

# KERNEL SPECTRAL CLUSTERING OF LARGE DIMENSIONAL DATA

ROMAIN COUILLET AND FLORENT BENAYCH-GEORGES

ABSTRACT. This article proposes a first analysis of kernel spectral clustering methods in the regime where the dimension  $p$  of the data vectors to be clustered and their number  $n$  grow large at the same rate. We demonstrate, under a  $k$ -class Gaussian mixture model, that the normalized Laplacian matrix associated with the kernel matrix asymptotically behaves similar to a so-called spiked random matrix. Some of the isolated eigenvalue-eigenvector pairs in this model are shown to carry the clustering information upon a separability condition classical in spiked matrix models. We evaluate precisely the position of these eigenvalues and the content of the eigenvectors, which unveil important properties concerning spectral clustering, in particular in simple toy models. Our results are then compared to the practical clustering of images from the MNIST database, thereby revealing an important match between theory and practice.

## CONTENTS

1. Introduction	1
2. Motivation and statement of main results	4
3. Preliminaries	7
4. Main Results	13
5. Special cases	25
6. Concluding Remarks	36
7. Proofs	38
Appendix: concentration lemmas	57
References	58

## 1. INTRODUCTION

Kernel spectral clustering encompasses a variety of algorithms meant to cluster data in an unsupervised manner based on the eigenvectors of certain data-driven matrices. These methods are so widely spread that they have become an essential ingredient of contemporary machine learning (see [V07] and references therein). This being said, the theoretical foundations of kernel spectral clustering are not unified as it can be obtained

---

2000 *Mathematics Subject Classification.* 62H30;60B20;15B52.

*Key words and phrases.* kernel methods; spectral clustering; random matrix theory.

Couillet's work is supported by the ANR Project RMT4GRAPH (ANR-14-CE28-0006).

from several independent ideas, hence the multiplicity of algorithms to meet the same objective.

Denote  $x_1, \dots, x_n \in \mathbb{R}^p$  the data vectors to be clustered in  $k$  similarity classes and  $\kappa : \mathbb{R}^p \times \mathbb{R}^p \rightarrow \mathbb{R}_+$  some data-affinity function (which by convention takes large values for resembling data pairs and small values for distinct vectors). We shall denote by  $K$  the kernel matrix defined by  $K_{ij} = \kappa(x_i, x_j)$ . One of the original approaches [V07] to data clustering consists in the relaxation of the following integer problem

$$\text{(RatioCut)} \quad \operatorname{argmin}_{\mathcal{S}_1 \cup \dots \cup \mathcal{S}_k = \{1, \dots, n\}} \sum_{i=1}^k \sum_{\substack{j \in \mathcal{S}_i \\ j \notin \mathcal{S}_i}} \frac{\kappa(x_i, x_j)}{|\mathcal{S}_i|}$$

where the minimum is taken over disjoint sets  $\mathcal{S}_i$ . Here the normalization by  $|\mathcal{S}_i|$  (with  $|\cdot|$  the set cardinality) ensures that each class is of approximately balanced weight. Letting  $M \in \mathbb{R}^{n \times k}$  be the matrix with  $[M]_{ij} = |\mathcal{S}_j|^{-\frac{1}{2}} \delta_{\{x_i \in \mathcal{S}_j\}}$  and  $\mathcal{M}$  the set of such matrices, this can be shown to be strictly the same as

$$\text{(RatioCut)} \quad \operatorname{argmin}_{M \in \mathcal{M}} \operatorname{tr} M^\top (D - K) M$$

where, with  $D \triangleq \mathcal{D}(K1_n)$  (with  $\mathcal{D}(\cdot)$  the diagonal operator),  $D - K$  is the so-called (unnormalized) Laplacian matrix of  $K$  and  $M$  contains the information about the  $\mathcal{S}_i$  classes. Observing that  $M^\top M = I_k$ , one may relax the above problem to

$$\text{(RelaxedRatioCut)} \quad \min_{\substack{M \in \mathbb{R}^{n \times k} \\ M^\top M = I_k}} \operatorname{tr} M^\top (D - K) M$$

which then reduces to an eigenvector problem. From the original form of  $M \in \mathcal{M}$ , the data clusters can then readily be retrieved from the entries of  $M$ .

Alternatively, the intuition from the popular Ng–Jordan–Weiss algorithm [NJW01] starts off by considering the so-called *normalized Laplacian*  $I_n - D^{-\frac{1}{2}} K D^{-\frac{1}{2}}$  of  $K$  and by noticing that, if ideally  $\kappa(x_i, x_j) = 0$  when  $x_i$  and  $x_j$  belong to distinct classes (and  $\kappa(x_i, x_j) \neq 0$  otherwise), then the vectors  $D^{\frac{1}{2}} 1_{\mathcal{S}_i} \in \mathbb{R}^n$ , where  $1_{\mathcal{S}_i}$  is the indicator vector of the class index  $\mathcal{S}_i$  (composed of ones for the indices of  $\mathcal{S}_i$  and zero otherwise), are eigenvectors for the zero eigenvalue of the normalized Laplacian. In practice,  $\kappa(x_i, x_j)$  is merely expected to take small values for vectors of distinct classes, so that the algorithm consisting in retrieving classes from the eigenvectors associated to the smallest eigenvalues of the (nonnegative definite) normalized Laplacian matrix will approximately perform the desired task.

Theoretically speaking, assuming  $x_1, \dots, x_n \in \mathbb{R}^p$  independently distributed as a  $k$ -mixture probability measure, it is proved in [VBB08] that the various spectral clustering algorithms are consistent as  $n \rightarrow \infty$  in the sense that they shall return the statistical clustering allowed by the kernel function  $\kappa$ . Despite this important result, it nonetheless remains unclear which clustering performance can be achieved for all finite  $n, p$  couples. This is all the more needed that spectral clustering methods are being increasingly used in settings where  $p$  can be of similar size, if not much larger, than  $n$ . In this article, we aim at providing a first understanding of the behavior of kernel spectral clustering as both  $n$

and  $p$  are large but of similar order of magnitude. To this end, we shall leverage the recent result from [EK10] on the limiting spectrum of kernel random matrices for independent and identically distributed zero mean (essentially Gaussian) vectors. Our approach is to generalize the latter to a  $k$ -class Gaussian mixture model for the normalized Laplacian of  $K$ , rather than for the kernel matrix  $K$  itself.

Our focus, for reasons discussed below, is precisely on the following version of the normalized Laplacian matrix

$$L \triangleq nD^{-\frac{1}{2}}KD^{-\frac{1}{2}} \quad (1)$$

which we shall from now on refer to, with a slight language abuse, as the Laplacian matrix of  $K$ . The kernel function  $\kappa$  will be such that  $\kappa(x, y) = f(\|x - y\|^2/p)$  for some sufficiently smooth  $f$  independent of  $n, p$ , that is

$$K = \left\{ f \left( \frac{1}{p} \|x_i - x_j\|^2 \right) \right\}_{1 \leq i, j \leq n}. \quad (2)$$

For  $x_i$  in class  $\mathcal{C}_a$ , we assume that  $x_i \sim \mathcal{N}(\mu_a, C_a)$  with  $(\mu_a, C_a) \in \{(\mu_1, C_1), \dots, (\mu_k, C_k)\}$  ( $k$  shall remain fixed while  $p, n \rightarrow \infty$ ), with, for each  $(a, b)$ ,  $\|\mu_a - \mu_b\| = O(1)$  while  $\|C_a\| = O(1)$  and  $\text{tr}(C_a - C_b) = O(\sqrt{p})$ . This setting can be considered as a critical growth rate regime in the sense that, supposing  $\text{tr} C_a$  and  $\text{tr} C_a^2$  to be of order  $O(p)$  (which is natural if  $\|C_a\| = O(1)$ ) and  $\|\mu_a - \mu_b\| = O(1)$ , the norms of the observations in each class  $\mathcal{C}_a$  fluctuate at rate  $\sqrt{p}$  around  $\text{tr} C_a$ , so that clustering ought to be possible so long as  $\text{tr}(C_a - C_b) = O(\sqrt{p})$ .

The contribution of this work is to provide a thorough analysis of the eigenvalues and eigenvectors of the matrix  $L$  in the aforementioned regime. In a nutshell, we shall demonstrate that there exist critical values for the means  $\mu_i$  and covariances  $C_i$  beyond which some (relevant) eigenvalues of  $L$  tend to isolate from the majority of the eigenvalues, thus inducing a so-called spiked model for  $L$ . When this occurs, the eigenvectors associated to these isolated eigenvalues will contain information about class clustering. Our objective is to precisely describe the structure of the individual eigenvectors as well as to evaluate correlation coefficients among these, keeping in mind that the ultimate interest is on harnessing spectral clustering methods.

Before delving concretely into our main results, some of which may seem quite cryptic on the onset, we introduce below a motivation example and some visual results of our work.

The proofs of some of the technical mathematical results are deferred to our companion paper [BC15].

*Notations:* The norm  $\|\cdot\|$  stands for the Euclidean norm for vectors and the associated operator norm for matrices. The notation  $\mathcal{N}(\mu, C)$  is the multivariate Gaussian distribution with mean  $\mu$  and covariance  $C$ . The vector  $1_m \in \mathbb{R}^m$  stands for the vector filled with ones. The delta Dirac measure is denoted  $\delta$ . The operator  $\mathcal{D}(v) = \mathcal{D}\{v_a\}_{a=1}^k$  is the diagonal matrix having  $v_1, \dots, v_k$  (scalars or vectors) as its ordered diagonal elements. The distance from a point  $x \in \mathbb{R}$  to a set  $\mathcal{X} \subset \mathbb{R}$  is denoted  $\text{dist}(x, \mathcal{X})$ . The support of a measure  $\nu$  is denoted  $\text{supp}(\nu)$ .

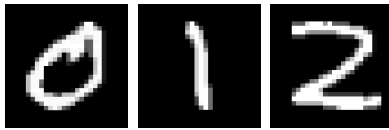


FIGURE 1. Samples from the MNIST database.

We shall often denote  $\{v_a\}_{a=1}^k$  a column vector with  $a$ -th entry (or block entry)  $v_a$  (which may be a vector itself), while  $\{V_{ab}\}_{a,b=1}^k$  denotes a square matrix with entry (or block-entry)  $(a, b)$  given by  $V_{ab}$  (which may be a matrix itself).

## 2. MOTIVATION AND STATEMENT OF MAIN RESULTS

Let us start by illustratively motivate our work. In Figure 2 are displayed in red the eigenvectors associated with the four largest eigenvalues of  $L$  (as defined in (1)) for  $x_1, \dots, x_n$  a set of (preprocessed<sup>1</sup>) vectorized images sampled from the popular MNIST database (handwritten digits) [LCB98]. The vector images are of size  $p = 784$  (for images are of size  $28 \times 28$ ) and we take  $n = 192$  samples, with the first 48  $x_i$ 's being images of zeros, next 48 images of ones, and last 96 images of twos. An example of these is displayed in Figure 1 (the vector values follow a grayscale from zero for black to one for white). The kernel function  $f$  (defining the kernel matrix  $K$  through (2)) is taken to be the standard  $f(x) = \exp(-x/2)$  Gaussian kernel.

As recalled earlier, performing spectral clustering on  $x_1, \dots, x_n$  consists here in extracting the leading eigenvectors from  $L$  and applying a procedure that identifies the three classes and properly maps every datum to its own class. As can be observed from Figure 2, the classes are visually clearly discriminated, although it also appears that some data stick out which, no matter how thorough the aforementioned procedure, are bound to be misclassified. A particularly common algorithm to cluster the data consists in extracting the entry  $i$  from each of the, say  $l$ , dominant eigenvectors to form a vector  $y_i \in \mathbb{R}^l$ . Clustering then resorts to applying some standard unsupervised classification technique, such as the popular k-means method [B06, Chapter 9], to the small dimensional vectors  $y_1, \dots, y_n$ . In Figure 3 are depicted the vectors  $y_i$  for  $l = 2$ , successively only accounting for the leading eigenvectors one and two (top figure) or two and three (bottom graph of Figure 2) of  $L$ . The crosses, each corresponding to a specific  $y_i \in \mathbb{R}^2$  vector, are colored according to their genuine class. As one can observe, for each class there exists a non trivial correlation between the entries of  $y_i$ , and thus between the eigenvectors of  $L$ .

In order to anticipate the performance of clustering methods, and to be capable of improving the latter, it is a fundamental first step to understand the behavior of the eigenvectors of Figure 2 along with their joint correlation, as exemplified in Figure 3. The present work intends to lay the theoretical grounds for such clustering performance understanding and improvement. Namely, we shall investigate the existence and position of isolated eigenvalues of  $L$  and shall show that some of them (not always all of them) carry

<sup>1</sup>The full MNIST database is preprocessed so that the (empirical) averaged squared vector norm is  $p$ .

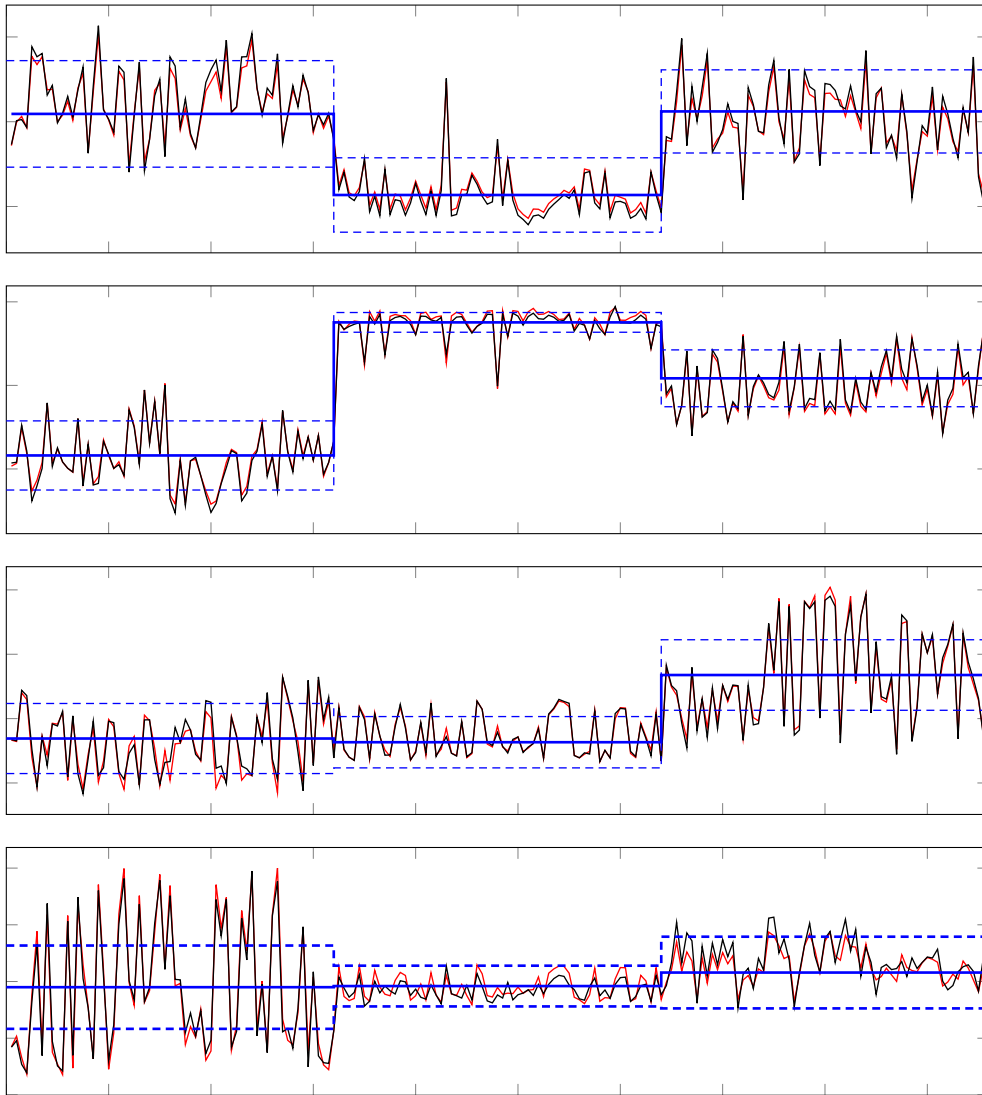


FIGURE 2. Leading four eigenvectors of  $L$  (red) versus  $\hat{L}$  (black) and theoretical class-wise means and standard deviations (blue); MNIST data.

information about the class structure of the problem. Then, since, by a clear invariance property of the model under consideration, each of the dominant eigenvectors of  $L$  can be divided class-wise into  $k$  chunks, each of which being essentially composed of independent realizations of a random variable with given mean and variance, we shall identify these means and variances. Finally, since eigenvectors are correlated, we shall evaluate the class-wise correlation coefficients.

As a first glimpse on the practical interest of our results, in Figure 2 are displayed in blue lines the theoretical means and standard deviations for each class-wise chunk of

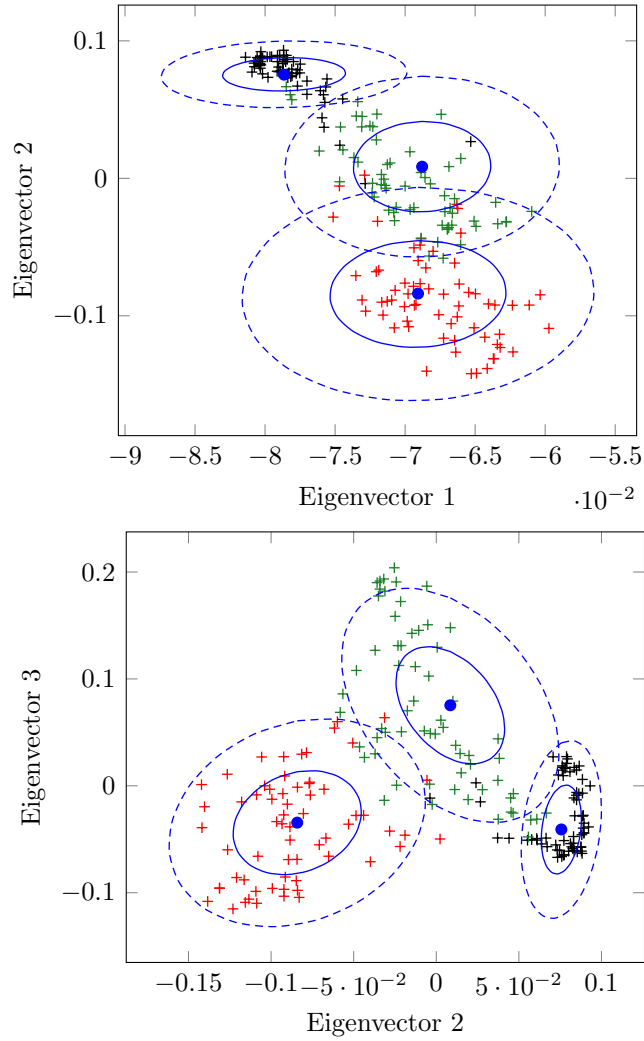


FIGURE 3. Two dimensional representation of the eigenvectors one and two (top) and two and three (bottom) of  $L$ , for the MNIST dataset. In blue, theoretical means and one- and two-standard deviations of fluctuations.

eigenvectors, obtained from the results of this article. That is, the means and standard deviations that one would obtain if the data were genuinely Gaussian (which here for the MNIST images they are obviously not). Also, Figure 3 proposes in blue ellipses the theoretical one- and two-standard deviations of the joint eigenvector entries, again if the data were to be Gaussian. It is quite interesting to see that, in spite of their evident non-Gaussianity, the theoretical findings visually conform to the data behavior. We are thus optimistic that the findings of this work, although restricted to Gaussian assumptions, can be applied to a large set of problems beyond strongly structured ones.

In more technical words, our main results, stated in complete detail in Section 4 below, may be summarized as follows:

- (1) as  $n, p \rightarrow \infty$  while  $n/p = O(1)$ ,  $\|L' - \hat{L}'\| \rightarrow 0$  (in operator norm) almost surely, where  $L'$  is a slight modification of  $L$  and  $\hat{L}'$  is a matrix which is an instance of the so-called spiked random matrix models, as introduced in [BBP05, BN12] (but closer to the model studied independently in [CH13]); that is, the spectrum of  $\hat{L}'$  is essentially composed of (one or several) clusters of eigenvalues and finitely many isolated ones;
- (2) as is standard in spiked models, there exists a phase transition phenomenon by which, the more distinct the classes, the more eigenvalues tend to isolate from the main eigenvalue bulk of  $\hat{L}'$  and the more information is contained within the eigenvectors associated with those eigenvalues. This statement is precisely accounted for by exhibiting conditions for the separability of the isolated eigenvalues from the main bulk, by exactly locating these eigenvalues, and by retrieving the asymptotic values of the class-wise means and variances of the isolated eigenvectors;
- (3) the eigenvectors associated to the isolated eigenvalues are correlated to one another and we precisely exhibit the asymptotic correlation coefficients.

Aside from these main expected results are some more subtle and somewhat unexpected outcomes:

- (4) the eigenvectors associated with some of the *non extreme* isolated eigenvalues of  $L'$  may contain information about the classes, and thus clustering may be performed not only based on extreme eigenvectors;
- (5) on the contrary, some of the eigenvectors associated to isolated eigenvalues, even the largest, may be purely noisy;
- (6) in some specific scenarios, the theoretical number of informative isolated eigenvalues cannot exceed two altogether, while in others as many as  $k - 1$  can be found in-between each pair of eigenvalue bulks of  $L'$ ;
- (7) in some other scenarios, two eigenvectors may be essentially the same, so that some eigenvectors may not always provide much information diversity.

We now turn to the detailed introduction of our model and to some necessary preliminary notions of random matrix theory.

### 3. PRELIMINARIES

Let  $x_1, \dots, x_n \in \mathbb{R}^p$  be independent vectors belonging to  $k$  distribution classes  $\mathcal{C}_1, \dots, \mathcal{C}_k$ , with  $x_{n_1+\dots+n_{a-1}+1}, \dots, x_{n_1+\dots+n_a} \in \mathcal{C}_a$  for each  $a \in \{1, \dots, k\}$  (so that each class  $\mathcal{C}_a$  has cardinality  $n_a$ ), where  $n_0 = 0$  and  $n_1 + \dots + n_k = n$ . We assume that  $x_i \in \mathcal{C}_a$  is given by

$$x_i = \mu_a + \sqrt{p}w_i$$

for some  $\mu_a \in \mathbb{R}^p$  and  $w_i \sim \mathcal{N}(0, p^{-1}C_a)$ , with  $C_a \in \mathbb{R}^{p \times p}$  nonnegative definite (the factors  $\sqrt{p}$  and  $p^{-1}$  will lighten the notations below).

We shall consider the large dimensional regime where both  $n$  and  $p$  are simultaneously large with the following growth rate assumption.

**Assumption 1** (Growth Rate). *As  $n \rightarrow \infty$ , the following conditions hold.*

(1) **Data scaling:** defining  $c_0 \triangleq \frac{p}{n}$

$$0 < \liminf_n c_0 \leq \limsup_n c_0 < \infty$$

(2) **Class scaling:** for each  $a \in \{1, \dots, k\}$ , defining  $c_a \triangleq \frac{n_a}{n}$ ,

$$0 < \liminf_n c_a \leq \limsup_n c_a < \infty.$$

We shall denote  $c \triangleq \{c_a\}_{a=1}^k$ .

(3) **Mean scaling:** let  $\mu^\circ \triangleq \sum_{i=1}^k \frac{n_i}{n} \mu_i$  and for each  $a \in \{1, \dots, k\}$ ,  $\mu_a^\circ \triangleq \mu_a - \mu^\circ$ , then

$$\limsup_n \max_{1 \leq a \leq k} \|\mu_a^\circ\| < \infty$$

(4) **Covariance scaling:** let  $C^\circ \triangleq \sum_{i=1}^k \frac{n_i}{n} C_i$  and for each  $a \in \{1, \dots, k\}$ ,  $C_a^\circ \triangleq C_a - C^\circ$ , then

$$\begin{aligned} \limsup_n \max_{1 \leq a \leq k} \|C_a^\circ\| &< \infty \\ \limsup_n \max_{1 \leq a \leq k} \frac{1}{\sqrt{n}} \operatorname{tr} C_a^\circ &< \infty. \end{aligned}$$

As discussed in the introduction, the growth rates above were chosen in such a way that the achieved clustering performance be non-trivial in the sense that: (i) the proportion of misclassification remains non-vanishing as  $n \rightarrow \infty$ , and (ii) there exists smallest values of  $\|\mu_a^\circ\|$  and  $\frac{1}{\sqrt{n}} \operatorname{tr} C_a^\circ$  below which no isolated eigenvector can be used to perform efficient spectral clustering.

For further use, we now define

$$\tau \triangleq \frac{2}{p} \operatorname{tr} C^\circ.$$

This quantity is important since it is easily shown that, under Assumption 1,

$$\max_{\substack{1 \leq i, j \leq n \\ i \neq j}} \left\{ \frac{1}{p} \|x_i - x_j\|^2 - \tau \right\} \rightarrow 0 \quad (3)$$

almost surely. The value  $\tau$ , which depends implicitly on  $n$ , is bounded but needs not converge as  $p \rightarrow \infty$ .

Let  $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  be a function satisfying the following assumptions.

**Assumption 2** (Kernel function). *The function  $f$  is three-times continuously differentiable in a neighborhood of the values taken by  $\tau$ . Moreover,  $\liminf_n f(\tau) > 0$ .*

Define now  $K$  to be the kernel matrix

$$K \triangleq \left\{ f \left( \frac{1}{p} \|x_i - x_j\|^2 \right) \right\}_{i, j=1}^n.$$



From (3), it appears that, while the diagonal elements of  $K$  are all equal to  $f(0)$ , the off-diagonal entries are jointly converging toward  $f(\tau)$ . This means that, up to  $(f(\tau) - f(0))I_n$ ,  $K$  is essentially a rank-one matrix. As spectral clustering is based on Laplacian matrices rather than on  $K$  itself, we shall focus here on the Laplacian matrix

$$L \triangleq nD^{-\frac{1}{2}}KD^{-\frac{1}{2}}$$

where  $D = \mathcal{D}(K1_n)$  is often referred to as the matrix of degrees of  $K$ . Aside from the arguments laid out in the introduction, the choice of studying the matrix  $L$  also follows from a better stability of clustering algorithms based on  $L$  versus  $K$  and  $D - K$  that we observed in various simulations.

The matrix  $L$  is also essentially a rank-one matrix but is rather simple to deal with since, unlike  $K$ , its dominant eigenvector is known precisely to be  $D^{\frac{1}{2}}1_n$  and it shall be shown that the projected matrix

$$L' \triangleq nD^{-\frac{1}{2}}KD^{-\frac{1}{2}} - n \frac{D^{\frac{1}{2}}1_n1_n^T D^{\frac{1}{2}}}{1_n^T D 1_n} \quad (4)$$

has bounded operator norm almost surely as  $n \rightarrow \infty$ . Note here that  $L'$  and  $L$  have the same eigenvalues and eigenvectors but for the eigenvalue-eigenvector pair  $(n, D^{\frac{1}{2}}1_n)$  of  $L$  turned into  $(0, D^{\frac{1}{2}}1_n)$  for  $L'$ . Under the aforementioned assumptions, the matrix  $L'$  will be subsequently shown to have its eigenvalues all of order  $O(1)$ .

Our first intermediary result shows that there exists a matrix  $\hat{L}'$  such that  $\|L' - \hat{L}'\| \rightarrow 0$  almost surely, where  $\hat{L}'$  follows an analytically tractable random matrix model. Before going into the result, a few notations need be introduced. In the remainder of the article, we shall use the following deterministic element notations<sup>2</sup>

$$\begin{aligned} M &\triangleq [\mu_1^\circ, \dots, \mu_k^\circ] \in \mathbb{R}^{p \times k} \\ t &\triangleq \left\{ \frac{1}{\sqrt{p}} \operatorname{tr} C_a^\circ \right\}_{a=1}^k \in \mathbb{R}^k \\ T &\triangleq \left\{ \frac{1}{p} \operatorname{tr} C_a^\circ C_b^\circ \right\}_{a,b=1}^k \in \mathbb{R}^{k \times k} \\ J &\triangleq [j_1, \dots, j_k] \in \mathbb{R}^{n \times k} \\ P &\triangleq I_n - \frac{1}{n} 1_n 1_n^T \in \mathbb{R}^{n \times n} \end{aligned}$$

---

<sup>2</sup>As a mental reminder, capital  $M$  stands here for *means* while  $t$ ,  $T$  account for vector and matrix of *traces*,  $P$  for a projection matrix (onto the orthogonal of the vector  $1_n$ ).

where  $j_a \in \mathbb{R}^n$  is the canonical vector of class  $\mathcal{C}_a$ , defined such that  $(j_a)_i = \delta_{x_i \in \mathcal{C}_a}$ , along with the random element notations

$$\begin{aligned} W &\triangleq [w_1, \dots, w_n] \in \mathbb{R}^{p \times n} \\ \Phi &\triangleq PW^\top M \in \mathbb{R}^{n \times k} \\ \psi &\triangleq \{\|w_i\|^2 - \mathbb{E}[\|w_i\|^2]\}_{i=1}^n \in \mathbb{R}^n. \end{aligned} \quad (5)$$

**Theorem 1** (Random Matrix Equivalent). *Let Assumptions 1 and 2 hold. Let  $L'$  be defined as in (4). Then, as  $n \rightarrow \infty$ ,*

$$\|L' - \hat{L}'\| \rightarrow 0$$

almost surely, where  $\hat{L}'$  is given by

$$\hat{L}' \triangleq -2 \frac{f'(\tau)}{f(\tau)} (PW^\top WP + UBU^\top) + 2 \frac{f'(\tau)}{f(\tau)} F(\lambda) I_n$$

with  $F(\lambda) = \frac{f(0) - f(\tau) + \tau f'(\tau)}{2f'(\tau)}$  and

$$\begin{aligned} U &\triangleq \left[ \frac{1}{\sqrt{p}} J, \Phi, \psi \right] \\ B &\triangleq \begin{bmatrix} B_{11} & I_k - 1_k c^\top & \left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right) t \\ I_k - c 1_k^\top & 0_{k \times k} & 0_{k \times 1} \\ \left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right) t^\top & 0_{1 \times k} & \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \end{bmatrix} \\ B_{11} &= M^\top M + \left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right) t t^\top - \frac{f''(\tau)}{f'(\tau)} T + \frac{p}{n} F(\tau) 1_k 1_k^\top \end{aligned}$$

and the case  $f'(\tau) = 0$  is obtained through extension by continuity ( $f'(\tau)B$  being well defined as  $f'(\tau) \rightarrow 0$ ).

From Theorem 1 it entails that the eigenvalues of  $L'$  and  $\hat{L}'$  converge to one another (as we have as an immediate corollary that  $\max_i |\lambda_i(L') - \lambda_i(\hat{L}')| \xrightarrow{\text{a.s.}} 0$ ) (see e.g., [HJ85, Theorem 4.3.7]), so that the determination of isolated eigenvalues in the spectrum of  $L'$  (or  $L$ ) can be studied from the equivalent problem for  $\hat{L}'$ . More importantly, from Theorem 1, it unfolds that, for every isolated eigenvector  $u$  of  $L'$  and its associated  $\hat{u}$  of  $\hat{L}'$ ,  $\|u - \hat{u}\| \xrightarrow{\text{a.s.}} 0$ . Thus, the spectral clustering performance based on the observable  $L'$  (or  $L$ ) may be asymptotically analyzed through that of  $\hat{L}'$ .

A few important remarks concerning Theorem 1 are in order before proceeding. From a mathematical standpoint, observe that, up to a scaled identity matrix and a constant scale factor, if  $f'(\tau) \neq 0$ ,  $\hat{L}'$  is a random matrix of the so-called spiked model family [BBP05] in that it equals the sum of a somewhat standard random matrix model  $PW^\top WP$  and of a small rank (here up to  $2k+1$ ) matrix  $UBU^\top$ . Nonetheless, it differs from classically studied spiked models in several aspects: (i)  $U$  is not independent of  $W$ , which is a technical issue

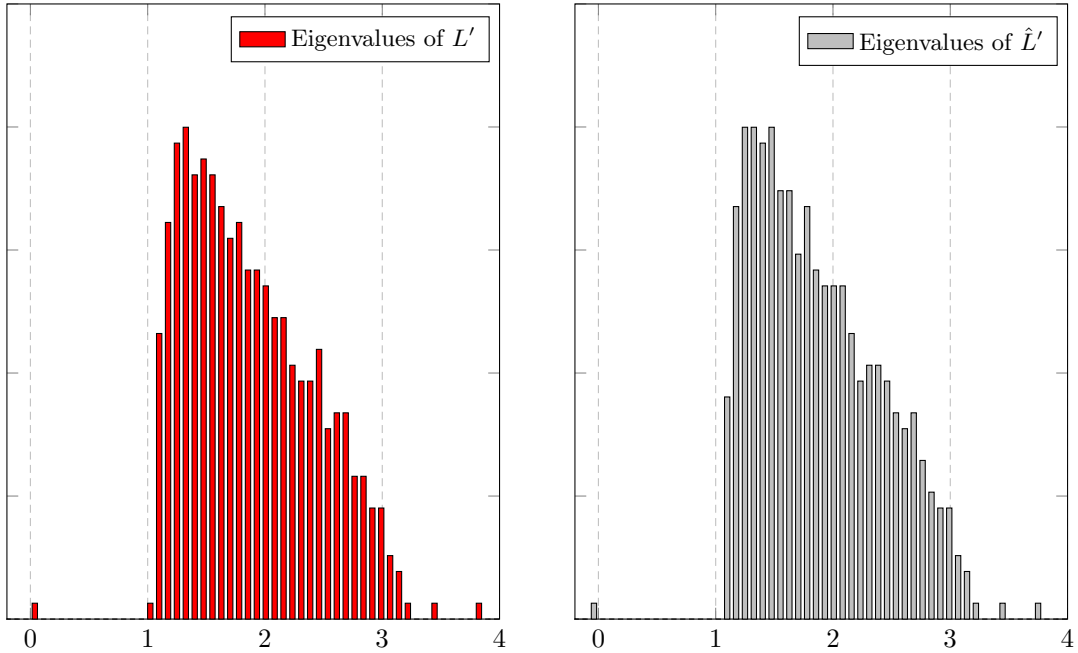


FIGURE 4. Eigenvalues of  $L'$  (left) and  $\hat{L}'$  (right), for  $p = 2048$ ,  $n = 512$ ,  $c_1 = c_2 = 1/4$ ,  $c_3 = 1/2$ ,  $[\mu_i]_j = 4\delta_{ij}$ ,  $C_i = (1 + 2(i - 1)/\sqrt{p})I_p$ ,  $f(x) = \exp(-x/2)$ .

that can fairly easily be handled, and (ii)  $PW^TWP$  itself constitutes a spiked model as  $P$  is a low rank perturbation of the identity matrix.<sup>3</sup>

As such, as  $n \rightarrow \infty$ , the eigenvalues of  $\hat{L}'$  are expected to be asymptotically the same as those of  $PW^TWP$  (which mainly gather in bulks) but possibly for finitely many of them which are allowed to wander away from the main eigenvalue bulks. As per classical spiked model results from random matrix theory, it is then naturally expected that, if some of the (finitely many) eigenvalues of  $UBU^T$  are sufficiently large, those shall induce isolated eigenvalues in the spectrum of  $\hat{L}'$ , the eigenvectors of which align to some extent to the eigenvectors of  $UBU^T$ . If instead  $f'(\tau) = 0$ , then  $\hat{L}' - \frac{f(0)-f(\tau)}{f(\tau)}I_n$  is of maximum rank  $k+1$  and is fully deterministic, hence has eigenvalue-eigenvector pairs immediately related to  $UBU^T$ .

From a practical aspect, observe that  $U$  is importantly constituted by the vectors  $j_a$ ,  $1 \leq a \leq k$ , while  $B$  contains the information about the inter-class mean deviations through  $M$ , and about the inter-class covariance deviations through  $t$  and  $T$ . As such, some of the

<sup>3</sup>Our choice of not breaking  $PW^TWP$  into  $W^TW$  plus small rank perturbation integrated to the perturbation  $UBU^T$  stands from the fact that the hypothetical isolated eigenvectors  $P$  engenders do not provide any clustering information, unlike  $UBU^T$ . Besides, by Remark 1 below or by interlacing inequalities,  $PW^TWP$  does not induce isolated eigenvalues on the right side of the support, where clustering algorithms look for eigenvalues.

aforementioned isolated eigenvalue-eigenvector pairs are expected to align to the canonical class basis  $J$  and we already intuit that this will be true all the more that the matrices  $M$ ,  $t$ ,  $T$  have sufficient energy (i.e., are sufficiently away from zero matrices). Theorem 1 thus already prefigures the behavior of spectral clustering methods thoroughly detailed in Section 4.

A more detailed analysis now sheds light on the behavior of the kernel function  $f$ . Note that, if  $f'(\tau) \rightarrow 0$ ,  $L'$  becomes essentially deterministic as  $f'(\tau)PW^\top WP \rightarrow 0$ , this having a positive effect of the alignment between  $L'$  and  $J$ . However, when  $f'(\tau) \rightarrow 0$ ,  $M$  vanishes from the expression of  $\hat{L}'$ , thus not allowing spectral clustering to rely on differences in means. Similarly, if  $f''(\tau) \rightarrow 0$ , then  $T$  vanishes, and thus differences in “shape” between the covariance matrices cannot be discriminated upon. Finally, if  $\frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \rightarrow 0$ , then differences in covariance traces are seemingly not exploitable.

An illustrative application of Theorem 1 is proposed in Figure 4, where a three-class example with Gaussian kernel function is considered. Note the extremely accurate spectrum approximation of  $L'$  by  $\hat{L}'$  in this example. Anticipating slightly our coming results, note here that, aside from the eigenvalue zero, two isolated (spiked) eigenvalues are observed, which shall presently be related to the eigenvalues of the small rank matrix  $UBU^\top$ .

Before introducing our technical approach, a few further random matrix notions are needed.

**Lemma 1** (Deterministic equivalent). *Let Assumptions 1 and 2 hold. Let  $z \in \mathbb{C}$  be at macroscopic distance from<sup>4</sup> the eigenvalues  $\lambda_1^p, \dots, \lambda_p^p$  of  $PW^\top WP$ . Let us define the resolvents  $Q_z \triangleq (PW^\top WP - zI_n)^{-1}$  and  $\tilde{Q}_z \triangleq (WPW^\top - zI_p)^{-1}$ . Then, as  $n \rightarrow \infty$ ,*

$$Q_z \leftrightarrow \bar{Q}_z \triangleq c_0 \mathcal{D} \{g_a(z)1_{n_a}\}_{a=1}^k - \left\{ \left( \frac{1}{z} + c_0 \frac{g_a(z)g_b(z)}{\sum_{i=1}^k c_i g_i(z)} \right) \frac{1_{n_a}1_{n_b}^\top}{n} \right\}_{a,b=1}^k$$

$$\tilde{Q}_z \leftrightarrow \bar{\tilde{Q}}_z \triangleq -z^{-1} \left( I_p + \sum_{a=1}^k c_a g_a(z) C_a \right)^{-1}$$

where the  $(g_1, \dots, g_k)$  is the unique vector of Stieltjes transforms solutions, for all such  $z$ , to the implicit equations

$$c_0 g_a(z) = -z^{-1} \left( 1 + \frac{1}{p} \text{tr} C_a \bar{\tilde{Q}}_z \right)^{-1}$$

and the notation  $A_n \leftrightarrow B_n$  stands for the fact that, as  $n \rightarrow \infty$ ,  $\frac{1}{n} \text{tr} D_n A_n - \frac{1}{n} \text{tr} D_n B_n \xrightarrow{\text{a.s.}} 0$  and  $d_{1,n}^\top(A_n - B_n)d_{2,n} \xrightarrow{\text{a.s.}} 0$  for all deterministic Hermitian matrix  $D_n$  and deterministic vectors  $d_{i,n}$  of bounded norms.

<sup>4</sup>Here and in the rest of the article, the phrase “at macroscopic distance” will mean that, as  $p \rightarrow \infty$ , as per Assumption 1, the distance between the objects under consideration stays lower bounded by a positive constant.

Denoting  $g^\circ(z) = c_0 \sum_{a=1}^k c_a g_a(z)$  and  $\mathbb{P}_p = \frac{1}{p} \sum_{i=1}^p \delta_{\lambda_i^p}$  the empirical spectral measure of  $PW^\top WP$ ,

$$\mathbb{P}_p - \bar{\mathbb{P}}_p \xrightarrow{\mathcal{D}} 0$$

almost surely, with  $\bar{\mathbb{P}}_p$  the probability measure having Stieltjes transform  $g^\circ(z)$ . Besides, denoting  $\mathcal{S}_p$  the support of  $\bar{\mathbb{P}}_p$  and  $\mathcal{G}_p = \{0\} \cup \{x \notin \mathcal{S}_p \mid g^\circ(x) = 0\}$ ,

$$\max_{1 \leq i \leq p} \text{dist}(\lambda_i^p, \mathcal{S}_p \cup \mathcal{G}_p) \rightarrow 0$$

almost surely.

Lemma 1 is a mere extension of Propositions 3 and 4 of the companion paper [BC15], where more details are given about the functions  $g_a$ . It is also an extension of standard random matrix results such as [SB95, SC95, CH14] to more involved structures (studied earlier in e.g., [CDS11, WCRS01]). This result mainly states that the bilinear forms as well as the linear statistics of the resolvent of  $PW^\top WP$  asymptotically tend to be deterministic in a controlled manner. Note in passing that, as proved in [BC15],  $g_1(z), \dots, g_k(z)$  can be evaluated by a provably converging fixed-point algorithm, for all  $z \in \mathbb{C} \setminus \mathcal{S}_p$ .

**Remark 1** (On  $\mathcal{S}_p$  and  $\mathcal{G}_p$ ). Since  $g^\circ(x)$  is strictly increasing on  $\mathbb{R} \setminus \mathcal{S}_p$ , in between each connected component of  $\mathcal{S}_p$ ,  $g^\circ(x) = 0$  can only have one solution. Besides,  $g^\circ(x) \uparrow 0$  as  $x \rightarrow \infty$  and  $g^\circ(x) \downarrow 0$  as  $x \rightarrow -\infty$ , so that, denoting  $\mathcal{S}_p^-$  and  $\mathcal{S}_p^+$  the leftmost and rightmost edges of  $\mathcal{S}_p$ ,  $(\mathcal{G}_p \setminus \{0\}) \cap (-\infty, \mathcal{S}_p^-) = \emptyset$  and  $\mathcal{G}_p \cap (\mathcal{S}_p^+, \infty) = \emptyset$ . In particular, if  $\mathcal{S}_p$  is constituted of a single connected component (as shall often be the case in practice when  $c_0 > 1$ ), then  $\mathcal{G}_p = \{0\}$ .

Besides, if  $g_1 = \dots = g_k = g^\circ$  (a scenario that shall be thoroughly studied in the course of this article), from [SC95], it appears that, for all  $x, y \in \mathbb{R} \setminus \mathcal{S}_p$ , if  $x > y > 0$ , then  $0 > g^\circ(x) > g^\circ(y)$ , and thus  $\mathcal{G}_p = \{0\}$ .

With those notations and remarks at hand, we are now ready to introduce our main results.

## 4. MAIN RESULTS

Before delving into the investigation of the eigenvalues and eigenvectors of  $L$ , recall from Theorem 1 that the behavior of  $L$  is strikingly different if  $f'(\tau)$  is away from zero or if instead  $f'(\tau) \rightarrow 0$ . We shall then systematically study both cases independently. In practice, if  $f'(\tau)$  only has limit points at zero (so is neither away nor converges to zero), then the following study will be valid up to extracting subsequences of  $p$ .

**4.1. Isolated Eigenvalues.** Assume first that  $f'(\tau)$  is away from zero, i.e.,  $\liminf_p |f'(\tau)| > 0$ . In order to study the isolated eigenvalues and associated eigenvectors of the model, we follow standard random matrix approaches as developed in e.g., [BN12, HLMNV13]. That is, to determine the isolated eigenvalues, we shall solve

$$\det(PW^\top WP + UBU^\top - \rho I_n) = 0$$

for  $\rho$  away from  $\mathcal{S}_p \cup \mathcal{G}_p$  defined in Lemma 1. Such  $\rho$  ensure the correct behavior of the resolvent  $Q_\rho = (PW^\top WP - \rho I_n)^{-1}$ . Factoring out  $PW^\top WP - \rho I_n$  and using Sylvester's identity, the above equation is then equivalent to

$$\det(BU^\top Q_\rho U + I_{2k+1}) = 0.$$

By Lemma 1 (and some arguments to handle the dependence between  $U$  and  $Q_\rho$ ),  $U^\top Q_\rho U$  tends to be deterministic in the large  $n$  limit, and thus, using a perturbation approach along with the argument principle, we find that the isolated eigenvalues of  $PW^\top WP + UBU^\top$  tend to be the values of  $\rho$  for which  $BU^\top Q_\rho U + I_{2k+1}$  has a zero eigenvalue, the multiplicity of  $\rho$  being asymptotically the same as that of the aforementioned zero eigenvalue.

All calculus made, we have the following first main results.

**Theorem 2** (Isolated eigenvalues,  $f'(\tau)$  away from zero). *Let Assumptions 1 and 2 hold and define the  $k \times k$  matrix*

$$G_z = h(\tau, z)I_k + D_{\tau, z}\Gamma_z$$

where

$$\begin{aligned} D_{\tau, z} &\triangleq -zh(\tau, z)M^\top \bar{Q}_z M - h(\tau, z)\frac{f''(\tau)}{f'(\tau)}T + \left(\frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)}\right)tt^\top \\ \Gamma_z &\triangleq \mathcal{D}\{c_a g_a(z)\}_{a=1}^k - \left\{ \frac{c_a g_a(z)c_b g_b(z)}{\sum_{i=1}^k c_i g_i(z)} \right\}_{a, b=1}^k \\ h(\tau, z) &\triangleq 1 + \left(\frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)}\right) \sum_{a=1}^k c_a g_a(z) \frac{2}{p} \text{tr} C_a^2. \end{aligned}$$

Besides, denoting  $\mathcal{H}_p = \{x \in \mathbb{R} \mid h(\tau, x) = 0\}$ , define  $\mathcal{S}'_p \triangleq \mathcal{S}_p \cup \mathcal{G}_p \cup \{F(\tau)\} \cup \mathcal{H}_p$ . Let  $\rho$  be at macroscopic distance from  $\mathcal{S}'_p$  and be such that  $G_\rho$  has a zero eigenvalue of multiplicity  $m_\rho$ . Then there exists  $\lambda_j^p \geq \dots \geq \lambda_{j+m_\rho-1}^p$  ( $j$  may depend on  $p$ ) eigenvalues of  $L$  such that

$$\max_{0 \leq i \leq m_\rho-1} \left| \left( -2\frac{f'(\tau)}{f(\tau)}\rho + 2\frac{f'(\tau)}{f(\tau)}F(\tau) \right) - \lambda_{j+i}^p \right| \xrightarrow{\text{a.s.}} 0.$$

As it shall turn out, the isolated eigenvalues identified in Theorem 2 are the only ones of practical interest for spectral clustering as they are strongly related to  $J$ . However, some other isolated eigenvalues may be found which we discuss here for completion (and thoroughly in the proof section).

**Remark 2** (Other isolated eigenvalues of  $L$ ,  $f'(\tau)$  away from zero). Under the conditions of Theorem 2, if there exists a  $\rho_+$  in a neighborhood of  $\mathcal{H}_p$  such that  $\det H_{\rho_+} = 0$  with  $H_z$  defined in (23), then there exist  $\lambda_j^p \geq \dots \geq \lambda_{j+m_{\rho_+}-1}^p$  eigenvalues of  $L$  satisfying

$$\max_{0 \leq i \leq m_{\rho_+}} \left| \left( -2\frac{f'(\tau)}{f(\tau)}\rho + 2\frac{f'(\tau)}{f(\tau)}F(\tau) \right) - \lambda_{j+i}^p \right| \xrightarrow{\text{a.s.}} 0$$

where  $m_\rho \geq 1$  is the multiplicity of zero as an eigenvalue of  $H_\rho$ . Note in particular that, if  $t = 0$  and  $\mathcal{H}_p \cap \mathcal{S}_p^c \neq \emptyset$ , then such  $\rho_+$  exists. Figure 7, commented later, provides an example where a  $\rho_+$  is found amongst the other isolated eigenvalues of  $L$  (emphasized here in blue). Note that  $\rho_+$  only depends on a weighted sum of the  $\text{tr } C_i^2$  and may even exist when  $M = 0$ ,  $t = 0$ , and  $T = 0$ . Intuitively, this already suggests that  $\rho_+$  is only marginally related to the spectral clustering problem.

Operating the variable change  $z \mapsto -f'(\tau)z/f(\tau)$  in the expression of  $G_z$ , we may form the matrix  $G_{-f(\tau)z/(2f'(\tau))}$  the null eigenvalues of which are achieved for the values  $-2f'(\tau)\rho/f(\tau)$  where  $\rho$  is defined in Theorem 2. While  $G_z$  is ill-defined as  $f'(\tau) \rightarrow 0$ ,  $G_{-f(\tau)z/(2f'(\tau))}$  has a non trivial limit, which we denote  $G_z^0$  and that allows for an extension of Theorem 2 to the case  $f'(\tau) \rightarrow 0$ . In particular,  $g_a(-f(\tau)z/(2f'(\tau)))$  behaves similar to  $2f'(\tau)/(c_0f(\tau)z)$  for all  $a \in \{1, \dots, k\}$  and we have the following simpler expression.

**Theorem 3** (Isolated eigenvalues,  $f'(\tau) \rightarrow 0$ ). *Let Assumptions 1–2 hold and define the  $k \times k$  matrix<sup>5</sup>*

$$G_z^0 \triangleq h^0(\tau, z)I_k + D_{\tau, z}^0 \Gamma_z^0$$

with

$$\begin{aligned} \Gamma_z^0 &\triangleq \frac{-1}{c_0 z} \mathcal{D}(c) \\ D_{\tau, z}^0 &\triangleq h^0(\tau, z) \frac{2f''(\tau)}{f(\tau)} T + \frac{f''(\tau)}{f(\tau)} tt^\top \\ h^0(\tau, z) &\triangleq 1 - \frac{1}{z} \frac{f''(\tau)}{f(\tau)} \sum_{i=1}^k \frac{c_i}{c_0} \frac{2}{p} \text{tr } C_i^2. \end{aligned}$$

Define also  $\mathcal{H}_p^0 = \{x \mid h^0(\tau, x) = 0\}$ . Then, for  $\rho^0$  at macroscopic distance from  $\mathcal{H}_p^0 \cup \{(f(0) - f(\tau))/f(\tau)\}$  such that  $G_{\rho^0}^0$  has a zero eigenvalue of multiplicity  $m_\rho^0$ , there exist  $\lambda_j^p \geq \dots \geq \lambda_{j+m_\rho^0-1}^p$  ( $j$  may depend on  $p$ ) eigenvalues of  $L$  satisfying

$$\max_{0 \leq i \leq m_\rho^0 - 1} \left| \left( \rho^0 + \frac{f(0) - f(\tau)}{f(\tau)} \right) - \lambda_{j+i}^p \right| \xrightarrow{\text{a.s.}} 0.$$

**Remark 3** (Full spectrum of  $L$ ,  $f'(\tau) \rightarrow 0$ ). Similar to Remark 2, under the conditions of Theorem 3, if there exists  $\rho_+^0$  in a small neighborhood of  $\mathcal{H}_p^0$  such that  $\det H_{\rho_+^0}^0 = 0$  with  $H_z^0$  defined in (25), then there exist  $\lambda_j^p \geq \dots \geq \lambda_{j+m_{\rho_+^0}^0-1}^p$ , eigenvalues of  $L$  satisfying

$$\max_{0 \leq i \leq m_{\rho_+^0}^0 - 1} \left| \left( \rho_+^0 + \frac{f(0) - f(\tau)}{f(\tau)} \right) - \lambda_{j+i}^p \right| \xrightarrow{\text{a.s.}} 0$$

<sup>5</sup>There,  $h^0(\tau, z) = \lim_{f'(\tau) \rightarrow 0} h(\tau, z)$ ,  $\Gamma_z^0 = \lim_{f'(\tau) \rightarrow 0} -f(\tau)/(2f'(\tau))\Gamma_{-f(\tau)z/(2f'(\tau))}$ , and  $D_{\tau, z}^0 = \lim_{f'(\tau) \rightarrow 0} -\frac{2f'(\tau)}{f(\tau)} D_{\tau, -f(\tau)z/(2f'(\tau))}$ .

with  $m_{\rho_+^0}$  the multiplicity of the zero eigenvalue of  $H_{\rho_+^0}^0$ . Along with the eigenvalue  $n$  and the eigenvalues identified in Theorem 3, this enumerates all the (asymptotic) isolated eigenvalues of  $L$ .

The two theorems above exhibit quite involved expressions that do not easily allow for intuitive interpretations. We shall see in Section 5 that these results greatly simplify in some specific scenarios of practical interest. We may nonetheless already extrapolate some elementary properties.

**Remark 4** (Large and small eigenvalues of  $D_{\tau,\rho}$ ). If an eigenvalue of  $D_{\tau,\rho}$  diverges to infinity as  $n, p \rightarrow \infty$ , by the boundedness property of Stieltjes transforms, we find that  $h(\tau, \rho)$  and  $\Gamma_\rho$  remain bounded and, thus, the value  $\rho$  cancelling the determinant of  $G_\rho$  must go to infinity as well. This is the expected behavior of spiked models. This implies in particular that, if, for some  $i, j$ ,  $\|\mu_i^\circ\| \rightarrow \infty$ , or  $t_i \rightarrow \infty$ , or  $T_{ij} \rightarrow \infty$ , as  $n, p \rightarrow \infty$  slowly (thus disrupting from our assumptions), there will exist an asymptotically unbounded eigenvalue in the spectrum of  $L$  (aside from the eigenvalue  $n$ ). On the opposite, if for all  $i, j$  those quantities vanish as  $n, p \rightarrow \infty$ , then  $D_{\tau,z}$  is essentially zero in the limit, and thus, aside from the  $\rho$ 's solution to  $h(\tau, \rho) = 0$  (and from the eigenvalue  $n$ ), no isolated eigenvalue can be found in the spectrum of  $L$ .

**Remark 5** (About the Kernel). It appears clearly from Theorem 3 that, as  $f'(\tau) = 0$ , the matrix  $M$  does not contribute to the spectrum of  $L$  and thus the  $\mu_i$ 's can be anticipated not to play any role in the resulting spectral clustering methods. Similarly, from Theorem 2, if  $f''(\tau) = 0$ , the cross-variances  $\frac{1}{p} \text{tr} C_i^\circ C_j^\circ$  will not intervene and thus cannot be discriminated over. Finally, letting  $\frac{5f'(\tau)}{8f(\tau)} = \frac{f''(\tau)}{2f'(\tau)}$  discards the impact of the traces  $\frac{1}{\sqrt{p}} \text{tr} C_i^\circ$ . This has interesting consequences in practice if one aims at discriminating data upon some specific properties.

**4.2. Eigenvectors.** Let us now turn to the central aspect of the article: the eigenvectors of  $L$  (being the same as those of  $L'$ , up to reordering).

To start with, note that, in both theorems, the eigenvalue  $n$  is associated with the eigenvector  $D^{\frac{1}{2}}1_n$ . Since the eigenvector is completely explicit (which shall not be the case of the next eigenvectors), it is fairly easy to study independently without resorting to any random matrix analysis. Precisely, we have the following result for it.

**Proposition 1** (Eigenvector  $D^{\frac{1}{2}}1_n$ ). *Let Assumptions 1–2 hold true. Then*

$$\frac{D^{\frac{1}{2}}1_n}{\sqrt{1_n^\top D 1_n}} = \frac{1_n}{\sqrt{n}} + \frac{1}{n\sqrt{c_0}} \frac{f'(\tau)}{2f(\tau)} \left[ \{t_a 1_{n_a}\}_{a=1}^k + \mathcal{D} \left\{ \sqrt{\frac{2}{p} \text{tr}(C_a^2)} 1_{n_a} \right\}_{a=1}^k \varphi \right] + o(n^{-1})$$

with  $\varphi \sim \mathcal{N}(0, I_n)$  and  $o(n^{-1})$  is meant entry-wise.



**Remark 6** (Precisions on  $D^{\frac{1}{2}}1_n$ ). We can precise the value of  $\varphi$  as follows. Recalling the definition (5) of  $\psi$ ,

$$\frac{D^{\frac{1}{2}}1_n}{\sqrt{1_n^\top D 1_n}} = \frac{1_n}{\sqrt{n}} + \frac{1}{n\sqrt{c_0}} \frac{f'(\tau)}{2f(\tau)} \left[ \{t_a 1_{n_a}\}_{a=1}^k + \psi \right] + o(n^{-1})$$

almost surely.

Note that the eigenvector  $D^{\frac{1}{2}}1_n$  may then be used directly for clustering, with increased efficiency when the entries  $t_a = \frac{1}{\sqrt{p}} \text{tr } C_a^\circ$  of  $t$  grow large for fixed  $\frac{1}{p} \text{tr } C_a^2$ . But the eigenvector (asymptotically) carries no information concerning  $M$  or  $T$  and is in particular of no use if all covariances  $C_i$  have the same trace.

For the other isolated eigenvectors, the study is much more delicate as we do not have an explicit expression as in Proposition 1. Instead, by statistical interchangeability of the class- $\mathcal{C}_a$  entries of, say, the  $i$ -th isolated eigenvector  $\hat{u}_i$  of  $L$ , we may write

$$\hat{u}_i = \sum_{a=1}^k \alpha_a^i \frac{j_a}{\sqrt{n_a}} + \sigma_a^i \omega_a^i \quad (6)$$

where  $\omega_a^i \in \mathbb{R}^n$  is a random vector, orthogonal to  $j_a$ , of unit norm, supported on the indices of  $\mathcal{C}_a$ , where its entries are identically distributed. The scalars  $\alpha_a^i \in \mathbb{R}$  and  $\sigma_a^i \geq 0$  are the coefficients of alignment to  $j_a$  and the standard deviation of the fluctuations around  $\alpha_a^i \frac{j_a}{\sqrt{n_a}}$ , respectively.

Assuming when needed unit multiplicity for the eigenvalue associated with  $\hat{u}_i$ , our objective is now threefold:

- (1) *Class-wise Eigenvector Means*. We first wish to retrieve the values of the  $\alpha_a^i$ 's. For this, note that

$$\alpha_a^i = \hat{u}_i^\top \frac{j_a}{\sqrt{n_a}}.$$

We shall evaluate these quantities by obtaining an estimator for the  $k \times k$  matrix  $\frac{1}{p} J^\top \hat{u}_i^\top \hat{u}_i^\top J$ . The diagonal entries of the latter will allow us to retrieve  $|\alpha_a^i|$  and the off-diagonal entries will be used to decide on the signs of  $\alpha_1^i, \dots, \alpha_k^i$  (up to a convention in the sign of  $\hat{u}_i$ ).

- (2) *Class-wise Eigenvector Fluctuations*. Next, we wish to retrieve the  $\sigma_a^i$ 's. To this end, using the knowledge of  $\alpha_a^i$ , it is easily seen that it suffices to evaluate  $\|\mathcal{D}(j_a)\hat{u}_i\|^2$ . Indeed,

$$(\sigma_a^i)^2 = \left( \hat{u}_i - \alpha_a^i \frac{j_a}{\sqrt{n_a}} \right)^\top \mathcal{D}(j_a) \left( \hat{u}_i - \alpha_a^i \frac{j_a}{\sqrt{n_a}} \right) = \hat{u}_i^\top \mathcal{D}(j_a) \hat{u}_i - (\alpha_a^i)^2.$$

This quantity will be evaluated using a similar approach as that developed to evaluate the  $\alpha_a^i$ 's.

- (3) *Class-wise Cross Correlations.* Finally, our last objective is to evaluate the correlation

$$\sigma_a^{i,j} \triangleq \left( \hat{u}_i - \alpha_a^i \frac{j_a}{\sqrt{n_a}} \right)^\top \mathcal{D}(j_a) \left( \hat{u}_j - \alpha_a^j \frac{j_a}{\sqrt{n_a}} \right) = \hat{u}_i^\top \mathcal{D}(j_a) \hat{u}_j - \alpha_a^i \alpha_a^j$$

between the fluctuations of two eigenvectors indexed by  $i$  and  $j$  on the subblock indexing  $\mathcal{C}_a$ . For this, it is sufficient to exploit the previous estimates and to evaluate the quantities  $\hat{u}_i^\top \mathcal{D}(j_a) \hat{u}_j$ . But, to this end, for lack of a better approach, we shall resort to estimating the more involved object  $\frac{1}{p} J^\top \hat{u}_i \hat{u}_i^\top \mathcal{D}(j_a) \hat{u}_j \hat{u}_j^\top J$ , from which  $\hat{u}_i^\top \mathcal{D}(j_a) \hat{u}_j$  can be extracted by division of any entry  $m, l$  by  $\alpha_m^i \alpha_l^j$ . The specific cross correlation between any eigenvector and the eigenvector  $D^{\frac{1}{2}} \mathbf{1}_n$  will be treated independently.

The three aforementioned steps are successively derived in the next three sections, starting with the evaluation of the coefficients  $\alpha_a^i$ .

4.2.1. *Eigenvector Means ( $\alpha_a^i$ ).* Consider the case where  $f'(\tau)$  is away from zero. First observe that, for  $\lambda_j^p, \dots, \lambda_{j+m_\rho-1}^p$  a group of the identified isolated eigenvalues of  $L$  all converging to the same limit (as per Theorem 2 or Remark 2), the corresponding eigenspace is (asymptotically) the same as the eigenspace associated with the corresponding deterministic eigenvalue  $\rho$  in the spectrum of  $PW^\top WP + UBU^\top$ . Denoting  $\hat{\Pi}_\rho$  a projector on the former eigenspace, we then have, by Cauchy's formula and our approximation of Theorem 1,

$$\frac{1}{p} j_a^\top \hat{\Pi}_\rho j_b = -\frac{1}{2\pi i} \oint_{\gamma_\rho} \frac{1}{p} j_a^\top (PW^\top WP + UBU^\top - zI_n)^{-1} j_b dz + o(1) \quad (7)$$

for a small (positively oriented) closed path  $\gamma_\rho$  circling around  $\rho$ , this being valid for all large  $n$ , almost surely.

Using matrix inversion lemmas, the right-hand side of (7) can be worked out and reduced to an expression involving the matrix  $G_z$  of Theorem 2. It then remains to perform a residue calculus on the final formulation which then leads to the following result.

**Theorem 4** (Eigenvector projections,  $f'(\tau)$  away from zero). *Let Assumptions 1 and 2 hold and assume  $f'(\tau)$  away from zero. Let also  $\lambda_j^p, \dots, \lambda_{j+m_\rho-1}^p$  be a group of isolated eigenvalues of  $L$  and  $\rho$  the associated deterministic approximate (of multiplicity  $m_\rho$ ) as per Theorem 2, and assume that  $\rho$  is uniformly away from any other eigenvalue retrieved in Theorem 2. Further denote  $\hat{\Pi}_\rho$  the projector on the eigenspace of  $L$  associated to these eigenvalues. Then,*

$$\frac{1}{p} J^\top \hat{\Pi}_\rho J = -h(\tau, \rho) \Gamma_\rho \Xi_\rho + o(1)$$

where

$$\Xi_\rho \triangleq \sum_{i=1}^{m_\rho} \frac{(V_{r,\rho})_i (V_{l,\rho})_i^\top}{(V_{l,\rho})_i^\top G'_\rho (V_{r,\rho})_i}$$

with  $V_{r,\rho} \in \mathbb{C}^{k \times m_\rho}$  and  $V_{l,\rho} \in \mathbb{C}^{k \times m_\rho}$  respectively sets of right and left eigenvectors of  $G_\rho$  associated with the eigenvalue zero, and  $G'_\rho$  the derivative of  $G_z$  along  $z$  evaluated at  $z = \rho$ .

Similarly, when  $f'(\tau) \rightarrow 0$ , we obtain, with the same limiting approach as for Theorem 3, the following estimate.

**Theorem 5** (Eigenvector projections,  $f'(\tau) \rightarrow 0$ ). *Let Assumptions 1 and 2 hold and assume  $f'(\tau) \rightarrow 0$ . Let  $\lambda_j^p, \dots, \lambda_{j+m_\rho-1}^p$  be a group of isolated eigenvalues of  $L$  and  $\rho^0$  the corresponding approximate (of multiplicity  $m_{\rho^0}$ ) defined in Theorem 3, and assume that  $\rho^0$  is uniformly away from any other eigenvalue retrieved in Theorem 3. Further denote  $\hat{\Pi}_{\rho^0}$  the projector on the eigenspace associated to these eigenvalues. Then,*

$$\frac{1}{p} J^\top \hat{\Pi}_{\rho^0} J = -h^0(\tau, \rho) \Gamma_\rho^0 \Xi_{\rho^0}^0 + o(1)$$

almost surely, where<sup>6</sup>

$$\Xi_{\rho^0}^0 \triangleq \sum_{i=1}^{m_\rho} \frac{(V_{r,\rho}^0)_i (V_{l,\rho}^0)_i^\top}{(V_{l,\rho}^0)_i^\top G_\rho^{0'} (V_{r,\rho}^0)_i}$$

with  $V_{r,\rho}^0 \in \mathbb{C}^{k \times m_\rho}$  and  $V_{l,\rho}^0 \in \mathbb{C}^{k \times m_\rho}$  sets of right and left eigenvectors of  $G_\rho^0$  associated with the eigenvalue zero, and  $G_\rho^{0'}$  the derivative of  $G_z^0$  along  $z$  evaluated at  $z = \rho$ .

Correspondingly to Remarks 2 and 3, we have the following complementary result for the isolated eigenvalues satisfying  $h(\tau, \rho) \rightarrow 0$ .

**Remark 7** (Irrelevance of the eigenvectors with  $h(\tau, \rho) \rightarrow 0$ ). In addition to Theorem 4, it can be shown (see the proof section) that, if  $\rho_+$  is an isolated eigenvalue as per Remark 2 having multiplicity one, then with similar notations as above

$$\frac{1}{p} J^\top \hat{\Pi}_{\rho_+} J = o(1)$$

almost surely. The same holds for  $\rho_+^0$  from Remark 3. As such, as far as spectral clustering is concerned, the eigenvectors forming the (dimension one) eigenspace  $\hat{\Pi}_{\rho_+}$  cannot be used for unsupervised classification.

From this remark, we shall from now on adopt the following convention. The finitely many isolated eigenvalue-eigenvector pairs  $(\rho_i, \hat{u}_i)$  of  $L$ , excluding those for which (at least on a subsequence)  $h(\tau, \rho) \rightarrow 0$ , will be denoted in the order  $\rho_1 \geq \rho_2 \geq \dots$ , with possibly equal values of  $\rho_i$ 's to account for multiplicity. In particular,  $\hat{u}_1 = \frac{D^{\frac{1}{2}} \mathbf{1}_n}{\sqrt{\mathbf{1}_n^\top D \mathbf{1}_n}}$ . The eigenvalue-eigenvector pairs  $(\rho, \hat{u})$  for which  $h(\tau, \rho) \rightarrow 0$  will no longer be listed.

<sup>6</sup>There,  $\Xi_\rho^0 = \lim_{f'(\tau) \rightarrow 0} -\frac{2f'(\tau)}{f(\tau)} \Xi_{-f(\tau)\rho/(2f'(\tau))}$ .

**Remark 8** (Eigenspace alignment). An alignment metric between the span of  $\hat{\Pi}_\rho$  and the sought-for subspace  $\text{span}(j_1, \dots, j_k)$  may be given by

$$0 \leq \text{tr} \left( \mathcal{D}(c^{-1}) \frac{1}{n} J^\top \hat{\Pi}_\rho J \right) \leq m_\rho$$

with  $(c^{-1})_i = 1/c_i$  and corresponds in practice to the extent to which the eigenvectors of  $L$  are close to linear combinations of the base vectors  $\frac{j_1}{\sqrt{n_1}}, \dots, \frac{j_k}{\sqrt{n_k}}$ .

Remark 8 may be straightforwardly applied to observe a peculiar phenomenon, when  $M = 0$ ,  $t = 0$  and the kernel is such that  $f'(\tau) \rightarrow 0$ .

**Remark 9** (Only  $T$  case and  $f'(\tau) \rightarrow 0$ ). If  $M = 0$ ,  $t = 0$ , and  $f'(\tau) \rightarrow 0$ , note that  $G_x^0 = h^0(\tau, x)I_k + D_{\tau,x}^0 \Gamma_x^0$  satisfies  $G_x^{0'} = \frac{h^0(\tau,x)'}{h^0(\tau,x)} G_x^0 \Gamma_x^0 - \frac{1}{x} (h^0(\tau, x)I_k - G_x^0)$  (the rightmost term arising from  $D_{\tau,x}^0 \Gamma_x^0 = G_x^0 - h^0(\tau, x)I_k$ ), so that, for  $x = \rho^0$ , since  $(V_{l,\rho^0}^0)^\top G_{\rho^0}^{0'} = 0$ , we find that

$$\text{tr} \left( \mathcal{D}(c^{-1}) \frac{1}{n} J^\top \hat{\Pi}_{\rho^0} J \right) = m_{\rho^0} + o(1)$$

almost surely. It shall also be seen through an example in Section 5 that this is no longer true in general if  $f'(\tau)$  is away from zero. As such, there is an asymptotic perfect alignment in the regime under consideration if only  $T$  is non vanishing, provided one takes  $f'(\tau) \rightarrow 0$ . In this case, it is theoretically possible, as  $n, p \rightarrow \infty$ , to correctly cluster all but a vanishing proportion of the  $x_i$ .

**Remark 10** (Evaluation of  $\alpha_a^i$ ). Let us consider the case where  $\rho$  is an isolated eigenvalue of unit multiplicity with associated eigenvector  $\hat{u}_i$  (thus  $\hat{\Pi}_\rho = \hat{u}_i \hat{u}_i^\top$ ). According to (6), we may now obtain the expression of  $\alpha_1^i, \dots, \alpha_k^i$  as follows:

- (1)  $\sqrt{\frac{1}{n_1} [J^\top \hat{u}_i \hat{u}_i^\top J]_{11}} = |\alpha_1^i|$  allows the retrieval of  $\alpha_1^i$  up to a sign shift. We may conventionally call this nonnegative value  $\alpha_1^i$  (if this turns out to be zero, we may proceed similarly with entry (2, 2) instead).
- (2) for all  $1 < a \leq k$ ,  $\sqrt{\frac{1}{n_a} [J^\top \hat{u}_i \hat{u}_i^\top J]_{aa}} \times \text{sign}([J^\top \hat{u}_i \hat{u}_i^\top J]_{1a})$  provides  $\alpha_a^i$  up to a sign shift which is consistent with the aforementioned convention, and thus we may redefine  $\alpha_a^i = \sqrt{\frac{1}{n_a} [J^\top \hat{u}_i \hat{u}_i^\top J]_{aa}} \times \text{sign}([J^\top \hat{u}_i \hat{u}_i^\top J]_{1a})$ .

4.2.2. *Eigenvector Fluctuations* ( $\sigma_a^i$ ). Let us now turn to the evaluation of the fluctuations of the isolated eigenvectors of  $L$  around their projection onto  $\text{span}(j_1, \dots, j_k)$ . As recalled earlier, our objective is to evaluate quantities of the type  $\hat{u}_i^\top \mathcal{D}(j_a) \hat{u}_i$  for  $\hat{u}_i$  an isolated eigenvector of  $L$  for which Theorem 4 provided a non trivial result. That is, we shall from now on disregard the isolated eigenvalues satisfying  $h(\tau, \rho) \rightarrow 0$ . First remark that, from Proposition 1, we already know the class-wise fluctuations of the eigenvector  $D^{\frac{1}{2}} 1_n$  (these are proportional to  $\frac{1}{p} \text{tr} C_a^2$  for class  $\mathcal{C}_a$ ), and thus we may simply work on the remaining eigenvectors.

Similar to Section 4.2.1 (and with the same notations), assuming  $f'(\tau)$  away from zero first, we follow a Cauchy integral approach by remarking that

$$\hat{u}_i^\top \mathcal{D}(j_a) \hat{u}_i = \text{tr } \mathcal{D}(j_a) \hat{u}_i \hat{u}_i^\top$$

and thus we can to the least evaluate

$$\text{tr } \mathcal{D}(j_a) \hat{\Pi}_\rho = -\frac{1}{2\pi i} \oint_{\gamma_\rho} \text{tr} (PW^\top WP + UBU^\top - zI_n)^{-1} + o(1)$$

this holding for all large  $n$ , almost surely. In case of unit multiplicity for  $\rho$ , we obtain the expected  $\hat{u}_i^\top \mathcal{D}(j_a) \hat{u}_i$ .

Before presenting the result, we need an additional technical result, mostly borrowed from [BC15]. Also, for readability, from now on, we shall use the shortcut notation  $\mathcal{D}_a \triangleq \mathcal{D}(j_a)$ .

**Lemma 2** (Further Deterministic Equivalents). *Under the conditions and notations of Lemma 1, for  $z_1, z_2$  (sequences of) complex numbers at macroscopic distance from the eigenvalues of  $PW^\top WP$ , as  $n \rightarrow \infty$ ,*

$$\begin{aligned} Q_{z_1} \mathcal{D}_a Q_{z_2} &\leftrightarrow \bar{Q}_{z_1} \mathcal{D}_a \bar{Q}_{z_2} + \sum_{b=1}^k R_{ab}^{z_1 z_2} \bar{Q}_{z_1} P \mathcal{D}_b P \bar{Q}_{z_2} \\ \tilde{Q}_{z_1} C_a \tilde{Q}_{z_2} &\leftrightarrow \bar{\tilde{Q}}_{z_1} C_a \bar{\tilde{Q}}_{z_2} + \sum_{b=1}^k R_{ba}^{z_1 z_2} \bar{\tilde{Q}}_{z_1} C_b \bar{\tilde{Q}}_{z_2} \\ \frac{1}{p} \tilde{Q}_{z_1} W \mathcal{D}_a W^\top \tilde{Q}_{z_2} &\leftrightarrow z_1 z_2 c_0 c_a g_a(z_1) g_a(z_2) \bar{\tilde{Q}}_{z_1} C_a \bar{\tilde{Q}}_{z_2} \end{aligned}$$

where  $R^{z_1 z_2} \in \mathbb{R}^{k \times k}$  is defined as  $R^{z_1 z_2} = (I_k - \Omega^{z_1 z_2})^{-1} \Omega^{z_1 z_2}$ , with

$$\Omega_{a,b}^{z_1 z_2} \triangleq z_1 z_2 c_0 c_b g_a(z_1) g_a(z_2) \frac{1}{p} \text{tr } C_a \bar{\tilde{Q}}_{z_1} C_b \bar{\tilde{Q}}_{z_2}.$$

In particular, as a consequence of Lemma 2, we have the following identities

$$\begin{aligned} \frac{1}{p} J^\top Q_{z_1} \mathcal{D}_a Q_{z_2} J &= E_{a; z_1 z_2}^J + o(1) \\ M^\top W Q_{z_1} \mathcal{D}_a Q_{z_2} W^\top J &= E_{a; z_1 z_2}^M + o(1) \\ \psi^\top Q_{z_1} \mathcal{D}_a Q_{z_2} \psi &= E_{a; z_1 z_2}^\psi + o(1) \end{aligned}$$

almost surely, where we define

$$\begin{aligned}
E_{a; z_1 z_2}^J &\triangleq \frac{1}{p} J^\top \bar{Q}_{z_1} \left[ \mathcal{D}_a + \sum_{b=1}^k R_{ab}^{z_1 z_2} P \mathcal{D}_b P \right] \bar{Q}_{z_2} J \\
E_{a; z_1 z_2}^M &\triangleq M^\top \bar{Q}_{z_1} \left[ \sum_{b=1}^k (R^{z_1 z_2} + I_k)_{ba} C_b \right] \bar{Q}_{z_2} M \\
E_{a; z_1 z_2}^\psi &\triangleq c_0 \sum_{b=1}^k c_b g_b(z_1) g_b(z_2) \frac{2}{p} \operatorname{tr} C_b^2 (R^{z_1 z_2} + I_k)_{ab}
\end{aligned} \tag{8}$$

**Remark 11** (About  $R^{z_1 z_2}$ ). The matrix  $R^{z_1 z_2}$  is strongly related to the derivative of the  $g_a(z)$ 's. In particular, we have that

$$\mathcal{D}(c) \{g'_a(z)\}_{a=1}^k = c_0 (R^{zz} + I_k) \mathcal{D}(c) \{g_a^2(z)\}_{a=1}^k = c_0 (I_k - \Omega^{zz})^{-1} \mathcal{D}(c) \{g_a^2(z)\}_{a=1}^k.$$

With this lemma at hand, we then have the following result on the class-wise fluctuations of the isolated eigenvectors.

**Theorem 6** (Eigenvector Fluctuations,  $f'(\tau)$  away from zero). *Under the assumptions and notations of Theorem 4 and (8), as  $n \rightarrow \infty$ ,*

$$\operatorname{tr} \mathcal{D}_a \hat{\Pi}_\rho = \operatorname{tr} \Xi_\rho \left( D_{\tau, \rho} E_{a; \rho \rho}^J - h(\tau, \rho) E_{a; \rho \rho}^M \Gamma_\rho - \frac{E_{a; \rho \rho}^\psi \left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right)^2}{h(\tau, \rho)} t t^\top \Gamma_\rho \right) + o(1)$$

almost surely.

Similarly, if  $f'(\tau) \rightarrow 0$ , we obtain

**Theorem 7** (Eigenvector Fluctuations,  $f'(\tau) \rightarrow 0$ ). *Under the assumptions and notations of Theorem 5, as  $n \rightarrow \infty$ ,*

$$\operatorname{tr} \mathcal{D}_a \hat{\Pi}_\rho^0 = \operatorname{tr} \Xi_\rho^0 \left( D_{\tau, \rho}^0 E_{a; \rho \rho}^{0; J} - \frac{E_{a; \rho \rho}^{0; \psi} \left( \frac{f''(\tau)}{f(\tau)} \right)^2}{h^0(\tau, \rho)} t t^\top \Gamma_\rho^0 \right) + o(1)$$

almost surely, where  $[E_{a; z_1 z_2}^{0; J}]_{ij} \triangleq \frac{c_a}{c_0} \frac{1}{z_1 z_2} \delta_{ia} \delta_{ja}$  and  $E_{a; z_1 z_2}^{0; \psi} \triangleq \frac{c_a}{c_0} \frac{2}{p} \operatorname{tr} C_a^2 \frac{1}{z_1 z_2}$ .

This result is much more involved than those previously obtained and does not lead itself to much insight. Nonetheless, we shall see in Section 5 that this greatly simplifies when considering special application cases.

**Remark 12** (Relation to  $\sigma_a^i$ ). Under the model (6), we may provide an estimate for the class  $\mathcal{C}_a$  fluctuation  $\sigma_a^i$  of the eigenvector  $\hat{u}_i$ , with  $\rho_i$  of unit multiplicity, by recalling that

$$(\sigma_a^i)^2 = \hat{u}_i^\top \mathcal{D}_a \hat{u}_i - (\alpha_a^i)^2$$

in which both quantities have now been estimated.

4.2.3. *Eigenvector Cross-Fluctuations* ( $\sigma_a^{i,j}$ ). We complete our eigenvector analysis by evaluating the cross-fluctuations between eigenvectors  $\hat{u}_i$  and  $\hat{u}_j$  of  $L$ , both assumed of unit multiplicity. We must divide this study into two cross-fluctuations: (i) those involving  $\hat{u}_1$  (that is  $(\mathbf{1}_n^\top D \mathbf{1}_n)^{-\frac{1}{2}} D^{\frac{1}{2}} \mathbf{1}_n$ ) versus  $\hat{u}_i$ ,  $i > 1$ , and (ii) those involving two eigenvectors  $\hat{u}_i$ ,  $\hat{u}_j$ ,  $i, j > 1$ .

For case (ii), to estimate

$$\sigma_{i,j}^a = \hat{u}_i^\top \mathcal{D}_a \hat{u}_j - \alpha_a^i \alpha_a^j$$

we need to evaluate  $\hat{u}_i^\top \mathcal{D}_a \hat{u}_j$ . However, unlike previous results,  $\hat{u}_i^\top \mathcal{D}_a \hat{u}_j$  cannot be directly estimated using the Cauchy integral approach. To work this around, we propose to estimate instead

$$\frac{1}{p} J^\top \hat{u}_i \hat{u}_i^\top \mathcal{D}_a \hat{u}_j \hat{u}_j^\top J = (\hat{u}_i^\top \mathcal{D}_a \hat{u}_j) \frac{1}{p} J^\top \hat{u}_i \hat{u}_j^\top J. \quad (9)$$

Indeed, if  $\hat{u}_i$  has a non-trivial projection onto  $\text{span}(j_1, \dots, j_k)$  (in the other case,  $\hat{u}_i$  is of no interest to clustering), then there exists at least one index  $a \in \{1, \dots, k\}$  for which  $\frac{1}{p} j_a^\top \hat{u}_i = \frac{c_a}{c_0} \alpha_a^i$  is non zero. The same holds for  $\hat{u}_j$ , and thus  $\hat{u}_i^\top \mathcal{D}_a \hat{u}_j$  can be retrieved by dividing a specific entry of (9) by the appropriate  $\alpha_a^i$  and  $\alpha_a^j$ .

Our approach to evaluate (9) is to operate a double-contour integration (two applications of Cauchy's formula) by which, for  $\rho, \tilde{\rho}$  two distinct isolated eigenvalues,

$$\frac{1}{p} J^\top \hat{\Pi}_\rho \mathcal{D}_a \hat{\Pi}_{\tilde{\rho}} J = -\frac{1}{4\pi} \oint_{\gamma_\rho} \oint_{\gamma_{\tilde{\rho}}} \frac{1}{p} J^\top \mathcal{Q}_z \mathcal{D}_a \mathcal{Q}_z J dz d\tilde{z} + o(1)$$

with obvious notations and with  $\mathcal{Q}_z = (PW^\top WP + UBU^\top - zI_n)^{-1}$ .

For case (i), note from Proposition 1 that, in the first order,  $\hat{u}_1$  is essentially the vector  $\frac{\mathbf{1}_n}{\sqrt{n}}$  of norm 1 plus fluctuations of norm  $O(n^{-\frac{1}{2}})$ . If one were to evaluate  $\hat{u}_1^\top \mathcal{D}_a \hat{u}_i$ ,  $i > 1$ , as previously done, this would thus provide an inaccurate estimate to capture the cross-fluctuations. However, since the fluctuating part of  $\hat{u}_1$  is well understood by Remark 6, and is in particular directly related to  $\psi$ , defined in (5), we shall here estimate instead  $\psi^\top \mathcal{D}_a \hat{u}_i$ , which can be obtained from  $\psi^\top \mathcal{D}_a \hat{u}_i \hat{u}_i^\top \frac{J}{\sqrt{p}}$ , the latter being in turn obtained, in case of unit multiplicity, from

$$\psi^\top \mathcal{D}_a \hat{\Pi}_{\rho_i} \frac{J}{\sqrt{p}} = \frac{1}{2\pi i} \oint_{\gamma_{\rho_i}} \psi^\top \mathcal{D}_a (PW^\top WP + UBU^\top - zI_n)^{-1} \frac{J}{\sqrt{p}} dz.$$

Using Lemma 2 in its full generality, we then obtain the following results.

**Theorem 8** (Joint eigenvector fluctuations,  $f'(\tau)$  away from zero). *Under the setting and notations of Theorem 4, as  $n \rightarrow \infty$ ,*

$$\begin{aligned} \psi^\top \mathcal{D}_a \hat{\Pi}_\rho \frac{J}{\sqrt{p}} &= c_a g_a(\rho) \frac{2}{p} \text{tr} C_a^2 \left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right) t^\top \Gamma_\rho \Xi_\rho + o(1) \\ \frac{1}{p} J^\top \hat{\Pi}_\rho \mathcal{D}_a \hat{\Pi}_{\tilde{\rho}} J &= h(\tau, \rho) h(\tau, \tilde{\rho}) \Xi_\rho^\top E_{a;\rho\tilde{\rho}} \Xi_{\tilde{\rho}} + o(1) \end{aligned}$$

almost surely, where

$$E_{a;\rho\tilde{\rho}} \triangleq E_{a;\rho\tilde{\rho}}^J + \Gamma_\rho E_{a;\rho\tilde{\rho}}^M \Gamma_{\tilde{\rho}} + \left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right)^2 E_{a;\rho\tilde{\rho}}^\psi \Gamma_\rho t t^\top \Gamma_{\tilde{\rho}}.$$

**Theorem 9** (Joint eigenvector fluctuations,  $f'(\tau) \rightarrow 0$ ). *Under the setting and notations of Theorem 5 and Theorem 7, as  $n \rightarrow \infty$ ,*

$$\begin{aligned} \psi^\top \mathcal{D}_a \hat{\Pi}_\rho^0 \frac{J}{\sqrt{p}} &= \frac{c_a}{c_0 \rho} \frac{2}{p} \operatorname{tr} C_a^2 \frac{f''(\tau)}{f(\tau)} t^\top \Gamma_\rho^0 \Xi_\rho^0 + o(1) \\ \frac{1}{p} J^\top \hat{\Pi}_\rho^0 \mathcal{D}_a \hat{\Pi}_{\tilde{\rho}}^0 J &= h^0(\tau, \rho) h^0(\tau, \tilde{\rho}) (\Xi_\rho^0)^\top E_{a;\rho\tilde{\rho}}^0 \Xi_{\tilde{\rho}}^0 + o(1) \end{aligned}$$

almost surely, where

$$E_{a;\rho\tilde{\rho}}^0 \triangleq E_{a;\rho\tilde{\rho}}^{0;J} + \left( \frac{f''(\tau)}{f(\tau)} \right)^2 E_{a;\rho\tilde{\rho}}^{0;\psi} \Gamma_\rho^0 t t^\top \Gamma_{\tilde{\rho}}^0.$$

**Remark 13** (Relation to  $\sigma_a^{i,j}$ ). For  $i, j > 1$ , from the convention on the signs of  $\hat{u}_i$  and  $\hat{u}_j$  given by Remark 10, we get immediately that

$$\sigma_a^{i,j} = \frac{\frac{1}{\sqrt{n_b}} j_b^\top \hat{u}_i \hat{u}_i^\top \mathcal{D}_a \hat{u}_j \hat{u}_j^\top \frac{1}{\sqrt{n_d}} j_d}{\alpha_b^i \alpha_d^j}$$

for any index  $b, d \in \{1, \dots, k\}$  for which  $\alpha_b^i, \alpha_d^j \neq 0$ . As for the case  $i = 1, j > 1$ , corresponding to  $\hat{u}_1 = (1_n^\top D 1_n)^{-\frac{1}{2}} D^{\frac{1}{2}} 1_n$ , we may similarly impose the convention that  $1_n^\top \hat{u}_1 > 0$  (for all large  $n$ ), which is easily ensured as  $\hat{u}_1 = \frac{1_n}{\sqrt{n}} + o(1)$  almost surely. Then we find that

$$\sigma_a^{1,j} = \frac{1}{n} \left[ \frac{1}{\sqrt{c_0}} \frac{f'(\tau)}{2f(\tau)} \frac{\psi^\top \hat{u}_j \hat{u}_j^\top \frac{1}{\sqrt{n_d}} j_d}{\alpha_j^d} + o(1) \right]$$

again for all  $d \in \{1, \dots, k\}$  for which  $\alpha_d^j \neq 0$ .

An illustration of the application of the results of Sections 4.2.1–4.2.3 to determine the class-wise means, fluctuations, and cross-fluctuations of the eigenvectors, as discussed in the early stage of Section 4.2, is depicted in Figure 5. There, under the same setting as in Figure 4 (that is, three classes with various means and covariances under Gaussian kernel), we display in class-wise colored crosses the  $n$  couples  $([\hat{u}_i]_a, [\hat{u}_j]_a)$ ,  $a = 1, \dots, n$ , for the  $i$ -th and  $j$ -th dominant eigenvectors  $\hat{u}_i$  and  $\hat{u}_j$  of  $L$ . The left figure is for  $(\hat{u}_1, \hat{u}_2)$  and the right figure for  $(\hat{u}_2, \hat{u}_3)$ . On top of these values are drawn the ellipses corresponding to the one- and two-dimensional standard deviations of the fluctuations, obtained from the covariance matrix  $\begin{bmatrix} (\sigma_i)_a^2 & (\sigma_{ij})_a \\ (\sigma_{ij})_a & (\sigma_j)_a^2 \end{bmatrix}$ .

In order to get a precise understanding of the behavior of the spectral clustering based on  $L$ , we shall successively constrain the conditions of Assumption 1 so to obtain meaningful results.



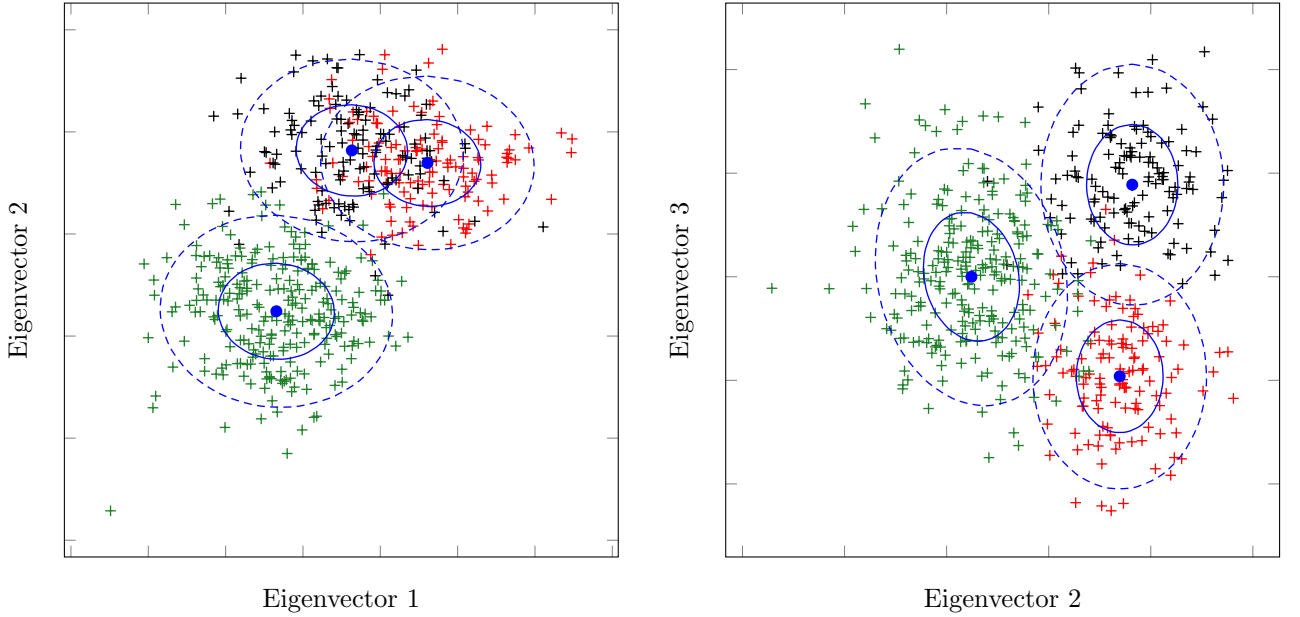


FIGURE 5. Two dimensional representation of the eigenvectors one and two (left) and two and three (right) of  $L$  in the setting of Figure 4.

## 5. SPECIAL CASES

The results obtained in Section 4 above are particularly difficult to analyze for lack of tractability of the functions  $g_1(z), \dots, g_k(z)$  (even though from a numerical point of view, these can be accurately estimated as the output of a provably converging fixed point algorithm; see [BC15]). In this section, we shall consider three scenarios for which all  $g_a(z)$  assume a constant value:

- (1) We first assume  $C_1 = \dots = C_k$ . In this setting, only  $M$  can be discriminated over for spectral clustering. We shall show that, in this case, up to  $k - 1$  isolated eigenvalues can be found in-between each pair of connected components in the limiting support of the empirical spectrum of  $PW^TWP$ , in addition to the isolated eigenvalue  $n$ . But more importantly, we shall show that, as long as  $f'(\tau)$  is away from zero, the kernel choice is asymptotically irrelevant (so one may take  $f(x) = x$  with the same asymptotic performance for instance).
- (2) Next, we shall consider  $\mu_1 = \dots = \mu_k$  and take  $C_a = (1 + p^{-1/2}\gamma_a)C_0$  for some fixed  $\gamma_1, \dots, \gamma_k$ . This ensures that  $T$  vanishes and thus only  $t$  can be used for clustering. There we shall surprisingly observe that a maximum of two isolated eigenvalues altogether can be found in the limiting spectrum and that  $\hat{u}_1$  (associated with eigenvalue  $n$ ) and the hypothetical  $\hat{u}_2$  are extremely correlated. This indicates here that clustering can be asymptotically performed based solely on  $\hat{u}_1$ .
- (3) Finally, to underline the effect of  $T$  alone, we shall enforce a model in which  $\mu_1 = \dots = \mu_k$ ,  $n_1 = \dots = n_k = n/k$ , and  $C_a$  is of the form  $\mathcal{D}(D_1, \dots, D_1, D_2, D_1, \dots, D_1)$

with  $D_2$  in the  $a$ -th position. There we shall observe that the hypothetical eigenvalues have multiplicity  $k - 1$ , thus precluding a detailed eigenvector analysis.

5.1. **Case  $C_a = C$  for all  $a$ .** Assume that for all  $a$ ,  $C_a = C$  (which we may further relax by merely requiring that  $t \rightarrow 0$  and  $T \rightarrow 0$  in the large  $p$  limit). For simplicity of exposition, we shall require in that case the following.

**Assumption 3** (Spectral convergence of  $C$ ). *As  $p \rightarrow \infty$ , the empirical spectral measure  $\frac{1}{p} \sum_{i=1}^p \delta_{\lambda_i(C)}$  converges weakly to  $\nu$ . Besides,*

$$\max_{1 \leq i \leq p} \text{dist}(\lambda_i(C), \text{supp}(\nu)) \xrightarrow{\text{a.s.}} 0.$$

This assumption ensures, along with the results from [SC95] and [BS98], that  $g_1(z), \dots, g_k(z)$  all converge towards a unique  $g(z)$ , solution to the implicit equation

$$g(z) = \frac{1}{c_0} \left( -z + \int \frac{u\nu(du)}{1 + g(z)u} \right)^{-1}. \quad (10)$$

This is the Stieltjes transform of a probability measure  $\bar{\mathbb{P}}$  having continuous density and support  $\mathcal{S}$ . Besides, as  $n, p \rightarrow \infty$ , none of the eigenvalues of  $PW^\top WP$  are found away from  $\mathcal{S} \cup \{0\}$ .

In this setting, only the matrix  $M$  can be discriminated upon to perform clustering and thus we need to take here  $f'(\tau)$  away from zero to obtain meaningful results which, since  $\tau \rightarrow 2 \int u\nu(du)$ , merely requires that  $f'(2 \int u\nu(du)) \neq 0$ . Starting then from Theorem 2 and using the fact that  $Mc = 0$ , we get

$$G_z = h(\tau, z)g(z) (I_k + M^\top (I_p + g(z)C)^{-1} M \mathcal{D}(c)) + o(1)$$

with<sup>7</sup>

$$h(\tau, z) = 1 + \left( \frac{5f'(\tau)}{4f(\tau)} - \frac{f''(\tau)}{f'(\tau)} \right) g(z) \int u^2 \nu(du).$$

The values of  $\rho \in \mathcal{S}' = \mathcal{S} \cup \{\rho; h(\tau, \rho) = 0\}$  for which  $G_\rho$  (asymptotically) has zero eigenvalues are such that  $I_k + M^\top [I_p + g(\rho)C]^{-1} M \mathcal{D}(c)$  is singular. By Sylvester's identity, this is equivalent to looking for such  $\rho$  satisfying

$$0 = \det \left( \frac{1}{g(\rho)} I_p + C + M \mathcal{D}(c) M^\top \right).$$

Hence, these  $\rho$ 's are such that  $-g(\rho)^{-1}$  coincides with one of the eigenvalues of  $C + M \mathcal{D}(c) M^\top$ . But from [MP67, SC95], the image of the restriction of  $-1/g$  to  $\mathbb{R}_+$  is defined on a subset of  $\mathbb{R}$  excluding the support of  $\nu$ . More precisely, the image of  $g : \mathbb{R} \setminus \mathcal{S} \rightarrow \mathbb{R}$ ,

<sup>7</sup>The  $o(1)$  term accounts for  $g(z)$  being here a limiting quantity rather than a finite  $p$  deterministic equivalent.

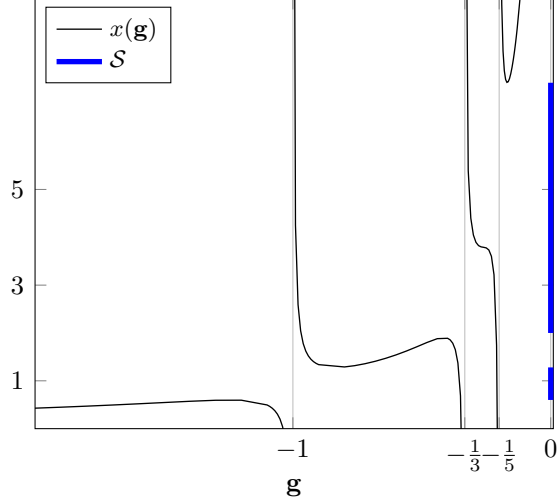


FIGURE 6. Representation of  $x(\mathbf{g})$  for  $\mathbf{g} \in \mathbb{R} \setminus \text{supp}(\nu)$ ,  $\nu = \frac{1}{3}(\delta_1 + \delta_3 + \delta_5)$ ,  $c_0 = 10$ . The support  $\mathcal{S}_p$  can be read on the right vertical axis and corresponds to the complementary to  $\{x(\mathbf{g}) \mid x'(\mathbf{g}) > 0\}$ .

$x \mapsto g(x)$ , is the set of values  $\mathbf{g}$  such that  $x'(\mathbf{g}) > 0$ , where  $x : \mathbb{R} \setminus \text{supp}(\nu) \rightarrow \mathbb{R}$ ,  $\mathbf{g} \mapsto x(\mathbf{g})$ , is defined by

$$x(\mathbf{g}) \triangleq -\frac{1}{c_0} \frac{1}{\mathbf{g}} + \int \frac{u}{1 + \mathbf{g}u} \nu(du).$$

A visual representation of  $x(\mathbf{g})$  is provided in Figure 6. Now, since  $Mc = 0$ ,  $M$  has maximum rank  $k - 1$  and thus, by Weyl's interlacing inequalities and Assumption 3, there asymptotically exist up to  $k - 1$  isolated eigenvalues in-between each connected component of  $\text{supp}(\nu)$ .

Additionally, since  $t = 0$ , as per Remark 2, if there exists  $\rho_+$  away from  $\mathcal{S}$  such that  $h(\tau, \rho_+) = 0$ , that is for which  $-g(\rho_+)^{-1} = \left( \frac{5f'(\tau)}{4f(\tau)} - \frac{f''(\tau)}{f'(\tau)} \right) \int u^2 \nu(du)$ , then an additional isolated eigenvalue of  $L$  may be found.

Gathering the above, we thus have the following corollary of Theorem 2.

**Corollary 1** (Eigenvalues for  $C_a = C$ ). *Let Assumptions 1–3 hold with  $f'(2 \int uv(du)) \neq 0$ . Denote  $\bigcup_{i=1}^s [l_i^\nu, r_i^\nu]$  the support of  $\nu$  and take the convention  $r_0^\nu = -\infty$ ,  $l_{s+1}^\nu = \infty$ . For each  $i \in \{0, \dots, s + 1\}$ , denote  $\ell_1^i, \dots, \ell_{k_i}^i$  the  $k_i \leq k - 1$  eigenvalues of  $C + M \mathcal{D}(c) M^\top$  in  $(r_i, l_{i+1})$  having non-vanishing distance from the  $r_i$  and  $l_{i+1}$ . For each  $j \in \{1, \dots, k_i\}$ , if*

$$1 > c_0 \int \frac{t^2 \nu(dt)}{(t - \ell_j^i)^2} \quad (11)$$

then there exists an isolated eigenvalue of  $L$ , which is asymptotically well approximated by  $-\frac{2f'(\tau)}{f(\tau)}\rho_j^i + \frac{f(0)-f(\tau)+\tau f'(\tau)}{f(\tau)}$ , where

$$\rho_j^i = \ell_j^i \left( \frac{1}{c_0} - \int \frac{u\nu(du)}{u - \ell_j^i} \right).$$

Besides, let  $\ell_+ = \left( \frac{5f'(\tau)}{4f(\tau)} - \frac{f''(\tau)}{f'(\tau)} \right) \int u^2\nu(du)$ . Then, if

$$1 > c_0 \int \frac{u^2\nu(du)}{(u - \ell_+)^2}$$

there is an additional corresponding isolated eigenvalue in the spectrum of  $L$  given by  $-\frac{2f'(\tau)}{f(\tau)}\rho_+ + \frac{f(0)-f(\tau)+\tau f'(\tau)}{f(\tau)}$  with

$$\rho_+ = \ell_+ \left( \frac{1}{c_0} - \int \frac{u\nu(du)}{u - \ell_+} \right).$$

These and  $n$  characterize all the isolated eigenvalues of  $L$ .

As a further corollary, for  $C = \beta I_p$ , we obtain

$$\rho_j^i = \frac{\ell_j^i}{c_0} + \beta \frac{\ell_j^i}{\ell_j^i - \beta}$$

under the condition that

$$|\ell_j^i - \beta| > \beta\sqrt{c_0} \quad (12)$$

which is a classical separability condition in standard spiked models [J01, BBP05, P07].

Before interpreting this result, let us next characterize the eigenvectors associated to these eigenvalues. The eigenvector attached to the eigenvalue  $n$  is  $D^{\frac{1}{2}}1_n$  and has already been analyzed and carries no information since  $t = 0$ . We also know that the hypothetical eigenvalue  $\rho_+$  does not carry any relevant clustering information. We are then left to study the eigenvalues  $\rho_j^i$ . For those (assuming they remain distant from one another),

$$\frac{1}{p} J^\top \hat{\Pi}_\rho J = -h(\tau, \rho)g(\rho) (\mathcal{D}(c) - cc^\top) \sum_{q=1}^{m_\rho} \frac{(V_{r,\rho})_q (V_{l,\rho})_q^\top}{(V_{l,\rho})_q^\top G'_\rho (V_{r,\rho})_q} + o(1)$$

almost surely, where  $\rho = \rho_j^i$  is understood to be any of the  $\rho_j^i$  eigenvalues from Corollary 1. It is easier here to use the fact that  $\bar{G}_\rho \triangleq G_\rho \mathcal{D}(c^{-1})$  is symmetric with  $V_\rho \triangleq \mathcal{D}(c^{-1})V_{r,\rho} = V_{l,\rho}$  as eigenvectors for its  $m_\rho$  zero eigenvalues. Thus

$$\frac{1}{p} J^\top \hat{\Pi}_\rho J = -h(\tau, \rho)g(\rho) (I_k - c1_k^\top) \sum_{q=1}^{m_\rho} \frac{(V_\rho)_q (V_\rho)_q^\top}{(V_\rho)_q^\top \bar{G}'_\rho (V_\rho)_q} + o(1).$$

Using  $\bar{G}_\rho V_\rho = 0$ , we get  $(V_\rho)_q^\top \bar{G}'_\rho (V_\rho)_q = -h(\tau, \rho)M^\top (I_k + g(\rho)C)^{-2} M g'(\rho)$ . Besides,  $V_\rho$  can be expressed as  $V_\rho = \mathcal{D}(c)M^\top \Upsilon_\rho$  with  $\Upsilon_\rho \in \mathbb{R}^{p \times m_\rho}$  the column-concatenated  $m_\rho$

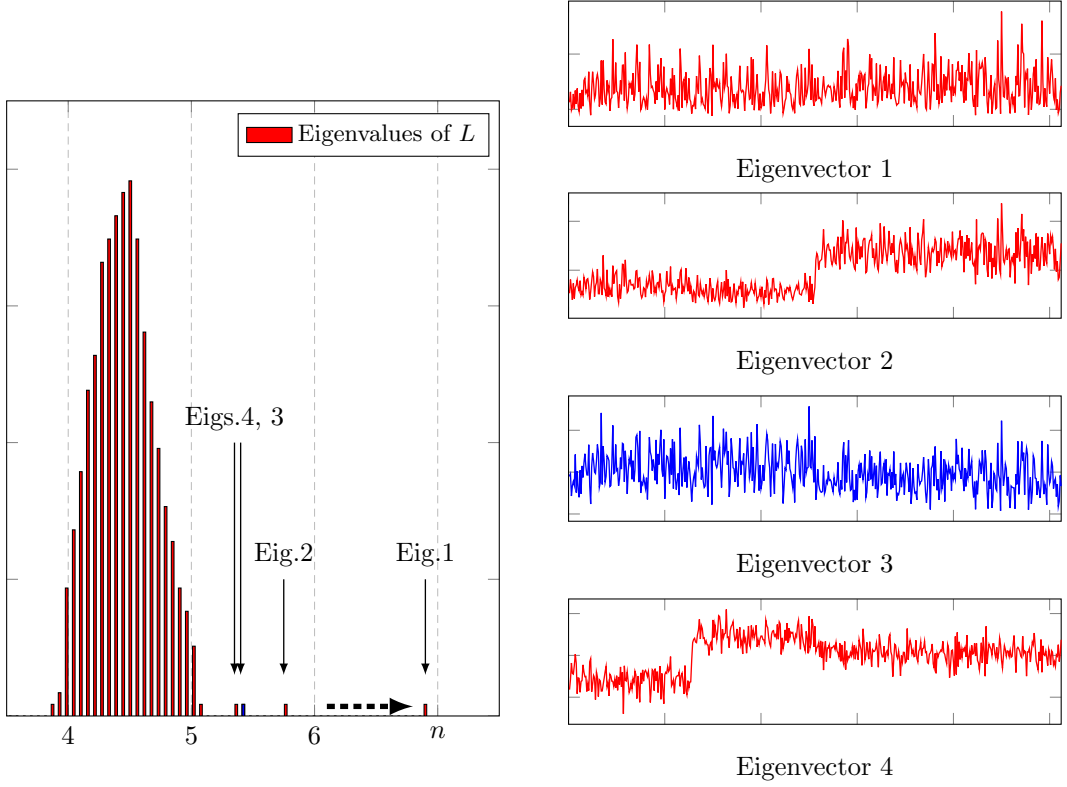


FIGURE 7. Eigenvalues of  $L'$  and dominant four eigenvectors of  $L$  for  $C_1 = \dots = C_k = I_p$ ,  $f(x) = 4(x - \tau)^2 - (x - \tau) + 4$  ( $\tau = 2$ ,  $f(0) = 22$ ,  $f(\tau) = 4$ ,  $f'(\tau) = -1$ ,  $f''(\tau) = 8$ ),  $p = 2048$ ,  $n = 512$ , three classes with  $n_1 = n_2 = 128$ ,  $n_3 = 384$ ,  $[\mu_i]_j = 5\delta_{ij}$ . Emphasis made on the “non-informative” eigenvalue-eigenvector pair, solution to  $h(\tau, \rho) = 0$ .

eigenvectors associated with the eigenvalue  $\ell = \ell_j^i = -1/g(\rho_j^i)$  of  $C + M\mathcal{D}(c)M^\top$ . Since  $\mathbf{1}_k^\top \mathcal{D}(c)M^\top = 0$  and  $(C - \ell I_p)^{-1}M\mathcal{D}(c)M^\top \Upsilon_\rho = -\Upsilon_\rho$ , we thus finally obtain

$$\frac{1}{p}J^\top \hat{\Pi}_\rho J = \frac{1}{\ell} \left( \frac{1}{c_0} - \int \frac{u^2 \nu(du)}{(u - \ell)^2} \right) \mathcal{D}(c)M^\top \Upsilon_\rho \Upsilon_\rho^\top M \mathcal{D}(c) + o(1).$$

Regarding fluctuations, note that the leftmost inverse matrix in the definition of  $R^{z_1 z_2}$  (Lemma 2) is merely a rank-one perturbation of the identity matrix, so that, after basic algebraic manipulations, we obtain

$$R^{z_1 z_2} = \frac{c_0 \frac{1}{p} \text{tr } C^2 (C + g(z_1)^{-1} I_p)^{-2}}{1 - c_0 \frac{1}{p} \text{tr } C^2 (C + g(z_2)^{-1} I_p)^{-2}} c \mathbf{1}_k^\top.$$

A careful (but straightforward) development of all the terms in Theorem 6 and Theorem 8, using the already made remarks, then allows one to obtain the following spectral clustering analysis for the setting under present concern.

**Corollary 2** (Spectral Clustering for  $C_a = C$ ). *For  $i \in \{0, \dots, s+1\}$ ,  $j \in \{1, \dots, k_i\}$ , let  $\rho_j^i$  and  $\ell_j^i$  be as defined in Corollary 1, assumed of unit multiplicity, with associated eigenvector  $\hat{u}_j^i$  in  $L$ . Under the model (6), denoted here*

$$\hat{u}_j^i = \sum_{a=1}^k (\alpha_j^i)_a \frac{j_a}{\sqrt{n_a}} + (\sigma_j^i)_a (\omega_j^i)_a$$

with  $(\omega_j^i)_a \in \mathbb{R}^n$  supported by  $\mathcal{C}_a$  of unit norm and orthogonal to  $j_a$ , we find that

$$\begin{aligned} (\alpha_j^i)_a^2 &= c_a \frac{1}{\ell_j^i} \Upsilon_{\rho_j^i}^\top \mu_a^\circ (\mu_a^\circ)^\top \Upsilon_{\rho_j^i} \left( 1 - c_0 \int \frac{u^2 \nu(du)}{(u - \ell_j^i)^2} \right) + o(1) \\ (\sigma_j^i)_a^2 &= c_a \left[ 1 - \frac{1}{\ell_j^i} \Upsilon_{\rho_j^i}^\top M \mathcal{D}(c) M^\top \Upsilon_{\rho_j^i} \left( 1 - c_0 \int \frac{u^2 \nu(du)}{(u - \ell_j^i)^2} \right) \right] + o(1) \end{aligned}$$

with  $\Upsilon_{\rho_j^i} \in \mathbb{R}^p$  the eigenvector of  $C + M \mathcal{D}(c) M^\top$  associated with the eigenvalue  $\ell_j^i$ . Besides, letting  $(\sigma_{j\bar{j}}^{\bar{i}})_a \triangleq (\sigma_j^i)_a (\sigma_{\bar{j}}^{\bar{i}})_a (\omega_j^i)_a^\top (\omega_{\bar{j}}^{\bar{i}})_a$ , we have

$$(\sigma_{j\bar{j}}^{\bar{i}})^2 = \frac{c_a^2}{\ell_j^i \ell_{\bar{j}}^{\bar{i}}} \left( 1 - c_0 \int \frac{u^2 \nu(du)}{(u - \ell_j^i)^2} \right) \left( 1 - c_0 \int \frac{u^2 \nu(du)}{(u - \ell_{\bar{j}}^{\bar{i}})^2} \right) \left( \Upsilon_{\rho_j^i}^\top C \Upsilon_{\rho_{\bar{j}}^{\bar{i}}} \right)^2 + o(1).$$

It is interesting to note, from Corollary 1 and Corollary 2, that all expressions of the relevant quantities in this setting are here completely explicit. This allows for easy interpretations of the results. Quite a few remarks can in particular be made.

**Remark 14** (Asymptotic kernel irrelevance). From the results of Corollaries 1 and 2, it appears that, aside from the hypothetical isolated eigenvalue  $\rho_+$  (the eigenvector of which carries in any case no information), when  $C_a = C$  for all  $a \in \{1, \dots, k\}$ , the choice of the kernel function  $f$  is asymptotically of no avail, so long that  $f'(\tau) \not\equiv 0$ . This can be interpreted in practice by the fact that, since the data  $x_i$  are linearly separable and that no difference aside location metrics can be exploited to discriminate them, the so-called “kernel trick”, which projects the data on a high dimensional space to improve separability, does not provide any additional gain for clustering.

**Remark 15** (On the possibility to cluster). Even though  $\nu$  would have unbounded support, (11) allows isolated eigenvalue-eigenvector pairs to emerge in-between successive clusters of eigenvalues and carry relevant clustering information, however necessarily with imperfect alignment to  $j_1, \dots, j_k$ . This disrupts from the standard assumption that only extreme eigenvectors may be exploited for spectral clustering.

**Remark 16** (On the interplay between  $C$  and  $M$ ). All quantities obtained in Corollary 2 rely on  $\Upsilon_{\rho_j^i}$  which is an “isolated” eigenvector of  $C + M \mathcal{D}(c) M^\top$ . When  $C = \beta I_p$ ,  $\Upsilon_{\rho_j^i}$  is merely an eigenvector of  $M \mathcal{D}(c) M^\top$  with eigenvalue  $\ell_j^i - \beta$ , so that  $(\sigma_j^i)_a^2$  simplifies as

$$(\sigma_j^i)_a^2 = c_a \left[ 1 - \frac{\ell_j^i - \beta}{\ell_j^i} \left( 1 - \frac{c_0 \beta^2}{(\beta - \ell_j^i)^2} \right) \right] + o(1).$$

Also, and possibly more importantly, since  $\Upsilon_{\rho_j^i}^\top \Upsilon_{\rho_j^{\tilde{i}}} = 0$  for distinct  $i, j$  and  $\tilde{i}, \tilde{j}$ , we find that  $(\sigma_{j\tilde{j}}^{i\tilde{i}})_a = o(1)$  for each  $a = 1, \dots, k$ . Therefore, the fluctuating parts of the isolated eigenvectors of  $L$  are asymptotically uncorrelated when  $C = \beta I_p$ .

**Remark 17** (On the existence of useless eigenvectors). Corollary 1 predicts the possibility that  $h(\tau, \rho) = 0$  for some  $\rho$ , the associated eigenvector of which does not align to  $j_1, \dots, j_k$ . In Figure 7, we present such a scenario in a  $k$ -class setting for which  $k+1$  (and not only  $k$ ) isolated eigenvalues of  $L$  are found. We depict in parallel the corresponding eigenvectors. As expected, the eigenvector  $D^{\frac{1}{2}}1_n$  does not visually align to  $j_1, \dots, j_k$ . For the other three, note that eigenvectors 2 and 4 do align to  $j_1, \dots, j_k$  and are thus the sought-for (at most)  $k-1$  eigenvectors in this connected component of  $\mathbb{R} \setminus \mathcal{S}'$ . As for eigenvector 3, it shows no strong alignment to  $j_1, \dots, j_k$ , and must therefore arise from the solution  $h(\tau, \rho_+) = 0$ .

**5.2. Case  $\mu_a = \mu$  and  $C_a = (1 + p^{-1/2}\gamma_a)C$  for all  $a$ .** Consider now the somewhat opposite scenario where  $\mu_1 = \dots = \mu_k$  and  $C_a = (1 + p^{-1/2}\gamma_a)C$ ,  $a \in \{1, \dots, k\}$ , for some  $\gamma_1, \dots, \gamma_k \in \mathbb{R}$  fixed. We shall further denote  $\gamma = (\gamma_1, \dots, \gamma_k)^\top$  and  $\gamma^2 = (\gamma_1^2, \dots, \gamma_k^2)^\top$ .

Similar to previously, we shall place ourselves for simplicity under Assumption 3. Although not necessary, it shall also be simpler to assume  $C^\circ = C$  (which is always possible up to modifying  $C, \gamma_1, \dots, \gamma_k$ ).<sup>8</sup> In this case, both scenarios  $f'(\tau)$  away or converging to zero are of interest. Let us focus first on the former. There  $G_z$  reduces to

$$G_z = I_k + g(z) \left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right) \left( 2 \int u^2 \nu(du) I_k + tt^\top \mathcal{D}(c) \right) + o(1)$$

with  $g(z)$  given by the implicit equation (10). Letting  $z = \rho \in \mathbb{R} \setminus \mathcal{S}$ , the roots of the right-hand matrix are the solutions (if any) to

$$-\frac{1}{g(\rho)} = \left( \frac{5f'(\tau)}{4f(\tau)} - \frac{f''(\tau)}{f'(\tau)} \right) \int u^2 \nu(du)$$

or to

$$-\frac{1}{g(\rho)} = \left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right) \left( 2 \int u^2 \nu(du) + \left( \int u \nu(du) \right)^2 c^\top \gamma^2 \right).$$

It is easily checked that the former value, corresponding to  $h(\tau, \rho) = 0$ , does not bring a zero eigenvalue in  $H_\rho$  and thus, as per Remark 2, does not map an isolated eigenvalue of  $L$ .

Assuming now  $f'(\tau) \rightarrow 0$ , a similar derivation leads to either  $\rho^0 = \frac{2f''(\tau)}{f(\tau)} \int u^2 \nu(du)$ , which corresponds to  $h^0(\tau, \rho^0) = 0$ , hence not a solution, or to

$$\rho^0 = \frac{f''(\tau)}{f(\tau)c_0} \left( 2 \int u^2 \nu(du) + \left( \int u \nu(du) \right)^2 c^\top \gamma^2 \right).$$

These results can then be gathered as follows.

<sup>8</sup>And again, we may here relax these assumptions further by merely requiring that  $M \rightarrow 0$  and  $T \rightarrow 0$ .

**Corollary 3** (Eigenvalues for  $C_a = (1 + p^{-1/2}\gamma_a)C$ ). *Let Assumptions 1–3 hold, with  $C_a = (1 + \frac{\gamma_a}{\sqrt{p}})C$  and  $\mu_a = \mu$  for all  $a$ . If  $f'(\tau) \not\rightarrow 0$ , denote*

$$\ell = \left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right) \left[ 2 \int u^2 \nu(du) + \left( \int u \nu(du) \right)^2 c^\top \gamma^2 \right]. \quad (13)$$

Then, if

$$1 > c_0 \int \frac{u^2 \nu(du)}{(u - \ell)^2}$$

there exists an isolated eigenvalue in the spectrum of  $L$  with value  $-2 \frac{f'(\tau)}{f(\tau)} \rho + \frac{f(0) + f(\tau) - \tau f'(\tau)}{f(\tau)}$  where

$$\rho = \ell \left( \frac{1}{c_0} - \int \frac{u \nu(du)}{u - \ell} \right). \quad (14)$$

This value, along with  $n$  are all the isolated eigenvalues of  $L$ . Otherwise, if  $f'(\tau) \rightarrow 0$ , then the non-zero spectrum of  $L$  is composed of the eigenvalue  $n$  and the eigenvalue  $\rho^0 + \frac{f(0) + f(\tau)}{f(\tau)}$  with

$$\rho^0 = \frac{f''(\tau)}{f(\tau)c_0} \left[ 2 \int u^2 \nu(du) + \left( \int u \nu(du) \right)^2 c^\top \gamma^2 \right]. \quad (15)$$

As for the eigenvector projections, assuming first  $f'(\tau) \not\rightarrow 0$ , note that as  $G_\rho$  is a rank-one perturbation of a scaled identity matrix, the left and right eigenvectors  $V_{l,\rho}, V_{r,\rho} \in \mathbb{R}^k$  are respectively proportional to  $t$  and  $\mathcal{D}(c)t$ , so that

$$\frac{V_{r,\rho} V_{l,\rho}^\top}{V_{l,\rho}^\top G'_\rho V_{r,\rho}} = - \frac{\gamma \gamma^\top \mathcal{D}(c) g(\rho)^2}{c^\top \gamma^2 g'(\rho)}.$$

We thus finally obtain,

$$\frac{1}{p} J^\top \hat{\Pi}_\rho J = \frac{1}{c_0} \frac{1 - c_0 \int \frac{u^2 \nu(du)}{(u - \ell)^2}}{2 \frac{\int u^2 \nu(du)}{(\int u \nu(du))^2} + c^\top \gamma^2} \mathcal{D}(c) \gamma \gamma^\top \mathcal{D}(c) + o(1)$$

and, similarly, for  $f'(\tau) \rightarrow 0$ ,

$$\frac{1}{p} J^\top \hat{\Pi}_{\rho^0} J = \frac{1}{c_0} \frac{1}{2 \frac{\int u^2 \nu(du)}{(\int u \nu(du))^2} + c^\top \gamma^2} \mathcal{D}(c) \gamma \gamma^\top \mathcal{D}(c) + o(1).$$

The result of Corollary 3 is quite surprising when compared to Corollary 1. Indeed, while the latter allowed for up to  $k - 1$  isolated eigenvalues to be found outside  $\mathcal{S}'$ , here a maximum of one eigenvalue is available, irrespective of  $C$ . This state of fact is obviously linked to  $tt^\top$  being of unit rank while  $MM^\top$  can be of rank up to  $k - 1$ . For practical purposes, there being no information diversity, the clustering task is increasingly difficult to achieve as  $k$  increases. But this becomes even worse when considering the cross-correlations



between  $D^{\frac{1}{2}}\mathbf{1}_n$  and (the hypothetical) eigenvector  $\hat{u}$  associated with  $\rho$ . Precisely, after some calculus, we obtain the counter-part of Corollary 2 as follows.

**Corollary 4** (Spectral Clustering for constant  $\mu_a$  and  $C_a = (1 + p^{-1/2}\gamma_a)C$ ). *Following the model (6), write here, for  $\hat{u}_1 = (\mathbf{1}_n^\top D \mathbf{1}_n)^{-\frac{1}{2}} D^{\frac{1}{2}} \mathbf{1}_n$ , and for  $\hat{u}_2$  the (hypothetical) isolated eigenvector of  $L$  associated with the limiting eigenvalue  $\rho$  defined in Corollary 3,*

$$\hat{u}_i = \sum_{a=1}^k (\alpha_i)_a \frac{j_a}{\sqrt{n_a}} + (\sigma_i)_a (\omega_i)_a$$

with  $(\omega_i)_a \in \mathbb{R}^n$  supported by  $\mathcal{C}_a$  of unit norm and orthogonal to  $j_a$ . Then, we find that, for  $f'(\tau) \not\rightarrow 0$ ,

$$\begin{aligned} (\alpha_1)_a^2 &= c_a \left( 1 + \frac{1}{\sqrt{p}} \frac{f'(\tau)}{2f(\tau)} \gamma_a + o(p^{-\frac{1}{2}}) \right)^2 \\ (\sigma_1)_a^2 &= c_a \left( \frac{f'(\tau)}{2f(\tau)} \right)^2 \frac{2}{p} \int u^2 \nu(du) + o(p^{-1}) \\ (\alpha_2)_a^2 &= c_a \frac{(\int u \nu(du))^2 \gamma_a^2}{2 \int u^2 \nu(du) + (\int u \nu(du))^2 c^\top \gamma^2} \left( 1 - c_0 \int \frac{u^2 \nu(du)}{(u-\ell)^2} \right) + o(1) \\ (\sigma_2)_a^2 &= c_a \left[ 1 - \frac{(\int u \nu(du))^2 c^\top \gamma^2}{2 \int u^2 \nu(du) + (\int u \nu(du))^2 c^\top \gamma^2} \left( 1 - c_0 \int \frac{u^2 \nu(du)}{(u-\ell)^2} \right) \right] + o(1) \end{aligned}$$

with  $\ell$  also defined in Corollary 3. Besides, for  $(\sigma_{12})_a \triangleq (\sigma_1)_a (\sigma_2)_a (\omega_1)_a^\top (\omega_2)_a$ , we have

$$(\sigma_{12})_a^2 = c_a^2 \frac{1}{p} \frac{\left( \frac{f'(\tau)}{2f(\tau)} \right)^2 (2 \int u^2 \nu(du))^2}{2 \int u^2 \nu(du) + (\int \nu(du))^2 c^\top \gamma^2} \left( 1 - c_0 \int \frac{u^2 \nu(du)}{(u-\ell)^2} \right) + o(p^{-1}).$$

If  $f'(\tau) \rightarrow 0$ , the results are unchanged but for the terms  $c_0 \int \frac{u^2 \nu(du)}{(u-\ell)^2}$  which vanish.

This leads us to the following important remark.

**Remark 18** (Irrelevance of  $\hat{u}_2$ ). From the expression of  $\ell$  in Corollary 3,  $c^\top \gamma^2$  is directly proportional to  $\ell$ . Thus, for sufficiently large values of  $c^\top \gamma^2$ , using a first order development in  $\ell$ , we get that  $(\sigma_{12})_a^2 (\sigma_1)_a^{-2} (\sigma_2)_a^{-2} \simeq 1$ . Since this is the equality case of the Cauchy-Schwarz inequality, we deduce that the eigenvectors  $\hat{u}_1$  and  $\hat{u}_2$  tend to have similar fluctuations. Besides, the useful part of  $(\alpha_1)_a$  and  $(\alpha_2)_a$  are directly proportional to  $c_a \gamma_a$ , thus making  $\hat{u}_1$  and  $\hat{u}_2$  eventually quite similar. Very little information is thus expected to be further extracted from  $\hat{u}_2$  beyond clustering based on  $\hat{u}_1$ . This remark is even more valid when considering  $f'(\tau) \rightarrow 0$ , for which the discussion holds true even for small  $c^\top \gamma^2$ .

Figure 8 provides an interpretation of Remark 18, by visually comparing the results of Section 5.3 and Section 5.2. Precisely here, we see, for the same choice of a kernel (tailored so that both cases exhibit the required number of informative eigenvectors), the

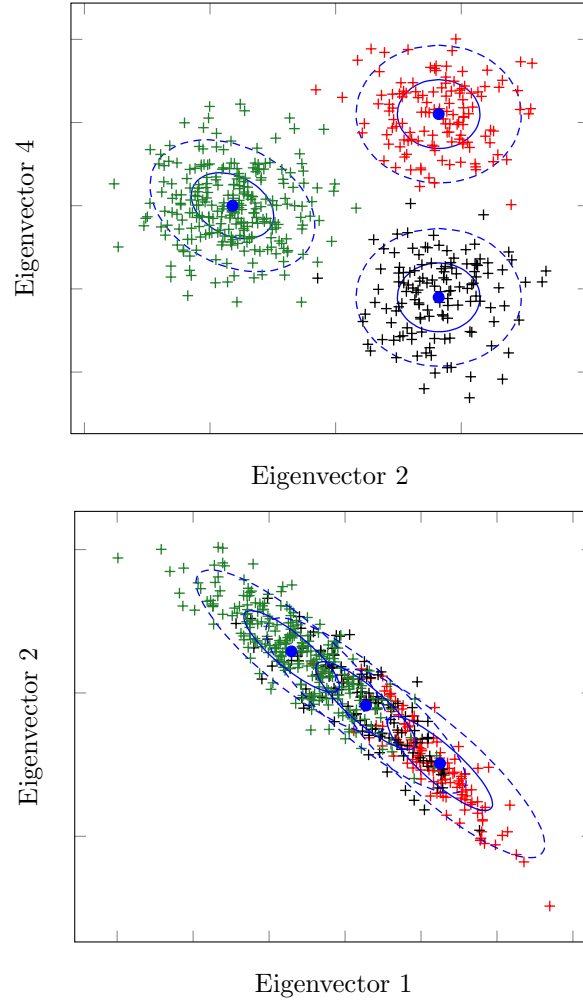


FIGURE 8. Comparison of two dimensional representation of (top): eigenvectors two and three of  $L$ ,  $[\mu_a]_j = 5\delta_{aj}$ ,  $C_1 = \dots = C_k = I_p$ , (bottom): eigenvectors one and two of  $L$ ,  $\mu_1 = \dots = \mu_k$ ,  $C_a = (1 + \frac{2(a-1)}{\sqrt{p}})I_p$ . In both cases,  $k = 3$ ,  $n_1 = n_2 = 192$ ,  $n_3 = 384$ ,  $f(x) = 1.5(x - \tau)^2 - 1(x - \tau) + 5$ . In blue, theoretical means and standard deviations of fluctuations.

distribution of the eigenvectors in the present scenario is quite concentrated along a one-dimensional direction, whereas the former scenario of different  $\mu_a$ 's exhibits a well scattered distribution for the data on the two-dimensional plane.<sup>9</sup>

<sup>9</sup>Note that we voluntarily considered a quite noisy scenario by setting  $C_a = (1 + \frac{2(a-1)}{p})I_p$ , hence  $\gamma_a^2 = 4(a-1)^2$ ; using larger values for  $\gamma_a^2$  rapidly leads  $\gamma_a^2$  to be close to  $\sqrt{p}$  for  $p = 2048$ , hence providing undesired approximation errors.

5.3. **Case  $\mu_a = \mu$ ,  $\text{tr} C_a$  constant.** We now study the effect of the matrix  $T$  alone (so imposing  $\mu_1 = \dots = \mu_k$  and  $\text{tr} C_1 = \dots = \text{tr} C_k$ ). Although it is possible to analyze completely such a scenario (to the least for  $k = 2$ ), it shall be simpler here to enforce  $g_1 = \dots = g_k$ . To this end, we shall assume the symmetric model  $C_i = \mathcal{D}([1_{i-1} \otimes D_1, D_2, 1_{k-i} \otimes D_1])$ , for  $i = 1, \dots, k$ , and for some nonnegative definite  $D_1, D_2 \in \mathbb{R}^{p/k \times p/k}$ . We also suppose that  $n_1 = \dots = n_k$ . A formulation of Assumption 3 adapted to the present setting comes in also handy.

**Assumption 4** (Spectral convergence of  $(k-1)D_1 + D_2$ ). *As  $p \rightarrow \infty$ , the empirical spectral measure  $\frac{1}{p} \sum_{i=1}^p \delta_{\lambda_i((k-1)D_1 + D_2)}$  converges weakly to  $\nu$ .*

This setting ensures that all  $g_i$ 's are identical and are asymptotically equivalent, under Assumption 4, to the implicitly defined

$$g(z) = \frac{1}{c_0} \left[ -z + \frac{1}{k} \int \frac{uv(du)}{1 + \frac{g(z)}{k}u} \right]^{-1}$$

to which we may associate the inverse

$$x(\mathbf{g}) = -\frac{1}{c_0} \frac{1}{\mathbf{g}} + \frac{1}{k} \int \frac{uv(du)}{1 + \frac{\mathbf{g}}{k}u}. \quad (16)$$

In this scenario,  $M = 0$  and  $t = 0$ , so that only  $T$  can be used to perform data clustering. Precisely, we have

$$T = \frac{1}{p} \text{tr}((D_1 - D_2)^2) \left[ I_k - \frac{1}{k} 1_k 1_k^\top \right].$$

Letting first  $f'(\tau) \not\rightarrow 0$ , it is rather immediate to apply Theorem 2 in which

$$G_z = h(\tau, z) \left( I_k - \frac{g(z)}{k} \frac{f''(\tau)}{f'(\tau)} \frac{1}{p} \text{tr}((D_1 - D_2)^2) \left[ I_k - \frac{1}{k} 1_k 1_k^\top \right] \right) + o(1).$$

The case  $f'(\tau) \rightarrow 0$  is handled similarly.

We then have the following result.

**Corollary 5** (Eigenvalues, constant mean and trace). *Let Assumptions 1–2 hold, with  $\mu_i = \mu$  and  $C_i = \mathcal{D}([1_{i-1} \otimes D_1, D_2, 1_{k-i} \otimes D_1])$ ,  $D_1, D_2 \in \mathbb{R}^{p/k \times p/k}$  satisfying Assumption 4, for each  $i \in \{1, \dots, k\}$ . Let also  $x(\mathbf{g})$  be defined as in (16). If  $f'(\tau) \not\rightarrow 0$ , denote*

$$\begin{aligned} \ell &= -\frac{1}{k} \frac{f''(\tau)}{f'(\tau)} \frac{1}{p} \text{tr}((D_1 - D_2)^2) \\ \ell_+ &= \left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right) \frac{2}{p} \text{tr}((k-1)D_1^2 + D_2^2). \end{aligned}$$

*Then, if  $x'(-1/\ell) > 0$  (resp.,  $x'(-1/\ell_+) > 0$ ),  $-2\frac{f'(\tau)}{f(\tau)}\rho + \frac{f(0)+f(\tau)-\tau f'(\tau)}{f(\tau)}$  (resp.,  $-2\frac{f'(\tau)}{f(\tau)}\rho_+ + \frac{f(0)+f(\tau)-\tau f'(\tau)}{f(\tau)}$ ) is asymptotically an isolated eigenvalue of  $L$ , with  $\rho = x(-1/\ell)$  (resp.,  $\rho_+ = x(-1/\ell_+)$ ) of multiplicity  $k-1$  (resp., 1). These and  $n$  form the whole asymptotic isolated spectrum of  $L$ .*

If instead  $f'(\tau) = 0$ , denote

$$\begin{aligned}\rho^0 &= 2 \frac{f''(\tau)}{f(\tau)} \frac{n}{pk} \frac{1}{p} \operatorname{tr}((D_1 - D_2)^2) \\ \rho_+^0 &= 2 \frac{f''(\tau)}{f(\tau)} \frac{n}{pp} \operatorname{tr}((k-1)D_1^2 + D_2^2).\end{aligned}$$

Then  $n$ ,  $\rho^0 + \frac{f(0)-f(\tau)}{f(\tau)}$  (with multiplicity  $k-1$ ) and  $\rho_+^0 + \frac{f(0)-f(\tau)}{f(\tau)}$  form the asymptotic isolated spectrum of  $L$ .

In terms of eigenvectors, the result is also quite immediate from the form of  $G_z$  and we have the following result.

**Corollary 6** (Eigenspace projections, constant mean and trace). *Under the assumptions and notations of Corollary 5, if  $f'(\tau) \neq 0$ ,  $\frac{1}{p} J^\top \hat{\Pi}_{\rho_+} J = o(1)$  while*

$$\frac{1}{p} J^\top \hat{\Pi}_\rho J = \frac{1}{kc_0} \left( 1 - c_0 \frac{1}{k^2} \int \frac{u^2 \nu(du)}{(u-\ell)^2} \right) \left[ I_k - \frac{1}{k} \mathbf{1}_k \mathbf{1}_k^\top \right] + o(1).$$

If instead  $f'(\tau) \rightarrow 0$ ,  $\frac{1}{p} J^\top \hat{\Pi}_{\rho_+^0} J = o(1)$  and

$$\frac{1}{p} J^\top \hat{\Pi}_{\rho^0} J = \frac{1}{kc_0} \left[ I_k - \frac{1}{k} \mathbf{1}_k \mathbf{1}_k^\top \right] + o(1).$$

From Corollary 6, we then now have, for  $f'(\tau) \neq 0$

$$\operatorname{tr} \mathcal{D}(c^{-1}) \frac{1}{n} J^\top \hat{\Pi}_\rho J = (k-1) \left( 1 - c_0 \frac{1}{k^2} \int \frac{u^2 \nu(du)}{(u-\ell)^2} \right) + o(1)$$

which is all the closer to  $k-1$  that  $\frac{f''(\tau)}{f'(\tau)}$  is small or that  $\frac{1}{p} \operatorname{tr}((D_1 - D_2)^2)$  is large (for a given  $\operatorname{tr}((k-1)D_1 + D_2)$ ). Thus, here the Frobenius norm of  $D_1 - D_2$  is the discriminative attribute. For  $f'(\tau) \rightarrow 0$ , we obtain simply  $\operatorname{tr} \mathcal{D}(c^{-1}) \frac{1}{n} J^\top \hat{\Pi}_\rho J = (k-1) + o(1)$ , hence a perfect alignment to  $j_1, \dots, j_k$ , consistently with Remark 9.

## 6. CONCLUDING REMARKS

**6.1. Practical considerations.** One important practical outcome of our study lies in the observation that clustering can be performed selectively on either  $M$ ,  $t$ , or  $T$  by properly setting the kernel function  $f$ . That is, assume the following hierarchical scenario in which a superclass  $\mathcal{C}_i$  is identified via a constant mean  $\mu_i$  but different subclasses of covariances  $C_{i,j}$ ,  $j = 1, \dots, i_k$  for some  $i_k$ . If the objective is to discriminate only the superclasses and not each individual class, then one may design  $f$  in such a way that both  $\frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)}$  and  $\frac{f''(\tau)}{f'(\tau)}$  are sufficiently small. Alternatively, if differences in mean appear due to an improper centering of the data, and that the relevant information is carried in the covariance structure, then it is useful to take  $f'(\tau) = 0$ . This may be useful in image classification for images with different lightness and contrast.

These considerations however assume the possibility to tailor the function  $f$  according to the value taken by its successive derivatives at  $\tau$ . To ensure this is possible, note that, under Assumption 1,

$$\frac{2}{np} \sum_{i=1}^n \left\| x_i - \frac{1}{n} \sum_{j=1}^n x_j \right\|^2 \xrightarrow{\text{a.s.}} \tau$$

and it is thus possible to consistently estimate  $\tau$  and, hence, to design  $f$  to one's purposes.

**6.2. Extension to non-Gaussian settings.** The present work strongly relies (mostly for mathematical tractability) on the Gaussian assumption on the data  $x_i$ . Nonetheless, as is often the case in random matrix theory, the results can be generalized to some extent beyond the Gaussian assumption. In particular, assume now that, for  $x_i$  in class  $\mathcal{C}_a$ , we take  $x_i = \mu_a + C_a^{\frac{1}{2}} z_i$ , where  $z_i$  is a random vector with independent zero mean unit variance entries  $z_{ij}$  having at least finite kurtosis  $\kappa \triangleq \mathbb{E}[z_{ij}^4] - 3$ . Then our results may be generalized by noticing that

$$|\psi_i|^2 = \frac{2}{p} \text{tr} C_a^2 + \kappa \frac{1}{p} \text{tr} (\text{diag}(C_a)^2) + o(1)$$

almost surely, for  $x_i$  in class  $\mathcal{C}_a$ , with  $\text{diag}(C)$  the operator which sets to zero all off-diagonal elements of  $C$ . Since  $\kappa \geq -2$ , the right-hand term can take any nonnegative value, which shall impact (positively or negatively) the detectability. In particular, this will impact the performances of both scenarios where  $M = 0$  studied in Section 5.

**6.3. On the growth rate.** To better understand our data setting, it is interesting to recall the intuition of Ng–Jordan–Weiss spelled out earlier in Section 1. In a perfectly discriminable setting and for an appropriate choice of  $f$ ,  $L$  would have  $k$  eigenvalues equal to  $n$  with associated eigenspace the span of  $\{j_1, \dots, j_k\}$  while all other eigenvalues would typically remain of order  $O(1)$ . When the data become less discriminable,  $k - 1$  of these eigenvalues will become smaller with associated eigenvectors only partially aligning to  $\{j_1, \dots, j_k\}$ . This remains valid until the  $k - 1$  eigenvalues become so small that they merge with the remaining spectrum. This therefore places our study at the cluster detectability limit beyond which clustering becomes (asymptotically) theoretically infeasible. Theorem 2 thus provides here the necessary and sufficient conditions for asymptotic detectability of classes in the Gaussian data setting, which are made explicit in Section 5 in several simple scenarios.

However, the above discussion on detectability assumes  $f$  fixed from the beginning. As it turns out from our results, it is often beneficial to take  $f'(\tau)/f(\tau)$  and  $f''(\tau)/f'(\tau)$  as small as possible (see how taking  $f'(0) = 0$  benefits to alignment to  $j_1, \dots, j_k$  for instance). Taking  $f$  to be the classically used Gaussian kernel  $f(x) = \exp(-\frac{x}{2\sigma^2})$ , this implies that  $\sigma^2$  should be taken small. In turn, this suggests that a more appropriate growth regime is when  $\sigma^2$  is not fixed but vanishes with  $n$  and thus that  $f$  should adapt to  $n$ . Another remark following the same suggestion is that the Taylor expansion of  $K_{ij}$  performed to obtain Theorem 1 naturally discards all the subsequent derivatives of  $f$  along with all the next

order properties of the data (i.e., only  $M$ ,  $t$ , and  $T$  can be discriminated upon). Allowing  $f$  to adapt to  $n$  would allow for a more flexible analysis. In particular, since  $\|x_i - x_j\|^2$  has  $O(\sqrt{p})$  fluctuations around  $\tau p$ , we may expect that, taking  $f(x) = \tilde{f}((x - \tau)\sqrt{p})$  for some fixed  $\tilde{f}$  function would provide a wider dynamics range to the kernel function. This setting, which is quite challenging to study as it does not lead itself to classical random matrix analysis, remains open.

## 7. PROOFS

**7.1. Preliminary definitions and remarks.** As we will regularly deal with uniform convergences of the entries of vectors of growing size, we shall call the union bound on many instances. To this end, we need the following definitions. For  $x \equiv x_n$  a random variable and  $u_n \geq 0$ , we write  $x = O(u_n)$  if for any  $\eta > 0$  and  $D > 0$ ,  $n^D \mathbb{P}(x \geq n^\eta u_n) \rightarrow 0$  as  $n \rightarrow \infty$ . Unless specified, when  $x$  depends, besides the implicit parameter  $n$ , on other parameters (if  $x = x_{ij}$  for example), the convergences will always be supposed to be uniform in the other parameters. As a consequence of this convention, this  $O(\cdot)$  notation, besides being compatible with sums and products, has the property that the maximum of a collection of at most  $n^C$  random variables of order  $O(u_n)$  is still  $O(u_n)$ , for any constant  $C$ .

As far as multidimensional objects are concerned, let us make the notation  $O(\cdot)$  a bit more precise.

- a) When  $v$  is a vector or a diagonal matrix,  $v = O(u_n)$  means that the maximal entry of  $v$  in absolute value is  $O(u_n)$ .
- b) When  $M$  is a square matrix,  $M = O(u_n)$  means that the operator norm of  $M$  is  $O(u_n)$ .
- c) When  $M$  is a square matrix, we shall write  $M = O_{1_n}(u_n)$  when the operator norm of  $M$  is  $O(u_n)$  and the vector  $M1_n$  is  $O(u_n)$  in the sense defined above.
- d) At last, for  $x$  a vector or a matrix,  $x = o(u_n)$  means that there is  $\alpha > 0$  such that  $x = O(n^{-\alpha}u_n)$ .

Note that definition c) below is quite surprising, as  $\|1_n\| = \sqrt{n}$ , but is in fact perfectly adapted to our context, where we have some matrix error terms which, thanks to some classical probabilistic phenomena, are of smaller order when observed through the vector  $1_n$  than through their maximal eigenvector.

Note also that for  $M = [m_{ij}]_{i,j=1}^n$  a random matrix, as  $\|M\|^2 \leq \text{tr} MM^*$ , we have

$$m_{ij} = O(u_n) \implies M = O_{1_n}(nu_n). \quad (17)$$

In the remainder of the proof, to make our expressions shorter and more readable, we shall regularly use the following conventions. Matrices will be indexed by blocks according to the partition  $\mathcal{C}_1, \dots, \mathcal{C}_k$ , with  $(a, b)$  being used for block  $a$  in rows and  $b$  in columns. In

particular

$$\begin{aligned}\{A_b\}_{b=1}^k &= [A_1, \dots, A_k] \\ \{A_a\}_{a=1}^k &= [A_1^\top, \dots, A_k^\top]^\top \\ \{A_{a,b}\}_{a,b=1}^k &= \{[A_{a,1}, \dots, A_{a,k}]\}_{a=1}^k.\end{aligned}$$

## 7.2. Proof of Theorem 1.

7.2.1. *Asymptotic equivalent of the matrix  $K$ .* The main idea of the proof is to exploit the fact that, due to  $p$  going to infinity, all off-diagonal entries of  $K$  converge to the same limit in the regime of Assumption 1, which we shall demonstrate in the course of the proof to be the only non-trivial one. This will allow us to Taylor-expand each entry of  $K$  to two non-vanishing orders, whereby “non-vanishing” means that the full matrix contribution (and not only its individual entries) in this order expansion has non-vanishing operator norm. It shall in particular appear that, while on the onset one may assume that higher Taylor orders ought to contribute less than smaller orders, the structure of the full matrix expansions underlies a more subtle reasoning.

To start, we shall operate a seemingly irrelevant centering operation on the  $x_i$ 's which shall considerably help simplifying the proof. Precisely, it will turn out convenient in the following to systematically recenter the vectors  $x_i$  and  $w_i$  around their empirical mean. As such, we shall define, for  $i = 1, \dots, n$ ,

$$w_i^\circ \triangleq w_i - \frac{1}{n} \sum_{j=1}^n w_j$$

and  $W^\circ \triangleq [w_1^\circ, \dots, w_n^\circ] = WP$ .

This centering procedure leads naturally one to introduce  $\tau^\circ = \frac{n+1}{n} \frac{2}{p} \text{tr } C^\circ = \tau + O(n^{-1})$ , i.e., the expected value of  $\|w_i^\circ\|^2$ , as well as  $\psi^\circ$  defined by  $\psi_i = \|w_i^\circ\|^2 - \mathbb{E}[\|w_i^\circ\|^2] = \psi_i + o(n^{-1})$ .

Using the fact that  $w_j - w_i = w_j^\circ - w_i^\circ$ , we have, for  $x_i \in \mathcal{C}_a$  and  $x_j \in \mathcal{C}_b$ ,

$$\begin{aligned}\frac{1}{p} \|x_j - x_i\|^2 &= \|w_j - w_i\|^2 + \frac{1}{p} \|\mu_b - \mu_a\|^2 + \frac{2}{\sqrt{p}} (\mu_b - \mu_a)^\top (w_j - w_i) \\ &= \underbrace{\tau^\circ + \frac{1}{p} \text{tr } C_a^\circ}_{\triangleq A} + \underbrace{\frac{1}{p} \text{tr } C_b^\circ}_{\triangleq B} + \underbrace{\psi_j^\circ}_{\triangleq C} + \underbrace{\psi_i^\circ}_{\triangleq D} - \underbrace{2(w_i^\circ)^\top w_j^\circ}_{\triangleq E} \\ &\quad + \underbrace{\frac{\|\mu_b^\circ - \mu_a^\circ\|^2}{p}}_{\triangleq F} + \underbrace{\frac{2}{\sqrt{p}} (\mu_b^\circ - \mu_a^\circ)^\top (w_j^\circ - w_i^\circ)}_{\triangleq G}\end{aligned}$$

It is easy to see, using for example Lemmas 3 and 4, that

$$\psi_i^\circ = O(n^{-1/2}), \quad (w_i^\circ)^\top w_j^\circ = O(n^{-\frac{1}{2}}), \quad \frac{1}{\sqrt{p}} (\mu_b^\circ - \mu_a^\circ)^\top (w_j^\circ - w_i^\circ) = O(n^{-1}). \quad (18)$$

Let us then Taylor-expand  $f(\frac{1}{p}\|x_j - x_i\|^2)$  around  $\tau^\circ$  and control, in operator norm, the order of each resulting matrix term in

$$K = \left[ f \left( \frac{1}{p} \|x_j - x_i\|^2 \right) \right]_{i,j=1}^n.$$

First, by (18), (17) allows one to claim that the error term of the entry-wise Taylor expansion will give rise to an error term, in  $K$ , with operator norm  $O(n^{-1/2})$ . Following this argument, let us precisely identify the non-vanishing terms in the Taylor expansion of  $K$ . By (17), the terms associated with the entries  $F^2, G^2, (A + B + C + D + E)(F + G)$  and  $(A + B + C + D)E$  will all give rise to a  $O_{1_n}(n^{-1/2})$  error term. Besides, similar to (18), we get

$$((w_i^\circ)^\top w_j^\circ)^2 - \frac{1}{p^2} \operatorname{tr} C_a C_b = O(n^{-3/2})$$

so that the term corresponding to entry  $E^2$  can be freely replaced by  $\frac{4}{p^2} \operatorname{tr} C_a C_b$  without affecting the resulting error operator norm in the large  $n$  limit. Considering also the (trivial) diagonal terms and using the notation introduced above, we finally get, with  $(WP)_a \in \mathbb{R}^{p \times n_a}$  the submatrix of  $WP$  constituted by the columns corresponding to class  $\mathcal{C}_a$ ,

$$\begin{aligned} K &= f(\tau^\circ) 1_n 1_n^\top + f'(\tau^\circ) \left[ \psi^\circ 1_n^\top + 1_n (\psi^\circ)^\top + \left\{ \|\mu_a^\circ - \mu_b^\circ\|^2 \frac{1_{n_a} 1_{n_b}^\top}{p} \right\}_{a,b=1}^k \right. \\ &\quad + \left\{ t_a \frac{1_{n_a}}{\sqrt{p}} \right\}_{a=1}^k 1_n^\top + 1_n \left\{ t_b \frac{1_{n_b}^\top}{\sqrt{p}} \right\}_{b=1}^k + 2 \left\{ \frac{1}{\sqrt{p}} (WP)_a^\top (\mu_b^\circ - \mu_a^\circ) 1_{n_b}^\top \right\}_{a,b=1}^k \\ &\quad - 2 \left\{ \frac{1}{\sqrt{p}} 1_{n_a} (\mu_b^\circ - \mu_a^\circ)^\top (WP)_b \right\}_{a,b=1}^k - 2PW^\top WP \left. \right] \\ &\quad + \frac{f''(\tau^\circ)}{2} \left[ (\psi^\circ)^2 1_n^\top + 1_n [(\psi^\circ)^2]^\top + \left\{ t_a^2 \frac{1_{n_a}}{p} \right\}_{a=1}^k 1_n^\top + 1_n \left\{ t_b^2 \frac{1_{n_b}^\top}{p} \right\}_{b=1}^k \right. \\ &\quad + 2 \left\{ t_a t_b \frac{1_{n_a} 1_{n_b}^\top}{p} \right\}_{a,b=1}^k + 2 \mathcal{D} \{ t_a I_{n_a} \}_{a=1}^k \psi^\circ \frac{1_n^\top}{\sqrt{p}} + 2 \psi^\circ \left\{ t_b \frac{1_{n_b}^\top}{\sqrt{p}} \right\}_{b=1}^k \\ &\quad + 2 \frac{1_n}{\sqrt{p}} (\psi^\circ)^\top \mathcal{D} \{ t_a 1_{n_a} \}_{a=1}^k + 2 \left\{ t_a \frac{1_{n_a}}{\sqrt{p}} \right\}_{a=1}^k (\psi^\circ)^\top + 4 \left\{ \operatorname{tr}(C_a C_b) \frac{1_{n_a} 1_{n_b}^\top}{p^2} \right\}_{a,b=1}^k \\ &\quad \left. + 2 \psi^\circ (\psi^\circ)^\top \right] + (f(0) - f(\tau^\circ) + \tau f'(\tau^\circ)) I_n + O_{1_n}(n^{-\frac{1}{2}}) \end{aligned}$$

where we denoted  $W_a \triangleq [w_{n_1+\dots+n_{a-1}+1}, \dots, w_{n_1+\dots+n_a}]$  the restriction of  $W$  to class  $\mathcal{C}_a$  elements,  $(\psi^\circ)^2 \triangleq [(\psi^\circ)_1^2, \dots, (\psi^\circ)_n^2]^\top$ , and we recall that  $t = \{ \frac{1}{\sqrt{p}} \operatorname{tr} C_a^\circ \}_{a=1}^k$ .

We shall now differentiate the study of the cases  $f'(\tau^\circ)$  away from zero and  $f'(\tau^\circ) \rightarrow 0$ .



7.2.2. *Case  $f'(\tau^\circ)$  away from zero.* As a consequence of the above, under the assumption that  $f'(\tau^\circ) \neq 0$ , we may write

$$K = -2f'(\tau^\circ) (PW^\top WP + VAV^\top) + (f(0) - f(\tau^\circ) + \tau f'(\tau^\circ))I_n + O_{1_n}(n^{-\frac{1}{2}}) \quad (19)$$

where  $V$  is the  $n \times (2k + 4)$  matrix defined by

$$\begin{aligned} V &\triangleq \left[ \frac{J}{\sqrt{p}}, v_1, \dots, v_k, \tilde{v}, \psi^\circ, \sqrt{p}(\psi^\circ)^2, \sqrt{p}\tilde{\psi}^\circ \right] \\ v_a &\triangleq PW^\top \mu_a^\circ = O(n^{-\frac{1}{2}}) \\ \tilde{v} &\triangleq \left\{ (WP)_a^\top \mu_a^\circ \right\}_{a=1}^k = O(n^{-\frac{1}{2}}) \\ \tilde{\psi}^\circ &\triangleq \mathcal{D} \left( \left\{ t_a \frac{1_{n_a}}{\sqrt{p}} \right\}_{a=1}^k \right) \psi^\circ = O(n^{-1}) \end{aligned}$$

and  $A \triangleq A_n + A_{\sqrt{n}} + A_1$ , with  $A_n$ ,  $A_{\sqrt{n}}$  and  $A_1$  are the symmetric matrices

$$\begin{aligned} A_n &\triangleq -\frac{f(\tau^\circ)}{2f'(\tau^\circ)}p \begin{bmatrix} 1_k 1_k' & 0_{k \times k+4} \\ * & 0_{k+4 \times k+4} \end{bmatrix} \\ A_{\sqrt{n}} &\triangleq -\frac{1}{2}\sqrt{p} \begin{bmatrix} \{t_a + t_b\}_{a,b=1}^k & 0_{k \times k} & 0_{k \times 1} & 1_k & 0_{k \times 1} & 0_{k \times 1} \\ * & 0_{k \times k} & 0_{k \times 1} & 0_{k \times 1} & 0_{k \times 1} & 0_{k \times 1} \\ * & * & 0 & 0 & 0 & 0 \\ * & * & * & 0 & 0 & 0 \\ * & * & * & * & 0 & 0 \\ * & * & * & * & * & 0 \end{bmatrix} \\ A_1 &\triangleq \begin{bmatrix} A_{1,11} & I_k & -1_k & -\frac{f''(\tau^\circ)}{2f'(\tau^\circ)}t & -\frac{f''(\tau)}{4f'(\tau)}1_k & -\frac{f''(\tau)}{2f'(\tau)}1_k \\ * & 0_{k \times k} & 0_{k \times 1} & 0_{k \times 1} & 0_{k \times 1} & 0_{k \times 1} \\ * & * & 0 & 0 & 0 & 0 \\ * & * & * & -\frac{f''(\tau)}{2f'(\tau)} & 0 & 0 \\ * & * & * & * & 0 & 0 \\ * & * & * & * & * & 0 \end{bmatrix} \\ A_{1,11} &= \left\{ -\frac{1}{2} \|\mu_b^\circ - \mu_a^\circ\|^2 - \frac{f''(\tau)}{4f'(\tau)} (t_a + t_b)^2 - \frac{f''(\tau)}{f'(\tau)} \frac{1}{p} \text{tr} C_a C_b \right\}_{a,b=1}^k. \end{aligned}$$

The division of  $A$  into  $A_n$ ,  $A_{\sqrt{n}}$ , and  $A_1$  is obviously related here to the fact that  $A_n$  has operator norm  $O(n)$ ,  $A_{\sqrt{n}}$  has operator norm  $O(\sqrt{n})$ , while  $A_1$  is of order  $O(1)$ . It is already interesting to note that, from the result above,  $K$  is asymptotically equivalent to a type of spiked models in the sense that  $VAV^\top$  is of finite rank and is summed up to  $PW^\top WP$  which, under reasonable conditions, does not exhibit spikes by itself.

**Taylor expansion of  $\sqrt{n}D^{-\frac{1}{2}}$ .** We next address the expansion of  $n^{-1}D = n^{-1}\mathcal{D}(K1_n)$ . For this, using Equation (19) and the convention for  $O_{1_n}(\cdot)$ , we write

$$\begin{aligned} n^{-1}D &= -\frac{2}{n}f'(\tau^\circ)\mathcal{D}(PW^\top WP1_n + VA_nV^\top 1_n + VA_{\sqrt{n}}V^\top 1_n + VA_1V^\top 1_n) \\ &\quad + O(n^{-\frac{3}{2}}). \end{aligned} \tag{20}$$

We shall importantly use in what follows the fact that  $P1_n = 0$  which shall help discard quite a few terms (and which is the main motivation for centering  $x_i$  and  $w_i$  in the first place).

Let us first provide estimates for the quantities involving  $A$  and  $V$  that we shall develop. In particular, with the estimate

$$V^\top 1_n = \begin{bmatrix} \left\{ \frac{n_a}{\sqrt{p}} \right\}_{a=1}^k \\ \left\{ (\mu_a^\circ)^\top WP1_n \right\}_{a=1}^k \\ \sum_{a=1}^k (\mu_a^\circ)^\top (WP)_a 1_n \\ (\psi^\circ)^\top 1_n \\ \sqrt{p}[(\psi^\circ)^2]^\top 1_n \\ \sqrt{p}(\tilde{\psi}^\circ)^\top 1_n \end{bmatrix} = \begin{bmatrix} O(n^{\frac{1}{2}}) \\ 0 \\ O(1) \\ O(1) \\ O(n^{\frac{1}{2}}) \\ O(1) \end{bmatrix},$$

we deduce

$$\begin{aligned} VA_nV^\top 1_n &= -\frac{f(\tau^\circ)}{2f'(\tau^\circ)}n1_n = O(n) \\ VA_{\sqrt{n}}V^\top 1_n &= \underbrace{-\frac{n}{2}\sum_{a=1}^k t_a \frac{j_a}{\sqrt{p}}}_{O(n^{\frac{1}{2}})} - \underbrace{\frac{1}{2}n\psi}_{O(n^{\frac{1}{2}})} - \underbrace{\frac{1}{2}(1_n^\top \psi^\circ)1_n}_{O(1)} \\ VA_1V^\top 1_n &= \underbrace{\frac{1}{p}\sum_{a,b=1}^k (A_{1,11})_{ab}n_b j_a}_{O(1)} + \underbrace{\frac{1}{\sqrt{p}}\sum_{a=1}^k n_a v_a}_{O(1)} - \underbrace{\frac{n}{\sqrt{p}}\tilde{v}}_{O(1)} - \underbrace{n\frac{f''(\tau^\circ)}{4f'(\tau^\circ)}(\psi^\circ)^2}_{O(1)} \\ &\quad - \underbrace{n\frac{f''(\tau^\circ)}{2f'(\tau^\circ)}\tilde{\psi}^\circ}_{O(1)} - \underbrace{\frac{f''(\tau^\circ)}{2f'(\tau^\circ)}\sum_{a=1}^k \frac{n_a}{\sqrt{p}}t_a \psi^\circ}_{O(1)} - \underbrace{\frac{f''(\tau^\circ)}{4f'(\tau^\circ)}(1_n^\top (\psi^\circ)^2)1_n}_{O(1)} + O(n^{-\frac{1}{2}}). \end{aligned}$$

Besides, applying  $\mathbf{1}_n^\top$  to the above estimates gives

$$\begin{aligned} \mathbf{1}_n^\top V A_{\sqrt{n}} V^\top \mathbf{1}_n &= 0 \\ \mathbf{1}_n^\top V A_1^\top V \mathbf{1}_n &= \frac{1}{p} \sum_{a,b=1}^k n_a n_b (A_{1,11})_{a,b} - n \frac{f''(\tau^\circ)}{2f'(\tau^\circ)} \mathbf{1}_n^\top (\psi^\circ)^2 + O(n^{\frac{1}{2}}) \\ &= -\frac{1}{p} \sum_{a,b=1}^n n_a n_b \left[ \frac{f''(\tau^\circ)}{4f'(\tau^\circ)} \left\{ (t_a + t_b)^2 + \frac{2}{p^2} \text{tr}(C_a + C_b)^2 \right\} \right. \\ &\quad \left. + \frac{1}{2} \|\mu_b^\circ - \mu_a^\circ\|^2 \right] + O(n^{\frac{1}{2}}). \end{aligned}$$

Finally, recalling that  $WW^\top = \frac{1}{p} \sum_{a=1}^k C_a^{\frac{1}{2}} Z_a Z_a^\top C_a^{\frac{1}{2}}$ , with  $\|C_a\|$  bounded and  $Z_a$  standard Gaussian independent across  $k$ , we get from [BS98] that  $\|WW^\top\| \leq \sum_{a=1}^k \|C_a^{\frac{1}{2}} Z_a Z_a^\top C_a^{\frac{1}{2}}\| = O(1)$ .

Getting back to  $n^{-1}D$ , with the above estimates, identifying  $\mathcal{D}(V A_n V^\top \mathbf{1}_n) = -\frac{f(\tau^\circ)}{2f'(\tau^\circ)} I_n$  as the leading order term, we have

$$n^{-1}D = f(\tau^\circ) \left[ I_n - \frac{2f'(\tau^\circ)}{nf(\tau^\circ)} \mathcal{D}(V A_{\sqrt{n}} V^\top \mathbf{1}_n + V A_1 V^\top \mathbf{1}_n) \right] + O(n^{-\frac{3}{2}})$$

so that, by a further Taylor expansion,

$$\begin{aligned} \sqrt{n}D^{-\frac{1}{2}} &= \frac{1}{\sqrt{f(\tau^\circ)}} \left[ I_n + \frac{f'(\tau^\circ)}{f(\tau^\circ)n} \mathcal{D}(V A_{\sqrt{n}} V^\top \mathbf{1}_n + V A_1 V^\top \mathbf{1}_n) \right. \\ &\quad \left. + \frac{3}{2} \left( \frac{f'(\tau^\circ)}{f(\tau^\circ)n} \right)^2 \mathcal{D}^2(V A_{\sqrt{n}} V^\top \mathbf{1}_n) - \frac{f(0) - f(\tau^\circ) + \tau f'(\tau^\circ)}{2f(\tau^\circ)n} I_n \right] + O(n^{-\frac{3}{2}}) \quad (21) \end{aligned}$$

where  $\mathcal{D}^2$  stands for the squared diagonal matrix.

**Taylor expansion of  $L$ .** With the Taylor expansions of  $K$  and  $D^{-\frac{1}{2}}$  in hand, we may now obtain the Taylor expansion of their left- and right-products, so to retrieve that of  $L$ . Using the sub-multiplicativity of the operator norm, we precisely find from (19) and (21)

$$\begin{aligned} &\sqrt{n}D^{-\frac{1}{2}}K \\ &= \frac{1}{\sqrt{f(\tau^\circ)}} \left[ -2f'(\tau^\circ) V A_n V^\top - 2f'(\tau^\circ) V A_{\sqrt{n}} V^\top - 2\frac{f'(\tau^\circ)^2}{nf(\tau^\circ)} \mathcal{D}(V A_{\sqrt{n}} V^\top \mathbf{1}_n) V A_n V^\top \right. \\ &\quad - 2f'(\tau^\circ) V A_1 V^\top + (f(0) + f(\tau^\circ) + \tau^\circ f'(\tau^\circ)) I_n - 2\frac{f'(\tau^\circ)^2}{nf(\tau^\circ)} \mathcal{D}(V A_{\sqrt{n}} V^\top \mathbf{1}_n) V A_{\sqrt{n}} V^\top \\ &\quad - \frac{2f'(\tau^\circ)^2}{nf(\tau^\circ)} \mathcal{D}(V A_1 V^\top \mathbf{1}_n) V A_n V^\top + \frac{f'(\tau^\circ)}{nf(\tau^\circ)} (f(0) - f(\tau^\circ) + \tau^\circ f'(\tau^\circ)) V A_n V^\top \\ &\quad \left. - 2f'(\tau^\circ) P W^\top W P - 3\frac{f'(\tau^\circ)^3}{n^2 f(\tau^\circ)^2} \mathcal{D}^2(V A_{\sqrt{n}} V^\top \mathbf{1}_n) V A_n V^\top \right] + O(n^{-\frac{1}{2}}). \end{aligned}$$

With the same strategy, we then have finally

$$\begin{aligned}
& nD^{-\frac{1}{2}}KD^{-\frac{1}{2}} \\
&= \frac{1}{f(\tau^\circ)} \left[ -2f'(\tau^\circ)VA_nV^\top - 2f'(\tau^\circ)VA_{\sqrt{n}}V^\top - \frac{2f'(\tau^\circ)^2}{nf(\tau^\circ)} \mathcal{D}(VA_{\sqrt{n}}V^\top 1_n)VA_nV^\top \right. \\
&\quad - \frac{2f'(\tau^\circ)^2}{nf(\tau^\circ)} VA_nV^\top \mathcal{D}(VA_{\sqrt{n}}V^\top 1_n) - 2f'(\tau^\circ)VA_1V^\top + (f(0) - f(\tau^\circ) + \tau^\circ f'(\tau^\circ))I_n \\
&\quad - 2\frac{f'(\tau^\circ)^2}{nf(\tau^\circ)} \mathcal{D}(VA_{\sqrt{n}}V^\top 1_n)VA_{\sqrt{n}}V^\top - 2\frac{f'(\tau^\circ)^2}{nf(\tau^\circ)} VA_{\sqrt{n}}V^\top \mathcal{D}(VA_{\sqrt{n}}V^\top 1_n) \\
&\quad - 2\frac{f'(\tau^\circ)^2}{nf(\tau^\circ)} \mathcal{D}(VA_1V^\top 1_n)VA_nV^\top - 2\frac{f'(\tau^\circ)^2}{nf(\tau^\circ)} VA_nV^\top \mathcal{D}(VA_1V^\top 1_n) \\
&\quad - 3\frac{f'(\tau^\circ)^3}{n^2f(\tau^\circ)^2} \mathcal{D}^2(VA_{\sqrt{n}}V^\top 1_n)VA_nV^\top - 3\frac{f'(\tau^\circ)^3}{n^2f(\tau^\circ)^2} VA_nV^\top \mathcal{D}^2(VA_{\sqrt{n}}V^\top 1_n) \\
&\quad \left. + 2\frac{f'(\tau^\circ)}{nf(\tau^\circ)} (f(0) - f(\tau^\circ) + \tau^\circ f'(\tau^\circ))VA_nV^\top - 2f'(\tau^\circ)PW^\top WP \right] + O(n^{-\frac{1}{2}}).
\end{aligned}$$

**Taylor expansion of  $L'$ .** At this point, it is worth mentioning that, although absolutely not fathomable from the expression above, computer simulations suggest that the spectrum of  $L = nD^{-\frac{1}{2}}KD^{-\frac{1}{2}}$  is composed of an isolated eigenvalue of magnitude  $O(n)$  and importantly of  $n-1$  eigenvalues of order  $O(1)$ . Nonetheless, surprisingly at first, the above approximation of  $L$  still contains terms of order  $O(\sqrt{n})$ ; this is explained by the fact that those terms result from the Taylor expansion of the leading eigenspace of dimension one. Fortunately, we precisely know the leading eigenvector of  $L$  to be  $D^{\frac{1}{2}}1_n$ , and thus we may project  $L$  orthogonally to it without affecting the eigenvalue-eigenvector pairs, but for this single isolated eigenvector. This will allow us to retrieve a matrix,  $L'$ , the eigenvalues of which are expected to be all of order  $O(1)$ .

Let us start by evaluating the vector  $D^{\frac{1}{2}}1_n$  (which is simply the diagonal matrix  $D^{\frac{1}{2}}$  turned to a vector). We have

$$\begin{aligned}
D^{\frac{1}{2}}1_n &= \sqrt{nf(\tau^\circ)} \left[ 1_n - \frac{f'(\tau^\circ)}{nf(\tau^\circ)} (VA_{\sqrt{n}}V^\top 1_n + VA_1V^\top 1_n) \right. \\
&\quad \left. - \frac{1}{2} \left( \frac{f'(\tau^\circ)}{nf(\tau^\circ)} \right)^2 \mathcal{D}(VA_{\sqrt{n}}V^\top 1_n)VA_{\sqrt{n}}V^\top 1_n + \frac{f(0) - f(\tau^\circ) + \tau^\circ f'(\tau^\circ)}{2nf(\tau^\circ)} 1_n + O(n^{-\frac{3}{2}}) \right].
\end{aligned}$$

Then, applying  $1_n^\top$  and  $1_n$  on each side of  $D$  (or alternatively, taking the squared norm of the above), we get

$$\begin{aligned}
1_n^\top D 1_n &= n^2 f(\tau) \left[ 1 - \frac{2f'(\tau)}{n^2 f(\tau)} (1_n^\top VA_{\sqrt{n}}V^\top 1_n + 1_n^\top VA_1V^\top 1_n) \right. \\
&\quad \left. + \frac{f(0) - f(\tau) + \tau f'(\tau)}{nf(\tau)} + O(n^{-\frac{3}{2}}) \right]
\end{aligned}$$

so that in particular

$$\begin{aligned} \frac{1}{1_n' D 1_n} &= \frac{1}{n^2 f(\tau)} \left[ 1 + \frac{2f'(\tau)}{n^2 f(\tau)} (1_n' V A_{\sqrt{n}} V' 1_n + 1_n' V A_1 V' 1_n) \right. \\ &\quad \left. - \frac{f(0) - f(\tau) + \tau f'(\tau)}{n f(\tau)} + \frac{4f'(\tau)^2}{n^4 f(\tau)^2} (1_n' V A_{\sqrt{n}} V' 1_n)^2 \right] + O(n^{-\frac{7}{2}}) \end{aligned}$$

and

$$\begin{aligned} &n \frac{D^{\frac{1}{2}} 1_n 1_n D^{\frac{1}{2}}}{1_n' D 1_n} \\ &= 1_n 1_n' - \frac{f'(\tau^\circ)}{n f(\tau^\circ)} (V A_{\sqrt{n}} V' 1_n 1_n' + 1_n 1_n' V A_{\sqrt{n}} V') + \frac{2f'(\tau^\circ)}{n^2 f(\tau^\circ)^2} 1_n' V A_{\sqrt{n}} V' 1_n \cdot 1_n 1_n' \\ &\quad - \frac{f'(\tau^\circ)}{n f(\tau^\circ)} (V A_1 V' 1_n 1_n' + 1_n 1_n' V A_1 V') + \frac{f'(\tau^\circ)^2}{n^2 f(\tau^\circ)^2} V A_{\sqrt{n}} V' 1_n 1_n' V A_{\sqrt{n}} V' \\ &\quad - \frac{1}{2} \frac{f'(\tau^\circ)^2}{n^2 f(\tau^\circ)^2} (\mathcal{D}(V A_{\sqrt{n}} V' 1_n) V A_{\sqrt{n}} V' 1_n 1_n' + 1_n 1_n' V A_{\sqrt{n}} V' \mathcal{D}(V A_{\sqrt{n}} V' 1_n)) \\ &\quad - \frac{2f'(\tau^\circ)^2}{n^3 f(\tau^\circ)^2} (V A_{\sqrt{n}} V' 1_n 1_n' + 1_n 1_n' V A_{\sqrt{n}} V') 1_n' V A_{\sqrt{n}} V' 1_n \\ &\quad + 2 \frac{f'(\tau^\circ)}{n^2 f(\tau^\circ)} (1_n' V A_1 V' 1_n) 1_n 1_n' + 4 \frac{f'(\tau^\circ)^2}{n^4 f(\tau^\circ)^2} (1_n' V A_{\sqrt{n}} V' 1_n)^2 1_n 1_n' + O(n^{-\frac{1}{2}}). \end{aligned}$$

From the above and the previously derived expression of  $L = n D^{-\frac{1}{2}} K D^{-\frac{1}{2}}$ , we finally retrieve the expression for  $L' = n D^{-\frac{1}{2}} K D^{-\frac{1}{2}} - n \frac{D^{\frac{1}{2}} 1_n 1_n D^{\frac{1}{2}}}{1_n' D 1_n}$ . Before proceeding to the full calculus, let us focus on the terms of operator norm of order  $n$  and  $\sqrt{n}$ .

The only term of order  $n$  arises from  $-2f'(\tau) V A_n V' = f(\tau) 1_n 1_n'$ , which is present in both  $L$  and  $n \frac{D^{\frac{1}{2}} 1_n 1_n D^{\frac{1}{2}}}{1_n' D 1_n}$  and thus vanishes in  $L'$ . More interesting are the terms of order  $n^{\frac{1}{2}}$ . These sum in  $L'$  as

$$\frac{2f'(\tau^\circ)}{f(\tau^\circ)} \left[ -V A_{\sqrt{n}} V' + \frac{1}{n} V A_{\sqrt{n}} V' 1_n 1_n' + \frac{1}{n} 1_n 1_n' V A_{\sqrt{n}} V' - \frac{1}{n^2} 1_n' V A_{\sqrt{n}} V' 1_n \cdot 1_n 1_n' \right].$$

Note here that  $-2\sqrt{p} V A_{\sqrt{n}} V'$  is the sum of the matrix  $\sqrt{p} \{(t_a + t_b) 1_{n_a} 1_{n_b}'\}_{a,b=1}^k$  with the matrix  $\{\psi_i + \psi_j\}_{i,j=1}^n$ . But, as can be checked by direct computation, for  $y \in \mathbb{R}^n$  a vector and  $Y$  a matrix defined by  $Y = \{y_i + y_j\}_{i,j=1}^n$ , we have  $Y - Y \frac{1_n 1_n'}{n} - \frac{1_n 1_n'}{n} Y + \frac{1}{n} 1_n' Y 1_n \frac{1_n 1_n'}{n} = 0$ . Hence, it follows that the terms of order  $\sqrt{n}$  in  $L'$  vanish.

We thus obtain a first conclusion, which corroborate the aforementioned simulations results, about the spectrum of  $L'$  being almost surely of operator norm  $O(1)$ . And thus, the spectrum of  $L$  is composed of an isolated eigenvalue equal to  $n$  and of  $n - 1$  remaining eigenvalues of order  $O(1)$ .

Let us now clarify the resulting expression for  $L'$ . Although the terms to be considered are apparently numerous, similar to  $K$ , they all contribute to a small rank matrix but for

$PW^\top WP$ , and thus we shall write

$$nD^{-\frac{1}{2}}KD^{-\frac{1}{2}} - n\frac{D^{\frac{1}{2}}1_n1_nD^{\frac{1}{2}}}{1_n^\top D 1_n} = -2\frac{f'(\tau^\circ)}{f(\tau^\circ)} [PW^\top WP + UBU^\top] + O(n^{-\frac{1}{2}})$$

for a small rank perturbation matrix  $UBU^\top$ ,  $B$  symmetric, with importantly  $O(1)$  operator norm. As such, note already that, since  $\tau^\circ = \tau + o(1)$  and  $\psi_i^\circ = \psi_i + o(n^{-1})$ , we may freely replace  $\tau^\circ$  by  $\tau$  and  $\psi_i^\circ$  by  $\psi_i$  in what follows, to the expense of  $o(1)$  in operator norm. We shall use this simpler notation from now on.

With this remark in mind, let us establish minimal descriptions for  $U$  and  $B$ . After development of the remaining subparts of  $L$ , excluding the term proportional to  $PW^\top WP$ , we obtain

$$\begin{aligned} & -\frac{2f'(\tau)}{f(\tau)}VA_1V^\top + 2\frac{f'(\tau)}{nf(\tau)}(VA_1V^\top 1_n 1_n^\top + 1_n 1_n^\top VA_1V^\top) \\ & -2\frac{f'(\tau)}{n^2f(\tau)}1_n^\top VA_1V^\top 1_n \cdot 1_n 1_n^\top \\ & = -2\frac{f'(\tau)}{f(\tau)}\sum_{a,b=1}^k\frac{j_a j_b^\top}{p}\left[[M^\top M]_{ab} + \frac{f''(\tau)}{2f'(\tau)}t_a t_b - \frac{f''(\tau)}{f'(\tau)}T_{ab}\right] \\ & -2\frac{f'(\tau)}{f(\tau)}\sum_{a,b=1}^k\left(\delta_{ab} - \frac{n_a}{n}\right)v_a\frac{j_b^\top}{\sqrt{p}} - 2\frac{f'(\tau)}{f(\tau)}\sum_{a,b=1}^k\left(\delta_{ab} - \frac{n_b}{n}\right)\frac{j_a}{\sqrt{p}}v_b^\top \\ & + \frac{f''(\tau)}{f(\tau)}\sum_{a=1}^k t_a\frac{j_a}{\sqrt{p}}\psi^\top + \frac{f''(\tau)}{f(\tau)}\sum_{a=1}^k t_a\psi\frac{j_a^\top}{\sqrt{p}} + \frac{f''(\tau)}{f(\tau)}\psi\psi^\top + O(n^{-\frac{1}{2}}) \end{aligned}$$

and for the remaining subparts of  $n\frac{D^{\frac{1}{2}}1_n1_n^\top D^{\frac{1}{2}}}{1_n^\top D 1_n}$ ,

$$\begin{aligned} & \frac{2f'(\tau)}{n^2f(\tau)^2}\left[\mathcal{D}(VA_{\sqrt{n}}V^\top 1_n)VA_{\sqrt{n}}V^\top 1_n 1_n^\top + 1_n 1_n^\top VA_{\sqrt{n}}V^\top \mathcal{D}(VA_{\sqrt{n}}V^\top 1_n) \right. \\ & - \frac{1}{2}VA_{\sqrt{n}}V^\top 1_n 1_n^\top VA_{\sqrt{n}}V^\top \frac{1}{n}(VA_{\sqrt{n}}V^\top 1_n 1_n^\top + 1_n 1_n^\top VA_{\sqrt{n}}V^\top)1_n^\top VA_{\sqrt{n}}V^\top 1_n \\ & \left. - \frac{2}{n^2}(1_n^\top VA_{\sqrt{n}}V^\top 1_n)^2 1_n 1_n^\top - n(\mathcal{D}(VA_{\sqrt{n}}V^\top 1_n)VA_{\sqrt{n}}V^\top + VA_{\sqrt{n}}V^\top \mathcal{D}(VA_{\sqrt{n}}V^\top 1_n))\right] \\ & = -\sum_{a,b=1}^k\frac{5f'(\tau)^2}{4f(\tau)^2}T_{ab}\frac{j_a j_b^\top}{p} - \sum_{a=1}^k\frac{5f'(\tau)^2}{4f(\tau)^2}t_a\left(\frac{j_a}{\sqrt{p}}\psi^\top + \psi\frac{j_a^\top}{\sqrt{p}}\right) - \frac{5f'(\tau)^2}{4f(\tau)^2}\psi\psi^\top + O(n^{-\frac{1}{2}}). \end{aligned}$$

Altogether, recalling the notations  $c_a = \frac{n_a}{n}$ ,  $c = \{c_a\}_{a=1}^k$ , and  $c_0 = \frac{p}{n}$ , this is finally

$$L' = -2\frac{f'(\tau)}{f(\tau)} [PW^\top WP + UBU^\top] + \frac{f(0) - f(\tau) + \tau f'(\tau)}{f(\tau)} I_n + O(n^{-\frac{1}{2}})$$

where

$$U = \left[ \frac{j_1}{\sqrt{p}}, \dots, \frac{j_k}{\sqrt{p}}, v_1, \dots, v_k, \psi \right]$$

and

$$B = \begin{bmatrix} B_{11} & I_k - \mathbf{1}_k c^\top & \left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right) t \\ I_k - c \mathbf{1}_k^\top & \mathbf{0}_{k \times k} & \mathbf{0}_{k \times 1} \\ \left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right) t^\top & \mathbf{0}_{1 \times k} & \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \end{bmatrix}$$

where

$$B_{11} = M^\top M + \left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right) t t^\top - \frac{f''(\tau)}{f'(\tau)} T + c_0 \frac{f(0) - f(\tau) + \tau f'(\tau)}{2f'(\tau)} \mathbf{1}_k \mathbf{1}_k^\top.$$

This concludes the proof for the case where  $f'(\tau^\circ)$  is away from zero.

7.2.3. *Case  $f'(\tau^\circ) \rightarrow 0$ .* Reproducing similar step as above (without factoring  $f'(\tau^\circ)$  in both  $A_n$  and  $A_{\sqrt{n}}$ ), when  $f'(\tau^\circ) \rightarrow 0$ , we have the simpler following expression (which happens to correspond to the  $f'(\tau) \rightarrow 0$  limit of the previous formula).

$$L = U_0 B_0 U_0^\top + \frac{f(0) - f(\tau)}{f(\tau)} I_n + O(n^{-\frac{1}{2}})$$

where

$$\begin{aligned} U_0 &= \left[ \frac{j_1}{\sqrt{p}}, \dots, \frac{j_k}{\sqrt{p}}, \psi \right] \\ B_0 &= \begin{bmatrix} B_{0;11} & \frac{f''(\tau)}{f(\tau)} t \\ \frac{f''(\tau)}{f(\tau)} t^\top & \frac{f''(\tau)}{f(\tau)} \end{bmatrix} \\ B_{0;11} &= \frac{f''(\tau)}{f(\tau)} t t^\top + \frac{2f''(\tau)}{f(\tau)} T - c_0 \frac{f(0) - f(\tau)}{f(\tau)} \mathbf{1}_k \mathbf{1}_k^\top. \end{aligned}$$

7.3. **Proof of Proposition 1.** Recall from our previous computations that

$$\begin{aligned} D^{\frac{1}{2}} \mathbf{1}_n &= \sqrt{nf(\tau^\circ)} \left[ \mathbf{1}_n - \frac{f'(\tau^\circ)}{nf(\tau^\circ)} (V A_{\sqrt{n}} V^\top \mathbf{1}_n + V A_1 V^\top \mathbf{1}_n) \right. \\ &\quad - \frac{1}{2} \left( \frac{f'(\tau^\circ)}{nf(\tau^\circ)} \right)^2 \mathcal{D} (V A_{\sqrt{n}} V^\top \mathbf{1}_n) V A_{\sqrt{n}} V^\top \mathbf{1}_n \\ &\quad \left. + \frac{f(0) - f(\tau^\circ) + \tau^\circ f'(\tau^\circ)}{2nf(\tau^\circ)} \mathbf{1}_n + O(n^{-\frac{3}{2}}) \right] \end{aligned}$$

and that

$$\frac{1}{\sqrt{1_n^\top D 1_n}} = \frac{1}{n\sqrt{f(\tau^\circ)}} \left[ 1 + \frac{f'(\tau^\circ)}{n^2 f(\tau^\circ)} (1_n^\top V A_{\sqrt{n}} V^\top 1_n + 1_n^\top V A_1 V^\top 1_n) - \frac{f(0) - f(\tau^\circ) + \tau^\circ f'(\tau^\circ)}{2n f(\tau^\circ)} + \frac{3}{2} \frac{f'(\tau^\circ)^2}{n^4 f(\tau^\circ)^2} (1_n^\top V A_{\sqrt{n}} V^\top 1_n)^2 \right] + O(n^{-\frac{5}{2}}).$$

Taking the product, we therefore find

$$\sqrt{np} \left( \frac{D^{\frac{1}{2}} 1_n}{\sqrt{1_n^\top D 1_n}} - \frac{1}{\sqrt{n}} 1_n \right) = \frac{f'(\tau^\circ)}{2f(\tau^\circ)} \left[ \{t_a 1_{n_a}\}_{a=1}^k + \sqrt{p} \psi \right] + O(n^{-\frac{1}{2}})$$

where the RHS dominant term is a random vector of independent entries with mean

$$\frac{f'(\tau^\circ)}{2f(\tau^\circ)} \{t_a 1_{n_a}\}_{a=1}^k$$

and covariance matrix

$$\left( \frac{f'(\tau^\circ)}{2f(\tau^\circ)} \right)^2 \mathcal{D} \left\{ \frac{2}{p} \text{tr}(C_a^2) I_{n_a} \right\}_{a=1}^k + o(n^{-1}).$$

Besides, each entry is asymptotically Gaussian (possibly with null variance) by the central limit theorem under Lindberg's condition. Here again, since  $\tau^\circ = \tau + o(1)$ , the result still holds if we replace  $\tau^\circ$  by  $\tau$ , and we obtain the sought for statement of Proposition 1.

As a next step for the understanding of the inner structure of  $L$ , we need to explore the behavior of  $PW^\top WP$ , which is provided by Lemma 1 and further by Lemma 2, which are proved next.

**7.4. Proofs of Lemma 1 and Lemma 2.** The derivation of the fundamental equations for  $z \in \mathbb{C}^+$  in both Lemma 1 and Lemma 2 follows from standard Gaussian calculus, introduced in [PS11]. In the companion paper [BC15], we provide in full the derivations for the matrix model  $W^\top W$ . The adaption to  $PW^\top WP$  follows the same lines and is thus not detailed any further here.

It remains to show the second part of Lemma 1. In [BC15], it is shown that the eigenvalues of  $W^\top W$  asymptotically do not escape the support  $\mathcal{S}_p$  of the measure  $\bar{\mathbb{P}}_p$ . As  $PW^\top WP$  is merely a rank-two perturbation of  $W^\top W$ , the spectrum analysis of the former then boils down to a standard (multiplicative) spiked model analysis, as carried out in e.g., [BN12]. It is there easily seen that, for  $\lambda$  at macroscopic distance from  $\mathcal{S}_p$ , the equation  $0 = \det(-\lambda I_p + PW^\top WP)$  is equivalent for all large  $n$ , almost surely, to  $0 = \lambda \frac{1}{n} 1_n^\top (W^\top W - \lambda I_n)^{-1} 1_n$  (see the proof of Theorem 2 for similar derivations). As the right-hand side expression is asymptotically equal to  $\lambda g^\circ(\lambda)$ , by the results of [BC15], the eigenvalues of  $PW^\top WP$  are therefore for all large  $n$  either in a close neighborhood of  $\mathcal{S}_p$  or close to  $\lambda$  such that  $\lambda g^\circ(\lambda) = 0$ . This proves Lemma 1.



**7.5. Proof of Theorem 2.** Suppose here that  $f'(\tau)$  is away from zero. According to Lemma 1, for all  $p$  large, almost surely, there exists no eigenvalue of  $PW^\top WP$  at macroscopic distance from  $\mathcal{S}_p \cup \mathcal{G}_p$ . Let  $z \in \mathbb{C}$  be away from  $\mathcal{S}_p \cup \mathcal{G}_p$ . We wish to solve the equation in  $z$

$$0 = \det(PW^\top WP + UBU^\top - zI_n)$$

which, up to multiplication by  $-\frac{2f'(\tau)}{f(\tau)}$  and addition of  $\frac{2f'(\tau)}{f(\tau)}F(\tau)$ , provides the isolated eigenvalues of  $\hat{L}'$ . Factoring out  $PW^\top WP - zI_n$  (of smallest absolute eigenvalue away from zero) and using Sylvester's identity, this is equivalent to solving for all large  $n$ , almost surely

$$0 = \det(I_{2k+1} + BU^\top Q_z U)$$

(with the notations from Lemma 1). We now need to retrieve a deterministic equivalent for  $I_{2k+1} + BU^\top Q_z U$ . We shall proceed by first approximating deterministically every subblock of  $U^\top Q_z U$ . From Lemma 1,  $\frac{1}{p}J^\top Q_z J$  is readily obtained as  $\frac{1}{p}J^\top \bar{Q}_z J + o(1)$ , with  $\frac{1}{p}J^\top \bar{Q}_z J = \Gamma_z - \frac{1}{zc_0}cc^\top$ . As for  $(\mu_i^\circ)^\top WQ_z W^\top \mu_j^\circ$ , note, with our previous estimators, that this is  $(\mu_i^\circ)^\top WPQ_z PW^\top \mu_j^\circ + o(1)$  (as  $\frac{1}{\sqrt{n}}1_n^\top W^\top \mu_j^\circ = o(1)$ ). Hence, using  $WPQ_z PW^\top = \tilde{Q}_z W P W^\top = I_p + z\tilde{Q}_z$ , we get  $(\mu_i^\circ)^\top WQ_z W^\top \mu_j^\circ = (\mu_i^\circ)^\top \mu_j^\circ + z(\mu_i^\circ)^\top \tilde{Q}_z \mu_j^\circ + o(1)$ , which then is easily estimated from Lemma 1.

The main technical difficulty arises for  $\psi$  which depends clearly on  $W$ , but which in fact behaves, as far as our estimators are concerned, as if it were independent. From Proposition 1 and Remark 6, denoting  $\mathcal{D} = \mathcal{D}(\sqrt{\frac{1}{p} \text{tr} C_a^2 1_{n_a}})_{a=1}^k$ ,  $\psi^\top Q_z \psi = \varphi^\top \mathcal{D} Q_z \mathcal{D} \varphi + o(1)$  for some  $\varphi$  having i.i.d. zero mean,  $1/p$ -variance entries. Although  $\varphi$  is not independent of  $W$ , we can show that  $\varphi^\top \mathcal{D} Q_z \mathcal{D} \varphi = \varphi^\top \mathcal{D} \bar{Q}_z \mathcal{D} \varphi + o(1)$ , which, by the independence of the entries of  $\varphi$ , all of variance  $1/p$ , leads to  $\varphi^\top \mathcal{D} Q_z \mathcal{D} \varphi = \frac{1}{p} \text{tr} \mathcal{D} \bar{Q}_z \mathcal{D} + o(1)$ , which is simply  $\sum_{i=1}^k \frac{c_i}{c_0} g_i(z) \frac{1}{p} \text{tr} C_i^2 + o(1)$ . To obtain this fact rigorously, we may, as in the proof of Lemmas 1 and 2 (see details in [BC15]), exploit the Gaussian integration-by-parts and Nash–Poincaré inequality method [PS11]; precisely, we obtain that  $\mathbb{E}[\psi^\top Q_z \psi] = \frac{1}{p} \text{tr} \mathcal{D} \bar{Q}_z \mathcal{D} + O(p^{-1})$  and  $\mathbb{E}[(\psi^\top Q_z \psi - \mathbb{E}[\psi^\top Q_z \psi])^m] = O(p^{-\frac{m}{2}})$ , from which the result unfolds. The calculus is however painstaking and is not further detailed.

Finally we show that all (block) cross-terms of  $U^\top Q_z U$  vanish. To this end, with  $W = p^{-\frac{1}{2}}[C_1^{\frac{1}{2}}Z_1, \dots, C_k^{\frac{1}{2}}Z_k]$ , one may use the polar decomposition  $Z_i = O_{r,i}\Delta_i O_{l,i}^\top$  with  $O_{r,i}, \Delta_i^\top, O_{l,i}$  independent and  $O_{l,i}, O_{r,i}$  Haar-distributed on the orthogonal group. With this notation, it is easily shown by conditioning over the  $\Delta_i$  and  $O_{l,i}^\top$  that  $u^\top Q_z W^\top v$  can be written as the inner product of a bounded norm deterministic vector and a unitarily invariant random vector, as long as  $u, v$  are independent of  $O_{r,i}$ , with  $\|u\| = O(1), \|v\| = O(1)$ . This implies by standard results that  $u^\top Q_z W^\top v = o(1)$ . This readily implies that  $\frac{1}{\sqrt{p}}J^\top Q_z W^\top M = o(1)$ . Similarly, with the same extra care as above to account for the dependence between  $\psi$  and  $W$ , we get  $\psi^\top Q_z W^\top M = o(1)$ , as well as  $\frac{1}{\sqrt{p}}\psi^\top Q_z J = o(1)$ .

Summarizing, this is

$$U^\top Q_z U = \begin{bmatrix} \Gamma_z - \frac{cc^\top}{zc_0} & 0_{k \times k} & 0_{k \times 1} \\ 0_{k \times k} & M^\top M + zM^\top \bar{Q}_z M & 0_{k \times 1} \\ 0_{1 \times k} & 0_{1 \times k} & \sum_{i=1}^k \frac{c_i}{c_0} g_i(z) \frac{1}{p} \operatorname{tr} C_i^2 \end{bmatrix} + o(1). \quad (22)$$

Proceeding to the complete calculus of the deterministic approximation for  $I_{2k+1} + BU^\top Q_z U$ , we obtain  $I_{2k+1} + BU^\top Q_z U = H_z + o(1)$ , where

$$H_z \triangleq \begin{bmatrix} H_{11} & M^\top (I_p + z\bar{Q}_z) M & (h(\tau, z) - 1)t \\ \Gamma_z & I_k & 0_{k \times 1} \\ \left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right) t^\top \Gamma_z & 0_{1 \times k} & h(\tau, z) \end{bmatrix} \quad (23)$$

$$H_{11} = I_k + \left[ M^\top M + \left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right) tt^\top - \frac{f''(\tau)}{f'(\tau)} T \right] \Gamma_z - F(\tau) 1_k c^\top.$$

Our objective is now to find the solutions to  $\det H_z = 0$  to then prove that the eigenvalues of  $PW^\top WP + UBU^\top$  are asymptotically those real  $z$ 's cancelling the determinant of  $H_z$ .

At this point, two cases must be differentiated, according to whether  $h(\tau, z) \rightarrow 0$  or not. Let us start with the more interesting  $h(\tau, z)$  away from zero case.

7.5.1.  *$h(\tau, z)$  away from zero.* In this scenario, using the Schur complement formula, with obvious block-wise notations following the structure of (23), we have

$$\begin{aligned} \det H_z &= H_{33} \det \left( H_{11} - [H_{12} \ H_{13}] \begin{bmatrix} I_k & 0_{k \times 1} \\ 0_{1 \times k} & H_{33}^{-1} \end{bmatrix} \begin{bmatrix} H_{21} \\ H_{31} \end{bmatrix} \right) \\ &= H_{33}^{1-k} \det (H_{11} H_{33} - H_{12} H_{21} H_{33} - H_{13} H_{31}). \end{aligned}$$

The  $k \times k$  matrix  $\underline{G}_z \triangleq H_{11} H_{33} - H_{12} H_{21} H_{33} - H_{13} H_{31}$  is explicitly given by

$$\begin{aligned} \underline{G}_z &= h(\tau, z) I_k + D_{\tau, z} \Gamma_z - h(\tau, z) \frac{1}{z} F(\tau) 1_k c^\top \\ &= G_z - h(\tau, z) \frac{1}{z} F(\tau) 1_k c^\top \end{aligned} \quad (24)$$

in the notations of Theorem 2. Note now that  $\underline{G}_z$  has right eigenvector  $1_k$  and left eigenvector  $c^\top$  both associated with the eigenvalue  $h(\tau, z)(1 - z^{-1}F(\tau))$ . Thus, provided such a  $z$  is away from  $\mathcal{S}_p$ ,  $\det H_z = 0$  when  $z = F(\tau)$ , that is, when  $-\frac{2f'(\tau)}{f(\tau)}z + \frac{2f'(\tau)}{f(\tau)}F(\tau) = 0$ . Therefore, we recover here precisely the (possibly isolated) zero eigenvalue of  $L'$  (as we should).

Now, note that, for all  $z$ 's distant from  $F(\tau)$ ,  $\underline{G}_z 1_k = h(\tau, z)(1 - z^{-1}F(\tau)) 1_k$  is away from zero, so that  $1_k$  is always a right eigenvector associated with a non-zero eigenvalue. Similarly, since  $c^\top D_{\tau, z} = 0$ ,  $c^\top \underline{G}_z = h(\tau, z)(1 - z^{-1}F(\tau)) c^\top$  and  $c^\top$  is always a left eigenvector with the same eigenvalue. Thus, since all other left-eigenvectors must be orthogonal to  $1_k$  and all other right-eigenvectors orthogonal to  $c^\top$ , the zero eigenvalues of  $\underline{G}_z$  must be the same as those of  $(I_k - 1_k c^\top) \underline{G}_z$  which is precisely  $G_z$ , except for the eigenvalue having eigenvectors  $c^\top$  and  $1_k$ . But  $G_z 1_k = h(\tau, z) 1_k$  and  $c^\top G_z = h(\tau, z) c^\top$ , which is away from

zero. To conclude, the sought-for isolated eigenvalues of  $L'$  (distinct from zero) correspond to those  $z$ 's away from  $\mathcal{S}_p$ , distant from  $F(\tau)$  and such that  $h(\tau, z)$  is away from zero, which are such that  $G_z$  has zero eigenvalues.

7.5.2.  $h(\tau, z) \rightarrow 0$ . In this scenario, as  $H_{33}^{1-k}$  diverges without bound, the study performed in the previous paragraph will no longer be valid in the final arguments of the proof (see next paragraph). As such, we are left with studying  $\det(H_z)$  from (23) directly. Although not easy to fully investigate, a few results already come. For instance, in the case where  $t = 0$ , note that if  $\mathcal{H}_p$  is not empty and thus contains at least a  $\rho_+$ ,  $\det(H_{\rho_+}) = 0$  with multiplicity one unless the same  $\rho_+$  coincidentally induces the upper-left  $2k \times 2k$  submatrix of  $H_{\rho_+}$  to be singular.

7.5.3. *Completion of the proof.* Using the fact that all our estimators above are analytical functions of  $z$  away from  $\mathcal{S}_p \cup \mathcal{G}_p$  and are uniform along  $z$  belonging to a bounded set away from  $\mathcal{S}_p \cup \mathcal{G}_p$  (taking for instance a union bound on finitely many points of the set and using the fact that  $\|Q_z - Q_{\tilde{z}}\| \leq |z - \tilde{z}|(\text{dist}(z, \mathcal{S}_p \cup \mathcal{G}_p) \text{dist}(\tilde{z}, \mathcal{S}_p \cup \mathcal{G}_p))^{-1}$  to control differences), by the argument principle, we find that, for any contour  $\gamma \subset \mathbb{C} \setminus \mathcal{S}_p$  enclosing some eigenvalues of  $PW^\top WP + UBU^\top$  kept at positive distance from the boundaries,

$$\frac{1}{2\pi i} \oint_\gamma \frac{\partial_z \det(PW^\top WP + UBU^\top - zI_n)}{\det(PW^\top WP + UBU^\top - zI_n)} - \frac{1}{2\pi i} \oint_\gamma \frac{\partial_z \det(H_z)}{\det(H_z)} \rightarrow 0$$

almost surely. As both sides are integers and correspond to the number of zeros of the respective denominators (which have no pole away from  $\mathcal{S}_p \cup \mathcal{G}_p$ ), we find that the multiplicity of an eigenvalue  $\lambda$  of  $PW^\top WP + UBU^\top$  is the same as that of its deterministic limit  $\rho$  leading to a root of  $H_z$ . Such  $\rho$ , if satisfying  $h(\tau, \rho) \not\rightarrow 0$ , must then have the same multiplicity as a root of  $G_\rho$ . If instead  $h(\tau, \rho) \rightarrow 0$ , then the multiplicity of  $\rho$  is the multiplicity of zero as an eigenvalue of  $H_\rho$  (which in general will be one).

Assuming now  $f'(\tau) \rightarrow 0$ , Theorem 1 gives

$$\hat{L}' = U_0 B_0 U_0^\top$$

where

$$U_0 = \begin{bmatrix} \frac{j_1}{\sqrt{p}} & \cdots & \frac{j_k}{\sqrt{p}} j_k & \psi \end{bmatrix}$$

$$B_0 = \begin{bmatrix} \frac{f''(\tau)}{f(\tau)} T + \frac{2f''(\tau)}{f(\tau)} t t^\top - c_0 \frac{f(0-f(\tau))}{f(\tau)} & \frac{f''(\tau)}{f(\tau)} t \\ \frac{f''(\tau)}{f(\tau)} t^\top & \frac{f''(\tau)}{f(\tau)} \end{bmatrix}.$$

Up to a shift by the constant  $\frac{f(0)-f(\tau)}{f(\tau)}$ , we are then to solve

$$0 = \det \left( L' - \frac{f(0) - f(\tau)}{f(\tau)} I_n - z I_n \right)$$

which, again by Sylvester's identity, is equivalent to solving, for  $z$  away from zero,

$$0 = \det(B_0 U_0^\top U_0 - z I_{k+1}).$$

If  $h^0(\tau, z)$  is at macroscopic distance from zero, this is asymptotically the same as finding  $z$  for which  $H_z^0$  has a zero eigenvalue, where

$$H_z^0 = \frac{f''(\tau)}{f(\tau)h^0(\tau, z)c_0} tt^\top \mathcal{D}(c) + \frac{2f''(\tau)}{f(\tau)c_0} T \mathcal{D}(c) - zI_k. \quad (25)$$

**7.6. Proof of Theorem 4.** Let  $\mathcal{I} \subset \mathbb{R}$  be a segment away from  $\mathcal{S}_p \cup \mathcal{G}_p$  and such that the eigenvalues  $\rho$  and  $\rho_+$  identified in Theorem 2 and Remark 2 are uniformly away from the boundaries of  $\mathcal{I}$  (either inside or outside). Let then  $\gamma_{\mathcal{I}}$  be a positively oriented contour of  $\mathbb{C}$  circling around and passing through the boundaries of  $\mathcal{I}$ . Then, from the results of Theorems 1 and 2, along with Cauchy's integration theorem, letting  $\hat{\Pi}_{\mathcal{I}}$  be the projector on the subspace associated with the eigenvalues of  $-\frac{f(\tau)}{2f'(\tau)}L + F(\tau)I_n$  in  $\mathcal{I}$  (this subspace being the same as that of  $L$  for the scaled eigenvalues),

$$\frac{1}{p} J^\top \hat{\Pi}_{\mathcal{I}} J = -\frac{1}{2\pi i} \oint_{\gamma_{\mathcal{I}}} \frac{1}{p} J^\top (PW^\top WP + UBU^\top - zI_n)^{-1} J dz + o(1).$$

By Woodbury's identity, this further reads

$$\begin{aligned} \frac{1}{p} J^\top \hat{\Pi}_{\mathcal{I}} J &= -\frac{1}{2\pi i} \oint_{\gamma_{\mathcal{I}}} \frac{1}{p} J^\top Q_z J dz \\ &+ \frac{1}{2\pi i} \oint_{\gamma_{\mathcal{I}}} \frac{1}{p} J^\top Q_z U (I_{2k+1} + BU^\top Q_z U)^{-1} BU^\top Q_z J dz + o(1). \end{aligned} \quad (26)$$

As  $\mathcal{I}$  is away from  $\mathcal{S}_p \cup \mathcal{G}_p$ , which asymptotically contains all the spectrum of  $PW^\top WP$ , the left right-hand side term is asymptotically zero, almost surely. We are then left with studying the rightmost term. This term comprises  $\frac{1}{\sqrt{p}} J^\top Q_z U$  which is a submatrix of  $U^\top Q_z U$  evaluated in (22) in the previous section, and  $(I_{2k+1} + BU^\top Q_z U)^{-1} B$  which we know also from the previous section to be  $H_z^{-1} B + o(1)$  (with  $H_z$  defined in (23)). We shall next evaluate  $H_z^{-1} B$ .

As for  $\Im[z] > 0$  (resp.,  $\Im[z] < 0$ ),  $\Im[g_a(z)] > 0$  (resp.,  $\Im[g_a(z)] < 0$ ) and that  $\gamma_{\mathcal{I}}$  intersects  $\mathbb{R}$  away from the zeroes of  $h(\tau, z)$ ,  $h(\tau, z)$  is away from zero (for all large  $n$ ) on  $\gamma_{\mathcal{I}}$ . Thus, remembering that  $H_{33} = h(\tau, z)$ , we can freely use a block inversion formula for  $H_z$  (pivoting around  $H_{11}$ ) to obtain

$$H_z^{-1} = \begin{bmatrix} H_{33} \underline{G}_z^{-1} & -H_{33} \underline{G}_z^{-1} H_{12} & -\underline{G}_z^{-1} H_{13} \\ -H_{33} H_{21} \underline{G}_z^{-1} & I_k + H_{33} H_{21} \underline{G}_z^{-1} H_{12} & H_{21} \underline{G}_z^{-1} H_{13} \\ -H_{31} \underline{G}_z^{-1} & H_{31} \underline{G}_z^{-1} H_{12} & H_{33}^{-1} + H_{33}^{-1} H_{31} \underline{G}_z^{-1} H_{13} \end{bmatrix}.$$

with  $\underline{G}_z$  defined in (24) as  $\underline{G}_z = H_{11}H_{33} - H_{12}H_{21}H_{33} - H_{13}H_{31}$ . Making these terms explicit and post multiplying by  $B$  then gives, in block definition

$$\begin{aligned}
 [H_z^{-1}B]_{11} &= \underline{G}_z^{-1} \underline{D}_{\tau,z} \\
 [H_z^{-1}B]_{12} &= h(\tau, z) \underline{G}_z^{-1} (I_k - 1_k c^\top) \\
 [H_z^{-1}B]_{13} &= \left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right) \underline{G}_z^{-1} t \\
 [H_z^{-1}B]_{21} &= -\Gamma_z \underline{G}_z^{-1} \underline{D}_{\tau,z} + I_k - c 1_k^\top \\
 [H_z^{-1}B]_{22} &= -h(\tau, z) \Gamma_z \underline{G}_z^{-1} (I_k - 1_k c^\top) \\
 [H_z^{-1}B]_{23} &= -\left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right) \Gamma_z \underline{G}_z^{-1} \\
 [H_z^{-1}B]_{31} &= \left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right) (h(\tau, z))^{-1} t^\top [I_k - \Gamma_z \underline{G}_z^{-1} \underline{D}_{\tau,z}] \\
 [H_z^{-1}B]_{32} &= -\left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right) t^\top \Gamma_z \underline{G}_z^{-1} (I_k - 1_k c^\top) \\
 [H_z^{-1}B]_{33} &= \left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right) (h(\tau, z))^{-1} \left[ 1 - \left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right) t^\top \Gamma_z \underline{G}_z^{-1} t \right] \quad (27)
 \end{aligned}$$

where, for  $D_{\tau,z}$  given in the statement of Theorem 4, we defined  $\underline{D}_{\tau,z} = D_{\tau,z} + h(\tau, z) c_0 F(\tau) 1_k 1_k^\top$ .

With  $B$  (in the statement of Theorem 2), (22), and (27) at hand (at this point, since  $\frac{1}{p} J^\top Q_z U$  vanishes outside the first block, we only need the blocks 11, 12, and 13 of  $H_z^{-1}B$ ), we then explicitly evaluate  $\frac{1}{p} J^\top Q_z U H_z^{-1} B U^\top Q J$  as

$$\frac{1}{p} J^\top Q_z U H_z^{-1} B U^\top Q J = \left( \Gamma_z - \frac{1}{z c_0} c c^\top \right) (I_k - h(\tau, z) \underline{G}_z^{-1}) + o(1).$$

As this is the integrand of the term of interest in (26), one must evaluate the associated residue and thus we are interested here in the values of  $z$  lying within  $\gamma_{\mathcal{I}}$  such that  $(h(\tau, z))^{-1} \underline{G}_z$  is singular. Let us first consider those  $z$ 's for which  $h(\tau, z)$  remains away from zero. As seen in the proof of Theorem 2, the poles of interest here are either  $z = F(\tau)$ , which is directly associated with the eigenvalue  $n$  of  $L$  and thus not of interest here, or the  $z$ 's such that  $\underline{G}_z$  has a zero eigenvalue with left-eigenvector orthogonal to  $1_k$  (and right-eigenvector orthogonal to  $c^\top$ ). Such left- and right-eigenvectors associated with the zero eigenvalues of  $\underline{G}_z$  are also those of  $G_z$  associated with the same eigenvalues. We therefore conclude that

$$\begin{aligned}
 \frac{1}{p} J^\top \hat{\Pi}_{\mathcal{I}} J &= - \sum_{\rho \in \mathcal{I}} \text{Res} \left( h(\tau, z) \left( \Gamma_z - \frac{1}{z c_0} c c^\top \right) \underline{G}_z^{-1} \right) + o(1) \\
 &= - \sum_{\rho \in \mathcal{I}} \text{Res} (h(\tau, z) \Gamma_z G_z^{-1}) + o(1)
 \end{aligned}$$

almost surely, where in the second equality we used the fact that the residue of  $G_z^{-1}$  must have right-eigenvectors orthogonal to  $c^\top$ , i.e., writing  $G_z = V_r \Lambda V_l^\top + h(\tau, z) 1_k c^\top$  in eigenvalue decomposition,  $c^\top V_r = 0$  and thus, since  $h(\tau, z)$  is not close to zero,  $c^\top \text{Res}(G_z^{-1}) = 0$ . Our result is then concluded by noticing that, for  $\rho$  such that  $G_\rho$  has a zero eigenvalue with multiplicity  $m_\rho$ , and with the previous notation (recalling also that  $[V_r \ 1_k] = [V_l \ c]^{-1}$ )

$$\begin{aligned} \lim_{z \rightarrow \rho} (z - \rho) G_z^{-1} &= \lim_{z \rightarrow \rho} (z - \rho) V_r \Lambda^{-1} V_l^\top \\ &= \lim_{z \rightarrow \rho} (z - \rho) V_{r,z} \Lambda_0^{-1} V_{l,z}^\top \\ &= \lim_{z \rightarrow \rho} (z - \rho) \sum_{i=1}^{m_\rho} \frac{(V_{r,z})_i (V_{l,z})_i^\top}{(V_{l,z})_i^\top G_z (V_{r,z})_i} \end{aligned}$$

with  $\Lambda_0 \in \mathbb{C}^{m_\rho \times m_\rho}$  the diagonal of  $m_\rho$  eigenvalues of  $G_z$  tending to zero as  $z \rightarrow \rho$ , and  $V_{r,z} = [(V_{r,z})_1, \dots, (V_{r,z})_{m_\rho}]$ ,  $V_{l,z} = [(V_{l,z})_1, \dots, (V_{l,z})_{m_\rho}]$  their associated right- and left-eigenvectors. From L'Hospital's rule, this further simplifies as

$$\lim_{z \rightarrow \rho} (z - \rho) G_z^{-1} = \lim_{z \rightarrow \rho} \sum_{i=1}^{m_\rho} \frac{(V_{r,z})_i (V_{l,z})_i^\top}{[\partial_z (V_{l,z})_i^\top G_z (V_{r,z})_i]_{z=\rho}}.$$

The derivative in the denominator above is  $(\partial_z (V_{l,z})_i^\top) G_z (V_{r,z})_i + (V_{l,z})_i^\top G_z (\partial_z (V_{r,z})_i) + (V_{l,z})_i^\top (\partial_z G_z) (V_{r,z})_i$ . But  $G_\rho (V_{r,\rho})_i = 0$  and  $(V_{l,\rho})_i^\top G_\rho = 0$ , so that finally

$$\lim_{z \rightarrow \rho} (z - \rho) G_z^{-1} = \sum_{i=1}^{m_\rho} \frac{(V_{r,\rho})_i (V_{l,\rho})_i^\top}{(V_{l,\rho})_i^\top G'_\rho (V_{r,\rho})_i}$$

with  $G'_\rho = [\partial_z G_z]_{z=\rho}$ . This concludes the proof in the case where no  $z$  satisfying  $h(\tau, z) = 0$  is found close to  $\mathcal{I}$ .

Let us now consider the residue associated to the hypothetical (real)  $\rho$ 's for which  $h(\tau, \rho) = 0$ . Then, if  $\|t\|$  is away from zero, as  $z \rightarrow \rho$ ,  $\underline{G}_z$  tends to a rank-one matrix proportional to  $tt^\top \Gamma_\rho$ . Thus, with the same reasoning as previously,

$$\lim_{z \rightarrow \rho} (z - \rho) h(\tau, z) \underline{G}_z^{-1} = \lim_{z \rightarrow \rho} \sum_{i=1}^{k-1} \frac{(h(\tau, z) + (z - \rho)h'(\tau, z))(P_{r,z})_i (P_{l,z})_i^\top}{(P_{l,z})_i^\top \underline{G}'_z (P_{r,z})_i}$$

with  $(P_{r,z})_i$  and  $(P_{l,z})_i$  the eigenvectors of  $\underline{G}_z$  associated to its vanishing eigenvalues. In the limit, the denominator is well defined, unless  $\rho$  coincides (or gets asymptotically close) to another of the  $\rho$ 's identified in Theorem 2. Discarding this situation, and realizing that the denominator must tend to zero, we thus find that the residue associated to  $\rho$  is zero. If instead  $\|t\| \rightarrow 0$ ,  $h(\tau, z) \underline{G}_z^{-1}$  is well defined by extension by continuity in  $z = \rho$ , and there is again no residue.

This completes the proof of Theorem 4 since, as  $\rho$  in the statement of the theorem is isolated from the other eigenvalues, one can set  $\mathcal{I}$  to be a segment containing solely  $\rho$ , all other eigenvalues being kept away.

The proof of Theorem 5 follows straightforwardly from the previous proof and is thus not commented any further.

**7.7. Proof of Theorem 6.** As it is seen in Theorem 4 that those  $\rho$ 's for which  $h(\tau, \rho) = 0$  are associated to eigenvectors which do not carry information about the classes, we shall not study them any longer, and thus restrict ourselves now to  $h(\tau, \rho)$  away from zero. Besides, as  $D^{\frac{1}{2}}1_n$  is perfectly known, we shall not consider this eigenvector.

Letting  $\mathcal{I} \subset \mathbb{R}$  be away from  $\mathcal{S}'_p$  and  $f'(\tau)$  away from zero, we have

$$\mathrm{tr} \mathcal{D}_a \hat{\Pi}_{\mathcal{I}} = -\frac{1}{2\pi i} \oint_{\gamma_{\mathcal{I}}} \mathrm{tr} \mathcal{D}_a (PW^{\top}WP + UBU^{\top} - zI_n)^{-1} dz.$$

As in the previous section, we find

$$\begin{aligned} \mathrm{tr} \mathcal{D}_a \hat{\Pi}_{\mathcal{I}} &= \frac{1}{2\pi i} \oint_{\gamma_{\mathcal{I}}} \mathrm{tr} \mathcal{D}_a Q_z U H_z^{-1} B U^{\top} Q_z dz + o(1) \\ &= \frac{1}{2\pi i} \oint_{\gamma_{\mathcal{I}}} \mathrm{tr} H_z^{-1} B U^{\top} Q_z \mathcal{D}_a Q_z U dz + o(1) \end{aligned}$$

where  $H_z^{-1} B U^{\top}$  has already been identified in (27) while a deterministic approximation for  $U^{\top} Q_z \mathcal{D}_a Q_z U$  can be obtained as in the early stages of the proof of Theorem 2, using now Lemma 2. Precisely, we find that

$$U^{\top} Q_z \mathcal{D}_a Q_z U = \begin{bmatrix} E_{a;zz}^J & 0_{k \times k} & 0_{k \times 1} \\ 0_{k \times k} & E_{a;zz}^M & 0_{k \times 1} \\ 0_{1 \times k} & 0_{1 \times k} & E_{a;zz}^{\psi} \end{bmatrix} + o(1) \quad (28)$$

where  $E_{a;z_1 z_2}^J$ ,  $E_{a;z_1 z_2}^M$ , and  $E_{a;z_1 z_2}^{\psi}$  are defined in (8). As a consequence,

$$\begin{aligned} &\mathrm{tr} \mathcal{D}_a Q_z U H_z^{-1} B U^{\top} Q_z \\ &= \mathrm{tr} (E_{a;z_1 z_2}^J \underline{G}_z^{-1} \underline{D}_{\tau,z}) - h(\tau, z) \mathrm{tr} (E_{a;zz}^M \Gamma_z \underline{G}_z^{-1} (I_k - 1_k c^{\top})) \\ &+ \frac{E_{a;zz}^{\psi}}{h(\tau, z)} \left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right) \left[ 1 - \left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right) t^{\top} \Gamma_z \underline{G}_z^{-1} t \right] + o(1). \end{aligned}$$

As we are not interested here in the eigenvector  $D^{\frac{1}{2}}1_n$ , we assume that  $\gamma_{\mathcal{I}}$  does not enclose zero.

Thus, taking the residue associated to  $\mathrm{tr} \mathcal{D}_a Q_z U H_z^{-1} B U^{\top} Q_z$ , we find, as previously, that we may freely turn  $\underline{G}_z^{-1}$  above into  $G_z^{-1}$  and that only the left-eigenvectors of  $G_z^{-1}$  orthogonal to  $1_k$  will be accounted for. As such, the expression above simplifies when taking the residues as

$$\begin{aligned} &\mathrm{Res} (\mathrm{tr} \mathcal{D}_a Q_z U H_z^{-1} B U^{\top} Q_z) \\ &= \mathrm{Res} (\mathrm{tr} E_{a;z_1 z_2}^J G_z^{-1} D_{\tau,z}) - \mathrm{Res} (h(\tau, z) \mathrm{tr} E_{a;zz}^M \Gamma_z G_z^{-1}) \\ &- \mathrm{Res} \left( \frac{E_{a;zz}^{\psi}}{h(\tau, z)} \left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right)^2 t^{\top} \Gamma_z G_z^{-1} t \right) + o(1). \end{aligned}$$

From this, the theorem unfolds. The case where  $f'(\tau) \rightarrow 0$  is treated similarly.

**7.8. Proof of Theorem 8.** The first part of the proof follows the arguments for the proofs of Theorem 4 and Theorem 6. Precisely, we have here, for  $\mathcal{I}$  a contour neither enclosing  $\rho$  such that  $h(\tau, \rho) = 0$  nor enclosing the eigenvalue  $n$  of  $L$ ,

$$\psi^\top \mathcal{D}_a \hat{\Pi}_{\mathcal{I}} \frac{J}{\sqrt{p}} = \frac{1}{2\pi i} \oint_{\gamma_{\mathcal{I}}} \psi^\top \mathcal{D}_a Q_z U H_z^{-1} B U^\top Q_z \frac{J}{\sqrt{p}} dz + o(1)$$

where we only need to evaluate the new term  $\psi^\top \mathcal{D}_a Q_z U$  since both  $H_z^{-1} B$  and  $U^\top Q_z \frac{J}{\sqrt{p}}$  (as a submatrix of  $U^\top Q_z U$ ) are known. For the former, as in previous derivations, we can show that  $\psi$  behaves as if it were independent of  $W$  when it comes to evaluating such bilinear forms. Since  $\psi$  has independent zero mean entries and  $\mathcal{D}_a \psi$  is supported on the indices of class  $\mathcal{C}_a$  and there has i.i.d. entries of variance  $2 \operatorname{tr} C_a^2 / p^2$ , applying Lemma 1 along with a quadratic-form-close-to-the-trace argument, we obtain

$$\psi^\top \mathcal{D}_a Q_z U = \begin{bmatrix} 0_{k \times k} & 0_{k \times k} & c_a g_a(z) \frac{2}{p} \operatorname{tr} C_a^2 \end{bmatrix} + o(1).$$

Together with the previous results on  $H_z^{-1} B$  and  $U^\top Q_z \frac{J}{\sqrt{p}}$ , this finally gives

$$\begin{aligned} & \psi^\top \mathcal{D}_a \hat{\Pi}_{\mathcal{I}} \frac{J}{\sqrt{p}} \\ &= \operatorname{Res} \left( -c_a g_a(z) \frac{2}{p} \operatorname{tr} C_a^2 \left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right) t^\top \Gamma_z \underline{G}_z^{-1} \underline{D}(\tau, z) \left( \Gamma_z - \frac{cc^\top}{zc_0} \right) \right) + o(1). \end{aligned}$$

Since  $\mathcal{I}$  does not contain the eigenvalue  $n$  of  $L$ , as usual, we may replace  $\underline{G}_z^{-1}$  by  $G_z^{-1}$  and  $\underline{D}(\tau, z)$  by  $D(\tau, z)$ , to obtain

$$\begin{aligned} & \psi^\top \mathcal{D}_a \hat{\Pi}_{\mathcal{I}} \frac{J}{\sqrt{p}} \\ &= \operatorname{Res} \left( -c_a g_a(z) \frac{2}{p} \operatorname{tr} C_a^2 \left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right) t^\top \Gamma_z G_z^{-1} D(\tau, z) \Gamma_z \right) + o(1) \end{aligned}$$

which, by the relation  $G_z^{-1} D(\tau, z) \Gamma_z = -h(\tau, z) G_z^{-1} + I_k$  (the latter leading to no residue), gives the result.

The second part of the proof follows once more very closely the proof of Theorem 6, now for two disjoint segments  $\mathcal{I}_1$  and  $\mathcal{I}_2$  of  $\mathbb{R} \setminus \mathcal{S}'_p$ ,

$$\frac{1}{p} J^\top \hat{\Pi}_{\mathcal{I}} \mathcal{D}_a \hat{\Pi}_{\mathcal{I}} J = \left( \frac{1}{2\pi i} \right)^2 \oint_{\gamma_{\mathcal{I}_1}} \oint_{\gamma_{\mathcal{I}_2}} \frac{1}{p} J^\top Q_{z_1} \mathcal{D}_a Q_{z_2} J + o(1)$$



where  $\mathcal{Q}_z = (PW^\top WP + UBU^\top - zI_n)^{-1}$ . This is further written as

$$\begin{aligned} & \frac{1}{p} J^\top \hat{\Pi}_{\mathcal{I}} \mathcal{D}_a \hat{\Pi}_{\mathcal{I}} J \\ &= \left( \frac{1}{2\pi i} \right)^2 \oint_{\gamma_{\mathcal{I}_1}} \oint_{\gamma_{\mathcal{I}_2}} \text{tr} U^\top \mathcal{Q}_{z_1} \mathcal{D}_a \mathcal{Q}_{z_2} U H_{z_2}^{-1} B U^\top \mathcal{Q}_{z_2} \mathcal{D}_a \mathcal{Q}_{z_1} U H_{z_1}^{-1} B dz_1 dz_2 + o(1). \end{aligned}$$

The integrand is essentially composed of the term  $U^\top \mathcal{Q}_{z_1} \mathcal{D}_a \mathcal{Q}_{z_2} U$  which is obtained as in (28) but for  $(z, z)$  changed into  $(z_1, z_2)$ , and of the terms  $H_{z_1}^{-1} B$  and  $H_{z_2}^{-1} B$ , given in (27). After an immediate calculus, similar to the previous section, and additionally using the fact that  $\Gamma_z G_z^{-1} = (G_z^{-1})^\top \Gamma_z$  or  $G_z^{-1} D_{\tau, z} = D_{\tau, z} (G_z^{-1})^\top$ , as well as  $G_z^{-1} D(\tau, z) \Gamma_z + h(\tau, z) G_z^{-1} = I_k$ , we find

$$\begin{aligned} & \text{Res} \left( \text{tr} U^\top \mathcal{Q}_{z_1} \mathcal{D}_a \mathcal{Q}_{z_2} U H_{z_2}^{-1} B U^\top \mathcal{Q}_{z_2} \mathcal{D}_a \mathcal{Q}_{z_1} U H_{z_1}^{-1} B \right) \\ &= \text{Res} \left( h(\tau, z_1) h(\tau, z_2) (G_{z_1}^{-1})^\top E_{a; z_1 z_2}^J G_{z_2}^{-1} + h(\tau, z_1) h(\tau, z_2) \Gamma_{z_1} G_{z_1}^{-1} E_{a; z_1 z_2}^M (G_{z_2}^{-1})^\top \right) \\ &+ \text{Res} \left( \left( \frac{5f'(\tau)}{8f(\tau)} - \frac{f''(\tau)}{2f'(\tau)} \right)^2 E_{a; z_1 z_2}^\psi \Gamma_{z_1} G_{z_1}^{-1} t t^\top (G_{z_2}^{-1})^\top \Gamma_{z_2} \right) + o(1). \end{aligned}$$

Taking the residues  $\text{Res}(G_{z_1}^{-1})$  and  $\text{Res}(G_{z_2}^{-1})$  over  $\gamma_{\mathcal{I}_1}$  and  $\gamma_{\mathcal{I}_2}$  successively, we retrieve the sought-for result. The same derivations can be performed for the case where  $f'(\tau) \rightarrow 0$ .

#### APPENDIX: CONCENTRATION LEMMAS

The following lemma is extracted from [E11, Lemma 2.12].

**Lemma 3.** *Let us fix  $\alpha, C > 0$  and consider  $y_1, \dots, y_p$  some independent complex centered random variables with variance 1 such that for each  $i$ , for all  $x \geq 0$ ,*

$$\mathbb{P}(|y_i| \geq x^\alpha) \leq C e^{-x}.$$

*Then for any deterministic  $a_1, \dots, a_p \in \mathbb{C}$ , we have*

$$\mathbb{P} \left( |a_1 y_1 + \dots + a_p y_p| \geq (\log p)^{\alpha + \frac{3}{2}} \sqrt{|a_1|^2 + \dots + |a_p|^2} \right) \leq C' p^{-\log \log p}$$

*where  $C'$  is a constant depending only on  $\alpha$  and  $C$ .*

The following lemma can be found in [RV13] (see also [HS71]). It states roughly that  $X^\top A X - \text{tr} A$  has order at most  $\max\{\sqrt{\text{tr} A A^\top}, \|A\|\} = \sqrt{\text{tr} A A^\top}$ .

**Lemma 4** (Hanson-Wright inequality). *Let  $X$  be a standard Gaussian vector in  $\mathbb{R}^d$  and let  $A$  be an  $n \times n$  real matrix. Then  $\mathbb{E}[X^\top A X] = \text{tr} A$ ,  $\text{Var}(X^\top A X) = 2 \text{tr} A A^\top$ , and there is a constant  $c$  independent of  $d$  and of  $A$  such that for any  $t > 0$ ,*

$$\mathbb{P}(|X^\top A X - \text{tr} A| > t) \leq 2 \exp \left( -c \min \left\{ \frac{t^2}{\text{tr} A A^\top}, \frac{t}{\|A\|} \right\} \right).$$

## REFERENCES

- [BS98] Bai, Z. D., Silverstein, J. W., 1998. No eigenvalues outside the support of the limiting spectral distribution of large dimensional sample covariance matrices. *The Annals of Probability* 26 (1), 316–345.
- [BBP05] Baik, J., Ben Arous, G., Péché, S., 2005. Phase transition of the largest eigenvalue for non-null complex sample covariance matrices. *The Annals of Probability* 33 (5), 1643–1697.
- [BC15] Benaych-Georges, F., Couillet, R., 2015. Spectral analysis of the Gram matrix of mixture models. arXiv.
- [BN12] Benaych-Georges, F., Nadakuditi, R. R., 2012. The singular values and vectors of low rank perturbations of large rectangular random matrices. *Journal of Multivariate Analysis* 111, 120–135.
- [B06] Bishop, C. M., 2006. *Pattern recognition and machine learning*. Springer.
- [CCHM14] Chapon, F., Couillet, R., Hachem, W., Mestre, X., 2014. The outliers among the singular values of large rectangular random matrices with additive fixed rank deformation. *Markov Processes and Related Fields* 20, 183–228.
- [CDS11] Couillet, R., Debbah, M., Silverstein, J. W., Jun. 2011. A deterministic equivalent for the analysis of correlated MIMO multiple access channels. *IEEE Transactions on Information Theory* 57 (6), 3493–3514.
- [CH13] Couillet, R., Hachem, W., 2013. Fluctuations of spiked random matrix models and failure diagnosis in sensor networks. *IEEE Transactions on Information Theory* 59 (1), 509–525.
- [CH14] Couillet, R., Hachem, W., 2014. Analysis of the limit spectral measure of large random matrices of the separable covariance type. *Random Matrix Theory and Applications* 3 (4), 1–23.
- [EK10] El Karoui, N., 2010. The spectrum of kernel random matrices. *The Annals of Statistics* 38 (1), 1–50.
- [E11] Erdős, L., 2011. Universality of wigner random matrices: a survey of recent results. *Russian Mathematical Surveys* 66 (3), 507.
- [HLMNV13] Hachem, W., Loubaton, P., Mestre, X., Najim, J., Vallet, P., 2013. A subspace estimator for fixed rank perturbations of large random matrices. *Journal of Multivariate Analysis* 114, 427–447.
- [HS71] Hanson, D. L., Wright, F. T., 1971. A bound on tail probabilities for quadratic forms in independent random variables. *The Annals of Mathematical Statistics*, 1079–1083.
- [HJ85] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.
- [J01] I.M. Johnstone. On the distribution of the largest Principal Component. *Ann. Statist.* 29 (2001) 295–327.
- [LCB98] LeCun, Y., Cortes, C., Burges, C., 1998. The mnist database of handwritten digits.
- [MP67] Marčenko, V. A., Pastur, L. A., 1967. Distribution of eigenvalues for some sets of random matrices. *Math USSR-Sbornik* 1 (4), 457–483.
- [NJW01] Ng, A. Y., Jordan, M., Weiss, Y., 2001. On spectral clustering: Analysis and an algorithm. *Proceedings of Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press 14, 849–856.
- [PS11] Pastur, L., Šerbina, M., 2011. Eigenvalue distribution of large random matrices. *American Mathematical Society*.
- [P07] D. Paul. Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. *Statist. Sinica*, 17(4):1617–1642, 2007.
- [RV13] Rudelson, M., Vershynin, R., 2013. Hanson-wright inequality and sub-gaussian concentration. *Electron. Commun. Probab* 18 (0).
- [SB95] Silverstein, J. W., Bai, Z. D., 1995. On the empirical distribution of eigenvalues of a class of large dimensional random matrices. *Journal of Multivariate Analysis* 54 (2), 175–192.
- [SC95] Silverstein, J. W., Choi, S., 1995. Analysis of the limiting spectral distribution of large dimensional random matrices. *Journal of Multivariate Analysis* 54 (2), 295–309.
- [V07] Von Luxburg, U., 2007. A tutorial on spectral clustering. *Statistics and computing* 17 (4), 395–416.

- [VBB08] Von Luxburg, U., Belkin, M., Bousquet, O., 2008. Consistency of spectral clustering. *The Annals of Statistics*, 555–586.
- [WCRS01] Wagner, S., Couillet, R., Debbah, M., Slock, D. T. M., 2012. Large system analysis of linear precoding in MISO broadcast channels with limited feedback. *IEEE Transactions on Information Theory* 58 (7), 4509–4537.

(R.C.) CENTRALESUPÉLEC – LSS – UNIVERSITÉ PARISUD, GIF SUR YVETTE, FRANCE  
*E-mail address:* `romain.couillet@supelec.fr`

(F.B.G.) MAP 5, UMR CNRS 8145 – UNIVERSITÉ PARIS DESCARTES, PARIS, FRANCE.  
*E-mail address:* `florent.benaych-georges@parisdescartes.fr`