Supplementary material

Supplement to "J P Williams, Y Xie, and J Hannig (2019+). The EAS approach for graphical selection consistency in vector autoregression models. *In review.*"

S1 Derivation of the generalized fiducial mass of G and the Jacobian

This section presents the details of deriving the generalized fiducial probability mass function (8) for each graph, G, beginning with (7). In particular, the closed-form mathematical expression of the Jacobian term, $J\big(Y, (\boldsymbol{\alpha}_G, \{\sigma_j\})\big)$, in (7) is worked out.

From Hannig et al. (2016),

$$J\big(Y, (\boldsymbol{\alpha}_G, \{\sigma_j\})\big) := |\mathcal{D}_g'\mathcal{D}_g|^{\frac{1}{2}} = \frac{n^{-\frac{|G|}{2}}}{\sigma_1 \cdots \sigma_p} |\widetilde{\mathcal{D}}_g'\widetilde{\mathcal{D}}_g|^{\frac{1}{2}},$$

where $\widetilde{\mathcal{D}}$ is $\mathcal{D}$ with the σ_i omitted and $n^{-\frac{1}{4}}$ factored out, the equality follows by the Cauchy-Binet formula, and

$$\mathcal{D} := n^{-\frac{1}{2}} \begin{pmatrix} \frac{\partial X^{(1)}}{\partial A_{11}} & \frac{\partial X^{(1)}}{\partial A_{12}} & \cdots & \frac{\partial X^{(1)}}{\partial A_{pp}} & \frac{\partial X^{(1)}}{\partial \sigma_1} & \cdots & \frac{\partial X^{(1)}}{\partial \sigma_p} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial X^{(n)}}{\partial A_{11}} & \frac{\partial X^{(n)}}{\partial A_{12}} & \cdots & \frac{\partial X^{(n)}}{\partial A_{pp}} & \frac{\partial X^{(n)}}{\partial \sigma_1} & \cdots & \frac{\partial X^{(n)}}{\partial \sigma_p} \end{pmatrix}.$$

Note that we have rescaled $\mathcal{D}$ from its definition in Hannig et al. (2016) by the factor of $n^{-\frac{1}{2}}$. This scaling is necessary for controlling the asymptotic rate of growth of the Jacobian. The first p^2 columns are partial derivatives with respect to the components A_{ij} of the transition matrix over all observed time instances of data. For each $t \in \{1, \dots, n\}$, these partial derivatives are expressed as

$$\begin{aligned} \frac{\partial X^{(t)}}{\partial A_{ij}} &= \frac{\partial A}{\partial A_{ij}} X^{(t-1)} + A \frac{\partial X^{(t-1)}}{\partial A_{ij}} \\ &= J^{ij} X^{(t-1)} + A\Big(\frac{\partial A}{\partial A_{ij}} X^{(t-2)} + A \frac{\partial X^{(t-2)}}{\partial A_{ij}}\Big) \\ &= J^{ij} X^{(t-1)} + AJ^{ij} X^{(t-2)} + A^2\Big(\frac{\partial A}{\partial A_{ij}} X^{(t-3)} + A \frac{\partial X^{(t-3)}}{\partial A_{ij}}\Big) \\ &\vdots \\ &= J^{ij} X^{(t-1)} + AJ^{ij} X^{(t-2)} + A^2 J^{ij} X^{(t-3)} + \cdots + A^{t-1} J^{ij} X^{(0)} + 0, \end{aligned}$$

where J^{ij} is the matrix whose only nonzero element is a 1 in the ijth coordinate. Thus, the column of $\mathcal{D}$ corresponding to the ijth partial derivative of A can be expressed as

$$\begin{aligned}
&\begin{pmatrix}
J^{ij}X^{(0)} & + & \cdots & + & 0 & + & 0 & + & 0 \\
\vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
A^{n-3}J^{ij}X^{(0)} & + & \cdots & + & J^{ij}X^{(n-3)} & + & 0 & + & 0 \\
A^{n-2}J^{ij}X^{(0)} & + & \cdots & + & AJ^{ij}X^{(n-3)} & + & J^{ij}X^{(n-2)} & + & 0 \\
A^{n-1}J^{ij}X^{(0)} & + & \cdots & + & A^2J^{ij}X^{(n-3)} & + & AJ^{ij}X^{(n-2)} & + & J^{ij}X^{(n-1)}
\end{pmatrix} \\
&= \Theta\big(I_n \otimes J^{ij}\big)\mathrm{vec}(\mathcal{X}),
\end{aligned}$$

where

$$\Theta := \begin{pmatrix}
I_p & \cdots & 0 & 0 & 0 \\
\vdots & & \vdots & \vdots & \vdots \\
A^{n-3} & \cdots & I_p & 0 & 0 \\
A^{n-2} & \cdots & A & I_p & 0 \\
A^{n-1} & \cdots & A^2 & A & I_p
\end{pmatrix}. \tag{S1}$$

Note that the notation Θ_g will be taken to mean the matrix Θ with A_g replacing every occurrence of A in (S1).

Next, the remaining p columns of $\mathcal{D}$ require the partial derivatives of the time instances of the data with respect to σ_i, for each $i \in \{1, \ldots, p\}$. Accordingly, observe that

$$\begin{aligned}
\frac{\partial X^{(t)}}{\partial \sigma_i} &= A\frac{\partial X^{(t-1)}}{\partial \sigma_i} + J^{ii}U^{(t)} \\
&= A\Big(A\frac{\partial X^{(t-2)}}{\partial \sigma_i} + J^{ii}U^{(t-1)}\Big) + J^{ii}U^{(t)} \\
&= A^2\Big(A\frac{\partial X^{(t-3)}}{\partial \sigma_i} + J^{ii}U^{(t-2)}\Big) + AJ^{ii}U^{(t-1)} + J^{ii}U^{(t)} \\
&\ \ \vdots \\
&= A^{t-1}J^{ii}U^{(1)} + A^{t-2}J^{ii}U^{(2)} + \cdots + AJ^{ii}U^{(t-1)} + J^{ii}U^{(t)},
\end{aligned}$$

and so the column of $\mathcal{D}$ corresponding to the partial derivative of σ_i can be expressed as

$$\begin{aligned}
&\begin{pmatrix}
J^{ii}U^{(1)} & + & 0 & + & \cdots & + & 0 & + & 0 \\
\vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots \\
A^{n-3}J^{ii}U^{(1)} & + & A^{n-4}J^{ii}U^{(2)} & + & \cdots & + & 0 & + & 0 \\
A^{n-2}J^{ii}U^{(1)} & + & A^{n-3}J^{ii}U^{(2)} & + & \cdots & + & J^{ii}U^{(n-1)} & + & 0 \\
A^{n-1}J^{ii}U^{(1)} & + & A^{n-2}J^{ii}U^{(2)} & + & \cdots & + & AJ^{ii}U^{(n-1)} & + & J^{ii}U^{(n)}
\end{pmatrix} \\
&= \Theta\big(I_n \otimes J^{ii}\sigma_i^{-1}\big)\big(Y - \mathcal{Z}\boldsymbol{\alpha}\big).
\end{aligned}$$

With expressions for all $p^2 + p$ columns of $\mathcal{D}$ now derived, the Jacobian can be written in closed-form.

Keeping the derivation of $J\big(Y,(\boldsymbol{\alpha}_G,\{\sigma_j\})\big)$ in mind, the marginal distribution,

$$
\begin{aligned}
r(G,\{\sigma_j\}\mid Y) &= \int r(\boldsymbol{\alpha}_G,\{\sigma_j\}\mid Y)\ d\boldsymbol{\alpha}_G \\
&\propto \int f(Y\mid \boldsymbol{\alpha}_G,\{\sigma_j\})J\big(Y,(\boldsymbol{\alpha}_G,\{\sigma_j\})\big)h\big(\boldsymbol{\alpha}_G,\{\sigma_j\}\big)\ d\boldsymbol{\alpha}_G \\
&= \int \frac{e^{-\frac{1}{2}(Y-\mathcal{Z}_G\boldsymbol{\alpha}_G)'\mathcal{W}^{-1}(Y-\mathcal{Z}_G\boldsymbol{\alpha}_G)}}{(2\pi)^{\frac{np}{2}}\left(\sigma_1^2\cdots\sigma_p^2\right)^{\frac{n+1}{2}}n^{\frac{|G|}{2}}}|\widetilde{\mathcal{D}}_g'\widetilde{\mathcal{D}}_g|^{\frac{1}{2}}h\big(\boldsymbol{\alpha}_G,\{\sigma_j\}\big)\ d\boldsymbol{\alpha}_G \\
&= \frac{e^{-\frac{1}{2}\mathcal{S}_G}(2\pi)^{-\frac{np}{2}}n^{-\frac{|G|}{2}}}{\left(\sigma_1^2\cdots\sigma_p^2\right)^{\frac{n+1}{2}}}\int e^{-\frac{1}{2}(\boldsymbol{\alpha}_G-\widehat{\boldsymbol{\alpha}}_g)'\mathcal{Z}_G'\mathcal{W}^{-1}\mathcal{Z}_G(\boldsymbol{\alpha}_G-\widehat{\boldsymbol{\alpha}}_g)}|\widetilde{\mathcal{D}}_g'\widetilde{\mathcal{D}}_g|^{\frac{1}{2}}h\big(\boldsymbol{\alpha}_G,\{\sigma_j\}\big)\ d\boldsymbol{\alpha}_G \\
&\propto \frac{e^{-\frac{1}{2}\mathcal{S}_G}n^{-\frac{|G|}{2}}}{\left(\sigma_1^2\cdots\sigma_p^2\right)^{\frac{n+1}{2}}}E_{\boldsymbol{\alpha}_G|\{\sigma_j\}}\Big(h\big(\boldsymbol{\alpha}_G,\{\sigma_j\}\big)|\widetilde{\mathcal{D}}_g'\widetilde{\mathcal{D}}_g|^{\frac{1}{2}}\Big)\cdot(2\pi)^{\frac{|G|}{2}}|\mathcal{Z}_G'\mathcal{W}^{-1}\mathcal{Z}_G|^{-\frac{1}{2}},
\end{aligned}
$$

where $\widehat{\boldsymbol{\alpha}}_g := \big(\mathcal{Z}_G'\mathcal{W}^{-1}\mathcal{Z}_G\big)^{-1}\mathcal{Z}_G'\mathcal{W}^{-1}Y$ is the weighted least squares estimator, $\mathcal{S}_G := Y'\mathcal{W}^{-1}\Big(I_{np}-\mathcal{Z}_G\big(\mathcal{Z}_G'\mathcal{W}^{-1}\mathcal{Z}_G\big)^{-1}\mathcal{Z}_G'\mathcal{W}^{-1}\Big)Y$ is the corresponding weighted sum-of-squared residuals, and the conditional expectation, $E_{\boldsymbol{\alpha}_G|\{\sigma_j\}}(\cdot) := E\big(\cdot\mid\{\sigma_j\}\big)$, is taken with respect to the $\mathrm{N}_{|G|}\big(\widehat{\boldsymbol{\alpha}}_g,(\mathcal{Z}_G'\mathcal{W}^{-1}\mathcal{Z}_G)^{-1}\big)$ distribution. Further, since $\mathcal{W}$ is a diagonal matrix it can be factored out of $\widehat{\boldsymbol{\alpha}}_g$ and from parts of $\mathcal{S}_G$. This is a well-known fact from linear model theory, and the details are as follows.

$$
\begin{aligned}
\mathcal{W}^{-1}\mathcal{Z}_G &= (I_n\otimes\Sigma^{-1})\begin{pmatrix} X^{(0)'}\otimes I_p \\ \\ X^{(n-1)'}\otimes I_p\end{pmatrix}_G \\
&= \begin{pmatrix} X^{(0)'}\otimes \Sigma^{-1} \\ \\ X^{(n-1)'}\otimes \Sigma^{-1}\end{pmatrix}_G \\
&= \begin{pmatrix} X^{(0)'}\otimes I_p \\ \\ X^{(n-1)'}\otimes I_p\end{pmatrix}_G (I_p\otimes\Sigma^{-1})_{G,G} \\
&= \mathcal{Z}_G(I_p\otimes\Sigma)^{-1}_{G,G},
\end{aligned}
$$

which gives $\mathcal{Z}_G'\mathcal{W}^{-1} = \big(\mathcal{W}^{-1}\mathcal{Z}_G\big)' = (I_p\otimes\Sigma)^{-1}_{G,G}\mathcal{Z}_G'$, and so

$$
\widehat{\boldsymbol{\alpha}}_g = \big((I_p\otimes\Sigma)^{-1}_{G,G}\mathcal{Z}_G'\mathcal{Z}_G\big)^{-1}(I_p\otimes\Sigma)^{-1}_{G,G}\mathcal{Z}_G'Y = \big(\mathcal{Z}_G'\mathcal{Z}_G\big)^{-1}\mathcal{Z}_G'Y,
$$

and $\mathcal{S}_G = Y'\mathcal{W}^{-1}(I_{np}-H_g)Y$, where $H_g := \mathcal{Z}_G\big(\mathcal{Z}_G'\mathcal{Z}_G\big)^{-1}\mathcal{Z}_G'$.

Next, simplify $|\mathcal{Z}_G'\mathcal{W}^{-1}\mathcal{Z}_G|$ as follows. Observe that

$$\begin{aligned}\mathcal{Z}_G'\mathcal{W}^{-1}\mathcal{Z}_G &= (\mathcal{X}'\otimes I_p)_G'(I_n\otimes\Sigma)^{-1}(\mathcal{X}'\otimes I_p)_G\\ &= \Big((\mathcal{X}'\otimes I_p)'(I_n\otimes\Sigma^{-1})(\mathcal{X}'\otimes I_p)\Big)_{G,G}\\ &= (\mathcal{X}'\otimes\Sigma^{-\frac{1}{2}})_G'(\mathcal{X}'\otimes\Sigma^{-\frac{1}{2}})_G.\end{aligned}$$

Permuting the columns of $(\mathcal{X}'\otimes\Sigma^{-\frac{1}{2}})_G$ gives the matrix

$$\begin{pmatrix} \sigma_1^{-1}X_{r_1^g}^{(0)'} & & \\ & \ddots & \\ & & \sigma_p^{-1}X_{r_p^g}^{(0)'} \\ & \vdots & \\ \sigma_1^{-1}X_{r_1^g}^{(n-1)'} & & \\ & \ddots & \\ & & \sigma_p^{-1}X_{r_p^g}^{(n-1)'} \end{pmatrix},$$

and so rearranging rows and columns gives

$$\begin{aligned}|\mathcal{Z}_G'\mathcal{W}^{-1}\mathcal{Z}_G| &= \begin{vmatrix} \sigma_1^{-2}\sum_{t=1}^n X_{r_1^g}^{(t-1)}X_{r_1^g}^{(t-1)'} & & \\ & \ddots & \\ & & \sigma_p^{-2}\sum_{t=1}^n X_{r_p^g}^{(t-1)}X_{r_p^g}^{(t-1)'} \end{vmatrix}\\ &= (\sigma_1^2)^{-|r_1^g|}\cdots(\sigma_p^2)^{-|r_p^g|}\prod_{j=1}^p\left|\sum_{t=1}^n X_{r_j^g}^{(t-1)}X_{r_j^g}^{(t-1)'}\right|.\end{aligned}$$

Note that rearranging the same number of rows as columns preserves the sign of the determinant. Accordingly,

$$r(G,\{\sigma_j\}\mid Y)\propto\frac{\prod_{j=1}^p\left|\sum_{t=1}^n X_{r_j^g}^{(t-1)}X_{r_j^g}^{(t-1)'}\right|^{-\frac{1}{2}}E_{\boldsymbol{\alpha}_G|\{\sigma_j\}}\Big(h\big(\boldsymbol{\alpha}_G,\{\sigma_j\}\big)|\widetilde{\mathcal{D}}_g'\widetilde{\mathcal{D}}_g|^{\frac{1}{2}}\Big)}{n^{\frac{|G|}{2}}(2\pi)^{-\frac{|G|}{2}}\big(\sigma_1^2\cdots\sigma_p^2\big)^{\frac{n+1}{2}}(\sigma_1^2)^{-\frac{|r_1^g|}{2}}\cdots(\sigma_p^2)^{-\frac{|r_p^g|}{2}}e^{\frac{1}{2}\mathcal{S}_G}}.$$

To further simplify this expression, recall that $\Sigma=\mathrm{diag}\{\sigma_1^2,\ldots\sigma_p^2\}$ and observe that

$$\mathcal{S}_G=\Big(X^{(1)'}\Sigma^{-1}\quad\cdots\quad X^{(n)'}\Sigma^{-1}\Big)\begin{pmatrix}(I_{np}-H_g)_1'Y\\ \vdots\\ (I_{np}-H_g)_{np}'Y\end{pmatrix}=\sum_{j=1}^p\sigma_j^{-2}m_j^g,$$

where as in (5), $m_j^g := \sum_{t=1}^n X_j^{(t)}(I_{np}-H_g)'_{(t-1)p+j}Y$. Hence,

$$\begin{aligned}
r(G \mid Y) &= \int_0^\infty \cdots \int_0^\infty r(G, \{\sigma_j\} \mid Y)\; d\sigma_1 \cdots d\sigma_p \\
&\propto \frac{\int_0^\infty \cdots \int_0^\infty E_{\boldsymbol{\alpha}_G|\{\sigma_j\}}\Big(h\big(\boldsymbol{\alpha}_G, \{\sigma_j\}\big)|\widetilde{\mathcal{D}}_g'\widetilde{\mathcal{D}}_g|^{\frac{1}{2}}\Big)\frac{e^{-\frac{m_1^g}{2}\sigma_1^{-2}}\cdots e^{-\frac{m_p^g}{2}\sigma_p^{-2}}}{(\sigma_1^2)^{\frac{n-|r_1^g|+1}{2}}\cdots(\sigma_p^2)^{\frac{n-|r_p^g|+1}{2}}}\; d\sigma_1\cdots d\sigma_p}{(\frac{n}{2\pi})^{\frac{|G|}{2}}\prod_{j=1}^p\Big|\sum_{t=1}^n X_{r_j^g}^{(t-1)}X_{r_j^g}^{(t-1)'}\Big|^{\frac{1}{2}}} \\
&= \frac{E\Big(h\big(\boldsymbol{\alpha}_G, \{\sigma_j\}\big)|\widetilde{\mathcal{D}}_g'\widetilde{\mathcal{D}}_g|^{\frac{1}{2}}\Big)\prod_{j=1}^p\big(\frac{m_j^g}{2}\big)^{-\frac{n-|r_j^g|}{2}}\Gamma\Big(\frac{n-|r_j^g|}{2}\Big)}{(\frac{n}{2\pi})^{\frac{|G|}{2}}\prod_{j=1}^p\Big|\sum_{t=1}^n X_{r_j^g}^{(t-1)}X_{r_j^g}^{(t-1)'}\Big|^{\frac{1}{2}}},
\end{aligned}$$

which gives (8).

S2 Additional lemmas

Lemma S2.1. *Assume that Condition 3.1 holds. Then for all $n \geq N_2$,*

$$P\Bigg(\frac{\big\|(\Gamma_n(0)\otimes(\Sigma^0)^{-1})_{G_o,G_o}\big\|_2^2}{\lambda_{\min}^2(\mathcal{Z}_{G_o}'\mathcal{W}^{-1}\mathcal{Z}_{G_o})} \geq \frac{1}{n^{2-\frac{\rho}{2}}p^5}\Bigg) \leq \frac{p^{\frac{7}{2}}}{n^{\frac{\rho}{4}}}\cdot\frac{(\sigma_{\max}^0)^2}{(\sigma_{\min}^0)^2}\cdot\frac{6\|\Gamma_n(0)\|_2/\delta}{(1-|G_o|/n-2/n)},$$

with probability exceeding $1-V_2-e^{-\frac{np}{4}}$, where V_2 is as in (10).

Lemma S2.2. *Take any $G \subset \{1,\ldots,p^2\}$. If $\|A_g\|_2 \leq c$, then*

$$|\widetilde{\mathcal{D}}_g'\widetilde{\mathcal{D}}_g|^{\frac{1}{2}} < \exp\left\{\frac{1}{2}(1-c)^{-2}\big(r_{\max}^g+(1+c)^2\big)\frac{\|Y\|^2}{\sqrt{n}}-\frac{|G|+p}{2}\right\},$$

where $r_{\max}^g := \max_{1\leq j\leq p}|r_j^g|$.

Lemma S2.3. *If $\|A\|_2 < 1$, then*

$$\|f_X(\Theta_g)\|_2 \leq \frac{\sigma_{\max}^2}{2\pi}\big(1-\|A\|_2\big)^{-2},$$

where f_X is the spectral density for a VAR(1) process with coefficient matrix A and error covariance matrix $\Sigma = diag\{\sigma_1^2,\ldots,\sigma_p^2\}$.

Lemma S2.4. *Assume Condition 3.1 holds. If $\|A_{g_o}\|_2 \leq c$, then for all $n \geq \max\{N_1,N_2\}$,*

$$P_x\Big(|\widetilde{\mathcal{D}}_{g_o}'\widetilde{\mathcal{D}}_{g_o}|^{\frac{1}{2}} \geq e^{\frac{|G_o|+p}{4}}\Big) \geq 1-V_3,$$

where V_3 is as in (11).

Lemma S2.5. *Assume that Condition 3.1 holds. Then for all $n \geq N_2$,*

$$P_x\left(\prod_{j=1}^{p} \frac{\left|(\mathcal{X}\mathcal{X}')_{r_j^{g_o}, r_j^{g_o}}\right|^{\frac{1}{2}}}{\left|(\mathcal{X}\mathcal{X}')_{r_j^{g}, r_j^{g}}\right|^{\frac{1}{2}}} \leq n^{\frac{|G_o|-|G|}{2}} e^{\frac{1}{2}\left(|G_o|[\delta+\lambda_{\max}(\Gamma_n(0))]+|G|2\delta^{-1}\right)}\right) \geq 1 - 2V_2,$$

where V_2 is as in (10).

Lemma S2.6. *Assume that Condition 3.1 holds, and that $1 \leq d \leq \min_{1\leq j\leq p} m_j^{g_o}$. Then for all $n \geq \max\{N_1, N_2\}$,*

$$P_x\left(\prod_{j=1}^{p}\left[\frac{(m_j^{g_o})^{\frac{n-|r_j^{g_o}|}{2}}}{(m_j^{g})^{\frac{n-|r_j^{g}|}{2}}}\right] \leq \left((\sigma^0_{\max})^2 3n\right)^{\frac{p^2}{2}} \cdot e^{(\sigma^0_{\max})^2 p^2 \sqrt{n}\cdot\frac{n}{2q}}\right)$$

$$\geq 1 - e^{-\frac{np}{4}} - \frac{2(\sigma^0_{\max})^2}{\delta(1-c^2)\sqrt{n}} - V_2,$$

where $q := \min_{1\leq j\leq p} m_j$ with $m_1, \ldots, m_p$ corresponding to the full model.

S3 Proofs

Proof of Theorem 3.9. Recall that

$$h\big(\boldsymbol{\alpha}_G, \{\sigma_j\}\big) := 1\left\{\frac{1}{2}\|\mathcal{Z}_G'\mathcal{W}^{-1}\mathcal{Z}_G(\boldsymbol{\alpha}_G - b_{\min})\|^2 \geq \varepsilon, \min_{1\leq j\leq p}\{m_j^g\} \geq d, \|A_g\|_2 \leq c\right\}$$

where $b_{\min}$ solves $\min_{b\in\mathbb{R}^{|G|}} \frac{1}{2}\|\mathcal{Z}_G'\mathcal{W}^{-1}\mathcal{Z}_G(\boldsymbol{\alpha}_G - b)\|^2$ subject to $\|b\|_0 \leq |G|-1$. By Lemma S2.2 the Jacobian term can be bounded to yield,

$$E\Big(h\big(\boldsymbol{\alpha}_G, \{\sigma_j\}\big)|\widetilde{\mathcal{D}}_g'\widetilde{\mathcal{D}}_g|^{\frac{1}{2}}\Big) \leq e^{\frac{1}{2}(1-c)^{-2}\left(r^g_{\max}+(1+c)^2\right)\frac{\|Y\|^2}{\sqrt{n}} - \frac{|G|+p}{2}} \cdot E\big[h\big(\boldsymbol{\alpha}_G, \{\sigma_j\}\big)\big].$$

Further,

$$\begin{aligned} E\big[h\big(\boldsymbol{\alpha}_G, \{\sigma_j\}\big)\big] &= P\Big(\frac{1}{2}\|\mathcal{Z}_G'\mathcal{W}^{-1}\mathcal{Z}_G(\boldsymbol{\alpha}_G - b_{\min})\|^2 \geq \varepsilon, \min_{1\leq j\leq p}\{m_j^g\} \geq d\Big) \\ &\leq P\Big(\frac{1}{2}\big\|\mathcal{Z}_G'\mathcal{W}^{-1}\mathcal{Z}_G\big(\boldsymbol{\alpha}_G - \widetilde{b}\big)\big\|^2 \geq \varepsilon, \min_{1\leq j\leq p}\{m_j^g\} \geq d\Big), \end{aligned}$$

where $\widetilde{b}$ solves $\min_{b\in\mathbb{R}^{|G|}} \big\|\big(E_x(\mathcal{Z}_G'\mathcal{Z}_G)\big)^{-1}E_x(\mathcal{Z}_G'Y) - b\big\|^2$ subject to $\|b\|_0 \leq |G|-1$. Recall that the least squares estimate $\widehat{\boldsymbol{\alpha}}_g := \big(\mathcal{Z}_G'\mathcal{Z}_G\big)^{-1}\mathcal{Z}_G'Y$, and apply the triangle inequality to get the following probabilistic bound.

$$
\begin{aligned}
&E\big[h\big(\boldsymbol{\alpha}_G,\{\sigma_j\}\big)\big] \\
&\quad\le P\Big(\|\mathcal{Z}_G'\mathcal{W}^{-1}\mathcal{Z}_G(\boldsymbol{\alpha}_G-\widehat{\boldsymbol{\alpha}}_g)\|\ge\sqrt{2\varepsilon}/3,\ \min_{1\le j\le p}\{m_j^g\}\ge d\Big) \\
&\qquad+P\bigg(\Big\|\mathcal{Z}_G'\mathcal{W}^{-1}\mathcal{Z}_G\Big[\widehat{\boldsymbol{\alpha}}_g-\big(E_x(\mathcal{Z}_G'\mathcal{Z}_G)\big)^{-1}E_x(\mathcal{Z}_G'Y)\Big]\Big\|\ge\sqrt{2\varepsilon}/3,\ \min_{1\le j\le p}\{m_j^g\}\ge d\bigg) \\
&\qquad+P\bigg(\Big\|\mathcal{Z}_G'\mathcal{W}^{-1}\mathcal{Z}_G\Big[\big(E_x(\mathcal{Z}_G'\mathcal{Z}_G)\big)^{-1}E_x(\mathcal{Z}_G'Y)-\widetilde{b}\Big]\Big\|\ge\sqrt{2\varepsilon}/3,\ \min_{1\le j\le p}\{m_j^g\}\ge d\bigg) \\
&\quad\le\frac{3|G|\sqrt{\Lambda_g}}{\sqrt{\pi\varepsilon}}e^{-\frac{\varepsilon}{9\Lambda_g}} \\
&\qquad+P\bigg(\sqrt{\Lambda_g}\cdot\big\|\widehat{\boldsymbol{\alpha}}_g-\big(E_x(\mathcal{Z}_G'\mathcal{Z}_G)\big)^{-1}E_x(\mathcal{Z}_G'Y)\big\|\ge\frac{\sqrt{2\varepsilon}}{3\sqrt{\Lambda_g}},\ \min_{1\le j\le p}\{m_j^g\}\ge d\bigg) \\
&\qquad+P\bigg(\sqrt{\Lambda_g}\cdot\big\|\big(E_x(\mathcal{Z}_G'\mathcal{Z}_G)\big)^{-1}E_x(\mathcal{Z}_G'Y)-\widetilde{b}\big\|\ge\frac{\sqrt{2\varepsilon}}{3\sqrt{\Lambda_g}},\ \min_{1\le j\le p}\{m_j^g\}\ge d\bigg),
\end{aligned}
$$

where the second inequality follows from Lemma 3.6, and recalling that $\Lambda_g := \mathrm{tr}\big(\mathcal{Z}_G'\mathcal{W}^{-1}\mathcal{Z}_G\big)$. Next, using the assumption that $\varepsilon=\Lambda_g\cdot\widetilde{\varepsilon}$ for some $\widetilde{\varepsilon}$ not depending on Σ or A_g, observe that

$$
\begin{aligned}
&E\big[h\big(\boldsymbol{\alpha}_G,\{\sigma_j\}\big)\big] \\
&\le\frac{3|G|\sqrt{\Lambda_g}}{\sqrt{\pi\varepsilon}}e^{-\frac{\varepsilon}{9\Lambda_g}}+2P\Big(\sqrt{\Lambda_g}\ge\sqrt{n^{1+\frac{\varrho}{2}}p^3},\ \min_{1\le j\le p}\{m_j^g\}\ge d\Big) \\
&+1\bigg\{\big\|\widehat{\boldsymbol{\alpha}}_g-\big(E_x(\mathcal{Z}_G'\mathcal{Z}_G)\big)^{-1}E_x(\mathcal{Z}_G'Y)\big\|\ge\frac{\sqrt{2\varepsilon}}{3\sqrt{n^{1+\frac{\varrho}{2}}p^3\Lambda_g}},\ \min_{1\le j\le p}\{m_j^g\}\ge d\bigg\} \\
&+1\bigg\{\big\|\big(E_x(\mathcal{Z}_G'\mathcal{Z}_G)\big)^{-1}E_x(\mathcal{Z}_G'Y)-\widetilde{b}\big\|\ge\frac{\sqrt{2\varepsilon}}{3\sqrt{n^{1+\frac{\varrho}{2}}p^3\Lambda_g}},\ \min_{1\le j\le p}\{m_j^g\}\ge d\bigg\},
\end{aligned}
$$

and applying Lemma 3.7 gives,

$$
\begin{aligned}
E\big[h\big(\boldsymbol{\alpha}_G,\{\sigma_j\}\big)\big]\le{}&\frac{3|G|\sqrt{\Lambda_g}}{\sqrt{\pi\varepsilon}}e^{-\frac{\varepsilon}{9\Lambda_g}}+e^{-\left(\frac{d\cdot n^{\frac{\varrho}{2}}p^2}{4\lambda_{\max}(\mathcal{X}\mathcal{X}'/n)}-\frac{np}{2}\right)}2^{-\frac{|G|}{2}+1} \\
&+1\bigg\{\big\|\widehat{\boldsymbol{\alpha}}_g-\big(E_x(\mathcal{Z}_G'\mathcal{Z}_G)\big)^{-1}E_x(\mathcal{Z}_G'Y)\big\|^2\ge\frac{2\varepsilon}{9n^{1+\frac{\varrho}{2}}p^3\Lambda_g}\bigg\} \\
&+1\bigg\{\big\|\big(E_x(\mathcal{Z}_G'\mathcal{Z}_G)\big)^{-1}E_x(\mathcal{Z}_G'Y)-\widetilde{b}\big\|^2\ge\frac{2\varepsilon}{9n^{1+\frac{\varrho}{2}}p^3\Lambda_g}\bigg\}.
\end{aligned}
$$

Hence, by Lemma 3.8 (and thus Condition 3.1) and Condition 3.3 it follows that for all $n\ge\max\{N_1,N_2\}$,

$$
E\big[h\big(\boldsymbol{\alpha}_G,\{\sigma_j\}\big)\big]\le\frac{3|G|\sqrt{\Lambda_g}}{\sqrt{\pi\varepsilon}}e^{-\frac{\varepsilon}{9\Lambda_g}}+e^{-\left(\frac{d\cdot n^{\frac{\varrho}{2}}p^2}{4\lambda_{\max}(\mathcal{X}\mathcal{X}'/n)}-\frac{np}{2}\right)}2^{-\frac{|G|}{2}+1},
$$

with probability exceeding $1 - V_1$ where V_1 is as defined in (9). ■

Proof of Theorem 3.10. Recall that

$$h\big(\boldsymbol{\alpha}_{G_o}, \{\sigma_j\}\big) := 1\left\{\frac{1}{2}\|\mathcal{Z}'_{G_o}\mathcal{W}^{-1}\mathcal{Z}_{G_o}(\boldsymbol{\alpha}_{G_o} - b_{\min})\|^2 \geq \varepsilon, \min_{1\leq j\leq p}\{m_j^{g_o}\} \geq d, \|A_{g_o}\|_2 \leq c\right\}$$

where $b_{\min}$ solves $\min_{b\in\mathbb{R}^{|G_o|}} \frac{1}{2}\|\mathcal{Z}'_{G_o}\mathcal{W}^{-1}\mathcal{Z}_{G_o}(\boldsymbol{\alpha}_{G_o} - b)\|^2$ s. t. $\|b\|_0 \leq |G_o| - 1$. To show the desired result, let $\widetilde{b}$ be the solution to

$$\min_{b\in\mathbb{R}^{|G_o|}} \big\|(\Gamma_n(0)\otimes(\Sigma^0)^{-1})_{G_o,G_o}(\boldsymbol{\alpha}^0_{G_o} - b)\big\|^2 \text{ subject to } \|b\|_0 \leq |G_o| - 1.$$

Then

$$\begin{aligned}
&\big\|(\Gamma_n(0)\otimes(\Sigma^0)^{-1})_{G_o,G_o}(\boldsymbol{\alpha}^0_{G_o} - \widetilde{b})\big\|^2 \\
&\quad \leq \big\|(\Gamma_n(0)\otimes(\Sigma^0)^{-1})_{G_o,G_o}(\boldsymbol{\alpha}^0_{G_o} - b_{\min})\big\|^2 \\
&\quad = \big\|(\Gamma_n(0)\otimes(\Sigma^0)^{-1})_{G_o,G_o}(\mathcal{Z}'_{G_o}\mathcal{W}^{-1}\mathcal{Z}_{G_o})^{-1}\mathcal{Z}'_{G_o}\mathcal{W}^{-1}\mathcal{Z}_{G_o}(\boldsymbol{\alpha}^0_{G_o} - b_{\min})\big\|^2 \\
&\quad \leq \big\|(\Gamma_n(0)\otimes(\Sigma^0)^{-1})_{G_o,G_o}(\mathcal{Z}'_{G_o}\mathcal{W}^{-1}\mathcal{Z}_{G_o})^{-1}\big\|_2^2 \cdot \|\mathcal{Z}'_{G_o}\mathcal{W}^{-1}\mathcal{Z}_{G_o}(\boldsymbol{\alpha}^0_{G_o} - b_{\min})\|^2 \\
&\quad \leq \frac{\big\|(\Gamma_n(0)\otimes(\Sigma^0)^{-1})_{G_o,G_o}\big\|_2^2}{\lambda^2_{\min}(\mathcal{Z}'_{G_o}\mathcal{W}^{-1}\mathcal{Z}_{G_o})} \cdot \|\mathcal{Z}'_{G_o}\mathcal{W}^{-1}\mathcal{Z}_{G_o}(\boldsymbol{\alpha}^0_{G_o} - b_{\min})\|^2,
\end{aligned}$$

and so for $\rho \in (0, \frac{1}{2})$,

$$\begin{aligned}
1&\left\{\big\|(\Gamma_n(0)\otimes(\Sigma^0)^{-1})_{G_o,G_o}(\boldsymbol{\alpha}^0_{G_o} - \widetilde{b})\big\|^2 \geq \frac{18\varepsilon}{n^{1-\rho}p^2\Lambda_{g_o}}\right\} \cdot P\big(\|A_{g_o}\|_2 \leq c\big) \\
&\leq P\left(\frac{\big\|(\Gamma_n(0)\otimes(\Sigma^0)^{-1})_{G_o,G_o}\big\|_2^2}{\lambda^2_{\min}(\mathcal{Z}'_{G_o}\mathcal{W}^{-1}\mathcal{Z}_{G_o})} \geq \frac{1}{n^{1-\rho}p^2\Lambda_{g_o}}, \|A_{g_o}\|_2 \leq c\right) \\
&\quad + P\Big(\|\mathcal{Z}'_{G_o}\mathcal{W}^{-1}\mathcal{Z}_{G_o}(\boldsymbol{\alpha}^0_{G_o} - b_{\min})\|^2 \geq 18\varepsilon, \|A_{g_o}\|_2 \leq c\Big) \\
&\leq P\left(\frac{\big\|(\Gamma_n(0)\otimes(\Sigma^0)^{-1})_{G_o,G_o}\big\|_2^2}{\lambda^2_{\min}(\mathcal{Z}'_{G_o}\mathcal{W}^{-1}\mathcal{Z}_{G_o})} \geq \frac{1}{n^{1-\rho}p^2\Lambda_{g_o}}\right) \\
&\quad + P\Big(\big\|\mathcal{Z}'_{G_o}\mathcal{W}^{-1}\mathcal{Z}_{G_o}(\boldsymbol{\alpha}^0_{G_o} - \widehat{\boldsymbol{\alpha}}_{g_o})\big\|^2 \geq 2\varepsilon\Big) \\
&\quad + P\Big(\big\|\mathcal{Z}'_{G_o}\mathcal{W}^{-1}\mathcal{Z}_{G_o}(\widehat{\boldsymbol{\alpha}}_{g_o} - \boldsymbol{\alpha}_{G_o})\big\|^2 \geq 2\varepsilon\Big) \\
&\quad + P\Big(\big\|\mathcal{Z}'_{G_o}\mathcal{W}^{-1}\mathcal{Z}_{G_o}(\boldsymbol{\alpha}_{G_o} - b_{\min})\big\|^2 \geq 2\varepsilon, \|A_{g_o}\|_2 \leq c\Big)
\end{aligned}$$

where $\widehat{\boldsymbol{\alpha}}_{g_o} := \big(\mathcal{Z}'_{G_o}\mathcal{Z}_{G_o}\big)^{-1}\mathcal{Z}'_{G_o}Y$ is the least squares estimator. Applying Lemma 3.6, multiplying both sides of the inequality by $1\big\{\min_{1\leq j\leq p}\{m_j^{g_o}\} \geq d\big\}$, and denoting $\omega := P\big(\|A_{g_o}\|_2 \leq c\big)$

gives,

$$
\begin{aligned}
1\Big\{ \big\|(\Gamma_n(0)\otimes(\Sigma^0)^{-1})_{G_o,G_o}(\boldsymbol{\alpha}^0_{G_o}-\widetilde{b})\big\|^2 \geq \frac{18\varepsilon}{n^{1-\rho}p^2\Lambda_{g_o}}, \min_{1\leq j\leq p}\{m_j^{g_o}\}\geq d\Big\}\cdot\omega \\
\leq P\bigg(\Lambda_{g_o}\cdot\frac{\big\|(\Gamma_n(0)\otimes(\Sigma^0)^{-1})_{G_o,G_o}\big\|_2^2}{\lambda^2_{\min}(\mathcal{Z}'_{G_o}\mathcal{W}^{-1}\mathcal{Z}_{G_o})}\geq\frac{1}{n^{1-\rho}p^2}, \min_{1\leq j\leq p}\{m_j^{g_o}\}\geq d\bigg) \\
+P\bigg(\Lambda_{g_o}\cdot\|\widehat{\boldsymbol{\alpha}}_{g_o}-\boldsymbol{\alpha}^0_{G_o}\|^2\geq\frac{2\varepsilon}{\Lambda_{g_o}}, \min_{1\leq j\leq p}\{m_j^{g_o}\}\geq d\bigg) \\
+\frac{|G_o|\sqrt{\Lambda_{g_o}}}{\sqrt{\pi\varepsilon}}e^{-\frac{\varepsilon}{\Lambda_{g_o}}}+E\big[h\big(\boldsymbol{\alpha}_{G_o},\{\sigma_j\}\big)\big] \\
\leq P\bigg(\frac{\big\|(\Gamma_n(0)\otimes(\Sigma^0)^{-1})_{G_o,G_o}\big\|_2^2}{\lambda^2_{\min}(\mathcal{Z}'_{G_o}\mathcal{W}^{-1}\mathcal{Z}_{G_o})}\geq\frac{1}{n^{2-\frac{\rho}{2}}p^5}, \min_{1\leq j\leq p}\{m_j^{g_o}\}\geq d\bigg) \\
+2P\bigg(\Lambda_{g_o}\geq n^{1+\frac{\rho}{2}}p^3, \min_{1\leq j\leq p}\{m_j^{g_o}\}\geq d\bigg) \\
+1\Big\{\|\widehat{\boldsymbol{\alpha}}_{g_o}-\boldsymbol{\alpha}^0_{G_o}\|^2\geq\frac{2\varepsilon}{9n^{1+\frac{\rho}{2}}p^3\Lambda_{g_o}}\Big\} \\
+\frac{|G_o|\sqrt{\Lambda_{g_o}}}{\sqrt{\pi\varepsilon}}e^{-\frac{\varepsilon}{\Lambda_{g_o}}}+E\big[h\big(\boldsymbol{\alpha}_{G_o},\{\sigma_j\}\big)\big].
\end{aligned}
$$

Next, by Condition 3.2 the indicator function on the left side is equal to 1, and for all $n\geq N_2$, by Lemmas 3.7 and S2.1 (and accordingly Condition 3.1),

$$
\begin{aligned}
\omega\leq\frac{p^{\frac{7}{2}}}{n^{\frac{\rho}{4}}}\cdot\frac{(\sigma^0_{\max})^2}{(\sigma^0_{\min})^2}\cdot\frac{6\|\Gamma_n(0)\|_2/\delta}{(1-|G_o|/n-2/n)}+e^{-\left(\frac{d\cdot n^{\frac{\rho}{2}}p^2}{4\lambda_{\max}(\mathcal{X}\mathcal{X}'/n)}-\frac{np}{2}\right)}2^{-\frac{|G_o|}{2}+1} \\
+1\Big\{\|\widehat{\boldsymbol{\alpha}}_{g_o}-\boldsymbol{\alpha}^0_{G_o}\|^2\geq\frac{2\varepsilon}{9n^{1+\frac{\rho}{2}}p^3\Lambda_{g_o}}\Big\}+\frac{|G_o|\sqrt{\Lambda_{g_o}}}{\sqrt{\pi\varepsilon}}e^{-\frac{\varepsilon}{\Lambda_{g_o}}}+E\big[h\big(\boldsymbol{\alpha}_{G_o},\{\sigma_j\}\big)\big],
\end{aligned}
$$

with probability exceeding $1-V_2-e^{-\frac{np}{4}}$, where V_2 is as in (10). Then, by observing that $\boldsymbol{\alpha}^0_{G_o}=\big(E_x(\mathcal{Z}'_{G_o}\mathcal{Z}_{G_o})\big)^{-1}E_x(\mathcal{Z}'_{G_o}Y)$ and applying Lemma 3.8 yields for all $n\geq\max\{N_1,N_2\}$,

$$
\begin{aligned}
\omega\leq\frac{p^{\frac{7}{2}}}{n^{\frac{\rho}{4}}}\cdot\frac{(\sigma^0_{\max})^2}{(\sigma^0_{\min})^2}\cdot\frac{6\|\Gamma_n(0)\|_2/\delta}{(1-|G_o|/n-2/n)}+e^{-\left(\frac{d\cdot n^{\frac{\rho}{2}}p^2}{4\lambda_{\max}(\mathcal{X}\mathcal{X}'/n)}-\frac{np}{2}\right)}2^{-\frac{|G_o|}{2}+1} \\
+\frac{|G_o|\sqrt{\Lambda_{g_o}}}{\sqrt{\pi\varepsilon}}e^{-\frac{\varepsilon}{\Lambda_{g_o}}}+E\big[h\big(\boldsymbol{\alpha}_{G_o},\{\sigma_j\}\big)\big],
\end{aligned}
\tag{S2}
$$

with probability exceeding $1-V_1-V_2-e^{-\frac{np}{4}}$, where V_1 is as in (9).

Finally, Condition 3.1 allows for the following probabilistic bound on ω.

$$\begin{aligned}
1-\omega &:= P\big(\|A_{g_o}\|_2 > c\big) \\
&\le P\big(\|A_{g_o} - \widehat{A}_{g_o}\|_2 > c\big) + 1\big\{\|\widehat{A}_{g_o} - A^0\|_2 > c\big\} + 1\big\{\|A^0\|_2 > c\big\} \\
&\le P\big(\|A_{g_o} - \widehat{A}_{g_o}\|_F > c\big) + 1\big\{\|\widehat{A}_{g_o} - A^0\|_F > c\big\} \\
&= P\big(\|\boldsymbol{\alpha}_{G_o} - \widehat{\boldsymbol{\alpha}}_{g_o}\|^2 > c^2\big) + 1\big\{\|\widehat{\boldsymbol{\alpha}}_{g_o} - \boldsymbol{\alpha}^0_{G_o}\|^2 > c^2\big\} \\
&= P\big(\|\boldsymbol{\alpha}_{G_o} - \widehat{\boldsymbol{\alpha}}_{g_o}\|^2 > c^2\big),
\end{aligned}$$

with probability exceeding $1-\widetilde{V}_1$ by Lemma 3.8 where $\widetilde{V}_1$ is V_1 as in (9) with ε replaced by $c^2 \cdot \frac{9n^{1+\frac{\rho}{2}}p^3\Lambda_{g_o}}{2}$. Further,

$$\begin{aligned}
P\big(\|\boldsymbol{\alpha}_{G_o} - \widehat{\boldsymbol{\alpha}}_{g_o}\|^2 > c^2\big) &= P\big(\|(\mathcal{Z}'_{G_o}\mathcal{W}^{-1}\mathcal{Z}_{G_o})^{-\frac{1}{2}}(\mathcal{Z}'_{G_o}\mathcal{W}^{-1}\mathcal{Z}_{G_o})^{\frac{1}{2}}(\boldsymbol{\alpha}_{G_o} - \widehat{\boldsymbol{\alpha}}_{g_o})\|^2 > c^2\big) \\
&\le P\Big(\lambda_{\max}\big((\mathcal{Z}'_{G_o}\mathcal{W}^{-1}\mathcal{Z}_{G_o})^{-1}\big)\|Z\|^2 > c^2\Big)
\end{aligned}$$

where, as in the proof of Lemma 3.6, $Z := (\mathcal{Z}'_{G_o}\mathcal{W}^{-1}\mathcal{Z}_{G_o})^{\frac{1}{2}}(\boldsymbol{\alpha}_{G_o} - \widehat{\boldsymbol{\alpha}}_{g_o}) \sim \mathrm{N}_{|G_o|}(0, I_{|G_o|})$. Then

$$\begin{aligned}
P\big(\|\boldsymbol{\alpha}_{G_o} - \widehat{\boldsymbol{\alpha}}_{g_o}\|^2 > c^2\big) &\le P\Big(\|Z\|^2 > c^2\lambda_{\min}(\mathcal{Z}'_{G_o}\mathcal{W}^{-1}\mathcal{Z}_{G_o})\Big) \\
&\le P\Big(\|Z\|^2 > c^2\sqrt{n}\Big) + P\Big(\lambda_{\min}(\mathcal{Z}'_{G_o}\mathcal{W}^{-1}\mathcal{Z}_{G_o}) < \sqrt{n}\Big).
\end{aligned}$$

Observe that

$$\begin{aligned}
P\Big(\|Z\|^2 > c^2\sqrt{n}\Big) &\le P\Big(|G_o| \max_{1\le i\le |G_o|}\{Z_i^2\} > c^2\sqrt{n}\Big) \\
&\le \sum_{i=1}^{|G_o|} P\Big(|Z_i| > cn^{\frac{1}{4}}|G_o|^{-\frac{1}{2}}\Big) \\
&= \sum_{i=1}^{|G_o|} 2P\Big(Z_i < -cn^{\frac{1}{4}}|G_o|^{-\frac{1}{2}}\Big) \\
&\le 2|G_o|\frac{|G_o|^{\frac{1}{2}}}{cn^{\frac{1}{4}}\sqrt{2\pi}} e^{\frac{-c^2\sqrt{n}}{2|G_o|}},
\end{aligned}$$

where the last inequality follows because for the standard normal CDF, Φ for $x > 0$, $\Phi(-x) \le \frac{1}{x\sqrt{2\pi}}e^{\frac{-x^2}{2}}$. Additionally, by the same arguments as in the proof of Lemma S2.1,

$$P\Big(\lambda_{\min}(\mathcal{Z}'_{G_o}\mathcal{W}^{-1}\mathcal{Z}_{G_o}) < \sqrt{n}\Big) \le \frac{6(\sigma^0_{\max})^2 p}{\delta\sqrt{n}(1-|G_o|/n - 2/n)},$$

with probability exceeding $1 - V_2 - e^{-\frac{np}{4}}$, where V_2 is as in (10). Thus,

$$\omega \ge 1 - \frac{\sqrt{2}|G_o|^{1+\frac{1}{2}}}{cn^{\frac{1}{4}}\sqrt{\pi}} e^{\frac{-c^2\sqrt{n}}{2|G_o|}} - \frac{6(\sigma^0_{\max})^2 p}{\delta\sqrt{n}(1-|G_o|/n-2/n)},$$

with probability exceeding $1 - \widetilde{V}_1 - V_2 - e^{-\frac{np}{4}}$.

Therefore, substituting back into (S2) gives for all $n \geq \max\{N_1, N_2\}$,

$$
\begin{aligned}
1 \leq \frac{\sqrt{2}|G_o|^{1+\frac{1}{2}}}{cn^{\frac{1}{4}}\sqrt{\pi}} e^{\frac{-c^2\sqrt{n}}{2|G_o|}} + \frac{6(\sigma^0_{\max})^2 p}{\delta\sqrt{n}(1 - |G_o|/n - 2/n)} \\
+ \frac{p^{\frac{7}{2}}}{n^{\frac{\rho}{4}}} \cdot \frac{(\sigma^0_{\max})^2}{(\sigma^0_{\min})^2} \cdot \frac{6\|\Gamma_n(0)\|_2/\delta}{(1 - |G_o|/n - 2/n)} + e^{-\left(\frac{d \cdot n^{\frac{\rho}{2}} p^2}{4\lambda_{\max}(\mathcal{X}\mathcal{X}'/n)} - \frac{np}{2}\right)} 2^{-\frac{|G_o|}{2}+1} \\
+ \frac{|G_o|\sqrt{\Lambda_{g_o}}}{\sqrt{\pi\varepsilon}} e^{-\frac{\varepsilon}{\Lambda_{g_o}}} + E\big[h\big(\boldsymbol{\alpha}_{G_o}, \{\sigma_j\}\big)\big],
\end{aligned}
$$

with probability exceeding $1 - V_1 - \widetilde{V}_1 - 2V_2 - 2e^{-\frac{np}{4}}$. Lastly, for any fixed $K_3 \in (0,1)$, by Condition 3.4, choose a positive constant N_3 such that for all $n \geq \max\{N_1, N_2, N_3\}$,

$$
1 \leq K_3 + E\big[h\big(\boldsymbol{\alpha}_{G_o}, \{\sigma_j\}\big)\big], \tag{S3}
$$

with probability exceeding $1 - V_1 - \widetilde{V}_1 - 2V_2 - 2e^{-\frac{np}{4}}$. Multiplying both sides of the inequality by $e^{\frac{|G_o|+p}{4}}$, and applying Lemma S2.4 yields the desired result. ∎

Proof of Lemma 3.6. It follows from the generalized fiducial distributional of $\boldsymbol{\alpha}_G$ that

$$
Z := (\mathcal{Z}_G'\mathcal{W}^{-1}\mathcal{Z}_G)^{\frac{1}{2}}(\boldsymbol{\alpha}_G - \widehat{\boldsymbol{\alpha}}_g) \sim \mathrm{N}_{|G|}(0, I_{|G|}).
$$

Thus,

$$
\|\mathcal{Z}_G'\mathcal{W}^{-1}\mathcal{Z}_G(\boldsymbol{\alpha}_G - \widehat{\boldsymbol{\alpha}}_g)\|^2 = \|LZ\|^2 = \|UDV'Z\|^2
$$

where $L := (\mathcal{Z}_G'\mathcal{W}^{-1}\mathcal{Z}_G)^{\frac{1}{2}}$ is a $|G| \times |G|$ matrix which has the singular value decomposition $L = QDV'$ for Q and V each orthogonal matrices. Since V is an orthogonal matrix and Z follows the standard multivariate normal distribution, $\widetilde{Z} := V'Z \sim \mathrm{N}(0, I_{|G|})$. Then

$$
\|LZ\|^2 = \|UD\widetilde{Z}\|^2 = \sum_{j=1}^{|G|} \widetilde{Z}_j^2 \lambda_j \leq \Lambda_g \max \widetilde{Z}_j^2
$$

where $\Lambda_g = \sum_{j=1}^{|G|} \lambda_j$, and λ_j is the jth eigenvalue of $L'L$. In other words,

$$
\Lambda_g = \mathrm{tr}(L'L) = \mathrm{tr}\big(\mathcal{Z}_G'\mathcal{W}^{-1}\mathcal{Z}_G\big).
$$

Then using the assumption that $\varepsilon = \Lambda_g \cdot \widetilde{\varepsilon}$ for some $\widetilde{\varepsilon}$ not depending on Σ or A_g,

$$\begin{aligned}
P\Big(\|LZ\|^2 \geq \varepsilon\Big) &\leq P\Big(\Lambda_g \max \widetilde{Z}_j^2 \geq \varepsilon\Big) \\
&\leq \sum_{j=1}^{|G|} P\Big(|Z_j| \geq \Lambda_g^{-\frac{1}{2}} \varepsilon^{\frac{1}{2}}\Big) \\
&= 2\sum_{j=1}^{|G|} P\Big(Z_j \leq -\Lambda_g^{-\frac{1}{2}} \varepsilon^{\frac{1}{2}}\Big) \leq \frac{|G|\sqrt{2\Lambda_g}}{\sqrt{\pi\varepsilon}} e^{-\frac{\varepsilon}{2\Lambda_g}},
\end{aligned}$$

where the last inequality follows because for the standard normal CDF, Φ for $x > 0$, $\Phi(-x) \leq \frac{1}{x\sqrt{2\pi}} e^{\frac{-x^2}{2}}$. ■

Proof of Lemma 3.7.

$$\begin{aligned}
&P\Big(\Lambda_g \geq n^{1+\frac{\varrho}{2}} p^3, \min_{1\leq j\leq p}\{m_j^g\} \geq d\Big) \\
&\qquad = P\Big(\mathrm{tr}\big(\mathcal{Z}_G' \mathcal{W}^{-1} \mathcal{Z}_G\big) \geq n^{1+\frac{\varrho}{2}} p^3\Big) 1\Big\{\min_{1\leq j\leq p}\{m_j^g\} \geq d\Big\} \\
&\qquad \leq P\Big(\mathrm{tr}\big(\mathcal{Z}' \mathcal{W}^{-1} \mathcal{Z}\big) \geq n^{1+\frac{\varrho}{2}} p^3\Big) 1\Big\{\min_{1\leq j\leq p}\{m_j^g\} \geq d\Big\} \\
&\qquad = P\Big(\mathrm{tr}\big((\mathcal{X}\mathcal{X}') \otimes (\Sigma^{-1})\big) \geq n^{1+\frac{\varrho}{2}} p^3\Big) 1\Big\{\min_{1\leq j\leq p}\{m_j^g\} \geq d\Big\} \\
&\qquad = P\Bigg(\mathrm{tr}(\mathcal{X}\mathcal{X}') \sum_{j=1}^{p} \sigma_j^{-2} \geq n^{1+\frac{\varrho}{2}} p^3\Bigg) 1\Big\{\min_{1\leq j\leq p}\{m_j^g\} \geq d\Big\} \\
&\qquad \leq P\Bigg(\lambda_{\max}(\mathcal{X}\mathcal{X}') \sum_{j=1}^{p} \sigma_j^{-2} \geq n^{1+\frac{\varrho}{2}} p^2\Bigg) 1\Big\{\min_{1\leq j\leq p}\{m_j^g\} \geq d\Big\}.
\end{aligned}$$

Then for any $t \in \big(0, \min_{1\leq j\leq p}\{m_j^g\}/2\big)$, recalling from the generalized fiducial density that $\sigma_j^2 \sim$ inv-gamma$\big((n - |r_j^g|)/2, m_j^g/2\big)$, the Chernoff bound gives,

$$\begin{aligned}
&P\Big(\Lambda_g \geq n^{1+\frac{\varrho}{2}} p^3, \min_{1\leq j\leq p}\{m_j^g\} \geq d\Big) \\
&\qquad \leq P\Big(e^{t\sum_{j=1}^{p}\sigma_j^{-2}} \geq e^{\frac{t\cdot n^{1+\frac{\varrho}{2}} p^2}{\lambda_{\max}(\mathcal{X}\mathcal{X}')}}\Big) 1\Big\{\min_{1\leq j\leq p}\{m_j^g\} \geq d\Big\} \\
&\qquad \leq e^{-\frac{t\cdot n^{1+\frac{\varrho}{2}} p^2}{\lambda_{\max}(\mathcal{X}\mathcal{X}')}} E\Big(e^{t\sum_{j=1}^{p}\sigma_j^{-2}}\Big) 1\Big\{\min_{1\leq j\leq p}\{m_j^g\} \geq d\Big\} \\
&\qquad = e^{-\frac{t\cdot n^{1+\frac{\varrho}{2}} p^2}{\lambda_{\max}(\mathcal{X}\mathcal{X}')}} \prod_{j=1}^{p} \left(1 - \frac{2t}{m_j^g}\right)^{-\frac{n-|r_j^g|}{2}} 1\Big\{\min_{1\leq j\leq p}\{m_j^g\} \geq d\Big\} \\
&\qquad \leq e^{-\frac{t\cdot n^{1+\frac{\varrho}{2}} p^2}{\lambda_{\max}(\mathcal{X}\mathcal{X}')}} \left(1 - \frac{2t}{d}\right)^{-\frac{np-|G|}{2}}.
\end{aligned}$$

Next, taking $t = d/4$ yields the desired result,

$$P\Big(\Lambda_g \geq n^{1+\frac{\rho}{2}} p^3, \min_{1\leq j\leq p}\{m_j^g\} \geq d\Big) \leq e^{-\frac{d \cdot n^{1+\frac{\rho}{2}} p^2}{4\lambda_{\max}(\mathcal{X}\mathcal{X}')}} 2^{\frac{np-|G|}{2}}$$
$$\leq e^{-\left(\frac{d\cdot n^{\frac{\rho}{2}} p^2}{4\lambda_{\max}(\mathcal{X}\mathcal{X}'/n)} - \frac{np}{2}\right)} 2^{-\frac{|G|}{2}}.$$

■

Proof of Lemma 3.8. First consider the expressions,

$$\begin{aligned}
E_x(\mathcal{Z}'Y) &= E_x(\mathcal{Z}'\mathcal{Z})\boldsymbol{\alpha}^0 + E_x\big(\mathcal{Z}'\mathrm{vec}\big((\Sigma^0)^{\frac{1}{2}}\mathcal{U}\big)\big)\\
&= \big(E_x(\mathcal{X}\mathcal{X}') \otimes I_p\big)\boldsymbol{\alpha}^0 + E_x\big(\mathrm{vec}\big((\Sigma^0)^{\frac{1}{2}}\mathcal{U}\mathcal{X}'\big)\big)\\
&= (n\Gamma_n(0)\otimes I_p)\boldsymbol{\alpha}^0 + \mathrm{vec}\Big((\Sigma^0)^{\frac{1}{2}} \sum_{t=1}^{n} E_x\big(U^{(t)}X^{(t-1)'}\big)\Big)\\
&= (n\Gamma_n(0)\otimes I_p)\boldsymbol{\alpha}^0,
\end{aligned} \tag{S4}$$

and

$$\begin{aligned}
E_x\big(Y'\mathcal{Z}\mathcal{Z}'Y\big) &= E_x\Big(\big(\boldsymbol{\alpha}^{0'}\mathcal{Z}'\mathcal{Z} + \mathrm{vec}\big((\Sigma^0)^{\frac{1}{2}}\mathcal{U}\big)'\mathcal{Z}\big)\big(\mathcal{Z}'\mathcal{Z}\boldsymbol{\alpha}^0 + \mathcal{Z}'\mathrm{vec}\big((\Sigma^0)^{\frac{1}{2}}\mathcal{U}\big)\big)\Big)\\
&= \boldsymbol{\alpha}^{0'}E_x\big(\mathcal{Z}'\mathcal{Z}\mathcal{Z}'\mathcal{Z}\big)\boldsymbol{\alpha}^0\\
&\quad + 2E_x\big(\mathrm{vec}\big((\Sigma^0)^{\frac{1}{2}}\mathcal{U}\big)'\mathcal{Z}\mathcal{Z}'\mathcal{Z}\big)\boldsymbol{\alpha}^0 + E_x\big(\mathrm{vec}\big((\Sigma^0)^{\frac{1}{2}}\mathcal{U}\big)'\mathcal{Z}\mathcal{Z}'\mathrm{vec}\big((\Sigma^0)^{\frac{1}{2}}\mathcal{U}\big)\big)\\
&\leq \lambda_{\max}\Big(E_x\big((\mathcal{X}\mathcal{X}')^2\big)\otimes I_p\Big)\|\boldsymbol{\alpha}^0\|^2\\
&\quad + 2E_x\big(\mathrm{vec}\big((\Sigma^0)^{\frac{1}{2}}\mathcal{U}\mathcal{X}'\mathcal{X}\mathcal{X}'\big)\big)'\boldsymbol{\alpha}^0 + E_x\big(\mathrm{tr}(\mathcal{X}\mathcal{U}'\Sigma^0\mathcal{U}\mathcal{X}')\big)\\
&\leq \lambda_{\max}\Big(E_x\big((\mathcal{X}\mathcal{X}')^2\big)\Big)\|\boldsymbol{\alpha}^0\|^2\\
&\quad + 2E_x\big(\mathrm{tr}(\mathcal{X}\mathcal{X}'\mathcal{X}\mathcal{U}'(\Sigma^0)^{\frac{1}{2}}A^0)\big) + (\sigma^0_{\max})^2 E_x\big(\mathrm{tr}(\mathcal{U}\mathcal{X}'\mathcal{X}\mathcal{U}')\big)
\end{aligned} \tag{S5}$$

These will be needed shortly.

Notice that $E_x(\mathcal{Z}_G'\mathcal{Z}_G) = (n\Gamma_n(0)\otimes I_p)_{G,G}$. Then for $K_g := (n\Gamma_n(0)\otimes I_p)_{G,G}\big(\mathcal{Z}_G'\mathcal{Z}_G\big)^{-1}$,

$$P_x\bigg(\|\widehat{\boldsymbol{\alpha}}_g - \big(E_x(\mathcal{Z}_G'\mathcal{Z}_G)\big)^{-1}E_x(\mathcal{Z}_G'Y)\|^2 \geq \frac{2\varepsilon}{9n^{1+\frac{\rho}{2}}p^3\Lambda_g}\bigg)$$
$$\leq P_x\bigg(\frac{\|K_g\mathcal{Z}_G'Y - E_x(\mathcal{Z}_G'Y)\|^2}{\lambda^2_{\min}\big((\Gamma_n(0)\otimes I_p)_{G,G}\big)} \geq \frac{2n^{1-\frac{\rho}{2}}\varepsilon}{9p^3\Lambda_g}\bigg)$$
$$\leq P_x\bigg(\frac{\|K_g\mathcal{Z}_G'Y - E_x(\mathcal{Z}_G'Y)\|^2}{\delta^2} \geq \frac{2n^{1-\frac{\rho}{2}}\varepsilon}{9p^3\Lambda_g}\bigg),$$

since by Condition 3.1,

$$\lambda^2_{\min}\big((\Gamma_n(0)\otimes I_p)_{G,G}\big) \geq \lambda^2_{\min}\big(\Gamma_n(0)\otimes I_p\big) = \lambda^2_{\min}\big(\Gamma_n(0)\big)\cdot 1 \geq \delta^2,$$

where the first inequality follows from the Poincaré separation theorem (see Rao (1979)) because $(\Gamma_n(0) \otimes I_p)_{G,G}$ is a principal submatrix of $\Gamma_n(0) \otimes I_p$.

Accordingly, for now set $\xi = \frac{2\delta^2}{9\Lambda_g}\varepsilon$. Then by the triangle and Markov inequalities,

$$
\begin{aligned}
P_x\bigg(\|K_g \mathcal{Z}_G' Y - E_x(\mathcal{Z}_G' Y)\|^2 \geq \frac{n^{1-\frac{\rho}{2}}\xi}{p^3} \bigg) &\leq P_x\bigg(\|(K_g - I_{|G|})\mathcal{Z}_G' Y\|^2 \geq \frac{n^{1-\frac{\rho}{2}}\xi}{4p^3} \bigg) \\
&\quad + P_x\bigg(\|\mathcal{Z}_G' Y - E_x(\mathcal{Z}_G' Y)\|^2 \geq \frac{n^{1-\frac{\rho}{2}}\xi}{4p^3} \bigg) \\
&\leq P_x\bigg(\|K_g - I_{|G|}\|_2^2 \cdot \|\mathcal{Z}_G' Y\|^2 \geq \frac{n^{1-\frac{\rho}{2}}\xi}{4p^3} \bigg) \\
&\quad + \frac{4p^3}{n^{1-\frac{\rho}{2}}\xi} E_x\big(\|\mathcal{Z}_G' Y - E_x(\mathcal{Z}_G' Y)\|^2\big).
\end{aligned} \tag{S6}
$$

Consider each right side quantity in turn. The first can be bounded as

$$
\begin{aligned}
&P_x\bigg(\|K_g - I_{|G|}\|_2^2 \cdot \|\mathcal{Z}_G' Y\|^2 \geq \frac{n^{1-\frac{\rho}{2}}\xi}{4p^3} \bigg) \\
&\qquad = P_x\bigg(\Big\| \Big[(n\Gamma_n(0) \otimes I_p)_{G,G} - \mathcal{Z}_G' \mathcal{Z}_G \Big] \big(\mathcal{Z}_G' \mathcal{Z}_G\big)^{-1} \Big\|_2^2 \cdot \|\mathcal{Z}_G' Y\|^2 \geq \frac{n^{1-\frac{\rho}{2}}\xi}{4p^3} \bigg) \\
&\qquad \leq P_x\bigg(\big\|(\Gamma_n(0) \otimes I_p)_{G,G} - \mathcal{Z}_G' \mathcal{Z}_G / n\big\|_2^2 \cdot \big\|\big(\mathcal{Z}_G' \mathcal{Z}_G / n\big)^{-1}\big\|_2^2 \geq n^{-2\rho} \bigg) \\
&\qquad\qquad + P_x\bigg(\|\mathcal{Z}_G' Y\|^2 \geq \frac{n^{1+\frac{3\rho}{2}}\xi}{4p^3} \bigg) \\
&\qquad \leq P_x\Big(\big\|(\Gamma_n(0) \otimes I_p)_{G,G} - \mathcal{Z}_G' \mathcal{Z}_G / n\big\|_2^2 \geq n^{-2\rho} \lambda_{\min}^2\big(\mathcal{Z}_G' \mathcal{Z}_G / n\big) \Big) \\
&\qquad\qquad + \frac{4p^3}{n^{1+\frac{3\rho}{2}}\xi} E_x\big(\|\mathcal{Z}_G' Y\|^2\big) \\
&\qquad \leq P_x\Big(\big\|(\Gamma_n(0) \otimes I_p)_{G,G} - \mathcal{Z}_G' \mathcal{Z}_G / n\big\|_2^2 \geq n^{-2\rho} \delta^2/4 \Big) \\
&\qquad\qquad + P_x\Big(\lambda_{\min}^2\big(\mathcal{Z}_G' \mathcal{Z}_G / n\big) < \delta^2/4 \Big) + \frac{4p^3}{n^{1+\frac{3\rho}{2}}\xi} E_x\big(\|\mathcal{Z}' Y\|^2\big)
\end{aligned}
$$

Then, again by the Poincaré separation theorem,

$$
\begin{aligned}
&P_x\bigg(\|K_g - I_{|G|}\|_2^2 \cdot \|\mathcal{Z}_G' Y\|^2 \geq \frac{n^{1-\frac{\rho}{2}}\xi}{4p^3} \bigg) \\
&\qquad \leq P_x\Big(\lambda_{\max}^2\big(\Gamma_n(0) \otimes I_p - \mathcal{Z}' \mathcal{Z} / n\big) \geq n^{-2\rho} \delta^2/4 \Big) \\
&\qquad\qquad + P_x\Big(\lambda_{\min}^2\big(\mathcal{Z}' \mathcal{Z} / n\big) < \delta^2/4 \Big) + \frac{4p^3}{n^{1+\frac{3\rho}{2}}\xi} E_x\big(\|\mathcal{Z}' Y\|^2\big) \\
&\qquad \leq P_x\Big(\lambda_{\max}^2\big(\Gamma_n(0) - \mathcal{X}\mathcal{X}' / n\big) \cdot 1 \geq n^{-2\rho} \delta^2/4 \Big) \\
&\qquad\qquad + P_x\Big(\lambda_{\min}^2\big(\mathcal{X}\mathcal{X}' / n\big) \cdot 1 < \delta^2/4 \Big) + \frac{4p^3}{n^{1+\frac{3\rho}{2}}\xi} E_x\big(\|\mathcal{Z}' Y\|^2\big),
\end{aligned}
$$

and by Lemma 3.19 (assuming Condition 3.1), for all $n \geq N_2$,

$$\begin{aligned}
P_x\bigg(\|K_g - I_{|G|}\|_2^2 \cdot \|\mathcal{Z}_G' Y\|^2 \geq \frac{n^{1-\frac{\rho}{2}}\xi}{4p^3}\bigg) \\
&\leq P_x\Big(\big\|\Gamma_n(0) - \mathcal{X}\mathcal{X}'/n\big\|_F^2 \geq n^{-2\rho}\delta^2/4\Big) \\
&\quad + 4\delta^{-2}\frac{(\sigma_{\max}^0)^4(1+c^2)}{(1-c^2)^3} \cdot \frac{2\min\{|G_o|, p\}^2}{n} + \frac{4p^3}{n^{1+\frac{3\rho}{2}}\xi} E_x\big(\|\mathcal{Z}' Y\|^2\big).
\end{aligned}$$

Next, applying the Markov inequality gives,

$$\begin{aligned}
P_x\Big(\big\|\Gamma_n(0) - \mathcal{X}\mathcal{X}'/n\big\|_F^2 \geq n^{-2\rho}\delta^2/4\Big) &\leq 4\delta^{-2}n^{2\rho}E_x\Big(\big\|\Gamma_n(0) - \mathcal{X}\mathcal{X}'/n\big\|_F^2\Big) \\
&= 4\delta^{-2}n^{2\rho}\mathrm{tr}\Big(\frac{1}{n^2}E_x\big((\mathcal{X}\mathcal{X}')^2\big) - \Gamma_n^2(0)\Big) \\
&\leq 4\delta^{-2}n^{2\rho}\frac{(\sigma_{\max}^0)^4(1+c^2)}{(1-c^2)^3} \cdot \frac{2\min\{|G_o|, p\}^2}{n},
\end{aligned}$$

for all $n \geq N_2$ by Lemma 3.17. Hence, for all $n \geq N_2$, and by (S5),

$$\begin{aligned}
P_x\bigg(\|K_g - I_{|G|}\|_2^2 \cdot \|\mathcal{Z}_G' Y\|^2 \geq \frac{n^{1-\frac{\rho}{2}}\xi}{4p^3}\bigg) \\
&\leq \Big(\frac{4\delta^{-2}}{n} + \frac{4\delta^{-2}n^{2\rho}}{n}\Big)\frac{(\sigma_{\max}^0)^4(1+c^2)}{(1-c^2)^3} \cdot 2\min\{|G_o|, p\}^2 + \frac{4p^3}{n^{1+\frac{3\rho}{2}}\xi}E_x\big(\|\mathcal{Z}'Y\|^2\big) \\
&\leq \frac{\delta^{-2}16(\sigma_{\max}^0)^4}{(1-c^2)^3} \cdot \frac{\min\{|G_o|, p\}^2}{n^{1-2\rho}} \\
&\quad + \frac{4p^3}{n^{1+\frac{3\rho}{2}}\xi}\bigg[\lambda_{\max}\Big(E_x\big((\mathcal{X}\mathcal{X}')^2\big)\Big)\|\boldsymbol{\alpha}^0\|^2 + (\sigma_{\max}^0)^2 E_x\big(\mathrm{tr}(\mathcal{U}\mathcal{X}'\mathcal{X}\mathcal{U}')\big) \\
&\quad\quad + 2E_x\big(\mathrm{tr}(\mathcal{X}\mathcal{X}'\mathcal{X}\mathcal{U}'(\Sigma^0)^{\frac{1}{2}}A^0)\big)\bigg].
\end{aligned}$$

Thus, by Lemmas 3.16, 3.17, and a slight modification of the proof of Lemma 3.18, for all

$n \geq \max\{N_1, N_2\}$,

$$
\begin{aligned}
P_x\bigg(&\|K_g - I_{|G|}\|_2^2 \cdot \|\mathcal{Z}_G' Y\|^2 \geq \frac{n^{1-\frac{\rho}{2}}\xi}{4p^3}\bigg) \\
&\leq \frac{\delta^{-2} 16 (\sigma^0_{\max})^4}{(1-c^2)^3} \cdot \frac{\min\{|G_o|, p\}^2}{n^{1-2\rho}} \\
&\qquad + \frac{4p^3 n^2 p^2}{n^{1+\frac{3\rho}{2}}\xi}\Bigg[\bigg(\frac{\mathrm{tr}\big(\Gamma_n^2(0)\big)}{p^2} + \frac{4(\sigma^0_{\max})^4 \min\{|G_o|, p\}^2}{(1-c^2)^3 n p^2}\bigg)\|\boldsymbol{\alpha}^0\|^2 \\
&\qquad\qquad + \frac{(\sigma^0_{\max})^4 \min\{|G_o|, p\}}{(1-c^2)np} + \frac{4(\sigma^0_{\max})^4 \min\{|G_o|, p\}^2}{(1-c^2)^2 n p^2}\Bigg] \\
&\leq 16\Bigg[\frac{\delta^{-2}(\sigma^0_{\max})^4 p^2}{n^{1-2\rho}(1-c^2)^3} + \frac{p^3 n^2 p^2 p}{n^{1+\frac{3\rho}{2}}\xi} \cdot \bigg(\frac{\|\Gamma_n(0)\|_2^2}{p} + \frac{(\sigma^0_{\max})^4}{n} \cdot \frac{3+c^4}{(1-c^2)^3}\bigg)\Bigg].
\end{aligned}
$$

This bounds the first right side term in (S6).

The second right side term in (S6) is bounded as follows. From (S4) and (S5),

$$
\begin{aligned}
\frac{4p^3}{n^{1-\frac{\rho}{2}}\xi}&E_x\big(\|\mathcal{Z}_G'Y - E_x(\mathcal{Z}_G'Y)\|^2\big) \\
&\leq \frac{4p^3}{n^{1-\frac{\rho}{2}}\xi}\Big(E_x(Y'\mathcal{Z}\mathcal{Z}'Y) - E_x(Y'\mathcal{Z})E_x(\mathcal{Z}'Y)\Big) \\
&\leq \frac{4p^3}{n^{1-\frac{\rho}{2}}\xi}\Bigg[\lambda_{\max}\Big(E_x\big((\mathcal{X}\mathcal{X}')^2\big)\Big)\|\boldsymbol{\alpha}^0\|^2 - \lambda_{\max}\Big(n^2\Gamma_n^2(0)\Big)\|\boldsymbol{\alpha}^0\|^2 \\
&\qquad + 2E_x\big(\mathrm{tr}(\mathcal{X}\mathcal{X}'\mathcal{X}\mathcal{U}'(\Sigma^0)^{\frac{1}{2}}A^0)\big) + (\sigma^0_{\max})^2 E_x\big(\mathrm{tr}(\mathcal{U}\mathcal{X}'\mathcal{X}\mathcal{U}')\big)\Bigg],
\end{aligned}
$$

and again by Lemmas 3.16, 3.17, and a slight modification of the proof of Lemma 3.18, for all $n \geq \max\{N_1, N_2\}$,

$$
\begin{aligned}
\frac{4p^3}{n^{1-\frac{\rho}{2}}\xi}&E_x\big(\|\mathcal{Z}_G'Y - E_x(\mathcal{Z}_G'Y)\|^2\big) \\
&\leq \frac{4p^3}{n^{1-\frac{\rho}{2}}\xi}\Big(E_x(Y'\mathcal{Z}\mathcal{Z}'Y) - E_x(Y'\mathcal{Z})E_x(\mathcal{Z}'Y)\Big) \\
&\leq \frac{4p^3 n^2 p^2}{n^{1-\frac{\rho}{2}}\xi}\Bigg[\frac{4(\sigma^0_{\max})^4 \min\{|G_o|, p\}^2}{(1-c^2)^3 n p^2}\|\boldsymbol{\alpha}^0\|^2 \\
&\qquad + \frac{4(\sigma^0_{\max})^4 \min\{|G_o|, p\}^2}{(1-c^2)^2 n p^2} + \frac{(\sigma^0_{\max})^4 \min\{|G_o|, p\}}{(1-c^2)np}\Bigg] \\
&\leq \frac{p^3 n^2 p^2}{n^{1-\frac{\rho}{2}}\xi} \cdot \frac{16(\sigma^0_{\max})^4 p}{n} \cdot \frac{3+c^4}{(1-c^2)^3}
\end{aligned}
$$

Therefore, returning back to the original inequality (S6), for all $n \geq \max\{N_1, N_2\}$,

$$P_x\Bigg(\|K_g \mathcal{Z}_G' Y - E_x(\mathcal{Z}_G' Y)\|^2 \geq \frac{n^{1-\frac{\rho}{2}}\xi}{p^3}\Bigg)$$
$$\leq 16(\sigma^0_{\max})^4 \Bigg[\frac{\delta^{-2}p^2}{(1-c^2)^3 n^{1-2\rho}} + \frac{p^6 n^{1-\frac{3\rho}{2}}}{\xi} \cdot \bigg(\frac{\|\Gamma_n(0)\|_2^2}{(\sigma^0_{\max})^4 p} + \frac{(3+c^4)}{(1-c^2)^3 n}\bigg)$$
$$+ \frac{(3+c^4)p^6 n^{\frac{\rho}{2}}}{(1-c^2)^3\xi}\Bigg]$$

Recalling the expression for ξ yields the desired result. ■

Proof of Lemma 3.15. Recalling that $\delta > 0$, an application of Chebyshev's inequality gives,

$$\begin{aligned}
P_x\Big(\big[\lambda_{\min}\big(\Omega - E_x(\Omega)\big)\big]^2 > \delta^2\Big) &\leq P_x\Big(\sum_{i=1}^{p} \lambda_i^2\big(\Omega - E_x(\Omega)\big) > \delta^2\Big) \\
&= P_x\Big(\mathrm{tr}\Big(\big(\Omega - E_x(\Omega)\big)'\big(\Omega - E_x(\Omega)\big)\Big) > \delta^2\Big) \\
&\leq \delta^{-2} E_x\Big[\mathrm{tr}\Big(\big(\Omega - E_x(\Omega)\big)'\big(\Omega - E_x(\Omega)\big)\Big)\Big] \\
&= \delta^{-2}\mathrm{tr}\Big(E_x(\Omega^2) - E_x^2(\Omega)\Big).
\end{aligned}$$

Observe that

$$\begin{aligned}
&E_x(\Omega^2) - E_x^2(\Omega) \\
&\quad = \frac{1}{n^2} E_x \begin{pmatrix} (\mathcal{X}\mathcal{X}')^2 + \mathcal{X}\mathcal{U}'\mathcal{U}\mathcal{X}' - n^2\Gamma_n^2(0) & \mathcal{X}\mathcal{X}'\mathcal{X}\mathcal{U}' + \mathcal{X}\mathcal{U}'\mathcal{U}\mathcal{U}' \\ (\mathcal{X}\mathcal{X}'\mathcal{X}\mathcal{U}')' + (\mathcal{X}\mathcal{U}'\mathcal{U}\mathcal{U}')' & (\mathcal{U}\mathcal{U}')^2 + \mathcal{U}\mathcal{X}'\mathcal{X}\mathcal{U}' - n^2 I_p \end{pmatrix},
\end{aligned}$$

and

$$\Gamma_n^2(0) := \frac{1}{n^2} E_x^2(\mathcal{X}\mathcal{X}') = \frac{1}{n^2} \sum_{t=1}^{n}\sum_{s=1}^{n}\sum_{k=0}^{t-2}\sum_{j=0}^{s-2} (A^0)^k \Sigma (A^0)^{k'} (A^0)^j \Sigma (A^0)^{j'}. \tag{S7}$$

Then since

$$\begin{aligned}
E_x(\mathcal{X}\mathcal{U}'\mathcal{U}\mathcal{U}') &= \sum_{t=1}^{n}\sum_{s=1}^{n} E_x\Big(X^{(t-1)} U^{(t)'} U^{(s)} U^{(s)'}\Big) \\
&= 0 + \sum_{t \neq s}\sum_{k=0}^{t-2} (A^0)^k \Sigma^{\frac{1}{2}} E_x\Big(U^{(t-1-k)} U^{(t)'} U^{(s)} U^{(s)'}\Big) \\
&= \sum_{t \neq s}\sum_{k=0}^{t-2} (A^0)^k \Sigma^{\frac{1}{2}} E_x\Big(E_x\Big(U^{(t-1-k)} U^{(t)'} \big| U^{(s)}\Big) U^{(s)} U^{(s)'}\Big) \\
&= 0,
\end{aligned}$$

and

$$\frac{1}{n^2}E_x\big((\mathcal{U}\mathcal{U}')^2\big) = \frac{1}{n^2}\sum_{t=1}^{n}\sum_{s=1}^{n}E_x\Big(U^{(t)}U^{(t)\prime}U^{(s)}U^{(s)\prime}\Big)$$
$$= \frac{1}{n^2}\sum_{t=1}^{n}E_x\Big(U^{(t)}U^{(t)\prime}U^{(t)}U^{(t)\prime}\Big) + \frac{n^2-n}{n^2}I_p$$
$$= \frac{1}{n^2}\cdot n\cdot(2+p)I_p + \frac{n-1}{n}I_p$$
$$= \frac{p+1}{n}I_p + I_p,$$

it follows that

$$E_x(\Omega^2) - E_x^2(\Omega)$$
$$= \begin{pmatrix} \frac{1}{n^2}E_x\big((\mathcal{X}\mathcal{X}')^2 + \mathcal{X}\mathcal{U}'\mathcal{U}\mathcal{X}'\big) - \Gamma_n^2(0) & \frac{1}{n^2}E_x(\mathcal{X}\mathcal{X}'\mathcal{X}\mathcal{U}') \\ \frac{1}{n^2}E_x(\mathcal{X}\mathcal{X}'\mathcal{X}\mathcal{U}')' & \frac{p+1}{n}I_p + \frac{1}{n^2}E_x(\mathcal{U}\mathcal{X}'\mathcal{X}\mathcal{U}') \end{pmatrix}.$$

Hence,

$$\mathrm{tr}\Big(E_x(\Omega^2) - E_x^2(\Omega)\Big) = \mathrm{tr}\Big(\frac{1}{n^2}E_x\big((\mathcal{X}\mathcal{X}')^2\big) - \Gamma_n^2(0)\Big)$$
$$+ \frac{2}{n^2}\mathrm{tr}\Big(E_x(\mathcal{X}\mathcal{U}'\mathcal{U}\mathcal{X}')\Big) + \frac{p(p+1)}{n}.$$

Applying Lemmas 3.16 and 3.17 gives the desired result. ■

Proof of Lemma 3.16. Begin by expressing

$$E_x(\mathcal{X}\mathcal{U}'\mathcal{U}\mathcal{X}')$$
$$= \sum_{t=1}^{n}\sum_{s=1}^{n}E_x\Big(X^{(t-1)}U^{(t)\prime}U^{(s)}X^{(s-1)\prime}\Big)$$
$$= \sum_{t=1}^{n}\sum_{s=1}^{n}\sum_{k=0}^{t-2}\sum_{j=0}^{s-2}(A^0)^k\Sigma^{\frac{1}{2}}E_x\Big(U^{(t-1-k)}U^{(t)\prime}U^{(s)}U^{(s-1-j)\prime}\Big)\Sigma^{\frac{1}{2}}(A^0)^{j\prime}.$$

Then

$$
\begin{aligned}
&E_x(\mathcal{X}\mathcal{U}'\mathcal{U}\mathcal{X}') \\
&= \sum_{t=1}^{n}\sum_{k=0}^{t-2}\sum_{j=0}^{t-2}(A^0)^k\Sigma^{\frac{1}{2}}E_x\Big(U^{(t)'}U^{(t)}\Big)E_x\Big(U^{(t-1-k)}U^{(t-1-j)'}\Big)\Sigma^{\frac{1}{2}}(A^0)^{j'} \\
&\quad + \sum_{t>s}\sum_{k=0}^{t-2}\sum_{j=0}^{s-2}(A^0)^k\Sigma^{\frac{1}{2}}E_x\Big(U^{(t-1-k)} \\
&\qquad\qquad \times E_x\Big(U^{(t)'}U^{(s)}U^{(s-1-j)'}\big|U^{(t-1-k)}\Big)\Big)\Sigma^{\frac{1}{2}}(A^0)^{j'} \\
&\quad + \sum_{s>t}\sum_{k=0}^{t-2}\sum_{j=0}^{s-2}(A^0)^k\Sigma^{\frac{1}{2}}E_x\Big(E_x\Big(U^{(t-1-k)}U^{(t)'}U^{(s)}\big|U^{(s-1-j)}\Big) \\
&\qquad\qquad \times U^{(s-1-j)'}\Big)\Sigma^{\frac{1}{2}}(A^0)^{j'},
\end{aligned}
$$

which gives

$$
E_x(\mathcal{X}\mathcal{U}'\mathcal{U}\mathcal{X}') = p\sum_{t=1}^{n}\sum_{k=0}^{t-2}(A^0)^k\Sigma(A^0)^{k'},
$$

and

$$
\begin{aligned}
\frac{1}{n^2}\mathrm{tr}\Big(E_x(\mathcal{X}\mathcal{U}'\mathcal{U}\mathcal{X}')\Big) &= \frac{p}{n^2}\sum_{t=1}^{n}\sum_{k=0}^{t-2}\big\|(A^0)^k\Sigma^{\frac{1}{2}}\big\|_F^2 \\
&\le \frac{p}{n^2}\sum_{t=1}^{n}\sum_{k=0}^{t-2}\mathrm{rank}((A^0)^k\Sigma^{\frac{1}{2}})\big\|(A^0)^k\Sigma^{\frac{1}{2}}\big\|_2^2 \\
&\le \frac{p}{n^2}\sum_{t=1}^{n}\sum_{k=0}^{t-2}\mathrm{rank}(A^0)\|A^0\|_2^{2k}\|\Sigma\|_2 \\
&\le \frac{p}{n^2}\sum_{t=1}^{n}\sum_{k=0}^{t-2}\min\{|G_o|,p\}c^{2k}(\sigma^0_{\max})^2.
\end{aligned}
$$

Then using the formula for a geometric series,

$$
\frac{1}{n^2}\mathrm{tr}\Big(E_x(\mathcal{X}\mathcal{U}'\mathcal{U}\mathcal{X}')\Big) \le \frac{(\sigma^0_{\max})^2}{(1-c^2)}\Big(1-\frac{1-c^{2n}}{n(1-c^2)}\Big)\frac{\min\{|G_o|,p\}p}{n},
$$

and so for all $n \ge N_1$,

$$
\frac{1}{n^2}\mathrm{tr}\Big(E_x(\mathcal{X}\mathcal{U}'\mathcal{U}\mathcal{X}')\Big) \le \frac{p(\sigma^0_{\max})^2\min\{|G_o|,p\}}{n(1-c^2)}.
$$

■

Proof of Lemma 3.17.

$$
\begin{aligned}
\frac{1}{n^2}E_x\big((\mathcal{X}\mathcal{X}')^2\big) &= \frac{1}{n^2}\sum_{t=1}^{n}\sum_{s=1}^{n}E_x\Big(X^{(t-1)}X^{(t-1)'}X^{(s-1)}X^{(s-1)'}\Big)\\
&= \frac{1}{n^2}\sum_{t=1}^{n}\sum_{s=1}^{n}T(t,s),
\end{aligned}
$$

where for $t \neq s$,

$$
\begin{aligned}
T(t,s) := \sum_{k=0}^{t-2}\sum_{l=0}^{s-2}&(A^0)^k\Sigma^{\frac{1}{2}}E_x\Big(U^{(t-1-k)}U^{(t-1-k)'}\Sigma^{\frac{1}{2}}(A^0)^{k'}\\
&\qquad\times (A^0)^l\Sigma^{\frac{1}{2}}U^{(s-1-l)}U^{(s-1-l)'}\Big)\Sigma^{\frac{1}{2}}(A^0)^{l'}\\
&+\sum_{k=0}^{t-2}\sum_{l\neq m}^{s-2}0+\sum_{k\neq j}^{t-2}\sum_{l=0}^{s-2}0\\
&+\sum_{k\neq j}^{t-2}\sum_{l\neq m}^{s-2}(A^0)^k\Sigma^{\frac{1}{2}}E_x\Big(U^{(t-1-k)}U^{(t-1-j)'}\Sigma^{\frac{1}{2}}(A^0)^{j'}\\
&\qquad\times (A^0)^l\Sigma^{\frac{1}{2}}U^{(s-1-l)}U^{(s-1-m)'}\Big)\Sigma^{\frac{1}{2}}(A^0)^{m'}.
\end{aligned}
$$

For $t > s$,

$$
\begin{aligned}
T(t,s) = \sum_{l=0}^{s-2}&(A^0)^{t-s+l}\Sigma(A^0)^{l'}(A^0)^{t-s+l}\Sigma(A^0)^{l'}\\
&\qquad+\operatorname{tr}\big((A^0)^l\Sigma(A^0)^{t-s+l'}\big)(A^0)^{t-s+l}\Sigma(A^0)^{l'}\\
&+\sum_{l=0}^{s-2}\sum_{k=0}^{t-2}(A^0)^k\Sigma(A^0)^{k'}(A^0)^l\Sigma(A^0)^{l'}\\
&+\sum_{l\neq m}^{s-2}(A^0)^{t-s+l}\Sigma(A^0)^{l'}(A^0)^{t-s+m}\Sigma(A^0)^{m'}\\
&+\sum_{l\neq m}^{s-2}\operatorname{tr}\big((A^0)^l\Sigma(A^0)^{t-s+l'}\big)(A^0)^{t-s+m}\Sigma(A^0)^{m'}
\end{aligned}
$$

which gives

$$
\begin{aligned}
T(t,s) = \sum_{l=0}^{s-2}\sum_{m=0}^{s-2} &(A^0)^{t-s+l}\Sigma(A^0)^{l'}(A^0)^{t-s+m}\Sigma(A^0)^{m'} \\
&+ \sum_{l=0}^{s-2}\sum_{m=0}^{s-2} \operatorname{tr}\big((A^0)^{l}\Sigma(A^0)^{t-s+l'}\big)(A^0)^{t-s+m}\Sigma(A^0)^{m'} \\
&+ \sum_{l=0}^{s-2}\sum_{k=0}^{t-2}(A^0)^{k}\Sigma(A^0)^{k'}(A^0)^{l}\Sigma(A^0)^{l'}.
\end{aligned}
$$

For $s > t$,

$$
\begin{aligned}
T(t,s) = \sum_{k=0}^{t-2}&(A^0)^{k}\Sigma(A^0)^{s-t+k'}(A^0)^{k}\Sigma(A^0)^{s-t+k'} \\
&\qquad + (A^0)^{k}\Sigma(A^0)^{s-t+k'}\operatorname{tr}\big((A^0)^{s-t+k}\Sigma(A^0)^{k'}\big) \\
&+ \sum_{k=0}^{t-2}\sum_{l=0}^{s-2}(A^0)^{k}\Sigma(A^0)^{k'}(A^0)^{l}\Sigma(A^0)^{l'} \\
&+ \sum_{k\neq j}^{t-2}(A^0)^{k}\Sigma(A^0)^{s-t+k'}(A^0)^{j}\Sigma(A^0)^{s-t+j'} \\
&+ \sum_{k\neq j}^{t-2}(A^0)^{k}\Sigma(A^0)^{s-t+k'}\operatorname{tr}\big((A^0)^{s-t+j}\Sigma(A^0)^{j'}\big)
\end{aligned}
$$

which yields

$$
\begin{aligned}
T(t,s) = \sum_{k=0}^{t-2}\sum_{j=0}^{t-2}&(A^0)^{k}\Sigma(A^0)^{s-t+k'}(A^0)^{j}\Sigma(A^0)^{s-t+j'} \\
&+ \sum_{k=0}^{t-2}\sum_{j=0}^{t-2}(A^0)^{k}\Sigma(A^0)^{s-t+k'}\operatorname{tr}\big((A^0)^{s-t+j}\Sigma(A^0)^{j'}\big) \\
&+ \sum_{k=0}^{t-2}\sum_{l=0}^{s-2}(A^0)^{k}\Sigma(A^0)^{k'}(A^0)^{l}\Sigma(A^0)^{l'}.
\end{aligned}
$$

And for $s = t$,

$$
\begin{aligned}
T(t,t) := \sum_{k=0}^{t-2}\sum_{l=0}^{t-2} & (A^0)^k \Sigma^{\frac{1}{2}} E_x\Big(U^{(t-1-k)}U^{(t-1-k)'}\Sigma^{\frac{1}{2}}(A^0)^{k'}(A^0)^l\Sigma^{\frac{1}{2}} \\
& \qquad \times E_x\Big(U^{(t-1-l)}U^{(t-1-l)'}\big|U^{(t-1-k)}\Big)\Big)\Sigma^{\frac{1}{2}}(A^0)^{l'} \\
& + \sum_{k\neq j}^{t-2}\sum_{l=0}^{t-2} 0 \\
& + \sum_{k\neq j}^{t-2}\sum_{l\neq m}^{t-2}(A^0)^k\Sigma^{\frac{1}{2}}E_x\Big(U^{(t-1-k)}U^{(t-1-j)'}\Sigma^{\frac{1}{2}}(A^0)^{j'} \\
& \qquad \times (A^0)^l\Sigma^{\frac{1}{2}}U^{(t-1-l)}U^{(t-1-m)'}\Big)\Sigma^{\frac{1}{2}}(A^0)^{m'}.
\end{aligned}
$$

Then

$$
\begin{aligned}
T(t,t) = \sum_{k=0}^{t-2} & (A^0)^k\Sigma^{\frac{1}{2}}\Big(2\Sigma^{\frac{1}{2}}(A^0)^{k'}(A^0)^k\Sigma^{\frac{1}{2}} \\
& \qquad + I_p \mathrm{tr}\big((A^0)^k\Sigma(A^0)^{k'}\big)\Big)\Sigma^{\frac{1}{2}}(A^0)^{k'} \\
& + \sum_{k\neq l}^{t-2}(A^0)^k\Sigma(A^0)^{k'}(A^0)^l\Sigma(A^0)^{l'} \\
& + \sum_{k\neq j}^{t-2}\Big((A^0)^k\Sigma(A^0)^{k'}(A^0)^j\Sigma(A^0)^{j'} \\
& \qquad + (A^0)^k\Sigma(A^0)^{k'}\mathrm{tr}\big((A^0)^j\Sigma(A^0)^{j'}\big)\Big),
\end{aligned}
$$

and so

$$
\begin{aligned}
T(t,t) = 2\sum_{k=0}^{t-2}\sum_{j=0}^{t-2} & (A^0)^k\Sigma(A^0)^{k'}(A^0)^j\Sigma(A^0)^{j'} \\
& + \sum_{k=0}^{t-2}\sum_{j=0}^{t-2}(A^0)^k\Sigma(A^0)^{k'}\mathrm{tr}\big((A^0)^j\Sigma(A^0)^{j'}\big).
\end{aligned}
$$

Finally, putting all the pieces together gives,

$$
\begin{aligned}
E_x\big((\mathcal{X}\mathcal{X}')^2\big) &= \sum_{t=1}^{n}T(t,t) + \sum_{t=1}^{n}\sum_{s=1}^{t-1}T(t,s) + \sum_{s=1}^{n}\sum_{t=1}^{s-1}T(t,s) \\
&= \tau_1 + \tau_2 + \sum_{t=1}^{n}\sum_{s=1}^{n}\sum_{k=0}^{t-2}\sum_{j=0}^{s-2}(A^0)^k\Sigma(A^0)^{k'}(A^0)^j\Sigma(A^0)^{j'},
\end{aligned}
$$

where from expression (S7) the last term is equivalent to $n^2\Gamma_n^2(0)$, and

$$\tau_1 := \sum_{t=1}^{n}\sum_{s=1}^{t}\sum_{k=0}^{s-2}\sum_{j=0}^{s-2}\Big[(A^0)^{t-s+k}\Sigma(A^0)^{k'}(A^0)^{t-s+j}\Sigma(A^0)^{j'} + (A^0)^{t-s+j}\Sigma(A^0)^{j'}\operatorname{tr}\big((A^0)^k\Sigma(A^0)^{t-s+k'}\big)\Big]$$

$$\tau_2 := \sum_{s=1}^{n}\sum_{t=1}^{s-1}\sum_{k=0}^{t-2}\sum_{j=0}^{t-2}\Big[(A^0)^{k}\Sigma(A^0)^{s-t+k'}(A^0)^{j}\Sigma(A^0)^{s-t+j'} + (A^0)^{k}\Sigma(A^0)^{s-t+k'}\operatorname{tr}\big((A^0)^{s-t+j}\Sigma(A^0)^{j'}\big)\Big].$$

To proceed, compute the trace of τ_1 and τ_2.

$$\begin{aligned}\operatorname{tr}(\tau_1) &= \sum_{t=1}^{n}\sum_{s=1}^{t}\sum_{k=0}^{s-2}\sum_{j=0}^{s-2}\Big[\operatorname{tr}\Big((A^0)^{t-s+k}\Sigma(A^0)^{k'}(A^0)^{t-s+j}\Sigma(A^0)^{j'}\Big) \\ &\qquad + \operatorname{tr}\Big((A^0)^{t-s+j}\Sigma(A^0)^{j'}\Big)\operatorname{tr}\Big((A^0)^{k}\Sigma(A^0)^{t-s+k'}\Big)\Big] \\ &\le \sum_{t=1}^{n}\sum_{s=1}^{t}\sum_{k=0}^{s-2}\sum_{j=0}^{s-2}\Big[\big\|(A^0)^{k}\Sigma(A^0)^{t-s+k'}\big\|_F\big\|(A^0)^{t-s+j}\Sigma(A^0)^{j'}\big\|_F \\ &\qquad + \big\|\Sigma^{\frac{1}{2}}(A^0)^{t-s+j'}\big\|_F\big\|\Sigma^{\frac{1}{2}}(A^0)^{j'}\big\|_F \\ &\qquad \times \big\|\Sigma^{\frac{1}{2}}(A^0)^{k'}\big\|_F\big\|\Sigma^{\frac{1}{2}}(A^0)^{t-s+k'}\big\|_F\Big],\end{aligned}$$

and so

$$\begin{aligned}\operatorname{tr}(\tau_1) &\le \sum_{t=1}^{n}\sum_{s=1}^{t}\sum_{k=0}^{s-2}\sum_{j=0}^{s-2}\|\Sigma\|_2^2\cdot\|A^0\|_2^{2(t-s+j+k)} \\ &\qquad \times\big(\min\{|G_o|,p\}+\min\{|G_o|,p\}^2\big) \\ &\le (\sigma^0_{\max})^4\sum_{t=1}^{n}\sum_{s=1}^{t}(c^2)^{t-s}\sum_{k=0}^{s-2}(c^2)^k\sum_{j=0}^{s-2}(c^2)^j \\ &\qquad \times\big(\min\{|G_o|,p\}+\min\{|G_o|,p\}^2\big).\end{aligned}$$

The formula for a geometric series further reduces this expression as

$$\begin{aligned}
\operatorname{tr}(\tau_1) &\leq (\sigma^0_{\max})^4 \sum_{t=1}^{n}\sum_{s=1}^{t}(c^2)^{t-s}\Big(\frac{1-(c^2)^{s-1}}{1-c^2}\Big)^2 \\
&\qquad\qquad\qquad \times \big(\min\{|G_o|,p\}+\min\{|G_o|,p\}^2\big) \\
&\leq \frac{\big(\min\{|G_o|,p\}+\min\{|G_o|,p\}^2\big)(\sigma^0_{\max})^4}{(1-c^2)^2}\sum_{t=1}^{n}(c^2)^t\sum_{s=1}^{t}\Big(\frac{1}{c^2}\Big)^s \\
&\leq \frac{\big(\min\{|G_o|,p\}+\min\{|G_o|,p\}^2\big)(\sigma^0_{\max})^4}{(1-c^2)^3}\Big(n-\frac{c^2-(c^2)^{n+1}}{1-c^2}\Big).
\end{aligned}$$

Similarly,

$$\operatorname{tr}(\tau_2) \leq \frac{\big(\min\{|G_o|,p\}+\min\{|G_o|,p\}^2\big)(\sigma^0_{\max})^4}{(1-c^2)^3}\Big(nc^2-\frac{c^2-(c^2)^{n+1}}{1-c^2}\Big),$$

and so

$$\begin{aligned}
&\operatorname{tr}\Big(\frac{1}{n^2}E_x\big((\mathcal{X}\mathcal{X}')^2\big)-\Gamma_n^2(0)\Big) \\
&\qquad\qquad \leq \frac{(\sigma^0_{\max})^4}{(1-c^2)^3}\Big(\frac{\min\{|G_o|,p\}+\min\{|G_o|,p\}^2}{n}\Big) \\
&\qquad\qquad\qquad \times\Bigg[\Big(1-\frac{c^2-(c^2)^{n+1}}{n(1-c^2)}\Big)+\Big(c^2-\frac{c^2-(c^2)^{n+1}}{n(1-c^2)}\Big)\Bigg].
\end{aligned}$$

Then for all $n \geq N_2$,

$$\begin{aligned}
&\operatorname{tr}\Big(\frac{1}{n^2}E_x\big((\mathcal{X}\mathcal{X}')^2\big)-\Gamma_n^2(0)\Big) \\
&\qquad\qquad \leq \frac{(\sigma^0_{\max})^4(1+c^2)}{(1-c^2)^3}\Big(\frac{\min\{|G_o|,p\}+\min\{|G_o|,p\}^2}{n}\Big).
\end{aligned}$$

■

Proof of Lemma 3.18. First,

$$\begin{aligned}
&E_x(\mathcal{X}\mathcal{X}'\mathcal{X}\mathcal{U}') \\
&= \sum_{t=1}^{n}\sum_{s=1}^{n}E_x\Big(X^{(t-1)}X^{(t-1)'}X^{(s-1)}U^{(s)'}\Big) \\
&= 0+\sum_{t\neq s}E_x\Big(X^{(t-1)}X^{(t-1)'}X^{(s-1)}U^{(s)'}\Big) \\
&= 0+\sum_{t>s}\sum_{k=0}^{t-2}\sum_{j=0}^{t-2}\sum_{l=0}^{s-2}(A^0)^k\Sigma^{\frac{1}{2}}E_x\Big(U^{(t-1-k)}U^{(t-1-j)'}\Sigma^{\frac{1}{2}}(A^0)^{j'} \\
&\qquad\qquad\qquad \times (A^0)^l\Sigma^{\frac{1}{2}}U^{(s-1-l)}U^{(s)'}\Big).
\end{aligned}$$

Next,

$$
\begin{aligned}
&E_x(\mathcal{X}\mathcal{X}'\mathcal{X}\mathcal{U}') \\
&\quad = \sum_{t>s}\sum_{k=0}^{t-2}\sum_{l=0}^{s-2}(A^0)^k\Sigma^{\frac{1}{2}}E_x\Big(U^{(t-1-k)}U^{(t-1-k)'}\Sigma^{\frac{1}{2}}(A^0)^{k'}(A^0)^l\Sigma^{\frac{1}{2}} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \times E_x\Big(U^{(s-1-l)}U^{(s)'}\big|U^{(t-1-k)}\Big)\Big) \\
&\qquad + \sum_{t>s}\sum_{k\neq j}\sum_{l=0}^{s-2}(A^0)^k\Sigma^{\frac{1}{2}}E_x\Big(U^{(t-1-k)} \\
&\qquad\qquad \times E_x\Big(U^{(t-1-j)'}\Sigma^{\frac{1}{2}}(A^0)^{j'}(A^0)^l\Sigma^{\frac{1}{2}}U^{(s-1-l)}U^{(s)'}\big|U^{(t-1-k)}\Big)\Big),
\end{aligned}
$$

and since the first sum is zero, breaking up the second sum gives

$$
\begin{aligned}
&E_x(\mathcal{X}\mathcal{X}'\mathcal{X}\mathcal{U}') \\
&= 0 + \sum_{t=1}^{n}\sum_{k\neq j}\sum_{l=0}^{t-1-k-2}(A^0)^k\Sigma^{\frac{1}{2}}E_x\Big(U^{(t-1-k)} \\
&\quad \times E_x\Big(U^{(t-1-j)'}\Sigma^{\frac{1}{2}}(A^0)^{j'}(A^0)^l\Sigma^{\frac{1}{2}}U^{(t-1-k-l-1)}\big|U^{(t-1-k)}\Big)U^{(t-1-k)'}\Big) \\
&\quad + \sum_{t=1}^{n}\sum_{k\neq j}\sum_{l=0}^{t-1-j-2}(A^0)^k\Sigma^{\frac{1}{2}}E_x\Big(U^{(t-1-k)} \\
&\quad \times E_x\Big(U^{(t-1-j)'}\Sigma^{\frac{1}{2}}(A^0)^{j'}(A^0)^l\Sigma^{\frac{1}{2}}U^{(t-1-j-l-1)}U^{(t-1-j)'}\big|U^{(t-1-k)}\Big)\Big).
\end{aligned}
$$

Further,

$$
\begin{aligned}
&E_x(\mathcal{X}\mathcal{X}'\mathcal{X}\mathcal{U}') \\
&\quad = \sum_{t=1}^{n}\sum_{j>k}(A^0)^k\Sigma^{\frac{1}{2}}E_x\Big(U^{(t-1-k)} \\
&\qquad\qquad\qquad \times E_x\Big(U^{(t-1-j)'}\Sigma^{\frac{1}{2}}(A^0)^{j'}(A^0)^{j-k-1}\Sigma^{\frac{1}{2}}U^{(t-1-j)}\Big)U^{(t-1-k)'}\Big) \\
&\quad + \sum_{t=1}^{n}\sum_{k>j}(A^0)^k\Sigma^{\frac{1}{2}}E_x\Big(U^{(t-1-k)} \\
&\quad \times E_x\Big(U^{(t-1-j)'}\Sigma^{\frac{1}{2}}(A^0)^{j'}(A^0)^{k-j-1}\Sigma^{\frac{1}{2}}U^{(t-1-k)}U^{(t-1-j)'}\big|U^{(t-1-k)}\Big)\Big),
\end{aligned}
$$

and so

$$
\begin{aligned}
&E_x(\mathcal{X}\mathcal{X}'\mathcal{X}\mathcal{U}') \\
&\quad= \sum_{t=1}^{n}\sum_{j>k}(A^0)^k\Sigma^{\frac{1}{2}}I_p\cdot E_x\Big(U^{(t-1-j)'}\Sigma^{\frac{1}{2}}(A^0)^{j'}(A^0)^{j-k-1}\Sigma^{\frac{1}{2}}U^{(t-1-j)}\Big) \\
&\qquad+\sum_{t=1}^{n}\sum_{k>j}(A^0)^k\Sigma^{\frac{1}{2}}E_x\Big(U^{(t-1-k)}U^{(t-1-k)'}\Sigma^{\frac{1}{2}}(A^0)^{k-j-1'}(A^0)^j\Sigma^{\frac{1}{2}}\Big) \\
&\quad= \sum_{t=1}^{n}\sum_{j>k}(A^0)^k\Sigma^{\frac{1}{2}}\operatorname{tr}\Big(\Sigma^{\frac{1}{2}}(A^0)^{j'}(A^0)^{j-k-1}\Sigma^{\frac{1}{2}}\Big) \\
&\qquad\qquad\qquad+\sum_{t=1}^{n}\sum_{k>j}(A^0)^k\Sigma(A^0)^{k-j-1'}(A^0)^j\Sigma^{\frac{1}{2}}.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&\operatorname{tr}\Big(E_x(\mathcal{X}\mathcal{X}'\mathcal{X}\mathcal{U}'A^0)\Big) \\
&\quad= \sum_{t=1}^{n}\sum_{k=0}^{t-2}\sum_{j=k+1}^{t-2}\operatorname{tr}\Big((A^0)^k\Sigma^{\frac{1}{2}}A^0\Big)\operatorname{tr}\Big(\Sigma(A^0)^{j'}(A^0)^{j-k-1}\Big) \\
&\qquad+\sum_{t=1}^{n}\sum_{j=0}^{t-2}\sum_{k=j+1}^{t-2}\operatorname{tr}\Big((A^0)^k\Sigma(A^0)^{k-j-1'}(A^0)^j\Sigma^{\frac{1}{2}}A^0\Big) \\
&\quad\le \min\{|G_o|,p\}^2(\sigma^0_{\max})^3\sum_{t=1}^{n}\sum_{k=0}^{t-2}\sum_{j=k+1}^{t-2}\left\|A^0\right\|_2^{2j} \\
&\qquad+\min\{|G_o|,p\}(\sigma^0_{\max})^3\sum_{t=1}^{n}\sum_{j=0}^{t-2}\sum_{k=j+1}^{t-2}\left\|A^0\right\|_2^{2k} \\
&\quad\le \Big(\min\{|G_o|,p\}^2+\min\{|G_o|,p\}\Big)(\sigma^0_{\max})^3\sum_{t=1}^{n}\sum_{k=0}^{t-2}\sum_{j=k+1}^{t-2}c^{2j},
\end{aligned}
$$

and applying the formula for a geometric series yields

$$
\begin{aligned}
&\operatorname{tr}\Big(E_x(\mathcal{X}\mathcal{X}'\mathcal{X}\mathcal{U}'A^0)\Big) \\
&\quad\le \Big(\min\{|G_o|,p\}^2+\min\{|G_o|,p\}\Big)(\sigma^0_{\max})^3\sum_{t=1}^{n}\sum_{k=0}^{t-2}\frac{c^{2(k+1)}-c^{2(t-1)}}{1-c^2} \\
&\quad\le \Big(\min\{|G_o|,p\}^2+\min\{|G_o|,p\}\Big)\frac{(\sigma^0_{\max})^3}{1-c^2}\sum_{t=1}^{n}c^2\frac{1-c^{2(t-1)}}{1-c^2} \\
&\quad\le \Big(\min\{|G_o|,p\}^2+\min\{|G_o|,p\}\Big)\frac{n(\sigma^0_{\max})^3c^2}{(1-c^2)^2}.
\end{aligned}
$$

■

Proof of Lemma 3.19. Define $\widehat{\Sigma} := \frac{1}{n}\mathcal{X}\mathcal{X}'$. By Condition 3.1, $\lambda_{\min}\big(\Gamma_n(0)\big) > \delta$ which implies that

$$\begin{aligned}
P_x\Big(\lambda_{\min}(\widehat{\Sigma}) \geq \delta/2\Big) &\geq P_x\Big(\lambda_{\min}\big(\Gamma_n(0)\big) + \lambda_{\min}\big(\widehat{\Sigma} - \Gamma_n(0)\big) \geq \delta - \delta/2\Big) \\
&\geq P_x\Big(\lambda_{\min}\big(\Gamma_n(0)\big) + \lambda_{\min}\big(\widehat{\Sigma} - \Gamma_n(0)\big) \geq \lambda_{\min}\big(\Gamma_n(0)\big) - \delta/2\Big) \\
&= P_x\Big(\lambda_{\min}\big(\widehat{\Sigma} - \Gamma_n(0)\big) \geq -\delta/2\Big) \\
&\geq P_x\Big(|\lambda_{\min}\big(\widehat{\Sigma} - \Gamma_n(0)\big)| \leq \delta/2\Big).
\end{aligned}$$

Further,

$$\begin{aligned}
P_x\Big(|\lambda_{\min}\big(\widehat{\Sigma} - \Gamma_n(0)\big)| > \delta/2\Big) &\leq P_x\Bigg(\sum_{i=1}^{p} \lambda_i^2\big(\widehat{\Sigma} - \Gamma_n(0)\big) > \delta^2/4\Bigg) \\
&= P_x\Big(\|\widehat{\Sigma} - \Gamma_n(0)\|_F^2 > \delta^2/4\Big) \\
&\leq 4\delta^{-2} E_x\Big(\|\widehat{\Sigma} - \Gamma_n(0)\|_F^2\Big) \\
&= 4\delta^{-2}\mathrm{tr}\Big(E_x\big(\widehat{\Sigma}^2\big) - \Gamma_n^2(0)\Big) \\
&\leq 4\delta^{-2}\frac{(\sigma_{\max}^0)^4(1+c^2)}{(1-c^2)^3} \cdot \frac{2\min\{|G_o|, p\}^2}{n}
\end{aligned} \tag{S8}$$

where the last inequality holds for all $n \geq N_2$ by Lemma 3.17. Therefore, for all $n \geq N_2$,

$$P_x\Big(\lambda_{\min}(\widehat{\Sigma}) \geq \delta/2\Big) \geq 1 - 4\delta^{-2}\frac{(\sigma_{\max}^0)^4(1+c^2)}{(1-c^2)^3} \cdot \frac{2\min\{|G_o|, p\}^2}{n}.$$

■

Proof of Lemma S2.1. By the Poincaré separation theorem (see Rao (1979)),

$$
\begin{aligned}
P\left(\frac{\left\|(\Gamma_n(0)\otimes(\Sigma^0)^{-1})_{G_o,G_o}\right\|_2^2}{\lambda_{\min}^2(\mathcal{Z}_{G_o}'\mathcal{W}^{-1}\mathcal{Z}_{G_o})}\geq\frac{1}{n^{2-\frac{\varrho}{2}}p^5}\right)
&\leq P\left(\frac{\left\|\Gamma_n(0)\otimes(\Sigma^0)^{-1}\right\|_2^2}{\lambda_{\min}^2(\mathcal{Z}'\mathcal{W}^{-1}\mathcal{Z})}\geq\frac{n^{\frac{\varrho}{2}}}{n^2p^5}\right)\\
&= P\left(\frac{\lambda_{\min}(\mathcal{Z}'\mathcal{W}^{-1}\mathcal{Z})}{\|\Gamma_n(0)\|_2(\sigma_{\min}^0)^{-2}}\leq\frac{np^{\frac{5}{2}}}{n^{\frac{\varrho}{4}}}\right)\\
&\leq P\left(\frac{\min\{\sigma_j^{-2}\}\lambda_{\min}(\mathcal{Z}'\mathcal{Z})}{\|\Gamma_n(0)\|_2(\sigma_{\min}^0)^{-2}}\leq\frac{np^{\frac{5}{2}}}{n^{\frac{\varrho}{4}}}\right)\\
&\leq \sum_{j=1}^p P\left(\sigma_j^{-2}\leq\frac{np^{\frac{5}{2}}}{n^{\frac{\varrho}{4}}}\frac{\|\Gamma_n(0)\|_2(\sigma_{\min}^0)^{-2}}{\lambda_{\min}(\mathcal{Z}'\mathcal{Z})}\right)\\
&= \sum_{j=1}^p P\left(\sigma_j^{2}\geq\frac{n^{\frac{\varrho}{4}}}{np^{\frac{5}{2}}}\frac{\lambda_{\min}(\mathcal{Z}'\mathcal{Z})}{\|\Gamma_n(0)\|_2(\sigma_{\min}^0)^{-2}}\right)\\
&\leq \frac{np^{\frac{5}{2}}}{n^{\frac{\varrho}{4}}}\frac{\|\Gamma_n(0)\|_2(\sigma_{\min}^0)^{-2}}{\lambda_{\min}(\mathcal{Z}'\mathcal{Z})}\sum_{j=1}^p\frac{m_j^{g_o}}{n-|r_j^{g_o}|-2}
\end{aligned}
$$

by the Markov inequality, recalling that $\sigma_j^2\sim\text{inv-gamma}\big((n-|r_j^{g_o}|)/2, m_j^{g_o}/2\big)$.

Next observe that

$$\sum_{j=1}^p\frac{m_j^{g_o}}{n-|r_j^{g_o}|-2}\leq\frac{1}{n-|G_o|-2}\sum_{j=1}^p m_j^{g_o},$$

and as in (S20),

$$P\left(\sum_{j=1}^p m_j^{g_o}>(\sigma_{\max}^0)^2 3np\right)\leq P\left(\sum_{t=1}^n\|U^{(t)}\|^2>3np\right)\leq P\big(\chi_{np}^2>3np\big)\leq e^{-\frac{np}{4}}.$$

Additionally, since $\lambda_{\min}(\mathcal{Z}'\mathcal{Z})=\lambda_{\min}(\mathcal{X}\mathcal{X}')\cdot 1$, it follows from Lemma 3.19 (assuming that Condition 3.1 holds) that for all $n\geq N_2$,

$$P_x\big(\lambda_{\min}(\mathcal{Z}'\mathcal{Z})\geq n\delta/2\big)\geq 1-4\delta^{-2}\frac{(\sigma_{\max}^0)^4(1+c^2)}{(1-c^2)^3}\cdot\frac{2\min\{|G_o|,p\}^2}{n}.$$

Therefore, for all $n\geq N_2$,

$$
\begin{aligned}
P\left(\frac{\left\|(\Gamma_n(0)\otimes(\Sigma^0)^{-1})_{G_o,G_o}\right\|_2^2}{\lambda_{\min}^2(\mathcal{Z}_{G_o}'\mathcal{W}^{-1}\mathcal{Z}_{G_o})}\geq\frac{1}{n^{2-\frac{\varrho}{2}}p^5}\right)
&\leq\frac{np^{\frac{5}{2}}}{n^{\frac{\varrho}{4}}}\frac{2\|\Gamma_n(0)\|_2(\sigma_{\min}^0)^{-2}}{n\delta}\frac{(\sigma_{\max}^0)^2 3np}{(n-|G_o|-2)}\\
&\leq\frac{p^{\frac{7}{2}}}{n^{\frac{\varrho}{4}}}\cdot\frac{(\sigma_{\max}^0)^2}{(\sigma_{\min}^0)^2}\cdot\frac{6\|\Gamma_n(0)\|_2/\delta}{(1-|G_o|/n-2/n)},
\end{aligned}
$$

with probability exceeding

$$1 - 4\delta^{-2}\frac{(\sigma^0_{\max})^4(1+c^2)}{(1-c^2)^3}\cdot\frac{2\min\{|G_o|,p\}^2}{n} - e^{-\frac{np}{4}}.$$

■

Proof of Lemma S2.2. Consider the following upper bound on the natural logarithm of the determinant of the positive definite matrix, $\widetilde{\mathcal{D}}_g'\widetilde{\mathcal{D}}_g$,

$$\log(|\widetilde{\mathcal{D}}_g'\widetilde{\mathcal{D}}_g|) \le \operatorname{tr}(\widetilde{\mathcal{D}}_g'\widetilde{\mathcal{D}}_g - I_{|G|+p}) = \operatorname{tr}(\widetilde{\mathcal{D}}_g'\widetilde{\mathcal{D}}_g) - |G| - p. \tag{S9}$$

The trace can be bounded as follows.

$$\begin{aligned}\operatorname{tr}(\widetilde{\mathcal{D}}_g'\widetilde{\mathcal{D}}_g) &= \frac{1}{\sqrt{n}}\sum_{j=1}^{p}\sum_{i\in r_j^g}\Big(\frac{\partial Y}{\partial A_{ij}}\Big)'\frac{\partial Y}{\partial A_{ij}} + \frac{1}{\sqrt{n}}\sum_{i=1}^{p}\Big(\frac{\partial Y}{\partial \sigma_i}\Big)'\frac{\partial Y}{\partial \sigma_i}\\ &= \frac{1}{\sqrt{n}}\sum_{j=1}^{p}\sum_{i\in r_j^g}\operatorname{vec}(\mathcal{X})'\big(I_n\otimes J^{ji}\big)\Theta_g'\Theta_g\big(I_n\otimes J^{ij}\big)\operatorname{vec}(\mathcal{X}) \qquad \text{(S10)}\\ &\quad + \frac{1}{\sqrt{n}}\sum_{i=1}^{p}\big(Y-\mathcal{Z}_G\boldsymbol{\alpha}_G\big)'\big(I_n\otimes J^{ii}\big)\Theta_g'\Theta_g\big(I_n\otimes J^{ii}\big)\big(Y-\mathcal{Z}_G\boldsymbol{\alpha}_G\big)\end{aligned}$$

where r_j^g is the set of active row indices of A_g for column $j\in\{1,\dots,p\}$, and Θ_g is as in (S1) with A_g in place of A.

Denote $\Upsilon_n := \operatorname{Cov}(Y,Y)$, and note that $\widetilde{\Upsilon}_n := \Theta_g\Theta_g'$ is the covariance matrix of a VAR(1) process with an identity contemporaneous error covariance matrix, and coefficient matrix A_g. Since Θ_g is a square matrix it must be the case that

$$\max_{v\in\mathbb{R}^{np}}\Big\{\frac{v'\Theta_g'\Theta_g v}{v'v}\Big\} = \max_{v\in\mathbb{R}^{np}}\Big\{\frac{v'\Theta_g\Theta_g' v}{v'v}\Big\} = \max_{v\in\mathbb{R}^{np}}\Big\{\frac{v'\widetilde{\Upsilon}_n v}{v'v}\Big\},$$

and so from (S10) it follows that

$$\begin{aligned}\operatorname{tr}(\widetilde{\mathcal{D}}_g'\widetilde{\mathcal{D}}_g) &\le \frac{1}{\sqrt{n}}\max_{v\in\mathbb{R}^{np}}\Big\{\frac{v'\widetilde{\Upsilon}_n v}{v'v}\Big\}\Bigg\{\sum_{j=1}^{p}\sum_{i\in r_j^g}\operatorname{vec}(\mathcal{X})'\big(I_n\otimes J^{ji}\big)\big(I_n\otimes J^{ij}\big)\operatorname{vec}(\mathcal{X})\\ &\qquad + \sum_{i=1}^{p}\big(Y-\mathcal{Z}_G\boldsymbol{\alpha}_G\big)'\big(I_n\otimes J^{ii}\big)\big(I_n\otimes J^{ii}\big)\big(Y-\mathcal{Z}_G\boldsymbol{\alpha}_G\big)\Bigg\}\\ &= \frac{\lambda_{\max}(\widetilde{\Upsilon}_n)}{\sqrt{n}}\Bigg\{\operatorname{vec}(\mathcal{X})'\sum_{j=1}^{p}|r_j^g|\big(I_n\otimes J^{jj}\big)\operatorname{vec}(\mathcal{X}) + \big\|Y-\mathcal{Z}_G\boldsymbol{\alpha}_G\big\|^2\Bigg\}\\ &= \frac{\lambda_{\max}(\widetilde{\Upsilon}_n)}{\sqrt{n}}\Bigg\{\sum_{j=1}^{p}|r_j^g|\sum_{t=0}^{n-1}(X_j^{(t)})^2 + \big\|Y-\mathcal{Z}_G\boldsymbol{\alpha}_G\big\|^2\Bigg\},\end{aligned}$$

where $|r_j|$ is the number of active rows of A_g for each column $j \in \{1, \ldots, p\}$.

By Proposition 2.3 in Basu et al. (2015), and by Lemma S2.3,

$$\lambda_{\max}(\widetilde{\Upsilon}_n) \leq 2\pi \operatorname*{ess\,sup}_{\theta \in [-\pi, \pi]} \|f_{\widetilde{X}}(\theta)\|_2 \leq (1 - \|A_g\|_2)^{-2},$$

recalling that $\widetilde{\Upsilon}_n$ is the covariance matrix for the VAR(1) process with an identity contemporaneous error covariance matrix, and coefficient matrix A_g. The function $f_{\widetilde{X}}$ denotes the spectral density of this process. Thus,

$$\operatorname{tr}(\widetilde{\mathcal{D}}_g' \widetilde{\mathcal{D}}_g) \leq \frac{(1 - \|A_g\|_2)^{-2}}{\sqrt{n}} \left\{ \sum_{j=1}^{p} |r_j^g| \sum_{t=0}^{n-1} (X_j^{(t)})^2 + \left\| Y - \mathcal{Z}_G \boldsymbol{\alpha}_G \right\|^2 \right\}.$$

Then

$$\begin{aligned}
\left\| Y - \mathcal{Z}_G \boldsymbol{\alpha}_G \right\| &= \left\| Y - \big(I_n \otimes A_g\big) \operatorname{vec}(\mathcal{X}) \right\| \\
&\leq \|Y\| + \left\| \big(I_n \otimes A_g\big) \operatorname{vec}(\mathcal{X}) \right\| \\
&= \|Y\| + \sqrt{\sum_{t=0}^{n-1} X^{(t)\prime} A_g' A_g X^{(t)}} \\
&\leq \|Y\| + \sqrt{\max_{v \neq 0} \left\{ \frac{v' A_g' A_g v}{v' v} \right\} \sum_{t=0}^{n-1} X^{(t)\prime} X^{(t)}} \\
&= \|Y\| + \|A_g\|_2 \|\operatorname{vec}(\mathcal{X})\| \\
&\leq \|Y\| (1 + \|A_g\|_2).
\end{aligned}$$

Hence,

$$\begin{aligned}
\operatorname{tr}(\widetilde{\mathcal{D}}_g' \widetilde{\mathcal{D}}_g) &\leq \frac{(1 - \|A_g\|_2)^{-2}}{\sqrt{n}} \left\{ \sum_{j=1}^{p} |r_j^g| \sum_{t=1}^{n-1} (X_j^{(t)})^2 + \|Y\|^2 (1 + \|A_g\|_2)^2 \right\} \\
&\leq \frac{(1 - c)^{-2}}{\sqrt{n}} \left\{ \sum_{j=1}^{p} |r_j^g| \sum_{t=1}^{n-1} (X_j^{(t)})^2 + \|Y\|^2 (1 + c)^2 \right\} \\
&\leq (1 - c)^{-2} \Big(r_{\max}^g + (1 + c)^2 \Big) \frac{\|Y\|^2}{\sqrt{n}}.
\end{aligned}$$

■

Proof of Lemma S2.3. For a stable VAR(1) process, i.e. $\|A\|_2 < 1$, the spectral density has the following form (Basu et al. 2015),

$$f_X(\theta) = \frac{1}{2\pi} \big(I_p - A e^{-i\theta}\big)^{-1} \Sigma \big[\big(I_p - A e^{-i\theta}\big)^{-1}\big]^*.$$

Then

$$\begin{aligned}
\|f_X(\theta)\|_2 &= \frac{1}{2\pi}\big\|\big(I_p - Ae^{-i\theta}\big)^{-1}\Sigma\big[\big(I_p - Ae^{-i\theta}\big)^{-1}\big]^*\big\|_2 \\
&\leq \frac{1}{2\pi}\big\|\big(I_p - Ae^{-i\theta}\big)^{-1}\big\|_2\|\Sigma\|_2\big\|\big[\big(I_p - Ae^{-i\theta}\big)^{-1}\big]^*\big\|_2 \\
&= \frac{1}{2\pi}\big\|\big(I_p - Ae^{-i\theta}\big)^{-1}\big\|_2^2 \max_{1\leq i\leq p}\sigma_i^2,
\end{aligned}$$

for $\Sigma = \mathrm{diag}\{\sigma_1^2, \ldots, \sigma_p^2\}$.

Note that, for any matrix B satisfying $\|B\|_2 < 1$,

$$(I_p - B)\sum_{k=0}^{\infty} B^k = I_p - \lim_{r\to\infty} B^{r+1} = I_p,$$

and so

$$\big\|(I_p - B)^{-1}\big\|_2 = \Big\|\sum_{k=0}^{\infty} B^k\Big\|_2 \leq \sum_{k=0}^{\infty}\|B^k\|_2 \leq \sum_{k=0}^{\infty}\|B\|_2^k = \frac{1}{1-\|B\|_2},$$

by the geometric series for real numbers. Therefore, taking $B = Ae^{-i\theta}$,

$$\|f_X(\theta)\|_2 \leq \frac{1}{2\pi}\big(1 - \|Ae^{-i\theta}\|_2\big)^{-2}\sigma_{\max}^2 = \frac{1}{2\pi}\big(1 - \|A\|_2\big)^{-2}\sigma_{\max}^2.$$

■

Proof of Lemma S2.4. Consider the inequality for positive definite matrices,

$$\log\big(|\widetilde{\mathcal{D}}_{g_o}'\widetilde{\mathcal{D}}_{g_o}|\big) \geq \mathrm{tr}\Big(I_{|G_o|+p} - (\widetilde{\mathcal{D}}_{g_o}'\widetilde{\mathcal{D}}_{g_o})^{-1}\Big),$$

which is easily proven with the fact that $\log(x) \leq x - 1$ for all $x > 0$ implies $\log(y) \geq 1 - \frac{1}{y}$ for all $y > 0$. Thus,

$$\begin{aligned}
\log\big(|\widetilde{\mathcal{D}}_{g_o}'\widetilde{\mathcal{D}}_{g_o}|\big) &\geq |G_o| + p - \mathrm{tr}\Big((\widetilde{\mathcal{D}}_{g_o}'\widetilde{\mathcal{D}}_{g_o})^{-1}\Big) \\
&\geq |G_o| + p - (|G_o| + p)\lambda_{\max}\Big((\widetilde{\mathcal{D}}_{g_o}'\widetilde{\mathcal{D}}_{g_o})^{-1}\Big) \\
&= (|G_o| + p)\Big[1 - \lambda_{\min}^{-1}(\widetilde{\mathcal{D}}_{g_o}'\widetilde{\mathcal{D}}_{g_o})\Big].
\end{aligned} \tag{S11}$$

A lower bound on $\lambda_{\min}(\widetilde{\mathcal{D}}_{g_o}'\widetilde{\mathcal{D}}_{g_o})$ can be derived as follows. Decompose

$$\widetilde{\mathcal{D}}_{g_o}'\widetilde{\mathcal{D}}_{g_o} = \frac{1}{\sqrt{n}}\mathcal{B}_{G_o}'\Theta_{g_o}'\Theta_{g_o}\mathcal{B}_{G_o}$$

where Θ_{g_o} is as in (S1) with A_{g_o} in place of A,

$$\mathcal{B}_{G_o} := \Big((I_n \otimes J^{11})\mathrm{vec}(\mathcal{X}) \quad (I_n \otimes J^{12})\mathrm{vec}(\mathcal{X}) \quad \cdots \quad (I_n \otimes J^{pp})\mathrm{vec}(\mathcal{U}) \Big)$$

$$= \begin{pmatrix} X^{(0)'}_{c_1^{g_o}} & & & U_1^{(1)} & & \\ & \ddots & & & \ddots & \\ & & X^{(0)'}_{c_p^{g_o}} & & & U_p^{(1)} \\ & \vdots & & & \vdots & \\ X^{(n-1)'}_{c_1^{g_o}} & & & U_1^{(n)} & & \\ & \ddots & & & \ddots & \\ & & X^{(n-1)'}_{c_p^{g_o}} & & & U_p^{(n)} \end{pmatrix}, \tag{S12}$$

and $c_i^{g_o}$ is the set of active column indices of A_{g_o} for row $i \in \{1, \dots, p\}$. Then

$$\begin{aligned} \lambda_{\min}(\widetilde{\mathcal{D}}'_{g_o}\widetilde{\mathcal{D}}_{g_o}) &= \frac{1}{\sqrt{n}} \min_{v \neq 0} \left\{ \frac{v'\mathcal{B}'_{G_o}\Theta'_{g_o}\Theta_{g_o}\mathcal{B}_{G_o}v}{v'\mathcal{B}'_{G_o}\mathcal{B}_{G_o}v} \cdot \frac{v'\mathcal{B}'_{G_o}\mathcal{B}_{G_o}v}{v'v} \right\} \\ &\geq \frac{1}{\sqrt{n}} \min_{v \neq 0} \left\{ \frac{v'\mathcal{B}'_{G_o}\Theta'_{g_o}\Theta_{g_o}\mathcal{B}_{G_o}v}{v'\mathcal{B}'_{G_o}\mathcal{B}_{G_o}v} \right\} \min_{v \neq 0} \left\{ \frac{v'\mathcal{B}'_{G_o}\mathcal{B}_{G_o}v}{v'v} \right\} \\ &\geq \frac{1}{\sqrt{n}} \lambda_{\min}(\Theta'_{g_o}\Theta_{g_o})\lambda_{\min}(\mathcal{B}'_{G_o}\mathcal{B}_{G_o}) \\ &\geq \frac{1}{\sqrt{n}} \lambda_{\min}(\Theta'_{g_o}\Theta_{g_o})\lambda_{\min}(\mathcal{B}'\mathcal{B}), \end{aligned}$$

where $\mathcal{B}$ is of the form (S12) corresponding to the graph with all components of A active, and the last inequality follows from the Poincaré separation theorem (see Rao (1979)) because $\mathcal{B}'_{G_o}\mathcal{B}_{G_o}$ is a principal submatrix of $\mathcal{B}'\mathcal{B}$.

The proof finishes by deriving lower bounds for both $\lambda_{\min}(\Theta'_{g_o}\Theta_{g_o})$ and $\lambda_{\min}(\mathcal{B}'\mathcal{B})$. First consider $\lambda_{\min}(\Theta'_{g_o}\Theta_{g_o})$, and observe that

$$\lambda_{\min}(\Theta'_{g_o}\Theta_{g_o}) = \frac{1}{\lambda_{\max}\big((\Theta'_{g_o}\Theta_{g_o})^{-1}\big)}, \tag{S13}$$

where

$$\Theta_{g_o}^{-1} = \begin{pmatrix} I_p & & & & \\ -A_{g_o} & I_p & & & \\ & -A_{g_o} & I_p & & \\ & & \ddots & \ddots & \\ & & & -A_{g_o} & I_p \end{pmatrix}.$$

Then for any $v := (v_1', \ldots, v_n')' \in \mathbb{R}^{np}$,

$$
\begin{aligned}
\frac{v'\left(\Theta_{g_o}'\Theta_{g_o}\right)^{-1}v}{v'v} &= \frac{\left\|(\Theta_{g_o}')^{-1}v\right\|^2}{v'v} \\
&= \frac{\|v_n\|^2 + \sum_{i=1}^{n-1}\|v_i - A_{g_o}'v_{i+1}\|^2}{v'v} \\
&\leq \frac{\|v_n\|^2 + \sum_{i=1}^{n-1}\left(\|v_i\| + \|A_{g_o}'v_{i+1}\|\right)^2}{v'v} \\
&\leq \frac{2\|v_n\|^2 + \sum_{i=1}^{n-1} 2\|v_i\|^2 + 2\|A_{g_o}'v_{i+1}\|^2}{v'v} \\
&\leq \frac{2v'v + 2\|A_{g_o}\|_2^2 v'v}{v'v},
\end{aligned}
$$

which gives

$$
\lambda_{\max}\left((\Theta_{g_o}'\Theta_{g_o})^{-1}\right) \leq 2(1 + \|A_{g_o}\|_2^2) \leq 2(1 + c^2),
$$

and thus from (S13),

$$
\lambda_{\min}(\Theta_{g_o}'\Theta_{g_o}) \geq \frac{1}{2}(1 + c^2)^{-1}. \tag{S14}
$$

It remains to derive a lower bound for $\frac{1}{\sqrt{n}}\lambda_{\min}(\mathcal{B}'\mathcal{B})$. For any $v \in \mathbb{R}^{p^2+p}$, expressing $v = (\vec{v}_1, \ldots, \vec{v}_p, v_{p^2+1}, \ldots, v_{p^2+p})$, it follows that for all v with $\|v\| = 1$,

$$
\begin{aligned}
v'\mathcal{B}'\mathcal{B}v &= \sum_{j=1}^{p}\sum_{t=1}^{n}\left(\begin{pmatrix} X^{(t-1)'} & U_j^{(t)} \end{pmatrix}\begin{pmatrix} \vec{v}_j \\ v_{p^2+j} \end{pmatrix}\right)^2 \\
&= \sum_{j=1}^{p}\sum_{t=1}^{n}\begin{pmatrix} \vec{v}_j' & v_{p^2+j} \end{pmatrix}\begin{pmatrix} X^{(t-1)} \\ U_j^{(t)} \end{pmatrix}\begin{pmatrix} X^{(t-1)'} & U_j^{(t)} \end{pmatrix}\begin{pmatrix} \vec{v}_j \\ v_{p^2+j} \end{pmatrix} \\
&= \sum_{j=1}^{p}\begin{pmatrix} \vec{v}_j' & 0 & \cdots & v_{p^2+j} \cdots & 0 \end{pmatrix} n\Omega \begin{pmatrix} \vec{v}_j \\ 0 \\ \vdots \\ v_{p^2+j} \\ \vdots \\ 0 \end{pmatrix}. \\
&\geq n\lambda_{\min}(\Omega)\underbrace{\sum_{j=1}^{p}\begin{pmatrix} \vec{v}_j' & v_{p^2+j} \end{pmatrix}\begin{pmatrix} \vec{v}_j \\ v_{p^2+j} \end{pmatrix}}_{=1},
\end{aligned}
$$

where

$$
\Omega := \frac{1}{n}\begin{pmatrix} \sum_{t=1}^{n} X^{(t-1)}X^{(t-1)'} & \sum_{t=1}^{n} X^{(t-1)}U^{(t)'} \\ \sum_{t=1}^{n} U^{(t)}X^{(t-1)'} & \sum_{t=1}^{n} U^{(t)}U^{(t)'} \end{pmatrix} = \frac{1}{n}\begin{pmatrix} \mathcal{X}\mathcal{X}' & \mathcal{X}\mathcal{U}' \\ \mathcal{U}\mathcal{X}' & \mathcal{U}\mathcal{U}' \end{pmatrix}.
$$

Then

$$\frac{1}{\sqrt{n}}\lambda_{\min}(\mathcal{B}'\mathcal{B}) \geq \sqrt{n}\lambda_{\min}(\Omega) \geq \sqrt{n}\Big(\lambda_{\min}\big(E_x(\Omega)\big) + \lambda_{\min}\big(\Omega - E_x(\Omega)\big)\Big). \tag{S15}$$

Having derived lower bounds for both $\lambda_{\min}(\Theta_g'\Theta_g)$ and $\frac{1}{\sqrt{n}}\lambda_{\min}(\mathcal{B}'\mathcal{B})$, from (S11) it follows that,

$$\begin{aligned} P_x\Big(\log\big(|\widetilde{\mathcal{D}}_{g_o}'\widetilde{\mathcal{D}}_{g_o}|\big) \geq \frac{|G_o|+p}{2}\Big) &\geq P_x\Big(1 - \lambda_{\min}^{-1}(\widetilde{\mathcal{D}}_{g_o}'\widetilde{\mathcal{D}}_{g_o}) \geq \frac{1}{2}\Big) \\ &= P_x\Big(\lambda_{\min}(\widetilde{\mathcal{D}}_{g_o}'\widetilde{\mathcal{D}}_{g_o}) \geq 2\Big) \\ &\geq P_x\Big(\frac{1}{2}(1+c^2)^{-1}\frac{1}{\sqrt{n}}\lambda_{\min}(\mathcal{B}'\mathcal{B}) \geq 2\Big), \end{aligned}$$

where the last inequality follows from (S14). Then by Condition 3.1,

$$\begin{aligned} P_x\Big(\log\big(|\widetilde{\mathcal{D}}_{g_o}'\widetilde{\mathcal{D}}_{g_o}|\big) \geq \frac{|G_o|+p}{2}\Big) &\geq P_x\Big(\frac{1}{\sqrt{n}}\lambda_{\min}(\mathcal{B}'\mathcal{B}) \geq \sqrt{n}\big[\lambda_{\min}\big(E_x(\Omega)\big) - \delta\big]\Big) \\ &\geq P_x\Big(\lambda_{\min}\big(\Omega - E_x(\Omega)\big) \geq -\delta\Big) \end{aligned}$$

where the final inequality comes from (S15). Finally, for all $n \geq \max\{N_1, N_2\}$, appealing to Lemma 3.15 yields the desired result,

$$\begin{aligned} &P_x\Big(\log\big(|\widetilde{\mathcal{D}}_{g_o}'\widetilde{\mathcal{D}}_{g_o}|\big) \geq \frac{|G_o|+p}{2}\Big) \\ &\quad \geq 1 - \delta^{-2}\Bigg[\frac{(\sigma_{\max}^0)^4(1+c^2)}{(1-c^2)^3}\Big(\frac{\min\{|G_o|,p\} + \min\{|G_o|,p\}^2}{n}\Big) \\ &\qquad + \frac{2p(\sigma_{\max}^0)^2\min\{|G_o|,p\}}{n(1-c^2)} + \frac{p(p+1)}{n}\Bigg]. \end{aligned}$$

■

Proof of Lemma S2.5. Observe

$$\prod_{j=1}^{p}\frac{\big|(\mathcal{X}\mathcal{X}')_{r_j^{g_o},r_j^{g_o}}\big|^{\frac{1}{2}}}{\big|(\mathcal{X}\mathcal{X}')_{r_j^{g},r_j^{g}}\big|^{\frac{1}{2}}} = n^{\frac{|G_o|}{2}-\frac{|G|}{2}}\prod_{j=1}^{p}\frac{\big|\widehat{\Sigma}_{r_j^{g_o},r_j^{g_o}}\big|^{\frac{1}{2}}}{\big|\widehat{\Sigma}_{r_j^{g},r_j^{g}}\big|^{\frac{1}{2}}}, \tag{S16}$$

where $\widehat{\Sigma} := \frac{1}{n}\mathcal{X}\mathcal{X}'$. First, establish the following lower bound for the denominator. For each

$j \in \{1, \ldots, p\}$,

$$\begin{aligned}
\log\left(|\widehat{\Sigma}_{r_j^g, r_j^g}|\right) &\geq \mathrm{tr}\Big(I_{|r_j^g|} - \widehat{\Sigma}^{-1}_{r_j^g, r_j^g}\Big) \\
&\geq |r_j^g| - |r_j^g| \lambda_{\max}\Big(\widehat{\Sigma}^{-1}_{r_j^g, r_j^g}\Big) \\
&= |r_j^g| \Big[1 - \lambda_{\min}^{-1}\Big(\widehat{\Sigma}_{r_j^g, r_j^g}\Big)\Big] \\
&\geq |r_j^g| \big[1 - \lambda_{\min}^{-1}(\widehat{\Sigma})\big],
\end{aligned} \tag{S17}$$

where the last inequality follows from the Poincaré separation theorem (Rao 1979). Then for all $n \geq N_2$,

$$P_x\bigg(\log\left(|\widehat{\Sigma}_{r_j^g, r_j^g}|\right) \geq |r_j^g|(1 - 2\delta^{-1})\bigg) \geq P_x\Big(\lambda_{\min}(\widehat{\Sigma}) \geq \delta/2\Big) \geq 1 - V_2,$$

where the last inequality follows by Lemma 3.19 (and thus requires Condition 3.1) with V_2 is as in (10).

Next, consider the numerator in (S16). The determinant can be bound from above as

$$\begin{aligned}
\log\left(|\widehat{\Sigma}_{r_j^{g_o}, r_j^{g_o}}|\right) &\leq \mathrm{tr}\Big(\widehat{\Sigma}_{r_j^{g_o}, r_j^{g_o}} - I_{|r_j^{g_o}|}\Big) \\
&= \mathrm{tr}\Big(\widehat{\Sigma}_{r_j^{g_o}, r_j^{g_o}}\Big) - |r_j^{g_o}| \\
&\leq |r_j^{g_o}| \Big[\lambda_{\max}\big(\widehat{\Sigma} - \Gamma_n(0)\big) + \lambda_{\max}\big(\Gamma_n(0)\big)\Big] - |r_j^{g_o}|,
\end{aligned}$$

where the final inequality follows again by the Poincaré separation theorem. Further, by the same argument as in (S8), for all $n \geq N_2$ Lemma 3.17 gives,

$$\begin{aligned}
P_x\Big(\log\left(|\widehat{\Sigma}_{r_j^{g_o}, r_j^{g_o}}|\right) \leq |r_j^{g_o}|\big[\delta + \lambda_{\max}\big(\Gamma_n(0)\big) - 1\big]\Big) &\geq P_x\Big(|\lambda_{\max}\big(\widehat{\Sigma} - \Gamma_n(0)\big)| \leq \delta\Big) \\
&\geq 1 - V_2.
\end{aligned} \tag{S18}$$

Therefore, substituting back into (S16) yields for all $n \geq N_2$,

$$\begin{aligned}
\prod_{j=1}^{p} \frac{\left|(\mathcal{X}\mathcal{X}')_{r_j^{g_o}, r_j^{g_o}}\right|^{\frac{1}{2}}}{\left|(\mathcal{X}\mathcal{X}')_{r_j^{g}, r_j^{g}}\right|^{\frac{1}{2}}} &= n^{\frac{|G_o|}{2} - \frac{|G|}{2}} \prod_{j=1}^{p} \frac{\left|\widehat{\Sigma}_{r_j^{g_o}, r_j^{g_o}}\right|^{\frac{1}{2}}}{\left|\widehat{\Sigma}_{r_j^{g}, r_j^{g}}\right|^{\frac{1}{2}}} \\
&\leq n^{\frac{|G_o|}{2} - \frac{|G|}{2}} \prod_{j=1}^{p} \frac{e^{\frac{1}{2}|r_j^{g_o}|\left[\delta + \lambda_{\max}\left(\Gamma_n(0)\right) - 1\right]}}{e^{\frac{1}{2}|r_j^{g}|(1 - 2\delta^{-1})}} \\
&\leq n^{\frac{|G_o|}{2} - \frac{|G|}{2}} e^{\frac{1}{2}\left(|G_o|[\delta + \lambda_{\max}(\Gamma_n(0))] + |G| 2\delta^{-1}\right)}
\end{aligned}$$

with probability exceeding $1 - 2V_2$. ∎

Proof of Lemma S2.6.

Assuming $1 \leq d \leq \min_{1 \leq j \leq p} m_j^{g_o}$,

$$\prod_{j=1}^{p}\left[\frac{(m_j^{g_o})^{\frac{n-|r_j^{g_o}|}{2}}}{(m_j^{g})^{\frac{n-|r_j^{g}|}{2}}}\right] \le \prod_{j=1}^{p}\frac{(m_j^{g_o})^{\frac{n}{2}}}{(m_j^{g})^{\frac{n-|r_j^{g}|}{2}}} \le \prod_{j=1}^{p}\frac{(m_j^{g_o})^{\frac{n}{2}}}{m_j^{\frac{n-|r_j^{g}|}{2}}}$$

where $m_1,\dots,m_p$ correspond to the full model. Then

$$\begin{aligned}
\prod_{j=1}^{p}\left[\frac{(m_j^{g_o})^{\frac{n-|r_j^{g_o}|}{2}}}{(m_j^{g})^{\frac{n-|r_j^{g}|}{2}}}\right] &\le \prod_{j=1}^{p}\frac{(m_j^{g_o})^{\frac{n}{2}}}{m_j^{\frac{n}{2}}}\cdot m_j^{\frac{|r_j^{g}|}{2}} \\
&= \left[\prod_{j=1}^{p} m_j^{\frac{|r_j^{g}|}{2}}\right]\cdot\prod_{j=1}^{p}\Big(\frac{m_j^{g_o}}{m_j}\Big)^{\frac{n}{2}} \\
&\le \left[\prod_{j=1}^{p}(m_j^{g_o})^{\frac{|r_j^{g}|}{2}}\right]\cdot\prod_{j=1}^{p}\Big(\frac{m_j^{g_o}-m_j+m_j}{m_j}\Big)^{\frac{n}{2}} \\
&\le \left[\prod_{j=1}^{p}(m_j^{g_o})^{\frac{p}{2}}\right]\cdot\prod_{j=1}^{p}\Big(\frac{m_j^{g_o}-m_j}{m_j}+1\Big)^{\frac{n}{2}} \\
&\le \left(\frac{1}{p}\sum_{j=1}^{p} m_j^{g_o}\right)^{\frac{p^2}{2}}\cdot\prod_{j=1}^{p}\Big(\frac{m_j^{g_o}-m_j}{q}+1\Big)^{\frac{n}{2}},
\end{aligned}$$

where the last inequality follows by the arithmetic-geometric inequality and $q := \min_{1\le j\le p} m_j$. Next, since $1+x\le e^x$ for real-valued x,

$$\begin{aligned}
\prod_{j=1}^{p}\left[\frac{(m_j^{g_o})^{\frac{n-|r_j^{g_o}|}{2}}}{(m_j^{g})^{\frac{n-|r_j^{g}|}{2}}}\right] &\le \left(\frac{1}{p}\sum_{j=1}^{p} m_j^{g_o}\right)^{\frac{p^2}{2}}\cdot\prod_{j=1}^{p} e^{\frac{m_j^{g_o}-m_j}{q}\cdot\frac{n}{2}} \\
&= (RSS_{g_o}/p)^{\frac{p^2}{2}}\cdot e^{(RSS_{g_o}-RSS)\cdot\frac{n}{2q}}
\end{aligned}$$

where $RSS_{g_o} := \sum_{j=1}^{p} m_j^{g_o}$ and $RSS := \sum_{j=1}^{p} m_j$.

From (5),

$$
\begin{aligned}
RSS_{g_o} &= \operatorname{tr}\big((\mathcal{Y} - \widehat{A}_{g_o}\mathcal{X})(\mathcal{Y} - \widehat{A}_{g_o}\mathcal{X})'\big) \\
&= \operatorname{tr}\bigg(\sum_{t=1}^{n} \big(X^{(t)} - \widehat{A}_{g_o}X^{(t-1)}\big)\big(X^{(t)} - \widehat{A}_{g_o}X^{(t-1)}\big)'\bigg) \\
&= \sum_{t=1}^{n} \big\|X^{(t)} - \widehat{A}_{g_o}X^{(t-1)}\big\|^2 \\
&= \big\|Y - \mathcal{Z}_{G_o}\widehat{\boldsymbol{\alpha}}_{g_o}\big\|^2 \\
&= \big\|(I_{np} - H_{g_o})Y\big\|^2 \\
&= \big\|(I_{np} - H_{g_o})(\mathcal{W}^0)^{\frac{1}{2}}\operatorname{vec}(\mathcal{U})\big\|^2 \\
&\leq (\sigma^0_{\max})^2 \sum_{t=1}^{n} \|U^{(t)}\|^2,
\end{aligned}
\tag{S19}
$$

where $H_{g_o} := \mathcal{Z}_{G_o}(\mathcal{Z}'_{G_o}\mathcal{Z}_{G_o})^{-1}\mathcal{Z}'_{G_o}$. Thus,

$$
\begin{aligned}
P\Big((RSS_{g_o}/p)^{\frac{p^2}{2}} > \big((\sigma^0_{\max})^2 3n\big)^{\frac{p^2}{2}}\Big) &\leq P\bigg(\sum_{t=1}^{n} \|U^{(t)}\|^2 > 3np\bigg) \\
&\leq P\big(\chi^2_{np} > 3np\big) \\
&\leq e^{-\frac{np}{4}},
\end{aligned}
\tag{S20}
$$

where the last inequality follows by the Chernoff bound (evaluating the moment generating function at 1/4).

Further, from (S19) it also follows that

$$
\begin{aligned}
RSS_{g_o} - RSS &= \big\|(I_{np} - H_{g_o})(\mathcal{W}^0)^{\frac{1}{2}}\operatorname{vec}(\mathcal{U})\big\|^2 - \big\|(I_{np} - H)(\mathcal{W}^0)^{\frac{1}{2}}\operatorname{vec}(\mathcal{U})\big\|^2 \\
&\leq \big\|(\mathcal{W}^0)^{\frac{1}{2}}\operatorname{vec}(\mathcal{U})\big\|^2 - \operatorname{vec}(\mathcal{U})'(\mathcal{W}^0)^{\frac{1}{2}}(I_{np} - H)(\mathcal{W}^0)^{\frac{1}{2}}\operatorname{vec}(\mathcal{U}) \\
&= \operatorname{vec}(\mathcal{U})'(\mathcal{W}^0)^{\frac{1}{2}}H(\mathcal{W}^0)^{\frac{1}{2}}\operatorname{vec}(\mathcal{U})
\end{aligned}
$$

where $H := \mathcal{Z}(\mathcal{Z}'\mathcal{Z})^{-1}\mathcal{Z}'$. Then

$$
\begin{aligned}
RSS_{g_o} - RSS &\leq \operatorname{vec}(\mathcal{U})'(\mathcal{W}^0)^{\frac{1}{2}}\mathcal{Z}(\mathcal{Z}'\mathcal{Z})^{-1}\mathcal{Z}'(\mathcal{W}^0)^{\frac{1}{2}}\operatorname{vec}(\mathcal{U}) \\
&\leq \lambda_{\max}\big((\mathcal{Z}'\mathcal{Z})^{-1}\big)\operatorname{vec}\big((\Sigma^0)^{\frac{1}{2}}\mathcal{U}\mathcal{X}'\big)'\operatorname{vec}\big((\Sigma^0)^{\frac{1}{2}}\mathcal{U}\mathcal{X}'\big) \\
&\leq \frac{1}{\lambda_{\min}(\mathcal{X}'\mathcal{X}) \cdot 1}(\sigma^0_{\max})^2\operatorname{tr}\big(\mathcal{X}\mathcal{U}'\mathcal{U}\mathcal{X}'\big),
\end{aligned}
$$

and so

$$
\begin{aligned}
P_x\Big(RSS_{g_o} - RSS \geq (\sigma^0_{\max})^2 p^2\sqrt{n}\Big)& \\
&\leq P_x\left(\frac{1}{\lambda_{\min}(\mathcal{X}'\mathcal{X})}(\sigma^0_{\max})^2 \mathrm{tr}\big(\mathcal{X}\mathcal{U}'\mathcal{U}\mathcal{X}'\big) \geq (\sigma^0_{\max})^2 p^2\sqrt{n}\right) \\
&= P_x\left(\frac{1}{n}\mathrm{tr}\big(\mathcal{X}\mathcal{U}'\mathcal{U}\mathcal{X}'\big) \geq p^2\sqrt{n}\lambda_{\min}(\mathcal{X}'\mathcal{X}/n)\right) \\
&\leq P_x\left(\frac{1}{n}\mathrm{tr}\big(\mathcal{X}\mathcal{U}'\mathcal{U}\mathcal{X}'\big) \geq p^2\sqrt{n}\cdot\frac{\delta}{2}\right) \\
&\qquad + P_x\Big(\lambda_{\min}(\mathcal{X}\mathcal{X}'/n) < \delta/2\Big) \\
&\leq \frac{1}{n}E_x\Big(\mathrm{tr}\big(\mathcal{X}\mathcal{U}'\mathcal{U}\mathcal{X}'\big)\Big)\frac{2}{p^2\sqrt{n}\delta} + V_2
\end{aligned}
$$

where the last inequality follows by the Markov inequality and Lemma 3.19 (assuming Condition 3.1) for all $n \geq N_2$, and V_2 is as in (10). Finally, by Lemma 3.16, for all $n \geq N_1$,

$$
\begin{aligned}
P_x\Big(RSS_{g_o} - RSS \geq (\sigma^0_{\max})^2 p^2\sqrt{n}\Big) &\leq \frac{p(\sigma^0_{\max})^2 \min\{|G_o|, p\}}{(1-c^2)}\frac{2}{p^2\sqrt{n}\delta} + V_2 \\
&\leq \frac{2(\sigma^0_{\max})^2}{\delta(1-c^2)\sqrt{n}} + V_2.
\end{aligned}
$$

Therefore, for all $n \geq \max\{N_1, N_2\}$,

$$
\begin{aligned}
P_x\left(\prod_{j=1}^{p}\left[\frac{(m_j^{g_o})^{\frac{n-|r_j^{g_o}|}{2}}}{(m_j^{g})^{\frac{n-|r_j^{g}|}{2}}}\right] \leq \big((\sigma^0_{\max})^2 3n\big)^{\frac{p^2}{2}} \cdot e^{(\sigma^0_{\max})^2 p^2\sqrt{n}\cdot\frac{n}{2q}}\right)& \\
\geq 1 - e^{-\frac{np}{4}} - \frac{2(\sigma^0_{\max})^2}{\delta(1-c^2)\sqrt{n}} - V_2.&
\end{aligned}
$$

■

References

Basu, S., Michailidis, G. et al. (2015), 'Regularized estimation in sparse high-dimensional time series models', *The Annals of Statistics* **43**(4), 1535–1567.

Hannig, J., Iyer, H., Lai, R. C. S. & Lee, T. C. M. (2016), 'Generalized fiducial inference: A reviev and new results', *Journal of American Statistical Association* **111**(515), 1346–1361.

Rao, C. R. (1979), 'Separation theorems for singular values of matrices and their applications in multivariate analysis', *Journal of Multivariate Analysis* **9**(3), 362–377.