Spectral Phase Transition and Optimal PCA in Block-Structured Spiked Models

Pierre Mergny [1] **Justin Ko** [2 3] **Florent Krzakala** [1]

Abstract

We discuss the inhomogeneous spiked Wigner model, a theoretical framework recently introduced to study structured noise in various learning scenarios, through the prism of random matrix theory, with a specific focus on its spectral properties. Our primary objective is to find an optimal spectral method and to extend the celebrated (Baik et al., 2005) (BBP) phase transition criterion — well-known in the homogeneous case— to our inhomogeneous, block-structured, Wigner model. We provide a thorough rigorous analysis of a transformed matrix and show that the transition for the appearance of 1) an outlier outside the bulk of the limiting spectral distribution and 2) a positive overlap between the associated eigenvector and the signal, occurs precisely at the optimal threshold, making the proposed spectral method optimal within the class of iterative methods for the inhomogeneous Wigner problem.

1. Introduction

The statistical challenge of inferring a low-dimensional signal from a noisy, high-dimensional observation is ubiquitous across statistics, probability, and machine learning. Spiked random matrix models have recently gained extensive interest, serving as a valuable platform for exploring this issue (Donoho & Johnstone, 1995; Péché, 2014; Lesieur et al., 2017). A prominent example is the spiked Wigner model, where a rank one matrix is observed through a componentwise homogeneous noise, that has been studied extensively in random matrix theory (Baik et al., 2005).

Most models, with the *spiked Wigner model* at the forefront, have focused however on scenarios where the noise is "homogeneous", aiming to understand how the performance of the inference depends on the noise level. Yet in practice, datasets are inherently structured and the exploration of inhomogeneity plays a pivotal role in unraveling their complexities. A prototypical model to study this phenomenon is to improve the aforementioned spiked Wigner model by introducing a block structure in the noise, a model which has been recently introduced in a series of papers (Behne & Reeves, 2022; Alberici et al., 2021a; 2022; Guionnet et al., 2022) and that arises in many different learning contexts such as community detection (Behne & Reeves, 2022; Guionnet et al., 2022), deep Boltzmann machines (Alberici et al., 2021b), or the dense limit of the celebrated degree-corrected stochastic block model (Guionnet et al., 2022; Karrer & Newman, 2011).

Our goal in this paper is to apply rigorous random matrix theory to such "inhomogenous" spiked models, and to provide an optimal reconstruction method from a spectral algorithm, to generalize the seminal work of (Baik et al., 2005) (BBP) to inhomogenous matrices.

Settings and open questions — The model is defined in practice by multiplying the Wigner matrix by a variance-profile matrix, namely one would like to infer in the high-dimensional setting $N \gg 1$, the underlying signal $\mathbf{x} \in \mathbb{R}^N$ from the noisy matrix observation:

$$\boldsymbol{Y} = \sqrt{\frac{1}{N}}\mathbf{x}\mathbf{x}^\top + \boldsymbol{H} \odot \left(\boldsymbol{\Delta}^{\odot 1/2}\right), \tag{1}$$

where $\boldsymbol{\Delta}$ is a block-constant matrix, $\odot$ denotes the Hadamard product $(\mathbf{A} \odot \mathbf{B})_{ij} = A_{ij}B_{ij}$, $\mathbf{A}^{\odot\alpha}$ is the Hadamard power ($(\mathbf{A}^{\odot\alpha})_{ij} = A_{ij}^{\alpha}$) and $\boldsymbol{H}$ is a real-valued symmetric matrix with independent Gaussian entries of unit variance above the diagonal. The precise assumptions on $\mathbf{x}$ and $\boldsymbol{\Delta}$ are postponed to the next section.

The study of the inhomogeneous model of Eq. (1) from a Bayes-optimal point of view has been performed in a series of work in the asymptotic limit $N \to \infty$ in (Guionnet et al., 2022; Behne & Reeves, 2022; Chen & Xia, 2022; Chen et al., 2021) who characterized the fundamental information-theoretic limit of reconstruction in this model. Recently (Pak et al., 2023) discussed algorithmic performances of the Approximate Message Passage (AMP in short) algorithm introduced in particular a simpler variant of this algorithm

[1]IdePHICS laboratory, École Fédérale Polytechnique de Lausanne, Switzerland [2]Department of Statistics and Actuarial Science, University of Waterloo, Canada [3]UMPA, ENS Lyon, France. Correspondence to: Pierre Mergny <mergny.pierre@gmail.com>.

Proceedings of the 41st International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024. Copyright 2024 by the author(s).

for this problem (linearized AMP, see also (Aubin et al., 2019; Mondelli & Montanari, 2018; Maillard et al., 2022)) and *conjectured* it to be optimal for detection. The linearized version of AMP can be shown to be equivalent to performing principal component analysis (PCA) on a linear transformation of the matrix $\boldsymbol{Y}$, namely estimate the signal $\mathbf{x}$ using the top eigenvector of the matrix:

$$\tilde{\boldsymbol{Y}} := \boldsymbol{Y} \odot \left(\boldsymbol{\Delta}^{\odot -1}\right) - \frac{1}{\sqrt{N}} \mathrm{Diag}\left(\boldsymbol{\Delta}^{\odot -1}\, \mathbf{1}\right) , \quad (2)$$

where $\mathbf{1} := (1, \dots, 1) \in \mathbb{R}^N$ and for $\mathbf{v} \in \mathbb{R}^N$, $\mathrm{Diag}(\mathbf{v})_{ij} := v_i \delta_{i,j}$ is the corresponding diagonal matrix made of the entries of its argument.

Due to the block structure of the noise in the random matrix $\tilde{\boldsymbol{Y}}$, tracking the (possible) outliers in its spectrum is a challenging problem and in (Pak et al., 2023) the associated phase transition for the existence of an outlier has only been conjectured to occur when a specific parameter (described below) reaches the value one. While it was hinted by (Pak et al., 2023) that the spectral method associated with the matrix (2) may be an efficient one for inhomogenous problems, it was left an open problem.

Our contributions — In this work, we step up to this challenge, and provide a rigorous analysis of the spectral method for the inhomogeneous spiked Wigner models of Eq. (2). More specifically,

1. We proved the conjecture of (Pak et al., 2023) and show that the spectral method for the matrix $\tilde{\boldsymbol{Y}}/\sqrt{N}$ defined by (2), has a phase transition at the predicted algorithmic threshold (Guionnet et al., 2022).
2. This makes PCA applied to (2) optimal in detection in the inhomogenous spiked model (at least in the absence of a statistical to computation gap, see (Guionnet et al., 2022)). Our results could also be used as spectral start (Mondelli & Venkataramanan, 2021) for AMP algorithms (Pak et al., 2023), making them optimal in terms of MMSE for these problems.
3. We obtained a complete characterization of the associated overlap with the hidden signal, and their phase transition, further generalizing the BBP results from homogeneous matrices to inhomogeneous ones.

In particular, we prove that this phase transition occurs when the top eigenvalue of a certain symmetric matrix $\boldsymbol{\Omega}_K$, containing all information of the model (proportion of each community, the noise level inside and between different communities), takes the value one. This critical value $\lambda_1(\boldsymbol{\Omega}_K) = 1$ corresponds to the phase transition of AMP for this model (Pak et al., 2023), and more generically to the algorithmic threshold of efficient inference of the problem (Guionnet et al., 2022) (see e.g. (Zdeborová & Krzakala, 2016; Bandeira et al., 2018; 2022) for discussion of the algorithmic to statistic gap in such high-dimensional problems). As a consequence, our result shows that PCA on the transformed matrix $\tilde{\boldsymbol{Y}}$ also achieves optimal efficient detection for the model (1).

To the authors' best knowledge, our work is the first to provide an exact asymptotic of the phase transition of heterogeneous spiked Wigner models, both for the top eigenvalue and the overlap. To obtain these results, our study relies on a detailed analysis of large block-structure random matrix models that go beyond the standard tools used in the spiked homogeneous models and that can be of independent interest for analyzing structured matrices in machine learning.

Other related works — The task of factorizing low-rank matrices to find a hidden signal in data, ranging from sparse PCA to community detection and sub-matrix localization. Many variants of the homogeneous problem have been studied in the high-dimensional limit, see e.g. (Deshpande et al., 2015; Lesieur et al., 2017; Barbier et al., 2018; Lesieur et al., 2017; Alaoui et al., 2020; Lelarge & Miolane, 2019; Barbier & Reeves, 2020). The inhomogeneous version was introduced and studied in (Behne & Reeves, 2022; Alberici et al., 2021a; 2022; Guionnet et al., 2022; Pak et al., 2023).

Spectral methods are a popular tool for solving rank-factorization problems (Donoho & Johnstone, 1995; Péché, 2014; Baik et al., 2005), and this has triggered a large amount of work in the random matrix theory (RMT) community. Spiked models with homogeneous noise are well understood and have been studied extensively in RMT, it is well known that as one decreases the noise level (that is the variance of the entries of the Wigner matrix), there is a phase transition in the behavior of the top eigenvalue which detaches from the semi-circle distribution, see for example (Péché, 2005; Féral & Péché, 2007; Capitaine et al., 2009; 2012; Pizzo et al., 2013; Renfrew & Soshnikov, 2013) and also (Baik & Silverstein, 2006; Paul, 2007; Bloemendal & Virág, 2013; Benaych-Georges & Nadakuditi, 2011; 2012; Guionnet et al., 2023) for other variants of these models.

The properties of the spectrum of variance-profile Wigner matrices (without any small-rank perturbation) are also well understood (Ajanki et al., 2017a;b; 2019; Alt et al., 2020) even though there is, in general, no closed-form solution to the limiting spectral distribution unlike the semi-circle distribution for Wigner matrices. To the best knowledge of the authors, only the two works (Bigot & Male, 2020; Lee & Lee, 2023) performed a spectral analysis of spiked block-structure Wigner models but in some specific regimes. (Bigot & Male, 2020) studied the spectrum of a (generalization of) models of the form given directly by Eq. (1)

(instead of the transform $\tilde{\boldsymbol{Y}}$ of Eq. (2) of this paper) using free probability tools but without studying the overlap or having a closed expression for the outlier, as it is done in this paper. (Lee & Lee, 2023) looked at the behavior of the top eigenvalue in a (restricted version) of (2×2) blocks and our result provides the extension to the $(K \times K)$ case with K fixed, together with an expression for the overlap. Eventually, we mention the series of work (Huang & Sellke, 2023a;b) on multi-species spherical models involving random matrix tools that are closely related to the ones of this paper.

Spectral methods are also important as providing a *warm start* for other algorithmic approaches. This is the case, for instance, for approximate message passing algorithms (AMP) (Mondelli & Venkataramanan, 2021). AMP has attracted a lot of attention in the high-dimensional statistics and machine learning community, see e.g. (Donoho et al., 2009; Bayati & Montanari, 2011; Rangan, 2011; Rangan & Fletcher, 2012), and was written, for the inhomogeneous spiked problem, in (Pak et al., 2023). AMP algorithms and the corresponding weak recovery conditions were later developed for the much larger class of matrix tensor product models (Rossetti & Reeves, 2023), which includes the inhomogeneous spiked problem among many others. AMP algorithms are optimal among first-order methods (Celentano et al., 2020). Equipped with the warm start provided by our theorem, their reconstruction threshold thus provides a bound on the algorithmic performances. The link between AMP and spectral methods was discussed, for instance, in (Saade et al., 2014; Lesieur et al., 2017; Aubin et al., 2019; Mondelli & Montanari, 2018; Mondelli et al., 2022; Maillard et al., 2022; Venkataramanan et al., 2022).

2. Main Theoretical Results

2.1. Assumptions and Notations

We denote by $\mathbb{H}_{\pm} := \{z \in \mathbb{C}, \pm\Im\mathfrak{m} z > 0\}$ and by $\overline{\mathbb{H}_{\pm}} := \mathbb{H}_{\pm} \cup \mathbb{R}$, the positive upper (resp. lower) complex plane. For any positive integer d, we denote by $[\![d]\!] := \{1, \dots, d\}$ and for any subset I of $[\![d]\!]$, we denote by $|I|$ the cardinality of this set.

For two vectors of equal dimension, we introduce the partial order relation $\mathbf{x} \succ \mathbf{y}$ (resp. $\mathbf{x} \succeq \mathbf{y}$) to denote the usual relations $x_i > y_i$ (resp. $x_i \geq y_i$) for all i. We say that a vector $\mathbf{x}$ is positive (resp. non-negative) if $\mathbf{x} \succ \mathbf{0}$ (resp. $\mathbf{x} \succeq \mathbf{0}$). We will repeatedly use the simple identity that if $\max_i x_i \leq c$ then $\mathbf{x} \preceq c\mathbf{1}$. We denote by $|\mathbf{x}| = (|x_i|)_i$ the component-wise absolute value of vector. We denote by $\mathbf{D}_{\mathbf{x}} = \mathrm{Diag}((x_i))$ the diagonal matrix obtained from the vector $\mathbf{x}$.

In the following, when considering a matrix $\mathbf{A}$ of size $(N \times N)$ (or a vector of in $\mathbb{R}^N$) we omit the dependency in the dimension N and statement involving such a matrix $\mathbf{A}$ with $N \to \infty$ implicitly refer to a sequence $\mathbf{A} \equiv \mathbf{A}_{(N)}$ of such matrices. $(K \times K)$ matrices and K-dimensional vectors with K fixed are usually (but not always when it is clear from the context) denoted with a K index (e.g. $\mathbf{A}_K$ and $\mathbf{a}_K$) to differentiate them with $(N \times N)$ matrices and N-dimensional vectors. In particular, $\mathbf{I}_K$ is the identity matrix of size $(K \times K)$, $\mathbf{0}_K = (0, \dots, 0)$ and $\mathbf{1}_K = (1, \dots, 1)$ are the K-dimensional vectors of zeroes and ones respectively.

For a symmetric matrix $\mathbf{A}$, we denote by $(\lambda_i(\mathbf{A}))_{1 \leq i \leq N}$ its eigenvalues in decreasing order, in particular $\lambda_1(\mathbf{A})$ is the largest eigenvalue. We also denote by $\mu_{\mathbf{A}} := \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i(\mathbf{A})}$ the *empirical spectral distribution*. We say that an eigenvalue $\lambda_i(\mathbf{A})$ of $\mathbf{A}$ separates from the bulk if $\mu_{\mathbf{A}}$ converges weakly almost surely to a *limiting spectral distribution* μ_A and $\lim_{N\to\infty} \min_{x \in \mathrm{Supp}(\mu_A)} \mathrm{dist}(x, \lambda_i(\mathbf{A})) > 0$.

For an absolutely continuous distribution μ, we denote by $\mathrm{Supp}(\mu) := \{x \in \mathbb{R} | \mu(x) > 0\}$ its support and by $\mathrm{r}(\mu) := \sup\{\mathrm{Supp}(\mu)\}$ the rightmost edge of the distribution.

We first formalize the inhomogeneous setting:

Assumption 2.1. *There exist an integer $K \geq 1$ and a partition of $[\![N]\!] = B_1 \cup \dots \cup B_K$, such that if we denote by $\rho_k(N) = |B_k|/N$ the proportion of each group (thus $\sum_{k=1}^{K} \rho_k(N) = 1$), we have for any $k \in [\![K]\!]$:*

$$\rho_k(N) \to \rho_k \in (0, 1)\,. \tag{3}$$

Since the eigenproblem is invariant by permutations, one can assume without loss of generality a block structure for the model: that is for all $k < l$, one can assume that we have $i < j$ if $i \in B_k$ and $j \in B_l$, and that inside each group B_k, the entries are in increasing order. For this reason, we will refer to the $(B_k \times B_l)$ as blocks, also our results are stated in the generic setting of Assumption 2.1 without assuming this ordering. This partition structure is used to characterize the noise profile. We will make the following assumption on the variance profile matrix to avoid degenerate cases, although we believe this assumption will not change our final result.

Assumption 2.2. *The entries Δ_{ij} of the variance profile matrix $\boldsymbol{\Delta}$ are independent of N and constant within any block $(B_k \times B_l)$ of the partition of assumption 2.1 with values given by*

$$\Delta_{ij} = \frac{1}{s_{kl}} \quad \textit{for any } i \in B_k, i \in B_l\,, \tag{4}$$

where $s_{kl} > 0$.

We will denote by $(K \times K)$ matrix $\mathbf{S}_K := (s_{kl})_{k,l}$, the $(K \times K)$ symmetric matrix encoding the variance structure.

Assumption 2.3. *$\boldsymbol{H} = (H_{ij})$ is a $(N \times N)$ symmetric matrix with independent entries admitting moments of all orders, with mean zero and variance $\mathbb{E}H_{ij}^2 = 1$.*

We emphasize that our assumption is universal in the sense that it does not require the H_{ij} to be Gaussian (as stated originally the conjecture of (Pak et al., 2023)). Our last assumption concerns the property of the signal vector $\mathbf{x}$:

Assumption 2.4. *Under the partition of Assumption 2.1, for any $k \in [\![K]\!]$, the signal $\mathbf{x}$ satisfies the asymptotic property:*

$$\frac{\|(x_i)_{i\in B_k}\|}{\sqrt{|B_k|}} \xrightarrow[N\to\infty]{\text{a.s.}} 1\,. \tag{5}$$

Note that by the law of large numbers, Assumption 2.4 is satisfied if $\mathbf{x} \sim \pi^{\otimes N}$ where π is a distribution with mean zero and variance one, as originally stated in the conjecture of (Pak et al., 2023).

With $\mathbf{\Delta}$ as in Assumption 2.3 and $\mathbf{\Sigma} = \mathbf{\Delta}^{\odot -1}$, if we introduce the matrices

$$\boldsymbol{X} := \boldsymbol{H} \odot \mathbf{\Sigma}^{\odot 1/2} - \mathbf{D}_{(\mathbf{\Sigma}\mathbf{1})}\,, \tag{6}$$

$$\mathbf{Z} := \left(\sqrt{\frac{1}{N}}\mathbf{x}\mathbf{x}^\top\right) \odot \mathbf{\Sigma}\,, \tag{7}$$

one can write the transformed output matrix $\tilde{\boldsymbol{Y}}$ of Eq. (2) as

$$\tilde{\boldsymbol{Y}} = \boldsymbol{X} + \mathbf{Z}\,. \tag{8}$$

The asymptotic spectral properties of the output matrix $\tilde{\boldsymbol{Y}}$ are naturally described by two $(K \times K)$ symmetric matrices encoding all parameters (ρ_k, s_{kl}) of the model:

$$\mathbf{\Omega}_K := \mathbf{D}_{(\boldsymbol{\rho}^{\odot 1/2})}\,\mathbf{S}_K \mathbf{D}_{(\boldsymbol{\rho}^{\odot 1/2})}\,, \tag{9}$$

$$\mathbf{\Gamma}_K := \mathbf{S}_K \mathbf{D}_{\boldsymbol{\rho}} = \mathbf{D}_{(\boldsymbol{\rho}^{\odot -1/2})}\,\mathbf{\Omega}_K \mathbf{D}_{(\boldsymbol{\rho}^{\odot 1/2})}\,. \tag{10}$$

The matrices $\mathbf{\Gamma}_K$ and $\mathbf{\Omega}_K$ are similar and thus share the same (real) eigenvalues. By Assumptions 2.1 and 2.2, all elements of $\mathbf{\Omega}_K$ are positive and by Perron-Frobenius theorem (see e.g. Chap. 8 of (Meyer & Stewart, 2023)) the top eigenvalue $\lambda_1(\mathbf{\Omega}_K)$ is simple and corresponds to the operator norm.

Eventually, we introduce the vector $\mathbf{q} \in \mathbb{R}^K$ of the overlap between the top eigenvector $\boldsymbol{u}_1$ of $\tilde{\boldsymbol{Y}}/\sqrt{N}$ and the vector in each community as:

$$\mathbf{q} := \left(\left\langle \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|}, \boldsymbol{u}_1 \right\rangle, \ldots, \left\langle \frac{\mathbf{x}_K}{\|\mathbf{x}_K\|}, \boldsymbol{u}_1 \right\rangle\right), \tag{11}$$

where $(\mathbf{x}_k)_i := x_i$ for $i \in B_k$ and $(\mathbf{x}_k)_i = 0$ for $i \notin B_k$. Note that the overlap vector $\mathbf{q}$ contains by definition the information on the recovery of the signal for each community $(x_i)_{i\in B_k}$. From it, one can obtain the overlap with the entire vector, $m := \langle \boldsymbol{u}_1, \mathbf{x}/\|\mathbf{x}\|\rangle$, as we have $m = \langle \mathbf{q}, \boldsymbol{\rho}^{\odot 1/2}\rangle$.

Remark 2.5. For simplicity, our results are stated in the case where $\boldsymbol{Y}$ in Eq. (1) has a rank 1 spike. However, this is not a crucial model assumption and the analysis can be extended to a spike of rank-R where $R > 1$. Indeed, when working in the eigenbasis of the rank-R spike, the computations are completely decoupled in the new eigenbasis and reduce to the rank-1 case. This is also the case for finite rank spiked *homogeneous* random matrix models, see for example (Benaych-Georges & Nadakuditi, 2011).

2.2. Phase Transition for Top Eigenvalue and Top Eigenvector

Our main theorem (Thm. 2.6 below) of this work indicates that the top eigenvalue $\lambda_1(\mathbf{\Omega}_K)$ plays the role of the effective signal-to-noise ratio (SNR) for this block-structure spiked model where the threshold value $\lambda_1(\mathbf{\Omega}_K) = 1$ separates a phase where the top eigenvector has zero overlap with the signal vector to a phase where it has a positive overlap with the signal.

The behavior of the SNR with respect to the parameters (ρ_k, s_{kl}) of the model as well as several explicit examples are postponed to App. I. In particular, we observe that the effective SNR for block-structure models satisfy many natural properties such as monotonicity with respect to the s parameters in Prop I.1. Under some additional structure on $\mathbf{\Omega}_K$, we have explicit forms of and recover classical results such as the SNR for homogeneous spiked Wigner models and the asymmetric rank-one homogeneous spiked models (see App. I.3).

Theorem 2.6. *Under Assumptions 2.1,2.3,2.2 and 2.4, one has the following phase transition for the outlier and the overlap*

- *For $\lambda_1(\mathbf{\Omega}_K) \le 1$:*
  1. *there is asymptotically no eigenvalue separating outside of the bulk at values higher than $\mathrm{r}(\mu_X)$, that is $\lambda_1(\tilde{\boldsymbol{Y}}/\sqrt{N}) \xrightarrow[N\to\infty]{\text{a.s.}} \mathrm{r}(\mu_X)$, the rightmost edge of the bulk distribution.*
  2. $\mathbf{q} \xrightarrow[N\to\infty]{\text{a.s.}} \mathbf{0}_K = (0, \ldots, 0)$.
- *For $\lambda_1(\mathbf{\Omega}_K) > 1$:*
  1. $\lambda_1(\tilde{\boldsymbol{Y}}/\sqrt{N}) \xrightarrow[N\to\infty]{\text{a.s.}} 1$ *and separates from the bulk;*
  2. $|\mathbf{q}| \xrightarrow[N\to\infty]{\text{a.s.}} C^{-1/2}\,\mathbf{D}_{(\boldsymbol{\rho}^{\odot 1/2})}(\mathbf{1}_K - \mathbf{g}(1)) \succ \mathbf{0}$ *with the positive constant C given by*

$$C := \langle \mathbf{1}_K - \mathbf{g}(1), \mathbf{\Gamma}_K^\top \mathbf{D}_{(\boldsymbol{\rho}\odot\mathbf{y})} \mathbf{\Gamma}_K (\mathbf{1}_K - \mathbf{g}(1))\rangle\,, \tag{12}$$

where $\mathbf{y} = (\mathbf{\Gamma}_K - \mathbf{D}^{-2}_{\mathbf{g}(1)})^{-1}\mathbf{1}_K$ and $\mathbf{g}(1)$ is the continuation at $z = 1$ of the solution of the quadratic vector equation of Prop. 3.1.

We highlight that the expression for the overlap after the transition involves a quantity $\mathbf{g}(1)$ which is a root of a quadratic vector equation, and in general the solution of this system of equations does not admit a closed-form expression but can be evaluated to any precision level.

In Fig. 1, we illustrated the behavior of the theoretical limiting spectral distribution and top eigenvalue of the matrix $\tilde{\boldsymbol{Y}}/\sqrt{N}$ before, at, and after the critical value $\lambda_1(\boldsymbol{\Omega}_K) = 1$ for specific values of the parameters. In Fig. 2, we plotted the value of the overlaps for different models and different values of the $\lambda_1(\boldsymbol{\Omega}_K)$.

We recall that inference on inhomogeneous models of the form given by Eq. (1) are known to exhibit a *statistical-to-computational gap*, see Lem. 2.6 of (Guionnet et al., 2022) and (Aubin et al., 2019; Bandeira et al., 2018) for other models exhibiting such phenomenon. Our main result proves Conjecture 2.6 in (Pak et al., 2023) and shows that the threshold for PCA on the transformed matrix $\tilde{\boldsymbol{Y}}$ matches the optimal algorithmic recovery threshold for these inhomogeneous models, see Part 2 of Lem. 2.6 in (Guionnet et al., 2022). We also note from Prop. 2.18 of (Guionnet et al., 2022) that the thresholds for spectral methods on the original matrix $\boldsymbol{Y}$ or on $\boldsymbol{Y} \odot (\boldsymbol{\Delta}^{\odot -1/2})$ (which *a priori* removes the inhomogeneity in $\boldsymbol{Y}$ at the cost of transforming the signal vector $\mathbf{x}$) are always strictly worse than this optimal algorithmic recovery. This, in particular, demonstrates the relevance of using $\tilde{\boldsymbol{Y}}$, and the interest in the guaranty we provide, for spectral methods, rather than those two matrices.

3. Outline of the Proof

We break down the proof of Thm. 2.6 into several steps by first establishing the spectral properties of the matrices $\boldsymbol{X}/\sqrt{N}$ defined in Eq. (6) and $\mathbf{Z}$ defined in Eq. (7), separately. Combining these properties, we prove that if there is an outlier in $\tilde{\boldsymbol{Y}}/\sqrt{N}$, then a certain $(K \times K)$ matrix must have an eigenvalue at one. For $\lambda_1(\boldsymbol{\Omega}_K) < 1$ we show that this condition is not achievable, while for $\lambda_1(\boldsymbol{\Omega}_K) > 1$, we show that this condition is satisfied and sufficient by constructing explicitly the corresponding eigenvector. We also describe how one can obtain the overlap vector $\mathbf{q}$.

3.1. Properties of the Limiting Spectral Distribution and its Stieltjes Transform

Quadratic Vector Equation — The matrix $\tilde{\boldsymbol{Y}}/\sqrt{N}$ is a fixed-rank perturbation of the variance-profile Wigner matrix $\boldsymbol{X}/\sqrt{N}$. By Weyl's interlacing theorem, as N goes to infinity its limiting spectral distribution (LSD) is thus equal to μ_X, the LSD of $\boldsymbol{X}/\sqrt{N}$. The latter is well-understood and has been studied extensively in the series of works (Ajanki et al., 2017a;b; 2019), which tackle a more general setting than the block structure of this paper. Even in the block-structure setting, there is, in general, no closed-form expression for the density μ_X. It is however known that this distribution is supported on a finite union of compact intervals, and the behavior of this distribution at the endpoints of those intervals is also well understood.

In the following, we will mainly need the following characterization of the Stieltjes transform of the LSD. We recall from random matrix theory that the Stietljes transform of a continuous measure μ is defined for any $z \in \mathbb{C} \setminus \mathrm{Supp}(\mu)$ as $g_\mu(z) := \int_{\mathrm{Supp}(\mu)} (z-\lambda)^{-1} \mathrm{d}\mu(\lambda)$ and this transform uniquely characterizes the distribution μ. This Stietljes transform is characterized by the following result for variance-profile Wigner matrices.

Proposition 3.1 (Quadratic Vector Equation for the Stieltjes Transform)**.** *Under Assumptions 2.1, 2.2, 2.3, one has:*

(a) the Stieltjes transform $g_{\mu_X}(z)$ of the limiting spectral distribution μ_X of the matrix $\boldsymbol{X}/\sqrt{N}$, admits the decomposition $g_{\mu_X}(z) := \sum_{k=1}^K \rho_k g_k(z)$ where for any $z \in \mathbb{H}_-$, $\mathbf{g}(z) := (g_1(z), \dots, g_K(z))$ is the unique solution of the quadratic vector equation (QVE):

$$\mathbf{1}_K = z\mathbf{g} - \mathbf{g} \odot \boldsymbol{\Gamma}_K(\mathbf{g} - \mathbf{1}_K)\,, \tag{13}$$

such that

$$\mathbf{g}(z) \in (\mathbb{H}_+)^K\,. \tag{14}$$

(b) $\lambda_1(\boldsymbol{X}/\sqrt{N}) \xrightarrow[N\to\infty]{\text{a.s.}} \mathrm{r}(\mu_X)$.

Proof. We refer to (Ajanki et al., 2017a;b; 2019) for the complete proof of part (a) and to App. A for a short sketch of the proof of this part. Part (b) follows rom Cor. 1.10 of (Ajanki et al., 2017b) (see also Thm. A.9 of (Husson, 2022)). □

Evaluation on the Real Line — Thm. 3.1 completely characterizes the Stieltjes transform *outside* the real line. Yet, in many cases, one is interested in understanding the behavior of this Stietljes transform precisely on this set. The property of the continuation on $\mathbb{R}$ is given the following result, from the same authors.

Proposition 3.2 (Continuation on the real line)**.** *The solution $\mathbf{g}(.)$ of the quadratic vectorial equation of Eq. (13) on $\mathbb{H}_-$ extends to $1/3$-Hölder continuous function $\mathbf{g} : \overline{\mathbb{H}_-} \to (\overline{\mathbb{H}_+})^K$ and is analytical on $\mathbb{C} \setminus \overline{\mathrm{Supp}(\mu_X)}$.*

Proof. See Corr. 2.7 of (Ajanki et al., 2017a) and Thm. A.2 of (Huang & Sellke, 2023b). □

Note that while Prop. 3.2 gives a regularity property for the continuation on the real line, it does not immediately give

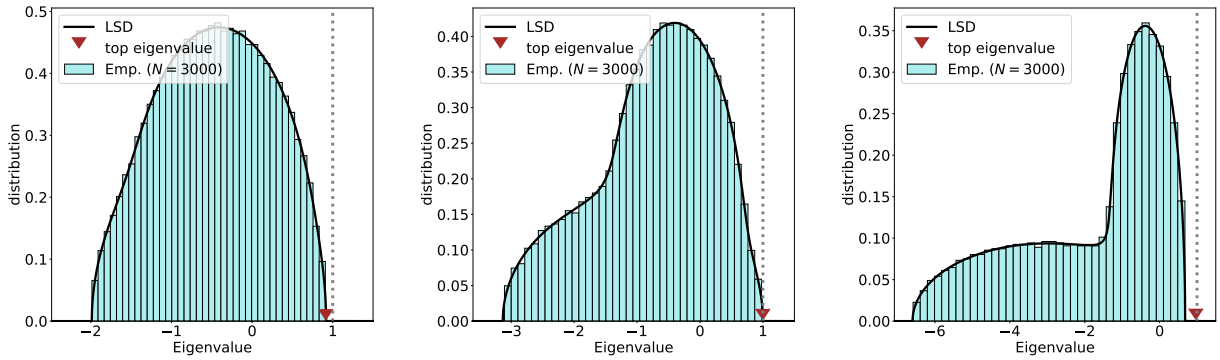

Figure 1. **Eigenvalue Distribution and Top Eigenvalue Position for Different Value of $\lambda_1(\mathbf{\Omega}_K)$.** The three figures correspond to a model with $K = 2$, $\boldsymbol{\rho} = (1/2, 1/2)$ and $\mathbf{S}_K = \begin{pmatrix} t & 1/2 \\ 1/2 & 1/4 \end{pmatrix}$ with a value of the parameter t different in each figure and chosen such that **(Left)** $\lambda_1(\mathbf{\Omega}_K) = 0.5 < 1$, **(Center)** $\lambda_1(\mathbf{\Omega}_K) = 1.0$ and **(Right)** $\lambda_1(\mathbf{\Omega}_K) = 3.0 > 1$. The black curve corresponds to the theoretical value of the limiting spectral distribution (LSD) obtained by solving the QVE of Eq. (13) numerically, while the colored histogram corresponds to the empirical distribution of a sample $\tilde{\boldsymbol{Y}}/\sqrt{N}$ with $N = 3000$, and the red triangle corresponds to the empirical value of its largest eigenvalue. The signal $\mathbf{x}$ has been sampled from a standard normal distribution. Before the transition (left), the rightmost edge is below one and there is no outlier; at the transition (center), the rightmost edge touches the value one; and after the transition (right) there is an outlier at one.

us a way to know if a certain solution of Eq. (13) evaluated on $\mathbb{R} \setminus \overline{\mathrm{Supp}(\mu_X)}$ corresponds to the proper analytical continuation of $\mathbf{g}(.)$. Take for example $z = 1$, then one can immediately check that $\mathbf{1}_K$ is always *one* solution (among the 2^K possible solutions) of Eq. (13), yet does it correspond to $\mathbf{g}(1)$ or to a spurious solution of Eq. (13)? The following result shows how to discriminate between the two cases.

Lemma 3.3. *Let $\tilde{\mathbf{g}}(\lambda) \in (\mathbb{R}_*)^K$ be any solution of Eq. (13) at $z = \lambda \in \mathbb{R} \setminus \overline{\mathrm{Supp}(\mu_X)}$ then $\tilde{\mathbf{g}}(\lambda) = \mathbf{g}(\lambda)$ is the proper solution of the QVE if and only if the linear system of equations*

$$(\mathbf{D}^{-2}_{\tilde{\mathbf{g}}(\lambda)} - \mathbf{\Gamma}_K)\mathbf{y} = \mathbf{1}_K \,, \tag{15}$$

admits a unique positive solution $\mathbf{y} \succ \mathbf{0}_K$.

Proof. The proof of this result is given in Appendix. B and relies on studying the behavior of Eq. (13) as we approach the value $z = \lambda$ from the lower complex plane. See also (Huang & Sellke, 2023b) for a similar statement and proof. □

Next, we show a necessary and sufficient condition for Eq. (15) to have a positive solution.

Proposition 3.4. *([Condition for Positive Solution] The linear system of Eq. (15) admits a unique positive solution if and only if $\lambda_1(\mathbf{D}^2_{\tilde{\mathbf{g}}(\lambda)}\mathbf{\Omega}_K) < 1$.*

Proof. The proof of (a generalized version of) this result is given in App. C and relies on properties of so-called M- and Z- matrices and the use of Farka's lemma to characterize positive solutions of (15). □

Behavior of the Rightmost Edge — To understand if the top eigenvalue separates from the bulk, we need to understand the behavior of the rightmost edge, which is given by the following result.

Proposition 3.5 (The rightmost edge is bounded by one)**.** *Let $\mathrm{r}(\mu_X)$ be the rightmost edge of the limiting spectral distribution μ_X of $\tilde{\boldsymbol{Y}}/\sqrt{N}$ then $\mathrm{r}(\mu_X) \leq 1$ with equality if and only if $\lambda_1(\mathbf{\Omega}_K) = 1$.*

Proof. This result is given in App. D. It follows from a characterization of the edge of the spectrum in terms of the singularities of the solution to the quadratic vector equations as in Thm. 2.6 of (Ajanki et al., 2017a) □

Eventually, we will use the following lemma, characterizing the Stieltjes transform above the rightmost edge.

Lemma 3.6 (Positivity and Monotonicity of $\mathbf{g}(.)$ above the rightmost edge)**.** *Let $\mathrm{r}(\mu_X)$ be the rightmost edge of μ_X, then for any $\lambda \in (\mathrm{r}(\mu_X), \infty)$ one has $\mathbf{g}(\lambda) \succ \mathbf{0}_K$ and this function is entrywise analytically decreasing on this interval.*

Proof. This is a standard result in complex analysis and follows from the fact (see (Ajanki et al., 2019)) that for

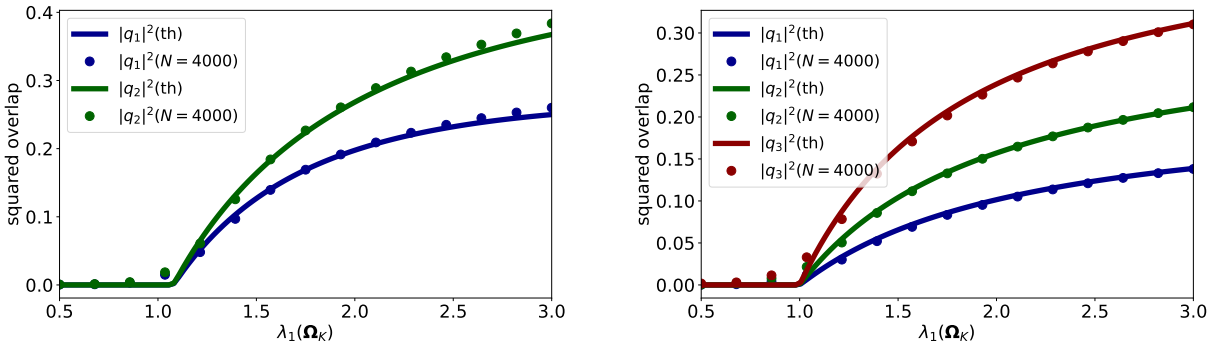

Figure 2. **Squared Overlap Vector for Different Value of $\lambda_1(\mathbf{\Omega}_K)$ and Different Models. (Left)** Value of the square of the overlaps for a model with $K = 2$, $\boldsymbol{\rho} = (1/3, 2/3)$ and $\mathbf{S}_K = \begin{pmatrix} 1/2 & t \\ t & 1/2 \end{pmatrix}$ where the range of the parameter t is set such that $\lambda_1(\mathbf{\Omega}_K)$ varies from 0.5 to 3. **(Right)** Value of the square of the overlaps for a model with $K = 3$, $\boldsymbol{\rho} = (2/9, 3/9, 4/9)$ and $\mathbf{S}_K = \begin{pmatrix} 3 & 1/2 & t \\ 1/2 & 1 & t \\ t & t & 1 \end{pmatrix}$ where the range of the parameter t is set such that $\lambda_1(\mathbf{\Omega}_K)$ varies from 0.5 to 3. In both cases, the dots represent an average over 30 samples of the empirical value of the overlaps for $N = 4000$, with $\mathbf{x}$ sampled from a standard Gaussian distribution.

any $k \in [\![K]\!]$, $g_k(z) := \int (z-\lambda)^{-1} \mathrm{d}\mu_k$ for some positive measure μ_k. Since $g_{\mu_X} = \sum_k \rho_k g_k$, we must have for any $k \in [\![K]\!]$, the rightmost edge $\mathrm{r}(\mu_k)$ of μ_k satisfies the bound $\mathrm{r}(\mu_k) \leq \mathrm{r}(\mu_X)$. Since $g_k'(z) = -\int (z-\lambda)^{-2} \mathrm{d}\mu_k < 0$ and $g_k(z) \underset{z\to\infty}{\sim} 1/z$, we have the desired result. □

Behavior at $z = 1$ – As stated in our main theorem, the position $z = 1$ is peculiar and the behavior of the Stietljes transform at this point changes depending on the effective SNR ($\lambda_1(\mathbf{\Omega}_K)$) value.

Proposition 3.7 (Behavior of the Stietljes at $z = 1$)**.** *We have*

(a) For $\lambda_1(\mathbf{\Omega}_K) \leq 1$: $\mathbf{g}(1) = \mathbf{1}_K$;

(b) For $\lambda_1(\mathbf{\Omega}_K) > 1$:

(i) $\mathbf{0} \prec \mathbf{g}(1) \prec \mathbf{1}_K$;

(ii) furthermore, if we introduce the vectors:

$$\mathbf{v}_1^{(r)} := \mathbf{D}_{(\boldsymbol{\rho}^{\odot 1/2})}(\mathbf{1}_K - \mathbf{g}(1)) \,, \tag{16}$$

$$\mathbf{v}_1^{(l)} := \mathbf{\Omega}_K \mathbf{v}_1^{(r)} \,, \tag{17}$$

$$\mathbf{w}_1 := \frac{\mathbf{D}_{\mathbf{g}(1)}^{-1/2} \mathbf{v}_1^{(r)}}{\|\mathbf{D}_{\mathbf{g}(1)}^{-1/2} \mathbf{v}_1^{(r)}\|} \,, \tag{18}$$

then $\mathbf{v}_1^{(l)}, \mathbf{v}_1^{(r)}$ are the (unnormalized) left and right top eigenvectors of $\mathbf{D}_{\mathbf{g}(1)}\mathbf{\Omega}_K$ associated to the simple top eigenvalue 1 and $\mathbf{w}_1$ is the orthonormal top eigenvector of $\mathbf{D}_{\mathbf{g}(1)}^{1/2}\mathbf{\Omega}_K\mathbf{D}_{\mathbf{g}(1)}^{1/2}$ associated with the simple top eigenvalue 1.

Proof. At $z = 1$, the vector $\mathbf{1}_K$ is one solution of the QVE (13).

(a): For $\lambda_1(\mathbf{\Omega}_K) \leq 1$, Prop. 3.4 gives us immediately Part *(a)* of the claim.

(b)-(i): Conversely for $\lambda_1(\mathbf{\Omega}_K) > 1$, we must have $\mathbf{g}(1) \neq \mathbf{1}_K$ by the same proposition, and since $\mathrm{r}(\mu_X) < 1$ (Prop. 3.5) and the positivity property of the Stieltjes (Lem. 3.6), we also have $\mathbf{g}(1) \succ \mathbf{0}_K$. An analysis of the solution of the QVE described in App. E allows one to conclude.

(b)-(ii): Using the identity $\mathbf{a} \odot \mathbf{b} \equiv \mathbf{D}_{\mathbf{a}}\mathbf{b}$ and the definition of the matrix $\mathbf{\Gamma}_K$ of Eq. (10), we can write the QVE (13) as:

$$(\mathbf{D}_{\mathbf{g}(1)}\mathbf{\Omega}_K)\mathbf{v}_1^{(r)} = \mathbf{v}_1^{(r)} \,, \tag{19}$$

and from Part *(b)-(i)* of this proposition, we have

$$\mathbf{v}_1^{(r)} \succ \mathbf{0}_K \,, \tag{20}$$

since $\mathbf{D}_{\mathbf{g}(1)}\mathbf{\Omega}_K$ is entrywise positive, $\mathbf{v}_1^{(r)}$ corresponds, by Perron Frobenius theorem, to its (right) top eigenvectorand the corresponding eigenvalue is equal to one (Eq. (19)) and is simple. The result for $\mathbf{v}_1^{(l)}$ (resp. $\mathbf{w}_1$) is obtained by right-multiplying Eq. (20) by $\mathbf{\Omega}_K$ (resp. by $\mathbf{D}_{\mathbf{g}(1)}^{-1/2}$). □

3.2. Eigendecomposition of the Small-Rank Matrix

Next, we turn our attention to the small-rank matrix $\mathbf{Z}$ given by Eq. (7). It is the Hadamard product of a rank-one matrix and an a priori rank-K matrix and thus is a priori also of rank-K by standard properties of Hadamard product (see e.g. (Johnson, 1990)). Our next result expresses $\mathbf{Z}$ as a rotation of the matrix $\mathbf{\Omega}_K$, up to a small error.

Proposition 3.8. *Under Assumptions 2.1 and 2.4, one has almost surely for N large enough,*

$$\mathbf{Z}/\sqrt{N} = \mathbf{V}\left(\boldsymbol{\Omega}_K + \boldsymbol{E}_K\right)\mathbf{V}^\top\,, \tag{21}$$

where $\boldsymbol{E}_K$ is a $(K \times K)$ matrix such that $\|\boldsymbol{E}_K\|_{\rm op} \xrightarrow[N\to\infty]{\rm a.s.} 0$, $\mathbf{V} := \left[\mathbf{v}^{(1)}, \dots, \mathbf{v}^{(K)}\right] \in \mathbb{R}^{N\times K}$ and the $\mathbf{v}^{(k)}$'s are orthonormal vectors satisfying the block structure property $(\mathbf{v}^{(k)})_i := x_i/\|(x_i)_{i\in B_k}\|$ for $i \in B_k$ and $(\mathbf{v}^{(k)})_i = 0$ for $i \notin B_k$.

Proof. By definition of $\mathbf{Z}/\sqrt{N}$,for N large enough if we denote for any $k \in [\![K]\!]$, the vector $\mathbf{x}_k$ where $(\mathbf{x}_k)_i := x_i$ for $x \in B_k$ and $(\mathbf{x}_k)_i = 0$ for $i \notin B_k$ as defined in (11), we have:

$$\frac{\mathbf{Z}}{\sqrt{N}} = \frac{1}{N}\sum_{k,l}^{K} s_{kl}\,\mathbf{x}_k\mathbf{x}_l^\top\,. \tag{22}$$

thus, if we normalize the vectors by their norms we get

$$\frac{\mathbf{Z}}{\sqrt{N}} = \sum_{k,l}^{K} s_{kl}\frac{\sqrt{|B_k||B_l|}}{N}\,\frac{\|\mathbf{x}_k\|\|\mathbf{x}_l\|}{\sqrt{|B_k||B_l|}}\,\frac{\mathbf{x}_k}{\|\mathbf{x}_k\|}\left(\frac{\mathbf{x}_l}{\|\mathbf{x}_l\|}\right)^\top, \tag{23}$$

where we have dropped the dependency in N for $B_k(N)$ for clarity. By Assumption 2.1 we have $|B_k|/N \to \rho_k \in (0,1)$ and by Assumption 2.4 $\|\mathbf{x}_k\|/\sqrt{|B_k|} \xrightarrow[N\to\infty]{\rm a.s} 1$. Since $(\boldsymbol{\Omega}_K)_{kl} = s_{kl}\sqrt{\rho_k}\sqrt{\rho_l}$, the remainder term converges almost surely to zero, and hence the operator norm of the associated matrix $\boldsymbol{E}_K$ converges also almost surely to zero by matrix norm equivalence and the fact its dimension K is fixed, thus we have the desired result. □

The condition N large enough is only needed to ensure the existence of the partition introduced in Assumption 2.1.

3.3. Proof of the Phase Transition for the Top Eigenvalue

Condition for the Existence of an Outlier – We are now ready to combine the previous results describing the spectrum and eigenvectors of the matrices $\boldsymbol{X}/\sqrt{N}$ and $\mathbf{Z}$ to describe the spectrum of the matrix $\tilde{\boldsymbol{Y}}/\sqrt{N}$ and in particular, its potential outliers. The following proposition gives our first result in this direction.

Proposition 3.9 (Equation for outliers)**.** *Under the same assumptions as in Thm. 2.6, z is the position of an outlier separating from the bulk, if and only if z is an almost-sure solution of*

$$\det\left(\mathbf{I}_K - \mathbf{D}_{\mathbf{g}(z)}\,\boldsymbol{\Omega}_K\right) = 0\,, \tag{24}$$

where $\mathbf{g}(z) = (g_1(z), \dots, g_K(z))$, is the analytical continuation on the real line of the solution of the QVE of Prop. 3.1.

Proof. The proof of this result is given in App. F. It relies on using (a generalized version of) the matrix determinant lemma, together with the eigendecomposition of Prop. 3.8 and deterministic equivalent for the resolvent of the matrix $\boldsymbol{X}/\sqrt{N}$. □

Existence of a (Top) Outlier at One in the Regime $\lambda_1(\boldsymbol{\Omega}_K) > 1$ – This is an immediate consequence of Part *(b)-(ii)* of Prop. 3.7, since for $\lambda_1(\boldsymbol{\Omega}_K) > 1$, 1 is the highest eigenvalue of $\mathbf{D}_{\mathbf{g}(1)}\boldsymbol{\Omega}_K$, hence $z = 1$ is the position of an outlier separating from the bulk since $\mathrm{r}(\mu_X) < 1$ (Prop. 3.5) in this regime. Since $\mathbf{g}$ is entrywise decreasing (Lem. 3.6) for any $z > 1$, $\lambda_1(\mathbf{D}_{\mathbf{g}(z)}\boldsymbol{\Omega}_K) < \lambda_1(\mathbf{D}_{\mathbf{g}(1)}\boldsymbol{\Omega}_K)$ and hence 1 is the position of the top outlier.

Non-Existence of an Outlier Above the Rightmost Edge in the Regime $\lambda_1(\boldsymbol{\Omega}_K) \le 1$ – We show that Prop. 3.9 is incompatible with the properties of the Stietljes transform derived in Sec. 3.1:

- On the one hand, we have that $z > \mathrm{r}(\mu_X)$, with $\mathrm{r}(\mu_X) \le 1$ (Prop. 3.5), is a position of an outlier if there is a $k \in [\![K]\!]$ such that $\lambda_k(\mathbf{D}_{\mathbf{g}(z)}\boldsymbol{\Omega}_K) = 1$. Since $\lambda_1(\mathbf{D}_{\mathbf{g}(z)}\boldsymbol{\Omega}_K) \ge \lambda_k(\mathbf{D}_{\mathbf{g}(z)}\boldsymbol{\Omega}_K)$, we must have by operator norm inequality:

$$1 \le \lambda_1(\mathbf{D}_{\mathbf{g}(z)}\boldsymbol{\Omega}_K) \le \lambda_1(\mathbf{D}_{\mathbf{g}(z)})\;\lambda_1(\boldsymbol{\Omega}_K)\,. \tag{25}$$

  Since $\mathbf{g}$ is entrywise decreasing on $(\mathrm{r}(\mu_X), \infty)$ (Lem. 3.6) and $\mathbf{g}(1) = \mathbf{1}_K$ (Prop. 3.7-*(a)*), this can only be satisfied if there exists a $z \in (\mathrm{r}(\mu_X), 1)$ for which we have

$$\mathbf{g}(z) \succ \mathbf{1}_K\,. \tag{26}$$

- On the other hand, for any $z \in (\mathrm{r}(\mu_X), 1)$, the analytical continuation of the Stietljes transform must satisfy by Prop. 3.4:

$$\lambda_1(\mathbf{D}^2_{\mathbf{g}(z)}\boldsymbol{\Omega}_K) < 1\,. \tag{27}$$

To see that these two conditions are incompatible we will make use of the following lemma, which is a consequence of the operator norm inequality:

Lemma 3.10 (Reversed Operator Norm Inequality)**.** *Let $\mathbf{A} \in \mathbb{R}^{d\times d}$ such that $\lambda_1(\mathbf{A}) \ge \dots \ge \lambda_d(\mathbf{A}) > 1$ and $\mathbf{B} \in \mathbb{R}^{d\times m}$ then we have $\|\mathbf{B}\|_{\rm op} < \|\mathbf{AB}\|_{\rm op}$.*

Proof. Since $\mathbf{A}$ is diagonalizable with non-zero eigenvalue, its inverse exists and we have by the operator norm inequality $\|\mathbf{B}\|_{\rm op} = \|\mathbf{A}^{-1}\mathbf{AB}\|_{\rm op} \le \lambda_1(\mathbf{A}^{-1})\|\mathbf{AB}\|_{\rm op} \le \frac{\|\mathbf{AB}\|_{\rm op}}{\lambda_d(\mathbf{A})} < \|\mathbf{AB}\|_{\rm op}$. □

If there exists an outlier at $z \in (\mathrm{r}(\mu_X), 1)$ with $\mathbf{g}(z)$ satisfying Eq. (26), we must have by Lem. 3.10:

$$\lambda_1(\mathbf{D}^2_{\mathbf{g}(z)}\boldsymbol{\Omega}_K) > \lambda_1(\mathbf{D}_{\mathbf{g}(z)}\boldsymbol{\Omega}_K) \geq 1\,, \tag{28}$$

where the last inequality follows from Eq. (25) and is incompatible with Eq. (27), such that there is no outlier above the rightmost edge in this regime.

3.4. Outline of the Proof for the Overlap Vector

To compute the limiting value of the overlap vector $\mathbf{q}$, let us first notice that this vector can be written in terms of eigenmatrix $\mathbf{V}$ of Prop. 3.8 as

$$\mathbf{q} = \mathbf{V}^\top \boldsymbol{u}_1\,, \tag{29}$$

where we recall that $\boldsymbol{u}_1$ is the eigenvector associated with the top eigenvalue of $\tilde{\boldsymbol{Y}}/\sqrt{N}$.

Next, the *resolvent* of $\tilde{\boldsymbol{Y}}/\sqrt{N}$ is defined for any $z \in \mathbb{C} \setminus \mathrm{Spec}(\tilde{\boldsymbol{Y}}/\sqrt{N})$ by

$$\mathbf{G}_{\frac{\tilde{\boldsymbol{Y}}}{\sqrt{N}}}(z) := \Big(z - \frac{\tilde{\boldsymbol{Y}}}{\sqrt{N}}\Big)^{-1}, \tag{30}$$

$$= \sum_{i=1}^{N} \Big(z - \lambda_i\big(\frac{\tilde{\boldsymbol{Y}}}{\sqrt{N}}\big)\Big)^{-1} \mathbf{u}_i\mathbf{u}_i^\top\,. \tag{31}$$

As $\lambda_1(\tilde{\boldsymbol{Y}}/\sqrt{N}) \xrightarrow[N\to\infty]{\text{a.s.}} 1$, separates from the bulk and is simple if $\lambda_1(\boldsymbol{\Omega}_K) > 1$, we have the following identity:

Lemma 3.11. *For $\lambda_1(\boldsymbol{\Omega}_K) > 1$, we have*

$$\lim_{N\to\infty} \mathbf{q}\mathbf{q}^\top = \lim_{z\to 1}(z-1) \cdot \lim_{N\to\infty} \mathbf{V}^\top \mathbf{G}_{\frac{\tilde{\boldsymbol{Y}}}{\sqrt{N}}}(z)\mathbf{V}\,. \tag{32}$$

From this identity, we first obtain the following intermediate result:

Proposition 3.12 (The Overlap Vector is Proportional to the Right Top Eigenvector)**.** *For $\lambda_1(\boldsymbol{\Omega}_K) > 1$, we have:*

$$\mathbf{q}\mathbf{q}^\top \xrightarrow[N\to\infty]{\text{a.s.}} \frac{\mathbf{v}_1^{(r)}(\mathbf{v}_1^{(r)})^\top}{-\phi_1'(1)\cdot C_0}\,, \tag{33}$$

where $\mathbf{v}_1^{(r)}$ is given by Eq. (16), $\phi_1(\lambda)$ is the top eigenvalue of $\mathbf{D}_{\mathbf{g}(\lambda)}\boldsymbol{\Omega}_K$, $\phi_1'(.)$ is the derivative of $\phi_1(.)$ with respect to λ and $C_0 = \langle \mathbf{v}_1^{(r)}, \mathbf{D}^{-1}_{\mathbf{g}(1)}\mathbf{v}_1^{(r)}\rangle$.

Proof. The proof of this result is given in Appendix G. It relies first on using Woodbury identity to express the resolvent of $\tilde{\boldsymbol{Y}}/\sqrt{N}$ in terms of the one of $\boldsymbol{X}/\sqrt{N}$, then using deterministic equivalent to simplify the expression and eventually identify the limit in Lem. 3.11 as the derivative of $\phi_1(.)$ evaluated at one. □

To finish the proof, we show that

Proposition 3.13. *For $\lambda_1(\boldsymbol{\Omega}_K) > 1$, with $\phi_1'(1)$ as in Prop. 3.12 we have*

$$-\phi_1'(1) = \frac{C}{C_0}\,, \tag{34}$$

where C is given by Eq. (12) and C_0 is given in Prop. 3.12.

Proof. The proof of this result is given in Appendix H and relies on perturbation theory for the top eigenvalue of $\mathbf{D}^{1/2}_{\mathbf{g}(1+\epsilon)}\boldsymbol{\Omega}_K\mathbf{D}^{1/2}_{\mathbf{g}(1+\epsilon)}$. □

Thanks to the definition (16) of $\mathbf{v}_1^{(r)}$, one has the desired result for the overlap.

The case $\lambda_1(\boldsymbol{\Omega}_K) \leq 1$ where there is no outlier, follows from standard results in RMT, see e.g. (Benaych-Georges & Nadakuditi, 2011).

4. Conclusion

In this work, we studied the behavior of a spiked block-Wigner matrix model given by Eq. (2). Our first contribution is the proof of a phase transition for the behavior of the top eigenvalue of this model, generalizing the seminal work of (Baik et al., 2005) to inhomogeneous problems, proving a conjecture stated in (Pak et al., 2023), and providing a sharp optimal method for detection for the inhomogenous spiked model (Behne & Reeves, 2022; Alberici et al., 2021a; 2022; Guionnet et al., 2022). Our results could also be used as spectral start (Mondelli & Venkataramanan, 2021) for the AMP algorithms of (Pak et al., 2023), making them optimal in terms of MMSE for these problems in the class of iterative algorithm (Guionnet et al., 2022).

Our second contribution is a sharp characterization of the overlap between the associated top eigenvector of this matrix and the original signal vector, undergoing a similar phase transition. While we have, for simplicity, considered the setting where the signal is a rank-one matrix, our computation can be carried to an arbitrarily fixed rank-R case (see Rem. 2.5).

An interesting venue for future work is to extend this computation to a general setting of general variance-profile shape (not necessarily of block type) and we leave this problem for future work. Finally, it would also be interesting to understand the fluctuations of the top eigenvalue around its limiting value in this inhomogeneous one, which are very well understood for a homogeneous problem.

Acknowledgements

The authors would like to thank Alice Guionnet and Aleksandr Pak and Lenka Zdeborová for fruitful discussions at the early stage of this project and Brice Huang and Mark Selke for discussions on their work (Huang & Sellke, 2023b). We also acknowledge funding from the Swiss National Science Foundation grant SNFS OperaGOST (grant number 200390), from the ERC Project LDRAM: ERC-2019-ADG Project 884584, and from the Natural Sciences and Engineering Research Council of Canada (NSERC) (RGPIN-2020-04597).

Impact Statement

This paper presents work whose goal is to advance the field of Machine Learning. There are many potential societal consequences of our work, none of which we feel must be specifically highlighted here.

References

Ajanki, O., Krüger, T., and Erdős, L. Singularities of Solutions to Quadratic Vector Equations on the Complex Upper Half-Plane. *Communications on Pure and Applied Mathematics*, 70(9):1672–1705, 2017a. ISSN 1097-0312. doi: 10.1002/cpa.21639. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/cpa.21639. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cpa.21639.

Ajanki, O., Erdos, L., and Krüger, T. Quadratic vector equations on complex upper half-plane. *Memoirs of the American Mathematical Society*, 261(1261):0–0, September 2019. ISSN 0065-9266, 1947-6221. doi: 10.1090/memo/1261. URL http://arxiv.org/abs/1506.05095. arXiv:1506.05095 [math-ph].

Ajanki, O. H., Erdős, L., and Krüger, T. Universality for general Wigner-type matrices. *Probability Theory and Related Fields*, 169(3):667–727, December 2017b. ISSN 1432-2064. doi: 10.1007/s00440-016-0740-2. URL https://doi.org/10.1007/s00440-016-0740-2.

Alaoui, A. E., Krzakala, F., and Jordan, M. Fundamental limits of detection in the spiked Wigner model. *The Annals of Statistics*, 48(2):863 – 885, 2020. doi: 10.1214/19-AOS1826. URL https://doi.org/10.1214/19-AOS1826.

Alberici, D., Camilli, F., Contucci, P., and Mingione, E. The multi-species mean-field spin-glass on the Nishimori line. *Journal of Statistical Physics*, 182:1–20, 2021a.

Alberici, D., Camilli, F., Contucci, P., and Mingione, E. The solution of the deep Boltzmann machine on the Nishimori line. *Comm. Math. Phys.*, 387(2): 1191–1214, 2021b. ISSN 0010-3616. doi: 10.1007/s00220-021-04165-0. URL https://doi.org/10.1007/s00220-021-04165-0.

Alberici, D., Camilli, F., Contucci, P., and Mingione, E. A statistical physics approach to a multi-channel Wigner spiked model. *Europhysics Letters*, 136(4):48001, 2022.

Alt, J., Erdős, L., and Krüger, T. The Dyson Equation with Linear Self-Energy: Spectral Bands, Edges and Cusps. *Documenta Mathematica*, 25:1421–1539, January 2020. ISSN 1431-0635. doi: 10.4171/dm/780. URL https://ems.press/journals/dm/articles/8965711.

Aubin, B., Loureiro, B., Maillard, A., Krzakala, F., and Zdeborová, L. The spiked matrix model with generative priors. *Advances in Neural Information Processing Systems*, 32, 2019.

Baik, J. and Silverstein, J. W. Eigenvalues of large sample covariance matrices of spiked population models. *Journal of Multivariate Analysis*, 97(6):1382–1408, July 2006. ISSN 0047-259X. doi: 10.1016/j.jmva.2005.08.003. URL https://www.sciencedirect.com/science/article/pii/S0047259X0500134X.

Baik, J., Ben Arous, G., and Péché, S. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab.*, 33(5): 1643–1697, 2005. ISSN 0091-1798. doi: 10.1214/009117905000000233. URL http://dx.doi.org/10.1214/009117905000000233.

Bandeira, A. S., Perry, A., and Wein, A. S. Notes on computational-to-statistical gaps: predictions using statistical physics. *Portugaliae mathematica*, 75(2):159–186, 2018.

Bandeira, A. S., El Alaoui, A., Hopkins, S., Schramm, T., Wein, A. S., and Zadik, I. The Franz-Parisi criterion and computational trade-offs in high dimensional statistics. *Advances in Neural Information Processing Systems*, 35: 33831–33844, 2022.

Barbier, J. and Reeves, G. Information-theoretic limits of a multiview low-rank symmetric spiked matrix model. In *2020 IEEE International Symposium on Information Theory (ISIT)*, pp. 2771–2776. IEEE, 2020.

Barbier, J., Dia, M., Macris, N., Krzakala, F., and Zdeborová, L. Rank-one matrix estimation: analysis of algorithmic and information theoretic limits by the spatial coupling method. *arXiv preprint arXiv:1812.02537*, 2018.

Bayati, M. and Montanari, A. The dynamics of message passing on dense graphs, with applications to compressed sensing. *IEEE Transactions on Information Theory*, 57 (2):764–785, 2011. doi: 10.1109/TIT.2010.2094817.

Behne, J. K. and Reeves, G. Fundamental limits for rank-one matrix estimation with groupwise heteroskedasticity. In *International Conference on Artificial Intelligence and Statistics*, pp. 8650–8672. PMLR, 2022.

Benaych-Georges, F. and Nadakuditi, R. R. The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Advances in Mathematics*, 227(1):494–521, May 2011. ISSN 0001-8708. doi: 10.1016/j.aim.2011.02.007. URL https://www.sciencedirect.com/science/article/pii/S0001870811000570.

Benaych-Georges, F. and Nadakuditi, R. R. The singular values and vectors of low rank perturbations of large rectangular random matrices. *Journal of Multivariate Analysis*, 111:120–135, October 2012. ISSN 0047-259X. doi: 10.1016/j.jmva.2012.04.019. URL https://www.sciencedirect.com/science/article/pii/S0047259X12001108.

Berman, A. and Plemmons, R. J. *Nonnegative matrices in the mathematical sciences*. SIAM, 1994.

Bigot, J. and Male, C. Freeness over the diagonal and outliers detection in deformed random matrices with a variance profile. *Information and Inference: A Journal of the IMA*, 10(3):863–919, July 2020. ISSN 2049-8772. doi: 10.1093/imaiai/iaaa012. _eprint: https://academic.oup.com/imaiai/article-pdf/10/3/863/40364247/iaaa012.pdf.

Bloemendal, A. and Virág, B. Limits of spiked random matrices I. *Probability Theory and Related Fields*, 156 (3):795–825, August 2013. ISSN 1432-2064. doi: 10.1007/s00440-012-0443-2. URL https://doi.org/10.1007/s00440-012-0443-2.

Boyd, S. P. and Vandenberghe, L. *Convex optimization*. Cambridge University Press, Cambridge, UK ; New York, 2004. ISBN 978-0-521-83378-3.

Capitaine, M., Donati-Martin, C., and Féral, D. The largest eigenvalues of finite rank deformation of large Wigner matrices: Convergence and nonuniversality of the fluctuations. *The Annals of Probability*, 37(1):1–47, January 2009. ISSN 0091-1798, 2168-894X. doi: 10.1214/08-AOP394. Publisher: Institute of Mathematical Statistics.

Capitaine, M., Donati-Martin, C., and Féral, D. Central limit theorems for eigenvalues of deformations of Wigner matrices. *Annales de l'I.H.P. Probabilités et statistiques*, 48(1):107–133, 2012. doi: 10.1214/10-AIHP410. URL http://www.numdam.org/articles/10.1214/10-AIHP410/.

Celentano, M., Montanari, A., and Wu, Y. The estimation error of general first order methods. In *Conference on Learning Theory*, pp. 1078–1141. PMLR, 2020.

Chen, H.-B. and Xia, J. Hamilton-Jacobi equations for inference of matrix tensor products. *Ann. Inst. Henri Poincaré Probab. Stat.*, 58(2):755–793, 2022. ISSN 0246-0203. doi: 10.1214/21-aihp1183. URL https://doi.org/10.1214/21-aihp1183.

Chen, H.-B., Mourrat, J.-C., and Xia, J. Statistical inference of finite-rank tensors, 2021. URL https://arxiv.org/abs/2104.05360.

Deshpande, Y., Abbe, E., and Montanari, A. Asymptotic mutual information for the two-groups stochastic block model. *CoRR*, abs/1507.08685, 2015. URL http://arxiv.org/abs/1507.08685.

Donoho, D. L. and Johnstone, I. M. Adapting to unknown smoothness via wavelet shrinkage. *Journal of the american statistical association*, 90(432):1200–1224, 1995.

Donoho, D. L., Maleki, A., and Montanari, A. Message-passing algorithms for compressed sensing. *Proceedings of the National Academy of Sciences*, 106(45):18914–18919, 2009.

Féral, D. and Péché, S. The Largest Eigenvalue of Rank One Deformation of Large Wigner Matrices. *Communications in Mathematical Physics*, 272(1):185–228, May 2007. ISSN 1432-0916. doi: 10.1007/s00220-007-0209-3. URL https://doi.org/10.1007/s00220-007-0209-3.

Guionnet, A., Ko, J., Krzakala, F., and Zdeborová, L. Low-rank matrix estimation with inhomogeneous noise. *arXiv preprint arxiv:2208.05918*, 2022.

Guionnet, A., Ko, J., Krzakala, F., Mergny, P., and Zdeborová, L. Spectral Phase Transitions in Non-Linear Wigner Spiked Models, October 2023. URL http://arxiv.org/abs/2310.14055. arXiv:2310.14055 [math].

Horn, R. A. and Johnson, C. R. *Matrix analysis*. Cambridge University Press, Cambridge, second edition, 2013. ISBN 978-0-521-54823-6.

Huang, B. and Sellke, M. Algorithmic Threshold for Multi-Species Spherical Spin Glasses, September 2023a. URL http://arxiv.org/abs/2303.12172. arXiv:2303.12172 [cond-mat, physics:math-ph].

Huang, B. and Sellke, M. Strong Topological Trivialization of Multi-Species Spherical Spin Glasses, September 2023b. URL http://arxiv.org/abs/2308.09677. arXiv:2308.09677 [cond-mat, physics:math-ph].

Husson, J. Large deviations for the largest eigenvalue of matrices with variance profiles. *Electronic Journal of Probability*, 27(none):1–44, January 2022. ISSN 1083-6489, 1083-6489. doi: 10.1214/22-EJP793. Publisher: Institute of Mathematical Statistics and Bernoulli Society.

Johnson, C. R. *Matrix theory and applications*, volume 40. American Mathematical Soc., 1990.

Karrer, B. and Newman, M. E. Stochastic blockmodels and community structure in networks. *Physical review E*, 83 (1):016107, 2011.

Lee, S. and Lee, J. O. Phase transition for the generalized two-community stochastic block model. *Journal of Applied Probability*, pp. 1–16, July 2023. ISSN 0021-9002, 1475-6072.

Lelarge, M. and Miolane, L. Fundamental limits of symmetric low-rank matrix estimation. *Probab. Theory Related Fields*, 173(3-4):859–929, 2019. ISSN 0178-8051. doi: 10.1007/s00440-018-0845-x. URL https://doi.org/10.1007/s00440-018-0845-x.

Lesieur, T., Krzakala, F., and Zdeborová, L. Constrained low-rank matrix estimation: Phase transitions, approximate message passing and applications. *Journal of Statistical Mechanics: Theory and Experiment*, 2017(7): 073403, 2017.

Maillard, A., Krzakala, F., Lu, Y. M., and Zdeborová, L. Construction of optimal spectral methods in phase retrieval. In *Mathematical and Scientific Machine Learning*, pp. 693–720. PMLR, 2022.

Meyer, C. D. and Stewart, I. *Matrix analysis and applied linear algebra*. SIAM, 2023.

Mondelli, M. and Montanari, A. Fundamental limits of weak recovery with applications to phase retrieval. In *Conference On Learning Theory*, pp. 1445–1450. PMLR, 2018.

Mondelli, M. and Venkataramanan, R. Approximate message passing with spectral initialization for generalized linear models. In *International Conference on Artificial Intelligence and Statistics*, pp. 397–405. PMLR, 2021.

Mondelli, M., Thrampoulidis, C., and Venkataramanan, R. Optimal combination of linear and spectral estimators for generalized linear models. *Foundations of Computational Mathematics*, 22(5):1513–1566, 2022.

Pak, A., Ko, J., and Krzakala, F. Optimal algorithms for the inhomogeneous spiked Wigner model. *arXiv preprint arXiv:2302.06665*, 2023. Neurips 2023.

Paul, D. Asymptotics of Sample Eigenstructure for a Large Dimensional Spiked Covariance Model. *Statistica Sinica*, 17(4):1617–1642, 2007. ISSN 10170405, 19968507. URL http://www.jstor.org/stable/24307692. Publisher: Institute of Statistical Science, Academia Sinica.

Péché, S. The largest eigenvalue of small rank perturbations of Hermitian random matrices. *Probability Theory and Related Fields*, 134(1):127–173, August 2005. doi: 10.1007/s00440-005-0466-z. URL https://doi.org/10.1007/s00440-005-0466-z.

Péché, S. Deformed ensembles of random matrices. In *Proceedings of the International Congress of Mathematicians, Seoul*, volume 3, pp. 1059–1174, 2014.

Pizzo, A., Renfrew, D., and Soshnikov, A. On Finite Rank Deformations of Wigner Matrices. *Annales de l'I.H.P. Probabilités et statistiques*, 49(1):64–94, 2013. ISSN 1778-7017. doi: 10.1214/11-AIHP459. URL http://www.numdam.org/item/AIHPB_2013__49_1_64_0/.

Plemmons, R. J. M-matrix characterizations.I—nonsingular M-matrices. *Linear Algebra and its Applications*, 18(2):175–188, January 1977. ISSN 0024-3795. doi: 10.1016/0024-3795(77)90073-8. URL https://www.sciencedirect.com/science/article/pii/0024379577900738.

Rangan, S. Generalized approximate message passing for estimation with random linear mixing. In *2011 IEEE International Symposium on Information Theory Proceedings*, pp. 2168–2172. IEEE, 2011.

Rangan, S. and Fletcher, A. K. Iterative estimation of constrained rank-one matrices in noise. In *2012 IEEE International Symposium on Information Theory Proceedings*, pp. 1246–1250. IEEE, 2012.

Renfrew, D. and Soshnikov, A. On Finite Rank Deformations of Wigner Matrices II: Delocalized Perturbations. *Random Matrices: Theory and Applications*, 02(01):1250015, 2013. doi: 10.1142/S2010326312500153. URL https://doi.org/10.1142/S2010326312500153.

Rossetti, R. and Reeves, G. Approximate message passing for the matrix tensor product model, 2023.

Saad, Y. *Iterative methods for sparse linear systems*. SIAM, 2003.

Saade, A., Krzakala, F., and Zdeborová, L. Spectral Clustering of Graphs with the Bethe Hessian. *Advances in Neural Information Processing Systems*, 27, 2014.

Tao, T. *Topics in Random Matrix Theory*. Graduate studies in mathematics. American Mathematical Soc., 2020. ISBN 9780821885079.

Venkataramanan, R., Kögler, K., and Mondelli, M. Estimation in rotationally invariant generalized linear models via approximate message passing. In *International Conference on Machine Learning*, pp. 22120–22144. PMLR, 2022.

Zdeborová, L. and Krzakala, F. Statistical physics of inference: Thresholds and algorithms. *Advances in Physics*, 65(5):453–552, 2016.

A. Sketch of Proof of the Derivation of the Quadratic Vector Equation

We briefly sketch the proof for the derivation of the Prop. 3.1 and refer the reader interested in the complete rigorous proof to the series of work (Ajanki et al., 2017a;b; 2019). For simplicity, we w

Let $z \in \mathbb{H}_-$ and denote by $\mathbf{G}_{\boldsymbol{X}/\sqrt{N}}(z) := (z - \boldsymbol{X}/\sqrt{N})^{-1}$ the resolvent of $\boldsymbol{X}/\sqrt{N}$. From Schur complement formula, we have for any $i \in [\![N]\!] := \{1, \dots, N\}$:

$$\frac{1}{G_{ii}(z)} = z - X_{ii} - \langle \vec{\boldsymbol{X}}_{\bullet i}, \mathbf{G}^{(-i)}(z)\vec{\boldsymbol{X}}_{\bullet i}\rangle \,, \tag{35}$$

where to ease notations we denoted by $G_{ij} \equiv (\mathbf{G}_{\boldsymbol{X}/\sqrt{N}})_{ij}$, $X_{ij} \equiv (\boldsymbol{X}/\sqrt{N})_{ij}$, $\vec{X}_{\bullet i} \in \mathbb{R}^{N-1}$ is the i-th column of $\boldsymbol{X}/\sqrt{N}$ without the (i,i) entry and $\mathbf{G}^{(-i)}(z)$ is the resolvent of the $\big((N-1)\times(N-1)\big)$ matrix $\boldsymbol{X}^{-(i)}/\sqrt{N}$ obtained by removing the i-th row and i-th column of $\boldsymbol{X}/\sqrt{N}$.

By definition of $\boldsymbol{X}/\sqrt{N}$, we have $X_{ii} = H_{ii}\sqrt{(\boldsymbol{\Sigma})_{ii}}/\sqrt{N} - (\boldsymbol{\Sigma}\mathbf{1})_{ii}/N$, with $H_{ii} \sim \mathcal{N}(0,1)$. Since the first term is of order $\mathcal{O}(N^{-1/2})$, for large N we have approximately

$$X_{ii} \approx -\frac{1}{N}\sum_{j=1}^{N}(\boldsymbol{\Sigma})_{ij} = -\frac{1}{N}\sum_{k=1}^{K}\sum_{j\in B_k}(\boldsymbol{\Sigma})_{ij} \,, \tag{36}$$

and by assumption $(\boldsymbol{\Sigma})_{ij} = s_{kl}$ for any $i \in B_k, j \in B_l$, such that for any k, the X_{ii} within B_k are (approximately) equal and given by:

$$X_{ii} \approx -\sum_{l} s_{kl}\rho_l \,. \tag{37}$$

The quadratic term in Eq. (35) is given by $\langle \vec{\boldsymbol{X}}_{\bullet i}, \mathbf{G}^{(-i)}(z)\vec{\boldsymbol{X}}_{\bullet i}\rangle = \sum_{j_1,j_2\neq i} X_{j_1,i}X_{j_2,i}(\mathbf{G}^{(-i)}(z))_{ij}$ and can be shown to concentrate to its average value. Since $X_{j_1,i}$ and $X_{j_2,i}$ are independent and centered (since $j_1, j_2 \neq i$), this averaged sum is equal to the sum over the $(N-1)$ term such that $j_1 = j_2 \neq i$. Furthermore, for large N, replacing $\mathbf{G}^{(-i)}(z)$ by $\mathbf{G}_{\boldsymbol{X}/\sqrt{N}}(z)$ and adding the term $j_1 = i$ in the sum only adds a negligible contribution to the sum such that, one has approximately:

$$\langle \vec{\boldsymbol{X}}_{\bullet i}, \mathbf{G}^{(-i)}(z)\vec{\boldsymbol{X}}_{\bullet i}\rangle \approx \frac{1}{N}\sum_{j=1}^{N}(\boldsymbol{\Sigma})_{ij}G_{jj}(z) = \frac{1}{N}\sum_{k=1}^{K}\sum_{j\in B_k}(\boldsymbol{\Sigma})_{ij}G_{jj}(z) \,, \tag{38}$$

and Eq. (35) is thus approximately given by

$$\frac{1}{G_{ii}(z)} \approx z + \sum_{l=1}^{K} s_{kl}\rho_l - \frac{1}{N}\sum_{j=1}^{N}(\boldsymbol{\Sigma})_{ij}G_{jj}(z) \,. \tag{39}$$

Using again the block structure of $\boldsymbol{\Sigma}$, one deduces from Eq. (39) the approximate identity $G_{ii}(z) \approx G_{jj}(z)$ if $i,j \in B_k$. Furthermore, $N^{-1}\mathrm{Tr}\,\mathbf{G}_{\boldsymbol{X}/\sqrt{N}}(z) = N^{-1}\sum_{i=1}^{N} G_{ii}(z) \to g_{\mu_X}(z)$, and if we denote by $g_k(z) := \lim_{N\to\infty} 1/|B_k|\sum_{i\in B_k} G_{ii}(z) \approx G_{ii}(z)$ for $i \in B_k$, we have $g_{\mu_X} = \sum_k \rho_k g_k$ with for any $k \in [\![K]\!]$, g_k is solution of

$$\frac{1}{g_k} = z + \sum_{l=1}^{K} s_{kl}\rho_l - \sum_{l=1}^{K} s_{kl}\rho_l g_l \,, \tag{40}$$

which is nothing else than the scalar version of the QVE of Prop. 3.1.

B. Proof of Prop. 3.3 (Stietljes transform on the Real Line and Positivity Condition for Solution of Associated Linear System)

For $z \in \mathbb{H}_-$ let's rewrite the QVE of Eq. (13) with solution $\mathbf{g} \in (\mathbb{H}_+)^K$ as:

$$z\mathbf{1}_K = \mathbf{f}(\mathbf{g}(z)) \,, \tag{41}$$

where $\mathbf{f} : (\mathbb{H}_+)^K \to (\mathbb{H}_-)^K$ is defined by

$$\mathbf{f}(\mathbf{g}) := \mathbf{g}^{\odot -1} + \mathbf{\Gamma}_K(\mathbf{g} - \mathbf{1}_K)\,. \tag{42}$$

Clearly, the function $\mathbf{f}(.)$ is continuous and in fact C^1 and invertible on $(\mathbb{H}_+)^K$ with Jacobian given by

$$(\nabla_{\mathbf{g}}\mathbf{f})(\mathbf{g}) = -\mathbf{D}_{\mathbf{g}}^{-2} + \mathbf{\Gamma}_K\,, \tag{43}$$

and the $g_k(z) := \int (z-\lambda')^{-1}\mathrm{d}\mu_k(\lambda')$ are analytical on $\mathbb{C} \setminus \overline{\mathrm{Supp}(\mu_X)}$ for any $k \in [\![K]\!]$ since $\mathrm{Supp}(\mu_k) \subseteq \mathrm{Supp}(\mu_X)$ for any $k \in [\![K]\!]$, from (Ajanki et al., 2019).

Fix any $\lambda \in \mathbb{R} \setminus \overline{\mathrm{Supp}(\mu_X)}$, as we approach λ from the lower complex plane $z \to \lambda$, $\mathbf{g}(z)$ approaches its analytical continuation $\mathbf{g}(\lambda)$ from the upper complex plane, which means that $\mathbf{g}'(\lambda)$ is well defined and belongs also to $(\mathbb{H}_+)^K$. By the inverse function theorem, if we set $\tilde{\mathbf{y}} = \mathbf{g}'(\lambda)$, we must have the equality:

$$-\mathbf{1}_K = (\nabla_{\mathbf{g}}\mathbf{f})(\mathbf{g}(\lambda))\tilde{\mathbf{y}}\,, \tag{44}$$

with the condition $\tilde{\mathbf{y}} \in (\mathbb{H}_+)^K$. Taking the imaginary part of this equation gives the desired result by setting $y_i = \mathfrak{Im}(\tilde{y}_i)$ since for $\lambda \in \mathbb{R} \setminus \overline{\mathrm{Supp}(\mu_X)}$, $\mathbf{g}(\lambda) \in \mathbb{R}^K$ and hence we also have $(\nabla_{\mathbf{g}}\mathbf{f})(\mathbf{g}(\lambda)) \in \mathbb{R}^{K\times K}$. Conversely, for any other solutions $\tilde{\mathbf{g}} \in (\mathbb{R}_*)^K$ of the QVE at $z = \lambda$, one cannot have $\mathbf{y} \succeq \mathbf{0}_K$ since $\mathbf{f}$ and $(\nabla_{\mathbf{g}}\mathbf{f})$ are well defined on $(\mathbb{C}_*)^K$ and $\mathbf{g} \in (\mathbb{H}_+)^K$ is the unique solution of the QVE on the upper complex plane.

C. Proof of Lem. 3.4 (Eigenvalue Condition for Positive Solution of the Linear System)

By multiplying Eq.(15) by $\mathbf{D}_{\tilde{\mathbf{g}}}^2$, we can re-write Eq. (15) as

$$(\mathbf{I}_K - \mathbf{D}_{\tilde{\mathbf{g}}}^2\mathbf{\Gamma}_K)\mathbf{y} = \tilde{\mathbf{g}}^{\odot 2}\,, \tag{45}$$

since $\mathbf{D}_{\tilde{\mathbf{g}}}^2\mathbf{\Gamma}_K$ and $\mathbf{D}_{\tilde{\mathbf{g}}}^2\mathbf{\Omega}_K$ are similar by definition of $\mathbf{\Gamma}_K$, we have $\lambda_1(\mathbf{D}_{\tilde{\mathbf{g}}}^2\mathbf{\Gamma}_K) = \lambda_1(\mathbf{D}_{\tilde{\mathbf{g}}}^2\mathbf{\Omega}_K)$. To conclude, it is enough to prove the following result.

Proposition C.1. *Let $\mathbf{Q} \in (\mathbb{R}_+)^{d\times d}$ with real eigenvalues, $\mathbf{r} \succ \mathbf{0} \in \mathbb{R}^d$ then the linear system $(\mathbf{I} - \mathbf{Q})\mathbf{z} = \mathbf{r}$ admits a unique positive solution $\mathbf{z} \succ \mathbf{0}$ if and only all eigenvalues of $\mathbf{Q}$ are lower than one.*

To prove this proposition, following (Plemmons, 1977) and Chap. 6 of (Berman & Plemmons, 1994) (see also Chap. 1 of (Saad, 2003)), we first introduce the two sets of matrix as:

Definition C.2 (Z- and M-matrices)**.** We say that a square matrix $\mathbf{A}$ is a Z-matrix and write $\mathbf{A} \in \mathfrak{Z}$ if all its off-diagonal elements are non-positive. If in addition, all its eigenvalues are positive we say that $\mathbf{A}$ is M-matrix and write $\mathbf{A} \in \mathfrak{M}$.

- Let's first consider the case $\lambda_1(\mathbf{Q}) < 1$. In this case, one can immediately check that $\mathbf{M} := \mathbf{I}_K - \mathbf{Q} \in \mathfrak{M}$ and one has the following lemma for M-matrices.

  Lemma C.3 (Inverse Positivity of M-matrices)**.** *If $\mathbf{A} \in \mathfrak{M}$ and is non-singular then $\mathbf{A}^{-1}$ has all it entries non-negative.*

  – *Proof of Lem. C.3.* See Chap. 6 of (Berman & Plemmons, 1994) for the complete proof. We briefly outline the main idea for completeness. If we denote by $\mathbf{a} := (A_{11}, \dots, A_{dd})$ the vector of diagonal elements, we have $\mathbf{a} \succ \mathbf{0}$. We can then write $\mathbf{A} = \mathbf{D}_{\mathbf{a}}(\mathbf{I} - \mathbf{B})$ with $\mathbf{B} = (\mathbf{I} - \mathbf{D}_{\mathbf{a}}^{-1}\mathbf{A})$ and since $(\mathbf{I} - \mathbf{B})^{-1} = \mathbf{A}^{-1}\mathbf{D}_{\mathbf{a}}$ can be shown to be entrywise non-negative and $\mathbf{D}_{\mathbf{a}}^{-1}$ is positive, we get the desired result for $\mathbf{A}^{-1}$. □

  As a consequence of this lemma, we have with the notations of Prop. C.1, $\mathbf{y} = \mathbf{M}^{-1}\mathbf{r} \succ \mathbf{0}_K$.

- Next for $\lambda_1(\mathbf{Q}) > 1$, the same matrix $\mathbf{M} := \mathbf{I}_K - \mathbf{Q} \in \mathfrak{Z} \setminus \mathfrak{M}$ (since it has at least one negative eigenvalue). The Perron-Frobenius theorem for Z-matrices writes

  Lemma C.4 (Perron-Frobenius Z-matrices)**.** *If $\mathbf{A} \in \mathfrak{Z}$, there exists a real number σ and a positive vector $\mathbf{u} \succ \mathbf{0}$ such that*

  (a) $\mathbf{A}^\top\mathbf{u} = \sigma\mathbf{u}$,

(b) $\mathfrak{Re}(\lambda) > \sigma$ *if* λ *is any characteristic root of* $\mathbf{A}$ *with* $\lambda \neq \sigma$.

– *Proof of Lem. C.4.* see Chap. 6 of (Berman & Plemmons, 1994). The proof is similar to the standard Perron-Frobenius theorem for entrywise positive matrix. □

As $1 - \lambda_1(\mathbf{Q})$ is the lowest eigenvalue of $\mathbf{M}^\top$, this Perron-Frobenius lemma implies that the corresponding eigenvector is positive $\mathbf{u} \succ \mathbf{0}$. In particular, we have $\langle -\mathbf{u}, \mathbf{1} \rangle < 0$ and $\mathbf{M}^\top(-\mathbf{u}) = (\lambda_1(\mathbf{Q}) - 1)\mathbf{u} \succeq \mathbf{0}$, which allows us to conclude thanks to Farka's lemma:

Lemma C.5 (Farka's lemma)**.** *Let* $\mathbf{A} \in \mathbb{R}^{d \times r}$, $\mathbf{b} \in \mathbb{R}^r$ *then either*

(a) there exists $\mathbf{y} \in \mathbb{R}^r$ *solution of* $\mathbf{A}\mathbf{y} = \mathbf{b}$ *with* $\mathbf{y} \succeq \mathbf{0}$,

(b) or there exists $\mathbf{z} \in \mathbb{R}^d$ *such that* $\mathbf{A}^\top \mathbf{z} \succeq \mathbf{0}$ *and* $\langle \mathbf{z}, \mathbf{b} \rangle < 0$.

– *Proof of Farka's lemma.* see Chap. 5 of (Boyd & Vandenberghe, 2004) where the proof follows from linear programming duality. □

Indeed, we have explicitly constructed $\mathbf{z} = -\mathbf{u}$ such that condition (b) holds for the linear system $(\mathbf{I}_K - \mathbf{Q})\mathbf{y} = \mathbf{r}$ and as such, it does not admit a positive solution $\mathbf{y}$ for $\lambda_1(\mathbf{Q}) > 1$.

This concludes the proof of Prop. C.1 since if $\lambda_1(\mathbf{Q}) = 1$ then the matrix $\mathbf{M} = \mathbf{I} - \mathbf{Q}$ is singular and one has no more uniqueness for the possible solution of the associated linear system, see e.g. Chap. 1 of (Saad, 2003).

D. Proof of Prop. 3.5 (Upper Bound of the Rightmost Edge)

Proposition D.1 (Condition for being an edge)**.** *If* $\lambda_\star$ *is the position of an edge then*

$$\lambda_\star \mathbf{1}_K = \mathbf{f}(\mathbf{g}_\star)\,, \tag{46}$$

for some $\mathbf{g}_\star \in (\mathbb{R}_*)^K$ *such that* $(\nabla_{\mathbf{g}}\mathbf{f})(\mathbf{g}_\star)$ *is singular.*

Proof. This follows from Thm. 2.6 of (Ajanki et al., 2017a). □

Lemma D.2. *Let* $\mathbf{g} \in (\mathbb{R}_*)^K$, *then* $(\nabla_{\mathbf{g}}\mathbf{f})(\mathbf{g})$ *is singular if there exists* $\mathbf{w} \succ \mathbf{0}_K$ *such that*

$$\mathbf{w}^\top(\mathbf{D}^2_{\mathbf{g}}\mathbf{\Gamma}_K) = \mathbf{w}^\top\,. \tag{47}$$

Proof. First since $\mathbf{g} \in (\mathbb{R}_*)^K$, $(\nabla_{\mathbf{g}}\mathbf{f})(\mathbf{g})$ is singular if $\mathbf{D}^2_{\mathbf{g}}(\nabla_{\mathbf{g}}\mathbf{f})(\mathbf{g})$ is, which means from definition (43) that there exists a $\mathbf{w} \neq \mathbf{0}_K$ such that Eq. (47) holds. Recalling Def. C.2, we have $\mathbf{M} = \mathbf{I}_K - (\mathbf{D}^2_{\mathbf{g}}\mathbf{\Gamma}_K) \in \mathfrak{Z}$ from which we deduce the positivity of $\mathbf{w}$ by Lem. C.4. □

For $\lambda_\star$ the position of any edge, let's denote by $\mathbf{w}_\star$ the corresponding left eigenvector in Lem. D.2. If we left multiply Eq. (46) by $\mathbf{w}_\star^\top \mathbf{D}^2_{\mathbf{g}_\star}$ we have

$$\lambda_\star \langle \mathbf{w}_\star, \mathbf{g}_\star^{\odot 2} \rangle = \langle \mathbf{w}_\star, \mathbf{g}_\star \rangle + \langle \mathbf{w}_\star, (\mathbf{D}^2_{\mathbf{g}_\star}\mathbf{\Gamma}_K)(\mathbf{g}_\star - \mathbf{1}_K) \rangle\,. \tag{48}$$

Using Eq. (47) this reads

$$\lambda_\star = \frac{\langle \mathbf{w}_\star, 2\mathbf{g}_\star - \mathbf{1}_K \rangle}{\langle \mathbf{w}_\star, \mathbf{g}_\star^{\odot 2} \rangle}\,, \tag{49}$$

and one gets the desired result thanks to $\mathbf{w}_\star \succ \mathbf{0}_K$ and the identity

$$\mathbf{1}_K + \mathbf{g}_\star^{\odot 2} - 2\mathbf{g}_\star = (\mathbf{1}_K - \mathbf{g}_\star)^{\odot 2} \succeq \mathbf{0}_K\,, \tag{50}$$

with equality if and only if $\mathbf{g}_\star = \mathbf{1}_K$ which is only attainable at $\lambda_1(\mathbf{\Omega}_K) = 1$ since $\mathbf{g}(1) = \mathbf{1}_K$ for $\lambda_1(\mathbf{\Omega}_K) \leq 1$ from Cor. 3.7 and $\mathbf{D}^2_{\mathbf{g}}\,(\nabla_{\mathbf{g}}\mathbf{f})(\mathbf{1}_K)$ is not singular for $\lambda_1(\mathbf{\Omega}_K) < 1$ from App. C, and is clearly singular for $\lambda_1(\mathbf{\Omega}_K) = 1$.

E. Proof of Prop. 3.7 (Behavior of the Stieltjes Transform at $z = 1$)

We provide the details of the QVE analysis required to complete the proof of Prop. 3.7 part (b)-(i). We want to prove the following claim

Proposition E.1. *If $\lambda_1(\mathbf{\Omega}_K) > 1$ then the correct solution of the QVE satisfies $\mathbf{0}_K \prec \mathbf{g}(1) \prec \mathbf{1}_K$.*

Note that the first inequality is trivial and the hard part is proving the second inequality from which we only know the condition $\mathbf{g}(1) \neq \mathbf{1}_K$ so far. Recall that the QVE at $z = 1$ can be written as:

$$\mathbf{0}_K = (\mathbf{1}_K - \mathbf{g}(1)) + \mathbf{g}(1) \odot \mathbf{\Gamma}_K(\mathbf{1}_K - \mathbf{g}(1)) \tag{51}$$

Let's introduce $\mathbf{y} = \mathbf{1}_K - \mathbf{g}(1)$, the condition we would like to prove is simply transferred to $\mathbf{y} \succ \mathbf{0}_K$. We can write the QVE for $\mathbf{y}$ as:

$$\mathbf{0}_K = \mathbf{y} + (\mathbf{1}_K - \mathbf{y}) \odot \mathbf{\Gamma}_K\, \mathbf{y} = \mathbf{y} + \mathbf{\Gamma}_K\, \mathbf{y} - \mathbf{y} \odot \mathbf{\Gamma}_K \mathbf{y} \tag{52}$$

We will analyze solutions to (52) as a function of the gap between the $\lambda_1(\mathbf{\Omega}_K)$ and 1. We have the following trivial fact

Lemma E.2. *If we denote by $\alpha = \lambda_1(\mathbf{\Omega}_K) - 1 > 0$, then we can write $\mathbf{\Omega}_K = (1 + \alpha) \cdot \mathbf{\Omega}_K^{(1)}$ such that $\lambda_1(\mathbf{\Omega}_K^{(1)}) = 1$.*

Proposition E.1 is immediate from the following uniform bound on the Stieltjes transform, which implies that $\mathbf{g}(\alpha, 1) \prec \mathbf{1}_K$ uniformly for all $\mathbf{\Omega}_K^{(1)}$ and $\alpha > 0$.

Lemma E.3. *Let $\mathbf{\Omega}_K^{(1)}$ be a fixed parameter matrix such that $\lambda_1(\mathbf{\Omega}_K^{(1)}) = 1$ and consider the QVE (52) with $\mathbf{\Omega}_K(\alpha) = (1+\alpha) \cdot \mathbf{\Omega}_K^{(1)}$. For any $\alpha > 0$, the proper solution $\mathbf{y}(\alpha) = \mathbf{1}_K - \mathbf{g}(\alpha, 1)$ of (52) satisfying (14) at $z = 1$ satisfies $\mathbf{y}(\alpha) \succ \mathbf{0}$.*

Proof. We need to show that $\mathbf{y}(\alpha) \succ \mathbf{0}$ for all α. If we write (52) in terms of α, we get

$$\mathbf{0}_K = \mathbf{y} + (1 + \alpha)(\mathbf{1} - \mathbf{y}) \odot \mathbf{\Gamma}_K^{(1)}\, \mathbf{y} \tag{53}$$

if we differentiate for α, we have

$$\mathbf{0}_K = \mathbf{y'} + (\mathbf{1} - \mathbf{y}) \odot \mathbf{\Gamma}_K^{(1)}\, \mathbf{y} - (1 + \alpha)\big(\mathbf{y'} \odot \mathbf{\Gamma}_K^{(1)}\, \mathbf{y} + (\mathbf{1} - \mathbf{y}) \odot \mathbf{\Gamma}_K^{(1)}\, \mathbf{y'}\big) \tag{54}$$

but using the previous equation, this simplifies to

$$\mathbf{0}_K = \mathbf{y'} - \frac{\mathbf{y}}{1 + \alpha} - (1 + \alpha)\big(\mathbf{y'} \odot \mathbf{\Gamma}_K^{(1)}\, \mathbf{y} + (\mathbf{1} - \mathbf{y}) \odot \mathbf{\Gamma}_K^{(1)}\, \mathbf{y'}\big) \tag{55}$$

Note that when $\alpha = 0$ we have $\mathbf{y}(0) = \mathbf{0}$ such that we end up with

$$\mathbf{\Gamma}_K^{(1)}\, \mathbf{y'}(0) = \mathbf{y'}(0) \tag{56}$$

which means that $\mathbf{y'}(0)$ is a multiple of the top eigenvector of $\mathbf{\Gamma}_K^{(1)}$ and since $\mathbf{\Gamma}_K^{(1)}$ is positive matrix, this eigenvector has positive entries by Perron Frobenius theorem from which we deduce that the entries of $\mathbf{y'}(0)$ have the same sign.

We claim that $\mathbf{y'}(0)$ must be non-negative. Suppose for the sake of contradiction that $\mathbf{y'}(0) \prec \mathbf{0}$, so that $\mathbf{y}(\alpha) \prec \mathbf{0}$ for some α sufficiently small. In particular, we assume that $\mathbf{g}(\alpha, 1) \succ \mathbf{1}$. Using the lower bound for congruent matrices (see Theo. 4.5.9 of (Horn & Johnson, 2013)), we have that

$$\lambda_1(\mathbf{D}_{\mathbf{g}(\alpha,1)^2} \mathbf{\Omega}_K) \geq \lambda_K(\mathbf{D}_{\mathbf{g}(\alpha,1)^2}) \lambda_1(\mathbf{\Omega}_K) > 1 \tag{57}$$

since $\lambda_1(\mathbf{\Omega}_K) > 1$ and all eigenvalues of $\mathbf{D}_{\mathbf{g}(\alpha,1)^2}$ are larger than 1. However, this contradicts the consistency equation in Prop. 3.4 which states that $\lambda_1(\mathbf{D}_{\mathbf{g}(\alpha,1)^2} \mathbf{\Omega}_K) \leq 1$, so $\mathbf{y'}(0)$ must be non-negative.

Furthermore, by Cor. 3.7 we have that $\mathbf{y} = \mathbf{1}_K - \mathbf{g}(1)$ is non-zero for all $\alpha > 0$, so $\mathbf{y}(0)$ cannot be the zero vector. This implies that $\mathbf{y}(\alpha) \succ \mathbf{0}$ for all α in a neighbourhood of zero. We will now show that $\mathbf{y}(\alpha) \succeq \mathbf{0}$ for all $\alpha \geq 0$. Let α^* denote the first crossing time for $\mathbf{y}(\alpha)$,

$$\alpha^* = \inf\{\alpha \geq 0 \mid \mathbf{y}(\alpha) \not\succeq \mathbf{0}\}. \tag{58}$$

Suppose for the sake of contradiction that $\alpha^* < \infty$. Since $\mathbf{y}(\alpha)$ is continuous, we must have that $\mathbf{y}(\alpha^*)_k = 0$ for some k and $\mathbf{y}(\alpha^*)_{k'} \geq 0$ for $k' \neq k$. However, examining the QVE for the kth entry of $y(\alpha^*)$ in (53) implies that

$$0 = (1+\alpha^*)(\mathbf{\Gamma}_K^{(1)}\,\mathbf{y})_k \tag{59}$$

so the fact that $(1+\alpha^*)\mathbf{\Gamma}_K^{(1)}$ has positive entries and $\mathbf{y}(\alpha^*) \succeq \mathbf{0}$ implies that $\mathbf{y}(\alpha^*) = \mathbf{0}$. However, by Cor. 3.7 we have that $\mathbf{y} = \mathbf{1}_K - \mathbf{g}(1)$ is non-zero for all $\alpha > 0$ which contradicts the fact that $\mathbf{y}(\alpha^*) = \mathbf{0}$. We can conclude that $\mathbf{y} \succ \mathbf{0}$ for all α as required, since the case that $\mathbf{y} \equiv \mathbf{0}$ is ruled out by Cor. 3.7. □

F. Proof of Prop. 3.9 (Equation for Outliers)

Let λ not in the spectrum of $\boldsymbol{X}/\sqrt{N}$ such that $\mathbf{G}_{\boldsymbol{X}/\sqrt{N}}(\lambda) := (\lambda - \boldsymbol{X}/\sqrt{N})^{-1}$ is well defined, then from the decomposition of Eq. (8), we have the identity

$$\Big(\lambda - \frac{\tilde{\boldsymbol{Y}}}{\sqrt{N}}\Big) = \Big(\lambda - \frac{\boldsymbol{X}}{\sqrt{N}}\Big)\Big(\mathbf{I} - \mathbf{G}_{\frac{\boldsymbol{X}}{\sqrt{N}}}(\lambda)\mathbf{Z}\Big)\,. \tag{60}$$

Taking the determinant of this equation with the eigendecomposition of Prop. 3.8 and cyclicity of the determinant, we get that the characteristic polynomial of $\tilde{\boldsymbol{Y}}/\sqrt{N}$ is given by:

$$\det\Big(z - \frac{\tilde{\boldsymbol{Y}}}{\sqrt{N}}\Big) = \det\Big(z - \frac{\boldsymbol{X}}{\sqrt{N}}\Big)\Big(\mathbf{I}_K - \mathbf{V}^\top \mathbf{G}_{\frac{\boldsymbol{X}}{\sqrt{N}}}(\lambda)\mathbf{V}\tilde{\mathbf{\Omega}}_K\Big)\,, \tag{61}$$

where we have a denoted by $\tilde{\mathbf{\Omega}}_K := \mathbf{\Omega}_K + \boldsymbol{E}_K$. Following (Benaych-Georges & Nadakuditi, 2011; 2012) as $N \to \infty$, an outlier of $\tilde{\boldsymbol{Y}}/\sqrt{N}$ separates from the bulk if it is a zero of the following limit of the secular function

$$h(\lambda) := \lim_{N\to\infty} \det\Big(\mathbf{I}_K - \mathbf{V}^\top \mathbf{G}_{\frac{\boldsymbol{X}}{\sqrt{N}}}(\lambda)\mathbf{V}\tilde{\mathbf{\Omega}}_K\Big)\,, \tag{62}$$

provided this limit is well-defined.

To tackle this limit, we will use deterministic equivalents for the resolvent.

Lemma F.1 (Anisotropic Deterministic Equivalent)**.** *Let $z \in \mathbb{C}\backslash\mathrm{Supp}(\mu_X)$, $\mathbf{w}_1, \mathbf{w}_2$ be two sequences of vectors independent of $\boldsymbol{X}$ such that for $i = 1, 2$, $\|\mathbf{w}_i\| \xrightarrow[N\to\infty]{\text{a.s}} C_i > 0$, denote by $\mathbf{P}$ the permutation matrix that ordered the groups $B_1, \dots, B_k$, and introduce $\overline{\mathbf{G}(z)}$ the $(N \times N)$ diagonal matrix*

$$\overline{\mathbf{G}(z)} := \mathbf{P}\begin{pmatrix} g_1(z)\mathbf{I}_{|B_1|} & & \\ & \ddots & \\ & & g_K(z)\mathbf{I}_{|B_K|}\end{pmatrix}\mathbf{P}^\top\,, \tag{63}$$

where the $g_k(.)$ are the solutions of the QVE of Prop. 3.1, then we have

$$\Big|\langle \mathbf{w}_1, \mathbf{G}_{\frac{\boldsymbol{X}}{\sqrt{N}}}(z)\mathbf{w}_2\rangle - \langle \mathbf{w}_1, \overline{\mathbf{G}(z)}\mathbf{w}_2\rangle\Big| \xrightarrow[N\to\infty]{\text{a.s}} 0. \tag{64}$$

Proof. This is a direct consequence of the anisotropic local law of Thm. 1.13 in (Ajanki et al., 2017b). □

Since $(\mathbf{V}^\top \mathbf{G}_{\boldsymbol{X}/\sqrt{N}}(\lambda)\mathbf{V})_{kl} = \langle \mathbf{v}^{(k)}, \mathbf{G}_{\boldsymbol{X}/\sqrt{N}}(\lambda)\mathbf{v}^{(l)}\rangle$, the use of Lemma F.1 with the block-structure of the $\mathbf{v}^{(k)}$ from Prop.3.8 implies

$$\mathbf{V}^\top \mathbf{G}_{\boldsymbol{X}/\sqrt{N}}(\lambda)\mathbf{V} \xrightarrow[N\to\infty]{\text{a.s}} \mathbf{D}_{\mathbf{g}(\lambda)}\,, \tag{65}$$

and since also $\tilde{\mathbf{\Omega}}_K \xrightarrow[N\to\infty]{\text{a.s}} \mathbf{\Omega}_K$ both in operator norm and entrywise by Prop. 3.8, by continuity of the determinant this leads to

$$h(\lambda) \overset{\text{a.s}}{=} \det\left(\mathbf{I}_K - \mathbf{D}_{\mathbf{g}(\lambda)}\,\mathbf{\Omega}_K\right)\,, \tag{66}$$

which concludes the proof.

G. Proof of Prop. 3.12 (The Overlap Vector is Proportional to the Right Top Eigenvector $\mathbf{v}_1^{(r)}$)

To ease notations, let's denote by $\check{\boldsymbol{Y}} = \tilde{\boldsymbol{Y}}/\sqrt{N}$ and $\check{\boldsymbol{X}} = \tilde{\boldsymbol{X}}/\sqrt{N}$ and $\lambda_i \equiv \lambda_i\Big(\frac{\tilde{\boldsymbol{Y}}}{\sqrt{N}}\Big)$, $\boldsymbol{u}_i$ are respectively the ordered eigenvalues and the associated eigenvectors of $\check{\boldsymbol{Y}}$. Thus, we have

$$\mathbf{V}^\top \mathbf{G}_{\check{\boldsymbol{Y}}}(z)\mathbf{V} = \sum_i^N (z-\lambda_i)^{-1}\mathbf{V}^\top \boldsymbol{u}_i \left(\mathbf{V}^\top \boldsymbol{u}_i\right)^\top \tag{67}$$

Thus if λ_1 is simple we have

$$\mathbf{q}\mathbf{q}^\top = \lim_{z\to\lambda_1} \left\{(z-\lambda_1)\cdot \mathbf{V}^\top \mathbf{G}_{\check{\boldsymbol{Y}}}(z)\mathbf{V}\right\} \tag{68}$$

Lemma G.1 (Resolvent Identities)**.** *With $\mathbf{G}_{\check{\boldsymbol{Y}}}(z) = (z\mathbf{I}-\check{\boldsymbol{Y}})^{-1}$ and $\mathbf{G}_{\check{\boldsymbol{X}}}(z) = (z\mathbf{I}-\check{\boldsymbol{X}})^{-1}$, we have for any $z \notin (\mathrm{Spec}(\check{\boldsymbol{Y}}) \cup \mathrm{Spec}(\check{\boldsymbol{X}}))$:*

$$\mathbf{G}_{\check{\boldsymbol{Y}}}(z) = \mathbf{G}_{\check{\boldsymbol{X}}}(z) + \mathbf{G}_{\check{\boldsymbol{X}}}(z)\mathbf{V}\boldsymbol{\Omega}_K\Big(\mathbf{I}-\mathbf{V}^\top\mathbf{G}_{\check{\boldsymbol{X}}}(z)\mathbf{V}\Big)^{-1}\mathbf{V}^\top\mathbf{G}_{\check{\boldsymbol{X}}}(z) \tag{69}$$

Proof. This is simply a consequence of the Woodbury matrix inequality. □

If we denote by $\tilde{\mathbf{G}} = \mathbf{V}^\top\mathbf{G}_{\check{\boldsymbol{X}}}(z)\mathbf{V}$, we get

$$\mathbf{V}^\top\mathbf{G}_{\check{\boldsymbol{Y}}}(z)\mathbf{V} = \tilde{\mathbf{G}} + \tilde{\mathbf{G}}\boldsymbol{\Omega}_K\Big(\mathbf{I}-\tilde{\mathbf{G}}\boldsymbol{\Omega}_K\Big)^{-1}\tilde{\mathbf{G}} \tag{70}$$

and so

$$\lim_{N\to\infty}\mathbf{q}\mathbf{q}^\top = \lim_{N\to\infty}\lim_{z\to\lambda_1}\left\{(z-\lambda_1)\cdot\Big(\tilde{\mathbf{G}}+\tilde{\mathbf{G}}\boldsymbol{\Omega}_K(\mathbf{I}-\tilde{\mathbf{G}}\boldsymbol{\Omega}_K)^{-1}\tilde{\mathbf{G}}\Big)\right\}, \tag{71}$$

$$\lim_{N\to\infty}\mathbf{q}\mathbf{q}^\top = \lim_{z\to 1}\left\{(z-1)\cdot\Big(\mathbf{D}_{\mathbf{g}(z)}+\mathbf{D}_{\mathbf{g}(z)}\boldsymbol{\Omega}_K(\mathbf{I}-\mathbf{D}_{\mathbf{g}(z)}\boldsymbol{\Omega}_K)^{-1}\mathbf{D}_{\mathbf{g}(z)}\Big)\right\}, \tag{72}$$

$$\lim_{N\to\infty}\mathbf{q}\mathbf{q}^\top = \mathbf{D}_{\mathbf{g}(1)}\boldsymbol{\Omega}_K\Big(\lim_{z\to 1}\Big\{(z-1)\cdot(\mathbf{I}-\mathbf{D}_{\mathbf{g}(z)}\boldsymbol{\Omega}_K)^{-1})\Big\}\Big)\mathbf{D}_{\mathbf{g}(1)}, \tag{73}$$

Without loss of generality, we write $z = 1+\epsilon$ with $\epsilon > 0$, such that $\mathbf{g}(1+\epsilon) \succ \mathbf{0}$ and thus

$$\mathbf{D}_{\mathbf{g}(1+\epsilon)}\boldsymbol{\Omega}_K = \mathbf{D}^{1/2}_{\mathbf{g}(1+\epsilon)}\cdot(\mathbf{D}^{1/2}_{\mathbf{g}(1+\epsilon)}\boldsymbol{\Omega}_K\mathbf{D}^{1/2}_{\mathbf{g}(1+\epsilon)})\cdot\mathbf{D}^{-1/2}_{\mathbf{g}(1+\epsilon)} \tag{74}$$

that is $\mathbf{D}_{\mathbf{g}(1+\epsilon)}\boldsymbol{\Omega}_K$ is similar to the symmetric matrix $\mathbf{D}^{1/2}_{\mathbf{g}(1+\epsilon)}\boldsymbol{\Omega}_K\mathbf{D}^{1/2}_{\mathbf{g}(1+\epsilon)}$ and we have

$$\lim_{\epsilon\to 0^+}\Big\{\epsilon\cdot(\mathbf{I}-\mathbf{D}_{\mathbf{g}(1+\epsilon)}\boldsymbol{\Omega}_K)^{-1}\Big\} = \mathbf{D}^{1/2}_{\mathbf{g}(1)}\cdot\lim_{\epsilon\to 0^+}\Big\{\epsilon\cdot(\mathbf{I}-\mathbf{D}^{1/2}_{\mathbf{g}(1+\epsilon)}\boldsymbol{\Omega}_K\mathbf{D}^{1/2}_{\mathbf{g}(1+\epsilon)})^{-1}\Big\}\cdot\mathbf{D}^{-1/2}_{\mathbf{g}(1)} \tag{75}$$

If we denote respectively by $\phi_i(1+\epsilon)$ and $\mathbf{w}_i(1+\epsilon)$, the ordered eigenvalues and corresponding orthonormal eigenvectors of the matrix $\mathbf{D}^{1/2}_{\mathbf{g}(1+\epsilon)}\boldsymbol{\Omega}_K\mathbf{D}^{1/2}_{\mathbf{g}(1+\epsilon)}$, we have

$$(\mathbf{I}-\mathbf{D}^{1/2}_{\mathbf{g}(1+\epsilon)}\boldsymbol{\Omega}_K\mathbf{D}^{1/2}_{\mathbf{g}(1+\epsilon)})^{-1} = \sum_{i=1}^K\frac{1}{1-\phi_i(1+\epsilon)}\mathbf{w}_i(1+\epsilon)\mathbf{w}_i^\top(1+\epsilon), \tag{76}$$

and since $\phi_1(1+\epsilon)\to 1$ and is simple, we get:

$$\lim_{\epsilon\to 0^+}\Big\{\epsilon\cdot(\mathbf{I}-\mathbf{D}_{\mathbf{g}(1+\epsilon)}\boldsymbol{\Omega}_K)^{-1}\Big\} = \mathbf{D}^{1/2}_{\mathbf{g}(1)}\cdot\lim_{\epsilon\to 0^+}\Big\{\sum_{i=1}^K\frac{\epsilon}{1-\phi_i(1+\epsilon)}\mathbf{w}_i(1+\epsilon)\mathbf{w}_i(1+\epsilon)^\top\Big\}\cdot\mathbf{D}^{-1/2}_{\mathbf{g}(1)}, \tag{77}$$

$$\lim_{\epsilon\to 0^+}\Big\{\epsilon\cdot(\mathbf{I}-\mathbf{D}_{\mathbf{g}(1+\epsilon)}\boldsymbol{\Omega}_K)^{-1}\Big\} = \mathbf{D}^{1/2}_{\mathbf{g}(1)}\cdot\lim_{\epsilon\to 0^+}\Big\{\frac{\epsilon}{1-\phi_i(1+\epsilon)}\mathbf{w}_1(1+\epsilon)\mathbf{w}_1(1+\epsilon)^\top\Big\}\cdot\mathbf{D}^{-1/2}_{\mathbf{g}(1)}, \tag{78}$$

$$\lim_{\epsilon\to 0^+}\Big\{\epsilon\cdot(\mathbf{I}-\mathbf{D}_{\mathbf{g}(1+\epsilon)}\boldsymbol{\Omega}_K)^{-1}\Big\} = \frac{-1}{\phi_1'(1)}\mathbf{D}^{1/2}_{\mathbf{g}(1)}\cdot\mathbf{w}_1(1)\mathbf{w}_1(1)^\top\cdot\mathbf{D}^{-1/2}_{\mathbf{g}(1)}, \tag{79}$$

thus we get

$$\lim_{N\to\infty} \mathbf{q}\mathbf{q}^\top = \frac{-1}{\phi_1'(1)} \cdot \mathbf{D}_{\mathbf{g}(1)}\mathbf{\Omega}_K \left(\mathbf{D}_{\mathbf{g}(1)}^{1/2}\mathbf{w}_1(1)\right)\left(\mathbf{D}_{\mathbf{g}(1)}^{1/2}\mathbf{w}_1(1)\right)^\top . \tag{80}$$

To conclude, let's remark that we can rewrite Eq. (19) as

$$(\mathbf{D}_{\mathbf{g}(1)}^{1/2}\mathbf{\Omega}_K\mathbf{D}_{\mathbf{g}(1)}^{1/2})\left(\mathbf{D}_{\mathbf{g}(1)}^{-1/2}\mathbf{v}_1^{(r)}\right) = \left(\mathbf{D}_{\mathbf{g}(1)}^{-1/2}\mathbf{v}_1^{(r)}\right) \tag{81}$$

from which we deduce that

$$\mathbf{w}_1(1) = \frac{\mathbf{D}_{\mathbf{g}(1)}^{-1/2}\mathbf{v}_1^{(r)}}{\|\mathbf{D}_{\mathbf{g}(1)}^{-1/2}\mathbf{v}_1^{(r)}\|} \tag{82}$$

and hence if we denote by $C_0 = \|\mathbf{D}_{\mathbf{g}(1)}^{-1/2}\mathbf{v}_1^{(r)}\|^2$ and $C = -\phi_1'(1)$, we have

$$\lim_{N\to\infty} \mathbf{q}\mathbf{q}^\top = \frac{1}{C\cdot C_0} \cdot \left(\mathbf{D}_{\mathbf{g}(1)}\mathbf{\Omega}_K\mathbf{v}_1^{(r)}\right)(\mathbf{v}_1^{(r)})^\top \tag{83}$$

which simplifies using Eq. (19) as

$$\lim_{N\to\infty} \mathbf{q}\mathbf{q}^\top = \frac{1}{C\cdot C_0} \cdot \mathbf{v}_1^{(r)}(\mathbf{v}_1^{(r)})^\top \tag{84}$$

H. Proof of Prop. 3.13 (Value of the Derivative of the Top Eigenvalue)

We recall the standard result concerning the first-order perturbation theory for eigenvalues of symmetric matrices.

Lemma H.1 (Hadamard First Variation Formula)**.** *Let $\mathbf{A}(t)$ be symmetric and smooth for its parameter t. For t_0, assume that $\lambda_i(t_0) \equiv \lambda_i\big(\mathbf{A}(t_0)\big)$ is simple with associated eigenvector $\mathbf{v}_i(t_0)$, then $\lambda_i'(t_0)$ is well-defined and given by $\lambda_i'(t_0) = \langle \mathbf{v}_i(t_0), \mathbf{A}'(t_0)\mathbf{v}_i(t_0)\rangle$.*

Proof. See for example Chap. 1 of (Tao, 2020). This is simply obtained by combining the differentiation of the eigenvalue equation $\mathbf{A}(t)\mathbf{v}_i(t) = \lambda_i(t)\mathbf{v}_i(t)$ and the differentiation of the normalization of the top eigenvector $\langle \mathbf{v}_i(t), \mathbf{v}_i(t)\rangle = 1$. □

In particular, since $\phi_1(1+\epsilon)$ is the top eigenvalue of the symmetric matrix $\mathbf{D}_{\mathbf{g}(1+\epsilon)}^{1/2}\mathbf{\Omega}_K\mathbf{D}_{\mathbf{g}(1+\epsilon)}^{1/2}$ and $\phi(1)$ is simple, we get after simplification, the simple formula:

$$\phi_1'(1) = \frac{1}{C_0} \cdot \langle \mathbf{v}_1^{(r)}, \mathbf{\Omega}_K\mathbf{D}_{\mathbf{g}'(1)}\mathbf{\Omega}_K\mathbf{v}_1^{(r)}\rangle , \tag{85}$$

with $C_0 = \langle \mathbf{v}_1^{(r)}, \mathbf{D}_{\mathbf{g}(1)}^{-1}\mathbf{v}_1^{(r)}\rangle$. Furthermore, we see from App. B, that $\mathbf{g}'(1)$ is by definition given by the solution of

$$\mathbf{1}_K = (\nabla_{\mathbf{g}} f)(\mathbf{g}(1))\, \mathbf{g}'(1) , \tag{86}$$

and since $(\nabla_{\mathbf{g}} f)(\mathbf{g}(1))$ is invertible for $\lambda_1(\mathbf{\Omega}_K) > 1$, we get using its expression (43):

$$\mathbf{g}'(1) = (\mathbf{D}_{\mathbf{g}(1)}^{-2} - \mathbf{\Gamma}_K)^{-1}\mathbf{1}_K . \tag{87}$$

Using the explicit expressions (17) (16) for the left and right eigenvector gives us:

$$-\phi_1'(1) = \frac{1}{C_0} \cdot \langle \mathbf{1}_K - \mathbf{g}(1), \mathbf{P}(\mathbf{1}_K - \mathbf{g}(1))\rangle , \tag{88}$$

with

$$\mathbf{P} = \mathbf{D}_{\boldsymbol{\rho}^{\odot 1/2}}\mathbf{\Omega}_K\mathbf{D}_{-\mathbf{g}'(1)}\mathbf{\Omega}_K\mathbf{D}_{\boldsymbol{\rho}^{\odot 1/2}} \tag{89}$$

which gives the desired result using the definition (10) of $\mathbf{\Gamma}_K$ and the commutativity of diagonal matrices.

I. Properties of the Signal-to-Noise Ratio $\lambda_1(\mathbf{\Omega}_K)$.

In this Appendix, we describe simple properties of the Signal-to-Noise Ratio. We recall the assumptions $\rho_k > 0$ and $s_{kl} > 0$ for any $1 \le k, l \le K$ which implies that $\lambda_1(\mathbf{\Omega}_K)$ is simple with top eigenvector $\mathbf{v}_1 \succ \mathbf{0}$ by Perron-Frobenius theorem.

I.1. Behavior of the SNR with the Block Noise Level

As described in Thm. 2.6, the top eigenvalue $\lambda_1(\mathbf{\Omega}_K)$ plays the role of the signal-to-noise ratio (SNR) for the inhomogeneous spiked model. We first prove that as we decrease the noise level in any block, which translates into increasing any s_{kl} (recall $Y_{ij} = (x_i x_j)/\sqrt{N} + H_{ij}\Sigma_{ij}^{-1/2}$ with $\Sigma_{ij} = s_{kl}$ for any $(i,j) \in B_k \times B_l$), we improve the SNR, as one should expect.

Proposition I.1. *For any $k, l \in \{1, \dots, K\}^2$, $\lambda_1(\mathbf{\Omega}_K)$ is an increasing function of s_{kl}.*

Proof. The matrix $\mathbf{\Omega}_K$ is continuous and differentiable for any s_{kl}, with derivative given by $\frac{\mathrm{d}}{\mathrm{d}(s_{kl})}\mathbf{\Omega}_K = \sqrt{\rho_k \rho_l}(\mathbf{e}_k \mathbf{e}_l^\top + \mathbf{e}_l \mathbf{e}_k^\top)$. Since the top eigenvalue is simple, we have by Hadamard's first variation formula (see Lem. H.1) that $\lambda_1(\mathbf{\Omega}_K)$ is differentiable for s_{kl} with derivative given by

$$\frac{\mathrm{d}}{\mathrm{d}(s_{kl})}\lambda_1(\mathbf{\Omega}_K) = \langle \mathbf{v}_1, \sqrt{\rho_k \rho_l}(\mathbf{e}_k \mathbf{e}_l^\top + \mathbf{e}_l \mathbf{e}_k^\top)\mathbf{v}_1 \rangle = 2\sqrt{\rho_k \rho_l} v_{1k} v_{1l} > 0, \tag{90}$$

where in the last equality, we have used that $\mathbf{v}_1 = (v_{11}, \dots, v_{1K}) \succ \mathbf{0}$ by the Perron-Frobenius theorem. □

I.2. Behavior of the SNR with Compression of the Data

Recall the partition of the signal into its K parts, $\mathbf{x} = \mathbf{x}_1 \oplus \cdots \oplus \mathbf{x}_K$ where for any $k \in \{1, \dots, K\}$, the *k-th species* is given as $\mathbf{x}_k = (\mathbf{x})_{i \in B_k}$ and for two vectors $\mathbf{x}, \mathbf{y} \in (\mathbb{R}^N, \mathbb{R}^M)$, $\mathbf{x} \oplus \mathbf{y} = (x_1, \dots, x_N, y_1, \dots, y_M) \in \mathbb{R}^{N+M}$. Next, let's consider that *one is not interested in retrieving the entire signal* $\mathbf{x}$ *but only a reduced version of it containing some $K' < K$ species*, that is explicitly one is interested in retrieving $\mathbf{x}^{(\mathrm{rdc})} = \bigoplus_{i \in I} \mathbf{x}_i$ for some $I \subset \{1, \dots, K\}$ such that $|I| = K'$. A typical example corresponds to the case $\mathbf{x}^{(\mathrm{rdc})} = \mathbf{x}_1$ where one is only interested in retrieving one species out of the K species. One may then ask if it is best

- to consider the transformed matrix $\tilde{\boldsymbol{Y}}$ on the *entire data matrix* $\boldsymbol{Y}$ of size $(N \times N)$, and then keep only the overlap with the reduced signal,
- or to consider the transformed matrix $\tilde{\boldsymbol{Y}}^{(\mathrm{rdc})}$ obtained from the reduced data matrix where one removes any component Y_{ij} such that $(i,j) \notin I^2$. The previous example with $K' = 1$ would correspond to looking at the *homogeneous spiked model* made of the top $(|B_1| \times |B_1|)$ left corner of the data matrix $\boldsymbol{Y}$, instead of the entire matrix $\boldsymbol{Y}$.

It is natural to expect the second option to be suboptimal to the first one as it loses the information on cross terms Y_{ij} where $i \in I$ and $j \notin I$ which might help retrieve the reduced signal $\mathbf{x}^{(\mathrm{rdc})}$. Our next result indicates that this is indeed the case, at the level of the SNR.

Proposition I.2. *If $\mathbf{\Omega}_{K'}^{(\mathrm{rdc})}$ is the parameter matrix for the reduced data matrix $\boldsymbol{Y}^{(\mathrm{rdc})}$, then we have:*

$$\lambda_1(\mathbf{\Omega}_{K'}^{(\mathrm{rdc})}) < \lambda_1(\mathbf{\Omega}_K)\,. \tag{91}$$

Proof. By symmetry and induction, it is enough to prove $\lambda_1(\mathbf{\Omega}_{K-1}^{(\mathrm{rdc})}) < \lambda_1(\mathbf{\Omega}_K)$, where $\mathbf{\Omega}_{K-1}^{(\mathrm{rdc})}$ corresponds to the parameter matrix for the reduced data matrix without the $|B_K|$ last rows and columns of the data matrix $\boldsymbol{Y}$. For $N' = N - |B_K|$, we thus have

$$(Y_{ij}^{(\mathrm{rdc})})_{1 \le i,j \le N'} = \Big(\frac{x_i x_j}{\sqrt{N}} + H_{ij}\Sigma_{ij}^{-1/2}\Big)_{1 \le i,j \le N'}\,, \tag{92}$$

In particular, note that in the large N limit, the non-zero eigenvalue of the rank-one matrix has been reduced by a factor $N'/N < 1$ since $\mathbf{x}^{(\mathrm{rdc})} \in \mathbb{R}^{N'}$. If we normalize the matrix

$$\boldsymbol{Y}^{(0)} = \sqrt{\frac{N}{N'}}\boldsymbol{Y}^{(\mathrm{rdc})}\,, \tag{93}$$

we get back to our original setting with $K-1$ species. If we denote by $\alpha := \lim_{N\to\infty} N'(N)/N \in (0,1)$, the parameters $(\rho_k^{(\mathrm{rdc})}, s_{kl}^{(\mathrm{rdc})})_{1\le k,l\le K-1}$ for this reduced model are related to the ones $(\rho_k, s_{kl})_{1\le k,l\le K}$ of the entire model by:

- $s_{k,l}^{(\mathrm{rdc})} = \frac{s_{k,l}}{\alpha}$ for any $1 \le k, l \le K-1$;
- $\rho_k^{(\mathrm{rdc})} = \alpha\rho_k$ for any $1 \le k \le K-1$.

Since $(\boldsymbol{\Omega}_{K-1}^{(\mathrm{rdc})})_{kl} = \sqrt{\rho_k^{(\mathrm{rdc})}\rho_l^{(\mathrm{rdc})}}s_{k,l}^{(\mathrm{rdc})} = \sqrt{\rho_k\rho_l}s_{kl}$, the parameter matrix $\boldsymbol{\Omega}_{K-1}^{(\mathrm{rdc})}$ is simply the top left $\big((K-1)\times(K-1)\big)$ minor of $\boldsymbol{\Omega}_K$. We can then use the same argument as for the proof of Cauchy interlacing theorem: if we denote by $\mathbf{v}_1^{(\mathrm{rdc})} \succ \mathbf{0}$ the top eigenvector of $\boldsymbol{\Omega}_{K-1}^{(\mathrm{rdc})}$, and $\mathbf{v}_1$ the one of $\boldsymbol{\Omega}_K$, we have:

- on the one hand:
$$\lambda_1(\boldsymbol{\Omega}_{K-1}^{(\mathrm{rdc})}) = \langle \mathbf{v}_1^{(\mathrm{rdc})}, \boldsymbol{\Omega}_{K-1}^{(\mathrm{rdc})}\mathbf{v}_1^{(\mathrm{rdc})}\rangle = \langle (\mathbf{v}_1^{(\mathrm{rdc})} \oplus 0), \boldsymbol{\Omega}_K(\mathbf{v}_1^{(\mathrm{rdc})} \oplus 0)\rangle\,, \tag{94}$$
where we recall $(\mathbf{v}_1^{(\mathrm{rdc})} \oplus 0) = (v_{1,1}^{(\mathrm{rdc})}, \dots, v_{1,K-1}^{(\mathrm{rdc})}, 0)$.
- on the other hand:
$$\lambda_1(\boldsymbol{\Omega}_K) = \max_{\mathbf{v},\|\mathbf{v}\|=1}\langle \mathbf{v}, \boldsymbol{\Omega}_K\mathbf{v}\rangle = \langle \mathbf{v}_1, \boldsymbol{\Omega}_K\mathbf{v}_1\rangle\,. \tag{95}$$

Since $\mathbf{v}_1 \succ \mathbf{0}$, we must have $\mathbf{v}_1 \neq \mathbf{v}_1^{(\mathrm{rdc})}$ leading to the desired result since $\lambda_1(\boldsymbol{\Omega}_K)$ is simple. □

Note that Prop. I.2 indicates that reducing the data matrix might lead to losing any correlation between the top eigenvector and the signal since for a model with parameters such that $\lambda_1(\boldsymbol{\Omega}_K) > 1$, the overlap between the top eigenvector and *each* species $\mathbf{x}_k$ is positive, and in particular one has also a positive overlap between the (entire) top eigenvector and the reduced signal vector $\mathbf{x}^{(\mathrm{rdc})}$. By reducing the data matrix, one may end up with $\lambda_1(\boldsymbol{\Omega}_{K'}^{(\mathrm{rdc})}) < 1$ such that there is no positive overlap between the associated top eigenvector $\mathbf{v}_1^{(\mathrm{rdc})}$ and the reduced signal $\mathbf{x}^{(\mathrm{rdc})}$.

I.3. Explicit and Semi-Explicit Form of the SNR in Specific Cases

For $K = 1$, one has $\boldsymbol{\Omega}_K = s$ and we retrieve the usual BBP threshold for homogeneous spiked Wigner model.
For $K = 2$ and $\boldsymbol{\rho} = (\rho, 1-\rho)$, one has

$$\lambda_1(\boldsymbol{\Omega}_2) := \frac{\rho s_{11} + (1-\rho)s_{22} + \sqrt{4\rho s_{12}^2 + s_{22}^2 - 2\rho s_{22}(s_{11}+s_{22}) + \rho^2((s_{11}+s_{22})^2 - 4s_{12}^2)}}{2}\,, \tag{96}$$

in particular, in the limit (not studied in this paper) where $s_{11} = s_{22} = 0$ and $s_{12} \equiv s$, one has $\lambda_1(\boldsymbol{\Omega}_2) = \sqrt{(1-\rho)\rho}\cdot s$, which is the transition for the *asymmetric* rank-one homogeneous channel, as one should expect.
Eventually, for a generic K with the same noise level on the diagonal $s_{kk} = s_{\mathrm{in}}$ for all $k \in [\![K]\!]$ and the same noise level off-diagonal $s_{kl} = s_{\mathrm{out}} \neq s_{\mathrm{in}}$ for all $k \neq l$, we have $\boldsymbol{\Omega}_K = (s_{\mathrm{in}} - s_{\mathrm{out}})\mathbf{D}_{\boldsymbol{\rho}} + s_{\mathrm{out}}(\boldsymbol{\rho}^{\odot 1/2})(\boldsymbol{\rho}^{\odot 1/2})^\top$, such that $\lambda_1(\boldsymbol{\Omega}_K)$ is the highest zero of the secular equation

$$1 + s_{\mathrm{out}}\sum_k \frac{\rho_k}{(s_{\mathrm{in}} - s_{\mathrm{out}})\rho_k - \lambda} = 0\,. \tag{97}$$

In particular, for the balanced model $\boldsymbol{\rho} = (1/K, \dots, 1/K)$, this is given explicitly by

$$\lambda_1(\boldsymbol{\Omega}_K) = s_{\mathrm{out}}\big(1 - \frac{1}{K}\big) + \frac{s_{\mathrm{in}}}{K}\,. \tag{98}$$