MATRIX CONCENTRATION INEQUALITIES AND FREE PROBABILITY II. TWO-SIDED BOUNDS AND APPLICATIONS

AFONSO S. BANDEIRA, GIORGIO CIPOLLONI, DOMINIK SCHRÖDER, AND RAMON VAN HANDEL

Abstract. The first paper in this series introduced a new family of nonasymptotic matrix concentration inequalities that sharply capture the spectral properties of very general random matrices in terms of an associated noncommutative model. These methods achieved matching upper and lower bounds for smooth spectral statistics, but only provided upper bounds for the spectral edges. Here we obtain matching lower bounds for the spectral edges, completing the theory initiated in the first paper.

The resulting two-sided bounds enable the study of problems that require an exact determination of the spectral edges to leading order, which is fundamentally beyond the reach of classical matrix concentration inequalities. To illustrate their utility, we develop two general results that explain phase transitions of spectral outliers of a large class of nonhomogeneous random matrices. This enables us to elucidate phase transition phenomena that arise in diverse applications, including decoding node labels on graphs, tensor PCA, contextual stochastic block models, and centered sample covariance matrices.

Contents

2010 *Mathematics Subject Classification.* 60B20; 60E15; 46L53; 46L54; 15B52.
Key words and phrases. Random matrices; matrix concentration inequalities; free probability.

1. Introduction

Let X be any $d \times d$ self-adjoint random matrix with jointly Gaussian entries. What can we say about its spectrum? As we made no assumptions on the mean and covariance of the entries, classical methods of random matrix theory shed little light on this question. Nonetheless, nontrivial bounds on the spectrum are achievable at this level of generality by means of operator-theoretic results that are often referred to as matrix concentration inequalities.

The classical such result, the noncommutative Khintchine inequality [36, 15], estimates any finite moment of a (centered) Gaussian matrix explicitly up to a constant factor in terms of the covariance of its entries. Stated precisely, for any self-adjoint Gaussian random matrix X with $\mathbf{E}X = 0$ we have

$$\operatorname{tr}[(\mathbf{E}X^2)^p]^{\frac{1}{2p}} \le \mathbf{E}[\operatorname{tr} X^{2p}]^{\frac{1}{2p}} \le \sqrt{2p}\,\operatorname{tr}[(\mathbf{E}X^2)^p]^{\frac{1}{2p}} \tag{1.1}$$

for $p \in \mathbb{N}$, where we define the normalized trace $\operatorname{tr} M := \frac{1}{d}\operatorname{Tr} M$ for any $M \in \mathrm{M}_d(\mathbb{C})$. Using the basic fact (see (1.5) below) that for any $M \in \mathrm{M}_d(\mathbb{C})$

$$\operatorname{tr}[|M|^{2p}]^{\frac{1}{2p}} = (1 + o(1))\|M\| \qquad \text{for } p \gg \log d, \tag{1.2}$$

the bound (1.1) also yields upper and lower bounds for the spectral norm $\|X\|$ up to a factor that grows logarithmically with dimension. Much work in the past two decades has been devoted to extending such bounds to a large class of non-Gaussian models, cf. [42] and the references therein.

Due to their generality and ease of use, matrix concentration inequalities have found numerous applications in pure and applied mathematics. At the same time, they can provide only rough bounds on the behavior of the spectrum that are often increasingly inaccurate in high dimension, in contrast to classical results in random matrix theory that become increasingly precise as $d \to \infty$. It has been a long-standing question whether there exist results at the level of generality of matrix concentration inequalities which can nonetheless sharply capture the spectral properties of many random matrix models.

Significant progress in this direction was achieved in the first part of this series [6] (inspired in part by [21, 43]), which introduced a new family of matrix concentration inequalities that optimally capture the behavior of very general Gaussian random matrices to leading order. A key feature of these inequalities is that they do not bound the spectrum of X directly, but rather quantify the deviation of the spectrum of X from that of an associated deterministic operator $X_{\rm free}$ defined by a matrix with entries in a C^*-probability space $(\mathcal{A},\tau)$ (we recall its definition in section 2). For example, [6, Theorem 2.7] yields the inequality

$$|\mathbf{E}[\operatorname{tr} X^{2p}]^{\frac{1}{2p}} - (\operatorname{tr}\otimes\tau)[X_{\rm free}^{2p}]^{\frac{1}{2p}}| \le 2p^{\frac{3}{4}}\tilde{v}(X) \tag{1.3}$$

for all $p\in\mathbb{N}$, where $\tilde{v}(X)^4 := \|\operatorname{Cov}(X)\|\,\|\mathbf{E}[(X-\mathbf{E}X)^2]\|$ and $\operatorname{Cov}(X)$ is the $d^2\times d^2$ covariance matrix of the entries of X. Here X need not be centered, that is, both mean and covariance of X are arbitrary.

Unlike (1.1), which upper and lower bounds the moments of X up to a constant factor, (1.3) computes the moments of X exactly to leading order when $\tilde{v}(X)$ is small. The latter situation is ubiquitous in applications. Such bounds extend also to non-Gaussian models by means of a universality principle [14]. At the same time, tools of free probability theory [21, 27] make it possible to compute or estimate the spectral statistics of $X_{\rm free}$ explicitly in terms of the mean and covariance of X, making (1.3) genuinely applicable to concrete situations.

The inequality (1.3) is just one example of the kind of results that are achieved by the theory of [6]; the same method of proof yields analogous two-sided bounds for many smooth spectral statistics. But arguably the most powerful aspect of this theory lies in its ability to capture the edges of the spectrum, which is considerably more delicate than the bulk spectral behavior. In this regard, however, the theory is incomplete. For example, it is shown in [6, Corollary 2.2] that

$$\mathbf{E}\|X\| \le \|X_{\rm free}\| + C\tilde{v}(X)(\log d)^{\frac{3}{4}} \tag{1.4}$$

for a universal constant C, which yields a sharp *upper* bound on $\|X\|$ whenever $\tilde{v}(X)$ is small. However, unlike in (1.3), the corresponding *lower* bound is missing. The reason for this discrepancy, as explained in [6, §8.2.3], lies in the elementary fact (1.2): while the norm of any $d\times d$ matrix can be approximated by moments of order $\log d$, it is far from clear whether the analogous property holds for the infinite-dimensional operator $X_{\rm free}$. This problem is resolved in this paper, which completes the theory of [6] and opens the door to new applications.

1.1. **Main results.** The simplest implication of the new results of this paper may be readily understood in the context of the above discussion: the following theorem extends the upper bound (1.4) to a two-sided bound.

Theorem 1.1. *For any $d\times d$ random matrix X with jointly Gaussian entries*

$$|\mathbf{E}\|X\| - \|X_{\rm free}\|| \le C\tilde{v}(X)(\log d)^{\frac{3}{4}},$$

where C is a universal constant. When X is self-adjoint, the same inequality holds if $\|X\|, \|X_{\rm free}\|$ are replaced by the upper edge of the spectrum $\lambda_{\max}(X), \lambda_{\max}(X_{\rm free})$.

However, our results are much more general than is suggested by Theorem 1.1. The main result of this paper provides a subgaussian matrix concentration inequality for the Hausdorff distance between the spectra of X and $X_{\rm free}$. This makes it possible both to achieve high probability results, and to detect interior edges of the

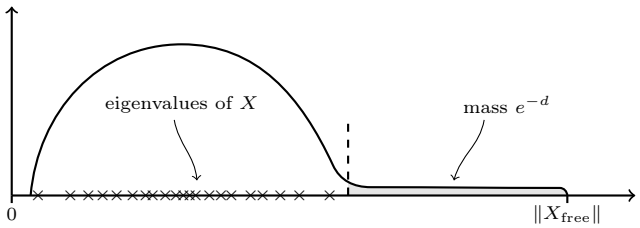

FIGURE 1.1. Illustration of a hypothetical obstruction to the validity of Theorem 1.1. The proof must show that this situation cannot occur.

spectrum in addition to exterior edges. Our results will be stated in full generality in section 2 after we recall the relevant definitions.

As was explained above, the key new ingredient that is needed in the proof of these results is that the norm of X_{free} (more generally, of its resolvent $(z - X_{\text{free}})^{-1}$) is well approximated by its moments. That the moments are upper bounded by the norm is trivial, which is the reason that the upper bound (1.4) was achievable in [6]. The converse direction is far from clear, however.

To understand where the difficulty lies, it is instructive to recall why (1.2) holds for a self-adjoint $d \times d$ matrix M with eigenvalues $\lambda_1(M), \ldots, \lambda_d(M)$: as

$$\operatorname{tr}[|M|^{2p}] = \frac{1}{d}\sum_{i=1}^{d} |\lambda_i(M)|^{2p} \geq \frac{1}{d} \max_i |\lambda_i(M)|^{2p} = \frac{1}{d}\|M\|^{2p}, \tag{1.5}$$

we have $\operatorname{tr}[|M|^{2p}]^{\frac{1}{2p}} \geq d^{-\frac{1}{2p}}\|M\| = (1 + o(1))\|M\|$ for $p \gg \log d$. Thus (1.2) holds because the empirical spectral distribution of M has mass $\frac{1}{d}$ at $\|M\|$. However, it is not clear why the spectral distribution of the infinite-dimensional operator X_{free} should also have large mass near its edges. For example, Figure 1.1 illustrates a hypothetical scenario where the spectral distribution of X_{free} only has mass e^{-d} near its upper edge; in this case, it would be extremely unlikely that any eigenvalue of X (each of which has mass $\frac{1}{d}$) is located near $\|X_{\text{free}}\|$, contradicting Theorem 1.1. Our proof must therefore show that such situations cannot occur.

We have in fact developed two distinct methods of proof to achieve this aim. The first method is based on the work of Alt, Erdős, and Kruger [2], who made a detailed study of the behavior of the spectral distribution of X_{free} near the edges of the spectrum under two strong regularity assumptions: flatness, which requires in particular that all entries of X have variance of the same order, and a uniform bound on $\|\mathbf{E}X\|$. Both assumptions are highly problematic in our setting, as they rule out precisely the kind of nonhomogeneous models that matrix concentration inequalities aim to capture. However, in fixed dimension d, any random matrix can be perturbed with negligible effect on the spectrum so that it satisfies these regularity assumptions with constants that diverge polynomially with d. One can therefore rule out situations as in Figure 1.1 by a combination of spectral perturbation theory and a quantitative refinment of the results of [2]. A proof of our main results by this approach appears in an early draft of this paper [7].

An entirely different approach arises, in a slightly different setting, in the work of Bordenave and Collins [12, §7.1]. The idea introduced there is that the comparison between moments and norm of X_{free} is closely connected to the ultracontractive properties of free operators. Exploiting this idea in the present setting considerably

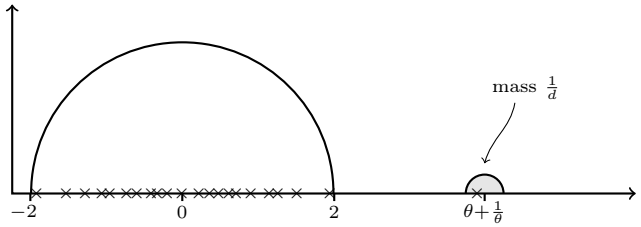

FIGURE 1.2. Spectral distribution of $X_{\rm free}$ for the spiked Wigner model.

shortens the proof of our main results, as it replaces the rather technical work based on [2] by operator-theoretic tools. We therefore present the latter approach in this paper. Despite this simplification, ultracontractivity does not suffice in itself to achieve satisfactory bounds, so that spectral perturbation arguments remain crucial for the proof. Details and further discussion are given in section 4.

While the above discussion focused on Gaussian models, random matrix models that arise in applications are often non-Gaussian. This will not present any additional complications, however, as the theory of this paper extends directly to many non-Gaussian situations using the universality principle of [14].

1.2. **Phase transitions.** The development of inequalities that exactly capture the spectral edges to leading order enables the study of problems that are fundamentally beyond the reach of classical matrix concentration inequalities, such as phase transition phenomena for spectral outliers of nonhomogeneous random matrices. We will develop this theme in some detail, both as a compelling illustration of our general theory and for its independent interest.

The classical study of phase transitions for spectral outliers due to Baik, Ben Arous, and Péché [4] has led to a large body of work, see, e.g., the survey [16]. Let us briefly recall one of the standard models in this area. Let G be a $d\times d$ self-adjoint matrix with i.i.d. standard Gaussian entries above the diagonal, and let

$$X = \theta\, vv^* + G$$

where $\theta \ge 0$ and $\|v\| = 1$. This is the spiked Wigner model. It is classical that the largest eigenvalue of G is $2+o(1)$. For the random matrix X, however, we observe a phase transition: its largest eigenvalue is still $2+o(1)$ when $\theta \le 1$, while an outlier eigenvalue emerges at $\theta + \frac{1}{\theta} + o(1)$ when $\theta > 1$.

While such a sharp transition is clearly inaccessible by classical matrix concentration inequalities, it can be recovered as an easy exercise from Theorem 1.1 using an explicit formula of Lehner (see (1.6) below) for the largest eigenvalue of $X_{\rm free}$.[1] As our theory applies to arbitrarily structured random matrices, however, it enables the study of such phenomena in far more general situations:

- We may replace G by a much more general nonhomogeneous model, including models that are highly sparse or whose entries exhibit strong dependence;
- We may replace $\theta\, vv^*$ by much more general perturbations whose rank may diverge at a rate determined by the fluctuations of the spectral statistics of G;

[1] It may appear surprising that $X_{\rm free}$, which has a continuous spectrum, can detect the presence of a single eigenvalue of X. This paradox is explained in Figure 1.2: the spectral distribution of $X_{\rm free}$ has a small connected component of mass $\frac{1}{d}$ which produces the outlier eigenvalue of X.

- We can obtain fully nonasymptotic results that provide explicit guarantees for the given random matrix X of fixed dimension d.

None of these features are readily accessible by prior work in this area. Most methods that have been used to study outliers of random matrices rely strongly on mean-field (that is, near-homogeneous) random matrix structure [16, 11, 26]. Some results for sparse models have appeared only very recently [40, 3], but rely on restrictive assumptions and make use of specialized tools.

In principle, our theory can be applied to an arbitrarily structured nonohomogeneous spiked model. However, the phase transition behavior will be determined in general by a complicated variational principle that does not have a simple analytic solution. Instead, we will primarily focus our attention on two general classes of models that can be understood in an explicit closed form:

1. Models of the form $X = A + G$ where A has rank $o(d)$ and the noise matrix G is *isotropic*, that is, $\mathbf{E}[G] = 0$ and $\mathbf{E}[G^2] = \mathbf{1}$.
2. A class of explicitly solvable *anisotropic* models due to Krzakala et al. (e.g., [32]) that arise by linearizing message passing algorithms of statistical physics.

These general results are established in strong nonasymptotic form that can be applied in a black box manner in complex situations.

It should be emphasized that the specific features of spiked models are completely irrelevant to our sharp matrix concentration theory: our theory reduces the study of arbitrarily structured random matrices to the question of understanding X_{free}. The difficulty in understanding the above models lies entirely in the latter deterministic question. To this end, we will make fundamental use of a formula of Lehner

$$\lambda_{\max}(X_{\mathrm{free}}) = \inf_{M>0} \lambda_{\max}\big(\mathbf{E}X + M^{-1} + \mathbf{E}[(X-\mathbf{E}X)M(X-\mathbf{E}X)]\big) \tag{1.6}$$

for arbitrary self-adjoint random matrices X (cf. [27, Corollary 1.5] and [6, §4.1]). In our study of spiked models, we will develop methods to capture the structure of this variational principle that may be useful also in other applications.

1.3. **Applications.** As a direct consequence of the general phase transition results described above, we are able to elucidate phase transition phenomena that arise in a diverse set of applied mathematical problems, including:

- We establish the sharp threshold for recovery of node labels on graphs from noisy observations of the edges by a simple spectral method [1, 18], under natural geometric assumptions on the underlying graph.
- We establish the sharp detection threshold of an important family of spectral methods for the tensor PCA problem [29, 44], providing the only sharp phase transition known to date for any tensor PCA algorithm for symmetric tensors.
- We prove a conjecture of Pak, Ko, and Krzakala [32, Conjecture 1.6] on the outlier phase transitions in a block-structured variant of the spiked Wigner model.[2]
- We prove a conjecture of Duranthon, Krzakala, and Zdeborová (personal communication) on the recovery threshold of contextual stochastic block models [19] by means of a spectral algorithm that is based on ideas of statistical physics.

[2]This conjecture is proved here in a strong nonasymptotic form. A weaker asymptotic form of this conjecture was proved independently in [28] concurrently with our work.

These applications will be introduced in detail in section 3. Each of these problems is of considerable interest in its own right and had remained open prior to this work. We emphasize that the theory of the present paper not only makes it possible to resolve such problems, but that each of these diverse applications is deduced with minimal effort from the general results described in section 1.2.

At the same time, problems that do not belong to one of the general classes described in the previous section can still be studied on a case-by-case basis using similar techniques. As an illustration, we will prove a conjecture of Han [22] on a phase transition phenomenon of centered sample covariance matrices.

1.4. **Organization of this paper.** This paper is organized as follows. In section 2, we formulate the main results of this paper: two-sided sharp matrix concentration inequalities for spectral edges, and "master theorems" for the two general classes of spiked models described above. In section 3, we will illustrate the applicability of these results in a diverse range of applied mathematical problems.

The rest of the paper is devoted to the proofs of these results. Section 4 develops the key ultracontractive estimates that are needed in the proofs of two-sided bounds on the spectral edges. The latter will subsequently be proved in section 5. Sections 6 and 7 are devoted to the proofs of the two "master theorems" for spiked models. Finally, section 8 contains proofs of some results needed in the applications.

1.5. **Notation.** The following notation will be used throughout this paper. For a bounded operator X, we denote by $\|X\|$ its operator norm and by $|X| := (X^*X)^{\frac{1}{2}}$ its modulus. If X is self-adjoint, we denote its spectrum as $\mathrm{sp}(X)$, and denote by $\lambda_{\max}(X) := \sup \mathrm{sp}(X)$ and $\lambda_{\min}(X) := \inf \mathrm{sp}(X)$ the upper and lower edges of the spectrum. The identity operator or matrix is denoted as $\mathbf{1}$. The algebra of $d \times d$ matrices with entries in a $*$-algebra $\mathcal{A}$ is denoted as $\mathrm{M}_d(\mathcal{A})$, and its subspace of self-adjoint elements is $\mathrm{M}_d(\mathcal{A})_{\mathrm{sa}}$. For $M \in \mathrm{M}_d(\mathbb{C})$, we denote by $\mathrm{Tr}\, M := \sum_{i=1}^d M_{ii}$ its unnormalized trace and by $\mathrm{tr}\, M := \frac{1}{d}\mathrm{Tr}\, M$ its normalized trace. Finally, we use the convention that when a functional is followed by square brackets, it is applied before any other operations; for example, $\mathbf{E}[X]^\alpha := (\mathbf{E}X)^\alpha$ and $\mathrm{tr}[M]^\alpha := (\mathrm{tr}\, M)^\alpha$.

2. Main results

2.1. **Basic model and matrix parameters.** Throughout this paper, we fix $d \geq 2$ and consider a $d \times d$ random matrix X with jointly Gaussian entries. Such a random matrix can always be represented (for some $n \in \mathbb{N}$) as

$$X = A_0 + \sum_{i=1}^n A_i g_i, \tag{2.1}$$

where $g_1, \ldots, g_n$ are i.i.d. standard Gaussian variables and $A_0, \ldots, A_n \in \mathrm{M}_d(\mathbb{C})$.

To the given random matrix X, we associate a corresponding noncommutative model $X_{\mathrm{free}} \in \mathrm{M}_d(\mathcal{A}) \simeq \mathrm{M}_d(\mathbb{C}) \otimes \mathcal{A}$ defined as

$$X_{\mathrm{free}} = A_0 \otimes \mathbf{1} + \sum_{i=1}^n A_i \otimes s_i, \tag{2.2}$$

where $s_1, \ldots, s_n$ is a free semicircular family in some C^*-probability space $(\mathcal{A}, \tau)$. We refer to [6, §4.1] for a very brief introduction and [30] for a pedagogical treatment of the basic notions of free probability. The main outcome of the sharp matrix

concentration theory of [6] and of the present paper is that, in many situations, the spectrum of X is well approximated by that of X_{free}.

Before we formulate our main results, we recall the definitions of the most common parameters that will appear in our bounds. In the following, we denote by

$$\mathrm{Cov}(X)_{ij,kl} := \mathbf{E}\big[(X-\mathbf{E}X)_{ij}\overline{(X-\mathbf{E}X)_{kl}}\big]$$

the $d^2 \times d^2$ covariance matrix of the entries of X. We now define the parameters

$$\begin{aligned}
\sigma(X)^2 &:= \|\mathbf{E}[(X-\mathbf{E}X)^*(X-\mathbf{E}X)]\| \vee \|\mathbf{E}[(X-\mathbf{E}X)(X-\mathbf{E}X)^*]\|,\\
v(X)^2 &:= \sup_{\mathrm{Tr}\,|M|^2=1} \mathbf{E}[|\mathrm{Tr}[M(X-\mathbf{E}X)]|^2] = \|\mathrm{Cov}(X)\|,\\
\sigma_*(X)^2 &:= \sup_{\|v\|=\|w\|=1} \mathbf{E}[|\langle v,(X-\mathbf{E}X)w\rangle|^2],
\end{aligned}$$

as well as the frequently appearing combination

$$\tilde{v}(X)^2 := v(X)\sigma(X).$$

We emphasize that these parameters depend only on $\mathrm{Cov}(X)$ and not on $\mathbf{E}X = A_0$. All these parameters are readily expressed explicitly in terms of $A_1,\dots,A_n$ and we have $\sigma_*(X) \le \sigma(X)$ and $\sigma_*(X) \le v(X)$, cf. [6, §2.1].

Remark 2.1. Roughly speaking, these parameters will play the following roles in our theory: $\sigma(X)$ controls the scale of the spectrum of $X-\mathbf{E}X$ (e.g., by (1.1)); $v(X)$ controls the degree to which the spectrum of X is approximated by X_{free}; and $\sigma_*(X)$ captures the fluctuations of the spectrum of X.

Finally, let us note that we will often restrict attention in the formulation and proofs of our results to self-adjoint random matrices X (that is, random matrices defined by $A_0,\dots,A_n \in \mathrm{M}_d(\mathbb{C})_{\mathrm{sa}}$). This entails no loss of generality, however, as results on the spectrum of self-adjoint operators extend directly to the singular value spectrum of non-self-adjoint operators by means of a standard dilation argument [6, Remark 2.6]. For this reason, we will formulate some of our main results for self-adjoint matrices whenever this leads to greater notational simplicity.

2.2. **Sharp matrix concentration inequalities.**

2.2.1. *Gaussian random matrices.* Recall that the Hausdorff distance between two subsets $A, B \subseteq \mathbb{R}$ of the real line is defined as

$$\mathrm{d_H}(A,B) := \inf\{\varepsilon > 0 : A \subseteq B + [-\varepsilon,\varepsilon] \text{ and } B \subseteq A + [-\varepsilon,\varepsilon]\}.$$

The following is the central result of this paper.

Theorem 2.2. *For any $d \times d$ self-adjoint Gaussian random matrix X, we have*

$$\mathbf{P}\big[\mathrm{d_H}(\mathrm{sp}(X),\mathrm{sp}(X_{\mathrm{free}})) > C\tilde{v}(X)(\log d)^{\frac{3}{4}} + C\sigma_*(X)t\big] \le e^{-t^2}$$

for all $t \ge 0$, where C is a universal constant.

Theorem 2.2 controls the entire spectrum of X and X_{free}. As the spectral edges are often of special interest, we spell out the following simple corollary.

Corollary 2.3. *Let X be an arbitrary (not necessarily self-adjoint) $d \times d$ Gaussian random matrix. Then we have*

$$\mathbf{P}\big[|\|X\| - \|X_{\mathrm{free}}\|| > C\tilde{v}(X)(\log d)^{\frac{3}{4}} + C\sigma_*(X)t\big] \le e^{-t^2}$$

for all $t \geq 0$ *and*

$$|\mathbf{E}\|X\| - \|X_{\rm free}\|| \leq C\tilde{v}(X)(\log d)^{\frac{3}{4}},$$

where C *is a universal constant. If* X *is self-adjoint, the same inequalities hold if* $\|X\|, \|X_{\rm free}\|$ *are replaced by* $\lambda_{\max}(X), \lambda_{\max}(X_{\rm free})$ *or by* $\lambda_{\min}(X), \lambda_{\min}(X_{\rm free})$.

Theorem 2.2 and Corollary 2.3 will be proved in section 5. Note that Theorem 1.1 in the introduction is merely a special case of Corollary 2.3.

2.2.2. *Non-Gaussian random matrices.* While the above inequalities are formulated for Gaussian random matrices, random matrices that arise in applications are often non-Gaussian. The Gaussian case is nonetheless of central importance, as the behavior of many non-Gaussian matrices can be understood in terms of an associated Gaussian model. For ease of reference, we presently state two general results of this kind that will be used in the applications in section 3.

One widely used non-Gaussian random matrix model is

$$Z = Z_0 + \sum_{i=1}^n Z_i, \tag{2.3}$$

where $Z_0 \in \mathrm{M}_d(\mathbb{C})$ is a nonrandom matrix and $Z_1, \ldots, Z_n$ are arbitrary independent $d \times d$ random matrices with $\mathbf{E}Z_i = 0$. Such models arise naturally in many applications [42, 14]. We presently state a universality principle [14, Theorem 2.6] that reduces the study of such models to the Gaussian case.

Theorem 2.4 ([14])**.** *Let* Z *be a* $d \times d$ *self-adjoint random matrix as in* (2.3)*, and suppose that* $\|Z_i\| \leq R$ *a.s. for* $i = 1, \ldots, n$*. Let* X *be the* $d \times d$ *Gaussian random matrix whose entries have the same mean and covariance as those of* Z*. Then*

$$\mathbf{P}\big[\mathrm{d_H}(\mathrm{sp}(Z), \mathrm{sp}(X)) > C\sigma_*(X)t^{\frac{1}{2}} + CR^{\frac{1}{3}}\sigma(X)^{\frac{2}{3}}t^{\frac{2}{3}} + CRt\big] \leq de^{-t}$$

for all $t \geq 0$*, where* C *is a universal constant.*

Theorem 2.4 shows that Z behaves as a Gaussian random matrix X, while Theorem 2.2 shows that X behaves as its noncommutative model $X_{\rm free}$. The combination of these two theorems therefore provides a powerful tool to study a large class of non-Gaussian random matrices. A variant of Theorem 2.4 that is applicable when Z_i are unbounded may be found in [6, Theorem 2.7].

A non-Gaussian model of a different nature arises in the study of sample covariance matrices, which requires an understanding of the spectra of quadratic polynomials of a Gaussian random matrix X such as $XX^* - \mathbf{E}XX^*$. Such models are captured by the following quadratic analogue of Theorem 2.2.

Theorem 2.5. *Let* X *be any (not necessarily self-adjoint)* $d \times d$ *Gaussian random matrix, and let* $B \in \mathrm{M}_d(\mathbb{C})_{\rm sa}$*. Then we have*

$$\mathbf{P}\big[\mathrm{d_H}(\mathrm{sp}(XX^*+B), \mathrm{sp}(X_{\rm free}X_{\rm free}^*+B\otimes\mathbf{1})) > C\{\|X_{\rm free}\|+\|B\|^{\frac{1}{2}}\}t+Ct^2\big] \leq e^{-\frac{t^2}{\sigma_*(X)^2}}$$

for all $t \geq \tilde{v}(X)(\log d)^{\frac{3}{4}}$*, where* C *is a universal constant.*

Theorem 2.5 will be proved in section 5.

Remark 2.6. Theorem 2.5 can be extended in two directions. On the one hand, we may consider models where X itself is non-Gaussian as in [14, §3.4]. On the other hand, the method of proof of Theorem 2.2 can be adapted to bound general noncommutative polynomials $P(X_1, \ldots, X_m)$ of Gaussian random matrices $X_1, \ldots, X_m$ in

terms of their noncommutative models $P(X_{1,\mathrm{free}}, \ldots, X_{m,\mathrm{free}})$. As such extensions digress from the main theme of this paper, we do not develop them further here.

2.3. **Phase transitions: isotropic case.** The theorems stated above explain the spectral properties of very general random matrices in terms of the noncommutative model X_{free}. To understand specific phenomena, it therefore remains to understand the corresponding behavior of X_{free}. We presently formulate a number of results that enable the study of spectral outliers in a broad range of models.

It will be convenient to define the function

$$\mathrm{B}(\theta) := \begin{cases} 2 & \text{for } \theta \le 1, \\ \theta + \frac{1}{\theta} & \text{for } \theta > 1. \end{cases}$$

As was discussed in section 1.2, this function describes the largest eigenvalue of the classical spiked Wigner model. The following result may be viewed as a far-reaching generalization of this phenomenon.

Theorem 2.7. *Let X be any $d \times d$ self-adjoint random matrix. Suppose that $\mathbf{E}[(X - \mathbf{E}X)^2] = \mathbf{1}$ and that $\mathbf{E}X$ has rank r with $\sigma_*(X)\sqrt{r} \le 1$. Then*

$$|\lambda_{\max}(X_{\mathrm{free}}) - \mathrm{B}(\lambda_{\max}(\mathbf{E}X))| \le 2\sigma_*(X)\sqrt{r}.$$

When combined with sharp matrix concentration inequalities, this yields a phase transition for a broad range of models: the isotropic assumption $\mathbf{E}[(X-\mathbf{E}X)^2] = \mathbf{1}$ holds in many (including sparse or dependent) applications, while $\sigma_*(X)\sqrt{r} = o(1)$ typically allows the rank to grow rapidly with dimension.

Remark 2.8. Let us briefly explain in what sense Theorem 2.7 captures an outlier of the spectrum. Denote by $\mu_X := \frac{1}{d}\sum_{i=1}^d \delta_{\lambda_i(X)}$ the empirical spectral distribution of X. We begin by recalling the deterministic fact (see, e.g., [25]) that

$$\left| \int f \, d\mu_X - \int f \, d\mu_{X-\mathbf{E}X} \right| = |\mathrm{tr}\, f(X) - \mathrm{tr}\, f(X - \mathbf{E}X)| \le \frac{r}{d} \|f'\|_{L^1(\mathbb{R})} = o(1)$$

for any $f \in C_0^\infty(\mathbb{R})$ when $\mathbf{E}X$ has rank $r = o(d)$.

Suppose the relevant matrix parameters are sufficiently small that the spectrum of X is well modelled by that of X_{free}. Then Theorem 2.7 (applied to $X \leftarrow X - \mathbf{E}X$) implies that $\lambda_{\max}(X - \mathbf{E}X) = 2 + o(1)$. Therefore, by the above deterministic fact, a fraction $1 - o(1)$ of the eigenvalues of X are bounded by $2 + o(1)$.

On the other hand, Theorem 2.7 (applied to X itself) shows that X has an eigenvalue at $\mathrm{B}(\lambda_{\max}(\mathbf{E}X)) > 2 + \varepsilon$ when $\lambda_{\max}(\mathbf{E}X) > 1$. This outlier eigenvalue is therefore bounded away from the bulk of the spectrum.

When there is an outlier, it is expected that the eigenvector associated to the largest eigenvalue of X yields information on the eigenvectors of $\mathbf{E}X$. This behavior is readily deduced from Theorem 2.7 by the following device. Here $1_A(M)$ is defined by functional calculus, that is, it is the projection onto the space spanned by the eigenvectors of $M \in \mathrm{M}_d(\mathbb{C})_{\mathrm{sa}}$ with eigenvalues in $A \subseteq \mathbb{R}$.

Theorem 2.9. *Let X be any $d \times d$ self-adjoint random matrix with $\lambda_{\max}(\mathbf{E}X) =: \theta$, and fix $0 < t \le \delta$. Define $X_s := X + s1_{(\theta-\delta,\theta]}(\mathbf{E}X)$, and suppose that for $s \in \{0, \pm t\}$*

$$\mathbf{P}\big[|\lambda_{\max}(X_s) - \mathrm{B}(\lambda_{\max}(\mathbf{E}X_s))| > \varepsilon\big] \le \rho.$$

Then any unit norm eigenvector $v_{\max}(X)$ of X with eigenvalue $\lambda_{\max}(X)$ satisfies

$$\mathbf{P}\left[\left|\langle v_{\max}(X), 1_{(\theta-\delta,\theta]}(\mathbf{E}X)v_{\max}(X)\rangle - \left(1-\frac{1}{\theta^2}\right)_+\right| > t + \frac{2\varepsilon}{t}\right] \le 3\rho.$$

For example, suppose the largest eigenvalue θ of $\mathbf{E}X$ is simple, and that there is a gap of size δ between the largest and second-largest eigenvalues of $\mathbf{E}X$. Then Theorem 2.9 yields $|\langle v_{\max}(\mathbf{E}X), v_{\max}(X)\rangle|^2 = (1-\frac{1}{\theta^2})_+ + o(1)$, provided that the lower-order terms that arise from the sharp matrix concentration inequalities and from Theorem 2.7 (i.e., ε in Theorem 2.9) are $o(\delta)$.

We have deliberately formulated the above results independently of any specific random matrix model so that they can be applied equally easily in the context of either Theorem 2.2 or Theorem 2.4; applications to concrete models will be illustrated in in section 3. The above results are proved in section 6.

Remark 2.10. While we have focused our treatment of outliers on the largest eigenvalue, one may also investigate other outliers in the spectrum using Theorem 2.2. As the above results already suffice for all the applications we will consider, we omit further development of such questions in the interest of space.

2.4. **Phase transitions: an anisotropic model.** In principle, there is nothing special about the isotropic assumption $\mathbf{E}[(X-\mathbf{E}X)^2] = \mathbf{1}$ made in Theorem 2.7: we can establish an analogous phase transition for an arbitrarily structured self-adjoint random matrix X so that $\mathbf{E}X$ has low rank. However, for anisotropic models the phase transition will generally not admit a simple description; it is determined by the solution to the variational problem (1.6), which cannot be expected to yield an analytic solution in the absence of some special structure. We presently discuss a class of anisotropic models where such special structure is present.

To define the model, we fix the following parameters:

- A partition of $[d] = C_1 \sqcup \cdots \sqcup C_q$ into q disjoint sets of size $|C_k| > 1$.
- A matrix $B \in \mathrm{M}_q(\mathbb{R})_{\mathrm{sa}}$ with nonnegative entries.
- A vector $z \in \mathbb{R}^d$ such that $\sum_{i\in C_k} z_i^2 = |C_k|$ for $k = 1, \ldots, q$.

Let $\mathbf{B} \in \mathrm{M}_d(\mathbb{R})_{\mathrm{sa}}$ be the block matrix defined by $\mathbf{B}_{ij} := B_{kl}$ for all $i \in C_k, j \in C_l$. Then we consider the $d \times d$ random matrix X defined as[3]

$$\begin{aligned} X &= \frac{1}{d}\operatorname{diag}(z)\mathbf{B}\operatorname{diag}(z) + X_\varnothing, \\ X_\varnothing &= -\operatorname{diag}\left(\frac{1}{d}\mathbf{B}1_d\right) + G, \end{aligned} \tag{2.4}$$

where G is a $d \times d$ real symmetric random matrix whose entries $(G_{ij})_{i\ge j}$ are independent with $\mathbf{E}[G_{ij}] = 0$ and $\mathbf{E}[G_{ij}^2] = \frac{1+1_{i=j}}{d}\mathbf{B}_{ij}$. To interpret the structure of this model, note that X is a low-rank perturbation of $X_\varnothing$ when $q \ll d$ as $\operatorname{rank}(\mathbf{B}) \le q$. We aim to understand the resulting outlier phase transition.

Remark 2.11. The significance of this model is that random matrices of the form (2.4) arise in applications as a linearization of message passing algorithms of statistical physics. Two such applications are discussed in sections 3.4 and 3.5.

[3]Here $1_d \in \mathbb{R}^d$ is the vector whose entries are all equal to one. As the analysis of this model involves both q- and d-dimensional vectors and matrices, we will indicate the dimension of the ones vector 1_d and of the identity matrix $\mathbf{1}_d$ in subscript to avoid confusion.

In the following, we will assume that the nonnegative matrix B is irreducible. This entails no loss of generality: if B is reducible, then X is block-diagonal and it suffices to consider its irreducible blocks. We denote by $c, b \in \mathbb{R}^q$ the vector with entries $c_k = \frac{|C_k|}{d}$ and the Perron-Frobenius (right) eigenvector $b > 0$ of $B\,\mathrm{diag}(c)$.

We are now ready to formulate the analogue of Theorem 2.7 in the present setting. Such a phase transition was first conjectured in [32] (see section 3.4).

Theorem 2.12. *Let $X, X_\varnothing$ be defined as in (2.4), and suppose all the above assumptions are in force. Then there exist $\lambda, \lambda_\varnothing \in \mathbb{R}$ so that*

$$|\lambda_{\max}(X_{\mathrm{free}}) - \lambda| \le \sqrt{\frac{8\|B1_q\|_\infty}{d}}, \qquad |\lambda_{\max}(X_{\varnothing,\mathrm{free}}) - \lambda_\varnothing| \le \sqrt{\frac{8\|B1_q\|_\infty}{d}},$$

where $\lambda_\varnothing$ satifies

$$\lambda_\varnothing \le 1 - \frac{\min_i b_i}{\max_i b_i}\big(1 - \lambda_{\max}(\mathrm{diag}(c)^{\frac12} B\,\mathrm{diag}(c)^{\frac12})^{\frac12}\big)^2,$$

while λ exhibits the following phase transition.

a. If $\lambda_{\max}(\mathrm{diag}(c)^{\frac12} B\,\mathrm{diag}(c)^{\frac12}) < 1$, then $\lambda_\varnothing = \lambda < 1$.
b. If $\lambda_{\max}(\mathrm{diag}(c)^{\frac12} B\,\mathrm{diag}(c)^{\frac12}) = 1$, then $\lambda_\varnothing = \lambda = 1$.
c. If $\lambda_{\max}(\mathrm{diag}(c)^{\frac12} B\,\mathrm{diag}(c)^{\frac12}) > 1$, then $\lambda_\varnothing < \lambda = 1$.

To interpret this result, note that by the same argument as in Remark 2.8, the bulk of the spectrum of X is bounded by $\lambda_\varnothing$. The largest eigenvalue of X is therefore an outlier precisely when $\lambda_{\max}(\mathrm{diag}(c)^{\frac12} B\,\mathrm{diag}(c)^{\frac12}) > 1$, and when the outlier appears it is always at $\lambda = 1$. Moreover, the bound on $\lambda_\varnothing$ yields an explicitly computable estimate on how far the outlier lies from the bulk, which is essential for the applicability of the result in nonasymptotic situations.

The proof of Theorem 2.12 in section 7 will provide explicit variational expressions for $\lambda, \lambda_\varnothing$ that are not analytically tractable in general. It is a remarkable feature of this model that we can nonetheless describe the phase transition in terms of the explictly computable parameter $\lambda_{\max}(\mathrm{diag}(c)^{\frac12} B\,\mathrm{diag}(c)^{\frac12})$. The proof of this fact requires a number of ideas and tools for analyzing Lehner-type variational principles that may be useful also in other applications.

Remark 2.13. It would be of interest to establish a counterpart of Theorem 2.9 in the present setting, which yields quantitative bounds on the overlap between z and the largest eigenvector of X. While it is rather easy to read off the correct behavior of the overlap from the proof of Theorem 2.12 in the asymptotic setting where q, B, c are fixed and $d \to \infty$ (see section 3.4), it has proved more challenging to obtain an explicitly computable nonasymptotic estimate for the overlap in the present setting. We leave this as an open problem.

3. Applications

3.1. **Simple examples.** For sake of illustration, we begin by spelling out the simplest form of the phase transition phenomenon that arises from section 2.3.

Theorem 3.1. *Let G be any $d \times d$ self-adjoint random matrix with $\mathbf{E}G = 0$ and $\mathbf{E}G^2 = \mathbf{1}$, which either has jointly Gaussian entries or is of the form (2.3). Let $\theta \ge 0$ and $v \in S^{d-1}$. Then $X = \theta\, vv^* + G$ satisfies*

$$\mathbf{P}\big[|\lambda_{\max}(X) - \mathrm{B}(\theta)| > C\varepsilon(t)\big] \le Cde^{-t}$$

for all $t > 0$, and

$$\mathbf{P}\Big[\Big||\langle v, v_{\max}(X)\rangle|^2 - \Big(1 - \tfrac{1}{\theta^2}\Big)_+\Big|^2 > C\varepsilon(t)\Big] \le Cde^{-t}$$

whenever $C\varepsilon(t) \le \theta^2$. Here $\varepsilon(t) = v(G)^{\frac12}(\log d)^{\frac34} + \sigma_*(G)t^{\frac12}$ in the Gaussian case and $\varepsilon(t) = v(G)^{\frac12}(\log d)^{\frac34} + \sigma_*(G)t^{\frac12} + R^{\frac13}t^{\frac23} + Rt$ in the setting of Theorem 2.4.

Proof. The first inequality follows by combining Corollary 2.3, Theorem 2.4, and Theorem 2.7, where we note that $\sigma(X) = \sigma(G) = 1$, $v(X) = v(G)$, $\sigma_*(X) = \sigma_*(G)$ and the parameter R in Theorem 2.4 are independent of θ, and that $\sigma_*(G) \le \tilde v(G)$. The second inequality follows from the first by applying Theorem 2.9 with $\delta = \theta$ and optimizing over the parameter t that appears in its statement. □

As a simple example, let us consider sparse Wigner matrices.

Example 3.2. Let $([d], E)$ be a k-regular graph with d vertices. Then we can define a $d\times d$ self-adjoint random matrix G with $G_{ij} = k^{-\frac12}\eta_{ij}1_{\{i,j\}\in E}$ for $i \ge j$, where η_{ij} are independent random variables such that $\mathbf{E}[\eta_{ij}] = 0$, $\mathbf{E}[|\eta_{ij}|^2] = 1$, $\|\eta_{ij}\|_\infty \le K$. In other words, G is a sparse Wigner matrix with an arbitrary deterministic sparsity pattern that has k nonzero entries in each row and column.

It is readily verified that we have $\mathbf{E}[G] = 0$, $\mathbf{E}[G^2] = \mathbf{1}$, $\sigma_*(G) \le v(G) \lesssim k^{-\frac12}$, and $R \lesssim Kk^{-\frac12}$. Thus choosing $t = (1+c)\log d$ in Theorem 3.1 shows that

$$\lambda_{\max}(\theta\, vv^* + G) = \mathrm{B}(\theta) + o(1), \qquad |\langle v, v_{\max}(\theta\, vv^* + G)\rangle|^2 = \Big(1 - \tfrac{1}{\theta^2}\Big)_+ + o(1)$$

with probability at least $1 - \frac{C}{d^c}$ whenever $k \gg K^2(\log d)^4$.

One very special case of this model is a periodic random band matrix with band width k, which is illustrated in Figure 3.1. In this case, as long as $K = O(1)$, we find that the classical phase transition for the spiked Wigner model (as described in section 1.2) extends to this highly sparse setting as soon as the width of the band grows at least polylogarithmically in the dimension of the matrix. This special case was recently investigated using entirely different methods in [3].

While Theorem 3.1 provides precise information on the largest eigenvalue and eigenvector, it does not in itself explain why this largest eigenvalue is an outlier of the spectrum when $\theta > 1$. The rough explanation given in Remark 2.8 can however readily be made precise in concrete situations such as the present one.

Corollary 3.3. *Consider the setting of Theorem 3.1, and denote by $\lambda_2(X)$ the second largest eigenvalue of X. Then we have for all $t > 0$*

$$\mathbf{P}\big[\lambda_2(X) > 2 + C\varepsilon(t)\big] \le Cde^{-t}.$$

Proof. Let $P = \mathbf{1} - vv^*$. Then the min-max theorem yields

$$\lambda_2(X) \le \lambda_{\max}(PXP) = \lambda_{\max}(PGP) \le \lambda_{\max}(G).$$

The conclusion follows by applying Theorem 3.1 with $\theta = 0$. □

Corollary 3.3 shows that whenever $\varepsilon(t) = o(1)$, at most one eigenvalue of X can exceed $2 + o(1)$. When $\theta > 1$, Theorem 3.1 then implies that the largest eigenvalue of X is simple and is separated from the rest of the spectrum.

Theorem 3.1 could be applied directly or with minimal modifications to various models that appear in applications (both with independent and dependent entries), such as adjacency matrices of nonhomogeneous random graphs [40], stochastic block

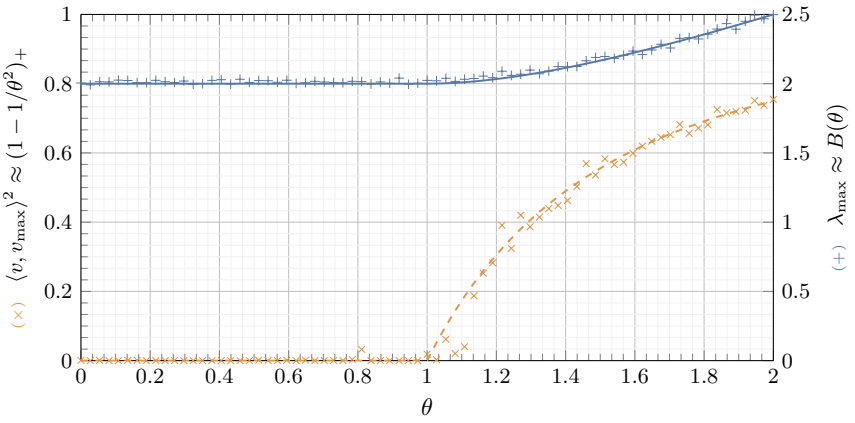

FIGURE 3.1. Outlier phase transition of a 2000×2000 random band matrix with band width 101. The markers represent the empricial result, the lines represent the theoretical prediction of Theorem 3.1.

models [26], or synchronization problems [39, §7]. In the following sections, we will investigate applications that exhibit more complex structures.

3.2. **Decoding node labels on graphs.** The following model is considered in [1]. Let $\Gamma = ([d], E)$ be a given k-regular graph with d vertices, and let $x \in \{-1,+1\}^d$ be an (unknown) binary labeling of the vertices. For each edge $\{i,j\} \in E$, we are given a noisy observation $Y_{ij} = x_i x_j \xi_{ij}$ of the correlation between the labels of the incident vertices, where $\xi_{ij} = \xi_{ji}$ are i.i.d. random variables for $i \geq j$ such that

$$\mathbf{P}[\xi_{ij} = 1] = 1-p, \qquad \mathbf{P}[\xi_{ij} = -1] = p$$

with $p \leq \frac{1}{2}$. The aim is to understand when it is possible to recover the vertex labels (up to a global sign) from the noisy observations.

Here we investigate a simple spectral method for this problem (see, e.g., [18]). Let us augment the above definitions by setting $Y_{ij} = 0$ for $\{i,j\} \notin E$, so that Y defines a $d \times d$ self-adjoint random matrix with independent entries. Note that

$$\mathbf{E}[Y_{ij}] = (1-2p) x_i x_j 1_{\{i,j\} \in E}, \qquad \mathrm{Var}(Y_{ij}) = 4p(1-p) 1_{\{i,j\} \in E}.$$

Thus $\mathbf{E}[(Y - \mathbf{E}Y)^2] = 4kp(1-p)\mathbf{1}$, where we used that Γ is k-regular. We may therefore apply Theorems 2.7 and 2.9 to a suitable modification of Y. In the following result, we will consider a sequence of graphs with $d, k \to \infty$ for simplicity of exposition; a nonasymptotic statement can be read off from the proof.

Theorem 3.4. *Let A be the adjacency matrix of Γ, and denote its singular values as $k = \mathrm{s}_1 \geq \mathrm{s}_2 \geq \cdots \geq \mathrm{s}_d$ and its spectral gap as $\lambda = k - \lambda_2(A)$. Parametrize the error probability as $p = \frac{1}{2} - \frac{1}{2} k^{-\frac{1}{2}} \theta$ with $0 \leq \theta \ll \sqrt{k}$. If*

$$\min_{1 \leq r \leq k} \big\{ \theta\, \mathrm{s}_{r+1} + \sqrt{rk} \big\} + k^{\frac{5}{6}} (\log d)^{\frac{2}{3}} \ll \min\{\theta\lambda, k\},$$

then

$$\frac{1}{d} |\langle x, v_{\max}(Y)\rangle|^2 = \Big(1 - \tfrac{1}{\theta^2}\Big)_+ + o(1)$$

with probability $1 - o(1)$. In particular, there exists an estimator $\hat{x}(Y) \in \{-1,+1\}^d$ so that $\frac{1}{d}|\langle x, \hat{x}(Y)\rangle| > \delta + o(1)$ for some $\delta > 0$ as soon as $\theta \geq 1 + \varepsilon$ for some $\varepsilon > 0$.

The proof of this result is given in section 8.1. The basic idea of the proof is approximate $\mathbf{E}Y$ by a matrix of rank r. We then apply Theorem 2.7 to the low-rank part of the model, and optimize the resulting bound over r to trade off between the width of the phase transition and the approximation error.

Let us note that the recovery condition $\theta > 1$ in Theorem 3.4 is the best possible in general: for example, when Γ is the complete graph, no estimator can recover a nontrivial fraction of the vertex labels when $\theta < 1$ (this follows, e.g., from [34, Theorem 6.3]). A key feature of Theorem 3.4 is that it enables us to achieve this recovery threshold for a large class of deterministic graphs. An analogous problem for Erdős-Rényi graphs was previously considered in [38].

Let us give two examples to illustrate the assumptions of Theorem 3.4.

Example 3.5 (Good expanders). Suppose that $\mathrm{s}_2(A) \le c\sqrt{k}$; this is the case, for example, when Γ is a random k-regular graph [41]. Under mild conditions, such graphs achieve the largest possible spectral gap by the Alon-Boppana theorem [31]. In this situation, we can choose $r \leftarrow 1$ in the assumption of Theorem 3.4. The conclusion of the theorem then follows for any $1 \le \theta \ll \sqrt{k}$ and $k \gg (\log d)^4$.

Example 3.6 (Graphs of intermediate degree). The previous example considered expanders, that is, graphs whose spectral gap λ is of order k. However, expansion is not necessary for the conclusion to hold when the degree k is sufficiently large, as we will presently illustrate. Note first that we can trivially estimate

$$\sum_{i=1}^{d} \mathrm{s}_i^2 = \mathrm{Tr}[A^2] = dk,$$

which implies $\mathrm{s}_i \le (\frac{dk}{i})^{\frac{1}{2}}$. Assuming $\theta \ge 1$ for simplicity, we can estimate

$$\min_{1\le r\le k} \bigl\{\theta\, \mathrm{s}_{r+1} + \sqrt{rk}\bigr\} \le \sqrt{k} \min_{1\le r\le k} \bigl\{r^{-\frac{1}{2}}\theta\sqrt{d} + r^{\frac{1}{2}}\bigr\} \lesssim d^{\frac{1}{4}}\sqrt{\theta k}$$

when $k \ge \theta\sqrt{d}$. Then the conclusion of Theorem 3.4 holds whenever

$$d^{\frac{1}{4}}\sqrt{\theta k} + k^{\frac{5}{6}}(\log d)^{\frac{2}{3}} \ll \min\{\theta\lambda, k\}.$$

For example, if $k \sim d^a$ for some $\frac{1}{2} < a \le 1$, then the conclusion of Theorem 3.4 holds whenever $1 \le \theta \ll d^{a-\frac{1}{2}}$ and $\frac{\lambda}{k} \gg \max\{\theta^{-1}d^{-\frac{a}{6}}(\log d)^{\frac{2}{3}}, \theta^{-\frac{1}{2}}d^{-\frac{1}{2}(a-\frac{1}{2})}\}$.

It is instructive to note that the spectral gap λ appears in Theorem 3.4 only in order to resolve the top eigenvector of Y. This is necessary: for example, if Γ were not connected, then the distribution of Y would be unchanged if we flip the signs of all vertex labels in one connected component, so that it is fundamentally impossible to recover x up to a global sign. On the other hand, if we were interested only in detecting the presence of an outlier eigenvalue in the spectrum of Y, no assumption on the spectral gap would be needed in the proof.

3.3. **Tensor PCA.** The following may be viewed as an analogue of the classical spiked Wigner model for tensors of order p. Fix $\lambda > 0$, a signal $x \in \{-1, +1\}^n$, and i.i.d. standard Gaussian variables $(Z_S)_{S\subseteq[n]:|S|=p}$. We are given a noisy observation tensor $Y = \lambda\, x^{\otimes p} + Z$; more precisely, we observe

$$Y_S := \lambda X_S + Z_S$$

for all $S \subseteq [n]$ with $|S| = p$, where $X_S := \prod_{i\in S} x_i$. This is the tensor PCA model, cf. [29, 44] and the references therein. The key problems in this context are

detecting whether a signal is present, and recovering the signal. In the interest of space we focus on detection, though recovery may be similarly investigated.

We presently describe a general family of spectral methods for the tensor PCA problem that was proposed in [44]. Let $p \geq 4$ be even, fix an integer $\ell \in [\frac{p}{2}, n - \frac{p}{2}]$, and define a symmetric $\binom{n}{\ell} \times \binom{n}{\ell}$ random matrix $M = (M_{S,T})_{S,T \subseteq [n] : |S| = |T| = \ell}$ as

$$M_{S,T} := \begin{cases} Y_{S \triangle T} & \text{if } |S \triangle T| = p, \\ 0 & \text{otherwise,} \end{cases}$$

where $\triangle$ denotes the symmetric difference. The presence of a signal is then detected by the presence of an outlier eigenvalue in the spectrum of M. It is shown in [44] that correct detection of the presence or absence of a signal with probability $1 - o(1)$ is achieved by this method when $\lambda \gg n^{-\frac{p}{4}} \sqrt{\log n}$.

Here we achieve a much more precise understanding of this method for a certain range of the design parameter ℓ of the detection algorithm.

Theorem 3.7. *Fix $\frac{p}{2} \leq \ell < \frac{3p}{4}$ and $\alpha > 0$, and define*

$$k_* := \binom{\ell}{p/2} \binom{n - \ell}{p/2}.$$

Then there is a constant C that depends only on p, ℓ, α so that

$$\mathbf{P}\big[|\lambda_{\max}(k_*^{-\frac{1}{2}} M) - \mathrm{B}(\lambda k_*^{\frac{1}{2}})| > C n^{-1} \lambda k_*^{\frac{1}{2}} + C n^{\frac{4\ell - 3p}{4}} + C n^{\frac{\ell - p}{4}} (\log n)^{\frac{3}{4}}\big] \leq \frac{C}{n^\alpha}.$$

In particular, for any $\varepsilon > 0$, the test

$$f(M) := \begin{cases} 1 & \text{if } \lambda_{\max}(k_*^{-\frac{1}{2}} M) > 2 + n^{-\frac{1}{5}}, \\ 0 & \text{otherwise,} \end{cases}$$

has $\mathbf{P}[f(M) = 1] = o(1)$ if $\lambda \leq k_*^{-\frac{1}{2}}$ and $\mathbf{P}[f(M) = 1] = 1 - o(1)$ if $\lambda \geq (1 + \varepsilon) k_*^{-\frac{1}{2}}$.

The order $\lambda \sim k_*^{-\frac{1}{2}} \sim c(p, \ell) n^{-\frac{p}{4}}$ of the signal strength is believed to be the weakest that can be detected by computationally efficient algorithms, cf. [23, 44]. To the best of our knowledge, however, Theorem 3.7 is the first result to establish a sharp phase transition for any tensor PCA algorithm for symmetric tensors.[4]

The proof of Theorem 3.7 is given in section 8.2. As in the previous section, the idea of the proof is to approximate $\mathbf{E}M$ by a low rank matrix. However, in the present case the entries of M exhibit a complicated dependence structure, which is nonetheless captured effortlessly by our main results.

Remark 3.8. The design parameter ℓ provides a tradeoff between computational cost and the detection threshold: the larger ℓ, the more costly is the computational method (as the dimension of M is of order n^ℓ) but the smaller is the detection threshold $c(p, \ell)$. It is conjectured in [44, Conjecture 3.6] that an arbitrarily small detection threshold $c(p, \ell)$ can be achieved by choosing ℓ sufficiently large. This regime is not captured by Theorem 3.7, however, as its validity is restricted to the range $\frac{p}{2} \leq \ell < \frac{3p}{4}$. When ℓ is large compared to p, the dependence structure of M is so strong that it is unclear whether it could be accurately modeled by M_{free}.

[4] For the asymmetric analogue of the tensor PCA model, a sharp transition was established in [29, §3.2] . This case is considerably simpler, as the natural counterpart of M with $\ell = \frac{p}{2}$ has independent entries and thus its analysis reduces to that of the classical spiked Wigner model.

Nonetheless, even the detection threshold achieved by Theorem 3.7 for $\ell = \frac{p}{2}$ is already of smaller order than is captured by the analysis of [44] for any value of ℓ.

Remark 3.9. The matrix M used here is known as a Kikuchi matrix. Such matrices have had a number of unexpected applications in recent years, such as to the study of Moore bounds for hypergraphs. We refer to [24] for more on this topic.

3.4. **Spike detection in block-structured models.** The above applications feature various nonhomogeneous but isotropic models. We now study an anisotropic model proposed by Pak, Ko, and Krzakala [32].

Let $x \in \{-1,+1\}^d$, and let H be a $d\times d$ self-adjoint random matrix whose entries $(H_{ij})_{i\ge j}$ are independent with $H_{ij} \sim N(0, \frac{1+1_{i=j}}{d}\mathbf{\Delta}_{ij})$; here $\mathbf{\Delta}_{ij} = \mathbf{\Delta}_{ji} \ge 0$ define an arbitrary variance profile of the entries of H. Then

$$\tilde{X} = \frac{1}{d}\,xx^* + H$$

is an anisotropic variant of the spiked Wigner model of section 1.2. The question is when it is possible to detect the presence of the spike xx^*, and to recover the entries of x, when we can only observe $\tilde{X}$. One may expect that this question can be addressed using the largest eigenvalue and eigenvector of $\tilde{X}$ as in section 3.1. Unfortunately, the variational principle (1.6) is generally not analytically tractable for anisotropic models. More surprisingly, the detection threshold achieved by this method turns out to be information-theoretically suboptimal [20].

Instead, [32] propose to consider the largest eigenvalue and eigenvector of a deterministic transformation of $\tilde{X}$ that is motivated by statistical physics:

$$X = \frac{1}{\mathbf{\Delta}} \odot \tilde{X} - \operatorname{diag}\left(\frac{1}{d\mathbf{\Delta}}1_d\right),$$

where $\frac{1}{\mathbf{\Delta}}$ denotes the elementwise inverse and $\odot$ denotes elementwise (Hadamard) product. This procedure can be implemented provided all variances $\mathbf{\Delta}_{ij} > 0$ are positive and known. It is conjectured in [32, Conjecture 1.6] that this approach achieves the optimal detection threshold when the variance profile $\mathbf{\Delta}$ has block structure. Here we prove this conjecture in a strong form.

In the following theorem, we denote by $X_\varnothing$ the null model associated to X, that is, the model where we replace $x \leftarrow 0$ in the definition of X.

Theorem 3.10. *Let $q \ge 1$, let Δ be a $q\times q$ self-adjoint matrix with positive entries, and let $C_1 \sqcup \cdots \sqcup C_q$ be a partition of $[d]$ into q disjoint sets of size $|C_k| =: c_k d > 1$. Assume the variance profile $\mathbf{\Delta}$ is defined by $\mathbf{\Delta}_{ij} = \Delta_{kl}$ for $i \in C_k, j \in C_l$. Let*

$$\mathrm{SNR}(\Delta) := \lambda_{\max}\big(\operatorname{diag}(c)^{\frac12}\tfrac{1}{\Delta}\operatorname{diag}(c)^{\frac12}\big).$$

Then there exists $\mu \le 1 - \kappa\,\big(1-\mathrm{SNR}(\Delta)^{\frac12}\big)^2$ so that the following hold.

a. If $\mathrm{SNR}(\Delta) > 1$, we have with probability $1-e^{-d^{\frac12}}$

$$|\lambda_{\max}(X) - 1| \le C\beta^{\frac12}\Big(\tfrac{(\log d)^{\frac34}}{d^{\frac14}} + \tfrac{q^{\frac12}}{d^{\frac12}}\Big), \quad |\lambda_{\max}(X_\varnothing) - \mu| \le C\beta^{\frac12}\Big(\tfrac{(\log d)^{\frac34}}{d^{\frac14}} + \tfrac{q^{\frac12}}{d^{\frac12}}\Big).$$

b. If $\mathrm{SNR}(\Delta) \le 1$, we have with probability $1-e^{-d^{\frac12}}$

$$|\lambda_{\max}(X) - \mu| \le C\beta^{\frac12}\Big(\tfrac{(\log d)^{\frac34}}{d^{\frac14}} + \tfrac{q^{\frac12}}{d^{\frac12}}\Big), \quad |\lambda_{\max}(X_\varnothing) - \mu| \le C\beta^{\frac12}\Big(\tfrac{(\log d)^{\frac34}}{d^{\frac14}} + \tfrac{q^{\frac12}}{d^{\frac12}}\Big).$$

Here $\beta := \max_{i,j}\frac{1}{\Delta_{ij}}$, $\kappa := \frac{\min_i b_i}{\max_i b_i}$ where b is the Perron eigenvector of $\frac{1}{\Delta}\operatorname{diag}(c)$.

The proof of this result in section 8.3 is a straightforward consequence of the fact that X is of the form (2.4) with $\mathbf{B} = \frac{1}{\mathbf{\Delta}}$ and $z = x$.

Theorem 3.10 shows that the largest eigenvalue of X can detect the presence or absence of a signal with probability $1 - o(1)$ when $\mathrm{SNR}(\Delta) \geq 1 + \varepsilon$ for any $\varepsilon > 0$, which coincides with the information-theoretic detection limit for this problem [20]. We emphasize that this is established here in a much stronger *nonasymptotic* form than was conjectured in [32] (and proved in [28] concurrently with our work). In particular, the conclusion is valid when $q \ll d$ and $\beta, \frac{1}{\kappa} \lesssim 1$, which means the number of blocks may be chosen to diverge rapidly as the dimension increases.

Remark 3.11. In [32, Conjecture 1.6], the largest eigenvalue of X is conjectured to detach from the bulk of the spectrum if and only if $\mathrm{SNR}(\Delta) > 1$. This formulation must be interpreted with care in the nonasymptotic setting, however, as it is possible when both $d, q \to \infty$ that even the spectrum of $X_\varnothing$ has components that detach from the bulk. The formulation of Theorem 3.10 avoids this pitfall.

Remark 3.12. We assumed that $x \in \{-1, 1\}^d$ primarily for simplicity of exposition. In [32], it is assumed instead that the entries of x are i.i.d. with $\mathbf{E}[x_i^2] = 1$. It is completely straightforward to extend Theorem 3.10 to this setting, see Remark 8.5. However, the quantitative rates in Theorem 3.10 must then depend on what assumptions are made on the distribution of x_i. As no new insights are obtained from such an extension, we have chosen to focus on the above concrete setting.

Beside the behavior of the top eigenvalue, [32] also conjectured a corresponding phase transition for the overlap between x and the largest eigenvector of X. As was explained in Remark 2.13, it has proved challenging to achieve nonasymptotic bounds for this quantity. However, in the asymptotic setting where all the model parameters are fixed as $d \to \infty$, the correct behavior may be readily established.

Theorem 3.13. *Consider the setting of Theorem 3.10, where q, Δ, c are taken to be fixed as $d \to \infty$. Then we have*

$$\frac{1}{d}|\langle x, v_{\max}(X)\rangle|^2 \to 0 \text{ in probability as } d \to \infty$$

if and only if $\mathrm{SNR}(\Delta) \leq 1$.

3.5. **Contextual stochastic block models.** We now discuss an entirely different anisotropic application: a spike detection problem with side information.

Let G be an $n \times n$ self-adjoint random matrix whose entries $(G_{ij})_{i \geq j}$ are independent with $G_{ij} \sim N(0, \frac{1+1_{i=j}}{n})$, H be a $p \times n$ random matrix all of whose entries are i.i.d. with distribution $N(0, \frac{1}{p})$, and $u \sim N(0, \frac{1}{p}\mathbf{1}_p)$ be an independent random vector. We further fix $v \in \{-1, +1\}^n$ and $\lambda, \mu \geq 0$. We now define

$$A = \frac{\lambda}{n} vv^* + G, \qquad Y = \sqrt{\frac{\mu}{n}}\, uv^* + H.$$

We aim to recover the signal v from observation of both A and Y. Note that A is precisely the classical spiked Wigner model, while Y provides additional information or "context" about v. The availability of side information should make it possible to detect weaker signals than is possible otherwise.

The above model was proposed in [19, eq. (8)–(9)] as a Gaussian counterpart of an analogous discrete model called the contextual stochastic block model. In particular, it is shown in the proof of [19, Theorem 4] by an indirect argument that

the existence of any nontrivial estimator of v in the Gaussian model will imply the existence of such an estimator in the discrete model. We will therefore focus our attention here for simplicity on the Gaussian model.

In [19, Theorem 6], the authors propose a detection algorithm that combines a spectral method with a univariate optimization problem. Here we investigate a simpler detection algorithm that is purely spectral in nature, which was suggested by O. Duranthon, F. Krzakala, and L. Zdeborová (personal communication) based on ideas of statistical physics. To define this algorithm, let

$$\hat{X} = \begin{bmatrix} \lambda A - (\lambda^2 + \frac{\mu p}{n})\mathbf{1}_n & Y^* \sqrt{\frac{\mu p}{n}} \\ Y\sqrt{\frac{\mu p}{n}} & -\mu \mathbf{1}_p \end{bmatrix},$$

and let $\hat{v} \in \mathbb{R}^n$ be the restriction of $v_{\max}(\hat{X}) \in \mathbb{R}^{n+p}$ to the first n coordinates. The following theorem shows that $\hat{v}$ has positive correlation with v up to the information-theoretic detection limit for this problem (cf. [19, Theorem 6]).

Theorem 3.14. *Let λ, μ be fixed as $n, p \to \infty$ with $\frac{n}{p} \to \gamma \in (0, \infty)$. Then*

$$\frac{1}{n}|\langle v, \hat{v}\rangle|^2 \geq \varepsilon - o(1) \text{ with probability } 1 - o(1) \text{ for some } \varepsilon > 0$$

if and only if $\lambda^2 + \frac{\mu^2}{\gamma} > 1$.

The proof of this result is given in section 8.4. The main idea of the proof is that we can identify $\hat{X}$ as another special instance of the model (2.4). The model differs from that of the previous section in that the matrix $\mathbf{B}$ now has many vanishing entries; this will, however, not cause any complications in our analysis.

Remark 3.15. Our analysis could be extended using Theorem 2.4 to achieve the same result directly for the discrete contextual stochastic block model, provided that its average degrees grow polylogarithmically in the dimension. The latter is necessary, as direct spectral methods fundamentally cannot work in the regime of constant average degrees due to the presence of high degree nodes (see, e.g., [8]). This does not contradict the indirect universality argument used in [19], which merely ensures the existence of an estimator in the discrete setting.

3.6. **Sample covariance error.** Let $X_1, \ldots, X_n$ be i.i.d. Gaussian random vectors in $\mathbb{R}^p$ with distribution $N(0, \Sigma)$. Then the sample covariance matrix

$$\hat{\Sigma} := \frac{1}{n}\sum_{i=1}^n X_i X_i^* \tag{3.1}$$

is a natural estimator of the covariance matrix Σ. If X is the $p \times n$ random matrix whose columns are $X_1, \ldots, X_n$, we may write $\hat{\Sigma} = \frac{1}{n} X X^*$.

Sample covariance matrices exhibit outlier phase transitions much like in the spiked Wigner model when the covariance matrix Σ is defined by a low-rank perturbation. This is in fact the original setting studied by Baik et al. [4, 5]. For sake of illustration, we will consider here the simplest such model where

$$\Sigma = \lambda\, vv^* + \mathbf{1}_p, \tag{3.2}$$

where $\lambda \geq 0$ and $\|v\| = 1$. In this case, it is shown in [4, 5] that in the asymptotic regime $n, p \to \infty$ with $\frac{p}{n} \to \delta$, the largest eigenvalue of $\hat{\Sigma}$ is a spectral outlier if and only if $\lambda > \sqrt{\delta}$ (cf. Theorem 3.16 below).

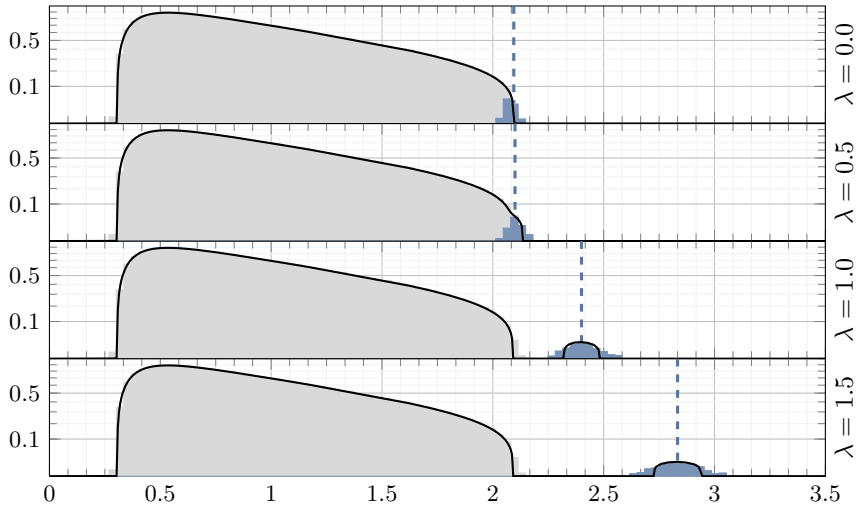

FIGURE 3.2. Illustration of 200 samples of $\hat{\Sigma}$ in Theorem 3.16 with $p = 400$ and $n = 2000$ (so that $\sqrt{\delta} \approx 0.45$). The light shaded area is the empirical histogram of all eigenvalues normalized to have total area 1, while the colored shaded area is the empirical histogram of the largest eigenvalue normalized to have area $\frac{1}{p}$. The solid line is the spectrum of the free model, and the dashed vertical line marks $\mathrm{S}(\lambda, \delta)$. The vertical axis follows a square-root scale to visualize the density of the outlier.

If we view $\hat{\Sigma}$ as an estimator of Σ, however, then it is often more relevant to understand the norm of the estimation error $\|\hat{\Sigma} - \Sigma\|$, rather than the norm of the estimator itself $\|\hat{\Sigma}\|$ as in [4, 5]. It was conjectured by Han [22] that the sample covariance error also exhibits a phase transition; however, this transition occurs at a larger value of the signal strength $\lambda > 1 + \sqrt{\delta}$.

We will prove a nonasymptotic form of the above phase transitions and of yet another transition for the smallest eigenvalue of $\hat{\Sigma} - \Sigma$. Here we define

$$\mathrm{S}(\lambda, \delta) := \begin{cases} (1+\sqrt{\delta})^2 & \text{for } \lambda \le \sqrt{\delta}, \\ (1+\lambda)(1+\frac{\delta}{\lambda}) & \text{for } \lambda > \sqrt{\delta}, \end{cases}$$

$$\mathrm{H}_+(\lambda, \delta) := \begin{cases} \delta + 2\sqrt{\delta} & \text{for } \lambda \le 1+\sqrt{\delta}, \\ \frac{1+\lambda}{2\lambda}(\sqrt{\delta} + \sqrt{\delta+4\lambda})\sqrt{\delta} & \text{for } \lambda > 1+\sqrt{\delta}, \end{cases}$$

$$\mathrm{H}_-(\lambda, \delta) := \begin{cases} \delta - 2\sqrt{\delta} & \text{for } \lambda \le 1-\sqrt{\delta}, \\ \frac{1+\lambda}{2\lambda}(\sqrt{\delta} - \sqrt{\delta+4\lambda})\sqrt{\delta} & \text{for } \lambda > 1-\sqrt{\delta}. \end{cases}$$

These functions play the same role for $\hat{\Sigma}$ and $\hat{\Sigma} - \Sigma$ as does $\mathrm{B}(\theta)$ in section 2.3. The following theorem is illustrated in Figures 3.2 and 3.3.

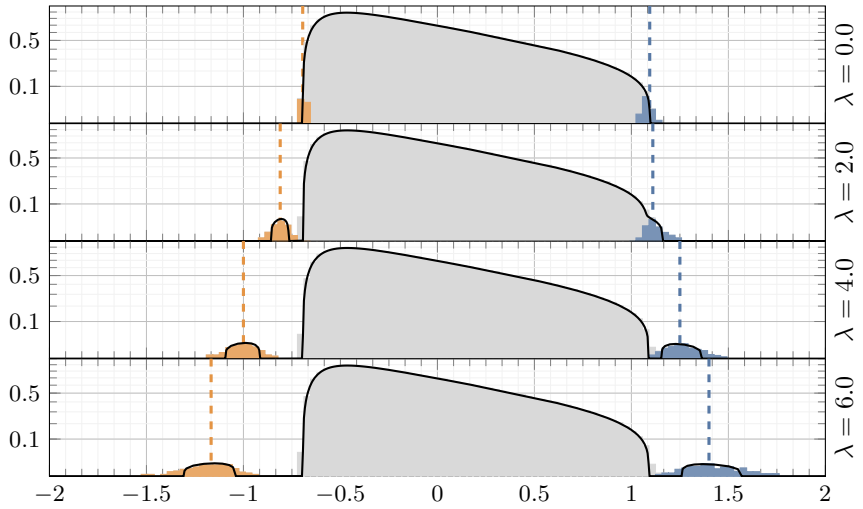

FIGURE 3.3. Illustration of 200 samples of $\hat{\Sigma} - \Sigma$ in Theorem 3.16 with the same parameters as in Figure 3.2. Here the two colored shaded areas are the empirical histograms of the smallest and largest eigenvalues, and the dashed vertical lines mark the locations of $\mathrm{H}_{\pm}(\lambda, \delta)$.

Theorem 3.16. *Let $\delta = \frac{p}{n}$ and $d = \max\{n, p\}$, and assume $n \geq (\log d)^3$. Then*

$$\mathbf{P}\big[\big|\|\hat{\Sigma}\| - \mathrm{S}(\lambda,\delta)\big| > C(1+\lambda+\delta)\, n^{-\frac{1}{4}} (\log d)^{\frac{3}{4}}\big] \leq e^{-Cn^{\frac{1}{2}}},$$

$$\mathbf{P}\big[\big|\lambda_{\max}(\hat{\Sigma} - \Sigma) - \mathrm{H}_{+}(\lambda,\delta)\big| > C(1+\lambda+\delta)\, n^{-\frac{1}{4}} (\log d)^{\frac{3}{4}}\big] \leq e^{-Cn^{\frac{1}{2}}},$$

$$\mathbf{P}\big[\big|\lambda_{\min}(\hat{\Sigma} - \Sigma) - \mathrm{H}_{-}(\lambda,\delta)\big| > C(1+\lambda+\delta)\, n^{-\frac{1}{4}} (\log d)^{\frac{3}{4}}\big] \leq e^{-Cn^{\frac{1}{2}}}.$$

Remark 3.17. Note that as $|\mathrm{H}_{-}(\lambda,\delta)| \leq \mathrm{H}_{+}(\lambda,\delta)$, Theorem 3.16 also implies that $\|\hat{\Sigma} - \Sigma\| = (1 + o(1))\mathrm{H}_{+}(\lambda,\delta)$ with probability $1 - o(1)$ as conjectured in [22].

The proof of Theorem 3.16 is given in section 8.5. We will use Theorem 2.5 to reduce the problem to the analysis of its deterministic model, which can be performed using a quadratic counterpart of the Lehner formula [33].

4. Ultracontractive bounds

The aim of this section is to prove that suitable analogues of the elementary fact (1.2) for matrices with scalar entries hold for matrices whose entries are polynomials of semicircular variables. In the following, we let $s_1, \ldots, s_n$ be a free semicircular family, and denote by $\|Z\|_{L^p(\tau)} := \tau(|Z|^p)^{\frac{1}{p}}$ the noncommutative L^p-norm.

The following may be viewed as a direct analogue of the scalar case.

Theorem 4.1. *Let $P \in \mathrm{M}_d(\mathbb{C}) \otimes \mathbb{C}\langle x_1, \ldots, x_n\rangle$ be any noncommutative polynomial of degree k with matrix coefficients. Then we have*

$$\|P(s_1, \ldots, s_n)\| \leq d^{\frac{3}{4q}} (2qk+1)^{\frac{3}{4q}} \|P(s_1, \ldots, s_n)\|_{L^{4q}(\mathrm{tr}\otimes\tau)}$$

for every $q \in \mathbb{N}$.

However, we will require such a property not for the norm of the matrix itself, but rather for the norm of its resolvent. For the resolvent of a self-adjoint polynomial, we can deduce an analogous result up to a small error.

Theorem 4.2. *Let $P \in \mathrm{M}_d(\mathbb{C}) \otimes \mathbb{C}\langle x_1, \ldots, x_n \rangle$ be any self-adjoint noncommutative polynomial of degree k with matrix coefficients. Then we have*

$$\|(z - P(s_1, \ldots, s_n))^{-1}\| \le d^{\frac{3}{4q}} (2qkr + 1)^{\frac{3}{4q}} \left(\|(z - P(s_1, \ldots, s_n))^{-1}\|_{L^{4q}(\mathrm{tr} \otimes \tau)} + \frac{12 r^{-1} \|P(s_1, \ldots, s_n)\|}{(\operatorname{Im} z)^2} \right)$$

for every $q, r \in \mathbb{N}$ and $z \in \mathbb{C}$, $\operatorname{Im} z > 0$.

Of primary interest in this paper is the noncommutative model X_{free} associated to a self-adjoint Gaussian random matrix X. It is evident from the definition (2.2) that X_{free} is a self-adjoint noncommutative polynomial of degree $k = 1$ with matrix coefficients. However, while Theorem 4.2 can be applied directly to X_{free} as a special case, this result is not adequate for our purposes because the error term in the resulting inequality is proportional to $\|X_{\mathrm{free}}\|$. Such a bound would give rise to sharp matrix concentration inequalities that become increasingly inaccurate as $\|\mathbf{E}X\| = \|A_0\| \to \infty$, which is unnatural in applications.

The following bound, which eliminates the dependence of the error term on A_0, is essential for the main results of this paper. Its proof is based on the fact that the resolvent $(z - X_{\mathrm{free}})^{-1}$ primarily captures the part of the spectrum of X_{free} that is close to $\operatorname{Re} z$, and is insensitive to eigenvalues of A_0 that are far from $\operatorname{Re} z$.

Theorem 4.3. *Let X_{free} be the noncommutative model associated to a self-adjoint random matrix X. For any $z \in \mathbb{C}$ with $0 < \operatorname{Im} z \le \sigma(X)$ and $q, r \in \mathbb{N}$, $K > 4$*

$$\|(z - X_{\mathrm{free}})^{-1}\| \le d^{\frac{3}{4q}} (2qr + 1)^{\frac{3}{4q}} \left(\|(z - X_{\mathrm{free}})^{-1}\|_{L^{4q}(\mathrm{tr} \otimes \tau)} + \left(\frac{K}{r} + \frac{1}{K - 2} \right) \frac{18 d \sigma(X)}{(\operatorname{Im} z)^2} \right).$$

We will apply this theorem in the following form.

Corollary 4.4. *Let X_{free} be the noncommutative model associated to a self-adjoint random matrix X. For any $z \in \mathbb{C}$ with $\operatorname{Re} z \in \operatorname{sp}(X_{\mathrm{free}})$ and $\operatorname{Im} z > 0$, we have*

$$\|(z - X_{\mathrm{free}})^{-1}\| \le C \left(\|(z - X_{\mathrm{free}})^{-1}\|_{L^{4q}(\mathrm{tr} \otimes \tau)} + \frac{\sigma_*(X)}{(\operatorname{Im} z)^2} \right)$$

for all $q \ge \log d$, where C is a universal constant.

The remainder of this section is devoted to the proofs of these results.

4.1. **Proof of Theorems 4.1 and 4.2.** The proof of Theorem 4.1 is essentially equivalent to that of [12, Corollary 7.2]. The basic tool we will use is ultracontractivity, due in the present setting to Bozejko and Biane [10].

Lemma 4.5. *Let $P \in \mathbb{C}\langle x_1, \ldots, x_n \rangle$ be any noncommutative polynomial of degree k with scalar coefficients. Then we have*

$$\|P(s_1, \ldots, s_n)\| \le (k + 1)^{\frac{3}{2}} \, \|P(s_1, \ldots, s_n)\|_{L^2(\tau)}.$$

Proof. We adopt the notation of [10, §1]. Let $e_1, \ldots, e_n$ be the coordinate basis of $\mathbb{C}^n$. Then we can represent $s_i = l(e_i) + l^*(e_i)$ in terms of the annihilation and creation operators on the free Fock space $F_0(\mathbb{C}^n)$. It follows readily from the definitions that any polynomial of $s_1, \ldots, s_n$ of degree k is a linear combination of eigenvectors of the number operator N^0 with eigenvalue at most k, that is,

$$P(s_1, \ldots, s_n) = P_0 + \cdots + P_k$$

where $N^0 P_r = r P_r$. Thus

$$\begin{aligned}\|P(s_1, \ldots, s_n)\| &\le \|P_0\| + \cdots + \|P_k\| \\ &\le (k+1)\{\|P_0\|_{L^2(\tau)} + \cdots + \|P_k\|_{L^2(\tau)}\} \\ &\le (k+1)^{\frac{3}{2}} \|P(s_1, \ldots, s_n)\|_{L^2(\tau)}\end{aligned}$$

where the first line uses the triangle inequality, the second line uses [10, Theorem 4], and the last line uses Cauchy-Schwarz and that the eigenspaces of N^0 are orthogonal with respect to the inner product of $L^2(\tau)$. □

We now recall a standard trick to handle matrix coefficients.

Lemma 4.6. *Let $(\mathcal{A}, \tau)$ be a noncommutative probability space, let $P \in \mathrm{M}_d(\mathbb{C}) \otimes \mathcal{A}$, and denote by $P_{ij} \in \mathcal{A}$ its matrix elements. Then we have*

$$\|P\| \le d \max_{ij} \|P_{ij}\|, \qquad \max_{ij} \|P_{ij}\|_{L^2(\tau)} \le d^{\frac{1}{2}} \|P\|_{L^2(\mathrm{tr} \otimes \tau)}.$$

Proof. For the first inequality, we may assume that $\mathcal{A}$ has been represented as a subalgebra of $B(H)$ for some Hilbert space H. Then by Cauchy-Schwarz

$$\|P\| = \sup \left| \sum_{i,j} \langle v_i, P_{ij} w_j \rangle \right| \le \sup \sum_{i,j} \|v_i\| \|P_{ij}\| \|w_j\| \le d \max_{ij} \|P_{ij}\|,$$

where the supremum is taken over $v_i, w_j \in H$ so that $\sum_i \|v_i\|^2 = \sum_j \|w_j\|^2 = 1$.

For the second inequality, we note that

$$d^{-\frac{1}{2}} \|P_{ij}\|_{L^2(\tau)} = \|e_1 e_1^* \otimes P_{ij}\|_{L^2(\mathrm{tr} \otimes \tau)} = \|(e_1 e_i^* \otimes \mathbf{1}) P (e_j e_1^* \otimes \mathbf{1})\|_{L^2(\mathrm{tr} \otimes \tau)}$$

and use that $\|e_1 e_i^* \otimes \mathbf{1}\| = \|e_j e_1^* \otimes \mathbf{1}\| = 1$. □

We can now prove Theorem 4.1.

Proof of Theorem 4.1. Fix $q \in \mathbb{N}$ and set $Q = |P(s_1, \ldots, s_n)|^{2q}$. Then the matrix elements Q_{ij} are polynomials of $s_1, \ldots, s_n$ of degree $2qk$. Thus

$$\|Q\| \le d \max_{i,j} \|Q_{ij}\| \le d\,(2qk+1)^{\frac{3}{2}} \max_{ij} \|Q_{ij}\|_{L^2(\tau)} \le d^{\frac{3}{2}} (2qk+1)^{\frac{3}{2}} \|Q\|_{L^2(\mathrm{tr} \otimes \tau)}$$

by Lemmas 4.5 and 4.6. It remains to note that

$$\|Q\| = \|P(s_1, \ldots, s_n)\|^{2q}, \qquad \|Q\|_{L^2(\mathrm{tr} \otimes \tau)} = \|P(s_1, \ldots, s_n)\|^{2q}_{L^{4q}(\mathrm{tr} \otimes \tau)},$$

and the conclusion follows immediately. □

To deduce a corresponding bound on the resolvent (in the case that the polynomial P is self-adjoint), we apply an approximation argument.

Proof of Theorem 4.2. Fix $z \in \mathbb{C}$, $\operatorname{Im} z > 0$, and write $P := P(s_1, \ldots, s_n)$. As $x \mapsto |(z-x)^{-1}|$ is $(\operatorname{Im} z)^{-2}$-Lipschitz, Jackson's theorem [37, Corollary 1.4.1] yields

$$\sup_{x \in [-\|P\|, \|P\|]} \big| |(z-x)^{-1}| - Q_r(x) \big| \le \frac{6r^{-1}\|P\|}{(\operatorname{Im} z)^2}$$

for a polynomial Q_r of degree r. Therefore

$$\begin{aligned}
\|(z-P)^{-1}\| &\le \|Q_r(P)\| + \frac{6r^{-1}\|P\|}{(\operatorname{Im} z)^2} \\
&\le d^{\frac{3}{4q}} (2qkr+1)^{\frac{3}{4q}} \|Q_r(P)\|_{L^{4q}(\operatorname{tr}\otimes\tau)} + \frac{6r^{-1}\|P\|}{(\operatorname{Im} z)^2} \\
&\le d^{\frac{3}{4q}} (2qkr+1)^{\frac{3}{4q}} \left(\|(z-P)^{-1}\|_{L^{4q}(\operatorname{tr}\otimes\tau)} + \frac{12r^{-1}\|P\|}{(\operatorname{Im} z)^2} \right)
\end{aligned}$$

by Theorem 4.1, where we used that $Q_r \circ P$ is a polynomial of degree at most kr and that $d^{\frac{3}{4q}}(2qkr+1)^{\frac{3}{4q}} \ge 1$. This completes the proof. □

4.2. **Proof of Theorem 4.3 and Corollary 4.4.** The difficulty in the proof of Theorem 4.3 is that we must capture the fact that the resolvent $(z - X_{\mathrm{free}})^{-1}$ is insensitive to the eigenvalues of A_0 that are far from $\operatorname{Re} z$. To implement this idea, we will use spectral perturbation theory to show that shrinking gaps in the spectrum of A_0 of size $\gg \sigma(X)$ will only result in a small perturbation of the resolvent.

We first recall an elementary fact.

Lemma 4.7. $\operatorname{sp}(X_{\mathrm{free}}) \subseteq \operatorname{sp}(A_0) + 2\sigma(X)[-1,1]$.

Proof. We obtain $\operatorname{sp}(X_{\mathrm{free}}) \subseteq \operatorname{sp}(A_0) + \|X_{\mathrm{free}} - A_0 \otimes \mathbf{1}\|[-1,1]$ as in the proof of [9, Theorem VI.3.3], while $\|X_{\mathrm{free}} - A_0 \otimes \mathbf{1}\| \le 2\sigma(X)$ follows from [36, p. 208]. □

We now implement the spectral perturbation argument for a single gap in $\operatorname{sp}(A_0)$. This bound will be iterated below in order to prove Theorem 4.3.

Proposition 4.8. *Fix $z \in \mathbb{C}$ with $0 < \operatorname{Im} z \le \sigma(X)$, fix $K > 4$, and fix $a, b \in \mathbb{R}$ with $\delta := b - a - K\sigma(X) \ge 0$. If $(a,b) \subset (\operatorname{Re} z, \infty) \backslash \operatorname{sp}(A_0)$ we have*

$$\|(z - X_{\mathrm{free}})^{-1} - (z - X_{\mathrm{free}} + \delta 1_{[b,\infty)}(A_0) \otimes \mathbf{1})^{-1}\| \le \frac{9\sigma(X)}{K-2} \frac{1}{(\operatorname{Im} z)^2},$$

while if $(a,b) \subset (-\infty, \operatorname{Re} z) \backslash \operatorname{sp}(A_0)$ we have

$$\|(z - X_{\mathrm{free}})^{-1} - (z - X_{\mathrm{free}} - \delta 1_{(-\infty,a]}(A_0) \otimes \mathbf{1})^{-1}\| \le \frac{9\sigma(X)}{K-2} \frac{1}{(\operatorname{Im} z)^2}.$$

Proof. We will only consider the case $(a,b) \subset (\operatorname{Re} z, \infty) \backslash \operatorname{sp}(A_0)$, as the complementary case follows in a completely analogous fashion. Throughout the proof, we assume without loss of generality that X_{free} is represented concretely as an operator on a Hilbert space, so that we may work with its spectral projections $1_I(X_{\mathrm{free}})$.

Define $X_{\mathrm{free}}^{(r)} := X_{\mathrm{free}} - r 1_{[b,\infty)}(A_0 \otimes \mathbf{1})$ for $r \in [0, \delta]$. Then

$$\begin{aligned}
\left\| \frac{d}{dr} (z - X_{\mathrm{free}}^{(r)})^{-1} \right\| &= \|(z - X_{\mathrm{free}}^{(r)})^{-1} 1_{[b,\infty)}(A_0 \otimes \mathbf{1}) (z - X_{\mathrm{free}}^{(r)})^{-1}\| \\
&\le \|(z - X_{\mathrm{free}}^{(r)})^{-1} 1_{[b,\infty)}(A_0 \otimes \mathbf{1})\| \, \|1_{[b,\infty)}(A_0 \otimes \mathbf{1}) (z - X_{\mathrm{free}}^{(r)})^{-1}\|.
\end{aligned}$$

Define $A_0^{(r)} := A_0 - r1_{[b,\infty)}(A_0)$, that is, $A_0^{(r)}$ is obtained from A_0 by subtracting r from all eigenvalues of A_0 that are greater than b, while leaving all other eigenvalues unchanged. As $r \le \delta < b-a$ and A_0 has no eigenvalues in (a,b), this implies

$$1_{[b,\infty)}(A_0 \otimes \mathbf{1}) = 1_{[b-r,\infty)}(A_0^{(r)} \otimes \mathbf{1}).$$

On the other hand, as $\mathrm{sp}(X_{\mathrm{free}}^{(r)}) \subseteq \mathrm{sp}(A_0^{(r)}) + 2\sigma(X)[-1,1]$ by Lemma 4.7, we have

$$(a+2\sigma(X), b-r-2\sigma(X)) \cap \mathrm{sp}(X_{\mathrm{free}}^{(r)}) = \varnothing,$$

where we note that $a+2\sigma(X) < b-r-2\sigma(X)$ as $r \le \delta$ and $K > 4$. Thus

$$\begin{aligned}
&\big\|1_{[b,\infty)}(A_0 \otimes \mathbf{1})(z - X_{\mathrm{free}}^{(r)})^{-1}\big\| \\
&\quad\le \big\|1_{[b-r,\infty)}(A_0^{(r)} \otimes \mathbf{1})1_{(-\infty,a+2\sigma(X)]}(X_{\mathrm{free}}^{(r)})(z - X_{\mathrm{free}}^{(r)})^{-1}\big\| \\
&\qquad + \big\|1_{[b-r,\infty)}(A_0^{(r)} \otimes \mathbf{1})1_{[b-r-2\sigma(X),\infty)}(X_{\mathrm{free}}^{(r)})(z - X_{\mathrm{free}}^{(r)})^{-1}\big\| \\
&\quad\le \frac{\big\|1_{[b-r,\infty)}(A_0^{(r)} \otimes \mathbf{1})1_{(-\infty,a+2\sigma(X)]}(X_{\mathrm{free}}^{(r)})\big\|}{\mathrm{Im}\, z} + \frac{1}{b-a-r-2\sigma(X)},
\end{aligned}$$

using $\frac{1}{|z-x|} \le \frac{1}{b-a-r-2\sigma(X)}$ for $x \ge b-r-2\sigma(X)$ (as $\mathrm{Re}\, z \le a \le b-r-2\sigma(X)$). The identical bound clearly holds for $\|(z - X_{\mathrm{free}}^{(r)})^{-1}1_{[b,\infty)}(A_0 \otimes \mathbf{1})\|$ as well.

We now apply the Davis-Kahan theorem [9, Theorem VII.3.1] to estimate

$$\begin{aligned}
\big\|1_{[b-r,\infty)}(A_0^{(r)} \otimes \mathbf{1})1_{(-\infty,a+2\sigma(X)]}(X_{\mathrm{free}}^{(r)})\big\| &\le \frac{\|X_{\mathrm{free}}^{(r)} - A_0^{(r)} \otimes \mathbf{1}\|}{b-a-r-2\sigma(X)} \\
&\le \frac{2\sigma(X)}{b-a-r-2\sigma(X)},
\end{aligned}$$

where we used the free Khintchine inequality [36, p. 208] in the second inequality. Putting together the above estimates and using the assumption $1 \le \frac{\sigma(X)}{\mathrm{Im}\, z}$, we get

$$\left\|\frac{d}{dr}(z - X_{\mathrm{free}}^{(r)})^{-1}\right\| \le \frac{9\sigma(X)^2}{(b-a-r-2\sigma(X))^2}\frac{1}{(\mathrm{Im}\, z)^2}.$$

The fundamental theorem of calculus yields

$$\big\|(z - X_{\mathrm{free}})^{-1} - (z - X_{\mathrm{free}}^{(\delta)})^{-1}\big\| \le \int_0^\delta \frac{9\sigma(X)^2}{(b-a-r-2\sigma(X))^2}\frac{1}{(\mathrm{Im}\, z)^2}\, dr,$$

and the proof is readily completed. □

We can now conclude the proof of Theorem 4.3.

Proof of Theorem 4.3. As $(z - X_{\mathrm{free}})^{-1}$ is unchanged if we replace $z \leftarrow i\,\mathrm{Im}\, z$ and $A_0 \leftarrow A_0 - (\mathrm{Re}\, z)\mathbf{1}$, we may assume without loss of generality that $\mathrm{Re}\, z = 0$. Now note that the connected components of $\mathbb{R}\backslash(\mathrm{sp}(A_0) \cup \{0\})$ contain at most d bounded intervals. If any such interval has length exceeding $K\sigma(X)$, we modify A_0 using Proposition 4.8 to shrink the size of that interval to $K\sigma(X)$ while incurring an error $\frac{9\sigma(X)}{K-2}\frac{1}{(\mathrm{Im}\, z)^2}$. Repeating this procedure for each such interval yields a new matrix A_0' such that $\|A_0'\| \le Kd\sigma(X)$ and $X_{\mathrm{free}}' := A_0' \otimes \mathbf{1} + \sum_{i=1}^n A_i \otimes s_i$ satisfies

$$\|(z - X_{\mathrm{free}})^{-1} - (z - X_{\mathrm{free}}')^{-1}\| \le \frac{9d\sigma(X)}{K-2}\frac{1}{(\mathrm{Im}\, z)^2}.$$

As $\|X'_{\rm free}\| \le (Kd+2)\sigma(X)$, we further obtain

$$\|(z-X'_{\rm free})^{-1}\| \le d^{\frac{3}{4q}}(2qr+1)^{\frac{3}{4q}}\left(\|(z-X'_{\rm free})^{-1}\|_{L^{4q}(\mathrm{tr}\otimes\tau)} + \frac{12r^{-1}(Kd+2)\sigma(X)}{(\mathrm{Im}\, z)^2}\right)$$

by Theorem 4.2. Combining the above bounds and using $d^{\frac{3}{4q}}(2qr+1)^{\frac{3}{4q}} \ge 1$ and $Kd+2 \le \frac{3Kd}{2}$ readily yields the conclusion. □

It remains to prove Corollary 4.4.

Proof of Corollary 4.4. Assume first that $\mathrm{Im}\, z \le \sigma(X)$. As

$$\sigma(X)^2 = \sup_{\|v\|=1}\sum_{i=1}^n\sum_{k=1}^d \langle e_k, A_i v\rangle^2 \le d \sup_{\|v\|=\|w\|=1}\sum_{i=1}^n \langle w, A_i v\rangle^2 = d\,\sigma_*(X)^2,$$

the conclusion follows from Theorem 4.3 with $K = 2d^{\frac{3}{2}}+2$, $r = \lceil 2d^{\frac{3}{2}}K\rceil \le 8d^3$.

We now consider the case that $\mathrm{Im}\, z > \sigma(X)$ and $\mathrm{Re}\, z \in \mathrm{sp}(X_{\rm free})$. Then

$$\begin{aligned}
\|(z-X_{\rm free})^{-1}\|^{4q}_{L^{4q}(\mathrm{tr}\otimes\tau)} &= (\mathrm{tr}\otimes\tau)[(|z-X_{\rm free}|^2)^{-2q}] \\
&\ge \mathrm{tr}[((\mathrm{id}\otimes\tau)|z-X_{\rm free}|^2)^{-2q}] \\
&= \mathrm{tr}[((\mathrm{Im}\, z)^2 + \mathbf{E}[(X-\mathbf{E}X)^2] + (\mathrm{Re}\, z - A_0)^2)^{-2q}] \\
&\ge \mathrm{tr}[((\mathrm{Im}\, z)^2 + \sigma(X)^2 + (\mathrm{Re}\, z - A_0)^2)^{-2q}],
\end{aligned}$$

where we used Jensen's inequality in C^*-algebras [35] in the first inequality, and that $\mathbf{E}[(X-\mathbf{E}X)^2] \le \sigma(X)^2\mathbf{1}$ and trace monotonicity [17, §2.2] in the second inequality. Now note that Lemma 4.7 ensures there is an eigenvalue of A_0 within distance $2\sigma(X)$ of $\mathrm{Re}\, z \in \mathrm{sp}(X_{\rm free})$. We can therefore estimate

$$\|(z-X_{\rm free})^{-1}\|^{4q}_{L^{4q}(\mathrm{tr}\otimes\tau)} \ge \frac{1}{d}\frac{1}{((\mathrm{Im}\, z)^2 + 5\sigma(X)^2)^{2q}} \ge \frac{1}{d}\frac{1}{(6(\mathrm{Im}\, z)^2)^{2q}}.$$

Thus we obtain

$$\|(z-X_{\rm free})^{-1}\| \le \frac{1}{\mathrm{Im}\, z} \le d^{\frac{1}{4q}}\sqrt{6}\,\|(z-X_{\rm free})^{-1}\|_{L^{4q}(\mathrm{tr}\otimes\tau)},$$

and the conclusion follows readily. □

5. Sharp matrix concentration inequalities

5.1. **Proof of Theorem 2.2.** As the upper bound on $\mathrm{sp}(X)$ was already proved in [6], we only need to prove the corresponding lower bound. To this end, we will follow the approach of [6, §6] with the crucial input of Corollary 4.4.

The basis for the proof is the following.

Lemma 5.1. *Fix $z \in \mathbb{C}$ with $\mathrm{Re}\, z \in \mathrm{sp}(X_{\rm free})$ and $\mathrm{Im}\, z > 0$. Then*

$$\mathbf{P}\left[c\|(z-X_{\rm free})^{-1}\| \ge \|(z-X)^{-1}\| + \frac{\tilde v(X)^4}{(\mathrm{Im}\, z)^5}(\log d)^3 + \frac{\sigma_*(X)}{(\mathrm{Im}\, z)^2}(\sqrt{\log d}+t)\right] \le e^{-t^2}$$

for all $t \ge 0$, where c is a universal constant.

Proof. We begin by noting that [6, Corollary 4.14 and (6.2)] implies

$$\mathbf{P}\left[\pm(\|(z-X)^{-1}\| - \mathbf{E}\|(z-X)^{-1}\|) \ge \frac{\sigma_*(X)}{(\mathrm{Im}\, z)^2}\,t\right] \le e^{-t^2/2} \tag{5.1}$$

for $t \ge 0$. This further implies

$$\mathbf{E}[\|(z-X)^{-1}\|^{2p}]^{\frac{1}{2p}} \le \mathbf{E}\|(z-X)^{-1}\| + \frac{\sigma_*(X)}{(\operatorname{Im} z)^2}\,2\sqrt{p} \tag{5.2}$$

for all $p \in \mathbb{N}$ by [13, Theorem 2.1]

On the other hand, [6, Theorem 6.1] yields for $q \in \mathbb{N}$

$$\|(z-X_{\rm free})^{-1}\|_{L^{4q}(\operatorname{tr}\otimes\tau)} \le \mathbf{E}[\operatorname{tr}|z-X|^{-4q}]^{\frac{1}{4q}} + \frac{8}{3}\frac{\tilde v(X)^4}{(\operatorname{Im} z)^5}(q+1)^3.$$

Applying (5.2) and Corollary 4.4 with $q = \lceil \log d \rceil$ yields

$$c\|(z-X_{\rm free})^{-1}\| \le \mathbf{E}\|(z-X)^{-1}\| + \frac{\tilde v(X)^4}{(\operatorname{Im} z)^5}(\log d)^3 + \frac{\sigma_*(X)}{(\operatorname{Im} z)^2}\sqrt{\log d}$$

for a universal constant c, where we used that $\operatorname{tr}|z-X|^{-4q} \le \|(z-X)^{-1}\|^{4q}$. The conclusion follows by applying (5.1). □

We now deduce a uniform analogue of the previous lemma.

Lemma 5.2. *Fix $\varepsilon > 0$. Then*

$$\mathbf{P}\Bigg[c\|(z-X_{\rm free})^{-1}\| \le \|(z-X)^{-1}\| + \frac{\tilde v(X)^4}{\varepsilon^5}(\log d)^3 + \frac{\sigma_*(X)}{\varepsilon^2}(\sqrt{\log d}+t) \\ \text{for all } z \in \operatorname{sp}(X_{\rm free}) + i\varepsilon\Bigg] \ge 1 - e^{-t^2}$$

for all $t \ge 0$, where c is a universal constant.

Proof. Using $\sigma(X) \le \sqrt{d}\,\sigma_*(X)$ as in the proof of Corollary 4.4, Lemma 4.7 yields

$$\operatorname{sp}(X_{\rm free}) \subseteq \operatorname{sp}(A_0) + 2\sqrt{d}\,\sigma_*(X)[-1,1].$$

As $|\operatorname{sp}(A_0)| \le d$, it follows that $\operatorname{sp}(X_{\rm free})$ can be covered by at most d intervals of length $4\sqrt{d}\,\sigma_*(X)$. We can therefore find $N \subset \operatorname{sp}(X_{\rm free})$ of cardinality $|N| \le 4d^{\frac{3}{2}}$ so that each point in $\operatorname{sp}(X_{\rm free})$ is within distance $\sigma_*(X)$ of a point in N.

Now note that for any $\lambda, \lambda' \in \mathbb{R}$ and $\varepsilon > 0$

$$\big|\|(\lambda+i\varepsilon-X)^{-1}\| - \|(\lambda'+i\varepsilon-X)^{-1}\|\big| \le \frac{|\lambda-\lambda'|}{\varepsilon^2},$$

and analogously when X is replaced by $X_{\rm free}$. We can therefore estimate

$$\mathbf{P}\Bigg[c'\|(z-X_{\rm free})^{-1}\| \ge \|(z-X)^{-1}\| + \frac{\tilde v(X)^4}{\varepsilon^5}(\log d)^3 + \frac{\sigma_*(X)}{\varepsilon^2}(\sqrt{\log d}+t) \\ \text{for some } z \in \operatorname{sp}(X_{\rm free}) + i\varepsilon\Bigg] \le$$

$$\mathbf{P}\Bigg[c\|(z-X_{\rm free})^{-1}\| \ge \|(z-X)^{-1}\| + \frac{\tilde v(X)^4}{\varepsilon^5}(\log d)^3 + \frac{\sigma_*(X)}{\varepsilon^2}(\sqrt{\log d}+t) \\ \text{for some } z \in N + i\varepsilon\Bigg] \le 4d^{\frac{3}{2}}e^{-t^2}$$

for any $t \ge 0$ using Lemma 5.1 and the union bound, where c, c' are universal constants. The conclusion follows by replacing $t \leftarrow t + 2\sqrt{\log d}$ and noting that $4d^{\frac{3}{2}}e^{-(t+2\sqrt{\log d})^2} \le e^{-t^2}$ for all $t \ge 0$ (recalling the standing assumption $d \ge 2$). □

We can now conclude the proof of Theorem 2.2.

Proof of Theorem 2.2. It was shown in [6, Theorem 2.1] that

$$\mathbf{P}\big[\mathrm{sp}(X) \subseteq \mathrm{sp}(X_{\mathrm{free}}) + C\{\tilde{v}(X)(\log d)^{\frac{3}{4}} + \sigma_*(X)t\}[-1,1]\big] \ge 1 - e^{-t^2}.$$

On the other hand, combining Lemma 5.2 with [6, Lemma 6.4] yields

$$\mathbf{P}\big[\mathrm{sp}(X_{\mathrm{free}}) \subseteq \mathrm{sp}(X) + C\{\tilde{v}(X)(\log d)^{\frac{3}{4}} + \sigma_*(X)t\}[-1,1]\big] \ge 1 - e^{-t^2},$$

where we used that $\sigma_*(X)\sqrt{\log d} \le \tilde{v}(X)(\log d)^{\frac{3}{4}}$. The union bound yields

$$\mathbf{P}\big[\mathrm{d_H}(\mathrm{sp}(X), \mathrm{sp}(X_{\mathrm{free}})) > C\{\tilde{v}(X)(\log d)^{\frac{3}{4}} + \sigma_*(X)t\}\big] \le 2e^{-t^2}.$$

We conclude by replacing $t \leftarrow t + \sqrt{\log 2}$ and using again that $\sigma_*(X) \le \tilde{v}(X)$. □

5.2. **Proof of Corollary 2.3.** The proof is entirely straightforward.

Proof of Corollary 2.3. Assume first that X is self-adjoint. Then the tail bound follows immediately from Theorem 2.2 as $|\|X\| - \|X_{\mathrm{free}}\|| \le \mathrm{d_H}(\mathrm{sp}(X), \mathrm{sp}(X_{\mathrm{free}}))$. To deduce the bound in expectation, we estimate

$$\begin{aligned}|\mathbf{E}\|X\| - \|X_{\mathrm{free}}\|| &\le \int_0^\infty \mathbf{P}\big[|\|X\| - \|X_{\mathrm{free}}\|| > x\big]\,dx \\ &\le C\tilde{v}(X)(\log d)^{\frac{3}{4}} + \int_0^\infty \mathbf{P}\big[|\|X\| - \|X_{\mathrm{free}}\|| > C\tilde{v}(X)(\log d)^{\frac{3}{4}} + x\big]\,dx\end{aligned}$$

and use that $\sigma_*(X) \le \tilde{v}(X)$. The corresponding results for $\lambda_{\max}(X), \lambda_{\max}(X_{\mathrm{free}})$ and $\lambda_{\min}(X), \lambda_{\min}(X_{\mathrm{free}})$ follow by an identical argument. Finally, the norm bounds extend directly to the non-self-adjoint case by [6, Remark 2.6]. □

5.3. **Proof of Theorem 2.5.** The basis for the proof is the following variant of the linearization argument of [6, Lemma 3.13].

Lemma 5.3. *Fix $\varepsilon > 0$, and define $B_\varepsilon := B + (\|B\| + 4\varepsilon^2)\mathbf{1}$ and*

$$\breve{X}_\varepsilon = \begin{bmatrix} 0 & X & B_\varepsilon^{\frac{1}{2}} \\ X^* & 0 & 0 \\ B_\varepsilon^{\frac{1}{2}} & 0 & 0 \end{bmatrix}, \qquad \breve{X}_{\mathrm{free},\varepsilon} = \begin{bmatrix} 0 & X_{\mathrm{free}} & B_\varepsilon^{\frac{1}{2}} \otimes \mathbf{1} \\ X_{\mathrm{free}}^* & 0 & 0 \\ B_\varepsilon^{\frac{1}{2}} \otimes \mathbf{1} & 0 & 0 \end{bmatrix}.$$

Then

$$\mathrm{d_H}(\mathrm{sp}(\breve{X}_\varepsilon), \mathrm{sp}(\breve{X}_{\mathrm{free},\varepsilon})) \le \varepsilon$$

implies

$$\mathrm{d_H}(\mathrm{sp}(XX^* + B), \mathrm{sp}(X_{\mathrm{free}}X_{\mathrm{free}}^* + B \otimes \mathbf{1}) \le 4\{\|X_{\mathrm{free}}\| + \|B\|^{\frac{1}{2}}\}\varepsilon + 5\varepsilon^2.$$

Proof. By [6, Remark 2.6], we have

$$\mathrm{sp}(\breve{X}_\varepsilon) \cup \{0\} = \mathrm{sp}((XX^* + B_\varepsilon)^{\frac{1}{2}}) \cup -\mathrm{sp}((XX^* + B_\varepsilon)^{\frac{1}{2}}) \cup \{0\},$$

and analogously for $\breve{X}_{\mathrm{free},\varepsilon}$.

Consider first any $\lambda \in \mathrm{sp}(XX^* + B_\varepsilon)$. Then $\lambda^{\frac{1}{2}} \ge 2\varepsilon$ by the definition of B_ε, and thus $\lambda^{\frac{1}{2}} \in \mathrm{sp}(\breve{X}_\varepsilon)$. By assumption, there exists $\mu \in \mathrm{sp}(\breve{X}_{\mathrm{free},\varepsilon})$ so that $|\lambda^{\frac{1}{2}} - \mu| \le \varepsilon$. This implies $\mu \ge \varepsilon$, so it must be that $\mu \in \mathrm{sp}((X_{\mathrm{free}}X_{\mathrm{free}}^* + B_\varepsilon \otimes \mathbf{1})^{\frac{1}{2}})$. Moreover,

$$|\lambda - \mu^2| = (\lambda^{\frac{1}{2}} + \mu)|\lambda^{\frac{1}{2}} - \mu| \le 2\mu\varepsilon + \varepsilon^2.$$

As $\mu^2 \in \mathrm{sp}(X_{\mathrm{free}}X_{\mathrm{free}}^* + B_\varepsilon \otimes \mathbf{1})$, we have shown that

$$\mathrm{sp}(XX^* + B_\varepsilon) \subseteq \mathrm{sp}(X_{\mathrm{free}}X_{\mathrm{free}}^* + B_\varepsilon \otimes \mathbf{1}) + (2\|X_{\mathrm{free}}X_{\mathrm{free}}^* + B_\varepsilon \otimes \mathbf{1}\|^{\frac{1}{2}}\varepsilon + \varepsilon^2)[-1,1].$$

Reversing the roles of $X, X_{\rm free}$ yields

$$\mathrm{d_H}(\mathrm{sp}(XX^* + B_\varepsilon), \mathrm{sp}(X_{\rm free}X_{\rm free}^* + B_\varepsilon \otimes \mathbf{1})) \le 2\|X_{\rm free}X_{\rm free}^* + B_\varepsilon \otimes \mathbf{1}\|^{\frac{1}{2}}\varepsilon + \varepsilon^2$$

by the identical argument.

To conclude the proof, note first that

$$\mathrm{d_H}(\mathrm{sp}(XX^*+B_\varepsilon), \mathrm{sp}(X_{\rm free}X_{\rm free}^*+B_\varepsilon\otimes\mathbf{1})) = \mathrm{d_H}(\mathrm{sp}(XX^*+B), \mathrm{sp}(X_{\rm free}X_{\rm free}^*+B\otimes\mathbf{1}))$$

as Hausdorff distance is translation-invariant $\mathrm{d_H}(I+t, J+t) = \mathrm{d_H}(I,J)$. On the other hand, we can estimate

$$\|X_{\rm free}X_{\rm free}^* + B_\varepsilon \otimes \mathbf{1}\| \le \|X_{\rm free}\|^2 + 2\|B\| + 4\varepsilon^2,$$

and the proof is readily completed. □

We can now complete the proof of Theorem 2.5.

Proof of Theorem 2.5. By Lemma 5.3, we can estimate

$$\begin{aligned}\mathbf{P}\big[\mathrm{d_H}(\mathrm{sp}(XX^* + B), \mathrm{sp}(X_{\rm free}X_{\rm free}^* + B \otimes \mathbf{1}) > 4\{\|X_{\rm free}\| + \|B\|^{\frac{1}{2}}\}\varepsilon + 5\varepsilon^2\big] \\ \le \mathbf{P}\big[\mathrm{d_H}(\mathrm{sp}(\breve{X}_\varepsilon), \mathrm{sp}(\breve{X}_{\mathrm{free},\varepsilon})) > \varepsilon\big].\end{aligned}$$

We now recall that $\sigma_*(\breve{X}_\varepsilon) = \sigma_*(X)$ and $\tilde{v}(\breve{X}_\varepsilon) \le 2^{\frac{1}{4}}\tilde{v}(X)$ by [6, Remark 2.6]. Thus $t \ge \tilde{v}(X)(\log d)^{\frac{3}{4}}$ implies $2t \ge 2^{-\frac{1}{4}}\tilde{v}(\breve{X}_\varepsilon)(\log d)^{\frac{3}{4}} + \sigma_*(\breve{X}_\varepsilon)\frac{t}{\sigma_*(X)}$. The conclusion follows from Theorem 2.2 by choosing $\varepsilon = C't$ for a universal constant C'. □

6. Phase transitions: isotropic case

6.1. **Proof of Theorem 2.7.** Theorem 2.7 is based on the Lehner formula (1.6). At its core, the reason that this variational principle exhibits phase transitions in the presence of low-rank structure is contained in the following simple observation: the last term in the Lehner formula is small when M has low rank.

Lemma 6.1. *Let $M \in \mathrm{M}_d(\mathbb{C})_{\rm sa}$ have rank r. Then*

$$\|\mathbf{E}[(X - \mathbf{E}X)M(X - \mathbf{E}X)]\| \le \sigma_*(X)^2 r\,\|M\|.$$

Proof. Writing $M = \sum_{i=1}^r \lambda_i v_i v_i^*$ with $|\lambda_i| \le \|M\|$ and $\|v_i\| = 1$, we obtain

$$\|\mathbf{E}[(X-\mathbf{E}X)M(X-\mathbf{E}X)]\| = \sup_{\|w\|=1}\left|\sum_{i=1}^r \lambda_i\,\mathbf{E}[|\langle v_i, (X-\mathbf{E}X)w\rangle^2]\right| \le \sigma_*(X)^2 r\|M\|$$

by the triangle inequality and the definition of $\sigma_*(X)$. □

We first prove the upper bound in Theorem 2.7.

Lemma 6.2. *Let X be any $d\times d$ self-adjoint random matrix with $\mathbf{E}[(X-\mathbf{E}X)^2] = \mathbf{1}$ and such that $\mathbf{E}X$ has rank r. Then we have*

$$\lambda_{\max}(X_{\rm free}) \le \mathrm{B}(\lambda_{\max}(\mathbf{E}X)) + 2\sigma_*(X)\sqrt{r}.$$

Proof. Denote by P the projection onto the range of $\mathbf{E}X$. Then we can upper bound $\lambda_{\max}(X_{\rm free})$ by restricting the infimum in (1.6) only to matrices of the form $M = sP + t(\mathbf{1}-P)$ for $s,t>0$, and using that for such M

$$\mathbf{E}[(X-\mathbf{E}X)M(X-\mathbf{E}X)] \le t + \sigma_*(X)^2 rs$$

by the isotropic assumption $\mathbf{E}[(X-\mathbf{E}X)^2] = \mathbf{1}$ and Lemma 6.1. This yields

$$\begin{aligned}\lambda_{\max}(X_{\rm free}) &\le \inf_{s,t>0} \lambda_{\max}\big(\mathbf{E}X + s^{-1}P + t^{-1}(\mathbf{1}-P) + t + \sigma_*(X)^2 rs\big)\\ &\le \inf_{t>0} \max\{\lambda_{\max}(\mathbf{E}X)+t, t^{-1}+t\} + 2\sigma_*(X)\sqrt{r},\end{aligned}$$

where we used $\mathbf{E}X \le \lambda_{\max}(\mathbf{E}X)P$ and $s^{-1}P \le s^{-1}$ in the second inequality. It remains to note that $\inf_{t>0}\max\{\theta+t, t^{-1}+t\} = \mathrm{B}(\theta)$. □

We now turn to the lower bound.

Lemma 6.3. *Let X be any $d\times d$ self-adjoint random matrix with $\mathbf{E}[(X-\mathbf{E}X)^2] = \mathbf{1}$ and such that $\mathbf{E}X$ has rank r with $\sigma_*(X)\sqrt{r} \le 1$. Then we have*

$$\lambda_{\max}(X_{\rm free}) \ge \mathrm{B}(\lambda_{\max}(\mathbf{E}X)) - 2\sigma_*(X)\sqrt{r}.$$

Proof. Let $r' \le r$ be the number of (strictly) negative eigenvalues of $\mathbf{E}X$, and assume without loss of generality that the associated eigenvectors are $e_{d-r'+1},\ldots,e_d$. Let $Q : \mathbb{C}^d \to \mathbb{C}^{d-r'}$ be the coordinate projection on the first $d-r'$ coordinate directions. Then the Lehner formula (1.6) yields

$$\begin{aligned}\lambda_{\max}(X_{\rm free}) &\ge \lambda_{\max}((Q\otimes\mathbf{1})X_{\rm free}(Q^*\otimes\mathbf{1}))\\ &= \inf_{M>0}\lambda_{\max}\big(Q\,\mathbf{E}X\,Q^* + M^{-1} + Q\mathbf{E}[(X-\mathbf{E}X)Q^*MQ(X-\mathbf{E}X)]Q^*\big),\end{aligned}$$

where the infimum is taken over $(d-r')$-dimensional matrices M.

To proceed, note that $\mathbf{1}-Q^*Q$ is a projection of rank $r' \le r$. Thus

$$Q\mathbf{E}[(X-\mathbf{E}X)Q^*Q(X-\mathbf{E}X)]Q^* \ge 1-\sigma_*(X)^2 r =: \beta$$

where we used the isotropic assumption $\mathbf{E}[(X-\mathbf{E}X)^2] = \mathbf{1}$, that $QQ^* = \mathbf{1}$, and Lemma 6.1. Moreover, note that $\beta \ge 0$ by assumption. We can therefore bound

$$\begin{aligned}&\lambda_{\max}\big(Q\,\mathbf{E}X\,Q^* + M^{-1} + Q\mathbf{E}[(X-\mathbf{E}X)Q^*MQ(X-\mathbf{E}X)]Q^*\big)\\ &\ge \lambda_{\max}\big(Q\,\mathbf{E}X\,Q^* + M^{-1}\big) + \beta\lambda_{\min}(M)\\ &\ge \max\{\lambda_{\max}(\mathbf{E}X), (\lambda_{\min}(M))^{-1}\} + \beta\lambda_{\min}(M),\end{aligned}$$

where we used $M^{-1} \ge 0$, $\lambda_{\max}(Q\,\mathbf{E}X\,Q^*) = \lambda_{\max}(\mathbf{E}X)$ and $Q\,\mathbf{E}X\,Q^* \ge 0$, respectively, to obtain the two terms in the maximum on the last line. Thus

$$\lambda_{\max}(X_{\rm free}) \ge \inf_{t>0}\max\{\lambda_{\max}(\mathbf{E}X)+\beta t, t^{-1}+\beta t\} = \beta^{\frac12}\mathrm{B}(\beta^{-\frac12}\lambda_{\max}(\mathbf{E}X)).$$

It remains to show that $\beta^{\frac12}\mathrm{B}(\beta^{-\frac12}\theta) \ge \mathrm{B}(\theta) - 2\sigma_*(X)\sqrt{r}$.

To this end, note first that $\beta^{\frac12} \ge 1-\sigma_*(X)\sqrt{r}$ as $\sqrt{1-x^2} \ge 1-x$ for $x\in[0,1]$. We now consider two regimes. If $\theta \le 1$, we have

$$\beta^{\frac12}\mathrm{B}(\beta^{-\frac12}\theta) \ge 2\beta^{\frac12} \ge 2 - 2\sigma_*(X)\sqrt{r} = \mathrm{B}(\theta) - 2\sigma_*(X)\sqrt{r}.$$

On the other hand, if $\theta > 1$, then we have

$$\beta^{\frac12}\mathrm{B}(\beta^{-\frac12}\theta) = \theta + \frac{\beta}{\theta} = \mathrm{B}(\theta) - \frac{\sigma_*(X)^2 r}{\theta} \ge \mathrm{B}(\theta) - \sigma_*(X)\sqrt{r}$$

as $\sigma_*(X)^2 r \le \sigma_*(X)\sqrt{r}$. The proof is complete. □

Theorem 2.7 follows immediately by combining Lemmas 6.2 and 6.3.

6.2. **Proof of Theorem 2.9.** Despite that we formulated Theorem 2.9 in the context of random matrices, the argument is entirely deterministic in nature. The proof is based on the following basic observation.

Lemma 6.4. *Let $X, P \in \mathrm{M}_d(\mathbb{C})_{\mathrm{sa}}$ and $t > 0$. Then*

$$\frac{\lambda_{\max}(X) - \lambda_{\max}(X - tP)}{t} \le \langle v_{\max}(X), P v_{\max}(X)\rangle \le \frac{\lambda_{\max}(X + tP) - \lambda_{\max}(X)}{t}$$

for any unit norm eigenvector $v_{\max}(X)$ of X with eigenvalue $\lambda_{\max}(X)$.

Proof. To prove the upper bound, note that we obtain

$$\begin{aligned}\lambda_{\max}(X + tP) - \lambda_{\max}(X) &= \sup_{\|v\|=1} \langle v, (X + tP)v\rangle - \langle v_{\max}(X), X v_{\max}(X)\rangle \\ &\ge t\langle v_{\max}(X), P v_{\max}(X)\rangle\end{aligned}$$

by choosing $v \leftarrow v_{\max}(X)$ in the supremum. The lower bound follows immediately if we replace $t \leftarrow -t$ in the above inequality. □

To exploit these inequalities in the setting of Theorem 2.7, we must estimate the bounded differences of the function $\mathrm{B}(\cdot)$.

Lemma 6.5. *For any $t > 0$, we have*

$$\frac{\mathrm{B}(\theta + t) - \mathrm{B}(\theta)}{t} \le \left(1 - \frac{1}{\theta^2}\right)_+ + t, \qquad \frac{\mathrm{B}(\theta) - \mathrm{B}(\theta - t)}{t} \ge \left(1 - \frac{1}{\theta^2}\right)_+ - t.$$

Proof. We readily compute

$$\frac{d\mathrm{B}(\theta)}{d\theta} = \left(1 - \frac{1}{\theta^2}\right)_+, \qquad \frac{d^2\mathrm{B}(\theta)}{d\theta^2} = \frac{2}{\theta^3} 1_{\theta \ge 1} \le 2.$$

Taylor expanding to first order yields

$$\frac{\mathrm{B}(\theta + t) - \mathrm{B}(\theta)}{t} = \left(1 - \frac{1}{\theta^2}\right)_+ + t \int_0^1 \frac{2}{(\theta + rt)^3} 1_{\theta + rt \ge 1} (1 - r)\, dr.$$

The upper bound in the statement follows using $\frac{2}{(\theta+rt)^3} 1_{\theta+rt\ge1} \le 2$. The lower bound follows by the identical argument once we replace $t \leftarrow -t$. □

We can now complete the proof.

Proof of Theorem 2.9. We begin by writing

$$\frac{\lambda_{\max}(X) - \lambda_{\max}(X_{-t})}{t} \le \langle v_{\max}(X), 1_{(\theta-\delta,\theta]}(\mathbf{E}X) v_{\max}(X)\rangle \le \frac{\lambda_{\max}(X_t) - \lambda_{\max}(X)}{t}$$

using Lemma 6.4. If in addition

$$|\lambda_{\max}(X_s) - \mathrm{B}(\lambda_{\max}(\mathbf{E}X_s))| \le \varepsilon \text{ for } s \in \{0, \pm t\},$$

Lemma 6.5 yields

$$\left|\langle v_{\max}(X), 1_{(\theta-\delta,\theta]}(\mathbf{E}X) v_{\max}(X)\rangle - \left(1 - \frac{1}{\theta^2}\right)_+\right| \le t + \frac{2\varepsilon}{t},$$

where we used that $\lambda_{\max}(\mathbf{E}X_t) = \theta + t$ and $\lambda_{\max}(\mathbf{E}X_{-t}) = \theta - t$ (because $t \le \delta$). It remains to note that the above condition holds with high probability

$$\mathbf{P}\big[|\lambda_{\max}(X_s) - \mathrm{B}(\lambda_{\max}(\mathbf{E}X_s))| \le \varepsilon \text{ for } s \in \{0, \pm t\}\big] \ge 1 - 3\rho$$

by the union bound, concluding the proof. □

7. Phase transitions: anisotropic case

The aim of this section is to prove Theorem 2.12. The proof consists of several distinct parts. In section 7.1, we apply a general reduction principle to reduce the dimension of the Lehner variational formula (1.6). In section 7.2, we approximate the Lehner formula for $X_{\mathrm{free}}, X_{\varnothing,\mathrm{free}}$ by simplified parameters $\lambda, \lambda_\varnothing$ using the low-rank structure of the model. We also obtain the quantitative bound on $\lambda_\varnothing$. We subsequently prove the phase transition of λ in section 7.3.

Notation. The following notations will be used primarily in this section. For any vector $v \in \mathbb{C}^d$ and matrix $M \in \mathrm{M}_d(\mathbb{C})$, we will denote

$$\frac{1}{v} := \begin{bmatrix} v_1^{-1} \\ \vdots \\ v_d^{-1} \end{bmatrix}, \qquad \mathrm{diag}(v) := \begin{bmatrix} v_1 & & \\ & \ddots & \\ & & v_d \end{bmatrix}, \qquad \mathrm{diag}^{-1}(M) := \begin{bmatrix} M_{11} \\ \vdots \\ M_{dd} \end{bmatrix}.$$

We will denote by $\mathrm{I}_{C_k} \in \mathbb{C}^d$ the indicator $(\mathrm{I}_{C_k})_i = 1_{i \in C_k}$ and by $\mathbf{I}_{C_k} := \mathrm{diag}(\mathrm{I}_{C_k})$. The elementwise (Hadamard) product of vectors or matrices is denoted as $\odot$.

7.1. Reduction.

7.1.1. *A general reduction principle.* The Lehner formula (1.6) is a minimization problem over $d \times d$ matrices. However, one can often reduce the dimension of the variational problem in models with invariant structure. The following general reduction principle greatly facilitates the analysis of such models.

Lemma 7.1 (Reduction principle)**.** *Let X be any $d \times d$ self-adjoint random matrix and let $\mathcal{A}$ be any $*$-subalgebra of $\mathrm{M}_d(\mathbb{C})$. Suppose that*

$$\mathbf{E}X \in \mathcal{A}, \qquad \mathbf{E}[(X - \mathbf{E}X)M(X - \mathbf{E}X)] \in \mathcal{A} \quad \text{for all} \quad M \in \mathcal{A}.$$

Then we have

$$\lambda_{\max}(X_{\mathrm{free}}) = \inf_{M \in \mathcal{A} : M > 0} \lambda_{\max}\big(\mathbf{E}X + M^{-1} + \mathbf{E}[(X - \mathbf{E}X)M(X - \mathbf{E}X)]\big).$$

Proof. That $\lambda_{\max}(X_{\mathrm{free}})$ is upper bounded by the expression in the statement is obvious from (1.6). It remains to prove the corresponding lower bound. To this end, let $\pi : \mathrm{M}_d(\mathbb{C}) \to \mathcal{A}$ be the conditional expectation given $\mathcal{A}$ (cf. [17, §4.3]). As conditional expectations are monotone, any $M \in \mathrm{M}_d(\mathbb{C})$ with $M > 0$ satisfies

$$\begin{aligned}\lambda_{\max}\big(\mathbf{E}X + M^{-1} + \mathbf{E}[(X - \mathbf{E}X)M(X - \mathbf{E}X)]\big) \geq \\ \lambda_{\max}\big(\mathbf{E}X + \pi(M)^{-1} + \pi(\mathbf{E}[(X - \mathbf{E}X)M(X - \mathbf{E}X)])\big),\end{aligned}$$

where we used $\mathbf{E}X \in \mathcal{A}$ and that $\pi(M^{-1}) \geq \pi(M)^{-1}$ by [17, Theorem 4.16].

We now claim that

$$\pi(\mathbf{E}[(X - \mathbf{E}X)M(X - \mathbf{E}X)]) = \mathbf{E}[(X - \mathbf{E}X)\pi(M)(X - \mathbf{E}X)].$$

Indeed, note that $\sigma : M \mapsto \mathbf{E}[(X - \mathbf{E}X)M(X - \mathbf{E}X)]$ is a self-adjoint linear map on $\mathrm{M}_d(\mathbb{C})$ with respect to the Hilbert-Schmidt inner product. As we assumed σ leaves $\mathcal{A}$ invariant and as it is self-adjoint, it leaves $\mathcal{A}^\perp$ invariant as well. Thus the claimed identity follows by writing $M = \pi(M) + M^\perp$ with $M^\perp \in \mathcal{A}^\perp$.

Combining the above observations with (1.6) yields

$$\lambda_{\max}(X_{\mathrm{free}}) \geq \inf_{M > 0} \lambda_{\max}\big(\mathbf{E}X + \pi(M)^{-1} + \mathbf{E}[(X - \mathbf{E}X)\pi(M)(X - \mathbf{E}X)]\big),$$

and the conclusion follows immediately. □

7.1.2. *The invariant algebra.* From now on we assume that $X, X_\varnothing$ are defined according to the model in section 2.4. The first step in our analysis will be to introduce a specific invariant $*$-algebra $\mathcal{A}$ for this model, to which Lemma 7.1 can be applied. To this end, define $f_k \in \mathbb{C}^d$ and $P_k \in \mathrm{M}_d(\mathbb{C})$ as

$$f_k := \frac{z \odot \mathrm{I}_{C_k}}{\sqrt{|C_k|}}, \qquad P_k := \mathbf{I}_{C_k} - f_k f_k^*.$$

The assumptions of section 2.4 imply that $f_1, \ldots, f_q$ are orthonormal, $P_1, \ldots, P_q$ are orthogonal projections onto nontrivial (as $|C_k| > 1$) orthogonal subspaces, and $P_1 + \cdots + P_q$ is the orthogonal projection onto $\{f_k : k \in [q]\}^\perp$.

Definition 7.2. Define the $*$-subalgebra

$$\mathcal{A} := \{\mathrm{A}(M, v) : M \in \mathrm{M}_q(\mathbb{C}),\ v \in \mathbb{C}^q\}$$

of $\mathrm{M}_d(\mathbb{C})$, where

$$\mathrm{A}(M, v) := \sum_{k,l=1}^q M_{kl} f_k f_l^* + \sum_{k=1}^q v_k P_k.$$

Remark 7.3. Note that $\mathrm{A}(M, v) \simeq M \oplus v_1 \mathbf{1}_{|C_1|-1} \oplus \cdots \oplus v_q \mathbf{1}_{|C_q|-1}$, so that we have $\lambda_{\max}(\mathrm{A}(M, v)) = \max\{\lambda_{\max}(M), \max_i v_i\}$. This will be used repeatedly below.

The following two lemmas show that $\mathcal{A}$ satisfies the assumptions of Lemma 7.1.

Lemma 7.4. $\mathbf{E}X = \mathrm{A}\big(\operatorname{diag}(c)^{\frac{1}{2}} B \operatorname{diag}(c)^{\frac{1}{2}} - \operatorname{diag}(Bc), -Bc\big) \in \mathcal{A}$.

Proof. Note that $\mathbf{B} = \sum_{k,l=1}^q B_{kl} \mathrm{I}_{C_k} \mathrm{I}_{C_l}^*$, so

$$\frac{1}{d} \operatorname{diag}(z) \mathbf{B} \operatorname{diag}(z) = \sum_{k,l=1}^q \sqrt{c_k} B_{kl} \sqrt{c_l}\, f_k f_l^*, \qquad \frac{1}{d} \mathbf{B} 1_d = \sum_{k,l=1}^q B_{kl} c_l\, \mathrm{I}_{C_k}.$$

The conclusion follows using $\mathbf{I}_{C_k} = P_k + f_k f_k^*$. □

Lemma 7.5. $\mathbf{E}[(X - \mathbf{E}X) A (X - \mathbf{E}X)] \in \mathcal{A}$ *for every* $A \in \mathcal{A}$. *More precisely,*

$$\begin{aligned}\mathbf{E}[(X - \mathbf{E}X)\, \mathrm{A}(M, v)\, (X - \mathbf{E}X)] = \\ \mathrm{A}\bigg(\operatorname{diag}\Big(B\Big(c \odot v + \frac{1}{d}(\operatorname{diag}^{-1}(M) - v)\Big)\Big) + \frac{1}{d} B \odot M^T, \\ B\Big(c \odot v + \frac{1}{d}(\operatorname{diag}^{-1}(M) - v)\Big) + \frac{1}{d} v \odot \operatorname{diag}^{-1}(B)\bigg)\end{aligned}$$

for all $M \in \mathrm{M}_q(\mathbb{C})$ *and* $v \in \mathbb{C}^q$, *where* M^T *denotes the transpose of* M.

Proof. Recall that $X - \mathbf{E}X = G$ (cf. section 2.4). For any $A \in \mathrm{M}_d(\mathbb{C})$, we compute

$$\begin{aligned}\mathbf{E}[GAG]_{kl} &= \sum_{r,s} A_{rs}\, \mathbf{E}[G_{kr} G_{sl}] = \frac{1}{d} \sum_{r,s} \mathbf{B}_{kr} A_{rs} (1_{k=s, r=l} + 1_{k=l, r=s}) \\ &= \Big(\frac{1}{d} \mathbf{B} \odot A^T + \frac{1}{d} \operatorname{diag}(\mathbf{B}(\operatorname{diag}^{-1}(A))\Big)_{kl}.\end{aligned}$$

We readily compute $\mathbf{B} \odot \mathrm{A}(M, v)^T = \mathrm{A}(B \odot M^T, v \odot \operatorname{diag}^{-1}(B))$, while

$$\frac{1}{d} \operatorname{diag}(\mathbf{B}(\operatorname{diag}^{-1}(\mathrm{A}(M, v))) = \sum_k \Big[B\Big(c \odot v + \frac{1}{d}(\operatorname{diag}^{-1}(M) - v)\Big)\Big]_k \mathbf{I}_{C_k}$$

using $\mathbf{B} = \sum_{k,l} B_{kl} \mathrm{I}_{C_k} \mathrm{I}_{C_l}^*$. The result follows as $\mathbf{I}_{C_k} = P_k + f_k f_k^*$. □

7.2. **The simplified parameters.** We now aim to approximate $\lambda_{\max}(X_{\mathrm{free}})$ and $\lambda_{\max}(X_{\varnothing,\mathrm{free}})$ by simplified parameters $\lambda, \lambda_\varnothing$: we will use the reduction principle of the previous section to reduce the variational principle (1.6) for d-dimensional matrices to a variational principle for q-dimensional vectors, and we will eliminate all the terms of order $\frac{1}{d}$ in Lemmas 7.4 and 7.5. We first consider λ.

Proposition 7.6. *Define*

$$\lambda := \inf_{v>0} \max\big\{\lambda_{\max}\big(\operatorname{diag}(c)^{\frac12} B \operatorname{diag}(c)^{\frac12} + \operatorname{diag}\big(B\operatorname{diag}(c)(v-1_q)\big)\big),\\ \lambda_{\max}\big(\operatorname{diag}(v)^{-1} + \operatorname{diag}\big(B\operatorname{diag}(c)(v-1_q)\big)\big)\big\}.$$

Then

$$|\lambda_{\max}(X_{\mathrm{free}}) - \lambda| \le \sqrt{\frac{8\|B1_q\|_\infty}{d}}.$$

Proof. The Lehner formula (1.6) and Lemmas 7.1, 7.4, and 7.5 yield

$$\lambda_{\max}(X_{\mathrm{free}}) = \inf_{M,v>0} \lambda_{\max}\big(\mathbf{E}X + \mathrm{A}(M^{-1}, \tfrac{1}{v}) + \mathbf{E}[G\,\mathrm{A}(M,v)\,G]\big)$$

using $\mathrm{A}(M,v)^{-1} = \mathrm{A}(M^{-1}, \frac{1}{v})$. Moreover,

$$\lambda = \inf_{v>0} \lambda_{\max}\big(\mathbf{E}X + \mathrm{A}(0, \tfrac{1}{v}) + \mathrm{A}(\operatorname{diag}(B\operatorname{diag}(c)v), B\operatorname{diag}(c)v)\big)$$

by Lemma 7.4. We must upper and lower bound $\lambda_{\max}(X_{\mathrm{free}})$ in terms of λ.

Upper bound. We can read off from Lemma 7.5 that

$$\mathbf{E}[G\,\mathrm{A}(\mathbf{1}_q, 0)\,G] \le \frac{2\|B1_q\|_\infty}{d}, \quad \mathbf{E}[G\,\mathrm{A}(0,v)\,G] \le \mathrm{A}(\operatorname{diag}(B\operatorname{diag}(c)v), B\operatorname{diag}(c)v)$$

for $v>0$, where we used $v \odot \operatorname{diag}^{-1}(B) - Bv \le 0$. Restricting the infimum over M in the variational principle for $\lambda_{\max}(X_{\mathrm{free}})$ to $M = \gamma \mathbf{1}_q$ for $\gamma > 0$ yields

$$\lambda_{\max}(X_{\mathrm{free}}) \le \inf_{v>0} \lambda_{\max}\big(\mathbf{E}X + \mathrm{A}(0,\tfrac{1}{v}) + \mathbf{E}[G\,\mathrm{A}(0,v)\,G]\big) + \inf_{\gamma>0}\left(\frac{1}{\gamma} + \frac{2\gamma\|B1_q\|_\infty}{d}\right)\\ \le \lambda + \sqrt{\frac{8\|B1_q\|_\infty}{d}},$$

where we used that $\mathrm{A}(M,v) = \mathrm{A}(M,0) + \mathrm{A}(0,v)$.

Lower bound. We can read off from Lemma 7.5 that

$$\mathbf{E}[G\,\mathrm{A}(0,v)\,G] \ge \mathrm{A}(\operatorname{diag}(B\operatorname{diag}(c - \tfrac{1}{d}1_q)v), B\operatorname{diag}(c-\tfrac{1}{d}1_q)v)$$

for $v>0$. Then we can estimate

$$\lambda = \inf_{v,w>0} \lambda_{\max}\big(\mathbf{E}X + \mathrm{A}(0, \tfrac{1}{v\wedge w}) + \tfrac{1}{d}\mathrm{A}(\operatorname{diag}(B(v\wedge w)), B(v\wedge w))\\ + \mathrm{A}(\operatorname{diag}(B\operatorname{diag}(c-\tfrac{1}{d}1_q)(v\wedge w)), B\operatorname{diag}(c-\tfrac{1}{d}1_q)(v\wedge w))\big)\\ \le \inf_{v,w>0} \lambda_{\max}\big(\mathbf{E}X + \mathrm{A}(0,\tfrac{1}{v}) + \mathrm{A}(0,\tfrac{1}{w}) + \tfrac{1}{d}\mathrm{A}(\operatorname{diag}(Bw), Bw) + \mathbf{E}[G\,\mathrm{A}(0,v)\,G]\big)\\ \le \lambda_{\max}(X_{\mathrm{free}}) + \inf_{w>0} \lambda_{\max}\big(\mathrm{A}(0,\tfrac{1}{w}) + \tfrac{1}{d}\mathrm{A}(\operatorname{diag}(Bw), Bw)\big),$$

where $v\wedge w$ denotes the elementwise minimum, and we used $\frac{1}{v\wedge w} \le \frac{1}{v} + \frac{1}{w}$. Choosing $w \leftarrow \sqrt{d}\|B1_q\|_\infty^{-\frac12} 1_q$ on the last line concludes the proof. □

The parameter $\lambda_\varnothing$ arises in a completely analogous fashion.

Proposition 7.7. *Define*

$$\lambda_\varnothing := \inf_{v>0} \lambda_{\max}\big(\operatorname{diag}(v)^{-1} + \operatorname{diag}\big(B\operatorname{diag}(c)(v-1_q)\big)\big).$$

Then

$$|\lambda_{\max}(X_{\varnothing,\mathrm{free}}) - \lambda_\varnothing| \le \sqrt{\frac{8\|B1_q\|_\infty}{d}}.$$

Proof. Note that the only difference between the definitions of X and $X_\varnothing$ is that $\mathbf{E}X$ is replaced by $\mathbf{E}X_\varnothing = \mathrm{A}(-\operatorname{diag}(Bc), -Bc) \in \mathcal{A}$ in Lemma 7.4. Thus the proof of Proposition 7.6 carries over verbatim to the present setting. □

We can now prove the upper bound on $\lambda_\varnothing$ in Theorem 2.12. Recall that $b > 0$ denotes the Perron-Frobenius (right) eigenvector of $B\operatorname{diag}(c)$.

Lemma 7.8. *We have*

$$\lambda_\varnothing \le 1 - \frac{\min_i b_i}{\max_i b_i}\big(1 - \lambda_{\max}(\operatorname{diag}(c)^{\frac12} B \operatorname{diag}(c)^{\frac12})^{\frac12}\big)^2.$$

Proof. Denote $\lambda_{\mathrm{cr}} := \lambda_{\max}(\operatorname{diag}(c)^{\frac12} B \operatorname{diag}(c)^{\frac12})$ for simplicity. As $\operatorname{diag}(c)^{\frac12} b$ is a positive eigenvector of $\operatorname{diag}(c)^{\frac12} B \operatorname{diag}(c)^{\frac12}$, the Perron-Frobenius theorem implies that its eigenvalue must be maximal, and thus $B\operatorname{diag}(c)b = \lambda_{\mathrm{cr}} b$. We now upper bound $\lambda_\varnothing$ by restricting the infimum in its definition to $v = 1_q + tb$. This yields

$$\lambda_\varnothing \le \inf_{t:1_q+tb>0} \max_i \left\{\frac{1}{1+tb_i} + t\lambda_{\mathrm{cr}} b_i\right\} = 1 - \sup_{t:1_q+tb>0} \min_i \left\{\frac{tb_i}{1+tb_i} - t\lambda_{\mathrm{cr}} b_i\right\}.$$

If we choose $t = (\lambda_{\mathrm{cr}}^{-\frac12} - 1)\frac{1}{\max_i b_i}$, then $1_q + tb > 0$ and

$$\min_i \left\{\frac{tb_i}{1+tb_i} - t\lambda_{\mathrm{cr}} b_i\right\} = \min_i \left\{\frac{1}{\lambda_{\mathrm{cr}}^{\frac12} + (1-\lambda_{\mathrm{cr}}^{\frac12})\frac{b_i}{\max_i b_i}} - \lambda_{\mathrm{cr}}^{\frac12}\right\}(1-\lambda_{\mathrm{cr}}^{\frac12})\frac{b_i}{\max_i b_i}.$$

The conclusion follows as

$$\left(\frac{1}{x+(1-x)a} - x\right)(1-x) \ge (1-x)^2$$

for all $x > 0$ and $0 \le a \le 1$. □

7.3. **The phase transition.** It remains to prove the phase transition for λ. To this end, we first develop in section 7.3.1 some basic properties of the minimizers in the definitions of λ and $\lambda_\varnothing$. While we will restrict attention to the present model, the methods used here are quite general and extend to other Lehner-type variational principles. We then exploit the special structure of the present model in section 7.3.2 to complete the proof of Theorem 2.12.

Before we begin the proof, let us make a minor simplification: while we assumed only that the matrix B is irreducible, we can assume without loss of generality that B has strictly positive entries in the remainder of the proof. Indeed, it is clear that all the quantities that appear in Theorem 2.12 are continuous in B. When $\lambda_{\max}(\operatorname{diag}(c)^{\frac12} B \operatorname{diag}(c)^{\frac12}) \ne 1$, we can apply the result for $B \leftarrow B + \varepsilon 1_q 1_q^*$ and let $\varepsilon \downarrow 0$ (the preservation of the strict inequality $\lambda_\varnothing < 1$ in the limit follows from the quantitative estimate on $\lambda_\varnothing$). The case $\lambda_{\max}(\operatorname{diag}(c)^{\frac12} B \operatorname{diag}(c)^{\frac12}) = 1$ now follows by applying the result to $B \leftarrow tB$ and letting $t \to 1$ from above and below.

7.3.1. *Basic properties of the minimizers.* The following basic but important result collects a number of general properties of the variational principle that defines $\lambda_\varnothing$: existence and uniqueness of a minimizer, and first-order optimality conditions.

Lemma 7.9. *The infimum in the definition of $\lambda_\varnothing$ (Proposition 7.7) is attained at a unique vector $v_\varnothing^* > 0$. Moreover, this minimizer satisfies the optimality conditions*

$$\frac{1}{v_\varnothing^*} + B\,\mathrm{diag}(c)(v_\varnothing^* - 1_q) = \lambda_\varnothing 1_q \tag{7.1}$$

and

$$\lambda_{\max}\big(\,\mathrm{diag}(c)^{\frac12} B\,\mathrm{diag}(c)^{\frac12} - \mathrm{diag}(v_\varnothing^*)^{-2}\big) = 0. \tag{7.2}$$

Proof. Let us write $\lambda_\varnothing = \inf_{v>0} f(v)$ with $f(v) = \max_i(\frac{1}{v} + B\,\mathrm{diag}(c)(v - 1_q))_i$. The existence of a minimizer $v_\varnothing^* > 0$ follows by a routine compactness argument and as $f(v)$ diverges if $v_i \to \{0, \infty\}$ for any i.

Next, we show that (7.1) must hold for *any* minimizer. Suppose $v > 0$ satisfies

$$\left(\frac{1}{v} + B\,\mathrm{diag}(c)(v - 1_q)\right)_j < \max_i \left(\frac{1}{v} + B\,\mathrm{diag}(c)(v - 1_q)\right)_i = \lambda_\varnothing$$

for some j. As B, c have positive entries, slightly decreasing v_j will strictly decrease all $(\frac{1}{v} + B\,\mathrm{diag}(c)(v-1_q))_i$ for $i \neq j$ while preserving $(\frac{1}{v} + B\,\mathrm{diag}(c)(v-1_q))_j < \lambda_\varnothing$. The perturbed v would therefore satisfy $f(v) < \lambda_\varnothing$, contradicting the definition of $\lambda_\varnothing$. We conclude that any minimizer must satisfy (7.1).

We now prove uniqueness. Let $\lambda_\varnothing = f(v_0) = f(v_1)$ and define $v_t = (1-t)v_0 + tv_1$ for $t \in [0,1]$. As f is convex, we have $\lambda_\varnothing \le f(v_t) \le (1-t)f(v_0) + tf(v_1) = \lambda_\varnothing$, so v_t is also a minimizer. As we have shown (7.1) holds for any minimizer, we have

$$0 = \frac{d^2}{dt^2}\left(\frac{1}{v_t} + B\,\mathrm{diag}(c)(v_t - 1_q)\right)_i = 2v_{t,i}^{-3}(v_1 - v_0)_i^2$$

for all i, which implies $v_0 = v_1$. Thus the minimizer is unique.

It remains to prove (7.2). Note that as B, c have positive entries, the Perron-Frobenius theorem yields[5] an eigenvector $w > 0$ associated to the maximal eigenvalue μ of $\mathrm{diag}(c)^{\frac12} B\,\mathrm{diag}(c)^{\frac12} - \mathrm{diag}(v_\varnothing^*)^{-2}$. Let $v_t = v_\varnothing^* - t\mu\,\mathrm{diag}(c)^{-\frac12} w$, so that

$$\frac{d}{dt}\left(\frac{1}{v_t} + B\,\mathrm{diag}(c)(v_t - 1_q)\right)\bigg|_{t=0} = -\mu^2\,\mathrm{diag}(c)^{-\frac12} w.$$

If $\mu \neq 0$, all entries of this vector are strictly negative, which would imply that $f(v_t) < f(v_0) = \lambda_\varnothing$ for t sufficiently small. This contradicts the definition of $\lambda_\varnothing$. We must therefore have $\mu = 0$, which is (7.2). □

We now prove a partial counterpart of the above lemma for the variational principle that defines λ. While more information could be obtained also in this case, we only prove the properties that will be needed below.

Lemma 7.10. *The infimum in the definition of λ (Proposition 7.6) is attained at a vector $v^* > 0$. Moreover, this minimizer satisfies*

$$\frac{1}{v^*} + B\,\mathrm{diag}(c)(v^* - 1_q) = \lambda 1_q \tag{7.3}$$

[5] If M is a self-adjoint matrix with nonnegative off-diagonal entries, $M + c\mathbf{1}$ is a nonnegative matrix for sufficiently large c. We can therefore apply the Perron-Frobenius theorem to the latter to deduce the existence of a positive eigenvector of M associated to its maximal eigenvalue.

and

$$\lambda_{\max}\big(\operatorname{diag}(c)^{\frac12} B \operatorname{diag}(c)^{\frac12} - \operatorname{diag}(v^*)^{-1}\big) \le 0. \tag{7.4}$$

Proof. The existence of a minimizer $v_* > 0$ follows as in the proof of Lemma 7.9. Now suppose there is a coordinate j so that v^* satisfies

$$\left(\frac{1}{v^*} + B\operatorname{diag}(c)(v^* - 1_q)\right)_j < \lambda.$$

As B, c have positive entries, we can reason as in the proof of Lemma 7.9 that slightly decreasing the jth coordinate of v^* will yield a strictly smaller value of the function being minimized in the definition of λ, contradicting the minimality of v^*. We conclude that v^* must satisfy (7.3). Finally, (7.4) follows from (7.3) and as

$$\lambda_{\max}\big(\operatorname{diag}(c)^{\frac12} B \operatorname{diag}(c)^{\frac12} + \operatorname{diag}\big(B\operatorname{diag}(c)(v^* - 1_q)\big)\big) \le \lambda$$

by the definition of λ. □

It is obvious from the definitions of $\lambda, \lambda_\varnothing$ that $\lambda_\varnothing \le \lambda$. The aim of the remainder of the proof is to characterize the phase transition from $\lambda_\varnothing < \lambda$ to $\lambda_\varnothing = \lambda$. A basic characterization of the phase regions follows directly from the variational principles.

Lemma 7.11. *$\lambda = \lambda_\varnothing$ if and only if $\lambda_{\max}(\operatorname{diag}(c)^{\frac12} B \operatorname{diag}(c)^{\frac12} - \operatorname{diag}(v_\varnothing^*)^{-1}) \le 0$.*

Proof. If $\lambda_{\max}(\operatorname{diag}(c)^{\frac12} B \operatorname{diag}(c)^{\frac12} - \operatorname{diag}(v_\varnothing^*)^{-1}) \le 0$, then choosing $v \leftarrow v_\varnothing^*$ in the definition of λ (cf. Proposition 7.6) and using (7.1) yields $\lambda \le \lambda_\varnothing$. As $\lambda_\varnothing \le \lambda$ holds trivially by the definitions of $\lambda, \lambda_\varnothing$, we conclude that $\lambda = \lambda_\varnothing$.

Now suppose that $\lambda = \lambda_\varnothing$. Then

$$\lambda_{\max}\big(\operatorname{diag}(v^*)^{-1} + \operatorname{diag}\big(B\operatorname{diag}(c)(v^* - 1_q)\big)\big) \le \lambda = \lambda_\varnothing$$

by the definition of λ, which implies that v^* is a minimizer in the definition of $\lambda_\varnothing$ (cf. Proposition 7.7). But Lemma 7.9 shows the latter is unique, so that $v^* = v_\varnothing^*$. Thus (7.4) yields $\lambda_{\max}(\operatorname{diag}(c)^{\frac12} B \operatorname{diag}(c)^{\frac12} - \operatorname{diag}(v_\varnothing^*)^{-1}) \le 0$. □

The difficulty in applying this lemma is that the phase transition criterion is not explicit as it involves $v_\varnothing^*$. In the rest of the proof, we will exploit the special properties of the present model to explicitly characterize the phase transition.

7.3.2. *Proof of Theorem 2.12.* The following fact could be viewed as the basic reason behind the special properties of the present model.

Lemma 7.12. *Suppose a vector $v > 0$ and $\mu \in \mathbb{R}$ satisfy*

$$\frac{1}{v} + B\operatorname{diag}(c)(v - 1_q) = \mu 1_q, \qquad \lambda_{\max}\big(\operatorname{diag}(c)^{\frac12} B \operatorname{diag}(c)^{\frac12} - \operatorname{diag}(v)^{-1}\big) = 0.$$

Then we must have $\mu = 1$.

Proof. The key idea is that the first equation in the statement is equivalent to

$$\big(\operatorname{diag}(c)^{\frac12} B \operatorname{diag}(c)^{\frac12} - \operatorname{diag}(v)^{-1}\big)\operatorname{diag}(c)^{\frac12}(v - 1_q) = (\mu - 1)\operatorname{diag}(c)^{\frac12} 1_q. \tag{7.5}$$

As B, c have positive entries, the Perron-Frobenius theorem and the second equation in the statement yield an eigenvector $w > 0$ of $\operatorname{diag}(c)^{\frac12} B \operatorname{diag}(c)^{\frac12} - \operatorname{diag}(v)^{-1}$ with eigenvalue 0. Taking the inner product of the above equation with w yields $0 = (\mu - 1)\langle w, \operatorname{diag}(c)^{\frac12} 1_q\rangle$, which implies $\mu = 1$ as $\langle w, \operatorname{diag}(c)^{\frac12} 1_q\rangle > 0$. □

Using this result, we can explicitly determine $v_\varnothing^*$ on the boundary of the phase region $\lambda = \lambda_\varnothing$ (cf. Lemma 7.11). This is the key step in the proof.

Lemma 7.13. *If* $\lambda_{\max}(\mathrm{diag}(c)^{\frac12} B\,\mathrm{diag}(c)^{\frac12} - \mathrm{diag}(v_\varnothing^*)^{-1}) = 0$, *then* $v_\varnothing^* = 1_q$.

Proof. By Lemma 7.12, the assumption and (7.1) imply that $\lambda_\varnothing = 1$. Thus

$$\big(\mathrm{diag}(c)^{\frac12} B\,\mathrm{diag}(c)^{\frac12} - \mathrm{diag}(v_\varnothing^*)^{-1}\big)\,\mathrm{diag}(c)^{\frac12}(v_\varnothing^* - 1_q) = 0$$

by (7.5). Now note that the Perron-Frobenius theorem and (7.2) provide an eigenvector $w > 0$ of $\mathrm{diag}(c)^{\frac12} B\,\mathrm{diag}(c)^{\frac12} - \mathrm{diag}(v_\varnothing^*)^{-2}$ with eigenvalue 0. Taking the inner product of the above equation with w yields

$$0 = \langle w, (\mathrm{diag}(v_\varnothing^*)^{-2} - \mathrm{diag}(v_\varnothing^*)^{-1})\,\mathrm{diag}(c)^{\frac12}(v_\varnothing^* - 1_q)\rangle = -\sum_i w_i c_i^{\frac12}\big((v_\varnothing^*)_i^{-1} - 1\big)^2,$$

which evidently implies $v_\varnothing^* = 1_q$. □

Lemmas 7.11 and 7.13 show that we must have $\lambda_{\max}(\mathrm{diag}(c)^{\frac12} B\,\mathrm{diag}(c)^{\frac12}) = 1$ on the boundary between the phase regions. This will enable us to fully characterize the phase regions by a continuity argument, completing the proof of Theorem 2.12

Proof of Theorem 2.12. The approximation by $\lambda, \lambda_\varnothing$ and the estimate on $\lambda_\varnothing$ were proved in Propositions 7.6 and 7.7 and in Lemma 7.8, respectively. The remainder of the proof will be completed in three steps to be proved below:

1. $\lambda_{\max}(\mathrm{diag}(c)^{\frac12} B\,\mathrm{diag}(c)^{\frac12}) > 1$ implies $\lambda_\varnothing < \lambda$.
2. $\lambda_\varnothing < \lambda$ implies $\lambda = 1$.
3. $\lambda_{\max}(\mathrm{diag}(c)^{\frac12} B\,\mathrm{diag}(c)^{\frac12}) < 1$ implies $\lambda < 1$.

Indeed, combining steps 1 and 2 yields part c of the theorem, while combining steps 2 and 3 yields part a of the theorem (as $\lambda_\varnothing \le \lambda$). Part b of the theorem now follows by applying the theorem to $B \leftarrow tB$ and letting $t \to 1$ from above and below.

It remains to prove each of the above steps.

Step 1. Suppose $\lambda_{\max}(\mathrm{diag}(c)^{\frac12} B\,\mathrm{diag}(c)^{\frac12}) > 1$. By Lemma 7.11 and as $\lambda_\varnothing \le \lambda$, it suffices to show that $\lambda_{\max}(\mathrm{diag}(c)^{\frac12} B\,\mathrm{diag}(c)^{\frac12} - \mathrm{diag}(v_\varnothing^*)^{-1}) > 0$.

Consider first the special case $B = 2\,1_q 1_q^*$, so that $\lambda_{\max}(\mathrm{diag}(c)^{\frac12} B\,\mathrm{diag}(c)^{\frac12}) = 2$ as $\sum_i c_i = 1$. Then (7.1) shows that $v_\varnothing^*$ is proportional to 1_q, so it suffices to minimize over $v \leftarrow t1_q$ in the definition of $\lambda_\varnothing$. A straightforward computation yields $v_\varnothing^* = 2^{-\frac12} 1_q$ and thus $\lambda_{\max}(\mathrm{diag}(c)^{\frac12} B\,\mathrm{diag}(c)^{\frac12} - \mathrm{diag}(v_\varnothing^*)^{-1}) > 0$.

For general B, choose a continuous family $t \mapsto B(t)$ so that $B(0) = 2\,1_q 1_q^*$, $B(1) = B$, and $\lambda_{\max}(\mathrm{diag}(c)^{\frac12} B(t)\,\mathrm{diag}(c)^{\frac12}) > 1$ to all $t \in [0,1]$. Denote by $v_\varnothing^*(t)$ the minimizer in the definition of $\lambda_\varnothing$ for $B \leftarrow B(t)$. As the minimizer $v_\varnothing^*(t)$ is unique by Lemma 7.9, it follows by a routine argument that $t \mapsto v_\varnothing^*(t)$ is continuous. On the other hand, Lemma 7.13 ensures that for all $t \in [0,1]$

$$\alpha(t) := \lambda_{\max}(\mathrm{diag}(c)^{\frac12} B(t)\,\mathrm{diag}(c)^{\frac12} - \mathrm{diag}(v_\varnothing^*(t))^{-1}) \neq 0 :$$

otherwise we would have $v_\varnothing^*(t) = 1_q$ and thus $\lambda_{\max}(\mathrm{diag}(c)^{\frac12} B(t)\,\mathrm{diag}(c)^{\frac12}) = 1$ for some t, which entails a contradiction. As we showed that $\alpha(0) > 0$ and $\alpha(t) \neq 0$ for all t, it follows by continuity that $\alpha(1) > 0$. This is the desired claim.

Step 2. Suppose that $\lambda_\varnothing < \lambda$. To show this implies $\lambda = 1$, it suffices by (7.3) and Lemma 7.12 to show that $\lambda_{\max}(\mathrm{diag}(c)^{\frac12} B \,\mathrm{diag}(c)^{\frac12} - \mathrm{diag}(v^*)^{-1}) = 0$.

Suppose the latter is not the case. Then (7.3) and the definition of λ imply

$$\lambda_{\max}\big(\mathrm{diag}(c)^{\frac12} B \,\mathrm{diag}(c)^{\frac12} + \mathrm{diag}(B\,\mathrm{diag}(c)(v^* - 1_q))\big) < \lambda \\ = \lambda_{\max}\big(\mathrm{diag}(v^*)^{-1} + \mathrm{diag}(B\,\mathrm{diag}(c)(v^* - 1_q))\big).$$

Then v^* must also be a minimizer of the quantity on the second line: otherwise we could slightly decrease the quantity on the second line while preserving the strict inequality on the first line, contradicting the definition of λ. This implies by the definition of $\lambda_\varnothing$ that $\lambda = \lambda_\varnothing$, which contradicts the assumption of step 2.

Step 3. Suppose that $\lambda_{\max}(\mathrm{diag}(c)^{\frac12} B \,\mathrm{diag}(c)^{\frac12}) < 1$. Then it follows readily that $\lambda \le 1$ by choosing $v \leftarrow 1_q$ in the definition of λ.

Now suppose that $\lambda = 1$. Then $v^* = 1_q$ would be a minimizer in the definition of λ. The same argument as in the proof of step 2 now shows that v^* must also be a minimizer in the definition of $\lambda_\varnothing$, so that $v^*_\varnothing = 1_q$. The latter contradicts (7.2). Thus we have shown that $\lambda < 1$, concluding the proof. □

8. Applications: proofs

8.1. **Decoding node labels on graphs.**

Proof of Theorem 3.4. Define

$$Y' := \frac{Y}{(4kp(1-p))^{\frac12}}, \qquad \theta' := \frac{k^{\frac12}(1-2p)}{(4p(1-p))^{\frac12}}.$$

Then we clearly have

$$\mathbf{E}Y' = \frac{\theta'}{k}\,\mathrm{diag}(x) A \,\mathrm{diag}(x), \qquad \mathbf{E}[(Y' - \mathbf{E}Y')^2] = \mathbf{1}.$$

Moreover, as $A1_d = k1_d$, the Perron-Frobenius theorem yields

$$\lambda_{\max}(\mathbf{E}Y') = \theta', \qquad v_{\max}(\mathbf{E}Y') = d^{-\frac12} x,$$

while $1_{(\theta'-\delta,\theta']}(\mathbf{E}Y') = d^{-1} x x^*$ for $\delta := \frac{\theta'}{k}\lambda$. Note for future reference that the assumptions of the theorem imply that $\theta' = (1 + o(1))\theta$ and $k \gg (\log d)^4$.

Let $A = \sum_{i=1}^d \lambda_i v_i v_i^*$ be an eigendecomposition of A so that $\lambda_1 = k$ and $|\lambda_i| = \mathrm{s}_i$. Then $A_r := \sum_{i=1}^r \lambda_i v_i v_i^*$ has rank at most r and $\|A - A_r\| \le \mathrm{s}_{r+1}$. Define

$$X := Y' - \mathbf{E}Y' + \frac{\theta'}{k}\,\mathrm{diag}(x) A_r \,\mathrm{diag}(x).$$

As $\mathbf{E}Y = (1-2p)\,\mathrm{diag}(x) A \,\mathrm{diag}(x)$, we can estimate

$$\|(4kp(1-p))^{\frac12} X - Y\| \le k^{-\frac12}\theta\,\mathrm{s}_{r+1}.$$

On the other hand, we have

$$\mathbf{P}\big[|\lambda_{\max}(X) - \mathrm{B}(\theta')| > Ck^{-\frac12}\sqrt{r} + Ck^{-\frac16}(\log d)^{\frac23}\big] \le \frac{C}{d^2}$$

by applying Theorems 2.2, 2.4, and 2.7 with $t = 3\log d$ and using that $\sigma(X) = 1$, $\sigma_*(X) \le v(X) \lesssim k^{-\frac12}$, $R \lesssim k^{-\frac12}$, and $k \gg (\log d)^4$. Therefore

$$\mathbf{P}\Big[|\lambda_{\max}(Y') - \mathrm{B}(\theta')| > Ck^{-1}\Big\{\min_{1\le r\le k}\big\{\theta\,\mathrm{s}_{r+1} + \sqrt{rk}\big\} + k^{\frac56}(\log d)^{\frac23}\Big\}\Big] \le \frac{C}{d^2},$$

where we optimized over the choice of r. The analogous estimates follow readily if we replace $Y' \leftarrow Y' + s1_{(\theta'-\delta,\theta']}(\mathbf{E}Y')$ and $\theta' \leftarrow \theta' + s$ for $|s| \le \delta$.

To proceed, note that the assumption of the theorem and $\theta' = (1+o(1))\theta$ imply

$$k^{-1}\left\{\min_{1\le r\le k}\left\{\theta\,\mathrm{s}_{r+1} + \sqrt{rk}\right\} + k^{\frac{5}{6}}(\log d)^{\frac{2}{3}}\right\} \ll \min\{\delta, 1\}.$$

We can therefore conclude using Theorem 2.9 that

$$\mathbf{P}\left[\left|\frac{1}{d}|\langle x, v_{\max}(Y)\rangle|^2 - \left(1 - \frac{1}{\theta^2}\right)_+\right| > t + \frac{o(\min\{\delta,1\})}{t}\right] \le \frac{C}{d^2}$$

for $0 < t \le \delta$. It remains to choose $0 < t \le \delta$ so that $t + \frac{o(\min\{\delta,1\})}{t} = o(1)$.

Finally, the existence of an estimator $\hat{x}(Y)$ follows from Lemma 8.1 below. □

At the end of the proof we used the following general rounding procedure.

Lemma 8.1. *Let $x \in \{-1,+1\}^d$ and $v \in S^{d-1}$ satisfy $\frac{1}{d}|\langle x, v\rangle|^2 \ge \varepsilon$. Then there exists a randomized estimator $\hat{x} \in \{-1,+1\}^d$, whose construction depends only on d, v, ε, such that $\frac{1}{d}|\langle x, \hat{x}\rangle| \ge \frac{\varepsilon}{8}$ with probability $1 - \frac{64}{d\varepsilon^2}$.*

Proof. Fix $c > 0$ that will be chosen shortly, and construct $\hat{x}$ by choosing each entry to be an independent random sign so that $\mathbf{E}[\hat{x}_i] = \frac{v_i\sqrt{d}}{c}1_{|v_i|\sqrt{d}\le c}$. Then

$$\frac{1}{d}|\mathbf{E}[\langle x, \hat{x}\rangle]| = \left|\sum_{i\in[d]}\frac{x_i v_i}{c\sqrt{d}}1_{|v_i|\sqrt{d}\le c}\right| \ge \frac{\sqrt{\varepsilon}}{c} - \left|\sum_{i\in[d]}\frac{x_i v_i}{c\sqrt{d}}1_{|v_i|\sqrt{d}> c}\right| \ge \frac{\sqrt{\varepsilon}}{c} - \frac{1}{c^2}.$$

Choosing $c = \frac{2}{\sqrt{\varepsilon}}$ yields $\frac{1}{d}\mathbf{E}|\langle x, \hat{x}\rangle| \ge \frac{1}{d}|\mathbf{E}[\langle x, \hat{x}\rangle]| \ge \frac{\varepsilon}{4}$.

Now note that $\mathrm{Var}(\frac{1}{d}|\langle x, \hat{x}\rangle|) \le \mathrm{Var}(\frac{1}{d}\langle x, \hat{x}\rangle) \le \frac{1}{d}$. We can therefore estimate

$$\mathbf{P}\left[\tfrac{1}{d}|\langle x, \hat{x}\rangle| < \tfrac{\varepsilon}{4} - t\right] \le \mathbf{P}\left[\left|\tfrac{1}{d}|\langle x, \hat{x}\rangle| - \tfrac{1}{d}\mathbf{E}|\langle x, \hat{x}\rangle|\right| > t\right] \le \frac{1}{dt^2}$$

by Chebyshev's inequality. Choosing $t = \frac{\varepsilon}{8}$ yields the conclusion. □

8.2. **Tensor PCA.** We begin with some basic observations.

Lemma 8.2. *M is a $d \times d$ self-adjoint random matrix with $d = \binom{n}{\ell}$, such that*

$$\mathbf{E}[(M - \mathbf{E}M)^2] = \sigma(M)^2\mathbf{1} \qquad \text{with} \qquad \sigma(M)^2 = \binom{\ell}{p/2}\binom{n-\ell}{p/2}.$$

Moreover, we have

$$v(M)^2 = \binom{p}{p/2}\binom{n-p}{\ell - p/2}.$$

In particular, $\sigma(M)^2 \asymp n^{\frac{p}{2}}$ and $v(M)^2 \asymp n^{\ell - \frac{p}{2}}$ as $n \to \infty$ (with p, ℓ fixed).

Proof. We readily compute for $|S| = |T| = \ell$

$$\mathbf{E}[(M - \mathbf{E}M)^2]_{S,T} = \sum_{|R|=\ell}\mathbf{E}[Z_{S\triangle R}Z_{T\triangle R}] = |\{R \subseteq [n] : |R| = \ell, |R\triangle S| = p\}|\,1_{S=T}$$

using $S\triangle R = T\triangle R$ if and only if $S = T$. As $|R| = |S|$, we have $|R\backslash S| = |S\backslash R| = \frac{1}{2}|R\triangle S|$, so each R satisfying $|R| = \ell, |R\triangle S| = p$ is formed by replacing $\frac{p}{2}$ elements of S by $\frac{p}{2}$ elements not in S. The number of all such R is evidently σ^2.

Now note that $|S| = |T| = \ell$ and $|U| = p$ satisfy $M_{S,T} - \mathbf{E}M_{S,T} = Z_U$ if and only if $S\triangle T = U$. Thus given U, all such S, T are formed by choosing $\frac{p}{2}$ elements of U to place in S (the remaining ones are placed in T), then choosing $\ell - \frac{p}{2}$ elements

not in U to place in $S \cap T$. Thus there are $m := \binom{p}{p/2}\binom{n-p}{\ell-p/2}$ matrix elements of $M - \mathbf{E}M$ that coincide with each independent standard Gaussian variable Z_U. As $\mathrm{Cov}(M)$ is block-diagonal with blocks of the form $1_m 1_m^*$, the conclusion follows. □

Next, we show that $\mathbf{E}M$ is approximately of low rank.

Lemma 8.3. *Let $\mathrm{s}_1 \ge \cdots \ge \mathrm{s}_d$ be the singular values of $\mathbf{E}M$. Then*

$$\lambda_{\max}(\mathbf{E}M) = \mathrm{s}_1 = \binom{\ell}{p/2}\binom{n-\ell}{p/2}\lambda, \qquad \mathrm{s}_{r+1} \le \frac{p}{n}\mathrm{s}_1,$$

where $r = \binom{n}{\ell-p/2} \asymp n^{\ell-\frac{p}{2}}$ as $n \to \infty$ (with p, ℓ fixed).

Proof. The first statement is [44, eq. (14)]. Next, by [44, Proposition A.1], the matrix $\frac{\mathbf{E}M}{\lambda}$ has $\ell+1$ *distinct* eigenvalues $\mu_0 > \cdots > \mu_\ell$ so that μ_m has multiplicity $\binom{n}{m} - \binom{n}{m-1}$. Moreover, it is shown in the proof of [44, Lemma A.3] that $|\mu_m| \le \frac{p}{n}\mu_0$ for $m > \ell - \frac{p}{2}$. It follows readily that $\mathrm{s}_{r+1} \le \frac{p}{n}\mathrm{s}_1$ for

$$r = \sum_{m=0}^{\ell-p/2}\left(\binom{n}{m} - \binom{n}{m-1}\right) = \binom{n}{\ell-p/2},$$

concluding the proof. □

We can now complete the proof of Theorem 3.7.

Proof of Theorem 3.7. Let $d = \binom{n}{\ell} \asymp n^\ell$ and $r = \binom{n}{\ell-p/2} \asymp n^{\ell-\frac{p}{2}}$. By Lemma 8.3, we can decompose $\mathbf{E}M = B + (\mathbf{E}M - B)$ so that B has rank r, $\lambda_{\max}(B) = \lambda k_*$, and $\|\mathbf{E}M - B\| \le \frac{\lambda p}{n} k_*$. Now define the random matrix

$$X := k_*^{-\frac{1}{2}}(M - \mathbf{E}M + B).$$

Then $\mathbf{E}[(X - \mathbf{E}X)^2] = \mathbf{1}$ and $v(X) \lesssim n^{\frac{\ell-p}{2}}$ by Lemma 8.2. We obtain

$$\mathbf{P}\big[|\lambda_{\max}(X) - \mathrm{B}(\lambda k_*^{\frac{1}{2}})| > Cn^{\frac{4\ell-3p}{4}} + Cn^{\frac{\ell-p}{4}}(\log d)^{\frac{3}{4}}\big] \le \frac{1}{d^\beta}$$

by applying Corollary 2.3 and Theorem 2.7 with $t \leftarrow \sqrt{\beta \log d}$, where we used that $\sigma_*(X) \le v(X)$ and C depends on β. Therefore

$$\mathbf{P}\big[|\lambda_{\max}(k_*^{-\frac{1}{2}}M) - \mathrm{B}(\lambda k_*^{\frac{1}{2}})| > Cpn^{-1}\lambda k_*^{\frac{1}{2}} + Cn^{\frac{4\ell-3p}{4}} + Cn^{\frac{\ell-p}{4}}(\log d)^{\frac{3}{4}}\big] \le \frac{1}{d^\beta}.$$

Finally, note that as p, ℓ are integers, $\ell < \frac{3p}{4}$ implies that $4\ell \le 3p - 1$. Thus

$$n^{\frac{4\ell-3p}{4}} \le n^{-\frac{1}{4}}, \qquad n^{\frac{\ell-p}{4}} \le n^{-\frac{p}{16}} \le n^{-\frac{1}{4}}.$$

Since $\mathrm{B}(\lambda k_*^{\frac{1}{2}}) = 2$ when $\lambda \le k_*^{-\frac{1}{2}}$ and $\mathrm{B}(\lambda k_*^{\frac{1}{2}}) \ge \mathrm{B}(1+\varepsilon) > 2$ when $\lambda \ge (1+\varepsilon)k_*^{-\frac{1}{2}}$, the final part of the theorem follows readily. □

Remark 8.4. It is readily read off from the proof that the final part of the statement of Theorem 3.7 can be impoved in various ways: the exponent of $n^{-\frac{1}{5}}$ in the definition of the test can be improved in a manner depending on the choice of ℓ, p, while the conclusion remains valid when $\varepsilon = n^{-\alpha}$ for a suitable choice of $\alpha > 0$. Since the precise exponents provided by the proof are not expected to be optimal, we have stated a slightly weaker result for simplicity of exposition.

8.3. **Spike detection in block-structured models.**

Proof of Theorem 3.10. Note that $X, X_\varnothing$ are precisely of the form (2.4) with $\mathbf{B} = \frac{1}{\mathbf{\Delta}}$ and $z = x$. Moreover, that $\sigma(X)^2 \le \frac{2}{d}\|\mathbf{B}1_d\|_\infty \le 2\beta$ and $\sigma_*(X)^2 \le v(X)^2 \le \frac{4\beta}{d}$ follows from a straightforward computation. Thus

$$\mathbf{P}\big[|\lambda_{\max}(X) - \lambda_{\max}(X_{\rm free})| > C\beta^{\frac12} d^{-\frac14} (\log d)^{\frac34}\big] \le e^{-d^{\frac12}}$$

by applying Corollary 2.3 with $t \leftarrow d^{\frac14}$, and the analogous bound holds for $X_\varnothing$. Now note that for $\lambda, \lambda_\varnothing$ as in Theorem 2.12, we have

$$|\lambda_{\max}(X_{\rm free}) - \lambda| \le \sqrt{\frac{8q\beta}{d}}, \qquad |\lambda_{\max}(X_{\varnothing,\rm free}) - \lambda_\varnothing| \le \sqrt{\frac{8q\beta}{d}}.$$

The conclusion follows from Theorem 2.12, where we set $\mu = \lambda_\varnothing$. □

Remark 8.5. The assumption that $x \in \{-1,+1\}^d$ was used in the proof only in order to ensure the assumption of section 2.4 that $\sum_{i\in C_k} z_i^2 = |C_k|$ for each k.

Let us consider instead the case that x is a random vector with i.i.d. entries such that $\mathbf{E}[x_i^2] = 1$, as is assumed in [32]. Then the above condition does not hold exactly for $z \leftarrow x$, but it holds approximately by the law of large numbers. In particular, in this case we may choose $z \in \mathbb{R}^d$ to be defined by

$$z_i = \frac{x_i}{\big(\frac{1}{|C_k|}\sum_{j\in C_k} x_j^2\big)^{\frac12}} \quad \text{for all } i \in C_k,\ k \in [q].$$

Then z satisfies the assumption of section 2.4 by construction, while the law of large numbers ensures that we have

$$\left\|\frac{1}{d}\operatorname{diag}(z)\mathbf{B}\operatorname{diag}(z) - \frac{1}{d}\operatorname{diag}(x)\mathbf{B}\operatorname{diag}(x)\right\| = o(1)$$

with probability $1-o(1)$ as long as $\min_k |C_k| \to \infty$ and q, β do not grow too rapidly. We can therefore replace x by z in the analysis up to a negligible error, and the remainder of the analysis proceeds verbatim as in the proof of Theorem 3.10.

The above argument is readily implemented without any further assumption in the asymptotic setting where q, Δ, c and the distribution of x_i are fixed as $d \to \infty$ by using the classical law of large numbers. However, implementing this procedure in a nonasymptotic setting would require us to quantify the error in the law of large numbers. This is readily accomplished in many situations, but the details of the bounds depend on the precise assumptions that are made on the distribution of x_i (for example, if they are subgaussian, we may use Bernstein's inequality to obtain nonasymptotic bounds.) As this part of the argument is completely independent of the random matrix analysis, we do not pursue it further here.

We now turn to the proof of Theorem 3.13.

Proof of Theorem 3.13. Let $B = \frac{1}{\Delta}$. We aim to apply Lemma 6.4 with $P = \frac{1}{d}xx^*$. Let us therefore define $X_t := X + \frac{t}{d}xx^*$, so that in the notation of Lemma 7.4

$$\mathbf{E}X_t = \mathrm{A}\big(\operatorname{diag}(c)^{\frac12}(B + t1_q1_q^*)\operatorname{diag}(c)^{\frac12} - \operatorname{diag}(Bc), -Bc\big).$$

Note that the precise form of $\mathbf{E}X$ played no role in the proof of Proposition 7.6, so that it transfers verbatim to the present setting. In particular, if we define

$$\lambda_t := \inf_{v>0} \max\big\{\lambda_{\max}\big(\operatorname{diag}(c)^{\frac12}(B+t1_q1_q^*)\operatorname{diag}(c)^{\frac12} + \operatorname{diag}\big(B\operatorname{diag}(c)(v-1_q)\big)\big),$$
$$\lambda_{\max}\big(\operatorname{diag}(v)^{-1} + \operatorname{diag}\big(B\operatorname{diag}(c)(v-1_q)\big)\big)\big\},$$

then the proofs of Proposition 7.6 and Theorem 3.10 readily yield

$$\mathbf{P}\Big[|\lambda_{\max}(X_t)-\lambda_t| > C\beta^{\frac12}\Big(\tfrac{(\log d)^{\frac34}}{d^{\frac14}} + \tfrac{q^{\frac12}}{d^{\frac12}}\Big)\Big] \le e^{-d^{\frac12}}$$

for every $t\in\mathbb{R}$. We therefore obtain for any $t>0$

$$\frac{\lambda_0-\lambda_{-t}}{t} - o(1) \le \frac1d|\langle x, v_{\max}(X)\rangle|^2 \le \frac{\lambda_t-\lambda_0}{t} + o(1)$$

with probability $1-o(1)$ as $d\to\infty$ using Lemma 6.4.

The major simplification of the asymptotic setting where q, B, c are fixed as $d\to\infty$ is that the definition of λ_t is then independent of d. To conclude the proof it therefore suffices to gain a qualitative, rather than quantitative, understanding of the behavior of $\lambda_t-\lambda_0$. In the remainder of the proof, we will consider the three cases $\mathrm{SNR}(\Delta)<1$, $\mathrm{SNR}(\Delta)>1$, and $\mathrm{SNR}(\Delta)=1$ separately.

Case 1. Suppose that $\mathrm{SNR}(\Delta)<1$, so that $\lambda_0 =: \lambda = \lambda_\varnothing$ by Theorem 2.12. Denote by v^* the minimizer in the definition of λ_0. Then it can be read off from the proof of Lemma 7.11 and from Lemma 7.13 that

$$\lambda_{\max}\big(\operatorname{diag}(c)^{\frac12}B\operatorname{diag}(c)^{\frac12} + \operatorname{diag}\big(B\operatorname{diag}(c)(v^*-1_q)\big)\big) < \lambda_0,$$
$$\lambda_{\max}\big(\operatorname{diag}(v^*)^{-1} + \operatorname{diag}\big(B\operatorname{diag}(c)(v^*-1_q)\big)\big) = \lambda_0.$$

Thus choosing $v\leftarrow v^*$ in the definition of λ_t shows that $\lambda_t\le\lambda_0$ when $t>0$ is sufficiently small. The conclusion follows immediately.

Case 2. Suppose that $\mathrm{SNR}(\Delta)>1$, so that $\lambda_0 =: \lambda > \lambda_\varnothing$ by Theorem 2.12. Denote by v^* the minimizer in the definition of λ_0. Then it follows from step 2 in the proof of Theorem 2.12 that we must have

$$\lambda_{\max}\big(\operatorname{diag}(c)^{\frac12}B\operatorname{diag}(c)^{\frac12} + \operatorname{diag}\big(B\operatorname{diag}(c)(v^*-1_q)\big)\big) = \lambda_0, \tag{8.1}$$
$$\lambda_{\max}\big(\operatorname{diag}(v^*)^{-1} + \operatorname{diag}\big(B\operatorname{diag}(c)(v^*-1_q)\big)\big) = \lambda_0. \tag{8.2}$$

Moreover, v^* is not a minimizer of the left-hand side of (8.2), as that would contradict $\lambda>\lambda_\varnothing$. Now note that the largest eigenvalue in (8.1) is simple and the associated eigenvector $w>0$ has strictly positive entries by the Perron-Frobenius theorem. Thus w is not orthogonal to $\operatorname{diag}(c)^{\frac12}1_q$, so we must have[6]

$$\lambda_{\max}\big(\operatorname{diag}(c)^{\frac12}(B-t1_q1_q^*)\operatorname{diag}(c)^{\frac12} + \operatorname{diag}\big(B\operatorname{diag}(c)(v^*-1_q)\big)\big) < \lambda_0 \tag{8.3}$$

for every $t>0$. But as v^* is not a minimizer of (8.2), we can slightly perturb v_* to decrease the latter while preserving the strict inequality in (8.3). This shows that $\lambda_{-t}<\lambda_0$ for every $t>0$. The conclusion follows immediately.

Case 3. Suppose that $\mathrm{SNR}(\Delta)=1$. Then $\lambda_0 =: \lambda = 1$ by Theorem 2.12, and $v^*=1_q$ is a minimizer in the definition of λ_0.

[6]Let M be a self-adjoint matrix whose top eigenvalue is simple with eigenvector w, and let x be a vector not orthogonal to w. Then we have $\frac{d}{dt}\lambda_{\max}(M-txx^*)|_{t=0} = -|\langle x,w\rangle|^2 < 0$.

Denote by $b > 0$ be the Perron-Frobenius eigenvector of $B\,\mathrm{diag}(c)$, and choose $s > 0$ sufficiently large that $\lambda_{\max}(\mathrm{diag}(c)^{\frac12} 1_q 1_q^* \mathrm{diag}(c)^{\frac12} - s\,\mathrm{diag}(b)) \le 0$. Then choosing $v \leftarrow 1_q - tsb$ in the definition of λ_t readily yields

$$\lambda_t - \lambda_0 \le \max\Big\{0, \max_i \Big\{\frac{1}{1-tsb_i} - 1 - tsb_i\Big\}\Big\}$$

for all sufficiently small $t > 0$, where we used that $\lambda_{\max}(\mathrm{diag}(c)^{\frac12} B\,\mathrm{diag}(c)^{\frac12}) =: \mathrm{SNR}(\Delta) = 1$ implies that the Perron-Frobenius eigenvalue of $B\,\mathrm{diag}(c)$ is 1.

To conclude, note that $\frac{1}{1-x} - 1 - x = \frac{x^2}{1-x}$, so we have shown that $\lambda_t - \lambda_0 \le O(t^2)$ for $t > 0$ sufficiently small. The conclusion follows immediately. □

8.4. **Contextual stochastic block models.**

Proof of Theorem 3.14. Let $d = n + p$, and partition $[d] = C_1 \sqcup C_2$ into $C_1 = \{1, \ldots, n\}$ and $C_2 = \{n+1, \ldots, n+p\}$. Define $B \in \mathrm{M}_2(\mathbb{R})_{\mathrm{sa}}$ and $\hat z \in \mathbb{R}^d$ as

$$B = \frac{d}{n}\begin{bmatrix}\lambda^2 & \mu \\ \mu & 0\end{bmatrix}, \qquad \hat z = \begin{bmatrix} v \\ u\sqrt{p}\end{bmatrix}.$$

Then the random matrix $\hat X$ is of the form (2.4) with $z \leftarrow \hat z$.

Next, define the random matrix X as in (2.4) with

$$z = \begin{bmatrix} v \\ \frac{u\sqrt{p}}{\|u\|}\end{bmatrix}.$$

The random matrix X satisfies all the assumptions of section 2.4. On the other hand, as $u \sim N(0, \frac1p \mathbf{1}_p)$, we have $\|u\| = 1 + o(1)$ with probability $1 - o(1)$ by the law of large numbers. Using $\|\mathrm{diag}(x) M\,\mathrm{diag}(y)\| \le \max_{ij} |M_{ij}|\,\|x\|\,\|y\|$, we obtain

$$\|X - \hat X\| \le \frac{1}{d} \max_{i,j} |B_{ij}|\,(\|z\| + \|\hat z\|)\,\|z - \hat z\| = o(1)$$

with probability $1 - o(1)$, where we used that $\max_{i,j} |B_{ij}| \to (1 + \frac1\gamma)\max\{\lambda^2, \mu\}$.

To reason about $\hat v$, we would like to apply Lemma 6.4 to $\hat X_t = \hat X + \frac{t}{n} \mathbf{I}_{C_1} z z^* \mathbf{I}_{C_1}^*$. Approximating $\hat X$ by X and reasoning as in the proof of Theorem 3.13, we have

$$\lambda_{\max}(\hat X_t) = \lambda_t + o(1) \text{ with probability } 1 - o(1),$$

where λ_t is defined by

$$\lambda_t := \inf_{v>0} \max\big\{\lambda_{\max}\big(te_1e_1^* + \mathrm{diag}(c)^{\frac12} \bar B\,\mathrm{diag}(c)^{\frac12} + \mathrm{diag}\left(\bar B\,\mathrm{diag}(c)(v - 1_2)\right)\big),$$
$$\lambda_{\max}\big(\mathrm{diag}(v)^{-1} + \mathrm{diag}\left(\bar B\,\mathrm{diag}(c)(v - 1_2)\right)\big)\big\}$$

where we define

$$\bar B = (1 + \tfrac1\gamma)\begin{bmatrix}\lambda^2 & \mu \\ \mu & 0\end{bmatrix}, \qquad c = \begin{bmatrix}\frac{\gamma}{1+\gamma} \\ \frac{1}{1+\gamma}\end{bmatrix}.$$

We can now follow the remainder of the proof of Theorem 3.13 verbatim to show that there is asymptotically positive overlap between v and $\hat v$ if and only if

$$\frac12\Big(\lambda^2 + \sqrt{\lambda^4 + \tfrac{4\mu^2}{\gamma}}\Big) = \lambda_{\max}\big(\mathrm{diag}(c)^{\frac12} \bar B\,\mathrm{diag}(c)^{\frac12}\big) > 1.$$

Now note that

$$\frac12\Big(\lambda^2 + \sqrt{\lambda^4 + \tfrac{4\mu^2}{\gamma}}\Big) = 1 \quad \text{if and only if} \quad \lambda^2 + \tfrac{\mu^2}{\gamma} = 1$$

and both $\frac{1}{2}\big(\lambda^2+\sqrt{\lambda^4+\frac{4\mu^2}{\gamma}}\big)$ and $\lambda^2+\frac{\mu^2}{\gamma}$ are monotone in λ^2 and μ^2, so

$$\frac{1}{2}\Big(\lambda^2+\sqrt{\lambda^4+\tfrac{4\mu^2}{\gamma}}\Big)>1 \quad \text{if and only if} \quad \lambda^2+\tfrac{\mu^2}{\gamma}>1.$$

This concludes the proof. □

8.5. **Sample covariance error.** In the following, we define $\hat\Sigma=\frac{1}{n}XX^*$ and Σ as in (3.1)—(3.2) and let $\delta=\frac{p}{n}$, $d=\max\{n,p\}$. We begin by applying Theorem 2.5.

Lemma 8.6. *For $n\ge(\log d)^3$, we have*

$$\mathbf{P}\big[\big|\|\hat\Sigma\|-\|\tfrac{1}{n}X_{\rm free}X_{\rm free}^*\|\big|>C(1+\lambda+\delta)\,n^{-\frac14}(\log d)^{\frac34}\big]\le e^{-Cn^{\frac12}},$$

$$\mathbf{P}\big[\mathrm{d_H}\big(\mathrm{sp}(\hat\Sigma-\Sigma),\mathrm{sp}(\tfrac{1}{n}X_{\rm free}X_{\rm free}^*-\Sigma\otimes\mathbf{1})\big)>C(1+\lambda+\delta)\,n^{-\frac14}(\log d)^{\frac34}\big]\le e^{-Cn^{\frac12}}.$$

Proof. We readily compute

$$\sigma_*(X)^2=v(X)^2=\|\Sigma\|=1+\lambda$$

and

$$\sigma(X)^2=n\max\{1+\lambda,\delta+\tfrac{\lambda}{n}\}\le 2n(1+\lambda+\delta).$$

Moreover, note that $\|X_{\rm free}\|\le 2\sigma(X)$ by [36, p. 208]. The conclusion now follows readily by applying Theorem 2.5 with $X\leftarrow n^{-\frac12}X$, $t\leftarrow 2(1+\lambda+\delta)^{1/2}n^{-\frac14}(\log d)^{\frac34}$, and either $B\leftarrow 0$ or $B\leftarrow-\Sigma$, respectively. (Note that while Theorem 2.5 is formulated for $d\times d$ matrices X, it is applicable here as we can always add enough zero rows or columns to X to make it $d\times d$ without changing the relevant norms.) □

We must now estimate the spectra of the free operators that appear in the above lemma. In principle, this can be achieved using methods that are sketched in the work of Lehner [27, §5], which requires some lengthy computations. These computations are considerably simplified by a quadratic counterpart of Lehner's formula obtained in [33]. For our present purposes, we may work with the following result that arises as a special case of the general formulas in [33].

Lemma 8.7. *Let X be the $p\times n$ random matrix whose columns are i.i.d. $N(0,\Sigma)$, and denote the eigenvalues of Σ as $\mu_1\ge\cdots\ge\mu_p\ge0$. Then*

$$\|\tfrac{1}{n}X_{\rm free}X_{\rm free}^*\|=\inf_{0<a<1}\inf_{x\in\Delta_p}\max_{i\in[p]}\left\{\frac{\mu_i}{nax_i}+\frac{\mu_i}{1-a}\right\},$$

$$\lambda_{\max}(\tfrac{1}{n}X_{\rm free}X_{\rm free}^*-\Sigma\otimes\mathbf{1})=\inf_{0<a<1}\inf_{x\in\Delta_p}\max_{i\in[p]}\left\{\frac{\mu_i}{nax_i}+\frac{a\mu_i}{1-a}\right\},$$

$$-\lambda_{\min}(\tfrac{1}{n}X_{\rm free}X_{\rm free}^*-\Sigma\otimes\mathbf{1})=\inf_{a>0}\inf_{x\in\Delta_p}\max_{i\in[p]}\left\{\frac{\mu_i}{nax_i}+\frac{a\mu_i}{1+a}\right\},$$

where $\Delta_p:=\{x\in\mathbb{R}^p:x>0,\ \sum_i x_i=1\}$.

Proof. Since the quantities to be computed are independent of the choice of basis in $\mathbb{R}^p$, we may assume without loss of generality that Σ is a diagonal matrix, so that X has independent entries $X_{ij}\sim N(0,\mu_i)$. We can also assume that $\mu_p>0$, since otherwise all quantities in the statement remain unchanged if we remove the zero rows. With these simplifications, [33, Corollary 1.5] yields

$$\|\tfrac{1}{n}X_{\rm free}X_{\rm free}^*\|=\inf_{\substack{v>0\\ \frac1n\sum_k\mu_kv_k<1}}\max_{i\in[p]}\left\{\frac{1}{v_i}+\frac{\mu_i}{1-\frac1n\sum_k\mu_kv_k}\right\},$$

and making the change of variables $x_i = \frac{\mu_i v_i}{\sum_k \mu_k v_k}$, $a = \frac{1}{n}\sum_k \mu_k v_k$ yields the first equation in the statement. The other two equations follow analogously. □

In our setting, we have $\mu_1 = 1+\lambda$ and $\mu_2 = \cdots = \mu_p = 1$. We claim that it then suffices to minimize in the above variational principles only over vectors x such that $x_2 = \cdots = x_p$. That the latter yields an upper bound is obvious (as we are restricting the infimum to a smaller set). For the lower bound, we may use that

$$\max_{2\le i\le n} \frac{1}{x_i} \ge \frac{1}{n-1}\sum_{i=2}^n \frac{1}{x_i} \ge \frac{1}{\frac{1}{n-1}\sum_{i=2}^n x_i}$$

by convexity to argue that for any vector x, the function being optimized can only decrease if we replace all $x_2,\ldots,x_n$ by their average.

Lemma 8.8. *In the setting of Theorem 3.16, we have*

$$\big|\|\tfrac{1}{n}X_{\rm free}X_{\rm free}^*\| - \mathrm{S}(\lambda,\delta)\big| \le C(1+\lambda+\delta)\,n^{-\frac{1}{2}},$$
$$\big|\lambda_{\max}(\tfrac{1}{n}X_{\rm free}X_{\rm free}^* - \Sigma\otimes\mathbf{1}) - \mathrm{H}_+(\lambda,\delta)\big| \le C(1+\lambda+\delta)\,n^{-\frac{1}{2}}.$$

Proof. Restricting the infimum over x in Lemma 8.7 to $x = (b, \frac{1-b}{p-1}, \cdots, \frac{1-b}{p-1})$ yields

$$\|\tfrac{1}{n}X_{\rm free}X_{\rm free}^*\| = \inf_{0<a,b<1} \max\left\{\frac{1+\lambda}{nab} + \frac{1+\lambda}{1-a}, \frac{p-1}{na(1-b)} + \frac{1}{1-a}\right\},$$
$$\lambda_{\max}(\tfrac{1}{n}X_{\rm free}X_{\rm free}^* - \Sigma\otimes\mathbf{1}) = \inf_{0<a,b<1} \max\left\{\frac{1+\lambda}{nab} + \frac{(1+\lambda)a}{1-a}, \frac{p-1}{na(1-b)} + \frac{a}{1-a}\right\}$$

by the above observation. We can rewrite the first line as

$$\begin{aligned}\|\tfrac{1}{n}X_{\rm free}X_{\rm free}^*\| &= \inf_{x\in\Delta_3}\sup_{0<\pi<1}\left\{\pi\left(\frac{1+\lambda}{nx_1}+\frac{1+\lambda}{x_3}\right)+(1-\pi)\left(\frac{p-1}{nx_2}+\frac{1}{x_3}\right)\right\}\\ &= \sup_{0<\pi<1}\left(\sqrt{\frac{\pi(1+\lambda)}{n}}+\sqrt{\frac{(1-\pi)(p-1)}{n}}+\sqrt{1+\pi\lambda}\right)^2,\end{aligned}$$

where we used the Sion minimax theorem and $\inf_{x\in\Delta_r}\sum_{i=1}^r \frac{a_i}{x_i} = \big(\sum_{i=1}^r \sqrt{a_i}\big)^2$. By exactly the same argument, the second line becomes

$$\lambda_{\max}(\tfrac{1}{n}X_{\rm free}X_{\rm free}^* - \Sigma\otimes\mathbf{1}) = \sup_{0<\pi<1}\left\{\left(\sqrt{\frac{\pi(1+\lambda)}{n}}+\sqrt{\frac{(1-\pi)(p-1)}{n}}+\sqrt{1+\pi\lambda}\right)^2 - (1+\pi\lambda)\right\},$$

where we used that $\frac{a}{1-a} = \frac{1}{1-a} - 1$.

To conclude the proof, we note that

$$\mathrm{S}(\lambda,\delta) = \sup_{0<\pi<1}\big(\sqrt{(1-\pi)\delta}+\sqrt{1+\pi\lambda}\big)^2,$$
$$\mathrm{H}_+(\lambda,\delta) = \sup_{0<\pi<1}\big\{\big(\sqrt{(1-\pi)\delta}+\sqrt{1+\pi\lambda}\big)^2 - (1+\pi\lambda)\big\}.$$

The conclusion now follows readily. □

It remains to estimate the smallest eigenvalue of the centered case. The proof is similar to that of Lemma 8.8, but differs in the details of the computation.

Lemma 8.9. *In the setting of Theorem 3.16, we have*

$$\big|\lambda_{\min}(\tfrac{1}{n}X_{\rm free}X_{\rm free}^* - \Sigma\otimes\mathbf{1}) - \mathrm{H}_-(\lambda,\delta)\big| \le C(1+\lambda+\delta)\,n^{-\frac{1}{2}}.$$

Proof. As in Lemma 8.8, we may restrict the infimum over x in the last equation display of Lemma 8.7 to $x = (b, \frac{1-b}{p-1}, \cdots, \frac{1-b}{p-1})$. This yields

$$-\lambda_{\min}(\tfrac{1}{n} X_{\text{free}} X_{\text{free}}^* - \Sigma \otimes \mathbf{1}) = \sup_{0<\pi<1} \inf_{a>0} \inf_{0<b<1} \left\{ \pi \left(\frac{1+\lambda}{nab} + \frac{(1+\lambda)a}{1+a} \right) + (1-\pi) \left(\frac{p-1}{na(1-b)} + \frac{a}{1+a} \right) \right\},$$

where we used the Sion minimax theorem as in the proof of Lemma 8.8.

Now note that for $u, v > 0$, we have $\inf_{0<b<1} \left(\frac{u}{b} + \frac{v}{1-b} \right) = (\sqrt{u} + \sqrt{v})^2$ and

$$\varphi(u,v) := \inf_{a>0} \left\{ \frac{a}{1+a} u + \frac{1}{a} v \right\} = \begin{cases} 2\sqrt{uv} - v & \text{if } v < u, \\ u & \text{otherwise.} \end{cases}$$

We therefore obtain

$$-\lambda_{\min}(\tfrac{1}{n} X_{\text{free}} X_{\text{free}}^* - \Sigma \otimes \mathbf{1}) = \sup_{0<\pi<1} \varphi\left(1 + \pi\lambda, \left(\sqrt{\frac{\pi(1+\lambda)}{n}} + \sqrt{\frac{(1-\pi)(p-1)}{n}} \right)^2 \right).$$

Using the explicit formula for φ, we readily estimate

$$\left| \lambda_{\min}(\tfrac{1}{n} X_{\text{free}} X_{\text{free}}^* - \Sigma \otimes \mathbf{1}) + \sup_{0<\pi<1} \varphi(1+\pi\lambda, (1-\pi)\delta) \right| \le C(1+\lambda+\delta) n^{-\frac{1}{2}}.$$

Now note that $(1-\pi)\delta < 1 + \pi\lambda$ holds if and only if $\pi > \frac{\delta-1}{\delta+\lambda}$. Therefore

$$\sup_{0<\pi<1} \varphi(1+\pi\lambda, (1-\pi)\delta) = \sup_{\frac{\delta-1}{\delta+\lambda} \le \pi < 1} \left\{ 2\sqrt{(1+\pi\lambda)(1-\pi)\delta} - (1-\pi)\delta \right\},$$

where we used that $\varphi(1+\pi\lambda, (1-\pi)\delta)$ is increasing for $0 < \pi < \frac{\delta-1}{\delta+\lambda}$. The supremum on the right-hand side equals $-\mathrm{H}_-(\lambda, \delta)$, concluding the proof. □

Combining Lemmas 8.6, 8.8, and 8.9 completes the proof of Theorem 3.16.

Acknowledgments. The authors thank Raphael Barboni, Charles Bordenave, Pravesh Kothari, Florent Krzakala, and Lenka Zdeborová for useful discussions, and an anonymous referee for several helpful suggestions. GC was supported in part by the MUR Excellence Department Project MatMod@TOV awarded to the Department of Mathematics, University of Rome Tor Vergata, CUP E83C18000100006. RvH was supported in part by NSF grants DMS-2054565 and DMS-2347954.

References

[1] E. Abbe, A. S. Bandeira, A. Bracher, and A. Singer. Decoding binary node labels from censored edge measurements: phase transition and efficient recovery. *IEEE Trans. Network Sci. Eng.*, 1(1):10–22, 2014.

[2] J. Alt, L. Erdős, and T. Krüger. The Dyson equation with linear self-energy: spectral bands, edges and cusps. *Doc. Math.*, 25:1421–1539, 2020.

[3] B. Au. BBP phenomena for deformed random band matrices, 2023. Preprint arxiv:2304.13047.

[4] J. Baik, G. Ben Arous, and S. Péché. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab.*, 33(5):1643–1697, 2005.

[5] J. Baik and J. W. Silverstein. Eigenvalues of large sample covariance matrices of spiked population models. *J. Multivariate Anal.*, 97(6):1382–1408, 2006.

[6] A. S. Bandeira, M. T. Boedihardjo, and R. van Handel. Matrix concentration inequalities and free probability. *Invent. Math.*, 234(1):419–487, 2023.

[7] A. S. Bandeira, G. Cipolloni, D. Schröder, and R. van Handel, March 2023. Unpublished.

[8] F. Benaych-Georges, C. Bordenave, and A. Knowles. Largest eigenvalues of sparse inhomogeneous Erdős-Rényi graphs. *Ann. Probab.*, 47(3):1653–1676, 2019.
[9] R. Bhatia. *Matrix analysis*, volume 169 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1997.
[10] P. Biane. Free hypercontractivity. *Comm. Math. Phys.*, 184(2):457–474, 1997.
[11] J. Bigot and C. Male. Freeness over the diagonal and outliers detection in deformed random matrices with a variance profile. *Inf. Inference*, 10(3):863–919, 2021.
[12] C. Bordenave and B. Collins. Norm of matrix-valued polynomials in random unitaries and permutations, 2023. Preprint arxiv:2304.05714v2.
[13] S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities*. Oxford University Press, Oxford, 2013. A nonasymptotic theory of independence, With a foreword by Michel Ledoux.
[14] T. Brailovskaya and R. van Handel. Universality and sharp matrix concentration inequalities. *Geom. Funct. Anal.*, 34(6):1734–1838, 2024.
[15] A. Buchholz. Operator Khintchine inequality in non-commutative probability. *Math. Ann.*, 319(1):1–16, 2001.
[16] M. Capitaine and C. Donati-Martin. Spectrum of deformed random matrices and free probability. In *Advanced topics in random matrices*, volume 53 of *Panor. Synthèses*, pages 151–190. Soc. Math. France, Paris, 2017.
[17] E. Carlen. Trace inequalities and quantum entropy: an introductory course. In *Entropy and the quantum*, volume 529 of *Contemp. Math.*, pages 73–140. Amer. Math. Soc., Providence, RI, 2010.
[18] M. Cucuringu. Synchronization over Z_2 and community detection in signed multiplex networks with constraints. *J. Complex Netw.*, 3(3):469–506, 2015.
[19] Y. Deshpande, S. Sen, A. Montanari, and E. Mossel. Contextual stochastic block models. In *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018.
[20] A. Guionnet, J. Ko, F. Krzakala, and L. Zdeborová. Low-rank matrix estimation with inhomogeneous noise. *Inf. Inference*, 14(2):Paper No. iaaf010, 80, 2025.
[21] U. Haagerup and S. Thorbjørnsen. A new application of random matrices: $\mathrm{Ext}(C^*_{\mathrm{red}}(F_2))$ is not a group. *Ann. of Math. (2)*, 162(2):711–775, 2005.
[22] Q. Han. Exact bounds for some quadratic empirical processes with applications, 2024. Preprint arxiv:2207.13594v3.
[23] S. B. Hopkins, J. Shi, and D. Steurer. Tensor principal component analysis via sum-of-square proofs. In *Conference on Learning Theory*, pages 956–1006. JMLR, 2015.
[24] J.-T. Hsieh, P. K. Kothari, and S. Mohanty. A simple and sharper proof of the hypergraph Moore bound. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2324–2344, 2023.
[25] D. Hundertmark and B. Simon. An optimal L^p-bound on the Krein spectral shift function. *J. Anal. Math.*, 87:199–208, 2002. Dedicated to the memory of Thomas H. Wolff.
[26] S. Lee and J. O. Lee. Phase transition for the generalized two-community stochastic block model. *J. Appl. Probab.*, 61(2):385–400, 2024.
[27] F. Lehner. Computing norms of free operators with matrix coefficients. *Amer. J. Math.*, 121(3):453–486, 1999.
[28] P. Mergny, J. Ko, and F. Krzakala. Spectral phase transition and optimal PCA in block-structured spiked models. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pages 35470–35491. PMLR, 21–27 Jul 2024.
[29] A. Montanari and E. Richard. A statistical model for tensor PCA. In *Proceedings of the 27th International Conference on Neural Information Processing Systems*, NIPS'14, pages 2897–2905, Cambridge, MA, USA, 2014. MIT Press.
[30] A. Nica and R. Speicher. *Lectures on the combinatorics of free probability*, volume 335 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2006.
[31] A. Nilli. On the second eigenvalue of a graph. *Discrete Math.*, 91(2):207–210, 1991.
[32] A. Pak, J. Ko, and F. Krzakala. Optimal algorithms for the inhomogeneous spiked wigner model. In *Advances in Neural Information Processing Systems*, volume 36, pages 76409–76424. Curran Associates, Inc., 2023.

[33] E. Parmaksiz and R. Van Handel. Computing extreme singular values of free operators, 2025. Preprint arxiv:2510.23987v1.
[34] A. Perry, A. S. Wein, A. S. Bandeira, and A. Moitra. Optimality and sub-optimality of PCA for spiked random matrices and synchronization, 2016. Preprint arxiv:1609.05573v2.
[35] D. Petz. Jensen's inequality for positive contractions on operator algebras. *Proc. Amer. Math. Soc.*, 99(2):273–277, 1987.
[36] G. Pisier. *Introduction to operator space theory*, volume 294 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2003.
[37] T. J. Rivlin. *An introduction to the approximation of functions*. Blaisdell Publishing Co. [Ginn and Co.], Waltham, Mass.-Toronto, Ont.-London, 1969.
[38] A. Saade, M. Lelarge, F. Krzakala, and L. Zdeborová. Spectral detection in the censored block model. In *2015 IEEE International Symposium on Information Theory (ISIT)*, pages 1184–1188, 2015.
[39] A. Singer. Angular synchronization by eigenvectors and semidefinite programming. *Appl. Comput. Harmon. Anal.*, 30(1):20–36, 2011.
[40] L. Stephan and L. Massoulié. Non-backtracking spectra of weighted inhomogeneous random graphs. *Math. Stat. Learn.*, 5(3-4):201–271, 2022.
[41] K. Tikhomirov and P. Youssef. The spectral gap of dense random regular graphs. *Ann. Probab.*, 47(1):362–419, 2019.
[42] J. A. Tropp. An introduction to matrix concentration inequalities. *Foundations and Trends in Machine Learning*, 8:1–230, 2015.
[43] J. A. Tropp. Second-order matrix concentration inequalities. *Appl. Comput. Harmon. Anal.*, 44(3):700–736, 2018.
[44] A. S. Wein, A. El Alaoui, and C. Moore. The Kikuchi hierarchy and tensor PCA. In *2019 IEEE 60th Annual Symposium on Foundations of Computer Science*, pages 1446–1468. IEEE Comput. Soc. Press, Los Alamitos, CA, 2019.

Department of Mathematics, ETH Zürich, Switzerland
Email address: bandeira@math.ethz.ch

Department of Mathematics, University of Rome Tor Vergata, Via della Ricerca Scientifica 1, 00133 Roma RM, Italy
Email address: cipolloni@axp.mat.uniroma2.it

Department of Mathematics, ETH Zürich, Switzerland
Email address: schroeder.dominik@gmail.com

Department of Mathematics, Princeton University, Princeton, NJ 08544, USA
Email address: rvan@math.princeton.edu