Two-way kernel matrix puncturing: towards resource-efficient PCA and spectral clustering

— Supplementary Material —

May 13, 2021

Abstract

This supplementary material provides the proofs of the main theorems of the core article.

1 Reminder of the main setting

For convenience, we first recall our main setting and assumptions. The central object of interest is the matrix

$$K = \left\{ \frac{1}{p} (X \odot S)^{\mathsf{H}} (X \odot S) \right\} \odot B \in \mathbb{C}^{n \times n} \tag{1}$$

under the large dimensional n, p regime. Here, X, S and B satisfy the following assumptions.

Assumption 1 (Data model)**.**

$$X = Z + P$$

where the $Z_{ij} \sim \mathcal{CN}(0,1)$ *are independent, and where* $P \in \mathbb{C}^{p \times n}$ *is a rank-*k *matrix for some* k*.*

Assumption 2 (Large p, n asymptotics)**.** *As* $n \to \infty$*,*

$$p/n \to c_0 \in (0, \infty)$$

and there exists a decomposition $P = LV^{\mathsf{H}}$ *of* P *with* $V \in \mathbb{C}^{n \times k}$ *isometric (i.e.,* $V^{\mathsf{H}}V = I_k$*) and*

$$\frac{1}{p} L^{\mathsf{H}} L \to \mathcal{L}$$

for some deterministic matrix $\mathcal{L} \in \mathbb{C}^{k \times k}$*. In particular, the eigenvalues of* $\mathcal{L}$ *are the limiting* k *non-trivial eigenvalues of* $\frac{1}{p} P^{\mathsf{H}} P$*. Besides,*

$$\limsup_n \max_{\substack{1 \le i \le n \\ 1 \le j \le k}} \{\sqrt{n} V_{ij}^2\} = 0.$$

2 The theorems

The spectral characterization of K is made through the study of its resolvent matrix

$$Q(z) = (K - zI_n)^{-1}.$$

The results are then as follows.

Theorem 1 (Deterministic equivalent for Q)**.** *Under Assumptions 1–2, let* $z \in \mathbb{C}$ *be away from the limsup of the union of supports of* $\nu_1, \nu_2, \ldots$*. Then, as* $n \to \infty$*,*

$$Q(z) \leftrightarrow m(z) \left[I_n + \frac{c_0^{-1} \varepsilon_S^2 \varepsilon_B m(z)}{1 + \varepsilon_B \varepsilon_S c_0^{-1} m(z)} V \mathcal{L} V^{\mathsf{H}} \right]^{-1}$$

where $m(\cdot)$ *is the unique Stieltjes transform solution to*[1]

$$z = \varepsilon_S b - \frac{1}{m(z)} - c_0^{-1} \varepsilon_B \varepsilon_S^2 m(z) + \frac{c_0^{-2} \varepsilon_B^3 \varepsilon_S^3 m(z)^2}{1 + c_0^{-1} \varepsilon_B \varepsilon_S m(z)}.$$

This theorem is in fact sufficiently exhaustive to characterize both the *macroscopic* spectrum of K (its limiting spectral measure) as well as the *microscopic* behavior of its dominant isolated eigenvalues and associated eigenvectors. The next result, which we name theorem in compliance with the core article, is in effect an (important) corollary of Theorem 1.

Theorem 2 (Phase transition, isolated eigenvalues and eigenvectors)**.** *Define the functions*

$$F(t) = t^4 + \frac{2}{\varepsilon_S} t^3 + \frac{1}{\varepsilon_S^2} \left(1 - \frac{c_0}{\varepsilon_B}\right) t^2 - \frac{2c_0}{\varepsilon_S^3} t - \frac{c_0}{\varepsilon_S^4}$$
$$G(t) = \varepsilon_S b + c_0^{-1} \varepsilon_B \varepsilon_S (1 + \varepsilon_S t) + \frac{\varepsilon_S}{1 + \varepsilon_S t} + \frac{\varepsilon_B}{t(1 + \varepsilon_S t)}$$

and $\Gamma \in \mathbb{R}$ *be the largest real solution to* $F(\Gamma) = 0$*. Further denote* $\ell_1 > \ldots > \ell_{\bar{k}}$ *the* $\bar{k} \le k$ distinct *eigenvalues of* $\mathcal{L}$ *of respective multiplicities* $L_1, \ldots, L_{\bar{k}}$*, and* $\Pi_1, \ldots, \Pi_{\bar{k}} \in \mathbb{R}^{k \times k}$ *the projectors on their respective associated eigenspaces. Similarly denote* $(\lambda_1, \hat{v}_1), \ldots, (\lambda_n, \hat{v}_n)$ *the eigenvalue-eigenvector pairs of* K *in descending order and gather the first* k *eigenvectors under the isometric matrices* $\hat{\mathcal{V}}_1 = [\hat{v}_1, \ldots, \hat{v}_{L_1}]$ *up to* $\hat{\mathcal{V}}_{\bar{k}} = [\hat{v}_{k - L_{\bar{k}} + 1}, \ldots, \hat{v}_k]$.

[1] We also recall that the notation $A \leftrightarrow B$ stands for the fact that, for any linear functional $u : \mathbb{C}^{n \times n} \to \mathbb{R}$ of bounded infinity norm, $u(A - B) \to 0$ almost surely as $n \to \infty$.

Then, for $i \in \{1,\ldots,\bar{k}\}$ *and for all* $j \in \{L_1+\ldots+L_{i-1}+1,\ldots,L_1+\ldots+L_i\}$*,*

$$\lambda_j \to \rho_i \equiv \begin{cases} F(\ell_i) & ,\ \ell_i > \Gamma \\ F(\Gamma) & ,\ \ell_i \leq \Gamma \end{cases}$$

almost surely, and

$$\hat{\mathcal{V}}_i \hat{\mathcal{V}}_i^{\mathsf{H}} \leftrightarrow \zeta_i V \Pi_i V^{\mathsf{H}},\ \textit{where}\ \zeta_i = \begin{cases} \frac{F(\ell_i)\varepsilon_S^3}{\ell_i(1+\varepsilon_S \ell_i)^3} & ,\ \ell_i > \Gamma \\ 0 & ,\ \ell_i \leq \Gamma \end{cases}$$

with the notation '$\leftrightarrow$*' introduced in Theorem 1. In particular, if the* ℓ_i*'s have unit multiplicities with associated population eigenvectors* v_i*, then*

$$|v_i^{\mathsf{H}} \hat{v}_i|^2 \to \zeta_i, \quad i = 1,\ldots,k.$$

3 Elements of proof

3.1 Rationale

The proof relies on the *Gaussian tools* for random matrices popularized in [] and consisting in exploiting Stein's lemma

$$\mathbb{E}[z\phi(z)] = \mathbb{E}\left[\frac{\partial}{\partial \bar{z}}\phi(z)\right]$$

for standard complex (or real) Gaussian random variables $z \sim \mathcal{CN}(0,1)$ along with the Nash-Poincaré inequality

$$\mathrm{Var}[f(z)] \leq \sum_{i=1}^{n} \left(\mathbb{E}\left[\left|\frac{\partial}{\partial z_i} f(z)\right|^2\right] + \mathbb{E}\left[\left|\frac{\partial}{\partial \bar{z}_i} f(z)\right|^2\right] \right)$$

for standard multivariate complex Gaussian $z \sim \mathcal{CN}(0, I_n)$.

As we shall see, Stein's lemma is used to "*unfold*" the a priori quite involved form of the expected value $\mathbb{E}[Q_{ij}]$ of the entries of the resolvent matrix Q of K. The Nash-Poincaré inequality is then subsequently used to control that the variance of Q_{ij} vanishes at a proper rate.

3.2 Proof of Theorem 1

In order to be in a position to apply Stein's lemma, we first exploit the straightforward *resolvent identity*: $KQ - zQ = I_n$, so to obtain

$$\mathbb{E}[Q_{ij}] = -\frac{1}{z}\delta_{ij} + \frac{1}{z}\mathbb{E}[[KQ]_{ij}].$$

By expanding $X = Z + P$, with P decomposed as $P = LV^{\mathsf{H}}$ ($L \in \mathbb{C}^{p\times k}$ and $V \in \mathbb{C}^{n\times k}$), we have to consider four terms in the expansion of $\mathbb{E}[[KQ]_{ij}]$.

Term 1: involving (Z^{H}, Z)

Anticipating coming results, instead of evaluating $\mathbb{E}[[KQ]_{ij}]$ directly, we rather evaluate a modified version in which matrix B is replaced by a deterministic matrix A with bounded operator norm and bounded entries: using Stein's lemma, we have

$$
\begin{aligned}
&\mathbb{E}\left[\left[\left(\left[\frac{1}{p}(Z\odot S)^{\mathsf{H}}(Z\odot S)\right]\odot A\right)Q\right]_{ij}\right] \\
&= \frac{1}{p}\sum_{l=1}^{p}\sum_{m=1}^{n} S_{li}S_{lm}A_{im}\mathbb{E}\left[\bar{Z}_{li}Z_{lm}Q_{mj}\right] \\
&= \frac{1}{p}\sum_{l=1}^{p}\sum_{m=1}^{n} S_{li}S_{lm}A_{im}\left(\mathbb{E}\left[\delta_{im}Q_{mj} + Z_{lm}\frac{\partial Q_{mj}}{\partial Z_{li}}\right]\right). \qquad (2)
\end{aligned}
$$

Using $\frac{\partial Q}{\partial Z_{ab}} = -Q\frac{\partial Q}{\partial Z_{ab}}Q$, it then comes

$$
\frac{\partial Q_{cd}}{\partial Z_{ab}} = -\frac{1}{p}\sum_{l,l'=1}^{n} Q_{il}\left[\left[(X\odot S)'(E_{ab}\odot S)\right]\odot B\right]_{ll'}Q_{l'd},
$$

for E_{ab} the matrix with all zero entries but at coordinate (a,b) where the entry equals 1. We further have that

$$
\left[\left[(X\odot S)^{\mathsf{H}}(E_{ab}\odot S)\right]\odot B\right]_{ll'} = \sum_{o=1}^{p}\bar{X}_{ol}S_{ol}\delta_{oa}S_{ab}B_{ll'} = \bar{X}_{al}S_{al}S_{ab}B_{ll'}\delta_{l'b}
$$

so that

$$
\frac{\partial Q_{cd}}{\partial Z_{ab}} = -\frac{1}{p}\left[QD_{B_{\cdot,b}}(X\odot S)^{\mathsf{H}}\right]_{ca}S_{ab}Q_{bd}.
$$

We then obtain for $T_1(A,S) \equiv \mathbb{E}\left[\left[\left(\left[\frac{1}{p}(Z\odot S)^{\mathsf{H}}(Z\odot S)\right]\odot A\right)Q\right]_{ij}\right]$ in (2):

$$
\begin{aligned}
T_1(A,S) &= \frac{1}{p}\sum_{l=1}^{p}\sum_{m=1}^{n} S_{li}S_{lm}A_{im}\mathbb{E}\left[\delta_{im}Q_{mj} - \frac{1}{p}Z_{lm}\left[QD_{B_{\cdot,i}}(X\odot S)^{\mathsf{H}}\right]_{ml}S_{li}Q_{ij}\right] \\
&= \mathbb{E}\left[\frac{1}{p}\left[S^{\mathsf{H}}S\right]_{ii}A_{ii}Q_{ij}\right] - \frac{1}{p^2}\sum_{l=1}^{p}\left[ZD_{S_{l,\cdot}}D_{A_{i,\cdot}}QD_{B_{\cdot,i}}(X\odot S)^{\mathsf{H}}D_{S_{\cdot,i}}D_{S_{\cdot,i}}\right]_{ll}Q_{ij}
\end{aligned}
$$

where D_x denotes the diagonal matrix with elements the entries of vector x.

Term 2: (Z', P)

We obtain for $T_2(A,P) \equiv \mathbb{E}\left[\left[\left(\left[\frac{1}{p}(Z\odot S)^{\mathsf{H}}(P\odot S)\right]\odot A\right)Q\right]_{ij}\right]$:

$$\begin{aligned}
T_2(A,R) &= \sum_{l=1}^{p}\sum_{m=1}^{n}\frac{1}{p}\mathbb{E}\left[\bar{Z}_{li}S_{li}P_{lm}S_{lm}A_{im}Q_{mj}\right],\\
&= \frac{1}{p}\sum_{l=1}^{p}\sum_{m=1}^{n}S_{li}P_{lm}S_{lm}A_{im}\mathbb{E}\left[\frac{\partial Q_{mj}}{\partial z_{li}}\right],\\
&= -\frac{1}{p^2}\sum_{l=1}^{p}\sum_{m=1}^{n}S_{li}P_{lm}S_{lm}A_{im}\mathbb{E}\left[\left(QD_{B_{\cdot,i}}(X\odot S)^{\mathsf{H}}\right)_{ml}S_{li}Q_{ij}\right]\\
&= -\frac{1}{p^2}\sum_{l=1}^{p}\mathbb{E}\left[\left[TD_{S_{l,\cdot}}D_{A_{i,\cdot}}QD_{B_{\cdot,i}}(X\odot S)^{\mathsf{H}}D_{S_{\cdot,i}}\right]_{ll}Q_{ij}\right]
\end{aligned}$$

where we used in particular the fact that $D_{S_{\cdot,i}}^2 = D_{S_{\cdot,i}}$.

Summation of $T_1(A,S)$ and $T_2(A,S)$:

Summing the two previous terms, we get

$$\begin{aligned}
T_1(A,S)+T_2(A,S) = \mathbb{E}&\left[\frac{1}{p}\left[S^{\mathsf{H}}S\right]_{ii}A_{ii}Q_{ij}\right]\\
&-\frac{1}{p^2}\sum_{l=1}^{p}\mathbb{E}\left[\left[XD_{S_{l,\cdot}}D_{A_{i,\cdot}}QD_{B_{\cdot,i}}(X\odot S)^{\mathsf{H}}D_{S_{\cdot,i}}\right]_{ll}Q_{ij}\right]
\end{aligned}$$

Due to the presence of the term $S_{l,\cdot}$ inside the matrix evaluated at position (l,l), the summation over l cannot be "turned into a trace", as conventionally done to prove e.g., the Marčenko-Pastur theorem [] (when $S_{ij}=1$ and $B_{ij}=1$ for all i,j). We therefore need to proceed otherwise by writing

$$\begin{aligned}
&\frac{1}{p^2}\sum_{l=1}^{p}\left[XD_{S_{l,\cdot}}D_{A_{i,\cdot}}QD_{B_{\cdot,i}}(X\odot S)^{\mathsf{H}}D_{S_{\cdot,i}}\right]_{ll}\\
&=\frac{1}{p^2}\sum_{l=1}^{p}X_{l,\cdot}D_{S_{l,\cdot}}D_{A_{i,\cdot}}QD_{B_{\cdot,i}}D_{S_{l,\cdot}}X_{l,\cdot}^{\mathsf{H}}S_{l,i}.
\end{aligned}$$

To evaluate the quadratic forms, we must "break" the dependence between $X_{l,\cdot}$ and Q. To this end, note that

$$Q = \left(\frac{1}{p}\sum_{i=1}^{p}\left\{\left[D_{S_{i,\cdot}}X_{i,\cdot}^{\mathsf{H}}X_{i,\cdot}D_{S_{i,\cdot}}\right]\odot B\right\}-zI_n\right)^{-1}$$

so that, applying Woodbury's identity,

$$Q = Q_{-l}-\frac{1}{p}Q_{-l}\left\{\left[D_{S_{l,\cdot}}X_{i,\cdot}^{\mathsf{H}}X_{i,\cdot}D_{S_{l,\cdot}}\right]\odot B\right\}\left(I_n+\frac{1}{p}Q_{-l}\left\{\left[D_{S_{l,\cdot}}X_{i,\cdot}^{\mathsf{H}}X_{i,\cdot}D_{S_{l,\cdot}}\right]\odot B\right\}\right)^{-1}Q_{-l}.$$

Plugged into the quadratic form over $X_{l,\cdot}$, this gives:

$$\frac{1}{p}X_{l,\cdot}D_{S_{l,\cdot}}D_{A_{i,\cdot}}QD_{B_{\cdot,i}}D_{S_{l,\cdot}}X_{l,\cdot}^{\mathsf{H}}$$
$$= \frac{1}{p}X_{l,\cdot}D_{S_{l,\cdot}}D_{A_{i,\cdot}}Q_{-l}D_{B_{\cdot,i}}D_{S_{l,\cdot}}X_{l,\cdot}^{\mathsf{H}}$$
$$- \frac{1}{p^2}X_{l,\cdot}D_{S_{l,\cdot}}D_{A_{i,\cdot}}Q_{-l}\left\{\left[D_{S_{l,\cdot}}X_{l,\cdot}^{\mathsf{H}}X_{l,\cdot}D_{S_{l,\cdot}}\right]\odot B\right\}\left(I_n + \frac{1}{p}Q_{-l}\left\{\left[D_{S_{l,\cdot}}X_{l,\cdot}^{\mathsf{H}}X_{l,\cdot}D_{S_{l,\cdot}}\right]\odot B\right\}\right)^{-1}$$
$$\times Q_{-l}D_{B_{\cdot,i}}D_{S_{l,\cdot}}X_{l,\cdot}^{\mathsf{H}}.$$

Recalling that $X = Z + LV^{\mathsf{H}}$ (so that $X_{l,\cdot} = Z_{l,\cdot} + L_{l,\cdot}V^{\mathsf{H}}$), we first find that, averaging over $Z_{l,\cdot}$,

$$\frac{1}{p}\mathbb{E}\left[X_{l,\cdot}D_{S_{l,\cdot}}D_{A_{i,\cdot}}Q_{-l}D_{B_{\cdot,i}}D_{S_{l,\cdot}}X_{l,\cdot}^{\mathsf{H}}\right]$$
$$= \frac{1}{p}\mathbb{E}\left[\mathrm{tr}D_{S_{l,\cdot}}D_{A_{i,\cdot}}Q_{-l}D_{B_{\cdot,i}}D_{S_{l,\cdot}}\right] + \frac{1}{p}\mathbb{E}\left[L_{l,\cdot}V^{\mathsf{H}}D_{S_{l,\cdot}}D_{A_{i,\cdot}}Q_{-l}D_{B_{\cdot,i}}D_{S_{l,\cdot}}VL_{l,\cdot}^{\mathsf{H}}\right].$$

A further application of the Nash-Poincaré inequality then shows that the variance of $X_{l,\cdot}D_{S_{l,\cdot}}D_{A_{i,\cdot}}Q_{-l}D_{B_{\cdot,i}}D_{S_{l,\cdot}}X_{l,\cdot}^{\mathsf{H}}$ vanishes as $O(1/p)$ while $\mathrm{Var}[Q_{ij}] = O(1)$, so that the above result extends into

$$\frac{1}{p}\mathbb{E}\left[X_{l,\cdot}D_{S_{l,\cdot}}D_{A_{i,\cdot}}Q_{-l}D_{B_{\cdot,i}}D_{S_{l,\cdot}}X_{l,\cdot}^{\mathsf{H}}Q_{ij}\right]$$
$$= \frac{1}{p}\mathbb{E}\left[\mathrm{tr}D_{S_{l,\cdot}}D_{A_{i,\cdot}}Q_{-l}D_{B_{\cdot,i}}D_{S_{l,\cdot}}\right]\mathbb{E}[Q_{ij}]$$
$$+ \frac{1}{p}\mathbb{E}\left[L_{l,\cdot}V^{\mathsf{H}}D_{S_{l,\cdot}}D_{A_{i,\cdot}}Q_{-l}D_{B_{\cdot,i}}D_{S_{l,\cdot}}VL_{l,\cdot}^{\mathsf{H}}\right]\mathbb{E}[Q_{ij}] + O(p^{-\frac{1}{2}}).$$

Now observe, for the second right-hand side term, that, by Cauchy-Schwarz's inequality and after summation over l,

$$\left(\frac{1}{p^2}\sum_{l=1}^{p}\mathbb{E}\left[L_{l,\cdot}V^{\mathsf{H}}D_{S_{l,\cdot}}D_{A_{i,\cdot}}Q_{-l}D_{B_{\cdot,i}}D_{S_{l,\cdot}}VL_{l,\cdot}^{\mathsf{H}}Q_{ij}\right]S_{l,i}\right)^2$$
$$\leq \frac{1}{p^2}\sum_{l'=1}^{n}\|L_{l',\cdot}\|^2\frac{1}{p^2}\sum_{l=1}^{p}\mathbb{E}[|Q_{ij}|^2L_{l,\cdot}D_{S_{\cdot,i}}V^{\mathsf{H}}D_{S_{l,\cdot}}D_{B_{\cdot,i}}$$
$$Q_{-l}^{\mathsf{H}}D_{S_{l,\cdot}}D_{A_{i,\cdot}}VV^{\mathsf{H}}D_{S_{l,\cdot}}D_{A_{i,\cdot}}Q_{-l}D_{B_{\cdot,i}}D_{S_{l,\cdot}}VD_{S_{\cdot,i}}L_{l,\cdot}^{\mathsf{H}}]$$
$$\leq \mathrm{tr}(L^{\mathsf{H}}L)\frac{1}{p^4}\sum_{l=1}^{p}\left\|\mathbb{E}[|Q_{ij}|^2D_{S_{\cdot,i}}V^{\mathsf{H}}D_{S_{l,\cdot}}D_{B_{\cdot,i}}Q_{-l}^{\mathsf{H}}D_{S_{l,\cdot}}D_{A_{i,\cdot}}VV^{\mathsf{H}}D_{S_{l,\cdot}}D_{A_{i,\cdot}}Q_{-l}\right.$$
$$\left.D_{B_{\cdot,i}}D_{S_{l,\cdot}}VD_{S_{\cdot,i}}]\right\|\|L_{l,\cdot}\|^2$$
$$\leq \frac{C}{p^2}(\mathrm{tr}(L^{\mathsf{H}}L))^2 = O(p^{-2})$$

for $C > 0$ a bound on the norm of the matrix in the expectation term. This bound holds because we imposed that $L^{\mathsf{H}}L \to \mathcal{L} = O_{\|\cdot\|}(1)$, because $\|Q_{-l}\| \leq 1/|\Im[z]|$ (or $\leq 1/z$ for $z < 0$) and because all entries of S, A, B are bounded.

Therefore, the term in the first line parentheses above is of order $O(p^{-1})$. As a consequence,

$$\begin{aligned}
&\frac{1}{p^2}\sum_{l=1}^{p}\mathbb{E}\left[X_{l,\cdot}D_{S_{l,\cdot}}D_{A_{i,\cdot}}Q_{-l}D_{B_{\cdot,i}}D_{S_{l,\cdot}}X_{l,\cdot}^{\mathsf{H}}Q_{ij}\right]S_{l,i}\\
&=\frac{1}{p^2}\sum_{l=1}^{p}\mathbb{E}\left[\mathrm{tr}D_{R_{l,\cdot}}D_{A_{i,\cdot}}Q_{-l}D_{B_{\cdot,i}}D_{S_{l,\cdot}}Q_{ij}\right]S_{l,i}+O(p^{-1}).
\end{aligned}$$

In the remainder of the derivations, we will often use the Cauchy-Schwarz and norm inequalities for more complex terms. We will not further develop them in detail when the result is immediate or close to the previous derivation.

Back to our original sum over (l,l) indices, we are now left to estimating the newly introduced quantity

$$\begin{aligned}
&-\frac{1}{p^2}X_{l,\cdot}D_{R_{l,\cdot}}D_{A_{i,\cdot}}Q_{-l}\left\{\left[D_{S_{l,\cdot}}X'_{l,\cdot}X_{l,\cdot}D_{S_{l,\cdot}}\right]\odot B\right\}\\
&\times\left(I_n+\frac{1}{p}Q_{-l}\left\{\left[D_{S_{l,\cdot}}X'_{l,\cdot}X_{l,\cdot}D_{S_{l,\cdot}}\right]\odot B\right\}\right)^{-1}Q_{-l}D_{B_{\cdot,i}}D_{S_{l,\cdot}}X'_{l,\cdot}.
\end{aligned}$$

This term is delicate as the dependence of the inner-matrix in $X_{l,\cdot}$ remains. Here the main observation to make is the following and depends on the nature of B:

$$\begin{aligned}
\frac{1}{p}\left\{\left[D_{S_{l,\cdot}}X_{l,\cdot}^{\mathsf{H}}X_{l,\cdot}D_{S_{l,\cdot}}\right]\odot B\right\}=&\frac{1}{p}\left\{\left[D_{S_{l,\cdot}}X_{l,\cdot}^{\mathsf{H}}X_{l,\cdot}D_{S_{l,\cdot}}\right]\odot\varepsilon_B 1_n 1_n^{\mathsf{T}}\right\}\\
&+\frac{1}{p}\left\{\left[D_{S_{l,\cdot}}X_{l,\cdot}^{\mathsf{H}}X_{l,\cdot}D_{S_{l,\cdot}}\right]\odot\mathring{B}\right\}\\
&+\frac{1}{p}\left\{\left[D_{S_{l,\cdot}}X_{l,\cdot}^{\mathsf{H}}X_{l,\cdot}D_{S_{l,\cdot}}\right]\odot(b-\varepsilon_B)I_p\right\}
\end{aligned}$$

where we wrote $\mathring{B} = B - \mathbb{E}[B]$ and used $\mathbb{E}[B] = \varepsilon_B 1_n 1_n^{\mathsf{T}} + (b - \varepsilon_B)I_p$.

Remark that

$$\frac{1}{p}\left[D_{S_{l,\cdot}}X_{l,\cdot}^{\mathsf{H}}X_{l,\cdot}D_{S_{l,\cdot}}\right]\odot\mathring{B}=D_{D_{S_{l,\cdot}}X_{l,\cdot}^{\mathsf{H}}}\mathring{B}D_{X_{l,\cdot}D_{S_{l,\cdot}}}$$

the spectral norm of which is bounded, for all large n, p with high probability by $O(\log p/\sqrt{p})$: this is because the spectrum of $\mathring{B}$ follows a semi-circle distribution in the limit with $\|\mathring{B}\|/\sqrt{2n} \to 1$, and $\|D_{X_{l,\cdot}}\|$ is the maximum of n independent Gaussian variables which, uniformly on X cannot grow faster than $O(\sqrt{\log(np)}) = O(\sqrt{\log p})$. This claim is confirmed by a further application of the Nash-Poincaré inequality. Similarly, $\frac{1}{p}\{[D_{S_{l,\cdot}}X_{l,\cdot}^{\mathsf{H}}X_{l,\cdot}D_{S_{l,\cdot}}]\odot(b-\varepsilon_B)I_p\}$ is bounded in norm by $O(\log p/p)$.

With these remarks at hand, we may freely replace B in the expression of $[D_{S_{l,\cdot}} X_{l,\cdot}^{\mathsf{H}} X_{l,\cdot} D_{S_{l,\cdot}}] \odot B$ above by $\varepsilon_B 1_n 1_n^{\mathsf{T}}$, so to obtain

$$
\begin{aligned}
&-\frac{1}{p^2} X_{l,\cdot} D_{S_{l,\cdot}} D_{A_{i,\cdot}} Q_{-l} \left\{ \left[D_{S_{l,\cdot}} X_{l,\cdot}^{\mathsf{H}} X_{l,\cdot} D_{S_{l,\cdot}} \right] \odot B \right\} \\
&\times \left(I_n + \frac{1}{p} Q_{-l} \left\{ \left[D_{S_{l,\cdot}} X_{l,\cdot}^{\mathsf{H}} X_{l,\cdot} D_{S_{l,\cdot}} \right] \odot B \right\} \right)^{-1} Q_{-l} D_{B_{\cdot,i}} D_{S_{l,\cdot}} X_{l,\cdot}^{\mathsf{H}} \\
&= -\frac{\varepsilon_B}{p^2} X_{l,\cdot} D_{R_{l,\cdot}} D_{A_{i,\cdot}} Q_{-l} D_{S_{l,\cdot}} X_{l,\cdot}^{\mathsf{H}} X_{l,\cdot} D_{S_{l,\cdot}} \\
&\times \left(I_n + \frac{\varepsilon_B}{p} Q_{-l} D_{S_{l,\cdot}} X_{l,\cdot}^{\mathsf{H}} X_{l,\cdot} D_{S_{l,\cdot}} \right)^{-1} Q_{-l} D_{B_{\cdot,i}} D_{S_{l,\cdot}} X_{l,\cdot}^{\mathsf{H}} + O_p(\sqrt{\log p}/\sqrt{p}).
\end{aligned}
$$

Using Sherman-Morrison's identity $u^{\mathsf{H}}(A+\lambda u v^{\mathsf{H}})^{-1} = \frac{u^{\mathsf{H}} A^{-1}}{1+\lambda v^{\mathsf{H}} A^{-1} u}$, this further simplifies into

$$
\begin{aligned}
&-\frac{1}{p^2} X_{l,\cdot} D_{S_{l,\cdot}} D_{A_{i,\cdot}} Q_{-l} \left\{ \left[D_{S_{l,\cdot}} X_{l,\cdot}^{\mathsf{H}} X_{l,\cdot} D_{S_{l,\cdot}} \right] \odot B \right\} \\
&\times \left(I_n + \frac{1}{p} Q_{-l} \left\{ \left[D_{S_{l,\cdot}} X_{l,\cdot}^{\mathsf{H}} X_{l,\cdot} D_{S_{l,\cdot}} \right] \odot B \right\} \right)^{-1} Q_{-l} D_{B_{\cdot,i}} D_{S_{l,\cdot}} X_{l,\cdot}^{\mathsf{H}} \\
&= -\frac{\varepsilon_B}{p^2} X_{l,\cdot} D_{S_{l,\cdot}} D_{A_{i,\cdot}} Q_{-l} D_{S_{l,\cdot}} X_{l,\cdot}^{\mathsf{H}} \frac{X_{l,\cdot} D_{S_{l,\cdot}} Q_{-l} D_{B_{\cdot,i}} D_{S_{l,\cdot}} X_{l,\cdot}^{\mathsf{H}}}{1 + \frac{\varepsilon_B}{p} X_{l,\cdot} D_{S_{l,\cdot}} Q_{-l} D_{S_{l,\cdot}} X_{l,\cdot}^{\mathsf{H}}} + O_p(\sqrt{\log p}/\sqrt{p}).
\end{aligned}
$$

The quadratic forms are now all accessible and all converge to their traces at uniform speed $O(\log(p)/\sqrt{p})$ (again by a control of their variances), so that

$$
\begin{aligned}
&-\frac{1}{p^2} X_{l,\cdot} D_{S_{l,\cdot}} D_{A_{i,\cdot}} Q_{-l} \left\{ \left[D_{S_{l,\cdot}} X_{l,\cdot}^{\mathsf{H}} X_{l,\cdot} D_{S_{l,\cdot}} \right] \odot B \right\} \\
&\times \left(I_n + \frac{1}{p} Q_{-l} \left\{ \left[D_{S_{l,\cdot}} X_{l,\cdot}^{\mathsf{H}} X_{l,\cdot} D_{S_{l,\cdot}} \right] \odot B \right\} \right)^{-1} Q_{-l} D_{B_{\cdot,i}} D_{S_{l,\cdot}} X_{l,\cdot}^{\mathsf{H}} \\
&= -\frac{\varepsilon_B}{p^2} \operatorname{tr} D_{S_{l,\cdot}} D_{A_{i,\cdot}} Q_{-l} \frac{\operatorname{tr} D_{B_{\cdot,i}} D_{S_{l,\cdot}}^2 Q_{-l}}{1 + \frac{\varepsilon_B}{p} \operatorname{tr} D_{S_{l,\cdot}} Q_{-l}} + O_p(\sqrt{\log p}/\sqrt{p}).
\end{aligned}
$$

With the same argument as above, one may freely replace Q_{-l} by Q up to a negligible cost of $O(1/\sqrt{p})$ in the above traces. Then, perturbing matrix K so to discard the contribution of $S_{l,\cdot}$ and $B_{\cdot,i}$ also comes at a negligible cost, so that, again with the same perturbation argument, we get

$$
\begin{aligned}
&-\frac{1}{p^2} X_{l,\cdot} D_{S_{l,\cdot}} D_{A_{i,\cdot}} Q_{-l} \left\{ \left[D_{S_{l,\cdot}} X_{l,\cdot}^{\mathsf{H}} X_{l,\cdot} D_{S_{l,\cdot}} \right] \odot B \right\} \\
&\times \left(I_n + \frac{1}{p} Q_{-l} \left\{ \left[D_{S_{l,\cdot}} X_{l,\cdot}^{\mathsf{H}} X_{l,\cdot} D_{S_{l,\cdot}} \right] \odot B \right\} \right)^{-1} Q_{-l} D_{B_{\cdot,i}} D_{S_{l,\cdot}} X_{l,\cdot}^{\mathsf{H}} \\
&= -\frac{\varepsilon_B}{p^2} \operatorname{tr} D_{S_{l,\cdot}} D_{A_{i,\cdot}} Q_{-l} \frac{\varepsilon_B \varepsilon_S \operatorname{tr} Q}{1 + \frac{\varepsilon_B \varepsilon_S}{p} \operatorname{tr} Q} + O_p(\sqrt{\log p}/\sqrt{p}).
\end{aligned}
$$

Summarizing the results above, we then get (with the same necessary controls by the Nash-Poincaré inequality as above),

$$T_1(A,R)+T_2(A,R)=\mathbb{E}\left[\frac{1}{p}[S^{\mathsf{H}}S]_{ii}A_{ii}Q_{ij}\right]-\mathbb{E}\left[\frac{1}{p^2}\sum_{l=1}^{p}\operatorname{tr}\left(D_{S_{l\cdot}}D_{A_{i\cdot}}Q_{-l}D_{B_{\cdot i}}D_{S_{l\cdot}}\right)S_{li}Q_{ij}\right]$$
$$+\frac{1}{p}\sum_{l=1}^{p}\mathbb{E}\left[\frac{\varepsilon_B}{p^2}\operatorname{tr}\left(D_{S_{l,\cdot}}D_{A_{i,\cdot}}Q_{-l}\right)\frac{\varepsilon_B\varepsilon_S\operatorname{tr}Q}{1+\frac{\varepsilon_B\varepsilon_S}{p}\operatorname{tr}Q}S_{li}Q_{ij}\right]+O\left(\frac{\log p}{\sqrt{p}}\right).$$

Term 3: (P^{H}, Z)

We consider now $T_3(A,S)\equiv\mathbb{E}\left[\left[\left(\left[\frac{1}{p}(P\odot S)^{\mathsf{H}}(Z\odot S)\right]\odot A\right)Q\right]_{ij}\right]$:

$$T_3(A,R)=\frac{1}{p}\sum_{l=1}^{p}\sum_{m=1}^{n}P_{li}S_{li}S_{lm}A_{im}\mathbb{E}\left[\frac{\partial Q_{mj}}{\partial \bar{Z}_{lm}}\right]$$

with

$$\begin{aligned}
\mathbb{E}\left[\frac{\partial Q_{cd}}{\partial \bar{Z}_{ab}}\right]&=-\frac{1}{p}\left[Q\left[(E_{ba}\odot S^{\mathsf{H}})(X\odot S)\cdot B\right]Q\right]_{cd}\\
&=-\frac{1}{p}\sum_{l=1}^{n}\sum_{m=1}^{p}Q_{cl}\left[(E_{ba}\odot S^{\mathsf{H}})(X\odot S)\cdot B\right]_{lm}Q_{md}\\
&=-\frac{1}{p}\sum_{l=1}^{n}\sum_{m=1}^{p}\sum_{o=1}Q_{cl}\delta_{bl}\delta_{ao}S^{\mathsf{H}}_{lo}[X\odot S]_{om}B_{lm}Q_{md}\\
&=-\frac{1}{p}\sum_{l=1}^{n}\sum_{m=1}^{p}Q_{cl}\delta_{bl}S^{\mathsf{H}}_{la}[X\odot S]_{am}B_{lm}Q_{md}\\
&=-\frac{1}{p}\sum_{m=1}^{p}Q_{cb}S^{\mathsf{H}}_{ba}[X\odot S]_{am}B_{bm}Q_{md}\\
&=-\frac{1}{p}Q_{cb}S^{\mathsf{H}}_{ba}\left[(X\odot S)D_{B_{b\cdot}}Q\right]_{ad}.
\end{aligned}$$

As a consequence,

$$\begin{aligned}
T_3(A,R)&=-\frac{1}{p^2}\sum_{l=1}^{p}\sum_{m=1}^{n}P_{li}S_{li}S_{lm}A_{im}Q_{mm}S_{ml}\mathbb{E}\left[(X\odot S)D_{B_{m\cdot}}Q\right]_{lj}\\
&=-\frac{1}{p^2}\sum_{l=1}^{p}P_{li}S_{li}\mathbb{E}\left[(X\odot S)\left(\sum_{m=1}^{p}D_{B_{m\cdot}}S_{lm}A_{im}Q_{mm}S_{ml}\right)Q\right]_{lj}\\
&=-\frac{1}{p^2}\sum_{l=1}^{p}P_{li}S_{li}\mathbb{E}\left[(X\odot S)D_{\{\operatorname{tr}QD_{B_{\cdot s}}D_{S_{l\cdot}}D_{A_{i\cdot}}\}_{s=1}^{n}}Q\right]_{lj}\\
&=-\frac{1}{p^2}\sum_{l=1}^{p}\mathbb{E}\left[[(T\odot S)^{\mathsf{H}}]_{il}\left[(X\odot S)D_{\{\operatorname{tr}QD_{B_{\cdot s}}D_{S_{l\cdot}}D_{A_{i\cdot}}\}_{s=1}^{n}}Q\right]_{lj}\right].
\end{aligned}$$

At this stage in the calculus, it is necessary to study the normalized trace

$$\frac{1}{p}\mathrm{tr}QD_{B_{\cdot s}}D_{A_{i\cdot}}D_{S_{l\cdot}}.$$

Note that, unless $B_{\cdot s}^{\mathsf{H}} = A_{i\cdot}$ (which could only occur for one value of s; typically $s = i$ if we take $A = B$), using similar perturbation arguments in the large n, p regime as above (the impact of column $B_{\cdot s}$ is negligible in Q), we obtain

$$\frac{1}{p}\mathrm{tr}QD_{B_{\cdot s}}D_{A_{i\cdot}}D_{S_{l\cdot}} = \frac{\varepsilon_B}{p}\mathrm{tr}QD_{A_{i\cdot}}D_{S_{l\cdot}} + o_p(1).$$

As such, $D_{\{\mathrm{tr}QD_{B_{\cdot s}}D_{A_{i\cdot}}D_{S_{l\cdot}}\}_{s=1}^p}$ is asymptotically close to a scaled identity matrix *depending on* i and we may then rewrite

$$T_3(A,R) = -\frac{\varepsilon_B}{p}\mathbb{E}\left[(P\odot R)^{\mathsf{H}} D_{\{\frac{1}{p}\mathrm{tr}QD_{A_{i\cdot}}D_{S_{l\cdot}}\}_{l=1}^p}(X\odot S)Q\right]_{ij} + o(1).$$

This further boils down to

$$\begin{aligned}T_3(A,S) &= -\frac{\varepsilon_B\varepsilon_S}{p}\mathbb{E}\left[D_{\{\frac{1}{p}\mathrm{tr}D_{A_{i\cdot}}Q\}_{i=1}^n}(P\odot S)^{\mathsf{H}}(X\odot S)Q\right]_{ij} + o(1)\\ &= -\frac{\varepsilon_B\varepsilon_S}{p}\mathbb{E}\left[\left\{\left[(P\odot S)^{\mathsf{H}}(X\odot S)\right]\odot d_{\{\frac{1}{p}\mathrm{tr}D_{A_{i\cdot}}Q\}_{i=1}^n}1_n^{\mathsf{T}}\right\}Q\right]_{ij} + o(1)\end{aligned}$$

where d_v is the (column) vector composed of the elements v_i.

Term 4: (P^{H}, P)

We finally add up the easiest term

$$T_4(A,R) \equiv \mathbb{E}\left[\left[\left(\left[\frac{1}{p}(P\odot S)^{\mathsf{H}}(P\odot S)\right]\odot A\right)Q\right]_{ij}\right].$$

Collecting the terms

Collecting all terms $T_i(A,S)$, it appears that the desired evaluation of $\mathbb{E}[Q_{ij}]$, which we obtain through that of

$$\mathbb{E}\left[\frac{1}{p}(X\odot S)^{\mathsf{H}}(X\odot S)Q\right]_{ij}$$

gives rise to two sets of “new” terms:

1. the traces

$$\mathrm{tr}\left(D_{S_{l\cdot}}D_{A_{i\cdot}}Q_{-l}\right)$$

2. the matrix expectation

$$\mathbb{E}\left[\left\{\left[(P\odot S)^{\mathsf{H}}(X\odot S)\right]\odot d_{\{\frac{1}{p}\mathrm{tr}D_{A_{i\cdot}}Q\}_{i=1}^n}1_n^{\mathsf{T}}\right\}Q\right]_{ij}.$$

Using perturbation arguments, the traces are easily analyzed and all lead to scaled versions of $\mathrm{tr}Q$ in the limit, thereby effectively not providing any new term.

The matrix expectation is less immediate and must be appropriately used to "close the loop" of the estimate of $\mathbb{E}[Q_{ij}]$. Specifically,

- letting $A = B$ in the initial equation leads to evaluating

$$\mathbb{E}\left[\left\{\left[(P\odot S)^{\mathsf{H}}(X\odot S)\right]\odot d_{\{\frac{1}{p}\mathrm{tr}D_{B_{i\cdot}}Q\}_{i=1}^{n}}1_n^{\mathsf{T}}\right\}Q\right]_{ij}$$

  which, from the fact that $\frac{1}{p}\mathrm{tr}D_{B_{i\cdot}}Q = \frac{\varepsilon_B}{p}\mathrm{tr}Q + o_p(1)$, is essentially

$$\mathbb{E}\left[\frac{\varepsilon_B}{p}\mathrm{tr}Q\left\{\left[(P\odot S)^{\mathsf{H}}(X\odot S)\right]\odot 1_n1_n^{\mathsf{T}}\right\}Q\right]_{ij}$$

  a term that we thus need to evaluate;

- letting then $A = 1_n1_n^{\mathsf{T}}$ leads instead to

$$\begin{aligned}&\mathbb{E}\left[\left\{\left[(P\odot S)^{\mathsf{H}}(X\odot S)\right]\odot d_{\{\frac{1}{p}\mathrm{tr}D_{1_n}Q\}_{i=1}^{n}}1_n^{\mathsf{T}}\right\}Q\right]_{ij}\\ &=\mathbb{E}\left[((1/p)\mathrm{tr}Q)\left\{\left[(P\odot S)^{\mathsf{H}}(X\odot S)\right]\odot 1_n1_n^{\mathsf{T}}\right\}Q\right]_{ij}\end{aligned}$$

  from which we may now close the loop.

Precisely, combining terms from $T_3(1_n1_n^{\mathsf{T}},S)$ and $T_4(1_n1_n^{\mathsf{T}},S)$, we first find that

$$\begin{aligned}&\left\{\left[\frac{1}{p}(P\odot S)^{\mathsf{H}}(X\odot S)\right]\odot 1_n1_n^{\mathsf{T}}\right\}Q\\ \leftrightarrow &-\left(\frac{\varepsilon_B\varepsilon_S}{p}\mathrm{tr}Q\right)\left\{\left[\frac{1}{p}(T\odot S)^{\mathsf{H}}(X\odot S)\right]\odot 1_n1_n^{\mathsf{T}}\right\}Q\\ &+\left\{\left[\frac{1}{p}(P\odot S)^{\mathsf{H}}(P\odot S)\right]\odot 1_n1_n^{\mathsf{T}}\right\}Q\end{aligned}$$

so that, letting $m(z)\in\mathbb{C}$ be such that $\frac{1}{n}\mathrm{tr}Q\leftrightarrow m(z)$,

$$\begin{aligned}&\left\{\left[\frac{1}{p}(P\odot S)^{\mathsf{H}}(X\odot S)\right]\odot 1_n1_n^{\mathsf{T}}\right\}Q\\ \leftrightarrow\ &\frac{1}{1+c_0^{-1}\varepsilon_B\varepsilon_S m(z)}\left\{\left[\frac{1}{p}(P\odot S)^{\mathsf{H}}(P\odot S)\right]\odot 1_n1_n^{\mathsf{T}}\right\}Q.\end{aligned}$$

Next, combining all $T_i(B,S)$, we get, using in particular $\frac{1}{p}[S^{\mathsf{H}}S]_{ii}\to\varepsilon_S$ almost surely,

$$\begin{aligned}Q\leftrightarrow &-\frac{1}{z}I_n+\frac{\varepsilon_S b}{z}Q-\varepsilon_S^2\varepsilon_B c_0^{-1}m(z)Q+\frac{\varepsilon_B^3\varepsilon_S^3c_0^{-2}m(z)^2}{1+c_0^{-1}\varepsilon_B\varepsilon_S m(z)}Q\\ &-\frac{\varepsilon_B^2\varepsilon_S c_0^{-1}m(z)}{1+c_0^{-1}\varepsilon_B\varepsilon_S m(z)}\frac{1}{p}(P\odot S)^{\mathsf{H}}(P\odot S)Q+\frac{1}{p}\left\{\left[(P\odot S)^{\mathsf{H}}(P\odot S)\right]\odot B\right\}Q.\end{aligned}$$

To complete the proof, we now need to handle the terms $(P\odot S)^{\mathsf{H}}(P\odot S)Q$ and $\frac{1}{p}\{[(P\odot S)^{\mathsf{H}}(P\odot S)]\odot B\}Q$ and relate then to Q directly. To this end, similar to previously, let us write $S=\varepsilon_S 1_p 1_n^{\mathsf{T}}+\mathring{S}$, $B=\varepsilon_B 1_n 1_n^{\mathsf{T}}+\mathring{B}$ and $P=\sum_{\ell=1}^k L_{\cdot,\ell}V_{\cdot,\ell}^{\mathsf{H}}$. Then, since we imposed the entries V_{ij} to be essentially of order $1/\sqrt{n}$,[2] observe that the matrix $\frac{1}{\sqrt{p}}L_{\cdot,\ell}V_{\cdot,\ell}^{\mathsf{H}}\odot\mathring{S}=\frac{1}{\sqrt{p}}D_{L_{\cdot,\ell}}\mathring{S}D_{V_{\cdot,\ell}^{\mathsf{H}}}$ has operator norm of order $1/\sqrt{p}$. The same reasoning applies to B, so that we may rewrite

$$\begin{aligned}zQ\leftrightarrow &-I_n+\varepsilon_S bQ-\varepsilon_S^2\varepsilon_B c_0^{-1}m(z)Q+\frac{\varepsilon_B^3\varepsilon_S^3 c_0^{-2}m(z)^2}{1+c_0^{-1}\varepsilon_B\varepsilon_S m(z)}Q\\ &-\frac{\varepsilon_B^2\varepsilon_S^3 c_0^{-1}m(z)}{1+c_0^{-1}\varepsilon_B\varepsilon_S m(z)}\frac{1}{p}P^{\mathsf{H}}PQ+\varepsilon_B\varepsilon_S^2\frac{1}{p}P^{\mathsf{H}}PQ\\ =&-I_n+\varepsilon_S bQ-\varepsilon_S^2\varepsilon_B c_0^{-1}m(z)Q+\frac{\varepsilon_B^3\varepsilon_S^3 c_0^{-2}m(z)^2}{1+c_0^{-1}\varepsilon_B\varepsilon_S m(z)}Q\\ &+\frac{\varepsilon_B\varepsilon_S^2}{1+c_0^{-1}\varepsilon_B\varepsilon_S m(z)}\frac{1}{p}P^{\mathsf{H}}PQ\end{aligned}$$

Further using

$$\frac{1}{p}P^{\mathsf{H}}P=\frac{1}{p}VL^{\mathsf{H}}LV^{\mathsf{H}}=V\mathcal{L}V^{\mathsf{H}}+o_{\|\cdot\|}(1)$$

along with the fact that $V^{\mathsf{H}}V=I_k$, we finally get the deterministic equivalent for Q

$$\begin{aligned}Q\leftrightarrow\Bigg[&(\varepsilon_S b-z)I_n-c_0^{-1}\varepsilon_S^2\varepsilon_B m(z)I_n+\frac{c_0^{-2}\varepsilon_S^3\varepsilon_B^3 m(z)^2}{1+\varepsilon_B\varepsilon_S c_0^{-1}m(z)}I_n\\ &+\frac{c_0^{-1}\varepsilon_S^2\varepsilon_B}{1+\varepsilon_B\varepsilon_S c_0^{-1}m(z)}V\mathcal{L}V^{\mathsf{H}}\Bigg]^{-1}.\end{aligned}$$

In particular, recalling that $m(z)$ is an asymptotic equivalent for $\frac{1}{n}\mathrm{tr}Q$, we have

$$m(z)=\frac{1}{(\varepsilon_S b-z)-c_0^{-1}\varepsilon_S^2\varepsilon_B m(z)+\frac{c_0^{-2}\varepsilon_S^3\varepsilon_B^3 m(z)^2}{1+\varepsilon_B\varepsilon_S c_0^{-1}m(z)}}$$

which unfolds from $V\mathcal{L}V^{\mathsf{H}}$ being of finite rank k (so that it does not affect the limiting normalized trace) and thus provides a *deterministic equivalent*. Equivalently, this is

$$z=\varepsilon_S b-\frac{1}{m(z)}-c_0^{-1}\varepsilon_B\varepsilon_S^2 m(z)+\frac{c_0^{-2}\varepsilon_B^3\varepsilon_S^3 m(z)^2}{1+c_0^{-1}\varepsilon_B\varepsilon_S m(z)}$$

[2]More specifically, it is enough to assume that $V_{ij}^2=o(1/\sqrt{n})$.

which, integrated in the previous expression of the random equivalent of Q, provides the shorter and final forms of the *deterministic equivalent*:

$$\boxed{Q \leftrightarrow m(z)\left[I_n + \frac{c_0^{-1}\varepsilon_S^2\varepsilon_B m(z)}{1+\varepsilon_B\varepsilon_S c_0^{-1}m(z)} V\mathcal{L}V^{\mathsf{H}}\right]^{-1}}$$

or possibly more expressively

$$Q \leftrightarrow m(z)V_\perp V_\perp^{\mathsf{H}} + m(z)V\left[I_k + \frac{c_0^{-1}\varepsilon_S^2\varepsilon_B m(z)}{1+\varepsilon_B\varepsilon_S c_0^{-1}m(z)}\mathcal{L}\right]^{-1} V^{\mathsf{H}}$$

where in this last equality $V_\perp$ is an orthonormal basis completing V (this last result follows from Woodbury's matrix inverse identity).

3.3 Proof of Theorem 2

With the previous result available, the proof of Theorem 2 follows from a classical random matrix approach.

Let $\mathcal{L} = \sum_{i=1}^{\bar{k}} \ell_i \mathcal{V}_i \mathcal{V}_i^{\mathsf{H}}$ be the spectral decomposition of $\mathcal{L}$ with $\mathcal{V}_i \in \mathbb{C}^{k\times L_i}$ isometric and such that $\Pi_i = \mathcal{V}_i\mathcal{V}_i^{\mathsf{H}}$ is a projector on the eigenspace associated to eigenvalue ℓ_i which we assume of multiplicity L_i greater or equal to 1.

Then, assuming asymptotic separability (that is, the existence of a spike associated to ℓ_i), such that the resulting associated eigenvalue(s) $\lambda_j,\ldots,\lambda_{j+L_i-1}$ of K converge to ρ_i with associated eigenspace $\hat{\mathcal{V}}_i$, we have, in the large n,p limit, almost surely (the limit is needed to ensure that $\lambda_j,\ldots,\lambda_{j+L_i-1}$ fall into the contour Γ_{ρ_i}),

$$\begin{aligned}
\hat{\mathcal{V}}_i\hat{\mathcal{V}}_i^{\mathsf{H}} &= -\frac{1}{2\pi\imath}\oint_{\Gamma_{\rho_i}} Q(z)dz \\
&\leftrightarrow -\frac{1}{2\pi\imath}\oint_{\Gamma_{\rho_i}} m(z)V\left[I_k + \frac{c_0^{-1}\varepsilon_S^2\varepsilon_B m(z)}{1+\varepsilon_B\varepsilon_S c_0^{-1}m(z)}\mathcal{L}\right]^{-1} V^{\mathsf{H}}dz
\end{aligned}$$

for Γ_x a positively oriented complex contour surrounding x closely. By residue calculus, we then find that

$$\begin{aligned}
\hat{\mathcal{V}}_i\hat{\mathcal{V}}_i^{\mathsf{H}} &\leftrightarrow -\lim_{z\in\mathbb{C}\to\rho_i}(z-\rho_i)m(z)U\left[I_k + \frac{c_0^{-1}\varepsilon_S^2\varepsilon_B m(z)}{1+\varepsilon_B\varepsilon_S c_0^{-1}m(z)}\mathcal{L}\right]^{-1} V^{\mathsf{H}} \\
&\leftrightarrow -\lim_{z\in\mathbb{C}\to\rho_i}(z-\rho_i)m(z)V\mathcal{V}_i\left[1 + \frac{c_0^{-1}\varepsilon_S^2\varepsilon_B m(z)}{1+\varepsilon_B\varepsilon_S c_0^{-1}m(z)}\ell_i\right]^{-1} \mathcal{V}_i^{\mathsf{H}}V^{\mathsf{H}}
\end{aligned}$$

where we exploited the fact that the denominator above must vanish as $z\to\rho_i$, thereby in passing *defining* ρ_i.

Specifically, we find that ρ_i, the limit of the empirical eigenvalues of K associated with ℓ_i, is solution to

$$1+\ell_i\frac{c_0^{-1}\varepsilon_B\varepsilon_S^2 m(\rho_i)}{1+c_0^{-1}\varepsilon_B\varepsilon_S m(\rho_i)} = 0 \Leftrightarrow \frac{1}{m(\rho_i)} = -c_0^{-1}\varepsilon_B\varepsilon_S(1+\varepsilon_S\ell_i).$$

In particular, we have the following convenient relation for what follows:

$$1 + c_0^{-1}\varepsilon_B\varepsilon_S m(\rho_i) = \frac{\varepsilon_S \ell_i}{1+\varepsilon_S\ell_i}.$$

Exploiting the relation $z = f(m(z))$ above, applied to $z = \rho_i$, this leads to the explicit value of the isolated "spike" ρ_i:

$$\boxed{\rho_i = \varepsilon_S b + c_0^{-1}\varepsilon_B\varepsilon_S(1+\varepsilon_S\ell_i) + \frac{\varepsilon_S}{1+\varepsilon_S\ell_i} + \frac{\varepsilon_B}{\ell_i(1+\varepsilon_S\ell_i)}.}$$

By l'Hospital's rule (or equivalently a first order Taylor expansion of both numerator and denominator in the inverse formula of the residue), we then have

$$\hat{\mathcal{U}}_i\hat{\mathcal{V}}_i^{\mathsf{H}} \leftrightarrow -V\mathcal{V}_i \frac{m(\rho_i)(1+c_0^{-1}\varepsilon_B\varepsilon_S m(\rho_i))^2}{\ell_i c_0^{-1}\varepsilon_B\varepsilon_S^2 m'(\rho_i)}\mathcal{V}_i^{\mathsf{H}}V^{\mathsf{H}}$$

where, exploiting the defining equation of $m(z)$, we find after mere algebraic calculus

$$\begin{aligned}\frac{1}{m'(z)} &= \frac{1}{m(z)^2} - c_0^{-1}\varepsilon_B\varepsilon_S^2 + c_0^{-2}\varepsilon_B^3\varepsilon_S^3 m(z)\frac{2+c_0^{-1}\varepsilon_B\varepsilon_S m(z)}{(1+c_0^{-1}\varepsilon_B\varepsilon_S m(z))^2} \\ &= \frac{1}{m(z)^2} - c_0^{-1}\varepsilon_B\varepsilon_S^2 + \frac{c_0^{-2}\varepsilon_B^3\varepsilon_S^3 m(z)}{1+c_0^{-1}\varepsilon_B\varepsilon_S m(z)} + \frac{c_0^{-2}\varepsilon_B^3\varepsilon_S^3 m(z)}{(1+c_0^{-1}\varepsilon_B\varepsilon_S m(z))^2}.\end{aligned}$$

Altogether, we finally find the fully explicit deterministic equivalent

$$\boxed{\hat{\mathcal{V}}_i\hat{\mathcal{V}}_i^{\mathsf{H}} \leftrightarrow \left(\frac{\varepsilon_S\ell_i}{1+\varepsilon_S\ell_i} - \frac{\varepsilon_S\ell_i}{c_0^{-1}\varepsilon_B(1+\varepsilon_S\ell_i)^3} - \frac{1}{c_0^{-1}(1+\varepsilon_S\ell_i)^3} - \frac{1}{c_0^{-1}\varepsilon_S\ell_i(1+\varepsilon_S\ell_i)^2}\right)V\mathcal{V}_i\mathcal{V}_i^{\mathsf{H}}V^{\mathsf{H}}.}$$

Equating the term in parentheses to zero then provides the phase transition condition: indeed, the asymptotic alignment of population and sample eigenspaces vanishes right at the position where the spike ρ_i escapes the limiting continuous part of the support of the eigenvalues of K. The phase transition for ρ_i then occurs when ℓ_i satisfies:

$$\boxed{0 = \ell_i^4 + \frac{2}{\varepsilon_S}\ell_i^3 + \frac{1}{\varepsilon_S^2}\left(1 - \frac{c_0}{\varepsilon_B}\right)\ell_i^2 - \frac{2c_0}{\varepsilon_S^3}\ell_i - \frac{c_0}{\varepsilon_S^4} \equiv F(\ell_i).}$$

This expression of F is convenient as it takes the form of a polynomial of order 4 with unit leading monomial coefficient. It then suffices to remark that the asymptotic alignment expression above expresses as $F(\rho_i)/\ell_i/(1+\varepsilon_S\ell_i)^3$ to conclude the proof of Theorem 2.