Spectral Estimators for Multi-Index Models: Precise Asymptotics and Optimal Weak Recovery

Filip Kovačević FILIP.KOVACEVIC@IST.AC.AT
Institute of Science and Technology Austria

Yihan Zhang YIHAN.ZHANG@BRISTOL.AC.UK
University of Bristol

Marco Mondelli MARCO.MONDELLI@IST.AC.AT
Institute of Science and Technology Austria

Editors: Nika Haghtalab and Ankur Moitra

Abstract

Multi-index models provide a popular framework to investigate the learnability of functions with low-dimensional structure and, also due to their connections with neural networks, they have been object of recent intensive study. In this paper, we focus on recovering the subspace spanned by the signals via spectral estimators – a family of methods routinely used in practice, often as a warm-start for iterative algorithms. Our main technical contribution is a precise asymptotic characterization of the performance of spectral methods, when sample size and input dimension grow proportionally and the dimension p of the space to recover is fixed. Specifically, we locate the top-p eigenvalues of the spectral matrix and establish the overlaps between the corresponding eigenvectors (which give the spectral estimators) and a basis of the signal subspace. Our analysis unveils a phase transition phenomenon in which, as the sample complexity grows, eigenvalues escape from the bulk of the spectrum and, when that happens, eigenvectors recover directions of the desired subspace. The precise characterization we put forward enables the optimization of the data preprocessing, thus allowing to identify the spectral estimator that requires the minimal sample size for weak recovery.

1. Introduction

Modern machine learning practices operate on high-dimensional datasets that are believed to possess low-dimensional structures, and multi-index models are a popular statistical framework for studying such scenarios (Li, 1991, 1992; Li and Duan, 1989). Specifically, for a labeled dataset $\mathcal{D} = \{(a_i, y_i)\}_{i=1}^n$, where $a_i \in \mathbb{R}^d$ and $y_i \in \mathbb{R}$ denote features and responses respectively, a multi-index model postulates that each data pair follows a generalized regression and the responses are function of a *low-dimensional* projection of the high-dimensional features. In formulas:

$$y_i = q\big(\langle a_i, w_1^* \rangle, \cdots, \langle a_i, w_p^* \rangle, \varepsilon_i\big), \tag{1.1}$$

where the link function q is a known nonlinearity operating on a p-dimensional linear transformation of a_i, given by $W^* = \begin{bmatrix} w_1^*, & \cdots, & w_p^* \end{bmatrix} \in \mathbb{R}^{d\times p}$, and on additional randomness from ε_i.

We study the problem of estimating the signals W^* given $\mathcal{D}$ under a Gaussian design where the a_i's are i.i.d. standard Gaussian. We focus on the proportional asymptotic regime where $n, d \to \infty$ with $n/d \to \delta \in]0, \infty[$, while keeping the dimension of the low-dimensional projection p fixed. Of particular interest in this work is the *weak recovery* of the subspace spanned by $w_1^*, \cdots, w_p^*$ which seeks an estimator $\widehat{W} = \begin{bmatrix} \widehat{w}_1, & \cdots, & \widehat{w}_p \end{bmatrix} \in \mathbb{R}^{d\times p}$ s.t. at least one of $\widehat{w}_i$'s is non-trivially correlated with some linear combination of the signals. Formally, there exists $v \in \mathbb{S}^{p-1}$ and some $\widehat{w}_i$ in $\widehat{W}$ such that the normalized correlation (or *overlap*) $\frac{|\langle \widehat{w}_i, W^* v\rangle|}{\|\widehat{w}_i\|_2 \|W^* v\|_2}$ is asymptotically non-vanishing.

To solve the aforementioned problem, this paper focuses on spectral estimators that construct the matrix $D = \frac{1}{n}\sum_{i=1}^n \mathcal{T}(y_i) a_i a_i^\top$, from the dataset $\mathcal{D}$ and then output its top-p eigenvectors. Here, $\mathcal{T}\colon \mathbb{R} \to \mathbb{R}$ refers to a user-defined preprocessing function, with popular choices including binary quantization and truncation (Wang et al., 2018; Chen and Candès, 2017). Spectral estimators are easy to design, efficient to compute, and effective in practice (Chen et al., 2021). However, such class of methods remains understudied for multi-index models, with existing results falling short of producing exact asymptotics (Chen and Meka, 2020) and being restricted to special cases, such as single-index (for which $p = 1$, corresponding to generalized linear models (McCullagh, 1984)) (Lu and Li, 2020; Mondelli and Montanari, 2019), polynomial link functions (Chen and Meka, 2020), and mixed regression (Zhang et al., 2022), see also Section 2. In contrast, this paper tackles the problem for fully general multi-index models, precisely identifying under what conditions spectral estimators are effective.

A compelling motivation for the precise analysis of spectral estimators comes from the choice of the preprocessing function $\mathcal{T}$. In the single-index case, its optimization has led to significant performance gains for Gaussian (Mondelli and Montanari, 2019; Luo et al., 2019), correlated (Zhang et al., 2024) and rotationally invariant (Maillard et al., 2022) designs. Remarkably, the related optimal spectral estimators have also been shown to achieve the computational limits of weak learnability for a class of link functions. In the proportional regime of interest in this work ($n = \Theta(d)$),[1] such limits have been connected to the performance of Approximate Message Passing (AMP) (Mondelli and Montanari, 2019; Maillard et al., 2020) and, more precisely, to the stability of its trivial fixed point. Here, AMP refers to a family of iterative algorithms that is provably optimal among first-order methods (Celentano et al., 2020; Montanari and Wu, 2024) and, in fact, there is significant evidence that AMP is optimal even among all polynomial-time algorithms, i.e., it achieves Bayes-optimal performance unless a statistical-to-computational gap is present (Barbier et al., 2019, 2023; Montanari and Venkataramanan, 2021; Rangan et al., 2019; Venkataramanan et al., 2022). Most recently, the multi-index case has been considered by Troiani et al. (2024) and a threshold capturing its computational limits has been computed by studying whether AMP is able to improve on a non-trivial initialization, see Remark 5 for a connection with these results.

Main contributions. This paper tackles the two main problems mentioned above, i.e., *(i)* the lack of a precise analysis of spectral estimators for multi-index models, and *(ii)* the design of an optimal preprocessing function $\mathcal{T}$. Specifically, our results are summarized below.

1. For any preprocessing function $\mathcal{T}$ satisfying mild assumptions, we precisely locate the top-p eigenvalues of the spectral matrix D and characterize the overlaps between the top-p eigenvectors and a basis of the subspace spanned by the signals. Our results describe a phase transition phenomenon, akin to the classical BBP transition in the spiked covariance model (Baik et al., 2005), with spectral outliers of D corresponding to eigenvectors employed for signal recovery.

2. Using the above characterization, we identify the optimal preprocessing function for weak recovery. Our optimality result guarantees that no other choice of preprocessing function results

1. Related, albeit different, perspectives are to consider a sample complexity polynomial in d (Damian et al., 2023, 2024) or information-theoretic limits (Barbier et al., 2019; Aubin et al., 2019).

in spectral estimators with a lower threshold, and therefore it also implies the suboptimality of existing heuristics to choose $\mathcal{T}$.

2. Related work

Multi-index models. Several approaches have been proposed to perform statistical inference in multi-index models, e.g., structural adaption via maximum minimization (Dalalyan et al., 2008), projection pursuit regression (Yuan, 2011) and techniques from compressed sensing (Fornasier et al., 2012). Andoni et al. (2014); Chen and Meka (2020) consider polynomial link functions, with the latter work proposing a spectral warm start that requires a sample size $n \gtrsim d(\log(d))^{\deg(q)}$, where $\deg(q)$ is the degree of the link. Due to the connection of multi-index models with two-layer neural networks, the area has witnessed a recent renewed interest with a focus on the performance of gradient-based methods. In particular, sample complexity bounds for gradient descent and statistical query lower bounds are provided by Damian et al. (2022) when the link function is polynomial and by Oko et al. (2024) when the multi-index model is the sum of single-index models. Abbe et al. (2022, 2023) introduce the concept of leap complexity and show that a class of staircase functions is learned via one-pass stochastic gradient descent (SGD) with $n = \Theta(d)$ samples. The leap exponent also appears in (Bietti et al., 2023) as the time required by gradient flow to escape a saddle point. Collins-Woodfin et al. (2024) prove a deterministic equivalent of the SGD dynamics. Ren and Lee (2024) provide an algorithm that recovers orthogonal multi-index models with a sample complexity matching the information exponent (Ben Arous et al., 2021). We note that none of these methods pin-points exactly the sample complexity required to recover a multi-index model, which constitutes the focus of our work and is achieved via the class of spectral methods reviewed below.

Spectral estimators. Spectral estimators have been applied to a variety of problems, e.g., community detection (Abbe, 2018), angular synchronization (Singer, 2011) and principal component analysis (PCA) (Montanari and Venkataramanan, 2021). In the setting of Gaussian design and proportional scaling between n and d, their asymptotic performance for single-index models is characterized by Lu and Li (2020). Optimal weak recovery thresholds and optimal overlaps are identified by Mondelli and Montanari (2019) and Luo et al. (2019), respectively. The above results are extended by Dudeja et al. (2020) to subsampled Haar designs and by Zhang et al. (2024) to correlated Gaussian designs. Rotationally invariant designs are considered by Maillard et al. (2022), which conjecture the form of the optimal spectral estimator using a linearization of AMP and the analysis of the Bethe Hessian. Such conjecture is partly addressed by Zhang et al. (2024), when the covariance of the a_i's is rotationally invariant. We note that, in the single-index case, optimal spectral methods match computational thresholds obtained from the stability of AMP (Mondelli and Montanari, 2019; Zhang et al., 2024) and, in special cases, information-theoretic thresholds as well (Mondelli and Montanari, 2019). An optimally-designed spectral estimator is able to meet the information-theoretic limits of weak recovery also for a class of heteroscedastic PCA problems (Zhang and Mondelli, 2024). Most closely related to our setting is work by Zhang et al. (2022): the authors consider mixtures of single-index models with independent signals and provide precise asymptotics for spectral estimators by using a mix of tools from random matrix theory and the theory of AMP. In contrast, our approach is purely random matrix theoretic, and it allows us to handle a general class of multi-index models with arbitrary correlation among the signals.

We finally note that a parallel paper by Defilippis et al. (2025) also considers a multi-index model with Gaussian design and introduces spectral estimators based on the linearization of AMP.

Defilippis et al. (2025) then conjecture that such spectral estimators are optimal in the sense that they achieve the computational threshold identified by Troiani et al. (2024), providing both numerical and rigorous evidence in favor of the conjecture. Our work focuses on a family of spectral estimators popular in the related literature (cf. Eq. (41) in Defilippis et al. (2025) and (3.2) in our work), and it resolves the conjecture of Defilippis et al. (2025) for a wide class of link functions q, including all permutation-invariant ones, see Remark 5 for details.

3. Preliminaries

Notation. Given a positive integer n, we use the shorthand $[n] := \{1, \dots, n\}$. We denote by 0_n the vector of length n with all zeros. For a symmetric matrix M, we denote its Moore-Penrose pseudo-inverse as $M^\dagger$, the set of all its eigenvalues as Λ^M, its i-th largest eigenvalue as λ_i^M and the corresponding i-th eigenvector of unit norm as v_i^M. We use $(e_i^{(d)})_{i\in[d]}$ to denote the canonical basis of $\mathbb{R}^d$ and suppress the superscript whenever there is no confusion. Unless otherwise specified, all limits of sequences of random quantities are computed in an almost sure sense as $n, d \to \infty$.

Multi-index models and weak recovery. We consider the problem of estimating p signals using n responses $(y_i)_{i\in[n]}$ generated i.i.d. from (1.1), where $\varepsilon_i \in \mathbb{R}$ accounts for noise and $q : \mathbb{R}^{p+1} \to \mathbb{R}$ is the link function. We denote by $A := \begin{bmatrix} a_1, & \cdots, & a_n \end{bmatrix}^\top \in \mathbb{R}^{n\times d}$ the design matrix, by $W^* := \begin{bmatrix} w_1^*, & \cdots, & w_p^* \end{bmatrix} \in \mathbb{R}^{d\times p}$ the signal matrix, and by $y := \begin{bmatrix} y_1, & \cdots, & y_n \end{bmatrix}^\top \in \mathbb{R}^n$ the response vector. Throughout the paper, we impose the following assumptions.

(A1) $A \in \mathbb{R}^{n\times d}$ contains i.i.d. standard Gaussian elements, i.e., $A_{i,j} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1)$.

(A2) $(w_i^*)_{i\in[p]}$ are linearly independent, have unit norm and, if random, they are independent of A.

(A3) $p \geq 1$ is fixed and $n, d \to \infty$ s.t. $n/d \to \delta \in]0, \infty[$.

(A4) $\varepsilon_1, \cdots, \varepsilon_n$ are independent of A, W^*, and they are i.i.d. according to a distribution P_ε on $\mathbb{R}$ with finite first two moments.

The linear independence requirement in Assumption (A2) is mild. For the purpose of subspace recovery (formally defined below), the presence of linearly dependent signals does not change the recoverability of a given estimator.

Definition 1 (Weak recovery) *Consider the model Equation* (1.1). *Let* $\widehat{W} \equiv \widehat{W}(A, y) = \begin{bmatrix} \widehat{w}_1, & \cdots, & \widehat{w}_p \end{bmatrix} \in \mathbb{R}^{d\times p}$ *be an estimator such that* $\|\widehat{w}_i\|_2 = 1$ *for all* $i \in [p]$. *We say that* $\widehat{W}$ *weakly recovers the subspace* $\operatorname{span}\{w_1^*, \cdots, w_p^*\}$ *if*

$$\max_{v\in\mathbb{S}^{p-1}} \left\{ \liminf_{d\to\infty} \frac{\left\|\widehat{W}^\top W^* v\right\|_2}{\|W^* v\|_2} \right\} > 0, \tag{3.1}$$

where the almost sure limit is taken with respect to the proportional scaling in Assumption (A3).

In words, $\widehat{W}$ weakly recovers $\mathcal{W} := \operatorname{span}\{w_1^*, \cdots, w_p^*\}$ if at least one of the $\widehat{w}_i$'s attains non-vanishing correlation with *some* linear combination of the signals $w_1^*, \cdots, w_p^*$. Other notions of weak recovery for multi-index models have been considered in the literature. For instance, Troiani et al. (2024) also discuss the weak recovery of the *whole* subspace $\mathcal{W}$, replacing the $\max_{v \in \mathbb{S}^{p-1}}$ in Equation (3.1) with $\min_{v \in \mathbb{S}^{p-1}}$. This requirement is clearly stronger than Definition 1 since *every* vector in $\mathcal{W}$ needs to be weakly recovered by some $\widehat{w}_i$ in $\widehat{W}$. In this work, we focus exclusively on Definition 1.

Spectral estimators. For a preprocessing function $\mathcal{T} : \mathbb{R} \to \mathbb{R}$, let $z_i := \mathcal{T}(y_i)$ for $i \in [n]$ and $Z := \operatorname{diag}\begin{bmatrix} z_1, & \cdots, & z_n \end{bmatrix} \in \mathbb{R}^{n \times n}$. Furthermore, we define the matrix $D \equiv D_n \in \mathbb{R}^{d \times d}$ as

$$D_n = \frac{1}{n} A^\top Z A = \frac{1}{n} \sum_{i=1}^{n} z_i a_i a_i^\top. \tag{3.2}$$

The spectral estimator then outputs the eigenvectors corresponding to the p largest eigenvalues of D, i.e., $\begin{bmatrix} v_1^D, & \cdots, & v_p^D \end{bmatrix}$.

We now make a simplifying assumption on W^* without loss of generality. Note that, for any W^* subject to Assumption (A2), by orthogonal invariance of the random design matrix A, the law of $\left\{ \frac{|\langle v_i^D, W^* v \rangle|}{\|W^* v\|_2} : i \in [p], v \in \mathbb{S}^{d-1} \right\}$ remains unchanged under the rotation mapping $W^* \mapsto W^* O$ for any orthogonal matrix $O \in \mathbb{R}^{p \times p}$. Therefore, for convenience, we take the unique rotation O that, for all i, maps w_i^* to $\sum_{j=1}^{i} c_{i,j} e_j$, with $\sum_{j=1}^{i} c_{i,j}^2 = 1$, and study $\{ |\langle v_i^D, e_j \rangle| : i, j \in [p] \}$, i.e., the overlap of the eigenvectors with the basis $\{e_1, \cdots, e_p\}$ spanned by the rotated signals. We note that, after this rotation, W^* can be written as

$$W^* = \begin{bmatrix} \widetilde{W}^* \\ 0_{(d-p) \times p} \end{bmatrix} \in \mathbb{R}^{d \times p}, \qquad \widetilde{W}^* = \begin{bmatrix} c_{1,1} & c_{2,1} & \cdots & c_{p,1} \\ & c_{2,2} & \cdots & c_{p,2} \\ & & \ddots & \vdots \\ & & & c_{p,p} \end{bmatrix} \in \mathbb{R}^{p \times p}. \tag{3.3}$$

We further define the random variables

$$(s, \varepsilon) \sim \mathcal{N}(0_p, I_p) \otimes P_\varepsilon, \quad y = q((\widetilde{W}^*)^\top s, \varepsilon), \quad z = \mathcal{T}(y), \tag{3.4}$$

and make the following assumptions on the preprocessing function $\mathcal{T}$.

(A5) $\mathcal{T}$ is bounded and $\mathbb{P}(z = 0) < 1$.

4. Main results

Consider the model Equation (1.1) under Assumptions (A1) to (A4). Our first result accurately locates, in the high-dimensional limit, the p largest eigenvalues of the matrix D in Equation (3.2) for any preprocessing function $\mathcal{T}$ subject to Assumption (A5), thereby unveiling a spectral phase transition phenomenon. To present it, a sequence of definitions is needed.

Let $\mathscr{T}$ be the collection of preprocessing functions $\mathcal{T}$ subject to Assumption (A5). For any $\mathcal{T} \in \mathscr{T}$, let $z = \mathcal{T}(y)$ as in Equation (3.4) and let $\tau = \inf\{c : \mathbb{P}(z \le c) = 1\}$ be the right edge of the support of z (by Assumption (A5), $\tau < \infty$). Define $\psi_\delta \colon]\tau, \infty[\to \mathbb{R}$ and its point of minimum as

$$\psi_\delta(\lambda) \coloneqq \lambda\left(\frac{1}{\delta} + \mathbb{E}\left[\frac{z}{\lambda - z}\right]\right), \qquad \bar{\lambda}_\delta \coloneqq \operatorname*{argmin}_{\lambda > \tau} \psi_\delta(\lambda).$$

Note that ψ_δ is convex and, thus, its minimum is unique. Finally, define $\zeta_\delta \colon \mathbb{R} \to \mathbb{R}$ and $R^\infty \colon]\tau, \infty[\to \mathbb{R}^{p\times p}$ as

$$\zeta_\delta(\lambda) \coloneqq \psi_\delta(\max\{\bar{\lambda}_\delta, \lambda\}), \qquad R^\infty(\alpha) \coloneqq \mathbb{E}\left[\frac{\alpha s s^\top z}{\alpha - z}\right], \tag{4.1}$$

where s, z are jointly distributed according to Equation (3.4).

Theorem 2 *Let $\mathcal{T} \colon \mathbb{R} \to \mathbb{R}$ be a preprocessing function subject to Assumption (A5), and let $D \in \mathbb{R}^{d\times d}$ be defined in Equation (3.2). Let $\alpha_1 \geq \ldots \geq \alpha_j > \tau$ (for some $j \in [p]$) be all the solutions to*

$$\det(\zeta_\delta(\alpha) I - R^\infty(\alpha)) = 0. \tag{4.2}$$

Then, for the top j eigenvalues of D, it holds that

$$\lambda_1^D, \ldots, \lambda_j^D \xrightarrow{\text{a.s.}} \zeta_\delta(\alpha_1), \ldots, \zeta_\delta(\alpha_j), \tag{4.3}$$

and for the remaining $p - j$ eigenvalues, it holds that

$$\lambda_{j+1}^D, \ldots, \lambda_p^D \xrightarrow{\text{a.s.}} \zeta_\delta(\bar{\lambda}_\delta).$$

In words, Theorem 2 shows a phase transition for the j largest eigenvalues of D as the (normalized) sample complexity δ varies. In fact, by the definition of $\zeta_\delta(\cdot)$ in Equation (4.1), Equation (4.3) implies that, for any $k \in [j]$, if $\alpha_k \leq \bar{\lambda}_\delta$, then λ_i^D asymptotically coincides with $\zeta_\delta(\bar{\lambda}_\delta)$, which corresponds to the right edge of the bulk of the spectrum of D; conversely, if $\alpha_k > \bar{\lambda}_\delta$, then the asymptotic value of λ_i^D is strictly larger than the right edge, thereby forming a spectral outlier. This phenomenon mirrors the classical BBP phase transition for the spiked covariance model (Baik et al., 2005).

We note that (4.2) has at most p solutions in $]\tau, +\infty[$. Furthermore, if $\mathcal{T}$ satisfies

$$\inf_{\|x\|_2 = 1} \lim_{\alpha \to \tau^+} \mathbb{E}\left[\frac{\alpha z \langle s, x\rangle^2}{\alpha - z}\right] = +\infty, \tag{4.4}$$

we are guaranteed that (4.2) has exactly p solutions. Both statements are proved in Proposition 17 deferred to Section B.3. The condition in (4.4) provides a natural generalization of the assumption made for $p = 1$ in previous work, see Equation (82) in (Mondelli and Montanari, 2019).

Our second result characterizes the asymptotic performance, in terms of overlaps, of the eigenvectors corresponding to the spectral outliers of D.

Theorem 3 *In the setting of Theorem 2, let $\alpha_k = \alpha_{k+1} = \cdots = \alpha_{k+m-1}$ be solutions to Equation (4.2) of multiplicity m, i.e., $\alpha_{k-1} > \alpha_k > \alpha_{k+m-2}$ ($k \geq 2, k + m - 2 \leq j$), and let $E_k^\infty \subset \mathbb{R}^p$ be the m-dimensional eigenspace of $R^\infty(\alpha_k)$. If $\lambda_i^D \xrightarrow{\text{a.s.}} \zeta_\delta(\alpha_k) > \zeta_\delta(\bar{\lambda}_\delta)$ ($i \in \{k, \ldots, k+m-1\}$), then*

$$\max_{l\in[p]} \liminf_{d\to\infty} \sum_{i=k}^{k+m-1} \left|\left\langle v_i^D, e_l^{(d)}\right\rangle\right|^2 > 0. \tag{4.5}$$

More precisely, under the additional assumption that either $m = 1$ or the eigenspace E_k^∞ stays invariant in a neighbourhood of α_k, for any $l \in [p]$,

$$\sum_{i=k}^{k+m-1} \left| \left\langle v_i^D, e_l^{(d)} \right\rangle \right|^2 \xrightarrow{\text{a.s.}} \frac{\zeta_\delta'(\alpha_k) \sum_{i=k}^{k+m-1} \left| \left\langle h_i^\infty, e_l^{(p)} \right\rangle \right|^2}{\zeta_\delta'(\alpha_k) + h_k^{\infty\top} \frac{d}{d\alpha} R^\infty(\alpha_k) h_k^\infty}, \tag{4.6}$$

where $\{h_i^\infty : k \le i \le k+m-1\} \subset \mathbb{R}^p$ is an orthonormal basis of E_k^∞ and $\frac{d}{d\alpha} R^\infty(\cdot)$ denotes the entry-wise derivative of the matrix $R^\infty(\cdot)$. Conversely, for all i s.t. $\lambda_i^D \xrightarrow{\text{a.s.}} \zeta_\delta(\bar{\lambda}_\delta)$, then

$$\max_{l \in [p]} \lim_{d \to \infty} \left| \left\langle v_i^D, e_l^{(d)} \right\rangle \right|^2 = 0. \tag{4.7}$$

In words, Theorem 3 shows that, if λ_k^D is an outlier (which happens when $\alpha_k > \bar{\lambda}_\delta$ by Theorem 2), the corresponding eigenvectors $V_k = \begin{bmatrix} v_k^D, & \cdots, & v_{k+m-1}^D \end{bmatrix} \in \mathbb{R}^{d \times m}$ weakly recover $\mathcal{W} = \text{span}\left\{ e_1^{(d)}, \cdots, e_p^{(d)} \right\}$ and Equation (4.6) expresses the asymptotic (squared) overlap $\left\| V_k^\top e_l^{(d)} \right\|_2^2$ in terms of p-dimensional equations. Conversely, if λ_k^D converges to the right edge of the bulk (which happens if either $\alpha_k \le \bar{\lambda}_\delta$ or k exceeds the number of solutions of Equation (4.2)), then the overlap vanishes.

Let us further elaborate on the invariance of the eigenspace E_k^∞ required for Equation (4.6) to hold. Specifically, if E_k^∞ is an eigenspace of $R^\infty(\alpha_k)$, then it is also an eigenspace of $R^\infty(\alpha_k + \Delta)$ for small enough Δ. This condition ensures that the denominator on the RHS of Equation (4.6) is independent of the choice of the vector from the eigenspace. If this was not the case, the norm of the overlap vectors, i.e. $\sum_{j=1}^p \left| \left\langle v_i^D, e_j \right\rangle \right|^2$, may not have an asymptotic limit, making analysis hard. The condition appears to be a proof artifact, and it is not even clear how to construct a case with eigenvalue multiplicity that violates the invariance condition. For example, the assumption is satisfied by models in which the eigenvalue multiplicity arises from the permutation-invariance of the link function q, see Proposition 19 in Appendix D.

Taken together, Theorems 2 and 3 generalize earlier work by Mondelli and Montanari (2019) on the single-index model and by Zhang et al. (2022) on mixtures of single-index models with independent signals. The independence of signals is crucial to the approach in (Zhang et al., 2022) which decomposes D_n in (3.2) as the free sum of matrices each corresponding to one of the signals. To circumvent this difficulty, we pursue a p-dimensional analog of the strategy in (Mondelli and Montanari, 2019), with additional adjustments. Namely, all results for eigenvalues and overlaps in (Mondelli and Montanari, 2019) (Eq. (95), (96), (97), and (99) therein) come from Eq. (94), for which there is no direct p-dimensional analogue. Our work identifies appropriate alternatives in the form of Equation (5.2) and Equation (5.4). We then show that the eigenvalues of D_n solve a fixed point equation (see Equation (5.3)) and that the eigenvectors are related to a $p \times p$ matrix (see Equation (5.4)). Finally, studying the limiting behavior (as $n, d \to \infty$) of the p-dimensional objects gives the claimed asymptotic characterizations. For $p > 1$, eigenvalue multiplicities complicate the analysis of the limits of eigenvalue derivatives and associated eigenspaces, which are needed for characterizing asymptotic overlaps. We handle such complications via tools from perturbation theory. The proof is outlined in Section 5, with several auxiliary results deferred to Appendices A to C.

Our result on overlaps in Theorem 3 concerns the basis vectors. However, unless the signals have vanishing correlation, one needs additional side information to assemble the overlaps with the basis into overlaps with the signals $(w_i^*)_{i\in[p]}$, due to the sign ambiguity of eigenvectors.

Finally, our third main result identifies an optimal preprocessing function allowing the spectral estimator to attain the lowest weak recovery threshold as per Definition 1. For any $\delta > 0$, let

$$\delta_c(\mathcal{T}) := \inf\left\{\delta > 0 : \max_{l\in[p]} \lim_{d\to\infty} \sum_{i=1}^{p} \left|\left\langle v_i^D, e_l^{(d)}\right\rangle\right|^2 > 0\right\}$$

be the recovery threshold for $\mathcal{T}$, that is, the smallest aspect ratio above which it achieves weak recovery. Then, define the optimal weak recovery threshold as

$$\delta_c := \inf_{\mathcal{T}\in\mathscr{T}} \delta_c(\mathcal{T}). \tag{4.8}$$

Finally, for random variables $(s, y) \in \mathbb{R}^p \times \mathbb{R}$ jointly distributed according to Equation (3.4), we use $p(y \mid s)$ to denote the conditional density of y given s.

Theorem 4 *The optimal weak recovery threshold δ_c equals*

$$\delta_c = \left[\max_{u\in\mathbb{S}^{p-1}} \int_{\mathbb{R}} \frac{\left(\mathbb{E}_s\left[p(y \mid s)\cdot(\langle s, u\rangle^2 - 1)\right]\right)^2}{\mathbb{E}_s[p(y \mid s)]} dy\right]^{-1}, \tag{4.9}$$

where expectations are intended over $s \sim \mathcal{N}(0_p, I_p)$. Furthermore, denoting by $u_c \in \mathbb{S}^{p-1}$ a maximizer in the above expression, for any $\delta > \delta_c$, weak recovery is achieved by taking $\mathcal{T}_\delta^*(y)$ as in

$$\mathcal{T}^*(y) := 1 - \frac{\mathbb{E}_s[p(y \mid s)]}{\mathbb{E}_s\left[p(y \mid s)\cdot\langle s, u_c\rangle^2\right]}, \qquad \mathcal{T}_\delta^*(y) := \frac{\sqrt{\delta_c}\cdot\mathcal{T}^*(y)}{\sqrt{\delta} - (\sqrt{\delta} - \sqrt{\delta_c})\cdot\mathcal{T}^*(y)}. \tag{4.10}$$

In words, Theorem 4 shows that the preprocessing in Equation (4.10) leads to weak recovery of the signal subspace with a sample complexity that is minimal *among all spectral methods*. As for Theorems 2 and 3, the result generalizes earlier work (Mondelli and Montanari, 2019; Zhang et al., 2022) to the multi-index case with arbitrarily correlated signals. The proof of Theorem 4 is in Appendix E.

The design of the optimal preprocessing function does not require the knowledge of $\widetilde{W}^*$, but only of the link function q and of the limiting sample covariance matrix of the signals $\Sigma := (\widetilde{W}^*)^\top \widetilde{W}^*$. In fact, $\mathcal{T}^*(y)$ remains unchanged under $\widetilde{W}^* \mapsto O\widetilde{W}^*$ for any orthogonal matrix $O \in \mathbb{R}^{p\times p}$. Indeed, one can verify that under the above mapping, $\mathbb{E}_s[p(y \mid s)]$ remains unchanged and $u_c \mapsto Ou_c$. We can then take $O = (\widetilde{W}^{*\top}\widetilde{W}^*)^{-1/2}\widetilde{W}^{*\top}$ and equivalently define

$$y = q\left((\widetilde{W}^*)^\top O^\top s, \varepsilon\right) = q\left(\left[(\widetilde{W}^*)^\top \widetilde{W}^*\right]^{1/2} s, \varepsilon\right).$$

We note that the link q and the sample covariance Σ are assumed to be known e.g. when y is the output of a neural network with p neurons in the teacher-student model. If q and Σ are unknown, the problem is in general much harder, and a separate set of samples may be used to first estimate q, Σ

(sample splitting, as in Sawaya et al. (2024)). If q is parameterized by θ of fixed dimension, then one can obtain θ and Σ from the moments of the random variable y in Eq. (3.4) in some cases (e.g., $y = s_1 s_2$ with generic Σ; $q = \sum_{i=1}^{p} s_i^2 + w$ with $\Sigma = I$ and w of unknown mean and variance). In such cases, θ and Σ can be consistently estimated using the empirical moments of the response vector with $o(n)$ samples. The same guarantees then continue to hold if one applies the optimal spectral estimator on the remaining $n - o(n)$ samples with the consistent estimate of q and Σ.

Remark 5 *Let $\mathcal{F}(M) := \mathbb{E}_y[E(y)ME(y)]$ with $E(y) := \mathbb{E}_s\big[ss^\top - I_p \mid y\big]$. Then, the threshold identified in Lemma 4.1 of Troiani et al. (2024) is given by $\left(\sup_{\substack{M \succ 0_{p\times p}\\ \|M\|_{\mathrm{F}}=1}} \|\mathcal{F}(M)\|_{\mathrm{F}}\right)^{-1}$. One can readily verify that this expression coincides with Equation (4.9) when the $\sup$ is achieved by a rank-1 matrix. This is the case when the matrices $E(y)$ are simultaneously diagonalizable for all y, a condition satisfied by permutation invariant link functions and all examples in Appendix F of (Troiani et al., 2024). The proof is outlined in Appendix G.*

Note that the simultaneous diagonalizability of $E(y)$ also implies the simultaneous diagonalizability of $R^\infty(\alpha)$. Namely,

$$R^\infty(\alpha) = \mathbb{E}\left[\frac{\alpha s s^\top z}{\alpha - z}\right] = \mathop{\mathbb{E}}_{y}\left[\frac{\alpha \mathcal{T}(y)}{\alpha - \mathcal{T}(y)} \mathop{\mathbb{E}}_{s}\Big[ss^\top \mid y\Big]\right] = \mathop{\mathbb{E}}_{y}\left[\frac{\alpha \mathcal{T}(y)}{\alpha - \mathcal{T}(y)}(E(y) + I_p)\right].$$

Thus, if $(E(y) + I_p)$ has the same eigenvectors for every y, so does $R^\infty(\alpha)$ for every α. This implies that the invariance condition on the eigenspace is satisfied, and Equation (4.6) holds for all eigenvectors with eigenvalues outside the bulk.

Numerical results. In Figures 1 and 2, we consider instances of Equation (1.1) with $d = 1500$ and $p = 2$, and we compute overlaps between v_i^D and w_j^* ($1 \le i, j \le 2$). Each data point is obtained by averaging over 10 i.i.d. trials, and error bars are reported at 1 standard deviation. The corresponding theoretical predictions are plotted using solid curves. Both numerical and theoretical values of overlaps are plotted as a function of the aspect ratio δ and captioned 'num' and 'thy' respectively.

In Figure 1, the link function is $q(s_1, s_2, \varepsilon) = s_1 s_2$ and the preprocessing function $\mathcal{T}$ is the optimal one, see Section F.1 for the derivation. The signals are sampled i.i.d. from independent isotropic Gaussians. Due to the permutation invariance of the model, for any eigenvector v_i^D, the overlap $\left|\left\langle v_i^D, w_j^*\right\rangle\right|$ is asymptotically the same for $j \in \{1, 2\}$. As there is a single outlier in the model, we only plot $\left|\left\langle v_1^D, w_1^*\right\rangle\right|$ and $\left|\left\langle v_1^D, w_2^*\right\rangle\right|$.

In Figure 2, we consider a two-component mixed phase retrieval model: $q(\xi_1, \xi_2, \varepsilon) = |\xi_\varepsilon|$, where the mixing variable ε is $\{1, 2\}$-valued with $\mathbb{P}(\varepsilon = 1) = 0.6$. The signals are sampled from a bivariate correlated Gaussian with correlation $\rho = 0.3$. The performance of spectral estimators is compared among five choices of preprocessing functions: *(i)* the quadratic function $\mathcal{T}(y) = \min\{y^2, 10\}$ used by Yi et al. (2014) for mixed linear regression; *(ii)* the trimming function $\mathcal{T}(y) = y^2 \mathbb{1}\{y^2 \le 7\}$ considered by Chen and

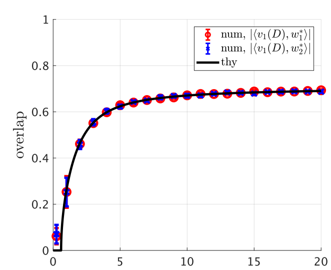

Figure 1: $q(s_1, s_2, \varepsilon) = s_1 s_2$. Overlaps $\left|\left\langle v_1^D, w_1^*\right\rangle\right|, \left|\left\langle v_1^D, w_2^*\right\rangle\right|$ and theoretical predictions from Theorem 3 are plotted as a function of δ.

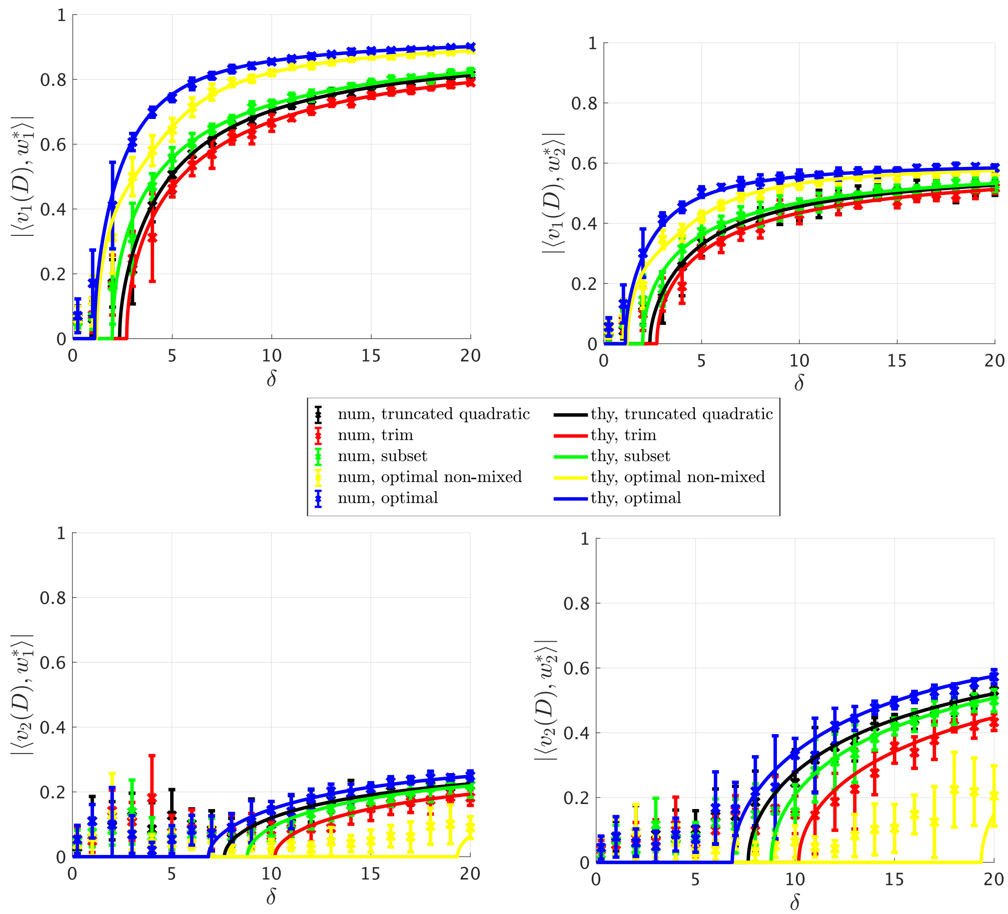

Figure 2: $q(\xi_1, \xi_2, \varepsilon) = |\xi_\varepsilon|$. Overlaps $\left\{\left|\left\langle v_i^D, w_j^* \right\rangle\right| : 1 \le i, j \le 2\right\}$ and theoretical predictions from Theorem 3 are plotted as a function of δ for five preprocessing functions. Our optimal preprocessing function in Equation (F.2) attains both the lowest weak recovery threshold (as ensured by Theorem 4) and the highest overlap.

Candès (2017) which zeros out labels with large magnitude; *(iii)* the subset function $\mathcal{T}(y) = \mathbb{1}\left\{y^2 > 2\right\}$ considered by Wang et al. (2018) which quantizes the labels to binary values; *(iv)* the optimal preprocessing function $\mathcal{T}(y) = \min\left\{1 - 1/y^2, -10\right\}$ for the *non-mixed* phase retrieval model derived by Luo et al. (2019); and *(v)* the optimal preprocessing function that is guaranteed by our theory to maximize the asymptotic overlap $\left|\left\langle v_1^D, w_1^* \right\rangle\right|$, see Section F.2 for details. As we assume boundedness of $\mathcal{T}$ (see Assumption (A5)), we truncate such functions whenever necessary. The truncation levels in *(i)* and *(iv)* are taken to be large enough so as not to significantly affect the performance; the truncation/quantization levels in *(ii)* and *(iii)* are taken from Mondelli and Montanari (2019).

A few remarks on Figures 1 and 2 are in order. First, the numerical and theoretical results exhibit accurate agreement even for $d = 1500$, suggesting a rapid rate of convergence. Second, for mixed phase retrieval, the mixing effect and the correlation between signals have crucial effects on the design of the optimal preprocessing function. Naively applying the optimal preprocessing function for *non-mixed* phase retrieval results in poor performance of v_2^D which achieves positive overlap with either signal only if $\delta > 19$ (see the bottom row of Figure 2). In contrast, the weak recovery threshold of v_2^D resulting from our proposed choice of preprocessing is less than 7. Third, while our choice of preprocessing is designed to minimize the weak recovery threshold of v_1^D, the performance is also competitive in terms of overlaps, outperforming all alternatives.

5. Proof techniques

Equivalent spectral characterization. Let us write $a_i = \begin{bmatrix} s_i & u_i \end{bmatrix}^\top$, with $s_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, I_p)$ and $u_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, I_{d-p})$. Thus, as $w_i^* \in \operatorname{span}\{e_1^{(d)}, \dots, e_p^{(d)}\}$ (see Equation (3.3)), y_i only depends on s_i and, more precisely, $y_i = q\big(\langle s_i, w_1^*\rangle, \cdots, \langle s_i, w_p^*\rangle, \varepsilon_i\big)$, with the abuse of notation that the s_i's are now vectors in $\mathbb{R}^d$ with the last $d-p$ coordinates set to 0. Extending the notation to matrices, we have that $A = \begin{bmatrix} S & U \end{bmatrix}$, where $S := \begin{bmatrix} s_1 & \cdots & s_n \end{bmatrix}^\top \in \mathbb{R}^{n\times p}$ and $U := \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix}^\top \in \mathbb{R}^{n\times(d-p)}$. Thus, we can re-write the spectral matrix D_n in (3.2) as

$$D_n = \frac{1}{n} A^\top Z A = \frac{1}{n}\begin{bmatrix} S^\top \\ U^\top \end{bmatrix} Z \begin{bmatrix} S & U \end{bmatrix} = \frac{1}{n}\begin{bmatrix} S^\top Z S & S^\top Z U \\ U^\top Z S & U^\top Z U \end{bmatrix} = \begin{bmatrix} a & q^\top \\ q & P \end{bmatrix}, \tag{5.1}$$

where $a := \frac{1}{n} S^\top Z S \in \mathbb{R}^{p\times p}$, $q := \frac{1}{n} U^\top Z S \in \mathbb{R}^{(d-p)\times p}$ and $P := \frac{1}{n} U^\top Z U \in \mathbb{R}^{(d-p)\times(d-p)}$.

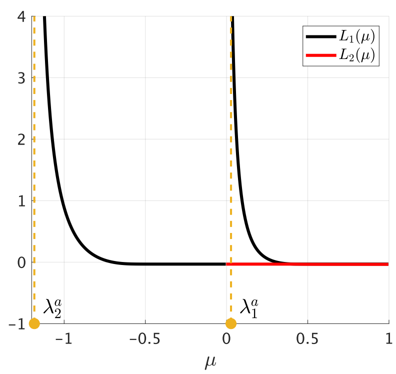

Figure 3: Plot of the functions $\widetilde{L}_1, \widetilde{L}_2$ (as well as L_1, L_2) defined in Equation (5.2) for the model $y_i = \langle a_i, w_1^*\rangle\langle a_i, w_2^*\rangle$ ($1 \le i \le n$), where $w_1^*, w_2^* \in \mathbb{R}^{1500}$ are orthogonal unit vectors and $n = 5d$. The preprocessing function is taken to be the optimal one given in Equation (F.1).

We start by working with a generic matrix D of the form in the RHS of (5.1), such that a, P are symmetric and, for all eigenvectors v_i^a of a, it holds that $q v_i^a \neq 0$ (this is the case for the matrix D_n in Equation (5.1) as showed in Lemma 7 in Appendix A). For simplicity, we assume that all eigenvalues of a are different (if any multiplicity exists, notation and exposition can easily be adjusted accordingly). Let $L_i : \mathbb{R} \setminus \Lambda^a \to \mathbb{R}$ (for $i \in [d-p]$) and $\tilde{L}_i :]\lambda_i^a, +\infty[\setminus \Lambda^a \to \mathbb{R}$ (for $i \in [p]$) be defined as

$$L_i(\mu) := \lambda_i(P - q(a - \mu I_p)^{-1} q^\top), \qquad \tilde{L}_i(\mu) := \begin{cases} L_1(\mu) & \text{if } \lambda_i^a < \mu < \lambda_{i-1}^a, \\ \quad\vdots & \\ L_i(\mu) & \text{if } \lambda_1^a < \mu < +\infty, \end{cases} \tag{5.2}$$

see Figure 3 for a plot. The $\tilde{L}_i$'s are piecewise continuous, since the L_i's are continuous in the respective domains. Furthermore, the domain of $\tilde{L}_i$ can be continuously extended to Λ^a (Lemma 8 in Appendix A), and $\tilde{L}_i(\mu)$ is non-increasing in this domain (Lemma 9 in Appendix A).

We use the functions in Equation (5.2) to characterize the top p eigenvalues $(\lambda_i^D)_{i\in[p]}$ of the matrix D. Specifically, Proposition 12 in Appendix A shows that, for $i \in [p]$, λ_i^D is the unique solution to

$$\tilde{L}_i(\mu) = \mu, \qquad \text{for } \mu \in]\lambda_i^a, \infty[. \tag{5.3}$$

Moving to eigenvectors, we write $v_i^D := \begin{bmatrix} h_i \\ g_i \end{bmatrix}$ ($h_i \in \mathbb{R}^p$, $g_i \in \mathbb{R}^{d-p}$) and characterize h_i in terms of the function $R(\lambda) := a - q^\top (P - \lambda I_{d-p})^{-1} q$ defined for $\lambda \in]\lambda_1^p, +\infty[$. Specifically, Proposition 13 in Appendix A shows that, for all i s.t. $\lambda_i^D > \lambda_1^P$,

$$h_i = \frac{\tilde{h}_i}{\sqrt{1 - \tilde{h}_i^\top \frac{d}{d\lambda} R(\lambda_i^D) \tilde{h}_i}}, \tag{5.4}$$

where $\tilde{h}_i$ is the eigenvector of $R(\lambda_i^D)$ and $\frac{d}{d\lambda} R(\lambda)$ is the entry-wise derivative of $R(\lambda)$. Finally, Lemma 14 in Appendix A expresses the entries of $R(\lambda)$ and $\frac{d}{d\lambda} R(\lambda)$ in terms of

$$\mathcal{L}_i(\mu) := \lambda_1(P + \mu q_i q_i^\top), \qquad \mathcal{L}_{i,j}(\mu) := \lambda_1(P + \mu (q_i + q_j)(q_i + q_j)^\top), \tag{5.5}$$

where q_i is the i-th column of q.

In summary, the eigenvalues of D are given by the fixed points of $\tilde{L}_i$ (see Equation (5.3)); the eigenvectors are computed in Equation (5.4) via the $p \times p$ matrices $R(\lambda)$ and $\frac{d}{d\lambda} R(\lambda)$, which are in turn related to $\mathcal{L}_i, \mathcal{L}_{i,j}$ (see Equation (5.5)). The rationale for this equivalent spectral characterization is that $\tilde{L}_i, \mathcal{L}_i, \mathcal{L}_{i,j}$ are all low-rank perturbations of P (the first is rank-p, the other two are rank-1). Thus, when considering the proportional asymptotic regime ($n, d \to \infty$ with n/d and p fixed), we can leverage the wide random matrix theoretic literature on low-rank perturbations and, specifically, results from (Bai and Yao, 2012). As a final, more technical note, the diagonal entries of $R(\lambda)$ and $\frac{d}{d\lambda} R(\lambda)$ are given by

$$R(\lambda)_{i,i} = a_{i,i} + \frac{1}{\mathcal{L}_i^{-1}(\lambda)}, \qquad \frac{d}{d\lambda} R(\lambda)_{i,i} = -\frac{1}{\left(\mathcal{L}_i^{-1}(\lambda)\right)^2 \mathcal{L}_i'(\mathcal{L}_i^{-1}(\lambda))}, \tag{5.6}$$

which is reminiscent of the expressions for the single-index case, see (Lu and Li, 2020, Lemma 3.1). In contrast, the off-diagonal entries are more cumbersome (see Equation (A.14)), and they are linked to the multi-index nature of the model.

Sketch of the proof of Theorem 2. We start with some manipulations: (4.2) is equivalent to $\prod_{i=1}^p (\zeta_\delta(\alpha) - \lambda_i^\infty(\alpha)) = 0$, where $\lambda_i^\infty(\alpha)$ is the i-th largest eigenvalue of $R^\infty(\alpha)$. Thus, the assumption of the theorem implies that α_i is the unique solution to $\zeta_\delta(\alpha) - \lambda_i^\infty(\alpha) = 0$. Let $\mu_i^* := \lambda_i^\infty(\alpha_i)$ and $\tilde{L}_i^\infty(\mu) := \zeta_\delta\big((\lambda_i^\infty)^{-1}(\mu)\big)$. Then, μ_i^* is the unique solution to $\tilde{L}_i^\infty(\mu) - \mu = 0$.

Next, recall that λ_i^D is the unique solution to (5.3). Proposition 16 in Appendix B establishes the limit of $\tilde{L}_i$, by relying on the classical results by Bai and Yao (2012) adapted to the analysis of single-index models in (Lu and Li, 2020; Mondelli and Montanari, 2019). Formally, this gives that

$$\tilde{L}_i(\mu) - \mu \xrightarrow{\text{a.s.}} \tilde{L}_i^\infty(\mu) - \mu. \tag{5.7}$$

As μ_i^* is the unique solution to $\tilde{L}_i^\infty(\mu) - \mu = 0$, we conclude that

$$\lambda_i^D \xrightarrow{\text{a.s.}} \tilde{L}_i^\infty(\mu_i^*). \tag{5.8}$$

Substituting $\mu_i^* = \lambda_i^\infty(\alpha_i)$ in Equation (5.8) gives the desired result in (4.3) for the top-j eigenvalues.

The passages to show the claim for the remaining $p-j$ eigenvalues are similar. As (4.2) has only j solutions by assumption, $\zeta_\delta(\alpha)-\lambda_i^\infty(\alpha) = \tilde{L}_i^\infty(\lambda_i^\infty(\alpha))-\lambda_i^\infty(\alpha) = 0$ has no solutions for $\alpha > \tau$ and $i > j$. Thus, $\tilde{L}_i^\infty(\mu) - \mu = 0$ has no solutions for $\mu \in]\lambda_i^{a^\infty}, t_i^\infty[$, with $t_i^\infty := \lim_{\alpha\to\tau} \lambda_i^\infty(\alpha)$. This in turn implies that, for large enough n, the solution to $\tilde{L}_i(\mu) - \mu = 0$ must be for $\mu > t_i^\infty$.

At this point, we use again Proposition 16 to show that, for $\mu > t_i^\infty$, $\tilde{L}_i(\mu) \xrightarrow{\text{a.s.}} \zeta_\delta(\bar{\lambda}_\delta)$, which implies that $\lambda_i^D \xrightarrow{\text{a.s.}} \zeta_\delta(\bar{\lambda}_\delta)$ for $i \in \{j+1, \dots, p\}$, thus proving the desired result. The detailed proof is in Section B.2.

Sketch of the proof of Theorem 3. As for the eigenvectors, we start with an asymptotic characterization of $R(\lambda)$ and $\frac{d}{d\lambda}R(\lambda)$. Specifically, Proposition 18 in Section C.1 shows that

$$R(\lambda_k^D) \xrightarrow{\text{a.s.}} R^\infty(\alpha_k), \qquad \frac{d}{d\lambda}R(\lambda_k^D) \xrightarrow{\text{a.s.}} \frac{1}{\zeta_\delta'(\alpha_k)}\frac{d}{d\alpha}R^\infty(\alpha_k), \tag{5.9}$$

where α_k is the k-th largest solution of Equation (4.2) and $\alpha_k > \bar{\lambda}_\delta$. To obtain Equation (5.9), the idea is to analyze the functions in Equation (5.5), as $n, d \to \infty$. We do so by using results from Bai and Yao (2012), which give

$$\mathcal{L}_i(\mu) \xrightarrow{\text{a.s.}} \zeta_\delta \circ {Q_i}^{-1} \circ G(\mu), \qquad \mathcal{L}_{i,j}(\mu) \xrightarrow{\text{a.s.}} \zeta_\delta \circ {Q_{i,j}}^{-1} \circ G(\mu), \tag{5.10}$$

where $G(\mu) = -1/\mu$, $Q_i(\alpha) := \mathbb{E}\left[\frac{s_i^2 z^2}{z-\alpha}\right]$ and $Q_{i,j}(\alpha) := \mathbb{E}\left[\frac{(s_i+s_j)^2 z^2}{z-\alpha}\right]$. As $\alpha_k > \bar{\lambda}_\delta$, Theorem 2 guarantees that $\lambda_k^D \xrightarrow{\text{a.s.}} \zeta_\delta(\alpha_k)$ and therefore

$$\mathcal{L}_i^{-1}(\lambda_k^D) \xrightarrow{\text{a.s.}} G \circ Q_i \circ \zeta_\delta^{-1} \circ \zeta_\delta(\alpha_k) = G \circ Q_i(\alpha_k). \tag{5.11}$$

By the law of large numbers, $a \xrightarrow{\text{a.s.}} a^\infty := \mathbb{E}\left[zss^\top\right]$. Thus, by combining Equations (5.6), (5.10) and (5.11) we conclude

$$R(\lambda_k^D)_{i,i} \xrightarrow{\text{a.s.}} a_{i,i}^\infty - Q_i(\alpha_k) = \mathbb{E}\left[s_i^2 z\right] - \mathbb{E}\left[\frac{s_i^2 z^2}{z-\alpha_k}\right] = \mathbb{E}\left[\frac{\alpha_k s_i^2 z}{\alpha_k - z}\right] = R^\infty(\alpha_k)_{i,i}. \tag{5.12}$$

Moreover, $\mathcal{L}_i(\mu)$ is differentiable (see Lemma 14 in Appendix A), so for its derivative it holds that

$$\mathcal{L}_i'(\mu) \xrightarrow{\text{a.s.}} \zeta_\delta' \circ {Q_i}^{-1} \circ G(\mu) \cdot \left(Q_i^{-1}\right)' \circ G(\mu) \cdot G'(\mu).$$

Plugging this into (5.9) we get

$$\frac{d}{d\lambda}R(\lambda_k^D)_{i,i} \xrightarrow{\text{a.s.}} \frac{\frac{d}{d\alpha}(R^\infty(\alpha_k)_{i,i})}{\zeta_\delta'(\alpha_k)}.$$

By performing similar calculations on the off-diagonal entries of $R(\lambda)$ and $\frac{d}{d\lambda}R(\lambda)$, Equation (5.9) follows (the complete proof is deferred to Section C.1).

We are now ready to prove Equation (4.6) when there is no multiplicity ($m = 1$). Having a multiplicity adds technical complications (and also the assumption on the invariance of the eigenspace E_k^∞), which are handled in Section C.2 where we also prove Equations (4.5) and (4.7).

First, note that the desired overlaps can be expressed as

$$\left|\left\langle v_k^D, e_j^{(d)}\right\rangle\right|^2 = (e_j^{(p)})^\top h_k h_k^\top (e_j^{(p)}). \tag{5.13}$$

Furthermore, recall that *(i)* h_k is related to the unit norm eigenvector $\tilde{h}_k$ of $R(\lambda_k^D)$ and to $\frac{d}{d\lambda}R(\lambda_k^D)$ via Equation (5.4), and that *(ii)* the limits of $R(\lambda_k^D)$ and to $\frac{d}{d\lambda}R(\lambda_k^D)$ are given by Equation (5.9). Then, applying the results from (Kato, 1995, II.1.4), the orthonormal projection to the eigenspace corresponding to the k-th eigenvalue also converges, i.e., $\Pi_{h_k} \xrightarrow{\text{a.s.}} \Pi_{h_k^\infty}$, where $\Pi_{h_k} = \frac{h_k h_k^\top}{\|h_k\|_2^2} = \tilde{h}_k \tilde{h}_k^\top$ and $\Pi_{h_k^\infty} = \frac{h_k^\infty h_k^{\infty\top}}{\|h_k^\infty\|_2^2} = h_k^\infty h_k^{\infty\top}$. As a consequence, we have that

$$\|h_k\|_2 = \frac{1}{\sqrt{1 + \tilde{h}_k^\top \frac{d}{d\lambda}R(\lambda_k^D)\tilde{h}_k}} \xrightarrow{\text{a.s.}} \frac{\sqrt{\zeta_\delta'(\alpha_k)}}{\sqrt{\zeta_\delta'(\alpha_k) + h_k^{\infty\top}\frac{d}{d\lambda}R^\infty(\alpha_k)h_k^\infty}}.$$

Combining these results, we obtain that

$$h_k h_k^\top \xrightarrow{\text{a.s.}} \frac{\zeta_\delta'(\alpha_k)h_k^\infty h_k^{\infty\top}}{\zeta_\delta'(\alpha_k) + h_k^{\infty\top}\frac{d}{d\lambda}R^\infty(\alpha_k)h_k^\infty},$$

which, together with Equation (5.13), proves the claim.

6. Concluding remarks

This paper provides the first asymptotic characterization of spectral estimators for multi-index models: we unveil a phase transition in the top-p eigenvalues of the spectral matrix D in Equation (3.2), giving a low-dimensional (and simple to check) condition for spikes to emerge from the bulk of the spectrum; the eigenvalue phase transition is associated to the recovery of the subspace spanned by the signals via the corresponding eigenvectors and, under some technical conditions, we prove a precise expression for the asymptotic overlap; finally, we optimize the data preprocessing and identify the spectral estimator that weakly recovers the signal subspace with the smallest sample complexity.

Spectral methods are commonly used as a warm start for other procedures, often of iterative nature. Thus, our analysis provides the starting point to combine spectral estimators either with a simple linear estimator (Mondelli et al., 2022; Zhang et al., 2022) or with AMP (Mondelli and Venkataramanan, 2022), with the objective of achieving – at least in absence of statistical-to-computational gaps – the Bayes-optimal limits of inference as characterized by Aubin et al. (2019).

Acknowledgements

This work was done when Y. Z. was at the Institute of Science and Technology Austria. Y. Z. and M. M. are funded by the European Union (ERC, INF^2, project number 101161364). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. The authors would like to acknowledge (in alphabetical order) discussions with Yatin Dandi, Leonardo Defilippis and Bruno Loureiro concerning their parallel work (Defilippis et al., 2025).

References

Emmanuel Abbe. Community detection and stochastic block models: Recent developments. *Journal of Machine Learning Research*, 18(177):1–86, 2018.

Emmanuel Abbe, Enric Boix Adsera, and Theodor Misiakiewicz. The merged-staircase property: a necessary and nearly sufficient condition for sgd learning of sparse functions on two-layer neural networks. In *Proceedings of Thirty Fifth Conference on Learning Theory*, volume 178 of *Proceedings of Machine Learning Research*, pages 4782–4887. PMLR, 02–05 Jul 2022.

Emmanuel Abbe, Enric Boix Adserà, and Theodor Misiakiewicz. Sgd learning on neural networks: leap complexity and saddle-to-saddle dynamics. In *Proceedings of Thirty Sixth Conference on Learning Theory*, volume 195 of *Proceedings of Machine Learning Research*, pages 2552–2623. PMLR, 12–15 Jul 2023.

Alexandr Andoni, Rina Panigrahy, Gregory Valiant, and Li Zhang. Learning sparse polynomial functions. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 500–510. ACM, New York, 2014.

Benjamin Aubin, Antoine Maillard, Jean Barbier, Florent Krzakala, Nicolas Macris, and Lenka Zdeborová. The committee machine: computational to statistical gaps in learning a two-layers neural network. *J. Stat. Mech. Theory Exp.*, (12):124023, 51, 2019.

Zhidong Bai and Jianfeng Yao. On sample eigenvalues in a generalized spiked population model. *J. Multivariate Anal.*, 106:167–177, 2012.

Jinho Baik, Gérard Ben Arous, and Sandrine Péché. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *Ann. Probab.*, 33(5):1643–1697, 2005.

Jean Barbier, Florent Krzakala, Nicolas Macris, Léo Miolane, and Lenka Zdeborová. Optimal errors and phase transitions in high-dimensional generalized linear models. *Proc. Natl. Acad. Sci. USA*, 116(12):5451–5460, 2019.

Jean Barbier, Francesco Camilli, Marco Mondelli, and Manuel Sáenz. Fundamental limits in structured principal component analysis and how to reach them. *Proc. Natl. Acad. Sci. USA*, 120(30): Paper No. e2302028120, 7, 2023.

Gerard Ben Arous, Reza Gheissari, and Aukosh Jagannath. Online stochastic gradient descent on non-convex losses from high-dimensional inference. *Journal of Machine Learning Research*, 22 (106):1–51, 2021.

Alberto Bietti, Joan Bruna, and Loucas Pillaud-Vivien. On learning gaussian multi-index models with gradient flow. *CoRR*, abs/2310.19793, 2023.

Michael Celentano, Andrea Montanari, and Yuchen Wu. The estimation error of general first order methods. In *Proceedings of Thirty Third Conference on Learning Theory*, volume 125 of *Proceedings of Machine Learning Research*, pages 1078–1141. PMLR, 09–12 Jul 2020.

Sitan Chen and Raghu Meka. Learning polynomials in few relevant dimensions. In *Proceedings of Thirty Third Conference on Learning Theory*, volume 125 of *Proceedings of Machine Learning Research*, pages 1161–1227. PMLR, 09–12 Jul 2020.

Yuxin Chen and Emmanuel J. Candès. Solving random quadratic systems of equations is nearly as easy as solving linear systems. *Comm. Pure Appl. Math.*, 70(5):822–883, 2017.

Yuxin Chen, Yuejie Chi, Jianqing Fan, and Cong Ma. Spectral methods for data science: A statistical perspective. *Foundations and Trends® in Machine Learning*, 14(5):566–806, 2021.

Elizabeth Collins-Woodfin, Courtney Paquette, Elliot Paquette, and Inbar Seroussi. Hitting the high-dimensional notes: an ODE for SGD learning dynamics on GLMs and multi-index models. *Inf. Inference*, 13(4):Paper No. iaae028, 107, 2024.

Arnak S. Dalalyan, Anatoly Juditsky, and Vladimir Spokoiny. A new algorithm for estimating the effective dimension-reduction subspace. *Journal of Machine Learning Research*, 9(53):1647–1678, 2008.

Alex Damian, Eshaan Nichani, Rong Ge, and Jason D Lee. Smoothing the landscape boosts the signal for sgd: Optimal sample complexity for learning single index models. In *Advances in Neural Information Processing Systems*, volume 36, pages 752–784. Curran Associates, Inc., 2023.

Alex Damian, Loucas Pillaud-Vivien, Jason Lee, and Joan Bruna. Computational-statistical gaps in gaussian single-index models (extended abstract). In *Proceedings of Thirty Seventh Conference on Learning Theory*, volume 247 of *Proceedings of Machine Learning Research*, pages 1262–1262. PMLR, 30 Jun–03 Jul 2024.

Alexandru Damian, Jason Lee, and Mahdi Soltanolkotabi. Neural networks can learn representations with gradient descent. In *Proceedings of Thirty Fifth Conference on Learning Theory*, volume 178 of *Proceedings of Machine Learning Research*, pages 5413–5452. PMLR, 02–05 Jul 2022.

Leonardo Defilippis, Yatin Dandi, Pierre Mergny, Florent Krzakala, and Bruno Loureiro. Optimal spectral transitions in high-dimensional multi-index models. *arXiv preprint arXiv:2502.02545v1*, 2025.

Klaus Deimling. *Nonlinear functional analysis*. Springer-Verlag, Berlin, 1985. ISBN 3-540-13928-1. doi: 10.1007/978-3-662-00547-7. URL https://doi.org/10.1007/978-3-662-00547-7.

Yihong Du. *Order structure and topological methods in nonlinear partial differential equations. Vol. 1*, volume 2 of *Series in Partial Differential Equations and Applications*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2006. ISBN 981-256-624-4. doi: 10.1142/9789812774446. URL https://doi.org/10.1142/9789812774446. Maximum principles and applications.

Rishabh Dudeja, Milad Bakhshizadeh, Junjie Ma, and Arian Maleki. Analysis of spectral methods for phase retrieval with random orthogonal matrices. *IEEE Trans. Inform. Theory*, 66(8):5182–5203, 2020.

Massimo Fornasier, Karin Schnass, and Jan Vybiral. Learning functions of few arbitrary linear parameters in high dimensions. *Found. Comput. Math.*, 12(2):229–262, 2012.

Roger A. Horn and Charles R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, second edition, 2013.

Tosio Kato. *Perturbation theory for linear operators*. Classics in Mathematics. Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition.

M. G. Kreĭn and M. A. Rutman. Linear operators leaving invariant a cone in a Banach space. *Uspehi Matem. Nauk (N.S.)*, 3(1(23)):3–95, 1948.

Peter Lancaster. On eigenvalues of matrices dependent on a parameter. *Numerische Mathematik*, 6 (1):377–387, 1964.

Ker-Chau Li. Sliced inverse regression for dimension reduction. *J. Amer. Statist. Assoc.*, 86(414): 316–342, 1991. With discussion and a rejoinder by the author.

Ker-Chau Li. On principal Hessian directions for data visualization and dimension reduction: another application of Stein's lemma. *J. Amer. Statist. Assoc.*, 87(420):1025–1039, 1992.

Ker-Chau Li and Naihua Duan. Regression analysis under link violation. *Ann. Statist.*, 17(3): 1009–1052, 1989.

Yue M. Lu and Gen Li. Phase transitions of spectral initialization for high-dimensional non-convex estimation. *Inf. Inference*, 9(3):507–541, 2020.

Wangyu Luo, Wael Alghamdi, and Yue M. Lu. Optimal spectral initialization for signal recovery with applications to phase retrieval. *IEEE Trans. Signal Process.*, 67(9):2347–2456, 2019.

Antoine Maillard, Bruno Loureiro, Florent Krzakala, and Lenka Zdeborová. Phase retrieval in high dimensions: Statistical and computational phase transitions. In *Advances in Neural Information Processing Systems*, volume 33, pages 11071–11082. Curran Associates, Inc., 2020.

Antoine Maillard, Florent Krzakala, Yue M. Lu, and Lenka Zdeborova. Construction of optimal spectral methods in phase retrieval. In *Proceedings of the 2nd Mathematical and Scientific Machine Learning Conference*, volume 145 of *Proceedings of Machine Learning Research*, pages 693–720. PMLR, 16–19 Aug 2022.

Peter McCullagh. Generalized linear models. *European J. Oper. Res.*, 16(3):285–292, 1984.

Marco Mondelli and Andrea Montanari. Fundamental limits of weak recovery with applications to phase retrieval. *Found. Comput. Math.*, 19(3):703–773, 2019.

Marco Mondelli and Ramji Venkataramanan. Approximate message passing with spectral initialization for generalized linear models. *J. Stat. Mech. Theory Exp.*, (11):Paper No. 114003, 41, 2022.

Marco Mondelli, Christos Thrampoulidis, and Ramji Venkataramanan. Optimal combination of linear and spectral estimators for generalized linear models. *Found. Comput. Math.*, 22(5):1513–1566, 2022.

Andrea Montanari and Ramji Venkataramanan. Estimation of low-rank matrices via approximate message passing. *Ann. Statist.*, 49(1):321–345, 2021.

Andrea Montanari and Yuchen Wu. Statistically optimal firstorder algorithms: a proof via orthogonalization. *Inf. Inference*, 13(4):Paper No. iaae027, 24, 2024.

Kazusato Oko, Yujin Song, Taiji Suzuki, and Denny Wu. Learning sum of diverse features: computational hardness and efficient gradient-based training for ridge combinations. In *Proceedings of Thirty Seventh Conference on Learning Theory*, volume 247 of *Proceedings of Machine Learning Research*, pages 4009–4081. PMLR, 30 Jun–03 Jul 2024.

Sundeep Rangan, Philip Schniter, and Alyson K. Fletcher. Vector approximate message passing. *IEEE Trans. Inform. Theory*, 65(10):6664–6684, 2019.

Yunwei Ren and Jason D. Lee. Learning orthogonal multi-index models: A fine-grained information exponent analysis. *CoRR*, abs/2410.09678, 2024.

Kazuma Sawaya, Yoshimasa Uematsu, and Masaaki Imaizumi. High-dimensional single-index models: Link estimation and marginal inference. *arXiv preprint arXiv:2404.17812*, 2024.

Jack W. Silverstein and Sang-Il Choi. Analysis of the limiting spectral distribution of large-dimensional random matrices. *J. Multivariate Anal.*, 54(2):295–309, 1995.

A. Singer. Angular synchronization by eigenvectors and semidefinite programming. *Appl. Comput. Harmon. Anal.*, 30(1):20–36, 2011.

Emanuele Troiani, Yatin Dandi, Leonardo Defilippis, Lenka Zdeborová, Bruno Loureiro, and Florent Krzakala. Fundamental limits of weak learnability in high-dimensional multi-index models. *CoRR*, abs/2405.15480, 2024.

Ramji Venkataramanan, Kevin Kögler, and Marco Mondelli. Estimation in rotationally invariant generalized linear models via approximate message passing. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pages 22120–22144. PMLR, 17–23 Jul 2022.

Gang Wang, Georgios B. Giannakis, and Yonina C. Eldar. Solving systems of random quadratic equations via truncated amplitude flow. *IEEE Trans. Inform. Theory*, 64(2):773–794, 2018.

Xinyang Yi, Constantine Caramanis, and Sujay Sanghavi. Alternating minimization for mixed linear regression. In *Proceedings of the 31st International Conference on Machine Learning*, volume 32 of *Proceedings of Machine Learning Research*, pages 613–621, Bejing, China, 22–24 Jun 2014. PMLR.

Ming Yuan. On the identifiability of additive index models. *Statist. Sinica*, 21(4):1901–1911, 2011.

Yihan Zhang and Marco Mondelli. Matrix denoising with doubly heteroscedastic noise: Fundamental limits and optimal spectral methods. *CoRR*, abs/2405.13912, 2024.

Yihan Zhang, Marco Mondelli, and Ramji Venkataramanan. Precise asymptotics for spectral methods in mixed generalized linear models. *CoRR*, abs/2211.11368, 2022.

Yihan Zhang, Hong Chang Ji, Ramji Venkataramanan, and Marco Mondelli. Spectral estimators for structured generalized linear models via approximate message passing (extended abstract). In *Proceedings of Thirty Seventh Conference on Learning Theory*, volume 247 of *Proceedings of Machine Learning Research*, pages 5224–5230. PMLR, 30 Jun–03 Jul 2024.

Appendix A. Formal statements and proofs for the equivalent spectral characterization

We start with two auxiliary results.

Lemma 6 *Under Assumption (A5), almost surely for all sufficiently large n it holds that*

$$\mathrm{rk}(ZS) = p.$$

Proof Notice first that, for all sufficiently large n, there are almost surely at least p elements in the array $[z_1, z_2, \ldots]$ that are non-zero. This follows from Assumption (A5) that $\mathbb{P}(Z = 0) < 1$, as done in the proof of (Lu and Li, 2020, Proposition 3.2). Now, the i-th column of ZS is obtained by scaling a standard p dimensional Gaussian by z_i. As the columns of S are almost surely independent and, for sufficiently large n, at least p elements z_i's are non-zero, it must be that at least p columns of ZS are linearly independent, which proves the claim. ∎

Lemma 7 *For every eigenvector v_i^a, it holds almost surely that*

$$qv_i^a \neq 0.$$

Proof By definition, it holds that $v_i^a \neq 0$. Thus, Lemma 6 implies that almost surely $ZSv_i^a \neq 0$. Furthermore, the elements of the matrix U are sampled independently from Z and S, so we can fix $x_i := ZSv_i^a \neq 0$ and conclude

$$\mathbb{P}\Big(U^\top x_i = 0\Big) = 0,$$

for the probability measure associated to the elements of U. This is due to the fact that, for $x_i \neq 0$,

$$\mathbb{P}\Big(U^\top x_i = 0\Big) = \mathbb{P}\big(\forall\, j \in [d-p],\ \langle u^j, x_i\rangle = 0\big) = \prod_{i=1}^{d-p} \mathbb{P}(\langle u_j, x_i\rangle = 0) = 0,$$

as each $u_j \in \mathbb{R}^n$ is sampled from a multi-variate Gaussian. Lastly, using the union bound, we have

$$\mathbb{P}\Big(\exists\, i \in [p], U^\top x_i = 0\Big) \leq \sum_{i=1}^{p} \mathbb{P}\Big(U^\top x_i = 0\Big) = 0,$$

implying that almost surely $qv_i^a \neq 0$ for every eigenvector of the matrix a. ∎

For i such that $p + 1 \leq i \leq d - p$, let us define

$$\tilde{L}_i(\mu) := \begin{cases} L_{i-p}(\mu) & \text{if } \lambda_p^a < \mu < \lambda_{p-1}^a, \\ \vdots & \\ L_i(\mu) & \text{if } \lambda_1^a < \mu < +\infty, \end{cases} \tag{A.1}$$

where we restrict the domain to $\mu \in]\lambda_p^a, +\infty[\setminus \Lambda^a$. This complements the definition of $\tilde{L}_i(\mu)$ for $i \in [p]$ in Equation (5.2). For $1 \leq i \leq d - p$, all functions $\tilde{L}_i$ can be continuously extended, as proved below.

Lemma 8 *For $i \in [p]$ and $j \geq 2$, it holds that*

$$\lim_{\mu_1 \to \lambda_i^{a+}} L_j(\mu_1) = \lim_{\mu_2 \to \lambda_i^{a-}} L_{j-1}(\mu_2).$$

Proof To simplify exposition, let us denote by $M_\mu := P - q(a - \mu I_p)^{-1} q^\top$, as well as by $r = \frac{q v_i^a}{\left\| q v_i^a \right\|_2}$ which is well defined as $q v_i^a \neq 0$. Using Weyl's inequality (see e.g. (Horn and Johnson, 2013, Section 4.3)), one has

$$\lambda_j(M_{\mu_1}) \leq \lambda_{j-1}(M_{\mu_2}) + \lambda_2(M_{\mu_1} - M_{\mu_2}).$$

Note that $M_{\mu_1} - M_{\mu_2} = q\big((a - \mu_2 I_p)^{-1} - (a - \mu_1 I_p)^{-1}\big) q^\top$ and that the SVD decomposition of $(a - \mu I_d)^{-1}$ is

$$(a - \mu I_p)^{-1} = \sum_{k=1}^{p} \frac{v_k^a v_k^{a\top}}{\lambda_k^a - \mu}.$$

Plugging that in, we get

$$(a - \mu_2 I_p)^{-1} - (a - \mu_1 I_p)^{-1} = v_i^a v_i^{a\top} \left(\frac{1}{\lambda_i^a - \mu_2} + \frac{1}{\mu_1 - \lambda_i^a} \right) + \sum_{k \neq i}^{p} v_k^a v_k^{a\top} \left(\frac{\mu_2 - \mu_1}{(\lambda_k^a - \mu_2)(\lambda_k^a - \mu_1)} \right).$$

Taking the limits $\mu_1 \to \lambda_i^{a+}$ and $\mu_2 \to \lambda_i^{a-}$, one readily obtains that $\lambda_2(M_{\mu_1} - M_{\mu_2}) := \varepsilon_{\mu_2} \to 0$. Before continuing, let us denote by $\Pi_r := r r^\top$ the orthogonal projection to the subspace defined by r, and by $\Pi_{r^\perp}$ the orthogonal projection to the subspace $r^\perp$. Obviously it holds that

$$\Pi_r + \Pi_{r^\perp} = I_{d-p}.$$

Thus,

$$\begin{aligned}
\lambda_{j-1}(M_{\mu_2}) &= \lambda_{j-1}((\Pi_r + \Pi_{r^\perp}) M_{\mu_2} (\Pi_r + \Pi_{r^\perp})) \\
&\leq \lambda_{j-1}(\Pi_{r^\perp} M_{\mu_2} \Pi_{r^\perp}) + \lambda_1(\Pi_r M_{\mu_2} \Pi_r + \Pi_{r^\perp} M_{\mu_2} \Pi_r + \Pi_r M_{\mu_2} \Pi_{r^\perp}),
\end{aligned}$$

where the last line is due to the fact that the matrix $\Pi_r M_{\mu_2} \Pi_r + \Pi_{r^\perp} M_{\mu_2} \Pi_r + \Pi_r M_{\mu_2} \Pi_{r^\perp}$ is symmetric and through a subsequent application of Weyl's inequality. Let us analyze the eigenvector corresponding to the largest eigenvalue. Namely, for an eigenvector $t = \Pi_{r^\perp} t + \Pi_r t =: t_{r^\perp} + t_r$ of the discussed matrix with corresponding eigenvalue λ, we have that

$$\begin{aligned}
\Pi_r M_{\mu_2} \Pi_r \, t_r + \Pi_r M_{\mu_2} \Pi_{r^\perp} \, t_{r^\perp} &= \lambda \, t_r, \\
\Pi_{r^\perp} M_{\mu_2} \Pi_r \, t_r &= \lambda \, t_{r^\perp}.
\end{aligned}$$

If t_r is the 0 vector, then $\lambda = 0$. Otherwise, as $\Pi_r M_{\mu_2} \Pi_{r^\perp} \, t_{r^\perp}$ has a convergent, finite limit as $\mu_2 \to \lambda_i^{a-}$ and $t_r^\top \Pi_r M_{\mu_2} \Pi_r \, t_r \to -\infty$, it must hold that $\lambda \to -\infty$. This gives the following inequality

$$\lambda_j(M_{\mu_1}) \leq \lambda_{j-1}(M_{\mu_2}) + \varepsilon_{\mu_2} \leq \lambda_{j-1}(\Pi_{r^\perp} M_{\mu_2} \Pi_{r^\perp}) + \varepsilon'_{\mu_2},$$

where $\varepsilon'_{\mu_2} \to 0$ as $\mu_2 \to \lambda_i^{a-}$.

Let us now lower bound $\lambda_j(M_{\mu_1})$. In a similar manner as before, we have that

$$\begin{aligned}\lambda_j(M_{\mu_1}) &= \lambda_j((\Pi_r + \Pi_{r^\perp})M_{\mu_1}(\Pi_r + \Pi_{r^\perp})) \\ &\geq \lambda_j\left(\frac{1}{2}\Pi_r M_{\mu_1}\Pi_r + \Pi_{r^\perp}M_{\mu_1}\Pi_{r^\perp}\right) + \lambda_{d-p}\left(\frac{1}{2}\Pi_r M_{\mu_1}\Pi_r + \Pi_{r^\perp}M_{\mu_1}\Pi_r + \Pi_r M_{\mu_1}\Pi_{r^\perp}\right) \\ &\geq \lambda_j\left(\frac{1}{2}\Pi_r M_{\mu_1}\Pi_r + \Pi_{r^\perp}M_{\mu_1}\Pi_{r^\perp}\right) - \varepsilon_{\mu_1},\end{aligned}$$

where as $\mu_1 \to \lambda_i^{a+}$ it holds that $\varepsilon_{\mu_1} \to 0$ with the same arguments as above.

Note that each of the eigenvectors of $\frac{1}{2}\Pi_r M_{\mu_1}\Pi_r$ corresponding to a non-zero eigenvalue is orthogonal to each of the eigenvectors of $\Pi_{r^\perp}M_{\mu_1}\Pi_{r^\perp}$ corresponding to a non-zero eigenvalue. Furthermore, $\lambda_1\left(\frac{1}{2}\Pi_r M_{\mu_1}\Pi_r\right) \to \infty$ as $\mu_1 \to \lambda_i^{a+}$. Lastly, since $\Pi_{r^\perp}M_{\mu_1}\Pi_{r^\perp}$ has a convergent, finite limit, it holds that

$$\lambda_j\left(\frac{1}{2}\Pi_r M_{\mu_1}\Pi_r + \Pi_{r^\perp}M_{\mu_1}\Pi_{r^\perp}\right) = \lambda_{j-1}(\Pi_{r^\perp}M_{\mu_1}\Pi_{r^\perp}),$$

as $\mu_1 \to \lambda_i^{a+}$. This allows us to conclude that

$$\lambda_{j-1}(\Pi_{r^\perp}M_{\mu_1}\Pi_{r^\perp}) - \varepsilon_{\mu_1} \leq \lambda_j(M_{\mu_1}) \leq \lambda_{j-1}(M_{\mu_2}) + \varepsilon_{\mu_2} \leq \lambda_{j-1}(\Pi_{r^\perp}M_{\mu_2}\Pi_{r^\perp}) + \varepsilon'_{\mu_2}.$$

Finally, by taking the limits $\mu_1 \to \lambda_i^{a+}$ and $\mu_2 \to \lambda_i^{a-}$ we get

$$\lim_{\mu_1 \to \lambda_i^{a+}} L_j(\mu_1) = \lim_{\mu_2 \to \lambda_i^{a-}} L_{j-1}(\mu_2) = \lambda_{j-1}\left(\Pi_{r^\perp}M_{\lambda_i^a}\Pi_{r^\perp}\right). \tag{A.2}$$

■

From now on, we will refer to the function $\tilde{L}_i$ as the one extended to the whole $]\lambda_i^a, +\infty[$ for $i \in [p]$, and to the whole $]\lambda_p^a, +\infty[$ for $i \geq p+1$. The next result characterizes the behavior of $\tilde{L}_i$ at the edges of its domain.

Lemma 9 *For $i \in [d-p]$, $\tilde{L}_i$ is non-increasing with the limit on the right edge of the domain given by*

$$\lim_{\mu \to \infty} \tilde{L}_i(\mu) = \lambda_i(P).$$

Moreover, for $i \in [p]$, the limit on the left edge of the domain is

$$\lim_{\mu \to \lambda_i^{a+}} \tilde{L}_i(\mu) = +\infty.$$

Proof First, we prove that each $L_i(\mu)$ is non-increasing in an arbitrary domain $]\lambda_j^a, \lambda_{j-1}^a[$. In fact, for any $\mu_1 > \mu_2$ in that interval, it holds that

$$\begin{aligned}L_i(\mu_1) - L_i(\mu_2) &= \lambda_i(P - q(a - \mu_1 I_p)^{-1}q^\top) - \lambda_i(P - q(a - \mu_2 I_p)^{-1}q^\top) \\ &\leq \lambda_1(q(a - \mu_2 I_p)^{-1}q^\top - q(a - \mu_1 I_p)^{-1}q^\top) \\ &= \lambda_1\left(q\sum_{k=1}^{p} v_k^a v_k^{a\top}\left(\frac{\mu_2 - \mu_1}{(\lambda_k^a - \mu_2)(\lambda_k^a - \mu_1)}\right)q^\top\right) \\ &\leq 0,\end{aligned}$$

where the first inequality is due to Weyl's inequality.

Thus, by the definition in (5.2) and by Lemma 8, we have that the function $\tilde{L}_i$ is non-increasing in $]\lambda_i^a, +\infty[$, for $i \in [p]$. In the same manner, for $i > p$, it also holds that $\tilde{L}_i$ is non-increasing in $]\lambda_p^a, +\infty[$. Moreover, since $qv_i^a \neq 0$ and

$$P - q(a - \mu I_p)^{-1}q^\top = P - \frac{(qv_i^a)(qv_i^a)^\top}{\lambda_i^a - \mu} - \sum_{k \neq i}^{p} \frac{(qv_k^a)(qv_k^a)^\top}{\lambda_k^a - \mu},$$

it holds that, for any $i \in [p]$,

$$\lim_{\mu \to \lambda_i^{a+}} \tilde{L}_i(\mu) = \lim_{\mu \to \lambda_i^{a+}} \lambda_1(P - q(a - \mu I_p)^{-1}q^\top) = +\infty.$$

Finally, using the same formula, one also obtains that, for $i \in [d-p]$,

$$\lim_{\mu \to \infty} \tilde{L}_i(\mu) = \lim_{\mu \to \infty} \lambda_i(P - q(a - \mu I_p)^{-1}q^\top) = \lambda_i(P),$$

which concludes the proof. ■

Having proven this properties, let us get back to discussing the eigenvalues λ_i^D. We do so by considering two cases.

Lemma 10 *An eigenvalue $\lambda_i^D \notin \Lambda^a$, $i \in [d-p]$, is a solution to*

$$L_k(\mu) = \mu, \tag{A.3}$$

for some k. Conversely, any solution to the previous equation is an eigenvalue of D that is also not an eigenvalue of a.

Proof All eigenvalues of D are exactly the solutions to

$$\det(D - \lambda I_d) = 0. \tag{A.4}$$

Applying the formula for the determinant of a block matrix implies

$$\det(D - \lambda I_d) = \det\Big(P - \lambda I_{d-p} - q(a - \lambda I_p)^{-1}q^\top\Big)\det(a - \lambda I_p).$$

As by assumption $\det(a - \lambda I_d) \neq 0$, (A.4) is equivalent to

$$\det\Big(P - \lambda I_{d-p} - q(a - \lambda I_p)^{-1}q^\top\Big) = 0.$$

Moreover by definition of the determinant and the fact that the matrix in questions is symmetric, it holds that

$$\begin{aligned}\det\Big(P - \lambda I_{d-p} - q(a - \lambda I_p)^{-1}q^\top\Big) &= \prod_{i=1}^{d-p} \lambda_i\Big(P - \lambda I_{d-p} - q(a - \lambda I_p)^{-1}q^\top\Big) \\ &= \prod_{i=1}^{d-p}(L_i(\lambda) - \lambda).\end{aligned}$$

Therefore, we have that

$$\det\Big(P - \lambda I_{d-p} - q(a - \lambda I_p)^{-1} q^\top\Big) = 0,$$

if and only if there exists a k and μ such that

$$L_k(\mu) = \mu.$$

■

The case in which the eigenvalues of D and a overlap is covered by the result below.

Lemma 11 *An arbitrary eigenvalue $\lambda_i^D \in \Lambda^a$ is equal to λ_j^a if and only if*

$$\lim_{\mu \to \lambda_j^{a}+} L_k(\mu) = \lambda_j^a, \tag{A.5}$$

for some $k \geq 2$.

Proof Let us first prove the if part of the statement. We denote the eigenvector corresponding to λ_i^D as $v_i^D = \begin{bmatrix} h_i \\ g_i \end{bmatrix}$, where $h_i \in \mathbb{R}^p$, $g \in \mathbb{R}^{d-p}$. It follows that

$$Dv_i^D = \begin{bmatrix} a & q^\top \\ q & P \end{bmatrix} \begin{bmatrix} h_i \\ g_i \end{bmatrix} = \lambda_j^a \begin{bmatrix} h_i \\ g_i \end{bmatrix}.$$

Splitting this equation into p and $d-p$ coordinates gives

$$ah_i + q^\top g_i = \lambda_j^a h_i, \tag{A.6}$$
$$qh_i + Pg_i = \lambda_j^a g_i. \tag{A.7}$$

Since $(a - \lambda_j^a I_p)$ is singular, its SVD decomposition is

$$(a - \lambda_j^a I_p) = \sum_{k=1}^{p} (\lambda_k^a - \lambda_j^a) v_k^a {v_k^a}^\top = \sum_{k \neq j}^{p} (\lambda_k - \lambda_j) v_k^a {v_k^a}^\top.$$

From (A.6), it follows that $(a - \lambda_j^a I_p) h_i = -q^\top g_i$. Then, plugging in the SVD, it holds that $\sum_{k\neq j}^{p} (\lambda_k - \lambda_j) v_k^a {v_k^a}^\top h_i = -q^\top g_i$. Multiplying both sides by v_j^a, due to the matrix being symmetric and thus eigenvectors orthogonal, it holds that

$$\sum_{k \neq j}^{p} (\lambda_k - \lambda_j) \langle v_k^a, h_i \rangle \big\langle v_k^a, v_j^a \big\rangle = 0 = -\Big\langle q^\top g_i, v_j^a \Big\rangle.$$

From there, we can conclude that $q^\top g_i \perp v_j^a$, which is equivalent to $g_i \perp q v_j^a$. Moreover, (A.6) can be rewritten as

$$h_i = -(a - \lambda_j^a I_p)^\dagger q^\top g_i + \alpha v_j^a,$$

for some α. Plugging this result into (A.7) gives

$$-q(a - \lambda_j^a I_p)^\dagger q^\top g_i + \alpha q v_j^a + P g_i = \lambda_j^a g_i, \tag{A.8}$$

for some α. Let us, as before, denote by $r = qv_j^a/\|qv_j^a\|_2$, and by $\Pi_{r^\perp}$ the orthogonal projection to the subspace defined by $r^\perp$. As noted before, $g_i \perp qv_j^a$, which implies that $g_i = \Pi_{r^\perp} g_i$. Plugging it into Equation (A.8) gives

$$-q(a - \lambda_j^a I_p)^\dagger q^\top \Pi_{r^\perp} g_i + \alpha q v_j^a + P\Pi_{r^\perp} g_i = \lambda_j^a \Pi_{r^\perp} g_i.$$

Multiplying the previous equation on the left by $\Pi_{r^\perp}$ results in

$$-\Pi_{r^\perp} q(a - \lambda_j^a I_p)^\dagger q^\top \Pi_{r^\perp} g_i + \Pi_{r^\perp} P\Pi_{r^\perp} g_i = \lambda_j^a \Pi_{r^\perp} g_i.$$

This means that $\Pi_{r^\perp} g_i$ is an eigenvector of the matrix $-\Pi_{r^\perp} q(a - \lambda_j^a I_p)^\dagger q^\top \Pi_{r^\perp} + \Pi_{r^\perp} P \Pi_{r^\perp}$ with the corresponding eigenvalue λ_j^a, i.e.,

$$\lambda_k(\Pi_{r^\perp}(P - q(a - \lambda_j^a I_p)^\dagger q^\top)\Pi_{r^\perp}) = \lambda_j^a,$$

for some k. From (A.2) we can see the LHS is exactly $\lim_{\mu \to \lambda_j^{a}+} L_k(\mu)$ for some k. As proved in Lemma 9, it holds that $\lim_{\mu \to \lambda_j^{a}+} L_1(\mu) = +\infty$, so it must be that $k \geq 2$.

Conversely, by following the same steps in reverse, if $\lim_{\mu \to \lambda_j^{a}+} L_k(\mu) = \lambda_j^a$, then there is a vector g_i that solves

$$-q(a - \lambda_j^a)^\dagger q^\top g_i + \alpha q v_j^a + P g_i = \lambda_j^a g_i,$$

for some α. By setting

$$h_i = -(a - \lambda_j^a I_p)^\dagger q^\top g_i + \alpha v_j^a,$$

it follows that $w = \begin{bmatrix} h_i \\ g_i \end{bmatrix}$ is an eigenvector of D with eigenvalue λ_j^a as stated. ■

Combining the previous properties, we are now ready to state and formally prove the characterization in Equation (5.3).

Proposition 12 *For $i \in [p]$, the eigenvalue λ_i^D is the unique solution to the equation*

$$\tilde{L}_i(\mu) = \mu, \tag{A.9}$$

in the respective domain $]\lambda_i^a, \infty[$.

Proof Let us first prove that (A.9) has a unique solution, for $i \in [p]$. Since $\tilde{L}_i(\mu)$ is non-increasing and continuous, we have that $\tilde{L}_i(\mu) - \mu$ is decreasing and continuous. Moreover, for $i \in [p]$ it has limits

$$\lim_{\mu \to \lambda_i^{a}+} \tilde{L}_i(\mu) - \mu = +\infty, \text{ and } \lim_{\mu \to \infty} \tilde{L}_i(\mu) - \mu = -\infty,$$

due to Lemma 9. Then, applying the intermediate value theorem implies that there must be a unique μ_i for which $\tilde{L}_i(\mu_i) - \mu_i = 0$.

Next, let us prove that the unique solution μ_i of (A.9) is indeed an eigenvalue of D. First, suppose that $\mu_i = \lambda_j^a$, for some $j \in [p]$. Then, (A.2) would imply that

$$\lim_{\mu \to \lambda_j^{a}+} L_k(\mu) = \lambda_j^a,$$

for some $k \geq 2$. Then, Lemma 11 implies that λ_j^a is an eigenvalue of D. Next, suppose that $\mu_i \notin \Lambda^a$. Then, by definition of $\tilde{L}_i$, it holds that

$$L_k(\mu_i) = \mu_i,$$

for some k. Lemma 10 then implies that μ_i must be an eigenvalue of D.

Finally, let us prove that μ_i is exactly the i-th eigenvalue of D. To do so, we first prove that every eigenvalue of D that is larger or equal to λ_p^a is a solution to the following equation in μ:

$$\tilde{L}_m(\mu) = \mu,$$

for some $m \in [d-p]$. This follows from Lemmas 10 and 11, which imply that any eigenvalue of D is covered by checking the conditions

$$L_k(\mu) = \mu \text{ or } \lim_{\mu \to \lambda_j^a +} L_k(\mu) = \lambda_j^a,$$

which are all covered by considering $\tilde{L}_m(\mu)$ for $m \in [d-p]$.

As $\tilde{L}_1(\mu) \geq \tilde{L}_2(\mu) \geq \cdots \geq \tilde{L}_p(\mu) \geq \cdots \geq \tilde{L}_{d-p}(\mu)$ and $\lambda_1^D \geq \lambda_2^D \geq \ldots \lambda_p^D$, it must be that the solution to (A.9) is exactly the i-th eigenvalue of the matrix D, and the proof is complete. ■

Next, we prove the eigenvector characterization in Equation (5.4).

Proposition 13 *Let $j \in [p]$ be s.t. $\lambda_i^D > \lambda_1^P$ for all $i \leq j$. Then, for all $i \leq j$, it holds that*

$$h_i = \frac{\tilde{h}_i}{\sqrt{1 - \tilde{h}_i^\top \frac{d}{d\lambda} R(\lambda_i^D) \tilde{h}_i}},$$

where $\tilde{h}_i$ is the eigenvector of $R(\lambda_i^D)$ and $\frac{d}{d\lambda} R(\lambda)$ is the entry-wise derivative of $R(\lambda)$.

Proof Note that the eigenvector equation is equivalent to the system of two equations

$$a h_i + q^\top g_i = \lambda_i^D h_i, \tag{A.10}$$

$$q h_i + P g_i = \lambda_i^D g_i. \tag{A.11}$$

As we consider only the eigenvectors v_i^D for $i \leq j$, the matrix $(P - \lambda_i^D I_{d-p})$ is invertible, and solving (A.10) gives

$$g_i = -(P - \lambda_i^D I_{d-p})^{-1} q h_i.$$

Substituting in (A.11) yields

$$a h_i - q^\top (P - \lambda_i^D I_{d-p})^{-1} q h_i = \lambda_i^D h_i.$$

Let us denote by $\tilde{h}_i = \frac{h_i}{\|h_i\|_2}$ the unit norm eigenvector of $a - q^\top (P - \lambda_i^D I_{d-p})^{-1} q$ corresponding to the eigenvalue λ_i^D, and also define $\tilde{g}_i := -(P - \lambda_i^D I_{d-p})^{-1} q \tilde{h}_i$. Then, $\tilde{h}_i$ and $\tilde{g}_i$ satisfy equations

(A.10) and (A.11), so $\tilde{v}_i^D = \begin{bmatrix} \tilde{h}_i \\ \tilde{g}_i \end{bmatrix}$ is aligned with an eigenvector corresponding to eigenvalue λ_i^D. However, $\tilde{v}_i^D$ does not necessarily have unit norm. It holds that

$$\begin{bmatrix} h_i \\ g_i \end{bmatrix} = v_i^D = \frac{\tilde{v}_i^D}{\|\tilde{v}_i^D\|_2} = \frac{\begin{bmatrix} \tilde{h}_i \\ \tilde{g}_i \end{bmatrix}}{\sqrt{\tilde{h}_i^\top \tilde{h}_i + \tilde{g}_i^\top \tilde{g}_i}} = \frac{\begin{bmatrix} \tilde{h}_i \\ \tilde{g}_i \end{bmatrix}}{\sqrt{1 + \tilde{h}_i^\top q^\top (P - \lambda_i^D I_{d-p})^{-2} q \tilde{h}_i}},$$

from which follows that

$$h_i = \frac{\tilde{h}_i}{\sqrt{1 + \tilde{h}_i^\top q^\top (P - \lambda_i^D I_{d-p})^{-2} q \tilde{h}_i}}.$$

The last thing to notice is that

$$q^\top (P - \lambda I_{d-p})^{-2} q = -\frac{d}{d\lambda}(a - q^\top (P - \lambda I_{d-p})^{-1} q),$$

from which the statement of the proposition follows. ■

We conclude by expressing the entries of the matrix $R(\lambda)$ and its derivative in terms of the auxiliary functions in Equation (5.5).

Lemma 14 *Let us assume that $\lambda > \lambda_1^P$. Then, for the diagonal elements of $R(\lambda)$, it holds that*

$$R(\lambda)_{i,i} = a_{i,i} + \frac{1}{\mathcal{L}_i^{-1}(\lambda)}, \tag{A.12}$$

$$\frac{d}{d\lambda} R(\lambda)_{i,i} = -\frac{1}{\left(\mathcal{L}_i^{-1}(\lambda)\right)^2 \mathcal{L}_i'(\mathcal{L}_i^{-1}(\lambda))}. \tag{A.13}$$

Moreover, for the off-diagonal elements, we have

$$2R(\lambda)_{i,j} = 2a_{i,j} + \frac{1}{\mathcal{L}_{i,j}^{-1}(\lambda)} - \frac{1}{\mathcal{L}_i^{-1}(\lambda)} - \frac{1}{\mathcal{L}_j^{-1}(\lambda)}, \tag{A.14}$$

$$2\frac{d}{d\lambda} R(\lambda)_{i,j} = \frac{1}{\left(\mathcal{L}_i^{-1}(\lambda)\right)^2 \mathcal{L}_i'(\mathcal{L}_i^{-1}(\lambda))} + \frac{1}{\left(\mathcal{L}_j^{-1}(\lambda)\right)^2 \mathcal{L}_j'(\mathcal{L}_j^{-1}(\lambda))} - \frac{1}{\left(\mathcal{L}_{i,j}^{-1}(\lambda)\right)^2 \mathcal{L}_{i,j}'(\mathcal{L}_{i,j}^{-1}(\lambda))}. \tag{A.15}$$

Proof Note that

$$R(\lambda)_{i,i} = a_{i,i} - q_i^\top (P - \lambda I_{d-p})^{-1} q_i, \qquad R(\lambda)_{i,j} = a_{i,j} - q_i^\top (P - \lambda I_{d-p})^{-1} q_j. \tag{A.16}$$

Transforming (A.16) with the matrix determinant lemma yields that, for any $\lambda > \lambda_1^P$,

$$R(\lambda)_{i,i} = a_{i,i} - q_i^\top (P - \lambda I_{d-p})^{-1} q_i = a_{i,i} + \frac{1}{\mathcal{L}_i^{-1}(\lambda)}. \tag{A.17}$$

Moreover, it holds that

$$R(\lambda)_{i,j} = a_{i,j} - \frac{(q_i+q_j)^\top (P-\lambda I_{d-p})^{-1}(q_i+q_j) - q_i^\top (P-\lambda I_{d-p})^{-1} q_i - q_j^\top (P-\lambda I_{d-p})^{-1} q_j}{2}.$$

From the same transformation with the matrix determinant lemma, it follows that

$$(q_i+q_j)^\top (P-\lambda I_{d-p})^{-1}(q_i+q_j) = -\frac{1}{\mathcal{L}_{i,j}^{-1}(\lambda)}.$$

Substituting the previous identity in Equation (A.17) gives Equation (A.14). Note that, for μ such that $\mathcal{L}_i(\mu) > \lambda_1^P$, it holds that $\mathcal{L}_i$ is an increasing differentiable function, so its inverse and derivative are well defined. Finally, by differentiating Equation (A.12) and Equation (A.14), we get the other two equations. ∎

Appendix B. Proofs for the characterization of eigenvalues

B.1. Auxiliary results

As a consequence of Proposition 12, the top p eigenvalues of D are entirely characterized by the functions $\tilde{L}_i$. As these functions are nothing more than patches of the functions L_i on different domains, we first direct our attention to analyzing asymptotic behavior of $L_i(\mu) = \lambda_i(P - q(a - \mu I_p)^{-1} q^\top)$. Notice that

$$P - q(a-\mu I_p)^{-1} q^\top = \frac{1}{n} U^\top M_n U, \tag{B.1}$$

where $M_n := Z - \frac{1}{n} ZS(a-\mu I_p)^{-1} S^\top Z$ is a rank p perturbation of the matrix Z.

Lemma 15 *For each $\mu > 0$, let $\alpha_1 \geq \ldots \geq \alpha_j > \tau$ be all the solutions to the equation*

$$\det(\mu I_p - R^\infty(\alpha)) = 0. \tag{B.2}$$

Then, for the top j eigenvalues of M_n, it holds that

$$\lambda_1^M, \ldots, \lambda_j^M \xrightarrow{\text{a.s.}} \alpha_1, \ldots, \alpha_j, \tag{B.3}$$

and for the remaining $p-j$ eigenvalues, it holds that

$$\lambda_{j+1}^M, \ldots, \lambda_p^M \xrightarrow{\text{a.s.}} \tau.$$

Proof Let us denote by $v := ZS$. An arbitrary eigenvalue λ_k^M of M_n satisfies the equation

$$\det\left(Z - \frac{1}{n} v(a-\mu I_p)^{-1} v^\top - \lambda_k^M I_n\right) = 0.$$

Thus, for $\alpha > \max\{z_i\}$, consider the following equation

$$\det\left(Z - \frac{1}{n} v(a-\mu I_p)^{-1} v^\top - \alpha I_n\right) = 0.$$

As $Z - \alpha I_n$ is invertible for $\alpha > \max\{z_i\}$, we can apply the matrix determinant lemma to obtain the equivalent equation

$$\det\left(\mu I_p - a + \frac{1}{n} v^\top (Z - \alpha I_n)^{-1} v\right) = 0. \tag{B.4}$$

Moreover,

$$a - \frac{1}{n} v^\top (Z - \alpha I_n)^{-1} v = \frac{1}{n}\sum_{i=1}^n z_i s_i s_i^\top - \frac{1}{n}\sum_{i=1}^n \frac{z_i^2 s_i s_i^\top}{z_i - \alpha} = \frac{1}{n}\sum_{i=1}^n \frac{\alpha z_i s_i s_i^\top}{\alpha - z_i}.$$

Thus, (B.4) becomes

$$\det\left(\mu I_p - \frac{1}{n}\sum_{i=1}^n \frac{\alpha z_i s_i s_i^\top}{\alpha - z_i}\right) = 0. \tag{B.5}$$

Let us prove that, for n large enough, this equation indeed has its top j solutions for $\alpha > \max\{z_i\}$. First, note that

$$\det\left(\mu I_p - \frac{1}{n}\sum_{i=1}^n \frac{\alpha z_i s_i s_i^\top}{\alpha - z_i}\right) = \prod_{i=1}^p (\mu - \lambda_i(\alpha)),$$

where through abuse of notation we define $\lambda_i(\alpha) :]\max\{z_i\}, +\infty[\to \mathbb{R}$ as $\lambda_i\left(\frac{1}{n}\sum_{i=1}^n \frac{\alpha z_i s_i s_i^\top}{\alpha - z_i}\right)$. Each function λ_i is continuous and strictly decreasing. This can be seen by taking arbitrary $\alpha_2 > \alpha_1$ to get

$$\lambda_i(\alpha_1) - \lambda_i(\alpha_2) \geq \lambda_p\left((\alpha_2 - \alpha_1)\frac{1}{n}\sum_{i=1}^n \frac{z_i^2 s_i s_i^\top}{(\alpha_1 - z_i)(\alpha_2 - z_i)}\right) > 0,$$

by using Weyl's inequality and the fact that there are almost surely at least p linearly independent vectors $z_i s_i$ by Lemma 6. Consequently, to prove that (B.5) has j solutions for $\alpha > \max\{z_i\}$, it equivalent to prove that, for each i,

$$\lambda_i(\beta_i') > \mu > \lambda_i(\beta_i''), \tag{B.6}$$

for some $\beta_i'' > \beta_i' > \max z_i$.

Note that, for any fixed α, it holds that

$$\frac{1}{n}\sum_{i=1}^n \frac{\alpha z_i s_i s_i^\top}{\alpha - z_i} \xrightarrow{\text{a.s.}} \mathbb{E}\left[\frac{\alpha z s s^\top}{\alpha - z}\right] = R^\infty(\alpha), \tag{B.7}$$

due to the law of large numbers. Due to the continuity of eigenvalues, it further follows that

$$\lambda_i\left(\frac{1}{n}\sum_{i=1}^n \frac{\alpha z_i s_i s_i^\top}{\alpha - z_i}\right) \xrightarrow{\text{a.s.}} \lambda_i\left(\mathbb{E}\left[\frac{\alpha z s s^\top}{\alpha - z}\right]\right) = \lambda_i^\infty(\alpha),$$

where $\lambda_i^\infty(\alpha)$ is continuous and strictly decreasing. This can be seen as, for any $\alpha > \tau$ and any arbitrary vector $x \in \mathbb{R}^p$, it holds that

$$\frac{d}{d\alpha}\left(x^\top \left(\mathbb{E}\left[\frac{\alpha z s s^\top}{\alpha - z}\right]\right) x\right) = -\mathbb{E}\left[\frac{\langle x, s\rangle^2 z^2}{(\alpha - z)^2}\right] < 0, \tag{B.8}$$

since $\mathbb{P}(z=0)<1$. Moreover, for $i \in [p]$,

$$\lim_{\alpha\to\infty} \lambda_i^{\infty}(\alpha) = \lambda_i^{a^{\infty}},$$

where the matrix $a^{\infty} = \mathbb{E}\left[zss^{\top}\right]$ is the limit of the matrix a. The condition of the lemma states that there exist $\alpha_1 \geq \ldots \geq \alpha_j > \tau$ such that

$$\det(\mu I_p - R^{\infty}(\alpha)) = 0.$$

Let us denote by $k \in \{0,\ldots,p\}$ the index such that $\lambda_{k+1}^{a^{\infty}} \leq \mu < \lambda_k^{a^{\infty}}$, with the abuse of notation $\lambda_0^{a^{\infty}} := +\infty$ and $\lambda_{p+1}^{a^{\infty}} := -\infty$. By assumption

$$\det(\mu I_p - R^{\infty}(\alpha)) = \prod_{i=0}^{p}(\mu - \lambda_i^{\infty}(\alpha)), \tag{B.9}$$

has j solutions in $\alpha \in]\tau,+\infty[$. Note that λ_i^{∞} is a strictly decreasing continuous function, so the only way that $\lambda_i^{\infty} - \mu$ does not have a solution in $\alpha \in]\tau,+\infty[$ is if either $\lim_{\alpha\to\tau^+}\lambda_i^{\infty}(\alpha) < \mu$ or $\lim_{\alpha\to\infty}\lambda_i^{\infty}(\alpha) > \mu$. Moreover, since $\lim_{\alpha\to\infty}\lambda_i^{\infty}(\alpha) = \lambda_i^{a^{\infty}}$, it will exactly hold for $i \in \{1,\ldots k\}$ that

$$\lim_{\alpha\to\infty}\lambda_i^{\infty}(\alpha) \geq \lambda_k^{a^{\infty}} > \mu. \tag{B.10}$$

The fact that there are only j solutions to (B.9) and $\lambda_i^{\infty}(\alpha) > \lambda_{i+1}^{\infty}(\alpha)$ implies that

$$\lambda_{i+k}^{\infty}(\alpha_i) = \mu,$$

for $i \in [j]$, as well as

$$\lim_{\alpha\to\infty}\lambda_i^{\infty}(\alpha) < \mu, \tag{B.11}$$

for $i \in \{j+k+1,\ldots,p\}$.

As each λ_i^{∞} is a strictly decreasing continuous function, this further implies that there exists some constant $\varepsilon > 0$ and $\alpha_1',\ldots,\alpha_j' > \tau$ such that

$$\lambda_{i+k}^{\infty}(\alpha_i') = \mu + \varepsilon.$$

Applying the convergence of (B.7) it further holds that

$$\lambda_{i+k}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\alpha_i' z_i s_i s_i^{\top}}{\alpha_i' - z_i}\right) \xrightarrow{\text{a.s.}} \lambda_{i+k}^{\infty}(\alpha_i') = \mu + \varepsilon.$$

Thus, for each i and $\varepsilon > \varepsilon_1 > 0$, there exists n_0 s.t. for $n > n_0$

$$\left|\lambda_{i+k}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\alpha_i' z_i s_i s_i^{\top}}{\alpha_i' - z_i}\right) - (\mu+\varepsilon)\right| < \varepsilon_1.$$

Developing the absolute value, it holds that

$$\lambda_{i+k}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\alpha_i' z_i s_i s_i^{\top}}{\alpha_i' - z_i}\right) > \mu + \varepsilon - \varepsilon_1 > \mu,$$

e.g. by taking $\varepsilon_1 = \varepsilon/2$. As $\lambda_{i+k}(\alpha) = \lambda_{i+k}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\alpha z_i s_i s_i^\top}{\alpha - z_i}\right)$ is a continuous decreasing function, starting from some n_0 there exists $\beta_i' > \tau$ s.t. $\lambda_{i+k}(\beta_i') > \mu$. Notice that by definition $\tau > \max z_i$ almost surely. In the same way as for β_i' it can be proved that there exists $\beta_i'' > \tau$ such that $\lambda_{i+k}(\beta_i'') < \mu$.

Thus, we conclude that, for large enough n, $\lambda_{i+k}(\alpha) = \mu$ has j solutions larger than $\max\{z_i\}$. These are indeed λ_i^M for $1 \leq i \leq j$. Due to monotonicity, each λ_i admits a functional inverse and it holds that

$$\lambda_i^M = \lambda_{i+k}^{-1}(\mu).$$

As $\lambda_{i+k} \xrightarrow{\text{a.s.}} \lambda_{i+k}^\infty$, applying (Lu and Li, 2020, Lemma A.1) implies that

$$\lambda_i^M \xrightarrow{\text{a.s.}} (\lambda_{i+k}^\infty)^{-1}(\mu), \tag{B.12}$$

for $1 \leq i \leq j$, which is exactly the statement (B.3) of the lemma.

Let us now prove the second part of the statement. To do so, we prove that, for large enough n, (B.5) has no more than j solution for $\alpha > \max\{z_i\}$. As stated in (B.10) and (B.11), it holds that $\lim_{\alpha\to\infty} \lambda_i^\infty(\alpha) > \mu$ for i s.t. $1 \leq i \leq k$ and that $\lim_{\alpha\to\tau^+} \lambda_i^\infty(\alpha) < \mu$ for i s.t. $j+1+k \leq i \leq p$. Thus, using the same argument as before, we also have that, for large enough n and any $\alpha > \tau$, $\lambda_i(\alpha) > \mu$ for i s.t. $1 \leq i \leq k$ and $\lambda_i(\alpha) < \mu$ for i s.t. $j+1+k \leq i \leq p$. Since $\max\{z_i\} \xrightarrow{\text{a.s.}} \tau$, as τ is the right edge of the support of z, such inequalities also hold for $\alpha > \max\{z_i\}$ (and n large enough). Hence, there cannot exist β_i' and β_i'' that satisfy (B.6), so (B.5) cannot have more than j solutions in $\alpha > \max\{z_i\}$.

From this, it directly follows that, for n large enough and any $l \in \{j+1\dots p\}$,

$$\lambda_l^M \leq \max\{z_i\},$$

almost surely. Furthermore, from the interlacing theorem, it holds that

$$\lambda_{p+1}^Z \leq \lambda_l^M,$$

for any $l \in \{j+1,\dots,p\}$. Thus, the λ_l^M's are sandwiched between the first and the p-th largest value of Z, both of which converge to the right edge of the bulk τ as a finite order statistic, which gives the desired result. ■

Proposition 16 *For any fixed $\mu \in]\lambda_i^{a^\infty}, t_i^\infty[$, it holds that*

$$\tilde{L}_i(\mu) \xrightarrow{\text{a.s.}} \tilde{L}_i^\infty(\mu) = \zeta_\delta\big((\lambda_i^\infty)^{-1}(\mu)\big). \tag{B.13}$$

Furthemore, for any fixed $\mu \in]t_i^\infty, +\infty[$, it holds that

$$\tilde{L}_i(\mu) \xrightarrow{\text{a.s.}} \zeta_\delta\big(\bar{\lambda}_\delta\big). \tag{B.14}$$

Proof Let k be such that $\lambda_{k+1}^{a^\infty} \leq \mu < \lambda_k^{a^\infty}$, with the abuse of notation $\lambda_0^{a^\infty} := +\infty$. Then, by the definition of $\tilde{L}_i$ in (5.2), it holds that

$$\tilde{L}_i(\mu) = L_{i-k}(\mu), \tag{B.15}$$

for n large enough. This is true since $\lambda_i^a \xrightarrow{\text{a.s.}} \lambda_i^{a^\infty}$, as $a \xrightarrow{\text{a.s.}} a^\infty$.

To obtain the convergence of the RHS in (B.15), we rely on the results from Bai and Yao (2012). Towards this end, we recall the definition of $L_{i-k}(\mu) = \lambda_{i-k}(P - q(a - \mu I_p)^{-1} q^\top)$ as in (5.2). Given the equality in (B.1), we turn our attention to the eigenvalues of M_n. Recall that the functions $\lambda_i^\infty(\alpha)$ are strictly decreasing with limits

$$\lim_{\alpha \to \tau^+} \lambda_i^\infty(\alpha) = t_i^\infty, \text{ and } \lim_{\alpha \to +\infty} \lambda_i^\infty(\alpha) = \lambda_i^{a^\infty}.$$

Then, as $\mu \in]\lambda_i^{a^\infty}, t_i^\infty[$, the equation

$$\lambda_i^\infty(\alpha) = \mu$$

has a unique solution in $\alpha > \tau$. Let us denote that solution by α_{i-k}. Due to the fact that

$$\lambda_i^{a^\infty} \cdots \leq \lambda_{k+1}^{a^\infty} \leq \mu < \lambda_k^{a^\infty} \text{ and } \mu < t_i^\infty \cdots < t_1^\infty,$$

using the same argument as in the proof of Lemma 15 below (B.9), we conclude that there are unique solutions $\alpha_1, \dots, \alpha_{i-k}$ s.t. $\lambda_{j+k}(\alpha_j) = \mu$ for $j \in [i-k]$. Then, it holds that $\alpha_1, \dots, \alpha_{i-k}$ satisfy the conditions of Lemma 15. From its proof, specifically (B.12), it follows that

$$\lambda_{i-k}^M \xrightarrow{\text{a.s.}} \lambda_i^{\infty -1}(\mu) = \alpha_{i-k}.$$

Furthermore, the empirical spectral distribution of M_n converges almost surely to the distribution of z. This claim follows from Cauchy's interlacing theorem, using the same argument as in the proof of (Lu and Li, 2020, Proposition 3.2). Assuming that the preprocessing function $\mathcal{T}$ is positive, we can apply (Bai and Yao, 2012, Theorems 4.1 and 4.2) to get

$$L_{i-k}(\mu) \xrightarrow{\text{a.s.}} \zeta_\delta\big((\lambda_i^\infty)^{-1}(\mu)\big),$$

following the steps in (Lu and Li, 2020, Proposition 3.3). Finally, the adjustment in (Mondelli and Montanari, 2019, Lemma 3) covers the case in which the preprocessing function is not necessarily positive, and the proof of Equation (B.13) is complete.

Let us now consider $\mu \in]t_i^\infty, +\infty[$ and prove Equation (B.14). Note that, in this interval of μ, the equation

$$\lambda_i^\infty(\alpha) = \mu$$

has no solutions in $\alpha > \tau$. Let us examine the equation (B.2) in Lemma 15:

$$\det(\mu I_p - R^\infty(\alpha)) = \prod_{l=1}^p (\mu - \lambda_l^\infty(\alpha)) = 0.$$

The previous equation has a solution $\mu - \lambda_l^\infty(\alpha) = 0$, as long as

$$\lambda_l^{a^\infty} < \mu < t_l^\infty, \tag{B.16}$$

due to the monotonicity of each $\lambda_l^\infty(\alpha)$. As

$$\lambda_i^{a^\infty} \cdots \leq \lambda_{k+1}^{a^\infty} \leq \mu < \lambda_k^{a^\infty} \text{ and } \mu > t_i^\infty \cdots \geq t_p^\infty,$$

it follows that (B.16), thus also (B.2), can have at most $i-(k+1)$ solutions. Thus, applying Lemma 15 it follows that

$$\lambda_{i-k}^{M} \xrightarrow{\text{a.s.}} \tau,$$

implying that M_n in limit has at most $i-(k+1)$ spikes outside the bulk. Moreover, the empirical spectral distribution of M_n converges almost surely to the distribution of z, so we conclude that

$$\tilde{L}_i(\mu) = L_{i-k}(\mu) \xrightarrow{\text{a.s.}} \zeta_\delta(\bar{\lambda}_\delta),$$

as the limit of the right edge of the spectral distribution of $\frac{1}{n}U^\top ZU$ by (Bai and Yao, 2012, Lemma 3.1), which is due to (Silverstein and Choi, 1995, Section 4). ■

B.2. Proof of Theorem 2

We start by recalling some definitions. For each i, let $\lambda_i^\infty(\alpha) = \lambda_i(R^\infty(\alpha))$ be the i-th largest eigenvalue of $R^\infty(\alpha)$ and let t_i^∞ be

$$t_i^\infty := \lim_{\alpha\to\tau} \lambda_i^\infty(\alpha).$$

Furthermore, due to the law of large numbers, we have

$$a \xrightarrow{\text{a.s.}} a^\infty := \mathbb{E}\left[zss^\top\right] \in \mathbb{R}^{p\times p}.$$

Lastly, consider

$$\tilde{L}_i^\infty(\mu) := \zeta_\delta\big((\lambda_i^\infty)^{-1}(\mu)\big), \tag{B.17}$$

on the domain $\mu \in]\lambda_i^{a^\infty}, t_i^\infty[$. Note that $\tilde{L}_i^\infty(\mu)$ is a continuous non-increasing function, as it is the composition of a non-decreasing function and a strictly increasing function. We are now ready to prove our characterization of the eigenvalues of D_n.

Proof Note that (4.2) can be reformulated as

$$\det(\zeta_\delta(\alpha)I_p - R^\infty(\alpha)) = \prod_{i=1}^{p}(\zeta_\delta(\alpha) - \lambda_i^\infty(\alpha)) = 0. \tag{B.18}$$

The assumption of the theorem implies that $\alpha_1 \geq \cdots \geq \alpha_j > \tau$ satisfy (B.18). Recall that each function $\zeta_\delta(\alpha) - \lambda_i^\infty(\alpha)$ is strictly increasing on $]\tau, +\infty[$, with the right edge limit $+\infty$. Therefore, the implicit assumption that $j \in [p]$ is justified as there could be at most 1 solution to the equation

$$\zeta_\delta(\alpha) - \lambda_i^\infty(\alpha) = 0,$$

for each $i \in [p]$. Moreover, it holds by definition that $\lambda_1^\infty(\alpha) \geq \cdots \geq \lambda_p^\infty(\alpha)$. Thus, it must be that each α_i is the unique solution to

$$\zeta_\delta(\alpha) - \lambda_i^\infty(\alpha) = 0,$$

for $i \in [j]$ and $\alpha > \tau$. Let us denote by $\mu_i^* := \lambda_i^\infty(\alpha_i) \in]\lambda_i^{a^\infty}, \tau_i^\infty[$. Then μ_i^* is a solution to the equation

$$\tilde{L}_i^\infty(\mu) - \mu = 0,$$

in the domain of definition $\tilde{L}_i^\infty(\mu)$, as $\tilde{L}_i^\infty(\mu_i^*) = \zeta_\delta((\lambda_i^\infty)^{-1}(\lambda_i^\infty(\alpha_i)))$. Moreover, μ_i^* is the unique such solution, due to the strict monotonicity of $\tilde{L}_i^\infty(\mu) - \mu$ on its domain $]\lambda_i^{a^\infty}, \tau_i^\infty[$.

Furthermore, Proposition 12 implies that each λ_i^D is the unique solution to (A.9). For each μ where $\tilde{L}_i^\infty(\mu)$ is defined, it holds that

$$\tilde{L}_i(\mu) - \mu \xrightarrow{\text{a.s.}} \tilde{L}_i^\infty(\mu) - \mu, \tag{B.19}$$

by Proposition 16 in Appendix B.

As both $\tilde{L}_i(\mu)$ and $\tilde{L}_i^\infty(\mu)$ are non-increasing, the functions $\tilde{L}_i(\mu) - \mu$ and $\tilde{L}_i^\infty(\mu) - \mu$ are strictly decreasing. Hence, by (Lu and Li, 2020, Lemma A.1), it holds that

$$\lambda_i^D \xrightarrow{\text{a.s.}} \tilde{L}_i^\infty(\mu_i^*). \tag{B.20}$$

Substituting $\mu_i^* = \lambda_i^\infty(\alpha_i)$ in Equation (B.20) gives (4.3) for $i \in [j]$.

It remains to show the claim for the remaining $p - j$ eigenvalues. As (4.2) has only j solutions by assumption, it follows that

$$\zeta_\delta(\alpha) - \lambda_i^\infty(\alpha) = \tilde{L}_i^\infty(\lambda_i^\infty(\alpha)) - \lambda_i^\infty(\alpha) = 0$$

has no solutions for $\alpha > \tau$ and $i > j$. Denoting $\mu = \lambda_i^\infty(\alpha)$, it further holds that

$$\tilde{L}_i^\infty(\mu) - \mu = 0$$

has no solutions for $\mu \in]\lambda_i^{a^\infty}, t_i^\infty[$. Since

$$\lim_{\mu \to \lambda_i^{a^\infty}} \tilde{L}_i^\infty(\mu) - \mu = +\infty,$$

it must be that $\tilde{L}_i^\infty(\mu) - \mu > 0$ for all $\mu \in]\lambda_i^{a^\infty}, t_i^\infty[$. Using (B.19), we have that

$$\tilde{L}_i(\mu) - \mu > 0,$$

for all $\mu \in]\lambda_i^{a^\infty}, t_i^\infty[$ and n large enough. As $\lambda_i^a \xrightarrow{\text{a.s.}} \lambda_i^{a^\infty}$ and each $\tilde{L}_i(\mu)$ is defined on $]\lambda_i^a, +\infty[$, the solution to the equation

$$\tilde{L}_i(\mu) - \mu = 0$$

must be for $\mu > t_i^\infty$. Then, applying Proposition 16, for any fixed μ it holds that

$$\tilde{L}_i(\mu) \xrightarrow{\text{a.s.}} \zeta_\delta(\bar{\lambda}_\delta).$$

Lastly, as both $\mu - \tilde{L}_i(\mu)$ and $\mu - \zeta_\delta(\bar{\lambda}_\delta)$ are increasing functions, Lemma A.1 in Lu and Li (2020) implies that

$$\lambda_i^D \xrightarrow{\text{a.s.}} \zeta_\delta(\bar{\lambda}_\delta),$$

for all $i > j$, which proves the claim. ■

B.3. Number of solutions to Equation (4.2)

Proposition 17 *The equation in Equation* (4.2) *has at most* p *solutions. Furthermore, if Equation* (4.4) *holds, then Equation* (4.2) *has exactly* p *solutions.*

Proof As stated in the proof of Theorem 2, it holds that

$$\det(\zeta_\delta(\alpha)I - R^\infty(\alpha)) = \prod_{i=1}^{p}(\zeta_\delta(\alpha) - \lambda_i^\infty(\alpha)). \tag{B.21}$$

Note that the function $\zeta_\delta(\alpha) - \lambda_i^\infty(\alpha)$ is continous and strictly increasing for $\alpha \in]\tau, +\infty[$, so that

$$\lim_{\alpha\to\infty} \zeta_\delta(\alpha) - \lambda_i^\infty(\alpha) = +\infty.$$

Thus, the equation in Equation (4.2) has at most p solutions. Furthermore, the assumption in Equation (4.4) implies that

$$\inf_{\|x\|_2=1} \lim_{\alpha\to\tau^+} x^\top R^\infty(\alpha)x = +\infty,$$

which is equivalent to

$$\lim_{\alpha\to\tau^+} \lambda_i^\infty(\alpha) = +\infty.$$

As $\lim_{\alpha\to\tau^+} \zeta_\delta(\alpha) = \bar{\lambda}_\delta < +\infty$, it then holds that

$$\lim_{\alpha\to\tau^+} \zeta_\delta(\alpha) - \lambda_i^\infty(\alpha) = -\infty,$$

proving there must be exactly p solutions to Equation (4.2) due to the intermediate value theorem. ■

Appendix C. Proofs for the characterization of eigenvectors

C.1. Auxiliary results

Proposition 18 *Let* $k \in [p]$ *be such that* $\alpha_k > \bar{\lambda}_\delta$, *where* α_k *is the* k*-th largest solution of Equation* (4.2). *Then, it holds that*

$$R(\lambda_k^D) \xrightarrow{\text{a.s.}} R^\infty(\alpha_k), \tag{C.1}$$

and

$$\frac{d}{d\lambda}R(\lambda_k^D) \xrightarrow{\text{a.s.}} \frac{1}{\zeta_\delta'(\alpha_k)}\frac{d}{d\alpha}R^\infty(\alpha_k). \tag{C.2}$$

Proof By Lemma 14, it suffices to understand the behavior of the functions in Equation (5.5), as $n, d \to \infty$. However, we first need to verify the assumption that $\lambda_k^D > \lambda_1^P$ almost surely. From Theorem 2, it follows that

$$\lambda_k^D \xrightarrow{\text{a.s.}} \zeta_\delta(\alpha_k),$$

and $\zeta_\delta(\alpha_k) > \zeta_\delta(\bar{\lambda}_\delta)$ as $\alpha_k > \bar{\lambda}_\delta$ and ζ_δ is strictly increasing on $]\bar{\lambda}_\delta, +\infty[$. Furthermore, $\lambda_1^P \xrightarrow{\text{a.s.}} \zeta_\delta(\bar{\lambda}_\delta)$, hence $\lambda_k^D > \lambda_1^P$ almost surely.

Let us denote by G the function

$$G(\mu) = -\frac{1}{\mu},$$

which we will use in the continuation of the proof. Using Bai and Yao (2012) as in the proof of Proposition 16, we get that

$$\mathcal{L}_i(\mu) \xrightarrow{\text{a.s.}} \zeta_\delta \circ Q_i{}^{-1} \circ G(\mu),$$

where $Q_i(\alpha) \coloneqq \mathbb{E}\left[\frac{s_i^2 z^2}{z-\alpha}\right]$. Notice that $Q_i(\alpha)$ is invertible by (Lu and Li, 2020, Remark 3.3), which is stated for the analogous function Q. In the same manner, it holds that

$$\mathcal{L}_{i,j}(\mu) \xrightarrow{\text{a.s.}} \zeta_\delta \circ Q_{i,j}{}^{-1} \circ G(\mu),$$

where $Q_{i,j}(\alpha) \coloneqq \mathbb{E}\left[\frac{(s_i+s_j)^2 z^2}{z-\alpha}\right]$. As $\alpha_k > \bar{\lambda}_\delta$, we have that ζ_δ is strictly increasing and invertible, hence

$$\mathcal{L}_i^{-1}(\lambda_k^D) \xrightarrow{\text{a.s.}} G \circ Q_i \circ \zeta_\delta^{-1} \circ \zeta_\delta(\alpha_k) = G \circ Q_i(\alpha_k),$$

which follows from (Lu and Li, 2020, Lemma A.1). Plugging this into (A.12) we get

$$R(\lambda_k^D)_{i,i} \xrightarrow{\text{a.s.}} a_{i,i}^\infty - Q_i(\alpha_k). \tag{C.3}$$

Note that

$$a_{i,i}^\infty - Q_i(\alpha_k) = \mathbb{E}\big[s_i^2 z\big] - \mathbb{E}\left[\frac{s_i^2 z^2}{z-\alpha_k}\right] = \mathbb{E}\left[\frac{\alpha_k s_i^2 z}{\alpha_k - z}\right] = R^\infty(\alpha_k)_{i,i}.$$

Similarly, it holds that

$$R(\lambda_k^D)_{i,j} \xrightarrow{\text{a.s.}} a_{i,j}^\infty - Q_{i,j}(\alpha_k),$$

which combined with (C.3) proves (C.1).

Moreover, we have that $\mathcal{L}_i(\mu)$ is differentiable (see Lemma 14), so for its derivative it holds that

$$\mathcal{L}_i'(\mu) \xrightarrow{\text{a.s.}} \zeta_\delta' \circ Q_i{}^{-1} \circ G(\mu) \cdot \big(Q_i^{-1}\big)' \circ G(\mu) \cdot G'(\mu),$$

which follows from (Lu and Li, 2020, Lemma A.2). Plugging this into (A.13) we get

$$\frac{d}{d\lambda} R(\lambda_k^D)_{i,i} \xrightarrow{\text{a.s.}} \frac{\frac{d}{d\alpha}(R^\infty(\alpha_k)_{i,i})}{\zeta_\delta'(\alpha_k)}.$$

Similarly, it holds that

$$\frac{d}{d\lambda} R(\lambda_k^D)_{i,j} \xrightarrow{\text{a.s.}} \frac{\frac{d}{d\alpha}(R^\infty(\alpha_k)_{i,j})}{\zeta_\delta'(\alpha_k)}.$$

Combining the last two equations we obtain (C.2). ■

C.2. Proof of Theorem 3

Proof We start by proving Equation (4.5). Let $v_i^D = \begin{bmatrix} h_i \\ g_i \end{bmatrix}$, for $i \in \{k, \dots, k+m-1\}$. Since $\alpha_k > \bar{\lambda}_\delta$, the conditions of Proposition 13 are satisfied as in the proof of Proposition 18. Thus, it holds that

$$h_i = \frac{\tilde{h}_i}{\sqrt{1 - \tilde{h}_i^\top \frac{d}{d\lambda} R(\lambda_i^D) \tilde{h}_i}}, \tag{C.4}$$

where $\tilde{h}_i = \frac{h_i}{\|h_i\|_2}$ is the unit norm eigenvector of $R(\lambda_i^D)$. Note that the vectors $\tilde{h}_i$ are orthogonal. Furthermore, Proposition 18 gives that

$$R(\lambda_k^D) \xrightarrow{\text{a.s.}} R^\infty(\alpha_k), \qquad \frac{d}{d\lambda} R(\lambda_k^D) \xrightarrow{\text{a.s.}} \frac{1}{\zeta_\delta'(\alpha_k)} \frac{d}{d\alpha} R^\infty(\alpha_k). \tag{C.5}$$

Then, applying the results from (Kato, 1995, II.1.4), it holds that the orthonormal projection to the eigenspace corresponding to the k-th eigenvalue also converges, that is

$$\Pi_{E_k} \xrightarrow{\text{a.s.}} \Pi_{E_k^\infty}, \tag{C.6}$$

where E_k is the space spanned by the eigenvectors $h_k, \dots, h_{k+m-1}$ and E_k^∞ is the eigenspace of the limiting matrix $R^\infty(\alpha_k)$, corresponding to the eigenvalue $\zeta_\delta(\alpha_k)$ of multiplicity m. Due to orthonormality of $\tilde{h}_i$, we can write the orthonormal projection more explicitly as

$$\Pi_{E_k} = \sum_{i=k}^{k+m-1} \frac{h_i h_i^\top}{\|h_i\|_2^2} = \sum_{i=k}^{k+m-1} \tilde{h}_i \tilde{h}_i^\top,$$

and

$$\Pi_{E_k^\infty} = \sum_{i=k}^{k+m-1} \frac{h_i^\infty h_i^{\infty\top}}{\|h_i^\infty\|_2^2} = \sum_{i=k}^{k+m-1} h_i^\infty h_i^{\infty\top}, \tag{C.7}$$

where $h_k^\infty \dots, h_{k+j-1}^\infty$ is any choice of the orthonormal eigenbasis of E_k^∞. From (C.4), it follows that

$$\sum_{i=k}^{k+m-1} \frac{1}{\|h_i\|_2^2} = m - \sum_{i=k}^{k+m-1} \tilde{h}_i^\top \frac{d}{d\lambda} R(\lambda_i^D) \tilde{h}_i \geq m - m \cdot \lambda_p \left(\frac{d}{d\lambda} R(\lambda_i^D) \right).$$

Moreover, due to (C.5) and the continuity of eigenvalues, the RHS has a convergent limit

$$m - m \cdot \lambda_p \left(\frac{d}{d\lambda} R(\lambda_i^D) \right) \xrightarrow{\text{a.s.}} m - m \frac{1}{\zeta_\delta'(\alpha_k)} \lambda_p \left(\frac{d}{d\alpha} R^\infty(\alpha_k) \right). \tag{C.8}$$

Note that the matrix $\frac{d}{d\alpha} R^\infty(\alpha_k) = -\mathbb{E}\left[\frac{ss^\top z^2}{(\alpha_k - z)^2}\right]$ is strictly negative definite for $\alpha_k > \tau$, which implies that $\lambda_p(\frac{d}{d\alpha} R^\infty(\alpha_k)) < 0$. As $\alpha_k > \bar{\lambda}_\delta$, it holds that $\zeta_\delta'(\alpha_k) > 0$ and the RHS of Equation (C.8) is finite. This further implies that, for each i s.t. $k \leq i \leq k+m-1$, it must hold

$$\liminf_{d \to \infty} \|h_i\|_2 > 0. \tag{C.9}$$

Note that

$$\begin{aligned}\sum_{i=k}^{k+m-1}\left|\left\langle v_i^D, e_l^{(d)}\right\rangle\right|^2 = \sum_{i=k}^{k+m-1}\left|\left\langle h_i, e_l^{(p)}\right\rangle\right|^2 &= \sum_{i=k}^{k+m-1}\|h_i\|_2^2\left|\left\langle \tilde{h}_i, e_l^{(p)}\right\rangle\right|^2 \\ &\geq \min_{t\in\{k,\dots,k+m-1\}}\|h_t\|_2^2\sum_{i=k}^{k+m-1}\left|\left\langle \tilde{h}_i, e_l^{(p)}\right\rangle\right|^2 \\ &= \min_{t\in\{k,\dots,k+m-1\}}\|h_t\|_2^2\cdot e_l^{(p)\top}\Pi_{E_k}e_l^{(p)}.\end{aligned} \tag{C.10}$$

Let us pick $e_l^{(d)}$ such that $\Pi_{E_k^\infty}(e_l^{(p)}) \neq 0$. Then, (C.5) implies that

$$\Pi_{E_k}(e_l^{(p)}) \neq 0, \tag{C.11}$$

for all d large enough. Finally, combining (C.9), (C.10) and (C.11) proves

$$\liminf_{d\to\infty}\sum_{i=k}^{k+m-1}\left|\left\langle v_i^D, e_l^{(d)}\right\rangle\right|^2 > 0,$$

which gives the claim in (4.5).

Next, we prove Equation (4.6) for $m > 1$. We assume, as in the statement, that E_k^∞ is also the eigenspace corresponding to the k-th eigenvalue of $R^\infty(\alpha+\Delta)$ for any small enough Δ. For arbitrary eigenvectors h_{i_1} and h_{i_2} from E_k^∞, it holds that

$$\begin{aligned}h_{i_1}^{\infty\top}\frac{d}{d\alpha}R^\infty(\alpha_k)h_{i_1}^\infty &= \lim_{\Delta\to 0}\frac{h_{i_1}^{\infty\top}R^\infty(\alpha_k+\Delta)h_{i_1}^\infty - h_{i_1}^{\infty\top}R^\infty(\alpha_k)h_{i_1}^\infty}{\Delta} \\ &= \lim_{\Delta\to 0}\frac{h_{i_2}^{\infty\top}R^\infty(\alpha_k+\Delta)h_{i_2}^\infty - h_{i_2}^{\infty\top}R^\infty(\alpha_k)h_{i_2}^\infty}{\Delta} \\ &= h_{i_2}^{\infty\top}\frac{d}{d\alpha}R^\infty(\alpha_k)h_{i_2}^\infty,\end{aligned} \tag{C.12}$$

since $h_{i_1}^{\infty\top}R^\infty(\alpha_k+\Delta)h_{i_1}^\infty = h_{i_2}^{\infty\top}R^\infty(\alpha_k+\Delta)h_{i_2}^\infty$ for any small enough Δ. Note that, for any ε and large enough d, it holds that

$$\left\|\Pi_{E_k} - \Pi_{E_k^\infty}\right\|_2 < \varepsilon. \tag{C.13}$$

due to (C.6). Let us now fix $\tilde{h}_i$, for some $i \in \{k,\dots,k+m-1\}$. As we can choose any orthonormal basis when writing out $\Pi_{E_k^\infty}$ in (C.7), let us choose one such that $h_i^\infty = \frac{\Pi_{E_k^\infty}(\tilde{h}_i)}{\left\|\Pi_{E_k^\infty}(\tilde{h}_i)\right\|_2}$. Then, (C.13) implies that

$$\left\|\sum_{i=k}^{k+m-1}\tilde{h}_i\tilde{h}_i^\top - \sum_{i=k}^{k+m-1}h_i^\infty h_i^{\infty\top}\right\|_2 < \varepsilon.$$

From the orthonormality of the chosen eigenbasis, it holds that

$$\begin{aligned}\left\|\tilde{h}_i - \Pi_{E_k^\infty}\tilde{h}_i\right\|_2 &= \left\|(\Pi_{E_k} - \Pi_{E_k^\infty})(\tilde{h}_i)\right\|_2 \\ &\leq \left\|\Pi_{E_k} - \Pi_{E_k^\infty}\right\|_2\left\|\tilde{h}_i\right\|_2 \\ &< \varepsilon.\end{aligned}$$

This also implies $1+\varepsilon > \left\|\Pi_{E_k^\infty}\tilde{h}_i\right\|_2 \geq 1-\varepsilon$, hence

$$\begin{aligned}\left\|\tilde{h}_i - h_i^\infty\right\|_2 &= \left\|\tilde{h}_i - \frac{\Pi_{E_k^\infty}(\tilde{h}_i)}{\left\|\Pi_{E_k^\infty}(\tilde{h}_i)\right\|_2}\right\|_2 \\ &= \left\|\frac{\tilde{h}_i - \Pi_{E_k^\infty}(\tilde{h}_i)}{\left\|\Pi_{E_k^\infty}(\tilde{h}_i)\right\|_2} + \tilde{h}_i \frac{1-\left\|\Pi_{E_k^\infty}(\tilde{h}_i)\right\|_2}{\left\|\Pi_{E_k^\infty}(\tilde{h}_i)\right\|_2}\right\|_2 \\ &\leq \frac{\varepsilon}{1-\varepsilon} + \frac{\varepsilon}{1-\varepsilon} < 4\varepsilon.\end{aligned}$$

Thus, (C.5) implies that, for large enough d,

$$\left\|\tilde{h}_i^\top \frac{d}{d\lambda} R(\lambda_k^D)\tilde{h}_i - \frac{1}{\zeta_\delta'(\alpha_k)} h_i^{\infty\top} \frac{d}{d\alpha} R^\infty(\alpha_k) h_i^\infty\right\|_2 < c \cdot \varepsilon,$$

for some constant c independent of ε. Plugging in the expression (C.12) makes h_i^∞ not depend on $\tilde{h}_i$ anymore, resulting in

$$\tilde{h}_i^\top \frac{d}{d\lambda} R(\lambda_k^D)\tilde{h}_i \xrightarrow{\text{a.s.}} \frac{1}{\zeta_\delta'(\alpha_k)} h_l^{\infty\top} \frac{d}{d\alpha} R^\infty(\alpha_k) h_l^\infty, \tag{C.14}$$

for an arbitrary unit norm eigenvector $h_l^\infty \in E_k^\infty$. Finally, combining (C.4) and (C.14) with (C.6) implies

$$\sum_{i=k}^{k+m-1} \tilde{h}_i\tilde{h}_i^\top \xrightarrow{\text{a.s.}} \frac{\zeta_\delta'(\alpha_k)\sum_{i=k}^{k+m-1} h_i^\infty h_i^{\infty\top}}{\zeta_\delta'(\alpha_k) + h_k^{\infty\top}\frac{d}{d\alpha}R^\infty(\alpha_k)h_k^\infty},$$

which proves Equation (4.6).

Finally, we prove the converse statement in Equation (4.7). Let us assume that $\zeta_\delta(\bar{\lambda}_\delta) \neq \lambda_l^{a^\infty}$ for all l. This is without loss of generality, as justified at the end of the argument. By assumption, we have

$$\lambda_i^D \xrightarrow{\text{a.s.}} \zeta_\delta(\bar{\lambda}_\delta). \tag{C.15}$$

Thus, for n large enough, it follows that $\lambda_i^D \notin \Lambda^a$, since $\lambda_l^a \xrightarrow{\text{a.s.}} \lambda_l^{a^\infty}$. We denote by $k \in [i]$ the index such that $\lambda_k^a < \lambda_i^D < \lambda_{k-1}^a$, with the abuse of notation that $\lambda_0^a = +\infty$. By Proposition 12, it holds that

$$\lambda_i^D = L_k(P - q(a - \lambda_i^D I_p)^{-1}q^\top). \tag{C.16}$$

Recall that the eigenvector condition for v_i^D is

$$Dv_i^D = \begin{bmatrix} a & q^\top \\ q & P \end{bmatrix}\begin{bmatrix} h_i \\ g_i \end{bmatrix} = \lambda_i^D \begin{bmatrix} h_i \\ g_i \end{bmatrix}.$$

Splitting this equation into p and $d-p$ coordinates gives

$$ah_i + q^\top g_i = \lambda_i^D h_i, \tag{C.17}$$

$$qh_i + Pg_i = \lambda_i^D g_i. \tag{C.18}$$

Solving (C.17) gives

$$h_i = -(a - \lambda_i^D I_p)^{-1} q^\top g_i,$$

where $(a - \lambda_i^D I_p)$ is invertible as $\lambda_i^D \notin \Lambda^a$. Substituting in (C.18) yields

$$P g_i - q(a - \lambda_i^D I_{d-p})^{-1} q^\top g_i = \lambda_i^D g_i.$$

Let us denote by $\tilde{g}_i = \frac{g_i}{\|g_i\|_2}$ the unit norm eigenvector of $P - q(a - \lambda_i^D I_p)^{-1} q^\top$ corresponding to the eigenvalue λ_i^D, and also define $\tilde{h}_i := -(a - \lambda_i^D I_p)^{-1} q^\top \tilde{g}_i$. By (C.16), it holds that $\tilde{g}_i$ is the eigenvector corresponding to the k-th eigenvalue of the matrix $P - q(a - \lambda_i^D I_p)^{-1} q^\top$. Moreover, $\tilde{h}_i$ and $\tilde{g}_i$ satisfy equations (C.17) and (C.18), so $\tilde{v}_i^D = \begin{bmatrix} \tilde{h}_i \\ \tilde{g}_i \end{bmatrix}$ is aligned with an eigenvector corresponding to eigenvalue λ_i^D. However, $\tilde{v}_i^D$ does not necessarily have unit norm. It holds that

$$\begin{bmatrix} h_i \\ g_i \end{bmatrix} = v_i^D = \frac{\tilde{v}_i^D}{\|\tilde{v}_i^D\|_2} = \frac{\begin{bmatrix} \tilde{h}_i \\ \tilde{g}_i \end{bmatrix}}{\sqrt{\tilde{g}_i^\top \tilde{g}_i + \tilde{h}_i^\top \tilde{h}_i}} = \frac{\begin{bmatrix} \tilde{h}_i \\ \tilde{g}_i \end{bmatrix}}{\sqrt{1 + \tilde{g}_i^\top q(a - \lambda_i^D I_p)^{-2} q^\top \tilde{g}_i}},$$

from which follows that

$$h_i = \frac{\tilde{h}_i}{\sqrt{1 + \tilde{g}_i^\top q(a - \lambda_i^D I_p)^{-2} q^\top \tilde{g}_i}}.$$

Moreover, notice that

$$\|h_i\|_2^2 = \frac{\tilde{g}_i^\top q(a - \lambda_i^D I_p)^{-2} q^\top \tilde{g}_i}{1 + \tilde{g}_i^\top q(a - \lambda_i^D I_p)^{-2} q^\top \tilde{g}_i} = 1 - \frac{1}{1 + \tilde{g}_i^\top q(a - \lambda_i^D I_p)^{-2} q^\top \tilde{g}_i}. \tag{C.19}$$

The term at the denominator can be simplified as

$$\begin{aligned} \tilde{g}_i^\top q(a - \lambda_i^D I_p)^{-2} q^\top \tilde{g}_i &= \tilde{g}_i^\top \left(\frac{d}{d\lambda} q(a - \lambda I_p)^{-1} q^\top \right) (\lambda_i^D) \tilde{g}_i \\ &= \left(\frac{d}{d\lambda} \lambda_k (q(a - \lambda I_p)^{-1} q^\top) \right) (\lambda_i^D) \\ &= \frac{d}{d\lambda} L_k(\lambda_i^D). \end{aligned} \tag{C.20}$$

Here, when the eigenvalue λ_i^D is simple, the second equality follows from (Lancaster, 1964, Theorem 5) and the fact that $\tilde{g}_i$ is the eigenvector corresponding to the k-th eigenvalue of the matrix $q(a - \lambda_i^D I_p)^{-1} q^\top$. The case in which the eigenvalue λ_i^D has multiplicity $m > 1$ is handled similarly via (Lancaster, 1964, Theorem 7), and it is discussed at the end of the argument.

Note that, by definition, it holds

$$\frac{d}{d\lambda} L_k(\lambda_i^D) = \frac{d}{d\lambda} \tilde{L}_i(\lambda_i^D).$$

We will prove that

$$\frac{d}{d\lambda} \tilde{L}_i(\lambda_i^D) \xrightarrow{\text{a.s.}} 0, \tag{C.21}$$

which implies $\|h_i\|_2^2 \xrightarrow{\text{a.s.}} 0$ and hence gives the desired statement in Equation (4.7). Recall that the assumption underlying Equation (4.7) is that Equation (C.15) holds. This happens if either the α_i associated to λ_i^D is s.t. $\alpha_i \le \bar{\lambda}_\delta$ or i exceeds the number of solutions of Equation (4.2) (i.e., λ_i^D is not associated to any solution of Equation (4.2)). We handle these two cases separately.

Case 1: $\lambda \in]\lambda_i^{a^\infty}, t_i^\infty[$. This corresponds to the case in which $\alpha_i \le \bar{\lambda}_\delta$. Applying Proposition 16, we have

$$\tilde{L}_i(\lambda) \xrightarrow{\text{a.s.}} \tilde{L}_i^\infty(\lambda) = \zeta_\delta((\lambda_i^\infty)^{-1}(\lambda)).$$

By combining Equation (C.15) and Lemma A.2 in Lu and Li (2020), it follows that

$$\frac{d}{d\lambda}\tilde{L}_i(\lambda_i^D) \xrightarrow{\text{a.s.}} \zeta_\delta'((\lambda_i^\infty)^{-1}(\zeta_\delta(\bar{\lambda}_\delta))) \cdot \frac{1}{\lambda_i'^\infty((\lambda_i^\infty)^{-1}(\zeta_\delta(\bar{\lambda}_\delta)))},$$

where the denominator is non-zero, as λ_i^∞ is strictly decreasing. As $(\lambda_i^\infty)^{-1}(\zeta_\delta(\bar{\lambda}_\delta)) = \alpha_i \le \bar{\lambda}_\delta$, the derivative ζ_δ' computed at that point is 0, which gives Equation (C.21).

Case 2: $\lambda \in]t_i^\infty, +\infty[$. This corresponds to the case in which λ_i^D is not associated to any α_i. Applying again Proposition 16, we have

$$\tilde{L}_i(\lambda) \xrightarrow{\text{a.s.}} \zeta_\delta(\bar{\lambda}_\delta).$$

Then, again by (Lu and Li, 2020, Lemma A.2) it follows that

$$\frac{d}{d\lambda}\tilde{L}_i(\lambda) \xrightarrow{\text{a.s.}} \frac{d}{d\lambda}(\zeta_\delta(\bar{\lambda}_\delta)) = 0,$$

as the LHS is a constant function that does not depend on λ.

We conclude by handling the two special cases mentioned during the argument above. If λ_i^D has a multiplicity $m > 1$, then the m derivatives $\frac{d}{d\lambda}L_k(\lambda_i^D)$ corresponding to the m eigenvalues vanish using the same argument. An application of (Lancaster, 1964, Theorem 7) gives that such derivatives are the eigenvalues of the $m \times m$ matrix obtained by replacing $\tilde{g}_i$ in the LHS of Equation (C.20) with a basis of the m-dimensional eigenspace associated to λ_i^D. As the eigenvalues of such matrix vanish, so does its trace and therefore the m norms $\|h_i\|_2$.

Finally, if $\zeta_\delta(\bar{\lambda}_\delta) = \lambda_l^{a^\infty}$ for some l, we can just add a vanishing perturbation to the matrix a such that $\zeta_\delta(\bar{\lambda}_\delta) \ne \lambda_l^{a^\infty}$. Applying the proved result, the overlaps vanish for any such small perturbation, and due to continuity, taking the perturbation to 0 would give the claim. ■

Appendix D. Invariance of the eigenspace for permutation-invariant link functions

Proposition 19 *If the link function q is permutation invariant in m coordinates, then the matrix $R^\infty(\alpha)$ has eigenspaces that do not change with α, of combined dimension $m-1$.*

Proof Let us denote by w_i the i-th column of the matrix $\widetilde{W}^*$ as in (3.3), representing the top-p entries of the reparametrized signal. Without loss of generality, we can assume that q is permutation invariant in the first m coordinates, i.e.,

$$q(t_1, \dots, t_m, t_{m+1} \dots, t_p, \varepsilon) = q(t_{\pi(1)}, \dots, t_{\pi(m)}, t_{m+1} \dots, t_p, \varepsilon),$$

for any permutation $\pi : [m] \to [m]$. Let E be the span of $\{w_i - w_{i+1} \ : \ i \in [m-1]\}$. Note that E has dimension $m-1$, due to the linear independence of the signals w_i. We will prove that E is a direct sum of eigenspaces of $R^\infty(\alpha)$, neither of which depends on α.

Let us define u_i as the image of $w_i - w_{i+1}$ under $R^\infty(\alpha)$, that is

$$u_i := R^\infty(\alpha)(w_i - w_{i+1}) = \alpha\, \mathbb{E}\left[\frac{s\langle s, w_i - w_{i+1}\rangle \mathcal{T}(y)}{\alpha - \mathcal{T}(y)}\right],$$

for $i \in [m-1]$. Let us denote by x_i the component of u_i that is orthogonal to w_i and w_{i+1}, i.e., $x_i := \Pi_{\{w_i, w_{i+1}\}^\perp} u_i$. Then, it holds that

$$u_i = a_1 w_i + a_2 w_{i+1} + x_i, \tag{D.1}$$

for some coefficients a_1 and a_2. We will first prove that $x_i = 0$. Towards that end, let $\mathcal{S}_i : \mathbb{R}^p \to \mathbb{R}^p$ be an isometric reflection that sends w_i to w_{i+1}, w_{i+1} to w_i, and keeps w_j fixed for $j \notin \{i, i+1\}$. Such a reflection exists due to the assumed linear independence of the w_i's. Notice that the normal distribution in $\mathbb{R}^p$ is invariant to the transformation $\mathcal{S}_i$. Thus, it follows that

$$\begin{aligned}
\langle u_i, x_i\rangle &= \alpha\, \mathbb{E}\left[\frac{\langle s, x_i\rangle\langle s, w_i - w_{i+1}\rangle \mathcal{T}(y)}{\alpha - \mathcal{T}(y)}\right] \\
&= \alpha\, \mathbb{E}\left[\frac{\langle s, x_i\rangle\langle s, w_i\rangle \mathcal{T}(y)}{\alpha - \mathcal{T}(y)}\right] - \alpha\, \mathbb{E}\left[\frac{\langle s, x_i\rangle\langle s, w_{i+1}\rangle \mathcal{T}(y)}{\alpha - \mathcal{T}(y)}\right] \\
&= \alpha\, \mathbb{E}\left[\frac{\langle \mathcal{S}_i s, x_i\rangle\langle \mathcal{S}_i s, w_i\rangle \mathcal{T}(q((\widetilde{W}^*)^\top \mathcal{S}_i s, \varepsilon))}{\alpha - \mathcal{T}(q((\widetilde{W}^*)^\top \mathcal{S}_i s, \varepsilon))}\right] - \alpha\, \mathbb{E}\left[\frac{\langle s, x_i\rangle\langle s, w_{i+1}\rangle \mathcal{T}(y)}{\alpha - \mathcal{T}(y)}\right] \\
&= \alpha\, \mathbb{E}\left[\frac{\langle s, x_i\rangle\langle s, w_{i+1}\rangle \mathcal{T}(q((\widetilde{W}^*)^\top s, \varepsilon))}{\alpha - \mathcal{T}(q((\widetilde{W}^*)^\top s, \varepsilon))}\right] - \alpha\, \mathbb{E}\left[\frac{\langle s, x_i\rangle\langle s, w_{i+1}\rangle \mathcal{T}(y)}{\alpha - \mathcal{T}(y)}\right] \\
&= 0,
\end{aligned}$$

due to the permutation invariance of q, the fact that $\mathcal{S}_i x_i = x_i$ and $\mathcal{S}_i w_i = w_{i+1}$. Therefore, it must be that $x_i = 0$. Moreover,

$$\begin{aligned}
\langle u_i, w_i\rangle &= \alpha\, \mathbb{E}\left[\frac{\langle s, w_i\rangle\langle s, w_i - w_{i+1}\rangle \mathcal{T}(y)}{\alpha - \mathcal{T}(y)}\right] \\
&= \alpha\, \mathbb{E}\left[\frac{\langle \mathcal{S}_i s, w_i\rangle\langle \mathcal{S}_i s, w_i - w_{i+1}\rangle \mathcal{T}(q((\widetilde{W}^*)^\top \mathcal{S}_i s)}{\alpha - \mathcal{T}(q((\widetilde{W}^*)^\top \mathcal{S}_i s)}\right] \\
&= \alpha\, \mathbb{E}\left[\frac{\langle s, w_{i+1}\rangle\langle s, w_{i+1} - w_i\rangle \mathcal{T}(q((\widetilde{W}^*)^\top s)}{\alpha - \mathcal{T}(q((\widetilde{W}^*)^\top s)}\right] \\
&= \alpha\, \mathbb{E}\left[\frac{\langle s, w_{i+1}\rangle\langle s, w_{i+1} - w_i\rangle \mathcal{T}(y)}{\alpha - \mathcal{T}(y)}\right] \\
&= -\langle u_i, w_{i+1}\rangle.
\end{aligned}$$

Combining this with the decomposition in (D.1) implies that $a_1 + a_2\langle w_i, w_{i+1}\rangle = -a_1\langle w_{i+1}, w_i\rangle - a_2$. As by assumption the signals w_i and w_{i+1} are linearly independent, it cannot be that $\langle w_i, w_{i+1}\rangle =$

-1, so it must be that $a_1 = -a_2$. This proves that $w_i - w_{i+1}$ are indeed eigenvectors for every α. By definition, it holds that

$$E = \bigoplus_{i=1}^{m-1} \mathrm{span}\{u_i\}.$$

Thus, it is left to prove that any pair of eigenvectors u_i and u_j, for $i, j \in [m-1]$, either have the same eigenvalue for all α, or for no α. We denote by $\lambda_{u_i}(\alpha)$ the eigenvalue that corresponds to the eigenvector u_i. Similar to before, let $\mathcal{S}^{(i)} : \mathbb{R}^p \to \mathbb{R}^p$ be an isometric reflection that sends w_i to w_{i+2}, w_{i+2} to w_i, and keeps w_j fixed for $j \notin \{i, i+2\}$. Such a reflection exists due to the assumed linear independence of the w_i's. Also, the normal distribution is invariant to the transformation $\mathcal{S}^{(i)}$. Then, it follows that

$$\begin{aligned}
\lambda_{u_i}(\alpha) \cdot \|w_i - w_{i+1}\|_2^2 = \langle u_i, w_i - w_{i+1} \rangle &= \alpha\, \mathbb{E}\left[\frac{\langle s, w_i - w_{i+1} \rangle^2 \mathcal{T}(y)}{\alpha - \mathcal{T}(y)} \right] \\
&= \alpha\, \mathbb{E}\left[\frac{\left\langle \mathcal{S}^{(i)} s, w_i - w_{i+1} \right\rangle^2 \mathcal{T}(q((\widetilde{W}^*)^\top \mathcal{S}^{(i)} s, \varepsilon))}{\alpha - \mathcal{T}(q((\widetilde{W}^*)^\top \mathcal{S}^{(i)} s, \varepsilon))} \right] \\
&= \alpha\, \mathbb{E}\left[\frac{\langle s, w_{i+2} - w_{i+1} \rangle^2 \mathcal{T}(q((\widetilde{W}^*)^\top s, \varepsilon))}{\alpha - \mathcal{T}(q((\widetilde{W}^*)^\top s, \varepsilon))} \right] \\
&= \langle u_{i+1}, w_{i+1} - w_{i+2} \rangle = \lambda_{u_{i+1}}(\alpha) \cdot \|w_{i+2} - w_{i+1}\|_2^2.
\end{aligned}$$

This proves that $\lambda_{u_i}(\alpha)/\lambda_{u_{i+1}}(\alpha)$ does not depend on α and, hence, $\lambda_{u_i}(\alpha)/\lambda_{u_j}(\alpha)$ does not depend on α for all $i, j \in [m-1]$, implying the existence of an eigenspace of dimension $m-1$ that does not change with α. ■

Appendix E. Proof of Theorem 4

Proof By Theorem 3, we have that

$$\max_{j \in [p]} \lim_{d \to \infty} \sum_{i=1}^{p} \left| \left\langle v_i^D, e_l^{(d)} \right\rangle \right|^2 > 0$$

holds if and only if there exists $\alpha > \bar{\lambda}_\delta$ that solves Equation (4.2).

Note that $\zeta_\delta(\alpha)$ is a strictly monotone function for $\alpha > \bar{\lambda}_\delta$, and it is constant for $\alpha \le \bar{\lambda}_\delta$. Thus, the existence of α solving Equation (4.2) s.t. $\alpha > \bar{\lambda}_\delta$ is equivalent to $\zeta'_\delta(\alpha_1) > 0$, where α_1 is the largest solution of (4.2). Thus, α_1 is the largest solution to

$$\det(\zeta_\delta(\alpha) I_p - R^\infty(\alpha)) = 0,$$

or equivalently

$$\lambda_1(R^\infty(\alpha)) = \zeta_\delta(\alpha).$$

This means that, for (4.2) to have solutions larger than $\bar{\lambda}_\delta$, there has to exist $\alpha_1 > \tau$ such that

$$\max_{\|u\|_2 = 1} u^\top R^\infty(\alpha_1) u = \zeta_\delta(\alpha_1),$$

$$\zeta'_\delta(\alpha_1) > 0,$$

or equivalently

$$\max_{\|u\|_2=1} u^\top R^\infty(\bar{\lambda}_\delta)u > \zeta_\delta(\bar{\lambda}_\delta). \tag{E.1}$$

This follows from the fact that $\zeta_\delta(\alpha)$ is strictly increasing for $\alpha > \bar{\lambda}_\delta$, and $u^\top R^\infty(\alpha)u$ is strictly decreasing, as proved in (B.8). Recall that $\bar{\lambda}_\delta$ is defined as the unique point that satisfies $\psi'_\delta(\bar{\lambda}_\delta) = 0$. This is equivalent to

$$\mathbb{E}\left[\frac{z^2}{(\bar{\lambda}_\delta - z)^2}\right] = \frac{1}{\delta}. \tag{E.2}$$

By definition, it holds that

$$R^\infty(\bar{\lambda}_\delta) = \mathbb{E}\left[\frac{\bar{\lambda}_\delta z}{\bar{\lambda}_\delta - z} s^\top s\right].$$

Therefore, the condition in (E.1) becomes

$$\max_{\|u\|_2=1} \mathbb{E}\left[\frac{\bar{\lambda}_\delta z}{\bar{\lambda}_\delta - z}\langle s, u\rangle^2\right] > \bar{\lambda}_\delta\left(\frac{1}{\delta} + \mathbb{E}\left[\frac{z}{\bar{\lambda}_\delta - z}\right]\right). \tag{E.3}$$

Note that D_n and $D'_n := D_n/\beta$ have the same principal eigenvector for an arbitrary scalar $\beta > 0$ so, without loss of generality, we can take $\bar{\lambda}_\delta = 1$. This transforms (E.3) and (E.2) into

$$\max_{\|u\|_2=1} \mathbb{E}\left[\frac{z(\langle s, u\rangle^2 - 1)}{1 - z}\right] > \frac{1}{\delta},$$
$$\mathbb{E}\left[\frac{z^2}{(1 - z)^2}\right] = \frac{1}{\delta}.$$

We turn our attention to finding the critical threshold δ_c such that no preprocessing function $\mathcal{T}$ exists which would satisfy these equations. To start, plugging in the definition of the expectation we get

$$\begin{aligned} \max_{\|u\|_2=1} \int_{\mathbb{R}} \frac{\mathcal{T}(y)}{1 - \mathcal{T}(y)} \mathbb{E}_s\Big[p(y \mid s) \cdot (\langle s, u\rangle^2 - 1)\Big] dy > \frac{1}{\delta}, \\ \int_{\mathbb{R}} \left(\frac{\mathcal{T}(y)}{1 - \mathcal{T}(y)}\right)^2 \mathbb{E}_s[p(y \mid s)] dy = \frac{1}{\delta}. \end{aligned} \tag{E.4}$$

Let us denote by $f(y) := \frac{\mathcal{T}(y)}{1-\mathcal{T}(y)}$. Note that

$$\begin{aligned} \frac{1}{\delta} &< \max_{\|u\|_2=1} \int_{\mathbb{R}} f(y) \mathbb{E}_s\Big[p(y \mid s) \cdot (\langle s, u\rangle^2 - 1)\Big] dy \\ &= \max_{\|u\|_2=1} \int_{\mathbb{R}} f(y) \sqrt{\mathbb{E}_s[p(y \mid s)]} \frac{\mathbb{E}_s\Big[p(y \mid s) \cdot (\langle s, u\rangle^2 - 1)\Big]}{\sqrt{\mathbb{E}_s[p(y \mid s)]}} dy \\ &\le \max_{\|u\|_2=1} \sqrt{\int_{\mathbb{R}} f^2(y) \mathbb{E}_s[p(y \mid s)] dy} \sqrt{\int_{\mathbb{R}} \frac{\Big(\mathbb{E}_s\Big[p(y \mid s) \cdot (\langle s, u\rangle^2 - 1)\Big]\Big)^2}{\mathbb{E}_s[p(y \mid s)]} dy} \\ &= \max_{\|u\|_2=1} \frac{1}{\sqrt{\delta}} \sqrt{\int_{\mathbb{R}} \frac{\Big(\mathbb{E}_s\Big[p(y \mid s) \cdot (\langle s, u\rangle^2 - 1)\Big]\Big)^2}{\mathbb{E}_s[p(y \mid s)]} dy}, \end{aligned} \tag{E.5}$$

where the third line follows from Hölder's inequality and the fourth line from the second condition in (E.4). This means that, regardless of which preprocessing function $\mathcal{T}$ was chosen, if it satisfies (E.4), then it must hold that

$$\frac{1}{\delta} < \max_{\|u\|_2=1} \int_{\mathbb{R}} \frac{\left(\mathbb{E}_s\left[p(y \mid s) \cdot (\langle s, u \rangle^2 - 1)\right]\right)^2}{\mathbb{E}_s[p(y \mid s)]} dy =: \frac{1}{\delta_t}.$$

This directly implies that

$$\delta_c \geq \delta_t. \tag{E.6}$$

Now, let us prove that, for any $\delta > \delta_t$, there exists a preprocessing function that achieves weak recovery. Towards this end, we turn our attention to when equality holds in (E.5), as this gives us the preprocessing function that exactly matches the upper bound. Namely, this is true if and only if almost everywhere

$$f^2(y) \mathop{\mathbb{E}}_s[p(y \mid s)] = c \cdot \frac{\left(\mathbb{E}_s\left[p(y \mid s) \cdot (\langle s, u_c \rangle^2 - 1)\right]\right)^2}{\left(\mathbb{E}_s[p(y \mid s)]\right)^2},$$

where $u_c = \operatorname{argmax}_{\tilde{u}} \int_{\mathbb{R}} \frac{\left(\mathbb{E}_s\left[p(y|s)\cdot(\langle s,\tilde{u}\rangle^2-1)\right]\right)^2}{\mathbb{E}_s[p(y|s)]} dy$ and c is some constant. Thus, we have that

$$f(y) = \sqrt{c} \frac{\mathbb{E}_s\left[p(y \mid s) \cdot (\langle s, u_c \rangle^2 - 1)\right]}{\mathbb{E}_s[p(y \mid s)]},$$

which in turn gives the choice

$$\bar{\mathcal{T}} := \frac{\sqrt{c}\mathcal{T}^*}{1 - (1 - \sqrt{c})\mathcal{T}^*}, \tag{E.7}$$

with

$$\mathcal{T}^*(y) = 1 - \frac{\mathbb{E}_s[p(y \mid s)]}{\mathbb{E}_s\left[p(y \mid s) \cdot \langle s, u_c \rangle^2\right]}.$$

The second condition in (E.4) determines c. Namely,

$$\frac{1}{\delta} = \int_{\mathbb{R}} \left(\frac{\bar{\mathcal{T}}(y)}{1 - \bar{\mathcal{T}}(y)}\right)^2 \mathop{\mathbb{E}}_s[p(y \mid s)] dy = \int_{\mathbb{R}} c \cdot \frac{\left(\mathbb{E}_s\left[p(y \mid s) \cdot (\langle s, u_c \rangle^2 - 1)\right]\right)^2}{\mathbb{E}_s[p(y \mid s)]} dy = c\frac{1}{\delta_t}. \tag{E.8}$$

Therefore, $c = \frac{\delta_t}{\delta}$. (E.8) immediately proves that the second condition in (E.4) is satisfied for $\bar{\mathcal{T}}$. The first condition in (E.4) is also satisfied since

$$\begin{aligned}\max_{\|u\|_2=1} \int_{\mathbb{R}} \frac{\bar{\mathcal{T}}}{1 - \bar{\mathcal{T}}} \mathop{\mathbb{E}}_s\left[p(y \mid s) \cdot (\langle s, u \rangle^2 - 1)\right] dy &\geq \sqrt{c} \cdot \int_{\mathbb{R}} \frac{\left(\mathbb{E}_s\left[p(y \mid s) \cdot (\langle s, u_c \rangle^2 - 1)\right]\right)^2}{\mathbb{E}_s[p(y \mid s)]} dy \\ &= \sqrt{\frac{\delta_t}{\delta}} \cdot \frac{1}{\delta_t} > \frac{1}{\delta}.\end{aligned}$$

Thus, $\bar{\mathcal{T}}$ achieves weak recovery, hence

$$\delta_c \leq \delta_t,$$

which combined with (E.6) proves $\delta_c = \delta_t$. This also implies that the expression in Equation (E.7) coincides with $\mathcal{T}_\delta^*$ as defined in Equation (4.10).

Lastly, we need to verify that $\mathcal{T}_\delta^*$ satisfies Assumption (A5). Let us first prove that $\mathcal{T}_\delta^*(y)$ is bounded. Since $\mathcal{T}^*(y) \leq 1$, it follows that

$$\sqrt{\delta} - (\sqrt{\delta} - \sqrt{\delta_c}) \cdot \mathcal{T}^*(y) \geq \sqrt{\delta_c},$$

and

$$\mathcal{T}_\delta^*(y) = \frac{\sqrt{\delta_c} \cdot \mathcal{T}^*(y)}{\sqrt{\delta} - (\sqrt{\delta} - \sqrt{\delta_c}) \cdot \mathcal{T}^*(y)} \leq \frac{\sqrt{\delta_c}\mathcal{T}^*(y)}{\sqrt{\delta_c}} \leq 1.$$

Furthermore, for $\mathcal{T}^*(y) \neq 0$, we have

$$\mathcal{T}_\delta^*(y) = \frac{\sqrt{\delta_c}}{\frac{\sqrt{\delta}}{\mathcal{T}^*(y)} - (\sqrt{\delta} - \sqrt{\delta_c})},$$

and it holds that $\frac{\sqrt{\delta}}{\mathcal{T}^*(y)} - (\sqrt{\delta} - \sqrt{\delta_c}) \in]-\infty, -(\sqrt{\delta} - \sqrt{\delta_c})[\cup]\sqrt{\delta_c}, +\infty[$. Thus, for $\mathcal{T}^*(y) \neq 0$,

$$\mathcal{T}_\delta^*(y) \geq -\frac{\sqrt{\delta_c}}{\sqrt{\delta} - \sqrt{\delta_c}},$$

whereas $\mathcal{T}_\delta^*(y) = 0$ for $\mathcal{T}^*(y) = 0$. This proves that $\mathcal{T}_\delta^*(y)$ is bounded.

Finally, by contradiction, suppose that $\mathbb{P}(\mathcal{T}_\delta^*(y) = 0) = 1$. Note that $\mathcal{T}_\delta^*(y) = 0$ if and only if $\mathcal{T}^*(y) = 0$. This holds whenever

$$\mathbb{P}\left(\mathbb{E}_s[p(y \mid s)] = \mathbb{E}_s\left[p(y \mid s) \cdot \langle s, u_c \rangle^2\right]\right) = 1. \tag{E.9}$$

However, Equation (E.9) implies that

$$\mathbb{P}\left(\mathbb{E}_s\left[p(y \mid s) \cdot (\langle s, u_c \rangle^2 - 1)\right] = 0\right) = 1,$$

giving that $\delta_c = +\infty$, for which the statement of the theorem trivially holds. Consequently, $\mathbb{P}(\mathcal{T}^*(y) = 0) \neq 1$ and the proof is complete. ■

Appendix F. Optimal preprocessing for the numerical experiments

F.1. $q(s) = \prod_{i=1}^2 s_i$

In the numerical experiment, we employ the function $\mathcal{T}^*(y)$ in place of $\mathcal{T}_\delta^*(y)$, as the latter is introduced for technical reasons. Recall that

$$\mathcal{T}^*(y) := 1 - \frac{\mathbb{E}_s[p(y \mid s)]}{\mathbb{E}_s\left[p(y \mid s) \cdot \langle s, u_c \rangle^2\right]}.$$

We calculate term by term. First, we have

$$\begin{aligned}
\mathop{\mathbb{E}}_{s}[p(y \mid s)] &= \frac{1}{2\pi}\iint_{\mathbb{R}^2} e^{-\frac{s_1^2}{2}} e^{-\frac{s_2^2}{2}} \delta(y - s_1 s_2) ds_1 ds_2 \\
&= \frac{1}{2\pi} 2 \int_0^\infty \int_0^\infty e^{-\frac{s_1^2}{2}} e^{-\frac{s_2^2}{2}} \delta(y - s_1 s_2) ds_1 ds_2,
\end{aligned}$$

where we have assumed that $y > 0$ (similar passages hold for $y \leq 0$). The constant 2 pops out, since there are two symmetric cases for $y > 0$, namely $s_1, s_2 > 0$ and $s_1, s_2 < 0$. Continuing with the change of variable $x_1 = s_1 s_2$ and $x_2 = \frac{s_1}{s_2}$, we have

$$\begin{aligned}
\frac{1}{2\pi} 2 \int_0^\infty \int_0^\infty e^{-\frac{s_1^2}{2}} e^{-\frac{s_2^2}{2}} \delta(y - s_1 s_2) ds_1 ds_2 &= \frac{1}{\pi} \int_0^\infty \int_0^\infty e^{-\frac{x_1 x_2}{2}} e^{-\frac{x_1/x_2}{2}} \delta(y - x_1) \frac{1}{2x_2} dx_1 dx_2 \\
&= \frac{1}{\pi} \int_0^\infty e^{-y\left(\frac{x_2}{2} + \frac{1}{2x_2}\right)} \frac{1}{2x_2} dx_2 \\
&= \frac{1}{2\pi} \int_{-\infty}^\infty e^{-y\cosh(t)} dt \\
&= \frac{1}{\pi} \int_0^\infty e^{-y\cosh(t)} dt \\
&= \frac{K_0(|y|)}{\pi},
\end{aligned}$$

where we did the change of variable $e^t = \frac{1}{x_2}$, and K_0 is the modified Bessel function of the second kind. Let us now calculate the second term for the choice $u_c = \frac{1}{\sqrt{2}}(1, 1)$, which can be verified to be optimal,

$$\begin{aligned}
\mathop{\mathbb{E}}_{s}\Big[p(y \mid s)\langle s, u_c\rangle^2\Big] &= \frac{1}{2\pi}\iint_{\mathbb{R}^2} e^{-\frac{s_1^2}{2}} e^{-\frac{s_2^2}{2}} \delta(y - s_1 s_2) \langle s, u_c\rangle^2 ds_1 ds_2 \\
&= \frac{1}{2\pi} 2 \int_0^\infty \int_0^\infty e^{-\frac{s_1^2}{2}} e^{-\frac{s_2^2}{2}} \delta(y - s_1 s_2) \frac{1}{2}(s_1^2 + s_2^2 + 2 s_1 s_2) ds_1 ds_2,
\end{aligned}$$

where, as before, we have assumed that $y > 0$ (similar passages hold for $y \leq 0$). Continuing with the change of variable $x_1 = s_1 s_2$ and $x_2 = \frac{s_1}{s_2}$, we have

$$\begin{aligned}
\mathop{\mathbb{E}}_{s}\Big[p(y \mid s)\langle s, u_c\rangle^2\Big] &= \frac{1}{\pi} \int_0^\infty \int_0^\infty e^{-\frac{s_1^2}{2}} e^{-\frac{s_2^2}{2}} \delta(y - s_1 s_2) \frac{1}{2}(s_1^2 + s_2^2 + 2 s_1 s_2) ds_1 ds_2 \\
&= \frac{1}{\pi} \int_0^\infty \int_0^\infty e^{-\frac{x_1 x_2}{2}} e^{-\frac{x_1/x_2}{2}} \delta(y - x_1) \frac{1}{2} x_1 \left(x_2 + \frac{1}{x_2} + 2\right) \frac{1}{2x_2} dx_1 dx_2 \\
&= \frac{1}{\pi} \int_0^\infty e^{-y\left(\frac{x_2}{2} + \frac{1}{2x_2}\right)} \frac{1}{2} y \left(x_2 + \frac{1}{x_2} + 2\right) \frac{1}{2x_2} dx_2 \\
&= \frac{1}{2\pi} \int_{-\infty}^\infty e^{-y\cosh(t)} y(\cosh(t) + 1) dx_2 \\
&= \frac{1}{\pi} \int_0^\infty e^{-y\cosh(t)} y(\cosh(t) + 1) dx_2 \\
&= \frac{yK_0(|y|) + |y|K_1(|y|)}{\pi},
\end{aligned}$$

where K_1 is the modified Bessel function with parameter equal to 1. Moreover, the absolute value in the final result is readily obtained after analyzing the case $y < 0$, which is analogous. Thus, we get

$$\mathcal{T}^*(y) = 1 - \frac{K_0(|y|)}{yK_0(|y|) + |y|K_1(|y|)}. \tag{F.1}$$

Finally, let us calculate the weak recovery threshold δ_c:

$$\frac{1}{\delta_c} = \int_{\mathbb{R}} \frac{\left(\mathbb{E}_s\left[p(y \mid s)(\langle s, u\rangle^2 - 1)\right]\right)^2}{\mathbb{E}_s[p(y \mid s)]} dy = \int_{\mathbb{R}} \frac{\left(\frac{yK_0(|y|)+|y|K_1(|y|)}{\pi} - \frac{K_0(|y|)}{\pi}\right)^2}{\frac{K_0(|y|)}{\pi}} dy \approx 1.68421 \approx (0.5937)^{-1},$$

where the exact value was calculated in WolframAlpha. We note that this value exactly matches the threshold in (Troiani et al., 2024, pg.9).

F.2. Mixed phase retrieval

Let $\eta = \mathbb{P}(\varepsilon = 1)$ and define the following auxiliary quantities:

$$\begin{aligned}
\gamma &= \frac{1}{2}\left[1 + \sqrt{4\rho^2\eta(1-\eta) + (2\eta - 1)^2}\right], \\
a_1 &:= 1 + \frac{2(\gamma - \eta)}{\eta} + \left(\frac{\gamma - \eta}{\eta\rho}\right)^2, \\
a_2 &:= \eta + (1-\eta)\rho^2 + \frac{2(\gamma - \eta)}{\eta} + \left(\frac{\gamma - \eta}{\eta\rho}\right)^2 \left[(1-\eta) + \eta\rho^2\right], \\
a_3 &:= (1-\eta)(1-\rho^2) + \left(\frac{\gamma - \eta}{\eta\rho}\right)^2 \eta(1-\rho^2), \\
b &:= a_1 - a_3.
\end{aligned}$$

Denote by ℓ^* the unique solution in $\ell \in](\gamma/\eta)^2 - b, \infty[$ to the following equation:

$$(\ell - a_3) \int_0^\infty \sqrt{\frac{2}{\pi}} e^{-y^2/2} \frac{(y^2-1)^2}{a_2 y^2 + \ell} \, dy = \frac{1}{\gamma^2 \delta}.$$

Then, the optimal preprocessing function is given by

$$\mathcal{T}(y) = \frac{y^2 - 1}{[a_2 + \gamma(\ell^* - a_3)]y^2 + \ell^* - \gamma(\ell^* - a_3)}. \tag{F.2}$$

This expression maximizes the asymptotic overlap $\left|\langle v_1^D, w_1^*\rangle\right|$ obtained by specializing our general formula in Theorem 4 to the mixed phase retrieval model. The derivations are along similar lines as detailed in Section F.1 for the model $y = s_1 s_2$, and we leave out the explicit calculations.

Appendix G. Equivalence to Troiani et al. (2024)

Proposition 20 *For simultaneously diagonalizable matrices $E(y)$ and orthogonal signals, the recovery threshold from Theorem 4 matches the one in Lemma 4.1 of Troiani et al. (2024). This equivalence holds for all permutation-invariant link functions.*

Proof The weak recovery threshold in Troiani et al. (2024) is characterized as

$$\frac{1}{\alpha_c} = \sup_{\substack{M \in \mathcal{S}_p^+ \\ \|M\|_F = 1}} \|\mathcal{F}(M)\|_F.$$

Here, $\mathcal{F}$ is an operator defined as

$$\mathcal{F}(M) \coloneqq \mathop{\mathbb{E}}_{y}[E(y) M E(y)],$$

where $E(y) \coloneqq \mathbb{E}_s\big[ss^\top - I_p \mid y\big]$. As noted in the proof of Lemma 4.1 in Troiani et al. (2024), the linear operator $\mathcal{F}$ is a self-map on the cone of positive semi-definite matrices. Indeed, for any positive semi-definite matrix M and any vector v, $v^\top \mathcal{F}(M) v = \mathbb{E}_y\big[v^\top E(y) M E(y) v\big] = \mathbb{E}_y\Big[\big\|M^{1/2} E(y) v\big\|_2^2\Big] \geq 0$, meaning that $\mathcal{F}(M)$ is also positive semi-definite. By the generalized Perron–Frobenius theorem for cone-preserving maps (Kreĭn and Rutman, 1948) (see also Theorem 1.1 in (Du, 2006), or Theorem 19.2 and Exercise 12 in (Deimling, 1985, §19)), it holds that $1/\alpha_c$ is the largest eigenvalue of the linear operator $\mathcal{F}$. That is, there exists a positive semi-definite matrix M^* with $\|M^*\|_F = 1$ for which it holds

$$\sup_{\substack{M \in \mathcal{S}_p^+ \\ \|M\|_F = 1}} \|\mathcal{F}(M)\|_F = \|\mathcal{F}(M^*)\|_F.$$

The corresponding eigen-equation $\mathcal{F}(M^*) = \lambda^* M^*$ implies $\langle \mathcal{F}(M^*), M^* \rangle_F = \lambda^*$. The LHS must be non-negative since $\langle \mathcal{F}(M^*), M^* \rangle_F = \sum_i \lambda_i^{M^*} \big\langle \mathcal{F}(M^*), v_i^{M^*} (v_i^{M^*})^\top \big\rangle_F \geq 0$ due to the cone-preserving property of $\mathcal{F}$ and non-negativity of each $\lambda_i^{M^*}$. Now we have $\|\mathcal{F}(M^*)\|_F^2 = \lambda^* \langle \mathcal{F}(M^*), M^* \rangle_F = \langle \mathcal{F}(M^*), M^* \rangle_F^2$, from which it follows that $\|\mathcal{F}(M^*)\|_F = \langle \mathcal{F}(M^*), M^* \rangle_F$, by non-negativity of the RHS just shown. So, one can rewrite the optimal threshold as

$$\sup_{\substack{M \in \mathcal{S}_p^+ \\ \|M\|_F = 1}} \|\mathcal{F}(M)\|_F = \sup_{\substack{M \in \mathcal{S}_p^+ \\ \|M\|_F = 1}} \langle \mathcal{F}(M), M \rangle_F.$$

Let us assume first that this maximum is obtained for a certain rank-1 matrix. Then, it would hold

$$\sup_{\substack{M \in \mathcal{S}_p^+ \\ \|M\|_F = 1}} \langle \mathcal{F}(M), M \rangle_F = \max_{\|u\|_2 = 1} \Big\langle \mathcal{F}(uu^\top), uu^\top \Big\rangle_F, \tag{G.1}$$

where $u \in \mathbb{R}^p$. Writing this out, we have

$$\begin{aligned}
\max_{\|u\|_2=1} \left\langle \mathcal{F}(uu^\top), uu^\top \right\rangle_F &= \max_{\|u\|_2=1} \operatorname{Tr}\left(\mathop{\mathbb{E}}_{y}\left[E(y)uu^\top E(y) \right] uu^\top \right) \\
&= \max_{\|u\|_2=1} \mathop{\mathbb{E}}_{y}\left[\operatorname{Tr}\left(u^\top E(y)uu^\top E(y)u \right) \right] \\
&= \max_{\|u\|_2=1} \mathop{\mathbb{E}}_{y}\left[\left(\mathop{\mathbb{E}}_{s}\left[\langle s, u\rangle^2 - 1 \mid y \right] \right)^2 \right] \\
&= \max_{\|u\|_2=1} \mathop{\mathbb{E}}_{y}\left[\left(\mathop{\mathbb{E}}_{s}\left[\frac{p(y \mid s)}{p(y)} (\langle s, u\rangle^2 - 1) \right] \right)^2 \right] \\
&= \max_{\|u\|_2=1} \mathop{\mathbb{E}}_{y}\left[\frac{\left(\mathbb{E}_s\left[p(y \mid s)(\langle s, u\rangle^2 - 1) \right] \right)^2}{\left(\mathbb{E}_s[p(y \mid s)] \right)^2} \right] \\
&= \max_{\|u\|_2=1} \int_{\mathbb{R}} p(y) \frac{\left(\mathbb{E}_s\left[p(y \mid s)(\langle s, u\rangle^2 - 1) \right] \right)^2}{\left(\mathbb{E}_s[p(y \mid s)] \right)^2} dy \\
&= \max_{\|u\|_2=1} \int_{\mathbb{R}} \frac{\left(\mathbb{E}_s\left[p(y \mid s)(\langle s, u\rangle^2 - 1) \right] \right)^2}{\mathbb{E}_s[p(y \mid s)]} dy \\
&= \frac{1}{\delta_c},
\end{aligned}$$

where δ_c is exactly the threshold from Theorem 4.

Let us prove that when $E(y)$ has the same eigenvectors regardless of y, the equality in (G.1) holds. Towards that end, we denote by u_i ($i \in [p]$) the orthonormal eigenbasis of $E(y)$. We introduce the matrices

$$A_i \coloneqq u_i u_i^\top, \quad B_{j,k} = \frac{u_j u_k^\top + u_k u_j^\top}{\sqrt{2}},$$

for $i, j, k \in [p]$ s.t. $j \neq k$. It holds that

$$\begin{aligned}
\langle A_{i_1}, A_{i_2} \rangle_F &= 0, \\
\langle A_{i_1}, A_{i_1} \rangle_F &= 1, \\
\langle A_{i_1}, B_{j_1,k_1} \rangle_F &= 0, \\
\langle B_{j_1,k_1}, B_{j_2,k_2} \rangle_F &= 0, \\
\langle B_{j_1,k_1}, B_{j_1,k_1} \rangle_F &= 1,
\end{aligned}$$

for any $i_1, i_2, j_1, j_2, k_1, k_2 \in [p]$ such that $i_1 \neq i_2$ and $(j_1, k_1) \neq (j_2, k_2)$. Due to the fact that $u_1, \dots, u_p$ form an orthonormal basis, the set

$$U \coloneqq \{A_i, B_{j,k}; i, j, k \in [p], j < k\}$$

is an orthonormal basis of the symmetric matrices $\mathcal{S}_p$ equipped with Frobenius inner product. Notice also that

$$\sup_{\substack{M \in \mathcal{S}_p^+ \\ \|M\|_F = 1}} \langle \mathcal{F}(M), M \rangle_F \leq \sup_{\substack{M \in \mathcal{S}_p \\ \|M\|_F = 1}} \langle \mathcal{F}(M), M \rangle_F,$$

since all positive definite matrices are by definition symmetric. By the previous discussion, we decompose the matrix $M \in \mathcal{S}_p$ such that $\|M\|_F = 1$, as

$$M = \sum_{i=1}^{p} \alpha_i A_i + \sum_{1 \le j < k \le p} \beta_{j,k} B_{j,k},$$

where $\alpha_i, \beta_{j,k} \in \mathbb{R}$ such that

$$\sum_{i=1} \alpha_i^2 + \sum_{1 \le j < k \le p} \beta_{j,k}^2 = 1.$$

Since the u_i's are eigenvectors of $E(y)$, it follows that every element of the set U is an eigenvector of the operator $\mathcal{F}$. Let us denote the corresponding eigenvalues as λ_{A_i} and $\lambda_{B_{j,k}}$. Furthermore, it holds that

$$\begin{aligned}
\lambda_{B_{j,k}} &= \langle \mathcal{F}(B_{j,k}), B_{j,k} \rangle \\
&= \mathop{\mathbb{E}}_{y}\Big[u_j^\top E(y) u_j \cdot u_k^\top E(y) u_k\Big] \\
&\le \sqrt{\mathop{\mathbb{E}}_{y}\Big[(u_j^\top E(y) u_j)^2\Big]} \cdot \sqrt{\mathop{\mathbb{E}}_{y}\big[(u_k^\top E(y) u_k)^2\big]} \\
&= \sqrt{\langle \mathcal{F}(A_j), A_j \rangle} \cdot \sqrt{\langle \mathcal{F}(A_k), A_k \rangle} \\
&= \sqrt{\lambda_{A_j} \lambda_{A_k}},
\end{aligned}$$

where Hölder's inequality was used to get the bound. Then, it holds that

$$\begin{aligned}
\langle \mathcal{F}(M), M \rangle_F &= \left\langle \mathcal{F}\left(\sum_{i=1}^{p} \alpha_i A_i + \sum_{1 \le j < k \le p} \beta_{j,k} B_{j,k}\right), \sum_{i=1}^{p} \alpha_i A_i + \sum_{1 \le j < k \le p} \beta_{j,k} B_{j,k} \right\rangle \\
&= \left\langle \sum_{i=1}^{p} \alpha_i \lambda_{A_i} A_i + \sum_{1 \le j < k \le p} \beta_{j,k} \lambda_{B_{j,k}} B_{j,k}, \sum_{i=1}^{p} \alpha_i A_i + \sum_{1 \le j < k \le p} \beta_{j,k} B_{j,k} \right\rangle \\
&= \sum_{i=1}^{p} \alpha_i^2 \lambda_{A_i} + \sum_{1 \le j < k \le p} \beta_{j,k}^2 \lambda_{B_{j,k}} \\
&\le \sum_{i=1}^{p} \alpha_i^2 \lambda_{A_i} + \sum_{1 \le j < k \le p} \beta_{j,k}^2 \sqrt{\lambda_{A_j} \lambda_{A_k}} \\
&\le \max_i \lambda_{A_i} \cdot \left(\sum_{i=1} \alpha_i^2 + \sum_{1 \le j < k \le p} \beta_{j,k}^2\right) \\
&= \max_i \lambda_{A_i} \\
&\le \max_{\|u\|_2 = 1} \left\langle \mathcal{F}(uu^\top), uu^\top \right\rangle_F,
\end{aligned} \tag{G.2}$$

which proves (G.1).

Lastly, we verify that $E(y)$ is diagonalizable when the link function q is permutation invariant. Towards that end, it holds for arbitrary $i, j \in [p]$ that

$$\begin{aligned}
(E(y))_{i,i} &= \mathop{\mathbb{E}}_{s}\big[s_i^2 - 1 \mid y\big] \\
&= \mathop{\mathbb{E}}_{s}\big[s_i^2 - 1 \mid q(\dots, s_i, \dots, s_j, \dots) = y\big] \\
&= \mathop{\mathbb{E}}_{s}\big[s_j^2 - 1 \mid q(\dots, s_j, \dots, s_i, \dots) = y\big] \\
&= \mathop{\mathbb{E}}_{s}\big[s_j^2 - 1 \mid q(\dots, s_i, \dots, s_j, \dots) = y\big] \\
&= (E(y))_{j,j},
\end{aligned}$$

since the link function q is permutation invariant. Similarly, for arbitrary $i_1, i_2, j_1, j_2 \in [p]$ such that $i_1 \neq j_1$ and $i_2 \neq j_2$, it holds that

$$\begin{aligned}
(E(y))_{i_1,j_1} &= \mathop{\mathbb{E}}_{s}[s_{i_1} s_{j_1} \mid y] \\
&= \mathop{\mathbb{E}}_{s}[s_{i_1} s_{j_1} \mid q(\dots, s_{i_1}, \dots, s_{i_2}, \dots, s_{j_1}, \dots, s_{j_2}, \dots) = y] \\
&= \mathop{\mathbb{E}}_{s}[s_{i_2} s_{j_2} \mid q(\dots, s_{i_2}, \dots, s_{i_1}, \dots, s_{j_2}, \dots, s_{j_1}, \dots) = y] \\
&= \mathop{\mathbb{E}}_{s}[s_{i_2} s_{j_2} \mid q(\dots, s_{i_1}, \dots, s_{i_2}, \dots, s_{j_1}, \dots, s_{j_2}, \dots) = y] \\
&= (E(y))_{i_2,j_2}.
\end{aligned}$$

Thus, if we denote by $a(y) \coloneqq E(y)_{i,i}$ for arbitrary $i \in [p]$ and by $b(y) \coloneqq E(y)_{i_1,j_1}$ for arbitrary $i_1, j_1 \in [p]$ s.t. $i_1 \neq j_1$, it follows that

$$E(y) = \begin{bmatrix} a(y) & b(y) & \cdots & b(y) \\ b(y) & a(y) & \cdots & b(y) \\ \vdots & \vdots & \ddots & \vdots \\ b(y) & b(y) & \cdots & a(y) \end{bmatrix}. \tag{G.3}$$

Notice that this matrix is going to have an eigenbasis

$$\begin{aligned}
u_1 &\coloneqq (1, 1, 1, \dots 1)/\sqrt{p}, \\
u_2 &\coloneqq (1, -1, 0, \dots 0)/\sqrt{2}, \\
&\vdots \\
u_p &\coloneqq (0, 0, \dots, 0, 1, -1)/\sqrt{2},
\end{aligned}$$

regardless of the value of y, which proves the final claim. ■