOE matrix A , that $\text{dist}\left(\overline{\phi(A, k)}, N(0, 1)\right) \rightarrow 0$ numerically, as seen in Figure 3b.

Given the numerical results and the above intuition, we raise the following modest conjecture.

Conjecture 1.6 (CLT). *Let A be a random $n \times n$ matrix drawn from the Gaussian Orthogonal Ensemble. Then, for any $k \in [n]$,*

$$\text{dist}\left(\overline{\phi(A, k)}, N(0, 1)\right) = o(1),$$

as n grows, uniformly in k , and using any reasonable distance between random variables.

A bolder conjecture would involve a much larger class of signed matrices. In line with the original conjecture in (1.1), one could consider any sequences of matrices with diverging Betti numbers. The paper [2] raises such a question and contains further discussion.

Our proof falls short of addressing Conjecture 1.6. One reason is that our estimate $\text{Var}[\phi(A, k)] = \tilde{O}(n^{5/2})$ from Theorem 1.3 is not sharp. Numerical simulations suggest that the true growth rate is $\text{Var}[\phi(A, k)] = \tilde{O}(n^2)$ as seen in Figure 3c.

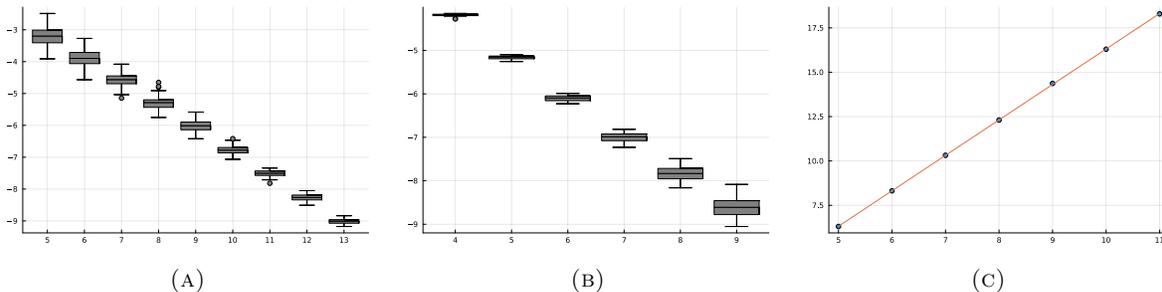


FIGURE 3. **Asymptotic statistics of the nodal count for GOE matrices.** (A) Log-log boxplot of the 1-Wasserstein distance between $\text{emp}(\phi_{\mathcal{N}}(A))$ and the semicircle distribution, as a function of $\log n$ for $n = 2^5, 2^6, \dots, 2^{13}$. Each box summarizes 100 independent samples of GOE_n matrices: the central line indicates the median, the box the interquartile range, whiskers extend to 1.5 times the interquartile range, and circles mark outliers. (B) Similar log-log boxplot of the Kolmogorov-Smirnov (KS) distance between $\phi(A, k)$ and a Gaussian distribution of the same mean and variance, for $n = 2^4, \dots, 2^9$. Each box summarizes the distribution of KS distances over all $k \in \{1, \dots, n\}$, computed from 10^6 independent GOE matrices. (C) Log-log plot of $\max_k \text{Var}[\phi(A, k)]$ versus $\log n$, estimated from 10^4 independent samples; the fitted slope 1.998 and intercept -3.670 suggest $\text{Var}[\phi(A, k)] = \tilde{O}(n^2)$.

A natural approach towards the conjecture would be to understand the moments, at least the lower order ones, of $\phi(A, k)$, which would require tightening our analysis. This is further emphasized in Figure 2b, which shows that there are lower-order dependencies of $\text{Var}[\phi(A, k)]$ with respect to k . Our coarse bounds are unable to capture this quadratic behavior. In light of this, we view a better understanding of this plot as a first step towards answering Conjecture 1.6. While, in principle, our approach is well-suited for making these finer estimates, it would make the proof much more technical. Thus, for the sake of readability and to reduce the technicality of this paper, we chose not to pursue this direction in the present work, and leave it as an interesting question for the future.

2. PRELIMINARIES AND PROOFS OF COROLLARIES

2.1. Some asymptotic notation. Throughout, we shall use the standard big O and Ω for asymptotic notation. Since we care about polynomial bounds, we shall use \tilde{O} to hide poly-logarithmic factors. Formally, $\tilde{O}(f(n))$ means $O(\log(n)^c f(n))$ for some $c > 0$. For vectors, we shall deviate slightly from the standard usage of the notation. Thus if ν is some vector, we shall write $\nu = \tilde{O}(f(n))$, when $\frac{\|\nu\|}{\sqrt{n}} = \tilde{O}(f(n))$. The main point is that with this definition, for *most* unit vectors θ ,

$$\langle \nu, \theta \rangle = \tilde{O}(f(n)). \tag{2.1}$$

To avoid possible confusion, we will recall this notation when using it. Similarly, if N is any random variable, we write $N = \tilde{O}(f(n))$ when $\sqrt{\mathbb{E}[N^2]} = \tilde{O}(f(n))$. In particular by the Cauchy-Schwartz inequality, if $N = \tilde{O}(f(n))$ and $N' = \tilde{O}(g(n))$, then $\mathbb{E}[N \cdot N'] = \tilde{O}(f(n)g(n))$.

2.2. Preliminaries on real orthogonal ensembles. Here, we shall give a brief overview of orthogonally invariant ensembles. In this work, our focus is on extracting meaningful statistics from an orthogonal ensemble when the eigenvalues of the ensemble are known. This is different from the traditional setting, where one wishes to understand the eigenvalues from the law of the random matrices.

Therefore, we shall introduce the orthogonally invariant ensembles in an ad hoc and somewhat nontraditional way. The reader is referred to [13] for a more comprehensive treatment, as well as references to all claims made in the following.

Let p be a probability distribution on \mathbb{R}^n . For $\mathbf{\Lambda} \sim p$ we order its entries by $\lambda_1 \leq \lambda_2, \dots \leq \lambda_n$. For $\Phi \sim \text{Haar}(\mathcal{O}(n))$, a random orthogonal matrix, independent from $\mathbf{\Lambda}$, we say that the random matrix $M = \Phi \text{diag}(\mathbf{\Lambda}) \Phi^T$ is drawn from the (real) orthogonal ensemble with eigenvalue distribution p , and denote it by $\text{OE}_n(p(\mathbf{\Lambda}))$. Equivalently, a symmetric M is drawn from some orthogonal ensemble if $\Psi M \Psi^*$ has the same distribution as M for any orthogonal matrix Ψ .

Perhaps the best-known orthogonal ensemble is the Gaussian Orthogonal Ensemble (GOE), in which up to symmetries M has independent Gaussian entries. Specifically, $M_{i,j} \sim N(0, \frac{1}{n})$ when $i \neq j \in [n]$ and $M_{i,i} \sim N(0, \frac{2}{n})$ for $i \in [n]$. In this case, the rotational invariance of the Gaussian immediately implies that M is drawn from some orthogonal ensemble. In fact, we can show that the eigenvalue distribution is explicitly given by

$$p_{\text{GOE}(n)}(\mathbf{\Lambda}) \propto \exp\left(-\frac{1}{4} \sum_{i=1}^n \lambda_i^2\right) \prod_{i < j} |\lambda_i - \lambda_j|. \quad (2.2)$$

A cornerstone result in random matrix theory states that the empirical eigenvalue distribution of the GOE stabilizes as $n \rightarrow \infty$. That is, if $\mathbf{\Lambda}_n \sim p_{\text{GOE}(n)}$, then

$$\text{emp}(\mathbf{\Lambda}_n) \xrightarrow[\text{a.s.}]{} \rho_{\text{s.c.}}, \quad (2.3)$$

where $\rho_{\text{s.c.}}$ is the semi-circle law, with explicit density $\frac{d\rho_{\text{s.c.}}}{dx}(x) = \frac{1}{2\pi} \sqrt{4 - x^2} \mathbf{1}_{\{|x| \leq 2\}}$.

Of course, there are many other examples of orthogonal ensembles. For example, if M has independent Gaussian entries, then MM^* is orthogonally invariant, and its empirical spectral distribution converges to the Marchenko-Pastur law. In general, the definition above allows the construction of many different examples, as we've demonstrated in Figure 1.

For our result, we will need to impose a mild regularity condition on the allowed eigenvalue distributions, which we now define.

Definition 2.1 (Spectral growth bound). *A random orthogonally invariant ensemble $\text{OE}_n(p(\mathbf{\Lambda}))$ is said to satisfy the spectral growth bound if*

$$\left\| \frac{\mathbf{\Lambda} - \text{avg}(\mathbf{\Lambda})}{\text{std}(\mathbf{\Lambda})} \right\|_{\infty} = \tilde{O}(1)$$

with probability at least $1 - \tilde{O}(n^{-3/2})$.

To expand a bit on Definition 2.1 and its uses, recall that if $M \sim \text{OE}_n(p(\mathbf{\Lambda}))$, then for any $k \geq 0$, $\text{Tr}(M^k) = \sum_{i=1}^n \lambda_i^k$. In that case, the condition dictates that once we normalize $\text{Tr}(M) = 0$ and $\text{Tr}(M^2) = n$, M must satisfy $\|M\|_{\text{op}} \leq C \log(n)^c$ and $\text{Tr}(M^k) \leq C \log(n)^{(k-2)c} n$, for some constants $C, c > 0$, with non-negligible probability. Below, we will use this consequence with $k = 4$.

For the GOE, it is straightforward to verify, using (2.2), that the condition is satisfied. A tighter analysis actually shows that we can choose $c = 1$, see [6, Chapter 3.1] for example. Similar estimates hold whenever the tails of the eigenvalue distribution are not too heavy.

2.3. Preliminaries on Gaussian concentration. Our proof of Theorem 1.3 will require us to estimate some integrals on the orthogonal group. As we shall demonstrate in the proof, it will be beneficial to reduce this calculation to Gaussian space. Towards that, we collect some useful results in this section.

The first one is a basic concentration inequality for low-degree polynomials in Gaussian variables.

Proposition 2.2 ([18, Theorem 6.7]). *Let p be a degree k polynomial and let G be a standard Gaussian vector. There exists a universal constant c_k such that*

$$\mathbb{P}\left(|p(X)| > t \mathbb{E}[|p(X)|^2]^{1/2}\right) \leq \exp\{-c_k t^{2/k}\} \quad \text{for all } t \geq 2.$$

From Proposition 2.2, we can deduce a general bound for quadratic forms in Gaussians. For that, if M is a matrix, we introduce the notation $\|M\|_{\max} := \max_{i,j} |M_{i,j}|$.

Lemma 2.3. *Let $\mathbf{x} \in \mathbb{R}^n$ be a fixed vector with $\|\mathbf{x}\|_2 \leq n^{1/2} \log^{c_1} n$, $\|\mathbf{x}\|_4 \leq n^{1/4} \log^{c_1} n$ and $G \in \mathbb{R}^{n \times p}$ be a standard Gaussian matrix, for some fixed constants c_1 and p . Then there exists a constant c_2 such that $\|G\|_{\max} \leq \log^{c_2} n$ and*

$$\|G^T \text{diag}(\mathbf{x})^k G - \mathbb{E}[G^T \text{diag}(\mathbf{x})^k G]\|_{\max} \leq \sqrt{n} \log^{c_2} n$$

for $k = 0, 1, 2$ with probability $1 - n^{-\log n}$.

Proof. By repeated application of Proposition 2.2, it suffices to show that for an arbitrary entry of $G^T \text{diag}(\mathbf{x})^k G$, its variance is at most $n \log^{c_3} n$ for some constant c_3 . This indeed follows immediately from the identity $\text{Var}[(G^T D G)_{ij}] = (1 + \delta_{ij}) \|D\|_F^2$. \square

As a typical example of Lemma 2.3, suppose that $\mathbf{\Lambda}$ satisfies the spectral growth condition, as in Definition 2.1, and is normalized as in Corollary 1.4. Then, if \mathbf{g} and $\hat{\mathbf{g}}$ are two independent standard Gaussians, we may consider G in Lemma 2.3 to have \mathbf{g} and $\hat{\mathbf{g}}$ as its two columns. In that case the Lemma implies the following bounds with high probability,

$$|\langle \mathbf{g}, \mathbf{g} \rangle - n|, \quad |\langle \mathbf{g}, \hat{\mathbf{g}} \rangle|, \quad |\langle \mathbf{g}, \text{diag}(\mathbf{\Lambda}) \mathbf{g} \rangle|, \quad |\langle \mathbf{g}, \text{diag}(\mathbf{\Lambda}) \hat{\mathbf{g}} \rangle| \leq \sqrt{n} \log^{c_2} n, \quad (2.4)$$

— where we have used that the spectral growth condition implies $\|\mathbf{\Lambda}\|_4 \leq n^{1/4} \log^{c_1} n$ for some $c_1 \geq 0$. Of course, the Lemma applies equally well to any finite sequence of independent standard Gaussians.

2.4. Proofs of corollaries. Having defined the necessary notions, we can now explain how to derive the Corollary 1.4 and Corollary 1.5 from Theorem 1.3.

Proof of Corollary 1.4. Suppose that $\mathbf{\Lambda}_n = \mathbf{\Lambda}$ satisfies

- (1) the assumption of Theorem 1.3,
- (2) $\text{avg}(\mathbf{\Lambda}_n) = 0$, $\text{std}(\mathbf{\Lambda}_n) = 1$ with probability 1, and
- (3) there exist some probability measure ρ on \mathbb{R} such that

$$\text{emp}(\mathbf{\Lambda}_n) \xrightarrow{\text{a.s.}} \rho, \quad \text{as } n \rightarrow \infty.$$

Let $x_n = \mathbb{E}[\mathbf{\Lambda}_n]$ and let φ be a bounded Lipschitz function on \mathbb{R} , with bound M and Lipschitz constant C . The above convergence implies $\frac{1}{n} \sum_{k=1}^n \varphi(x_{n,k}) \rightarrow \int \varphi d\rho$. For $A \in \text{OE}_n(p(\mathbf{\Lambda}_n))$, the normalization and Theorem 1.3 gives

$$\mathbb{E}[\phi_{\mathcal{N}}(A, k)] = x_{n,k} + \tilde{O}(n^{-1}), \quad \text{and} \quad \text{Var}[\phi_{\mathcal{N}}(A, k)] = \tilde{O}(n^{-1/2}),$$

uniformly on k . Fix $0 < \varepsilon < \frac{1}{4}$, Chebyshev's inequality (and the triangle inequality) now implies

$$\mathbb{P}[|\phi_{\mathcal{N}}(A, k) - x_{n,k}| > n^{-\varepsilon}] = \tilde{O}(n^{2\varepsilon - \frac{1}{2}}) \leq \delta_n,$$

for some sequence $\delta_n \rightarrow 0$ (independent of k). For any $t > 0$, Markov's inequality (and the triangle inequality) gives

$$\begin{aligned} \mathbb{P}\left[\left|\frac{1}{n} \sum_{k=1}^n \varphi(\phi_{\mathcal{N}}(A, k)) - \varphi(x_{n,k})\right| > t\right] &\leq \frac{1}{t} \frac{1}{n} \sum_{k=1}^n \mathbb{E}\left[|\varphi(\phi_{\mathcal{N}}(A, k)) - \varphi(x_{n,k})|\right] \\ &\leq \frac{1}{t} (C n^{-\varepsilon} + 2M \delta_n) \rightarrow 0. \end{aligned}$$

In particular, the estimates continues to hold for $t = \frac{1}{\sqrt{n^\varepsilon}}$. So, $\frac{1}{n} \sum_{k=1}^n \varphi(x_{n,k}) \rightarrow \int \varphi d\rho$, and we conclude that

$$\mathbb{P} \left[\frac{1}{n} \sum_{k=1}^n \varphi(\phi_{\mathcal{N}}(A, k)) \rightarrow \int \varphi d\rho \right] = 1.$$

Using the Portmanteau Theorem (to go from any bounded Lipschitz to any bounded continuous) we conclude that $\text{emp}(\phi_{\mathcal{N}}(A)) \xrightarrow{\text{a.s.}} \rho$, as needed.

Since $\phi_{\mathcal{N}}(A, k) - \sigma_{\mathcal{N}}(A, k) = c(k-1) \frac{\sqrt{n}}{\binom{n}{2}}$ is deterministic and is uniformly bounded by $\frac{c'}{\sqrt{n}}$, the same argument gives $\text{emp}(\sigma_{\mathcal{N}}(A)) \xrightarrow{\text{a.s.}} \rho$ \square

Corollary 1.5 follows immediately from Corollary 1.4, Wigner's Semi-Circle Theorem (2.3), and the following two facts:

- (1) if A is GOE then A_{sgn} is GOE,
- (2) if $\text{emp}(\sigma_{\mathcal{N}}(A)) \xrightarrow{\text{a.s.}} \rho$ and ρ is symmetric around 0, then $\text{emp}_{\mathcal{N}}(\sigma(A)) \xrightarrow{\text{a.s.}} \rho$.

\square

3. SIGNS OF QUADRATIC FORMS ON THE ORTHOGONAL GROUP

Our main tool for the proof of Theorem 1.3 is a precise estimate on the possible correlation between the signs of certain quadratic forms on the orthogonal group. As mentioned, the main difficulty in handling such integrals lies in the inclusion of the sign function. To address this difficulty, we reduce the computation on the orthogonal group to a computation in Gaussian space. For Gaussian variables, there is an explicit formula for the correlation of signs, sometimes called Grothendieck's identity or Sheppard's formula.

Lemma 3.1. *Let $(X, Y) \in \mathbb{R}^2$ be a centered Gaussian vector with unit variances and $\mathbb{E}[XY] = \rho$. Then*

$$\mathbb{E}[\text{sgn}(X) \text{sgn}(Y)] = \frac{2}{\pi} \arcsin(\rho) = \frac{2}{\pi} \left(\rho + \frac{\rho^3}{6} + \frac{3\rho^5}{40} + O(\rho^7) \right).$$

The reader can find a proof in [17, Lemma 3.2] for example, while the identity $\frac{2}{\pi} \arcsin(\rho) = \frac{2}{\pi} \left(\rho + \frac{\rho^3}{6} + \frac{3\rho^5}{40} + O(\rho^7) \right)$ is an immediate consequence of the Taylor expansion. We can now state and prove our main technical tool.

Proposition 3.2. *Let B be a real $n \times n$ matrix and let P be a rank-one orthogonal projection. Assume that $\|(I - P)B(I - P)\|_F^2 = n + \tilde{O}(1)$ and that the operator norms $\|B\|$, and $\|B^T P\|$ are $\tilde{O}(1)$. Further assume $\text{Tr}(B) = \tilde{O}(1)$. Let $\mathbf{u}, \hat{\mathbf{u}}$ denote the first two columns of a Haar-random orthogonal matrix in $\mathcal{O}(n)$. Then*

$$\mathbb{E}[\text{sgn}(\langle P\mathbf{u}, \hat{\mathbf{u}} \rangle \langle B\mathbf{u}, \hat{\mathbf{u}} \rangle)] = \frac{2^{3/2}}{\pi^{3/2}} \text{Tr}(PB) n^{-1/2} + \tilde{O}(n^{-3/2}).$$

Before proving the proposition, we mention that for symmetric matrices, with no loss of generality one can assume that B is a diagonal. In that case, we can take P to be the orthogonal projection on a coordinate e_k , for some $k = 1, \dots, k$, which corresponds to an eigenspace of B with, say, eigenvalue λ_k . With these choices, as long as A is drawn from an orthogonally invariant ensemble, we shall show that

$$\mathbb{E}[\phi(A, k)] = \mathbb{E}[\text{sgn}(\langle P\mathbf{u}, \hat{\mathbf{u}} \rangle \langle B\mathbf{u}, \hat{\mathbf{u}} \rangle)],$$

where B is the diagonalization of A . Thus Proposition 3.2 provides a direct formula for the expected value of $\phi(A, k)$. Building in this representation and the ideas that will appear in the proof, we shall also use this formula to control the variance of $\phi(A, k)$.

Proof of Proposition 3.2. Our proof is conducted in several steps. We first prove an analog result for quadratic forms in Gaussians. We then explain how to reduce, through conditioning, the computation on the orthogonal group to Gaussian space. We then estimate the error by integrating over the conditioned variables.

Step 0: Introduce the Gaussian setting. Let \mathbf{g} and $\hat{\mathbf{g}}$ be two independent standard Gaussian vectors in \mathbb{R}^n . By rotational invariance of $(\mathbf{g}, \hat{\mathbf{g}})$, we may and will assume $P = e_1 e_1^T$ is the projection onto the first coordinate. The well-known construction of the first two Haar columns is

$$(\mathbf{u}, \hat{\mathbf{u}}) \stackrel{\text{law}}{=} \left(\frac{\mathbf{g}}{\|\mathbf{g}\|}, \frac{Q\hat{\mathbf{g}}}{\|Q\hat{\mathbf{g}}\|} \right), \quad Q = Q_{\mathbf{g}} := I_n - \frac{\mathbf{g}\mathbf{g}^T}{\|\mathbf{g}\|^2}.$$

The map $(u, v) \mapsto \text{sgn}(\langle Pu, v \rangle \langle Bu, v \rangle)$ is invariant under *independent scalings* of u and v . Hence

$$\text{sgn}(\langle Pu, \hat{\mathbf{u}} \rangle \langle Bu, \hat{\mathbf{u}} \rangle) \stackrel{\text{law}}{=} \text{sgn}(\langle P\mathbf{g}, Q\hat{\mathbf{g}} \rangle \langle B\mathbf{g}, Q\hat{\mathbf{g}} \rangle) = \text{sgn}(g_1 \langle Qe_1, \hat{\mathbf{g}} \rangle \langle QB\mathbf{g}, \hat{\mathbf{g}} \rangle).$$

Step 1: Condition on \mathbf{g} and reduce to two linear forms of $\hat{\mathbf{g}}$. Fix \mathbf{g} and set

$$u := \text{sgn}(g_1)Qe_1, \quad v := QB\mathbf{g}, \quad X := \frac{\langle u, \hat{\mathbf{g}} \rangle}{\|u\|}, \quad Y := \frac{\langle v, \hat{\mathbf{g}} \rangle}{\|v\|}.$$

For almost any value of \mathbf{g} , $(X, Y) \mid \mathbf{g}$ is bivariate normal, centered, with

$$\text{Var}(X \mid \mathbf{g}) = 1, \quad \text{Var}(Y \mid \mathbf{g}) = 1, \quad \rho := \text{Cov}(X, Y \mid \mathbf{g}) = \frac{\langle u, v \rangle}{\|u\|\|v\|} = \text{sgn}(g_1) \frac{\langle Qe_1, QB\mathbf{g} \rangle}{\|Qe_1\|\|QB\mathbf{g}\|},$$

so that, by Lemma Lemma 3.1,

$$\mathbb{E}_{\hat{\mathbf{g}}}(\text{sgn}(\langle Pu, \hat{\mathbf{u}} \rangle \langle Bu, \hat{\mathbf{u}} \rangle)) = \mathbb{E}_{\hat{\mathbf{g}}}(\text{sgn}(X)\text{sgn}(Y)) = \frac{2}{\pi} \left(\rho + \frac{\rho^3}{6} \right) + O(\rho^5).$$

It is now a matter of bounding the random correlation ρ . Notice that $\langle Qe_1, QB\mathbf{g} \rangle = \langle Qe_1, B\mathbf{g} \rangle$ due to $Q^2 = Q = Q^T$. The fact that $\|Qx\|^2 = \|x\|^2 - \frac{|\langle x, \mathbf{g} \rangle|^2}{\|\mathbf{g}\|^2}$ for any $x \in \mathbb{R}^n$ allows to write

$$\text{sgn}(g_1)\rho = \frac{\|\mathbf{g}\|^2 \langle Qe_1, B\mathbf{g} \rangle}{\sqrt{(\|\mathbf{g}\|^2 - g_1^2)(\|\mathbf{g}\|^2 \|B\mathbf{g}\|^2 - |\langle \mathbf{g}, B\mathbf{g} \rangle|^2)}} = \frac{\|\mathbf{g}\|^2 (B\mathbf{g})_1 - g_1 \langle \mathbf{g}, B\mathbf{g} \rangle}{\sqrt{(\|\mathbf{g}\|^2 - g_1^2)(\|\mathbf{g}\|^2 \|B\mathbf{g}\|^2 - |\langle \mathbf{g}, B\mathbf{g} \rangle|^2)}}.$$

Define the fluctuations

$$\varepsilon := \frac{\|\mathbf{g}\|^2 - n}{n}, \quad \delta := \frac{\|B\mathbf{g}\|^2 - n}{n}, \quad \eta := \frac{\langle \mathbf{g}, B\mathbf{g} \rangle}{n}. \quad (3.1)$$

Notice that $\mathbb{E}[\varepsilon] = 0$, $\mathbb{E}[\delta] = n^{-1}(\|B\|_F^2 - n)$, and $\mathbb{E}[\eta] = n^{-1}\text{Tr}(B)$. The assumptions that $\|B\|_F^2 - n = O(1)$ and $\text{Tr}(B) = \tilde{O}(1)$, together with the Gaussian concentration from Lemma 2.3, show that with high probability (e.g. $1 - n^{-\log n}$) the random variables $\varepsilon, \delta, \eta$ are $\tilde{O}(n^{-1/2})$. Define the Gaussian

$$Z = (B\mathbf{g})_1 = \langle B^T e_1, \mathbf{g} \rangle.$$

Notice that $\|B^T e_1\| = \|B^T P e_1\| \leq \|B^T P\| = \tilde{O}(1)$ by assumption. So g_1 and Z are centered Gaussians with variances 1 and $\tilde{O}(1)$ respectively. In particular, with high probability g_1 and Z are $\tilde{O}(1)$. We

may conclude that with high probability (e.g. $1 - n^{-\log n}$),

$$\begin{aligned}
\operatorname{sgn}(g_1)\rho &= \frac{n(1+\varepsilon)Z - n\eta g_1}{\sqrt{(n(1+\varepsilon) - g_1^2)(n^2(1+\varepsilon)(1+\delta) - n^2\eta^2)}} \\
&= n^{-1/2} \frac{Z - \eta g_1 + \tilde{O}(n^{-3/2})}{\sqrt{(1+\varepsilon + g_1^2 n^{-1})(1+\varepsilon + \delta + \varepsilon\delta - \eta^2)}} \\
&= n^{-1/2} \frac{Z - \eta g_1 + \tilde{O}(n^{-3/2})}{\sqrt{1 + 2\varepsilon + \delta + g_1^2 n^{-1} + \varepsilon\delta - \eta^2 + \tilde{O}(n^{-3/2})}} \\
&= n^{-1/2} (Z - \eta g_1) \left(1 - \frac{1}{2}(2\varepsilon + \delta + g_1^2 n^{-1} + \varepsilon\delta - \eta^2) + \frac{3}{8}(2\varepsilon + \delta)^2 + \tilde{O}(n^{-3/2})\right) \\
&= n^{-1/2} \left[Z \left(1 - \varepsilon - \frac{\delta}{2} - \frac{1}{2}(g_1^2 n^{-1} + 4\varepsilon\delta - \eta^2 + 3\varepsilon^2 + \delta^2/4)\right) - \eta g_1 \left(1 - \varepsilon - \frac{\delta}{2}\right) \right] + \tilde{O}(n^{-2}) \\
&= n^{-1/2} Z - n^{-1/2} \left(Z \left(\varepsilon + \frac{\delta}{2}\right) + \eta g_1 \right) + W + \tilde{O}(n^{-2}),
\end{aligned}$$

where W is with high probability $\tilde{O}(n^{-3/2})$

$$W := n^{-1/2} \left(\eta g_1 \left(\varepsilon + \frac{\delta}{2}\right) - Z \frac{1}{2} (g_1^2 n^{-1} + 4\varepsilon\delta - \eta^2 + 3\varepsilon^2 + \delta^2/4) \right).$$

In particular, with probability at least $1 - n^{-\log n}$,

$$\rho = \tilde{O}(n^{-1/2}), \quad \text{and} \quad \rho = n^{-1/2} \operatorname{sgn}(g_1) \left(Z - Z\varepsilon - Z\frac{\delta}{2} + \eta g_1 \right) + \tilde{O}(n^{-3/2}).$$

Since $|\operatorname{sgn}| \leq 1$, we can now calculate up to $\tilde{O}(n^{-3/2})$,

$$\mathbb{E}[\operatorname{sgn}(\langle P\mathbf{u}, \hat{\mathbf{u}} \rangle \langle B\mathbf{u}, \hat{\mathbf{u}} \rangle)] = \frac{2}{\pi} \mathbb{E}_{\mathbf{g}}[\rho] + \tilde{O}(n^{-3/2}).$$

Step 2: Integration over $\mathbf{w} = (\mathbf{I}_n - e_1 e_1^T) \mathbf{g}$. To do so, we let $\mathbf{w} = (0, g_2, \dots, g_n)$, so that its $n-1$ non-zero entries form a standard $n-1$ dimensional Gaussian independent of g_1 , and we first integrate over \mathbf{w} . We can write

$$\varepsilon = \frac{\|\mathbf{w}\|^2 + g_1^2 - n}{n}, \quad \delta := \frac{\|B\mathbf{w}\|^2 + g_1 B_{11} - n}{n}, \quad \eta := \frac{B_{11} g_1^2 + \langle \mathbf{w}, B\mathbf{w} \rangle + \langle \mathbf{w}, B e_1 \rangle + \langle e_1, B\mathbf{w} \rangle}{n},$$

and

$$Z = B_{11} g_1 + \langle B^T e_1, \mathbf{w} \rangle.$$

Notice that by our assumptions, $\mathbb{E}_{\mathbf{w}} \|B\mathbf{w}\|^2 = \|(\mathbf{I} - P)B(\mathbf{I} - P)\|_F^2 = n + \tilde{O}(1)$ and $\mathbb{E}_{\mathbf{w}} \langle \mathbf{w}, B\mathbf{w} \rangle = \operatorname{Tr}((\mathbf{I} - P)B(\mathbf{I} - P)) = \operatorname{Tr}(B) - \operatorname{Tr}(PB) = \tilde{O}(1)$. Using these observations and the symmetry $\mathbf{w} \mapsto -\mathbf{w}$, we get

$$\begin{aligned}
\mathbb{E}_{\mathbf{w}}[Z] &= B_{11} g_1, \\
\mathbb{E}_{\mathbf{w}}[Z\varepsilon] &= \mathbb{E}_{\mathbf{w}}[B_{11} g_1 \varepsilon] = \frac{B_{11} g_1 (g_1^2 - 1)}{n} = \tilde{O}(1/n), \\
\mathbb{E}_{\mathbf{w}}[Z\delta] &= \mathbb{E}_{\mathbf{w}}[B_{11} g_1 \delta] = \frac{B_{11} g_1 (B_{11} g_1 + \tilde{O}(1))}{n} = \tilde{O}(1/n), \\
\mathbb{E}_{\mathbf{w}}[\eta] &= \frac{B_{11} g_1^2 + \tilde{O}(1)}{n} = \tilde{O}(1/n).
\end{aligned}$$

We conclude that,

$$\begin{aligned}\mathbb{E}_{\mathbf{g}}[\rho] &= \mathbb{E}_{g_1}[\mathbb{E}_{\mathbf{w}}(\rho)] = n^{-1/2} \text{Tr}(PB) \mathbb{E}_{g_1}[|g_1|] + \tilde{O}(n^{-3/2}) \\ &= \sqrt{\frac{2}{\pi}} n^{-1/2} \text{Tr}(PB) + \tilde{O}(n^{-3/2}),\end{aligned}$$

and therefore

$$\mathbb{E}[\text{sgn}(\langle P\mathbf{u}, \hat{\mathbf{u}} \rangle \langle B\mathbf{u}, \hat{\mathbf{u}} \rangle)] = \frac{2}{\pi} \mathbb{E}_{\mathbf{g}}[\rho] + \tilde{O}(n^{-3/2}) = \frac{2^{3/2}}{\pi^{3/2}} \text{Tr}(PB) n^{-1/2} + \tilde{O}(n^{-3/2}).$$

This proves the claim. \square

4. STATISTICS OF THE NODAL COUNT

In this section, we employ Proposition 3.2 and prove Theorem 1.3. Recall the setting: $A = \Phi \text{diag}(\Lambda) \Phi^T$, where $\Phi \sim \text{Haar}(\mathcal{O}(n))$, and $\Lambda = (\lambda_1, \dots, \lambda_n)$ can be random, but Φ and $\Lambda = (\lambda_1, \dots, \lambda_n)$ are independent. We can and will assume without loss of generality that $\text{avg}(\Lambda) = 0$ and $\text{std}(\Lambda) = 1$ with probability 1. We also assume the spectral growth bound from Definition 2.1, according to which $\|\Lambda\|_{\infty} = \tilde{O}(1)$ with probability $1 - \tilde{O}(n^{-3/2})$. Our goal is to prove Theorem 1.3, which states that

$$\mathbb{E}_{\Phi, \Lambda}[\phi(A, k)] = \binom{n}{2} \left(\frac{1}{2} + \frac{\sqrt{2}}{\pi^{3/2}} \mathbb{E}_{\Lambda}[\lambda_k] n^{-1/2} + \tilde{O}(n^{-3/2}) \right), \quad \text{and} \quad (4.1)$$

$$\text{Var}_{\Phi, \Lambda}[\phi(A, k)] = \tilde{O}(n^{5/2}). \quad (4.2)$$

The first part of the proof is a simple reduction based on the permutation symmetry of rows in $\mathcal{O}(n)$. Let $\Phi_j = (\Phi_{j,1}, \dots, \Phi_{j,n})$ denote the j -th row of Φ , let $P_k = e_k e_k^T$ be the projection on the k -th coordinate, and for $i < j$ and $r < s$ define the random variable

$$M_{ij} := A_{ij} \Phi_{i,k} \Phi_{j,k} = \langle \Phi_i, P_k \Phi_j \rangle \langle \Phi_i, \text{diag}(\Lambda) \Phi_j \rangle,$$

so that

$$\phi(A, k) = \frac{1}{2} \sum_{i < j} (1 + \text{sign}(M_{ij})).$$

Notice that the permutation symmetry of rows in $\mathcal{O}(n)$ gives

$$\begin{aligned}M_{ij} &\stackrel{\text{law}}{=} M_{12}, \\ M_{ij} M_{rs} &\stackrel{\text{law}}{=} \begin{cases} M_{12}^2 & (i, j) = (r, s) \\ M_{12} M_{34} & i \neq r, j \neq s \\ M_{13} M_{23} & i \neq r, j = s \quad \text{or} \quad i = r, j \neq s \end{cases}.\end{aligned}$$

Linearity of expectation and bilinearity of covariance allow us to conclude

$$\begin{aligned}\mathbb{E}[\phi(A, k)] &= \frac{1}{2} \binom{n}{2} (1 + \mathbb{E}[\text{sign}(M_{12})]), \quad \text{and} \\ \text{Var}(\phi(A, k)) &= \text{Cov} \left(\frac{1}{2} \sum_{i < j} \text{sign}(M_{ij}), \frac{1}{2} \sum_{r < s} \text{sign}(M_{rs}) \right) \\ &= \frac{1}{4} \binom{n}{2} \text{Var}(\text{sgn}(M_{12})) + \frac{1}{2} (n-2) \binom{n}{2} \text{Cov}(\text{sgn}(M_{13}), \text{sgn}(M_{23})) \\ &\quad + 2 \binom{n}{2} \binom{n-2}{2} \text{Cov}(\text{sgn}(M_{12} M_{34})).\end{aligned}$$

We now prove Theorem 1.3 based on three Lemmas. The second part of the proof, which is the technical part, is to prove these Lemmas.

Lemma 4.1 (Expected value for an edge). *Let $\mathbf{\Lambda}$ be fixed, and assume it satisfies $\text{avg}(\mathbf{\Lambda}) = 0$, $\text{std}(\mathbf{\Lambda}) = 1$, and $\|\mathbf{\Lambda}\|_\infty = \tilde{O}(1)$. Then,*

$$\mathbb{E}_\Phi [\text{sgn}(M_{12})] = \frac{2^{3/2}}{\pi^{3/2}} \lambda_k n^{-1/2} + \tilde{O}(n^{-3/2}).$$

This lemma is straightforward from Proposition 3.2.

Proof. The assumption on $\mathbf{\Lambda}$ implies that $B = \text{diag}(\mathbf{\Lambda})$ and $P = e_k e_k^T$ satisfy the conditions of Proposition 3.2, and $\text{Tr}(BP) = \lambda_k$. which tells us that

$$\mathbb{E}_\Phi [\text{sgn}(M_{12})] = \mathbb{E}_\Phi [\text{sgn}(\langle \Phi_1, P\Phi_2 \rangle \langle \Phi_1, \text{diag}(\mathbf{\Lambda})\Phi_2 \rangle)] = \frac{2^{3/2}}{\pi^{3/2}} n^{-1/2} \lambda_k + \tilde{O}(n^{-3/2}).$$

□

Lemma 4.2 (Covariance of Adjacent Edges). *Let $\mathbf{\Lambda}$ be fixed, and assume it satisfies $\text{avg}(\mathbf{\Lambda}) = 0$, $\text{std}(\mathbf{\Lambda}) = 1$, and $\|\mathbf{\Lambda}\|_\infty = \tilde{O}(1)$. Then,*

$$\text{Cov}_\Phi (\text{sgn}(M_{13}), \text{sgn}(M_{23})) = \tilde{O}(n^{-1/2}).$$

Lemma 4.3 (Covariance of Non-Adjacent Edges). *Let $\mathbf{\Lambda}$ be fixed, and assume it satisfies $\text{avg}(\mathbf{\Lambda}) = 0$, $\text{std}(\mathbf{\Lambda}) = 1$, and $\|\mathbf{\Lambda}\|_\infty = \tilde{O}(1)$. Then,*

$$\text{Cov}_\Phi (\text{sgn}(M_{12}), \text{sgn}(M_{34})) = \tilde{O}(n^{-3/2})$$

The proofs of Lemma 4.3 and Lemma 4.2 are in Subsection 4.1. We now use these lemmas to prove Theorem 1.3.

Proof of Theorem 1.3 based on Lemmas 4.1, 4.2, 4.3. Let \mathcal{E}_T be the event of $\|\mathbf{\Lambda}\|_\infty = \tilde{O}(1)$ and let $\mathbf{1}_{\{T\}}$ be the indicator of \mathcal{E}_T . By assumption, $\mathbb{E}(1 - \mathbf{1}_{\{T\}}) = \tilde{O}(n^{-3/2})$, so for any x, y random variables

$$\mathbb{E}[\text{sign}(x)] = \mathbb{E}[\text{sign}(x)\mathbf{1}_{\{T\}} + \text{sign}(x)(1 - \mathbf{1}_{\{T\}})] = \mathbb{E}[\text{sign}(x)\mathbf{1}_{\{T\}}] + \tilde{O}(n^{-3/2}),$$

and similarly,

$$\text{Cov}(\text{sign}(x), \text{sign}(y)) = \text{Cov}(\text{sign}(x)\mathbf{1}_{\{T\}}, \text{sign}(y)) + \tilde{O}(n^{-3/2}).$$

By Lemma 4.1, we conclude the needed expectation result

$$\mathbb{E}[\phi(A, k)] = \frac{1}{2} \binom{n}{2} \left(1 + \mathbb{E}[\text{sign}(M_{12})\mathbf{1}_{\{T\}}] + \tilde{O}(n^{-3/2}) \right) = \binom{n}{2} \left(\frac{1}{2} + \frac{\sqrt{2}}{\pi^{3/2}} n^{-1/2} \mathbb{E}_\Lambda[\lambda_k] + \tilde{O}(n^{-3/2}) \right).$$

Using that $\text{Var}(\text{sign}(M_{12})) \leq 1$ by definition, together with Lemma 4.3 and Lemma 4.2, we get the needed variance bound

$$\begin{aligned} \text{Var}(\phi(A, k)) &\leq \frac{1}{4} \binom{n}{2} + \frac{1}{2} (n-2) \binom{n}{2} \left(\text{Cov}(\text{sgn}(M_{13})\mathbf{1}_{\{T\}}, \text{sgn}(M_{23})) + \tilde{O}(n^{-3/2}) \right) \\ &\quad + 2 \binom{n}{2} \binom{n-2}{2} \left(\text{Cov}(\text{sgn}(M_{12})\mathbf{1}_{\{T\}}, \text{sgn}(M_{34})) + \tilde{O}(n^{-3/2}) \right) \\ &= \tilde{O}(n^{5/2}). \end{aligned}$$

□

4.1. **Estimating** $\text{Cov}(\text{sign}(M_{ij}), \text{sign}(M_{rs}))$ - **Proofs of Lemma 4.2 and Lemma 4.3.** For convenience denote $(\mathbf{u}, \hat{\mathbf{u}}, \mathbf{v}, \hat{\mathbf{v}}) = (\Phi_1, \Phi_2, \Phi_3, \Phi_4)$, the four columns of a Haar random orthogonal matrix Φ .¹ We assume $\mathbf{\Lambda} = (\lambda_1, \dots, \lambda_n)$ is fixed with $\text{avg}(\mathbf{\Lambda}) = 0$, $\text{std}(\mathbf{\Lambda}) = 1$, and $\|\mathbf{\Lambda}\|_\infty = \tilde{O}(1)$. Define the quartic polynomial $M : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ by

$$M(x, y) := x_k y_k \sum_{j=1}^n \lambda_j x_j y_j = \langle e_k e_k^T x, y \rangle \langle \text{diag}(\mathbf{\Lambda}) x, y \rangle, \quad (4.3)$$

so that

$$M_{12} = M(\mathbf{u}, \hat{\mathbf{u}}), \quad M_{13} = M(\mathbf{u}, \mathbf{v}), \quad M_{23} = M(\hat{\mathbf{u}}, \mathbf{v}), \quad M_{34} = M(\hat{\mathbf{u}}, \hat{\mathbf{v}}). \quad (4.4)$$

To help with the conditioning, we shall reduce some of the variables to Gaussian space. For that, we also introduce the (random) isometry $U : \mathbb{R}^{n-2} \rightarrow \text{span}(\mathbf{u}, \hat{\mathbf{u}})^\perp \subset \mathbb{R}^n$. In other words, U is the unique, up to a change of basis, matrix such that $U^T U = I_{n-2}$ and $U U^T = I_n - \mathbf{u} \mathbf{u}^T - \hat{\mathbf{u}} \hat{\mathbf{u}}^T$.

Proof of Lemma 4.2. Let $\mathbf{g} \sim N(0, I_{n-2})$ be a standard Gaussian vector independent of $\mathbf{u}, \hat{\mathbf{u}}$. By the definition of the matrix U , we have

$$\begin{aligned} \mathbf{v} \mid (\mathbf{u}, \hat{\mathbf{u}}) &\stackrel{\text{law}}{=} \frac{U \mathbf{g}}{\|U \mathbf{g}\|} \\ \text{sgn}(M_{13} M_{23}) \mid (\mathbf{u}, \hat{\mathbf{u}}) &\stackrel{\text{law}}{=} \text{sgn}(M(\mathbf{u}, U \mathbf{g}) M(\hat{\mathbf{u}}, U \mathbf{g})) \\ &= \text{sgn}(\langle e_k e_k^T \mathbf{u}, U \mathbf{g} \rangle \langle \text{diag}(\mathbf{\Lambda}) \mathbf{u}, U \mathbf{g} \rangle \langle e_k e_k^T \hat{\mathbf{u}}, U \mathbf{g} \rangle \langle \text{diag}(\mathbf{\Lambda}) \hat{\mathbf{u}}, U \mathbf{g} \rangle) \\ &= \text{sgn}(u_k \hat{u}_k) \text{sgn}(\langle U^T e_k, \mathbf{g} \rangle \langle U^T \text{diag}(\mathbf{\Lambda}) \mathbf{u}, \mathbf{g} \rangle \langle U^T e_k, \mathbf{g} \rangle \langle U^T \text{diag}(\mathbf{\Lambda}) \hat{\mathbf{u}}, \mathbf{g} \rangle) \\ &= \text{sgn}(u_k \hat{u}_k) \text{sgn}(\langle U^T \text{diag}(\mathbf{\Lambda}) \mathbf{u}, \mathbf{g} \rangle \langle U^T \text{diag}(\mathbf{\Lambda}) \hat{\mathbf{u}}, \mathbf{g} \rangle) \end{aligned}$$

Denote the following $(\mathbf{u}, \hat{\mathbf{u}})$ -dependent parameters

$\alpha_{\mathbf{u}} := \|\text{diag}(\mathbf{\Lambda}) \mathbf{u}\|$, $\alpha_{\hat{\mathbf{u}}} := \|\text{diag}(\mathbf{\Lambda}) \hat{\mathbf{u}}\|$, $\beta_{\mathbf{u}} = \langle \mathbf{u}, \text{diag}(\mathbf{\Lambda}) \mathbf{u} \rangle$, $\beta_{\hat{\mathbf{u}}} = \langle \hat{\mathbf{u}}, \text{diag}(\mathbf{\Lambda}) \hat{\mathbf{u}} \rangle$, $\gamma := \langle \mathbf{u}, \text{diag}(\mathbf{\Lambda}) \hat{\mathbf{u}} \rangle$, and define

$$\begin{aligned} \rho &:= \frac{\langle U^T \text{diag}(\mathbf{\Lambda}) \mathbf{u}, U^T \text{diag}(\mathbf{\Lambda}) \hat{\mathbf{u}} \rangle}{\|U^T \text{diag}(\mathbf{\Lambda}) \mathbf{u}\| \|U^T \text{diag}(\mathbf{\Lambda}) \hat{\mathbf{u}}\|} \\ &= \frac{\langle \text{diag}(\mathbf{\Lambda}) U U^T \text{diag}(\mathbf{\Lambda}) \mathbf{u}, \hat{\mathbf{u}} \rangle}{\|U U^T \text{diag}(\mathbf{\Lambda}) \mathbf{u}\| \|U U^T \text{diag}(\mathbf{\Lambda}) \hat{\mathbf{u}}\|} \\ &= \frac{\langle \text{diag}(\mathbf{\Lambda})^2 \mathbf{u}, \hat{\mathbf{u}} \rangle - \beta_{\mathbf{u}} \gamma - \beta_{\hat{\mathbf{u}}} \gamma}{\sqrt{(\alpha_{\mathbf{u}}^2 - \beta_{\mathbf{u}}^2 - \gamma^2)(\alpha_{\hat{\mathbf{u}}}^2 - \beta_{\hat{\mathbf{u}}}^2 - \gamma^2)}}, \end{aligned}$$

so that Lemma 3.1 gives

$$\mathbb{E}[\text{sgn}(M_{13} M_{23}) \mid (\mathbf{u}, \hat{\mathbf{u}})] = \frac{2}{\pi} \text{sgn}(u_k \hat{u}_k) \arcsin \rho = \frac{2}{\pi} \text{sgn}(u_k \hat{u}_k) \rho + O(\rho^3).$$

Lemma 2.3 allows to estimate the above parameters. First consider $\xi, \hat{\xi}$ independent standard Gaussian vectors of size n , and let $Q = I_n - \frac{\xi \xi^T}{\|\xi\|^2}$, so that

$$(\mathbf{u}, \hat{\mathbf{u}}) \stackrel{\text{law}}{=} \left(\frac{\xi}{\|\xi\|}, \frac{Q \hat{\xi}}{\|Q \hat{\xi}\|} \right).$$

Denote the following $(\xi, \hat{\xi})$ -dependent small parameters

$$\varepsilon_1 := \frac{\|\xi\|^2 - n}{n}, \quad \varepsilon_2 := \frac{\|\hat{\xi}\|^2 - n}{n}, \quad \varepsilon_3 := \frac{\|\text{diag}(\mathbf{\Lambda}) \xi\|^2 - n}{n}, \quad \varepsilon_4 := \frac{\|\text{diag}(\mathbf{\Lambda}) \hat{\xi}\|^2 - n}{n},$$

¹Equivalently, $(\mathbf{u}, \hat{\mathbf{u}}, \mathbf{v}, \hat{\mathbf{v}})$ is a random element of the Stiefel manifold $V_4(\mathbb{R}^n)$ with the the uniform probability measure.

$$\varepsilon_5 = \frac{\langle \xi, \text{diag}(\mathbf{\Lambda})\xi \rangle}{n}, \quad \varepsilon_6 = \frac{\langle \hat{\xi}, \text{diag}(\mathbf{\Lambda})\hat{\xi} \rangle}{n}, \quad \varepsilon_7 := \frac{\langle \xi, \text{diag}(\mathbf{\Lambda})\hat{\xi} \rangle}{n}, \quad \varepsilon_8 := \frac{\langle \xi, \text{diag}(\mathbf{\Lambda})^2\hat{\xi} \rangle}{n}.$$

Notice that $\mathbb{E}[\varepsilon_j] = 0$ for all j because of the $\xi, \hat{\xi}$ independence and the normalization $\text{avg}(\mathbf{\Lambda}) = 0$, $\text{std}(\mathbf{\Lambda}) = 1$. Moreover, each ε_j is a quadratic form in the variables $\xi, \hat{\xi}$ and so, together with the bound on $\|\mathbf{\Lambda}\|_\infty$, the Gaussian concentration from Lemma 2.3 applies, as in (2.4). This means that the typical event \mathcal{E}_T for which $\varepsilon_j = \tilde{O}(n^{-1/2})$ for all j uniformly has $\mathbb{P}(\mathcal{E}_T) \geq 1 - n^{-\log n}$ by Lemma 2.3. We can write

$$\alpha_{\mathbf{u}}^2 = \frac{1 + \varepsilon_3}{1 + \varepsilon_1}, \quad \alpha_{\hat{\mathbf{u}}}^2 = \frac{1 + \varepsilon_4}{1 + \varepsilon_2}, \quad \beta_{\mathbf{u}} = \frac{\varepsilon_5}{(1 + \varepsilon_1)}, \quad \beta_{\hat{\mathbf{u}}} = \frac{\varepsilon_6}{(1 + \varepsilon_2)}, \quad \gamma := \frac{\varepsilon_7}{\sqrt{(1 + \varepsilon_1)(1 + \varepsilon_2)}},$$

which means that under the typical event

$$\rho \mathbf{1}_{\{T\}} = \frac{\frac{\varepsilon_8}{\sqrt{(1+\varepsilon_1)(1+\varepsilon_2)}} - \frac{\varepsilon_5 \varepsilon_7}{(1+\varepsilon_1)\sqrt{(1+\varepsilon_1)(1+\varepsilon_2)}} - \frac{\varepsilon_6 \varepsilon_7}{(1+\varepsilon_2)\sqrt{(1+\varepsilon_1)(1+\varepsilon_2)}}}{\sqrt{\frac{1+\varepsilon_3}{1+\varepsilon_1} - \frac{\varepsilon_5^2}{(1+\varepsilon_1)^2} - \frac{\varepsilon_7^2}{(1+\varepsilon_1)(1+\varepsilon_2)}} \sqrt{\frac{1+\varepsilon_4}{1+\varepsilon_2} - \frac{\varepsilon_6^2}{(1+\varepsilon_2)^2} - \frac{\varepsilon_7^2}{(1+\varepsilon_1)(1+\varepsilon_2)}}} \mathbf{1}_{\{T\}} = \tilde{O}(n^{-1/2}).$$

Therefore

$$\mathbb{E}[\text{sgn}(M_{13}M_{23})] = \mathbb{E}_{\mathbf{u}, \hat{\mathbf{u}}} \left[\frac{2}{\pi} \text{sgn}(u_k \hat{u}_k) \arcsin \rho \mathbf{1}_{\{T\}} \right] + \tilde{O}(n^{-\log n}) = \tilde{O}(n^{-1/2}).$$

Since $\mathbb{E}[\text{sgn}(M_{13})] \mathbb{E}[\text{sgn}(M_{23})] = \mathbb{E}[\text{sgn}(M_{12})]^2 = \tilde{O}(n^{-1})$ by Lemma 4.1, we conclude that

$$\text{Cov}(\text{sgn}(M_{13}), \text{sgn}(M_{23})) = \tilde{O}(n^{-1/2}).$$

□

Proof of Lemma 4.3. Our desired result is $I = \tilde{O}(n^{-3/2})$, where

$$I := \mathbb{E}[\text{sgn}(M(\mathbf{u}, \hat{\mathbf{u}})M(\mathbf{v}, \hat{\mathbf{v}}))] - \mathbb{E}[\text{sgn}(M(\mathbf{u}, \hat{\mathbf{u}}))]\mathbb{E}[\text{sgn}(M(\mathbf{v}, \hat{\mathbf{v}}))].$$

Notice that, by (1.2),

$$\begin{aligned} I &= \mathbb{E}_{\mathbf{u}, \hat{\mathbf{u}}} [\text{sgn}(M(\mathbf{u}, \hat{\mathbf{u}})) \mathbb{E}_{\mathbf{v}, \hat{\mathbf{v}}} [\text{sgn}(M(\mathbf{v}, \hat{\mathbf{v}})) - \text{sgn}(M(\mathbf{u}, \hat{\mathbf{u}}))]] \\ &= \mathbb{E}_{\mathbf{u}, \hat{\mathbf{u}}} \left[\text{sgn}(M(\mathbf{u}, \hat{\mathbf{u}})) \mathbb{E}_{\mathbf{v}, \hat{\mathbf{v}}} \left[\text{sgn}(M(\mathbf{v}, \hat{\mathbf{v}})) - \frac{2^{3/2}}{\pi^{3/2}} n^{-1/2} \lambda_k \right] \right] + \tilde{O}(n^{-3/2}). \end{aligned}$$

For a fixed choice of $\mathbf{u}, \hat{\mathbf{u}}$, assuming only that $(\mathbf{u})_k, (\hat{\mathbf{u}})_k$ are $\tilde{O}(n^{-1/2})$, we recall the isometry $U : \mathbb{R}^{n-2} \rightarrow \text{span}(\mathbf{u}, \hat{\mathbf{u}})^\perp$ as above. Let $\xi, \hat{\xi}$ be the first two columns of a Haar-random orthogonal $(n-2) \times (n-2)$ matrix. Define $P = \frac{1}{\|U^T e_k\|^2} U^T e_k e_k^T U$ and $B = U^T \text{diag}(\mathbf{\Lambda})U$. Then,

$$\begin{aligned} (\mathbf{v}, \hat{\mathbf{v}}) \mid (\mathbf{u}, \hat{\mathbf{u}}) &\stackrel{\text{law}}{=} (U\xi, U\hat{\xi}) \\ \text{sgn}(M(\mathbf{v}, \hat{\mathbf{v}})) \mid (\mathbf{u}, \hat{\mathbf{u}}) &\stackrel{\text{law}}{=} \text{sgn}(\langle e_k e_k^T U\xi, U\hat{\xi} \rangle \langle \text{diag}(\mathbf{\Lambda})U\xi, U\hat{\xi} \rangle) \\ &= \text{sgn}(\langle P\xi, \hat{\xi} \rangle \langle B\xi, \hat{\xi} \rangle) \end{aligned}$$

If we can show that P and B satisfy the assumptions of Proposition 3.2, and that $\text{Tr}(PB) = \lambda_k + \tilde{O}(1/n)$, then we get that with high probability in $\mathbf{u}, \hat{\mathbf{u}}$,

$$\mathbb{E}_{\mathbf{v}, \hat{\mathbf{v}}} [\text{sgn}(M(\mathbf{v}, \hat{\mathbf{v}}))] - \frac{2^{3/2}}{\pi^{3/2}} n^{-1/2} \lambda_k = \tilde{O}(n^{-3/2}),$$

and therefore,

$$I = \mathbb{E}_{\mathbf{u}, \hat{\mathbf{u}}} \left[\text{sgn}(M(\mathbf{u}, \hat{\mathbf{u}})) \left[\tilde{O}(n^{-3/2}) \right] \right] + \tilde{O}(n^{-3/2}) = \tilde{O}(n^{-3/2}).$$

We are left with showing that P and B satisfy the assumptions of Proposition 3.2, and that $\text{Tr}(PB) = \lambda_k + \tilde{O}(1/n)$. Notice that the operator norm of B is bounded by $\|B\| \leq \|\mathbf{\Lambda}\|_\infty = \tilde{O}(1)$ by assumption.

Similarly, the assumptions $\|\mathbf{\Lambda}\|_\infty = \tilde{O}(1)$ and $\text{avg}(\mathbf{\Lambda}) = 0$ gives $\text{Tr}(B) = \text{Tr}(\text{diag}(\mathbf{\Lambda})) - \langle \mathbf{u}, \text{diag}(\mathbf{\Lambda})\mathbf{u} \rangle - \langle \hat{\mathbf{u}}, \text{diag}(\mathbf{\Lambda})\hat{\mathbf{u}} \rangle = \tilde{O}(1)$. The assumption that $\text{std}(\mathbf{\Lambda}) = 1$ allows to bound the Frobenius norm: $|\|B\|_F^2 - \|\mathbf{\Lambda}\|_2^2| \leq 2\|\mathbf{\Lambda}\|_\infty^2$ means $\|B\|_F^2 = n + \tilde{O}(1)$. Similarly, $|\|(I_{n-2} - P)B(I_{n-2} - P)\|_F^2 - \|B\|_F^2| \leq 2\|\mathbf{\Lambda}\|_\infty^2$ so $\|(I_{n-2} - P)B(I_{n-2} - P)\|_F^2 = n + \tilde{O}(1)$. Finally, $\|B^T P\| \leq \|B\| = \tilde{O}(1)$. We conclude that P and B satisfy the assumptions of Proposition 3.2. We now calculate

$$\begin{aligned} \text{Tr}(PB) &= \frac{1}{\|U^T e_k\|^2} \text{Tr}(U^T e_k e_k^T U U^T \text{diag}(\mathbf{\Lambda}) U) \\ &= \frac{1}{1 - (\mathbf{u}_k^2 - (\hat{\mathbf{u}}_k^2)} \text{Tr}(e_k e_k^T (I_n - \mathbf{u}\mathbf{u}^T - \hat{\mathbf{u}}\hat{\mathbf{u}}^T) \text{diag}(\mathbf{\Lambda}) (I_n - \mathbf{u}\mathbf{u}^T - \hat{\mathbf{u}}\hat{\mathbf{u}}^T)) \\ &= \frac{\lambda_k - 2\lambda_k((\mathbf{u}_k^2 + (\hat{\mathbf{u}}_k^2) + \langle \text{diag}(\mathbf{\Lambda})(\mathbf{u}\mathbf{u}^T + \hat{\mathbf{u}}\hat{\mathbf{u}}^T)e_k, (\mathbf{u}\mathbf{u}^T + \hat{\mathbf{u}}\hat{\mathbf{u}}^T)e_k \rangle)}{1 - (\mathbf{u}_k^2 - (\hat{\mathbf{u}}_k^2)} \end{aligned}$$

so

$$\begin{aligned} |\text{Tr}(PB) - \lambda_k| &= \left| \frac{-\lambda_k((\mathbf{u}_k^2 + (\hat{\mathbf{u}}_k^2) + \langle \text{diag}(\mathbf{\Lambda})(\mathbf{u}\mathbf{u}^T + \hat{\mathbf{u}}\hat{\mathbf{u}}^T)e_k, (\mathbf{u}\mathbf{u}^T + \hat{\mathbf{u}}\hat{\mathbf{u}}^T)e_k \rangle)}{1 - (\mathbf{u}_k^2 - (\hat{\mathbf{u}}_k^2)} \right| \\ &\leq 2\|\mathbf{\Lambda}\|_\infty \left| \frac{(\mathbf{u}_k^2 + (\hat{\mathbf{u}}_k^2)}{1 - (\mathbf{u}_k^2 - (\hat{\mathbf{u}}_k^2)} \right| = \tilde{O}(1/n). \end{aligned}$$

□

ACKNOWLEDGEMENTS

The authors thank the Department of Mathematics at the Massachusetts Institute of Technology, where this project was initiated. L.A. was supported by the Department of Mathematics at MIT and by the Simons Foundation Grant No. 601948 (D.J.). D.M. was partially supported by the Brian and Tiffinie Pang Faculty Fellowship. This material is based upon work supported by the National Science Foundation under Grant No. DMS-2513687. J.U. thanks Mehtaab Sawhney for many stimulating conversations on the subject. The authors are also grateful to Louisa Thomas for helpful suggestions that improved the presentation.

REFERENCES

- [1] Lior Alon, Ram Band, and Gregory Berkolaiko. Universality of nodal count distribution in large metric graphs. *Experimental Mathematics*, 33(2):301–335, 2024.
- [2] Lior Alon, Gregory Berkolaiko, and Mark Goresky. Smooth critical points of eigenvalues on the torus of magnetic perturbations of graphs. *arXiv preprint arXiv:2505.17215*, 2025.
- [3] Lior Alon and Mark Goresky. Morse theory for discrete magnetic operators and nodal count distribution for graphs. *Journal of Spectral Theory*, 13(4):1225–1260, 2023.
- [4] Lior Alon and Mark Goresky. Nodal count for a random signing of a graph with disjoint cycles. *arXiv preprint arXiv:2403.01033*, 2024.
- [5] Lior Alon and John Urschel. Average nodal count and the nodal count condition for graphs. *arXiv preprint arXiv:2404.03151*, 2024.
- [6] Greg W. Anderson, Alice Guionnet, and Ofer Zeitouni. *An introduction to random matrices*, volume 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.
- [7] R Band, I Oren, and U Smilansky. Nodal domains on graphs: how to count them and why?, analysis on graphs and its applications, volume 77 of *proc. In Sympos. Pure Math*, pages 5–27.
- [8] Ram Band. The nodal count $\{0, 1, 2, 3, \dots\}$ implies the graph is a tree. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 372(2007):20120504, 2014.
- [9] Gregory Berkolaiko. A lower bound for nodal count on discrete and metric graphs. *Communications in mathematical physics*, 278:803–819, 2008.
- [10] Gregory Berkolaiko. Nodal count of graph eigenfunctions via magnetic perturbation. *Analysis & PDE*, 6(5):1213–1233, 2013.
- [11] Michael V Berry. Statistics of nodal lines and points in chaotic quantum billiards: perimeter corrections, fluctuations, curvature. *Journal of Physics A: Mathematical and General*, 35(13):3025, 2002.

- [12] Yves Colin de Verdière. Magnetic interpretation of the nodal defect on graphs. *Analysis & PDE*, 6(5):1235–1242, 2013.
- [13] Percy Deift and Dimitri Gioev. *Random matrix theory: invariant ensembles and universality*, volume 18 of *Courant Lecture Notes in Mathematics*. Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 2009.
- [14] Harold Donnelly and Charles Fefferman. Nodal sets of eigenfunctions on reimannian manifolds. *Inventiones mathematicae*, 93(1):161–183, 1988.
- [15] Miroslav Fiedler. Eigenvectors of acyclic matrices. *Czechoslovak Mathematical Journal*, 25(100):607–618, 1975.
- [16] Sven Gnutzmann, Uzy Smilansky, and Joachim Weber. Nodal counting on quantum graphs. *Waves in random media*, 14(1):S61, 2003.
- [17] Michel X. Goemans and David P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. Assoc. Comput. Mach.*, 42(6):1115–1145, 1995.
- [18] Svante Janson. *Gaussian hilbert spaces*. Number 129. Cambridge university press, 1997.
- [19] Manjunath Krishnapur, Pär Kurlberg, and Igor Wigman. Nodal length fluctuations for arithmetic random waves. *Annals of Mathematics*, pages 699–737, 2013.
- [20] A. Logunov. Nodal sets of laplace eigenfunctions: polynomial upper estimates of the hausdorff measure. *Annals of Mathematics (2)*, 187(1):221–239, 2018.
- [21] A. Logunov. Nodal sets of laplace eigenfunctions: proof of nadirashvili’s conjecture and of the lower bound in yau’s conjecture. *Annals of Mathematics (2)*, 187(1):241–262, 2018.
- [22] A. Logunov and E. Malinnikova. Nodal sets of laplace eigenfunctions: Estimates of the hausdorff measure in dimensions two and three. In *50 Years with Hardy Spaces: A Tribute to Victor Havin*, volume 261 of *Operator Theory: Advances and Applications*, pages 333–344. Birkhäuser, 2018.
- [23] Domenico Marinucci, Giovanni Peccati, Maurizia Rossi, and Igor Wigman. Non-universality of nodal length distribution for arithmetic random waves. *Geometric and Functional Analysis*, 26:926–960, 2016.
- [24] Fedor Nazarov and Mikhail Sodin. On the number of nodal domains of random spherical harmonics. *American Journal of Mathematics*, 131(5):1337–1357, 2009.
- [25] Fedor Nazarov and Mikhail Sodin. Fluctuations in the number of nodal domains. *Journal of Mathematical Physics*, 61(12), 2020.
- [26] Ivan Nourdin, Giovanni Peccati, and Maurizia Rossi. Nodal statistics of planar random waves. *Communications in Mathematical Physics*, 369:99–151, 2019.
- [27] Zeév Rudnick and Igor Wigman. On the volume of nodal sets for eigenfunctions of the laplacian on the torus. In *Annales Henri Poincare*, volume 9, pages 109–130. Springer, 2008.
- [28] C. Sturm. Mémoire sur les équations différentielles linéaires du second ordre. *Journal de Mathématiques Pures et Appliquées*, 1:106–186, 1836.
- [29] Shing-Tung Yau. Problem section. In *Seminar on Differential Geometry*, volume 102 of *Annals of Mathematics Studies*, pages 669–706. Princeton University Press, Princeton, N.J., 1982.