# Asymptotics of diagonal elements of projection matrices under many instruments/regressors

Stanislav Anatolyev
CERGE-EI, Czech Republic & New Economic School, Russia
Pavel Yaskov
Steklov Mathematical Institute of RAS & NUST 'MISIS'

## A Online Appendix: Technical lemmas and proofs

**Proof of Lemma 3.1.** The result follows from Corollary 3.4 in Yaskov (2014). $\square$

**Lemma A.1** *Let $\mathbb{P}(\lambda_{\min}(Z'_{-1}Z_{-1}) < \delta n) = o(1/n)$, and let $\mathbb{P}(\|z_l\| > K\sqrt{l}) = o(1/l)$ for some $\delta, K > 0$. If (MI) holds, then there exists a constant $C \in (0,1)$ such that, as $n \to \infty$, we have*

$$\mathbb{P}(P_{kk} \leq C \textit{ for all } 1 \leq k \leq n) \to 1.$$

**Proof of Lemma A.1.** Denoting $M_n = \max_{1\leq k\leq n} \|z_k\|^2$ and $\lambda_n = \min_{1\leq k\leq n} \lambda_{\min}(Z'_{-k}Z_{-k})$ we can see that $\mathbb{P}(M_n > K^2 l) \leq \sum_{k=1}^n \mathbb{P}(\|z_k\| > K\sqrt{l}) = n\mathbb{P}(\|z_l\| > K\sqrt{l}) = o(1)$ and

$$\mathbb{P}(\lambda_n < \delta n) \leq \sum_{k=1}^{n} \mathbb{P}(\lambda_{\min}(Z'_{-k}Z_{-k}) < \delta n) = n\mathbb{P}(\lambda_{\min}(Z'_{-1}Z_{-1}) < \delta n) = o(1).$$

In addition, by the Sherman-Morrison formula, we have

$$P_{kk} = f(z'_k(Z'_{-k}Z_{-k})^{-1}z_k) \leq f(\|z_k\|^2/\lambda_{\min}(Z'_{-k}Z_{-k})) \leq f(M_n/\lambda_n) \leq f(K^2 l/(\delta n))$$

on $\{M_n \leq K^2 l, \lambda_n \geq \delta n\}$ for all $1 \leq k \leq n$, where $f(x) = x/(1+x)$, $x \geq 0$. Since $l = O(n)$ and $\mathbb{P}(M_n \leq K^2 l, \lambda_n \geq \delta n) = 1 - o(1)$, we get the desired result. $\square$

**Lemma A.2** *For $l > k \geq 1$, let $u_{l-k}$ and $v_k$ be random vectors in $\mathbb{R}^{l-k}$ and $\mathbb{R}^k$, respectively, such that $\mathbb{E}u_{l-k}u'_{l-k} = I_{l-k}$, $\mathbb{E}v_k v'_k = I_k$, and $(\{v_k\}_{k\geq 1}, d)$ satisfies Property P for some $d$. If $z_l = (u'_{l-k}, v'_k)'$ and $k = k(l) = l - o(l)$, $l \to \infty$, then $(\{z_l\}_{l\geq 2}, d)$ satisfies Property P.*

**Proof of Lemma A.2.** For each $l > 1$, let $A_l$ be an $l \times l$ symmetric positive semi-definite matrix such that $\lambda_{\max}(A_l)$ is bounded over $l$. Write $A_l$ as

$$A_l = \begin{pmatrix} B_m & C_{mk} \\ C'_{mk} & D_k \end{pmatrix},$$

where $m = l - k$, $B_m$, $C_{mk}$, and $D_k$ are $m \times m$, $m \times k$, and $k \times k$ matrices, respectively. Since $\lambda_{\max}(A_l)$ is uniformly bounded, $m = o(l)$, and $(\{v_k\}_{k \geq 1}, d)$ satisfies Property P, we have

$$\frac{z'_l A_l z_l - d\,\mathrm{tr}(A_l)}{l} = \frac{u'_m B_m u_m + 2u'_m C_{mk} v_k - d\,\mathrm{tr}(B_m)}{m} \cdot o(1) + o_p(1).$$

Therefore we only need to show that $d\,\mathrm{tr}(B_m)/m$, $u'_m B_m u_m/m$, and $u'_m C_{mk} v_k/m$ are bounded in probability.

Any random variable $d$ is bounded in probability. In addition, $\mathrm{tr}(B_m) \leq m\lambda_{\max}(B_m) \leq m\lambda_{\max}(A_l) = O(m)$. We also have $\mathbb{E}u'_m B_m u_m = \mathrm{tr}(B_m) = O(m)$. By the Cauchy inequality, $2\mathbb{E}|u'_m C_{mk} v_k| \leq \mathbb{E}u'_m u_m + \mathbb{E}(C_{mk} v_k)'(C_{mk} v_k) = m + \mathrm{tr}(C'_{mk} C_{mk}) = m + \mathrm{tr}(C_{mk} C'_{mk}) \leq m + m\lambda_{\max}(C_{mk} C'_{mk}) \leq m + m\lambda_{\max}(A_l A'_l) = m + m\lambda_{\max}(A_l)^2 = O(m)$. □

**Lemma A.3** *Let $\lambda_{\max}(\mathbb{E}z_l z'_l) \leq \lambda$ and $\mathbb{E}|a' z_l|\mathbb{I}_{\{d>0\}} \geq c$ for some $\lambda, c > 0$, any $l \geq 1$, and all $a \in \mathbb{R}^l$ with $a'a = 1$. If Property P holds for $(\{z_l\}_{l \geq 1}, d)$ and $\alpha < \mathbb{P}(d > 0)$, then Condition A holds.*

**Proof of Lemma A.3.** The result follows from Corollary 3.2 in Yaskov (2016). □

**Lemma A.4** *Let $z_l$ be a random vector in $\mathbb{R}^l$ for any $l \geq 1$ and $\{z_{lk}\}_{k=1}^n$ be IID copies of $z_l$. If $\varepsilon_n \sqrt{n} \to \infty$, then $S_n - \mathbb{E}S_n \xrightarrow{p} 0$ as $n \to \infty$ for all $l = O(n)$, where*

$$S_n = \mathrm{tr}\left(\sum_{k=1}^n z_{lk} z'_{lk} + \varepsilon_n n I_l\right)^{-1}.$$

**Proof of Lemma A.4.** In what follows, we will write $z_k$ instead of $z_{lk}$. Denote $\mathbb{E}[\cdot|z_k, \ldots, z_n]$, $1 \leq k \leq n$, by $\mathbb{E}_k$, and $\mathbb{E}$ by $\mathbb{E}_{n+1}$. Then

$$S_n - \mathbb{E}S_n = \sum_{k=1}^n (\mathbb{E}_k - \mathbb{E}_{k+1}) S_n = \sum_{k=1}^n (\mathbb{E}_k - \mathbb{E}_{k+1})(S_n - S_n^k),$$

where $S_n^k = \mathrm{tr}(C_k + \varepsilon_n n I_l)^{-1}$ and $C_k = \sum_{j \neq k} z_j z'_j$. By (1), we get

$$|S_n - S_n^k| = \left|\frac{z'_k (C_k + \varepsilon_n n I_l)^{-2} z_k}{1 + z'_k (C_k + \varepsilon_n n I_l)^{-1} z_k}\right| \leq \frac{1}{\varepsilon_n n}\left|\frac{z'_k (C_k + \varepsilon_n n I_l)^{-1} z_k}{1 + z'_k (C_k + \varepsilon_n n I_l)^{-1} z_k}\right| \leq \frac{1}{\varepsilon_n n}.$$

Since $\{(\mathbb{E}_k - \mathbb{E}_{k+1})(S_n - S_n^k)\}_{k=1}^n$ is a martingale difference sequence,

$$\mathbb{E}(S_n - \mathbb{E}S_n)^2 = \sum_{k=1}^n \mathbb{E}|(\mathbb{E}_k - \mathbb{E}_{k+1})(S_n - S_n^k)|^2 \leq \frac{n}{(\varepsilon_n n)^2} = o(1).$$

Hence, we obtain the desired result. $\square$

**Lemma A.5** *Let $l = l(n) \to \infty$, $\lambda^*_{\min}(Z'Z)/\sqrt{n} \xrightarrow{p} \infty$, and $\lambda_{\max}(\mathbb{E}z_l z_l') = O(1)$ as $n \to \infty$. Then for any continuous function $f$ on $[0,1]$,*

$$\frac{1}{n}\sum_{k=1}^n f(P_{kk}) - \mathbb{E}f(P_{11}) \xrightarrow{p} 0.$$

**Proof of Lemma A.5.** Any continuous function on $[0,1]$ could be approximated by a smooth function. Therefore, we may consider only smooth functions for $f$. In what follows, we will omit the index $l$ and write $z_i$ instead of $z_{li}$. The proof consists of verification of several claims.

*Claim 1.* There is a sequence $\lambda_n > 0$ such that $\lambda_n \to \infty$ and $n^{-1}\sum_{i=1}^n [f(P_{ii}) - f_i] \xrightarrow{p} 0$, where $f_i = f(z_i'(Z'Z + \lambda_n I_l)^{-1} z_i)$.

Since $\lambda^*_{\min}(Z'Z) \xrightarrow{p} \infty$, there are $\lambda_n > 0$ that grow to infinity slower than $\lambda^*_{\min}(Z'Z)$ (i.e. $\lambda_n/\lambda^*_{\min}(Z'Z) \xrightarrow{p} 0$). If $Z'Z$ is non-degenerate, then $\lambda_{\min}(Z'Z) = \lambda^*_{\min}(Z'Z)$ and

$$\begin{aligned}
|z_i'(Z'Z)^{-1}z_i - z_i'(Z'Z + \lambda_n I_l)^{-1} z_i| &= \lambda_n z_i'(Z'Z)^{-1}(Z'Z + \lambda_n I_l)^{-1} z_i \\
&\leq \lambda_n\, z_i'(Z'Z)^{-1/2}(Z'Z + \lambda_n I_l)^{-1}(Z'Z)^{-1/2} z_i \\
&\leq \frac{\lambda_n}{\lambda_{\min}(Z'Z + \lambda_n I_l)} z_i'(Z'Z)^{-1} z_i \\
&\leq \frac{\lambda_n}{\lambda^*_{\min}(Z'Z)}. \qquad (9)
\end{aligned}$$

We now show that the last inequalities still hold for degenerate $Z'Z$. There is an $l \times l$ orthogonal matrix $C$ and an $l \times l$ diagonal matrix $D$ such that $Z'Z = CDC'$. Therefore, setting $v_i = C'z_i$ and $V$ to be a matrix with rows $v_1', \ldots, v_n'$, we see that $P_{ii} = v_i'(V'V)^+ v_i$ and $D = V'V = \sum_{i=1}^n v_i v_i'$. In particular, if $d_k$, the $k^{th}$ diagonal entry of $D$, is zero, then the $k^{th}$ entry of each $v_i$ is zero. Assume without loss of generality that $d_1 \geq \ldots \geq d_m > d_{m+1} = \ldots = d_l = 0$ for some $m < l$. Then for all $i$, $v_i = (u_i', 0, \ldots, 0)'$ with $l - m$ zeros for some

$u_i \in \mathbb{R}^m$. As a result, $P_{ii} = u_i'(U'U)^{-1}u_i$ and $\lambda_{\min}(U'U) = \lambda^*_{\min}(V'V) = \lambda^*_{\min}(Z'Z)$, where $U$ is a matrix with rows $u_1', \dots, u_n'$. By (9),

$$\begin{aligned}|z_i'(Z'Z)^+z_i - z_i'(Z'Z+\lambda_n I_l)^{-1}z_i| &= |v_i'(V'V)^+v_i - v_i'(V'V+\lambda_n I_l)^{-1}v_i| \\ &= |u_i'(U'U)^{-1}u_i - u_i'(U'U+\lambda_n I_m)^{-1}u_i| \\ &\leq \frac{\lambda_n}{\lambda_{\min}(U'U)} = \frac{\lambda_n}{\lambda^*_{\min}(Z'Z)}.\end{aligned}$$

Because of the smoothness of $f$, the latter yields Claim 1.

*Claim 2.* $n^{-1}\sum_{i=1}^n [f_i - \mathbb{E}_{-i}f_i] \xrightarrow{p} 0$, where $\mathbb{E}_{-i} = \mathbb{E}[\cdot|z_j, j \neq i]$.

Since $|f_i|$ is bounded and $\{f_i - \mathbb{E}_{-i}f_i\}_{i=1}^n$ are exchangeable random variables,

$$\mathbb{E}\left|\frac{1}{n}\sum_{i=1}^n [f_i - \mathbb{E}_{-i}f_i]\right|^2 = O(n^{-1}) + O(1)\cdot \mathbb{E}[f_1 - \mathbb{E}_{-1}f_1][f_2 - \mathbb{E}_{-2}f_2].$$

Hence, we only need to show that $\mathbb{E}[f_1 - \mathbb{E}_{-1}f_1][f_2 - \mathbb{E}_{-2}f_2] = o(1)$.

By (2),

$$z_i'(Z'Z+\lambda_n I_l)^{-1}z_i = g(z_i(Z_{-i}'Z_{-i}+\lambda_n I_l)^{-1}z_i)$$

with $g(x) = x/(1+x)$, $x \geq 0$, and $Z_{-i}$ is obtained from $Z$ by deleting its $i^{th}$ row. In addition, the function $h(x) = f(g(x))$ is second-order smooth on $\mathbb{R}_+$, and there is $C_0 > 0$ such that $|h^{(k)}(x)|^2 \leq C_0$ on $\mathbb{R}_+$ for each $k = 0, 1$. Put $f_{ij} = h(z_i'(Z_{-ij}'Z_{-ij}+\lambda_n I_l)^{-1}z_i)$ and $\mathbb{E}_{-ij} = \mathbb{E}[\cdot|Z_{-ij}]$, $i \neq j$, for $Z_{-ij}$ ($= Z_{-ji}$) that is obtained by deleting $i^{th}$ and $j^{th}$ rows in $Z$. Since

$$\mathbb{E}[f_{12} - \mathbb{E}_{-12}f_{12}][f_{21} - \mathbb{E}_{-12}f_{21}] = \mathbb{E}\left(\mathbb{E}_{-12}[f_{12} - \mathbb{E}_{-12}f_{12}][f_{21} - \mathbb{E}_{-12}f_{21}]\right) = 0$$

and $\mathbb{E}_{-1}f_{12} = \mathbb{E}_{-12}f_{12} = \mathbb{E}_{-12}f_{21} = \mathbb{E}_{-2}f_{21}$, it follows from Claim 2 and Claim 3 below that $\mathbb{E}[f_1 - \mathbb{E}_{-1}f_1][f_2 - \mathbb{E}_{-2}f_2] = o(1)$. Indeed,

$$\begin{aligned}|\mathbb{E}[f_1 - \mathbb{E}_{-1}f_1][f_2 - \mathbb{E}_{-2}f_2]| &= |\mathbb{E}[(f_1 - f_{12}) + (f_{12} - \mathbb{E}_{-12}f_{12}) + (\mathbb{E}_{-1}f_{12} - \mathbb{E}_{-1}f_1)][f_2 - \mathbb{E}_{-2}f_2]| \\ &\leq |\mathbb{E}[f_{12} - \mathbb{E}_{-12}f_{12}][(f_2 - f_{21}) + (f_{21} - \mathbb{E}_{-12}f_{21}) + (\mathbb{E}_{-2}f_{21} - \mathbb{E}_{-2}f_2)]| \\ &\quad + 2C_0[\mathbb{E}|f_1 - f_{12}| + \mathbb{E}|\mathbb{E}_{-1}f_{12} - \mathbb{E}_{-1}f_1|] \\ &\leq 2C_0[\mathbb{E}|f_1 - f_{12}| + \mathbb{E}|\mathbb{E}_{-1}f_{12} - \mathbb{E}_{-1}f_1| + \mathbb{E}|f_2 - f_{21}| + \mathbb{E}|\mathbb{E}_{-2}f_{21} - \mathbb{E}_{-2}f_2|] = o(1).\end{aligned}$$

*Claim 3.* $\mathbb{E}|f_i - f_{ij}| \to 0$ and $\mathbb{E}|\mathbb{E}_{-i}f_i - \mathbb{E}_{-i}f_{ij}| \to 0$ for any fixed $i, j$, $i \neq j$.

Formula (1) yields

$$\Delta_{ij} = z_i'[(Z_{-i}'Z_{-i} + \lambda_n I_l)^{-1} - (Z_{-ij}'Z_{-ij} + \lambda_n I_l)^{-1}]z_i = \frac{|z_i'(Z_{-ij}'Z_{-ij} + \lambda_n I_l)^{-1}z_j|^2}{1 + z_j'(Z_{-ij}'Z_{-ij} + \lambda_n I_l)^{-1}z_j} \geq 0.$$

If $\Delta_{ij} \leq 1$, then, by the mean value theorem, $|f_i - f_{ij}| \leq C_0\, \Delta_{ij}$. If $\Delta_{ij} > 1$, then $|f_i - f_{ij}| \leq 2C_0$. By conditional Jensen's inequality,

$$\mathbb{E}|\mathbb{E}_{-i}(f_i - f_{ij})| \leq \mathbb{E}|f_i - f_{ij}| \leq 2C_0\, \mathbb{E}\min\{\Delta_{ij}, 1\}$$

and

$$\mathbb{E}\min\{\Delta_{ij}, 1\} = \mathbb{E}\mathbb{E}_{-i}\min\{\Delta_{ij}, 1\} \leq \mathbb{E}\min\{\mathbb{E}_{-i}\Delta_{ij}, 1\}.$$

It follows from the equality $\mathbb{E}_{-i}z_i z_i' = \mathbb{E}zz'$ that

$$\begin{aligned}
\mathbb{E}_{-i}\Delta_{ij} &= \mathbb{E}_{-i}\frac{z_j'(Z_{-ij}'Z_{-ij} + \lambda_n I_l)^{-1}z_i z_i'(Z_{-ij}'Z_{-ij} + \lambda_n I_l)^{-1}z_j}{1 + z_j'(Z_{-ij}'Z_{-ij} + \lambda_n I_l)^{-1}z_j} \\
&\leq \lambda_{\max}(\mathbb{E}zz')\frac{z_j'(Z_{-ij}'Z_{-ij} + \lambda_n I_l)^{-2}z_j}{1 + z_j'(Z_{-ij}'Z_{-ij} + \lambda_n I_l)^{-1}z_j} \leq \frac{\lambda_{\max}(\mathbb{E}zz')}{\lambda_n} = o(1).
\end{aligned}$$

Hence, Claim 3 obtains.

*Claim 4.* $\mathbb{E}|n^{-1}\sum_{i=1}^n \mathbb{E}_{-i}f_i - \mathbb{E}_{-1}f_1| \to 0$.

Using that $\mathbb{E}_{-1}f_{12} = \mathbb{E}_{-12}f_{12} = \mathbb{E}_{-12}f_{21} = \mathbb{E}_{-2}f_{21}$, Claim 3, and the exchangeability of $\{\mathbb{E}_{-i}f_i - \mathbb{E}_{-1}f_1\}_{i=2}^n$, we derive that

$$\begin{aligned}
\mathbb{E}\left|\frac{1}{n}\sum_{i=1}^n \mathbb{E}_{-i}f_i - \mathbb{E}_{-1}f_1\right| &\leq \frac{1}{n}\sum_{i=1}^n \mathbb{E}|\mathbb{E}_{-i}f_i - \mathbb{E}_{-1}f_1| \leq \mathbb{E}|\mathbb{E}_{-1}f_1 - \mathbb{E}_{-2}f_2| \\
&= \mathbb{E}|\mathbb{E}_{-1}f_1 - \mathbb{E}_{-1}f_{12} + \mathbb{E}_{-2}f_{21} - \mathbb{E}_{-2}f_2| \\
&\leq \mathbb{E}|\mathbb{E}_{-1}f_1 - \mathbb{E}_{-1}f_{12}| + \mathbb{E}|\mathbb{E}_{-2}f_{21} - \mathbb{E}_{-2}f_2| = o(1).
\end{aligned}$$

Thus, Claim 4 is proven.

*Claim 5.* If $\lambda_{\min}^*(Z'Z) \xrightarrow{p} \infty$, then $n^{-1}\sum_{i=1}^n f(P_{ii}) - \mathbb{E}_{-1}f(P_{11}) \xrightarrow{p} 0$.

This follows from Claims 1–4.

*Claim 6.* $\mathbb{E}|\mathbb{E}_{-1}f_1 - \mathbb{E}f_1|^2 \to 0$.

To prove Claim 6 we need the assumption $\lambda^*_{\min}(Z'Z)/\sqrt{n} \xrightarrow{p} \infty$. Going back to the definition of $\lambda_n$ in Claim 1, we can initially take $\lambda_n$ growing faster than $\sqrt{n}$ and slower than $\lambda^*_{\min}(Z'Z)$ (i.e. $\lambda_n/\lambda^*_{\min}(Z'Z) \xrightarrow{p} 0$). Let $\mathbb{E}_i = \mathbb{E}[\cdot|z_2,\dots,z_i]$ and $\mathbb{E}_1 = \mathbb{E}$. Using that $\mathbb{E}_i(\mathbb{E}_{-1}f_{1i}) = \mathbb{E}_{i-1}(\mathbb{E}_{-1}f_{1i})$, we represent $\mathbb{E}_{-1}f_1 - \mathbb{E}f_1$ as the sum of martingale differences

$$\mathbb{E}_{-1}f_1 - \mathbb{E}f_1 = \sum_{i=2}^{n}(\mathbb{E}_i - \mathbb{E}_{i-1})\mathbb{E}_{-1}f_1 = \sum_{i=2}^{n}(\mathbb{E}_i - \mathbb{E}_{i-1})\mathbb{E}_{-1}(f_1 - f_{1i}),$$

where, by (1) and the inequalities given in the proof of Claim 3,

$$|\mathbb{E}_{-1}(f_1-f_{1i})| \le \mathbb{E}_{-1}|f_1-f_{1i}| \le 2C_0\mathbb{E}_{-1}\min\{\Delta_{1i},1\} \le 2C_0\min\{\mathbb{E}_{-1}\Delta_{1i},1\} \le \frac{2C_0}{\lambda_n}\lambda_{\max}(\mathbb{E}zz').$$

Claim 6 now follows from

$$\mathbb{E}|\mathbb{E}_{-1}f_1 - \mathbb{E}f_1|^2 = \sum_{i=2}^{n}\mathbb{E}|(\mathbb{E}_i - \mathbb{E}_{i-1})\mathbb{E}_{-1}(f_1 - f_{1i})|^2 \le \frac{4C_0^2\lambda_{\max}(\mathbb{E}zz')^2 n}{\lambda_n^2} = o(1).$$

We finish the proof of the lemma by noting that $\mathbb{E}f_1 - \mathbb{E}f(P_{11}) = o(1)$ (see the proof of Claim 1). □

**Proof of Theorem 3.2.** For the sake of simplicity, we further omit index $l$ when writing $z_{lk}$. Fix $k$. By Property P, for any $\varepsilon > 0$,

$$z_k'(Z_{-k}'Z_{-k} + \varepsilon n I_l)^{-1}z_k - d_k\mathrm{tr}(Z_{-k}'Z_{-k} + \varepsilon n I_l)^{-1} \xrightarrow{p} 0$$

because of the independence of $z_k$ and $Z_{-k}$, where $Z_{-k}$ is obtained by removing $k^{th}$ row in $Z$. Hence, there exist $\{\varepsilon_n\}_{n=1}^\infty$ tending to zero arbitrarily slowly, such that

$$z_k'(Z_{-k}'Z_{-k} + \varepsilon_n n I_l)^{-1}z_k - d_k S_{nk} \xrightarrow{p} 0,$$

where $S_{nk} = \mathrm{tr}(Z_{-k}'Z_{-k} + \varepsilon_n n I_l)^{-1}$. In particular, we can take $\varepsilon_n\sqrt{n} \to \infty$. Lemma A.4 now yields $z_k'(Z_{-k}'Z_{-k} + \varepsilon_n n I_l)^{-1}z_k - d_k\mathbb{E}S_{nk} \xrightarrow{p} 0$.

By Condition A, $\varepsilon_n n/\lambda_{\min}(Z'Z) \xrightarrow{p} 0$. Arguing as in Claim 1 in the proof of Lemma A.5, we derive that

$$|P_{kk} - z_k'(Z'Z + \varepsilon_n n I_l)^{-1}z_k| \le \min\{\varepsilon_n n/\lambda_{\min}(Z'Z), 1\} = o_p(1).$$

By (2) and the above arguments,

$$z_k'(Z'Z + \varepsilon_n n I_l)^{-1}z_k = g(z_k'(Z_{-k}'Z_{-k} + \varepsilon_n n I_l)^{-1}z_k) = g(d_k\mathbb{E}S_{nk}) + e_n,$$

where $g(x) = x/(x+1)$, $e_n \xrightarrow{p} 0$, and $|e_n| \leq 2$ a.s. Since $\mathbb{P}(\lambda_{\min}(Z'Z) > 0) \to 1$ and $P_{kk}$ are identically distributed over $k$, we have

$$\mathbb{E}P_{kk} = \frac{1}{n}\mathbb{E}\sum_{j=1}^{n} P_{jj} = \frac{l}{n} + o(1) \to \alpha.$$

As a result, $\mathbb{E}g(d_k\mathbb{E}S_{nk}) = \mathbb{E}g(d\mathbb{E}S_{nk}) \to \alpha$. Note that $f(s) = \mathbb{E}g(sd)$ is a strictly increasing continuous function with $f(0) = 0$ and $f(s) \to \mathbb{P}(d > 0)$, $s \to \infty$, whenever $\mathbb{P}(d > 0) > 0$. Therefore, $\mathbb{E}S_{nk} \to c$ for $c > 0$ solving $f(c) = \alpha$. Such $c$ exists when $\alpha \in (0, \mathbb{P}(d > 0))$. Combining the above estimates, we infer that $P_{kk} \xrightarrow{p} g(cd_k) = cd_k/(1 + cd_k)$. $\square$

**Lemma A.6** *Under the conditions of Lemma 3.4(a) or (b), there is $C > 0$ such that, for any $l \times l$ positive semi-definite symmetric matrix $A_l$ and $b > 1$,*

$$\mathbb{E}|x_l'A_lx_l - \operatorname{tr}(A_l)| \leq Cb\sqrt{l}\lambda_{\max}(A_l) + Cl\lambda_{\max}(A_l)\max_{k\geq 1}\mathbb{E}e_k^2\mathbb{I}_{\{|e_k^2-1|>b^2\}}. \tag{10}$$

**Proof of Lemma A.6.** First, assume that $\xi_l = e_l$, $l \geq 1$. Write $A_l = (a_{ij})_{i,j=1}^{l}$. Then

$$x_l'A_lx_l - \operatorname{tr}(A_l) = \sum_{k=1}^{l} a_{kk}(e_k^2 - 1) + 2\sum_{1\leq j<k\leq l} a_{jk}e_je_k = \sum_{k=1}^{l} a_{kk}(e_k^2 - 1) + 2\sum_{k=2}^{l} E_k,$$

where

$$E_k = \left(\sum_{j=1}^{k-1} a_{jk}e_j\right)e_k,$$

$2 \leq k \leq l$. Note that $\{E_k\}_{k=2}^{l}$ and $\{a_{kk}(e_k^2 - 1)\}_{k=1}^{l}$ are martingale difference sequences. By the Cauchy-Schwartz inequality,

$$\left(\mathbb{E}\left|\sum_{k=2}^{l} E_k\right|\right)^2 \leq \mathbb{E}\left|\sum_{k=2}^{l} E_k\right|^2 = \sum_{k=2}^{l}\mathbb{E}E_k^2 = \sum_{k=2}^{l}\sum_{j=1}^{k-1} a_{jk}^2 \leq \operatorname{tr}(A_l^2).$$

By the Burkholder-Davis-Gundy inequality,

$$\mathbb{E}\left|\sum_{k=1}^{l} a_{kk}(e_k^2 - 1)\right| \leq C\mathbb{E}\left|\sum_{k=1}^{l} a_{kk}^2(e_k^2 - 1)^2\right|^{1/2},$$

where $C > 0$ is an absolute constant. Since $\sqrt{x+y} \leq \sqrt{x} + \sqrt{y}$ for $x, y \geq 0$,

$$\mathbb{E}\left|\sum_{k=1}^{l} a_{kk}^2(e_k^2 - 1)^2\right|^{1/2} \leq I_1 + I_2,$$

where

$$
\begin{aligned}
I_1 &= \mathbb{E}\left|\sum_{k=1}^{l} a_{kk}^2(e_k^2-1)^2\mathbb{I}_{\{|e_k^2-1|\le b^2\}}\right|^{1/2}, \\
I_2 &= \mathbb{E}\left|\sum_{k=1}^{l} a_{kk}^2(e_k^2-1)^2\mathbb{I}_{\{|e_k^2-1|> b^2\}}\right|^{1/2}.
\end{aligned}
$$

By Jensen's inequality,

$$
I_1 \le \left|\sum_{k=1}^{l} a_{kk}^2\mathbb{E}(e_k^2-1)^2\mathbb{I}_{\{|e_k^2-1|\le b^2\}}\right|^{1/2} \le \sqrt{2b^2\mathrm{tr}(A_l^2)}.
$$

Here we also used $\mathbb{E}(e_k^2-1)^2\mathbb{I}_{\{(|e_k^2-1|\le b^2\}} \le b^2\mathbb{E}|e_k^2-1| \le 2b^2$. In addition,

$$
I_2 \le \sum_{k=1}^{l} |a_{kk}|\mathbb{E}e_k^2\mathbb{I}_{\{|e_k^2-1|>b^2\}} = \mathrm{tr}(A_l)\max_{k\ge1}\mathbb{E}e_k^2\mathbb{I}_{\{|e_k^2-1|>b^2\}},
$$

where we also have used that $\sqrt{x+y} \le \sqrt{x}+\sqrt{y}$ for $x, y \ge 0$ and $|e_k^2-1| \le e_k^2$ when $b > 1$ and $|e_k^2-1| > b^2$. The above estimates yield

$$
\mathbb{E}|x_l' A_l x_l - \mathrm{tr}(A_l)| \le Cb\sqrt{\mathrm{tr}(A_l^2)} + C\mathrm{tr}(A_l)\max_{k\ge1}\mathbb{E}e_k^2\mathbb{I}_{\{|e_k^2-1|>b^2\}}, \tag{11}
$$

where $x_l = (e_1, \ldots, e_l)'$ and $C > 0$ is an absolute constant.

Consider the case with $x_l = (\xi_1, \ldots, \xi_l)'$. By the definition of $\xi_j$, there are $l \times k$ matrices $\Gamma_{lk}$ such that $\Gamma_{lk}v_k \to x_l$ in probability and in mean square as $k \to \infty$ for $v_k = (e_1, \ldots, e_k)'$. Since $\{e_k\}_{k\ge1}$ is an orthonormal sequence, we have

(1) $\Gamma_{lk}\Gamma_{lk}' = \mathbb{E}(\Gamma_{lk}v_k)(\Gamma_{lk}v_k)' \to \mathbb{E}x_l x_l' = I_l$,

(2) $v_k'(\Gamma_{lk}'A_l\Gamma_{lk})v_k = (\Gamma_{lk}v_k)'A_l(\Gamma_{lk}v_k) \xrightarrow{p} x_l'A_l x_l$,

(3) $\mathrm{tr}(\Gamma_{lk}'A_l\Gamma_{lk}) = \mathrm{tr}(\Gamma_{lk}\Gamma_{lk}'A_l) \to \mathrm{tr}(A_l)$,

(4) $\mathrm{tr}((\Gamma_{lk}'A_l\Gamma_{lk})^2) = \mathrm{tr}(\Gamma_{lk}\Gamma_{lk}'A_l\Gamma_{lk}\Gamma_{lk}'A_l) \to \mathrm{tr}(A_l^2)$.

We need a version of the Fatou lemma that states that $\mathbb{E}|\zeta| \le \underline{\lim}_{k\to\infty}\mathbb{E}|\zeta_k|$ if $\zeta_k \xrightarrow{p} \zeta$. Put $B_k = \Gamma_{lk}'A_l\Gamma_{lk}$. By the Fatou lemma and (11),

$$
\begin{aligned}
\mathbb{E}|x_l'A_l x_l - \mathrm{tr}(A_l)| &\le \underline{\lim}_{k\to\infty}\mathbb{E}|v_k'B_k v_k - \mathrm{tr}(B_k)| \\
&\le \underline{\lim}_{k\to\infty}[Cb\sqrt{\mathrm{tr}(B_k^2)} + C\mathrm{tr}(B_k)\max_{j\ge1}\mathbb{E}e_j^2\mathbb{I}_{\{|e_j^2-1|>b^2\}}] \\
&\le Cb\sqrt{\mathrm{tr}(A_l^2)} + C\mathrm{tr}(A_l)\max_{k\ge1}\mathbb{E}e_k^2\mathbb{I}_{\{|e_k^2-1|>b^2\}} \\
&\le Cb\lambda_{\max}(A_l)\sqrt{l} + Cl\lambda_{\max}(A_l)\max_{k\ge1}\mathbb{E}e_k^2\mathbb{I}_{\{|e_k^2-1|>b^2\}}.
\end{aligned}
$$

Hence, we get the desired inequality. □

**Proof of Lemma 3.4.** If $\{e_k\}_{k\geq 1}$ are IID and $x_l$ is given in (a), then

$$\max_{k\geq 1} \mathbb{E}e_k^2 \mathbb{I}_{\{|e_k^2-1|>b^2\}} = \mathbb{E}e_1^2 \mathbb{I}_{\{|e_1^2-1|>b^2\}}$$

and the desired result follows from Lemma A.6. Indeed, dividing both sides of (10) by $l$, letting $l \to \infty$ and then $b \to \infty$, we infer that $(\{x_l\}_{l\geq 1}, 1)$ satisfies Property P. Multiplying by $d$, we conclude that $(\{dx_l\}_{l\geq 1}, d^2)$ satisfies Property P.

If $\{e_k\}_{k\geq 1}$ are independent with $\mathbb{E}|e_k|^{2+\delta} \leq C$ and $x_l$ is as in (b), then, for $b > 1$,

$$\max_{k\geq 1} \mathbb{E}e_k^2 \mathbb{I}_{\{|e_k^2-1|>b^2\}} \leq \max_{k\geq 1} \frac{\mathbb{E}e_k^{2+\delta}}{(b^2+1)^{\delta/2}} \leq \frac{C}{(b^2+1)^{\delta/2}}.$$

The rest of the proof follows the same argument as above.

Consider (c), where $x_l$ is a centered random vector with a log-concave density and $\mathbb{E}x_l x_l' = I_l$. By Lemma 2.5 in Pajor and Pastur (2009), $\mathrm{var}(x_l' A_l x_l / l) \leq \delta_l$ for some $\delta_l = o(1)$ and all $l \times l$ symmetric positive semi-definite matrices $A_l$ with $\lambda_{\max}(A_l) \leq 1$. Obviously, this implies that $(\{x_l\}_{l\geq 1}, 1)$ satisfies Property P. Multiplying by $d$, we get the desired result.

Suppose $x_l = F_l(v_m)$, where $F_l$ and $v_m$ are as in (d). Then, $f = \varphi \circ F_l$ is a $c$-Lipschitz function for any 1-Lipschitz function $\varphi : \mathbb{R}^l \to \mathbb{R}$. Indeed, for all $u, v \in \mathbb{R}^m$,

$$|\varphi(F_l(u)) - \varphi(F_l(v))| \leq \|F_l(u) - F_l(v)\| \leq c\|u - v\|.$$

Since $\lambda_{\max}(\mathrm{var}(v_m)) \leq C$ for all $m$, the density of $v_m$ has the form $\exp\{-U(v)\}$ for a convex function $U = U(v)$ such that $\partial^2 U(v) - (1/C)I_m = \mathrm{var}(v_m)^{-1} - (1/C)I_m$ is positive semi-definite for all $v \in \mathbb{R}^m$.

Hence, by Theorem 2.7 and Proposition 1.3 in Ledoux (2001) (see also examples in Section 3.2 in El Karoui, 2009), there is $C_1 = C_1(C, c) > 0$ such that, for any 1-Lipschitz function $\varphi : \mathbb{R}^l \to \mathbb{R}$ and $f = \varphi \circ F_l$,

$$\mathbb{P}(|\varphi(x_l) - \mathrm{med}(\varphi(x_l))| > t) = \mathbb{P}(|f(v_m) - \mathrm{med}(f(v_m))| > t) \leq 2\exp\{-C_1 t^2\}, \quad t > 0,$$

where $\mathrm{med}(\xi)$ is a median of a random variable $\xi$.[8] Now, by Lemma 7 in El Karoui (2009), $(\{x_l\}_{l\geq 1}, 1)$ satisfies Property P. Multiplying by $d$, we finish the proof. □

[8] $\mathrm{med}(\xi)$ is any such point $\mu$ that $\mathbb{P}(\xi < \mu) \leq 1/2 \leq \mathbb{P}(\xi \leq \mu)$.

**Lemma A.7** *Let $\{e_k\}_{k\geq 1}$ be independent random variables with $\mathbb{E}e_k = 0$ and $\mathbb{E}e_k^2 = 1$. If $\mathbb{E}|e_k| \geq c$ for some $c > 0$ and all $k \geq 1$, then, for any $\{a_k\}_{k\geq 0}$ with $\sum_{k\geq 0} a_k^2 = 1$,*

$$\mathbb{E}\left|a_0 + \sum_{k\geq 1} a_k e_k\right| \geq \frac{c}{\sqrt{32 + c^2}}.$$

**Proof of Lemma A.7.** Note that $\mathbb{E}\left|a_0 + \sum_{k\geq 1} a_k e_k\right|^2 = \sum_{k\geq 0} a_k^2 = 1$. We may assume without loss of generality that there is a finite set of non-zero $a_k$ (otherwise, we can take a limit). By Jensen's inequality,

$$|a_0| = \mathbb{E}\left|a_0 + \mathbb{E}\sum_{k\geq 1} a_k e_k\right| \leq \mathbb{E}\left|a_0 + \sum_{k\geq 1} a_k e_k\right| = I.$$

In addition,

$$\sqrt{1 - a_0^2}\mathbb{E}\left|\sum_{k\geq 1} \tilde{a}_k e_k\right| - |a_0| \leq I,$$

where $\tilde{a}_k = a_k/\sqrt{1 - a_0^2}$, $k \geq 1$, and $\sum_{k\geq 1} \tilde{a}_k^2 = 1$. If we prove that

$$\mathbb{E}\left|\sum_{k\geq 1} \tilde{a}_k e_k\right| \geq \frac{c}{2\sqrt{2}}, \tag{12}$$

then we obtain the desired bound:

$$I \geq \inf_{b\in[0,1]} \max\left\{\frac{c}{2\sqrt{2}}\sqrt{1 - b^2} - b, b\right\} = \frac{c}{\sqrt{32 + c^2}}.$$

Let us prove (12). Write $a_k$ instead of $\tilde{a}_k$ and let $\{\tilde{e}_k\}_{k\geq 1}$ be an independent copy of $\{e_k\}_{k\geq 1}$. Then

$$\mathbb{E}\left|\sum_{k\geq 1} a_k(e_k - \tilde{e}_k)\right| \leq \mathbb{E}\left|\sum_{k\geq 1} a_k e_k\right| + \mathbb{E}\left|\sum_{k\geq 1} a_k \tilde{e}_k\right| = 2\mathbb{E}\left|\sum_{k\geq 1} a_k e_k\right|.$$

In addition, by Jensen's inequality, $\mathbb{E}|e_k - \tilde{e}_k| \geq \mathbb{E}|e_k - \mathbb{E}[\tilde{e}_k|e_k]| = \mathbb{E}|e_k|$ for all $k \geq 1$. Since $\{e_k - \tilde{e}_k\}_{k\geq 1}$ are independent symmetric random variables, then $\{e_k - \tilde{e}_k\}_{k\geq 1} = \{d_k|e_k - \tilde{e}_k|\}_{k\geq 1}$ in distribution, where $\{d_k\}_{k\geq 1}$ are IID random variables that have $\mathbb{P}(d_k = \pm 1) = 1/2$ and are independent of $\{|e_k - \tilde{e}_k|\}_{k\geq 1}$. By Jensen's inequality,

$$\mathbb{E}\left|\sum_{k\geq 1} a_k d_k \mathbb{E}|e_k - \tilde{e}_k|\right| \leq \mathbb{E}\left|\sum_{k\geq 1} a_k d_k |e_k - \tilde{e}_k|\right| = \mathbb{E}\left|\sum_{k\geq 1} a_k(e_k - \tilde{e}_k)\right|.$$

By Khinchin's inequality with explicit constants (see Theorem 1 in Szarek, 1975),

$$\frac{c}{\sqrt{2}} \leq \frac{1}{\sqrt{2}}\left(\sum_{k\geq 1} a_k^2 (\mathbb{E}|e_k|)^2\right)^{1/2} \leq \frac{1}{\sqrt{2}}\left(\sum_{k\geq 1} a_k^2 (\mathbb{E}|e_k - \tilde{e}_k|)^2\right)^{1/2} \leq \mathbb{E}\left|\sum_{k\geq 1} a_k d_k \mathbb{E}|e_k - \tilde{e}_k|\right|.$$

□

# Additional references

El Karoui, N (2009): "Concentration of measure and spectra of random matrices: applications to correlation matrices, elliptical distributions and beyond," *Annals of Applied Probability*, 19, 2362–2405.

Pajor, A. and L. Pastur (2009): "On the limiting empirical measure of eigenvalues of the sum of rank one matrices with log-concave distribution," *Studia Mathematica*, 195, 11–29.

Szarek, V. (1975): "On the best constants in the Khinchin inequality," *Studia Mathematica*, 58, 197–208.