

Online Appendix to “Selecting the Relevant Variables for Factor Estimation in a Factor-Augmented VAR Model”

John C. Chao* and Norman R. Swanson[†]

April 11, 2022

Abstract

This Online Appendix is comprised of two sections. In the first section, we give a theorem which shows that, even in the case where the conventional factor pervasiveness assumption does not hold, consistent estimation of the factors can be achieved if the variables used for factor estimation are selected based on the variable selection procedure introduced in our main paper. Section 2 of this Online Appendix then gives a proof of this theorem as well as the proofs of some supporting lemmas.

1 Consistent Estimation of Factors after Variable Pre-Screening

In this section of the Online Appendix, we give a theorem which shows that if the variable selection procedure introduced in our main paper is used to pre-screen the variables prior to factor estimation; then, consistent factor estimation, up to an invertible matrix transformation, can be achieved. It should be noted that being able to estimate the factors consistently up to an invertible matrix transformation is already sufficient for us to be able to consistently estimate the conditional mean function of a factor-augmented forecast equation. This has been shown in an earlier version of our paper, Chao and Swanson (2022a); see Theorem 5 of that paper. Hence, we do not make further identifying assumptions here to

*Department of Economics, 7343 Preinkert Drive, University of Maryland, jcchao@umd.edu

[†]Department of Economics, 9500 Hamilton Street, Rutgers University, nswanson@econ.rutgers.edu.

[‡]The authors are grateful to Simon Freyaldenhoven, Yuan Liao, Minchul Shin, Xiye Yang, and seminar participants at the Federal Reserve Bank of Philadelphia for useful comments received on earlier versions of this paper. Chao thanks the University of Maryland for research support.

try to estimate the factors consistently, say, up to a sign change. However, consistent factor estimation up to a sign change could be achieved within our framework, if we were to make additional identifying assumptions such as those given in Stock and Watson (2002).

For the factor estimation problem, we need to impose some further conditions, in addition to the assumptions given in the body of the main paper. These additional assumptions are stated below.

Assumption OA-1: There exists a constant $\underline{C} > 0$ such that $\inf_t \lambda_{\min} \{E[\varepsilon_t \varepsilon_t']\} \geq \underline{C} > 0$.

Assumption OA-2: There exists a positive constant C such that $\sup_t \left(\frac{1}{N_1} \sum_{i \in H^c} \sum_{k \in H^c} |E[u_{i,t} u_{k,t}]| \right) \leq C < \infty$ for every positive integer N_1 , where $H^c = \{k \in \{1, \dots, N\} : \gamma_k \neq 0\}$.

Assumption OA-3: There exists a positive constant \overline{C} , such that:

$$0 < \frac{1}{\overline{C}} \leq \lambda_{\min} \left(\frac{\Gamma' \Gamma}{N_1} \right) \leq \lambda_{\max} \left(\frac{\Gamma' \Gamma}{N_1} \right) \leq \overline{C} < \infty \text{ for all } N_1, N_2 \text{ sufficiently large,}$$

where N_1 is the number of relevant variables or the number of components of the subvector $Z_t^{(1)}$ and N_2 is the number of irrelevant variables or the number of components of the subvector $Z_t^{(2)}$, with $Z_t^{(1)}$ and $Z_t^{(2)}$ as defined in expressions (9) and (10) of the main paper.

Assumption OA-4: Let $\mathbb{S}_{i,T}^+$ denote either the statistic $\max_{1 \leq \ell \leq d} |S_{i,\ell,T}|$ or the statistic $\sum_{\ell=1}^d \varpi_\ell |S_{i,\ell,T}|$, and let φ be the tuning parameter of the variable selection decision rule

$$i \in \begin{cases} \widehat{H}^c & \text{if } \mathbb{S}_{i,T}^+ \geq \Phi^{-1} \left(1 - \frac{\varphi}{2N} \right) \\ \widehat{H} & \text{if } \mathbb{S}_{i,T}^+ < \Phi^{-1} \left(1 - \frac{\varphi}{2N} \right) \end{cases}, \quad (1)$$

as described in section 2 of the main paper. In addition, let $N = N_1 + N_2$, and assume that φ satisfies the following three conditions: (a) $\varphi \rightarrow 0$ as $N_1, N_2 \rightarrow \infty$, (b) there exists some constant $a > 0$, such that $\varphi \geq \frac{1}{N^a}$ for all N_1, N_2 sufficiently large, and (c)

$$\max \left\{ \frac{N^{\frac{2}{7}} \varphi^{\frac{5}{7}}}{N_1}, \frac{N^{\frac{1}{3}} \varphi}{N_1 T} \right\} \rightarrow 0 \text{ as } N_1, N_2, T \rightarrow \infty.$$

Remark OA1.1:

- (a) Since the factor loading matrix Γ is an $N \times Kp$ matrix, the matrix $\Gamma' \Gamma$ will have order of magnitude equal to N under the conventional assumption of factor pervasiveness. Much of the factor analysis literature in both econometrics and statistics has studied the case where factors are pervasive in this sense. Assumption OA-3 above allows for possible violations of this conventional pervasiveness assumption, which will occur in our setup when $N_1/N \rightarrow 0$.
(b) Note that the rate condition given in part (c) of Assumption OA-4 depends on N_1 .

However, if we choose φ so that $\varphi N^{\frac{2}{5}} = O(1)$; then,

$$\frac{N^{\frac{2}{7}} \varphi^{\frac{5}{7}}}{N_1} = O\left(\frac{1}{N_1}\right) = o(1) \quad \text{and} \quad \frac{N^{\frac{1}{3}} \varphi}{N_1 T} = O\left(\frac{1}{N_1 N^{\frac{1}{15}} T}\right) = o\left(\frac{1}{N_1}\right).$$

Hence, with this choice of φ , Assumption OA-4 part (c) will be satisfied as long as $N_1 \rightarrow \infty$, and there is no need to impose any further condition on the rate at which N_1 grows. Requiring that $N_1 \rightarrow \infty$ is a minimal condition, since if $N_1 \not\rightarrow \infty$; then consistent factor estimation, even up to an invertible matrix transformation, is impossible. Moreover, Monte Carlo results reported in Section 3 of the main paper show that our variable selection procedure performs very well in finite samples, under the tuning parameter choice $\varphi = N^{-\frac{2}{5}}$, both in terms of controlling the probability of a false positive (or Type I) error and in terms of controlling the probability of a false negative (or Type II) error.

Next, consider the post-variable-selection principal component estimator of $\underline{F}_t = (F'_t, F'_{t-1}, \dots, F'_{t-p+1})$ given by

$$\hat{\underline{F}}_t = \frac{\hat{\Gamma}' Z_{t,N}(\widehat{H}^c)}{\hat{N}_1}, \quad (2)$$

where $Z_{t,N}(\widehat{H}^c) = \left[Z_{1,t} \mathbb{I}\{1 \in \widehat{H}^c\}, Z_{2,t} \mathbb{I}\{2 \in \widehat{H}^c\}, \dots, Z_{N,t} \mathbb{I}\{N \in \widehat{H}^c\} \right]'$, with

$$\mathbb{I}\{i \in \widehat{H}^c\} = \begin{cases} 1 & \text{if } i \in \widehat{H}^c, \text{ i.e., if } \mathbb{S}_{i,T}^+ \geq \Phi^{-1}\left(1 - \frac{\varphi}{2N}\right) \\ 0 & \text{if } i \in \widehat{H}, \text{ i.e., if } \mathbb{S}_{i,T}^+ < \Phi^{-1}\left(1 - \frac{\varphi}{2N}\right) \end{cases},$$

and where $\hat{N}_1 = \#(\widehat{H}^c)$, i.e., the cardinality of the set \widehat{H}^c . Here, $\hat{\Gamma}$ denotes the principal component estimator of the loading matrix Γ , constructed from taking $\sqrt{\hat{N}_1}$ times a matrix whose columns are the eigenvectors associated with the Kp largest eigenvalues of the post-variable-selection sample covariance matrix $\hat{\Sigma}(\widehat{H}^c)$, where, in this case,

$$\hat{\Sigma}(\widehat{H}^c) = \frac{Z(\widehat{H}^c)' Z(\widehat{H}^c)}{\hat{N}_1 T_0} = \frac{1}{\hat{N}_1 T_0} \sum_{t=p}^T Z_{t,N}(\widehat{H}^c) Z_{t,N}(\widehat{H}^c)', \quad (3)$$

with $T_0 = T - p + 1$.

Our next result shows that the estimator given in expression (2) consistently estimates the unobserved factors \underline{F}_t , up to an invertible $Kp \times Kp$ matrix transformation.

Theorem 3: *Suppose that Assumptions 2-1, 2-2, 2-3, 2-4, 2-5, 2-6, 2-7, 2-8, and 2-9 of the main paper hold. Suppose, in addition, that Assumptions OA-1, OA-2, OA-3, and OA-4*

also hold. Let $\widehat{\underline{F}}_t$ be as defined in expression (2). Then,

$$\left\| \widehat{\underline{F}}_t - Q' \underline{F}_t \right\|_2 = o_p(1), \text{ for all fixed } t,$$

where $Q = (\Gamma' \Gamma / N_1)^{\frac{1}{2}} \Xi \widehat{V}$, and where \widehat{V} is the $Kp \times Kp$ orthogonal matrix given in expression (6) below, and Ξ is a $Kp \times Kp$ orthogonal matrix whose columns are the eigenvectors of the matrix

$$M_{FF}^* = \left(\frac{\Gamma' \Gamma}{N_1} \right)^{1/2} M_{FF} \left(\frac{\Gamma' \Gamma}{N_1} \right)^{1/2} = \left(\frac{\Gamma' \Gamma}{N_1} \right)^{1/2} \frac{1}{T_0} \sum_{t=p}^T E [\underline{F}_t \underline{F}_t'] \left(\frac{\Gamma' \Gamma}{N_1} \right)^{1/2}.$$

Remark OA1.2:

(a) If we examine the proof of Theorem 3 as well as the supporting arguments given in the proof of Lemma OA-2 (both of which can be found in the next section of this Online Appendix), we see that two of the key components of the proof involve showing that $\left\| \left(\Gamma \left(\widehat{H}^c \right) - \Gamma \right) / \sqrt{N_1} \right\|_2 \xrightarrow{p} 0$ and that $\left(\widehat{N}_1 - N_1 \right) / N_1 \xrightarrow{p} 0$. This is one of the reasons why in Remark 2.3(b) of the main paper, we have argued that initial variable selection should focus on determining which variables load strongly on the factors without worrying specifically at that stage about the related issues of predictability or, for that matter, any other issue. By contrast, if we make our initial variable selection based on some more stringent criterion that takes into consideration not only variable relevance but also other concerns such as predictability, then, we may end up with a much smaller set \widetilde{H}^c of selected variables relative to the set \widehat{H}^c selected under our procedure. In particular, in this case, it may be possible that even in large samples a significant number of the rows of $\Gamma \left(\widetilde{H}^c \right)$ may contain only zero elements even though the corresponding rows of Γ are not zero vectors, so that the desired result $\left\| \left(\Gamma \left(\widetilde{H}^c \right) - \Gamma \right) / \sqrt{N_1} \right\|_2 \xrightarrow{p} 0$ may not hold. For the same reason, if we let \widetilde{N}_1 denote the cardinality of the set of selected indices based on an alternative, more stringent variable selection procedure, then, the result $\left(\widetilde{N}_1 - N_1 \right) / N_1 \xrightarrow{p} 0$ also may not hold, since, by definition, N_1 is the number of rows of Γ which have at least one non-zero element.

(b) Note that, with the proper specification of the tuning parameter φ , the consistent factor estimation result given in Theorem 3 can be attained without the condition $N / (TN_1) \rightarrow 0$ given in Bai and Ng (2021). Theorem 3 does require that $N_1 \rightarrow \infty$, which is a minimal condition as previously argued in Remark OA1.1(b).

2 Proofs of Theorem 3 and of Supporting Lemmas

This section gives the proof of Theorem 3 followed by the statement and proof of two supporting lemmas: Lemmas OA-1 and Lemma OA-2.

Proof of Theorem 3:

To proceed, note first that the principal component estimator of \underline{E}_t can be written as $\hat{\underline{E}}_t = \hat{\Gamma}' Z_{t,N}(\widehat{H}^c) / \hat{N}_1$, where $\hat{\Gamma} = \sqrt{\hat{N}_1} \hat{B}$ and where the columns of the matrix \hat{B} are the eigenvectors associated with the Kp largest eigenvalues of the (post-variable-selection) sample covariance matrix $\hat{\Sigma}(\widehat{H}^c) = Z(\widehat{H}^c)' Z(\widehat{H}^c) / (\hat{N}_1 T_0)$. Moreover, by the result of part (d) of Lemma OA-1, the matrix \hat{B} has the representation $\hat{B} = \hat{G}_1 \hat{V}$, where \hat{G}_1 is an $N \times Kp$ matrix, whose columns define an orthonormal basis for an invariant subspace of $\hat{\Sigma}(\widehat{H}^c)$ and where \hat{V} is a $Kp \times Kp$ orthogonal matrix as defined in expression (6) in part (c) of Lemma OA-1. Making use of this representation, we can further write

$$\begin{aligned} \hat{\underline{E}}_t - Q' \underline{E}_t &= \frac{\sqrt{\hat{N}_1} \hat{V}' \hat{G}_1' Z_{t,N}(\widehat{H}^c)}{\hat{N}_1} - Q' \underline{E}_t \\ &= \frac{\hat{V}' \hat{G}_1' \Gamma(\widehat{H}^c) \underline{E}_t}{\sqrt{\hat{N}_1}} + \frac{\hat{V}' \hat{G}_1' U_{t,N}(\widehat{H}^c)}{\sqrt{\hat{N}_1}} - Q' \underline{E}_t \\ &= \left(\frac{\hat{V}' \hat{G}_1' \Gamma(\widehat{H}^c)}{\sqrt{\hat{N}_1}} - Q' \right) \underline{E}_t + \frac{\hat{V}' \hat{G}_1' U_{t,N}(\widehat{H}^c)}{\sqrt{\hat{N}_1}} \end{aligned}$$

where $\Gamma(\widehat{H}^c) = (\mathbb{I}\{1 \in \widehat{H}^c\} \gamma_1, \dots, \mathbb{I}\{N \in \widehat{H}^c\} \gamma_N)'$ and

$U_{t,N}(\widehat{H}^c) = (\mathbb{I}\{1 \in \widehat{H}^c\} u_{1,t}, \dots, \mathbb{I}\{N \in \widehat{H}^c\} u_{N,t})'$. Next, write

$\left(\hat{V}' \hat{G}_1' \Gamma / \sqrt{\hat{N}_1} \right) - Q' = \left[\left(1 + [\hat{N}_1 - N_1] / N_1 \right)^{-\frac{1}{2}} - 1 \right] \left(\hat{V}' \hat{G}_1' \Gamma / \sqrt{N_1} \right) + \left(\hat{V}' \hat{G}_1' \Gamma / \sqrt{N_1} \right) - Q'$
and $\left(\Gamma(\widehat{H}^c) - \Gamma \right) / \sqrt{\hat{N}_1} = \left(1 + [\hat{N}_1 - N_1] / N_1 \right)^{-\frac{1}{2}} \left(\left[\Gamma(\widehat{H}^c) - \Gamma \right] / \sqrt{N_1} \right)$, so that

$$\begin{aligned} \frac{\hat{V}' \hat{G}_1' \Gamma(\widehat{H}^c) \underline{E}_t}{\sqrt{\hat{N}_1}} &= Q' \underline{E}_t + \left(\frac{\hat{V}' \hat{G}_1' \Gamma}{\sqrt{\hat{N}_1}} - Q' \right) \underline{E}_t + \hat{V}' \hat{G}_1' \left(\frac{\Gamma(\widehat{H}^c) - \Gamma}{\sqrt{\hat{N}_1}} \right) \underline{E}_t \\ &= Q' \underline{E}_t + \left(\frac{\hat{V}' \hat{G}_1' \Gamma}{\sqrt{N_1}} - Q' \right) \underline{E}_t + \left[\left(1 + \frac{\hat{N}_1 - N_1}{N_1} \right)^{-\frac{1}{2}} - 1 \right] \frac{\hat{V}' \hat{G}_1' \Gamma}{\sqrt{N_1}} \underline{E}_t \\ &\quad + \left[\left(1 + \frac{\hat{N}_1 - N_1}{N_1} \right)^{-\frac{1}{2}} \right] \hat{V}' \hat{G}_1' \left(\frac{\Gamma(\widehat{H}^c) - \Gamma}{\sqrt{N_1}} \right) \underline{E}_t \end{aligned}$$

It follows that

$$\begin{aligned}
\widehat{\underline{E}}_t - Q' \underline{E}_t &= \left(\frac{\widehat{V}' \widehat{G}'_1 \Gamma(\widehat{H}^c)}{\sqrt{\widehat{N}_1}} - Q' \right) \underline{E}_t + \frac{\widehat{V}' \widehat{G}'_1 U_{t,N}(\widehat{H}^c)}{\sqrt{\widehat{N}_1}} \\
&= \left(\frac{\widehat{V}' \widehat{G}'_1 \Gamma}{\sqrt{\widehat{N}_1}} - Q' \right) \underline{E}_t + \left[\left(1 + \frac{\widehat{N}_1 - N_1}{N_1} \right)^{-\frac{1}{2}} - 1 \right] \frac{\widehat{V}' \widehat{G}'_1 \Gamma}{\sqrt{\widehat{N}_1}} \underline{E}_t \\
&\quad + \left[\left(1 + \frac{\widehat{N}_1 - N_1}{N_1} \right)^{-\frac{1}{2}} \right] \widehat{V}' \widehat{G}'_1 \left(\frac{\Gamma(\widehat{H}^c) - \Gamma}{\sqrt{\widehat{N}_1}} \right) \underline{E}_t + \frac{\widehat{V}' \widehat{G}'_1 U_{t,N}(\widehat{H}^c)}{\sqrt{\widehat{N}_1}}
\end{aligned}$$

Hence, applying the triangle inequality as well as parts (a)-(c), (g), and (h) of Lemma OA-2 along with the Slutsky's theorem, we obtain

$$\begin{aligned}
&\left\| \widehat{\underline{E}}_t - Q' \underline{E}_t \right\|_2 \\
&\leq \left\| \frac{\widehat{V}' \widehat{G}'_1 \Gamma}{\sqrt{\widehat{N}_1}} - Q' \right\|_2 \|\underline{E}_t\|_2 + \left| \left(1 + \frac{\widehat{N}_1 - N_1}{N_1} \right)^{-\frac{1}{2}} - 1 \right| \left\| \frac{\widehat{V}' \widehat{G}'_1 \Gamma}{\sqrt{\widehat{N}_1}} \right\|_2 \|\underline{E}_t\|_2 \\
&\quad + \left| \left(1 + \frac{\widehat{N}_1 - N_1}{N_1} \right)^{-\frac{1}{2}} \right| \left\| \widehat{V}' \widehat{G}'_1 \right\|_2 \left\| \frac{\Gamma(\widehat{H}^c) - \Gamma}{\sqrt{\widehat{N}_1}} \right\|_2 \|\underline{E}_t\|_2 + \left\| \frac{\widehat{V}' \widehat{G}'_1 U_{t,N}(\widehat{H}^c)}{\sqrt{\widehat{N}_1}} \right\|_2 \\
&= \left\| \frac{\widehat{V}' \widehat{G}'_1 \Gamma}{\sqrt{\widehat{N}_1}} - Q' \right\|_2 \|\underline{E}_t\|_2 + \left| \left(1 + \frac{\widehat{N}_1 - N_1}{N_1} \right)^{-\frac{1}{2}} - 1 \right| \left\| \frac{\widehat{V}' \widehat{G}'_1 \Gamma}{\sqrt{\widehat{N}_1}} \right\|_2 \|\underline{E}_t\|_2 \\
&\quad + \left| \left(1 + \frac{\widehat{N}_1 - N_1}{N_1} \right)^{-\frac{1}{2}} \right| \left\| \frac{\Gamma(\widehat{H}^c) - \Gamma}{\sqrt{\widehat{N}_1}} \right\|_2 \|\underline{E}_t\|_2 + \left\| \frac{\widehat{V}' \widehat{G}'_1 U_{t,N}(\widehat{H}^c)}{\sqrt{\widehat{N}_1}} \right\|_2 \\
&\quad \left(\text{since } \left\| \widehat{V}' \widehat{G}'_1 \right\|_2 = \sqrt{\lambda_{\max}(\widehat{G}_1 \widehat{V} \widehat{V}' \widehat{G}'_1)} = \sqrt{\lambda_{\max}(\widehat{V}' \widehat{G}'_1 \widehat{G}_1 \widehat{V})} = \sqrt{\lambda_{\max}(I_{Kp})} = 1 \right) \\
&= o_p(1) O_p(1) + o_p(1) O_p(1) O_p(1) + O_p(1) o_p(1) O_p(1) + o_p(1) \\
&= o_p(1). \quad \square
\end{aligned}$$

Lemma OA-1: Let $\widehat{\Sigma}(\widehat{H}^c)$ be the post-variable-selection sample covariance matrix as defined in expression (3), and let $T_0 = T - p + 1$, as before. Decompose $\widehat{\Sigma}(\widehat{H}^c)$ as $\widehat{\Sigma}(\widehat{H}^c) =$

$A + E$, where $A = \Gamma M_{FF} \Gamma' / N_1$ and where

$$\begin{aligned}
E &= \widehat{\Sigma}(\widehat{H}^c) - \frac{\Gamma M_{FF} \Gamma'}{N_1} \\
&= \left(\Gamma(\widehat{H}^c) M_{FF} \Gamma(\widehat{H}^c)' / \widehat{N}_1 - \Gamma M_{FF} \Gamma' / N_1 \right) + \frac{1}{\widehat{N}_1} \Gamma(\widehat{H}^c) \left[\frac{F' F}{T_0} - M_{FF} \right] \Gamma(\widehat{H}^c)' \\
&\quad + \frac{U(\widehat{H}^c)' \underline{F} \Gamma(\widehat{H}^c)'}{\widehat{N}_1 T_0} + \frac{\Gamma(\widehat{H}^c) \underline{F}' U(\widehat{H}^c)}{\widehat{N}_1 T_0} + \frac{U(\widehat{H}^c)' U(\widehat{H}^c)}{\widehat{N}_1 T_0},
\end{aligned} \tag{4}$$

with $U(\widehat{H}^c) = (\mathbb{I}\{1 \in \widehat{H}^c\} u_1, \dots, \mathbb{I}\{N \in \widehat{H}^c\} u_N)$ and $M_{FF} = T_0^{-1} \sum_{t=p}^T E[F_t F_t']$. Let Assumptions 2-1, 2-2, 2-3, 2-4 2-5, 2-6, 2-7, 2-9, OA-1, OA-2, OA-3, and OA-4 all hold; and define $G_{N \times N} = \begin{bmatrix} G_1 & G_2 \\ N \times Kp & N \times (N-Kp) \end{bmatrix}$ to be an orthogonal matrix whose columns are the eigenvectors of the matrix A . Without loss of generality, suppose also that the columns of G_1 are the eigenvectors associated with the non-zero eigenvalues of A , whereas G_2 contains the eigenvectors associated with the zero eigenvalue which has an algebraic multiplicity of $N - Kp$ in this case¹. Partition the matrices $G'AG$ and $G'EG$ as follows:

$$\begin{aligned}
G'AG &= \begin{pmatrix} \Lambda_1 & 0 \\ Kp \times Kp & Kp \times (N-Kp) \\ 0 & \Lambda_2 \\ (N-Kp) \times Kp & (N-Kp) \times (N-Kp) \end{pmatrix} = \begin{pmatrix} \Lambda_1 & 0 \\ Kp \times Kp & Kp \times (N-Kp) \\ 0 & 0 \\ (N-Kp) \times Kp & (N-Kp) \times (N-Kp) \end{pmatrix} \text{ and} \\
G'EG &= \begin{pmatrix} E_{11} & E'_{21} \\ Kp \times Kp & Kp \times (N-Kp) \\ E_{21} & E_{22} \\ (N-Kp) \times Kp & (N-Kp) \times (N-Kp) \end{pmatrix},
\end{aligned}$$

where Λ_1 is a diagonal matrix whose diagonal elements are the Kp largest eigenvalues of the matrix A .²

Under the assumed conditions, the following statements are true.

- (a) There exists a $(N - Kp) \times Kp$ matrix R such that the columns of the matrix $\widehat{G}_1 = (G_1 + G_2 R)(I_{Kp} + R' R)^{-1/2}$ define an orthonormal basis for a subspace that is invariant for $\widehat{\Sigma}(\widehat{H}^c) = A + E$ w.p.a.1. Moreover, $\|R\|_2 = o_p(1)$ as N_1, N_2 , and $T \rightarrow \infty$.
- (b) $\|\widehat{G}_1 - G_1\|_2 = o_p(1)$ as N_1, N_2 , and $T \rightarrow \infty$.

¹An explicit proof that 0 is an eigenvalue of the matrix $A = \Gamma M_{FF} \Gamma' / N_1$ with algebraic multiplicity equaling $N - Kp$ is given in the Technical Appendix of an earlier version of this paper, Chao and Swanson (2022b) which is available at http://econweb.umd.edu/~chao/Research/research_files/AppConFacVarSel-03-18-2022.pdf. In particular, this result is shown in part (a) of Lemma D-11 in Appendix D of Chao and Swanson (2022b).

²An explicit proof showing that $G'AG$ can be partitioned in the manner given here is also provided in the proof of Lemma D-11 of Chao and Swanson (2022b).

(c) There exists a unique symmetric matrix L such that $(A + E)\hat{G}_1 = \hat{G}_1 L$. Moreover, let

$$\hat{\Lambda} = \text{diag}(\hat{\lambda}_1, \dots, \hat{\lambda}_{Kp}) \quad (5)$$

denote a diagonal matrix whose diagonal elements are the eigenvalues of the matrix L , and let

$$\hat{V} = (\hat{v}_1 \ \hat{v}_2 \ \dots \ \hat{v}_{Kp}) \quad (6)$$

be a $Kp \times Kp$ matrix whose ℓ^{th} column (i.e., \hat{v}_ℓ) is an eigenvector of L associated with the eigenvalue $\hat{\lambda}_\ell$ for $\ell = 1, \dots, Kp$. Then, \hat{V} is an orthogonal matrix and $(\hat{G}_1 \hat{v}_\ell, \hat{\lambda}_\ell)$ is an eigenpair for the matrix $A + E$ for $\ell = 1, \dots, Kp$.

(d) The columns of the matrix $\hat{G}_1 \hat{V} = (\hat{G}_1 \hat{v}_1 \ \hat{G}_1 \hat{v}_2 \ \dots \ \hat{G}_1 \hat{v}_{Kp})$ are the eigenvectors associated with the Kp largest eigenvalues of the post-variable-selection sample covariance matrix $A + E = \hat{\Sigma}(\hat{H}^c)$.

Proof of Lemm OA-1:

To show part (a), we first verify that two key conditions of Theorem 8.1.10 of Golub and van Loan (1996), i.e.,

$$\text{sep}(\Lambda_1, \Lambda_2) = \min_{X \neq 0} \frac{\|\Lambda_1 X - X \Lambda_2\|_F}{\|X\|_F} > 0 \quad (7)$$

and

$$\|E\|_2 \leq \frac{\text{sep}(\Lambda_1, \Lambda_2)}{5} = \frac{1}{5} \min_{X \neq 0} \frac{\|\Lambda_1 X - X \Lambda_2\|_F}{\|X\|_F}, \quad (8)$$

are satisfied here³. To proceed, let $\text{ran}(G_1)$ denote the range space of G_1 , i.e., $\text{ran}(G_1) = \{g \in \mathbb{R}^N : g = G_1 b \text{ for some } b \in \mathbb{R}^{Kp}\}$; and, by definition, Λ_1 is a $Kp \times Kp$ diagonal matrix whose diagonal elements are the non-zero eigenvalues of the matrix $A = \Gamma M_{FF} \Gamma' / N_1$. Now, for any $\tilde{g} \in \text{ran}(G_1)$, note that there exists $b \in \mathbb{R}^{Kp}$ such that $g^* = A \tilde{g} = (\Gamma M_{FF} \Gamma' / N_1) G_1 b = G_1 \Lambda_1 b = G_1 b^*$, where $b^* = \Lambda_1 b$, from which it follows that $g^* \in \text{ran}(G_1)$, so that $\text{ran}(G_1)$ is an invariant subspace of A . Next, given Assumptions 2-1, 2-2(a)-(b), 2-5, 2-6, OA-1, and OA-3, it is straightforward to show that there exists a positive constant \underline{c} such that

$$\text{sep}(\Lambda_1, \Lambda_2) = \text{sep}(\Lambda_1, 0) = \min_{X \neq 0} \frac{\|\Lambda_1 X\|_F}{\|X\|_F} \geq \lambda_{\min}(\Lambda_1) = \lambda_{\min} \left(\frac{M_{FF}^{1/2} \Gamma' \Gamma M_{FF}^{1/2}}{N_1} \right) \geq \underline{c} > 0 \quad (9)$$

for all N_1 and N_2 sufficiently large⁴, so that the condition given in expression (7) is fulfilled.

³It should be noted that Golub and van Loan (1996) use slightly different notations than we do here. In particular, the two conditions given in expressions (7) and (8), if stated in their notations, would be $\text{sep}(D_1, D_2) > 0$ and $\|E\|_2 \leq \frac{\text{sep}(D_1, D_2)}{5}$.

⁴A more detailed proof of the result that $\text{sep}(\Lambda_1, \Lambda_2) \geq \underline{c} > 0$ can be found in the Technical Appendix of an earlier version of this paper, Chao and Swanson (2022b). In particular, this result is shown in part (c) of Lemma D-11 in Appendix D of Chao and Swanson (2022b).

In addition, by the result given in Lemma D-10 of Chao and Swanson (2022b), we have $\|E\|_2 = \left\| \widehat{\Sigma} \left(\widehat{H}^c \right) - (\Gamma M_{FF} \Gamma' / N_1) \right\|_2 = o_p(1)$ as N_1, N_2 , and $T \rightarrow \infty$; from which it follows that $\|E\|_2 \leq \text{sep}(\Lambda_1, 0) / 5$ w.p.a.1 as N_1, N_2 , and $T \rightarrow \infty$, so that the condition given in expression (8) is also satisfied w.p.a.1. Hence, application of Theorem 8.1.10 of Golub and van Loan (1996) allows us to conclude that there exists a $(N - Kp) \times Kp$ matrix R such that the columns of the matrix $\widehat{G}_1 = (G_1 + G_2 R) (I_{Kp} + R' R)^{-1/2}$ define an orthonormal basis for a subspace that is invariant for $A + E$ w.p.a.1. In addition,

$$\begin{aligned} \|R\|_2 &\leq \frac{4}{\text{sep}(\Lambda_1, 0)} \|E\|_2 \leq 4 \left[\lambda_{\min} \left(\frac{M_{FF}^{1/2} \Gamma' \Gamma M_{FF}^{1/2}}{N_1} \right) \right]^{-1} \|E\|_2 \\ &\leq \frac{4}{\underline{c}} \|E\|_2 \quad (\text{for some } \underline{c} > 0 \text{ given expression (9)}) \\ &= o_p(1), \end{aligned}$$

which shows result (a).

To show that $\left\| \widehat{G}_1 - G_1 \right\|_2 = o_p(1)$, we first show that an explicit representation for G_1 can be given as $G_1 = (\Gamma / \sqrt{N_1}) (\Gamma' \Gamma / N_1)^{-1/2} \Xi = \Gamma (\Gamma' \Gamma)^{-1/2} \Xi$, where Ξ is an orthogonal matrix whose columns are eigenvectors of the matrix $M_{FF}^* = (\Gamma' \Gamma / N_1)^{1/2} M_{FF} (\Gamma' \Gamma / N_1)^{1/2}$. To see that this representation satisfies the various properties we require of G_1 , note first that $G_1' G_1 = \Xi' (\Gamma' \Gamma / N_1)^{-1/2} (\Gamma' \Gamma / N_1) (\Gamma' \Gamma / N_1)^{-1/2} \Xi = I_{Kp}$; hence, G_1 so represented does have orthonormal columns. Moreover, note that

$$\begin{aligned} \frac{\Gamma M_{FF} \Gamma'}{N_1} G_1 &= \frac{\Gamma}{\sqrt{N_1}} M_{FF} \frac{\Gamma' \Gamma}{N_1} \left(\frac{\Gamma' \Gamma}{N_1} \right)^{-1/2} \Xi = \frac{\Gamma}{\sqrt{N_1}} \left(\frac{\Gamma' \Gamma}{N_1} \right)^{-1/2} M_{FF}^* \Xi = \Gamma (\Gamma' \Gamma)^{-1/2} \Xi \Lambda_1 \\ &= G_1 \Lambda_1 \end{aligned} \tag{10}$$

where Λ_1 is a $Kp \times Kp$ diagonal matrix whose diagonal elements are the eigenvalues of the matrix M_{FF}^* , which also happen to be the non-zero eigenvalues of the matrix $A = \Gamma M_{FF} \Gamma' / N_1$. Pre-multiplying the above equation by G_1' , we obtain $G_1' (\Gamma M_{FF} \Gamma' / N_1) G_1 = G_1' G_1 \Lambda_1 = \Lambda_1$. Since equation (10) shows that the columns of $\Gamma (\Gamma' \Gamma)^{-1/2} \Xi$ are indeed the eigenvectors of the matrix $A = \Gamma M_{FF} \Gamma' / N_1$, by the argument given previously in the proof of part (a) above, we can then deduce that $\text{ran}(G_1)$, the range space of G_1 with $G_1 = \Gamma (\Gamma' \Gamma)^{-1/2} \Xi$, is an invariant subspace of A . It follows that setting $G_1 = \Gamma (\Gamma' \Gamma)^{-1/2} \Xi$ fulfills all the required properties which we have previously specified for G_1 .

Next, write

$$\begin{aligned} \widehat{G}_1 - G_1 &= (G_1 + G_2 R) (I_{Kp} + R' R)^{-1/2} - G_1 \\ &= G_1 \left[(I_{Kp} + R' R)^{-1/2} - I_{Kp} \right] + G_2 R (I_{Kp} + R' R)^{-1/2} \end{aligned}$$

$$= \frac{\Gamma}{\sqrt{N_1}} \left(\frac{\Gamma' \Gamma}{N_1} \right)^{-1/2} \Xi \left[(I_{Kp} + R' R)^{-1/2} - I_{Kp} \right] + G_2 R (I_{Kp} + R' R)^{-1/2}$$

Applying the submultiplicative property of matrix norms and the triangle inequality, we obtain

$$\begin{aligned} \left\| \hat{G}_1 - G_1 \right\|_2 &\leq \left\| \frac{\Gamma}{\sqrt{N_1}} \left(\frac{\Gamma' \Gamma}{N_1} \right)^{-1/2} \right\|_2 \left\| \Xi \right\|_2 \left\| (I_{Kp} + R' R)^{-1/2} - I_{Kp} \right\|_2 \\ &\quad + \left\| G_2 \right\|_2 \left\| R \right\|_2 \left\| (I_{Kp} + R' R)^{-1/2} \right\|_2 \\ &= \left\| I_{Kp} - (I_{Kp} + R' R)^{-1/2} \right\|_2 + \left\| R \right\|_2 \left\| (I_{Kp} + R' R)^{-1/2} \right\|_2 \end{aligned}$$

where the last equality follows from the fact that $\left\| \Xi \right\|_2 = \sqrt{\lambda_{\max}(\Xi' \Xi)} = \sqrt{\lambda_{\max}(I_{Kp})} = 1$, $\left\| G_2 \right\|_2 = \sqrt{\lambda_{\max}(G_2' G_2)} = \sqrt{\lambda_{\max}(I_{N-Kp})} = 1$, $\left\| (I_{Kp} + R' R)^{-1/2} - I_{Kp} \right\|_2 = \left\| I_{Kp} - (I_{Kp} + R' R)^{-1/2} \right\|_2$, and

$$\left\| \frac{\Gamma}{\sqrt{N_1}} \left(\frac{\Gamma' \Gamma}{N_1} \right)^{-1/2} \right\|_2 = \sqrt{\lambda_{\max} \left\{ \left(\frac{\Gamma' \Gamma}{N_1} \right)^{-1/2} \frac{\Gamma' \Gamma}{N_1} \left(\frac{\Gamma' \Gamma}{N_1} \right)^{-1/2} \right\}} = \sqrt{\lambda_{\max} \{I_{Kp}\}} = 1.$$

Now, if (λ, ρ) is an eigen-pair of $R' R$; then, by definition, $R' R \rho = \lambda \rho$ with $\lambda \geq 0$, since $R' R$ is a positive semidefinite matrix. It follows by elementary properties of eigenvalues and eigenvectors that, in this case, $(\sqrt{1 + \lambda} - 1) / \sqrt{1 + \lambda}$ will be an eigenvalue of the matrix $I_{Kp} - (I_{Kp} + R' R)^{-1/2}$ associated with the eigenvector ρ , so that

$$\left[I_{Kp} - (I_{Kp} + R' R)^{-1/2} \right] \rho = \frac{\sqrt{1 + \lambda} - 1}{\sqrt{1 + \lambda}} \rho$$

Next, let $g(\lambda) = (\sqrt{1 + \lambda} - 1) / \sqrt{1 + \lambda}$; and, by taking derivative of $g(\lambda)$ with respect to λ , we obtain

$$g'(\lambda) = \frac{1}{2} \frac{1}{1 + \lambda} - \frac{1}{2} \frac{\sqrt{1 + \lambda} - 1}{(1 + \lambda)^{3/2}} = \frac{1}{2(1 + \lambda)^{3/2}} > 0 \text{ for all } \lambda \geq 0,$$

so that, in particular, $g(\lambda)$ is an increasing function of λ for $\lambda \geq 0$. Making use of these results, we see that

$$\begin{aligned} &\left\| \hat{G}_1 - G_1 \right\|_2 \\ &\leq \left\| I_{Kp} - (I_{Kp} + R' R)^{-1/2} \right\|_2 + \left\| R \right\|_2 \left\| (I_{Kp} + R' R)^{-1/2} \right\|_2 \end{aligned}$$

$$\begin{aligned}
&= \sqrt{\lambda_{\max} \left(\left[I_{Kp} - (I_{Kp} + R'R)^{-1/2} \right]' \left[I_{Kp} - (I_{Kp} + R'R)^{-1/2} \right] \right)} \\
&\quad + \|R\|_2 \sqrt{\lambda_{\max} \left((I_{Kp} + R'R)^{-1/2} (I_{Kp} + R'R)^{-1/2} \right)} \\
&= \lambda_{\max} \left[I_{Kp} - (I_{Kp} + R'R)^{-1/2} \right] + \|R\|_2 \lambda_{\max} \left[(I_{Kp} + R'R)^{-1/2} \right] \\
&\leq \frac{\sqrt{1 + \lambda_{\max}(R'R)} - 1}{\sqrt{1 + \lambda_{\min}(R'R)}} + \frac{\|R\|_2}{\sqrt{1 + \lambda_{\min}(R'R)}} \\
&\leq \sqrt{1 + \|R\|_2^2} - 1 + \|R\|_2 \\
&= o_p(1) \text{ as } N_1, N_2, \text{ and } T \rightarrow \infty \text{ (since } \|R\|_2 = o_p(1) \text{)}.
\end{aligned}$$

To show part (c), note that, by the result given in part (a) above, the columns of $\widehat{G}_1 = (G_1 + G_2R)(I_r + R'R)^{-1/2}$ form an orthonormal basis for a subspace that is invariant for $A + E$. It then follows from applying Theorem 3.9 on page 22 of Stewart and Sun (1990) that there exists a unique matrix L such that $(A + E)\widehat{G}_1 = (A + E)(G_1 + G_2R)(I_r + R'R)^{-1/2} = (G_1 + G_2R)(I_r + R'R)^{-1/2}L = \widehat{G}_1L$. Note further that, since by assumption $G = \begin{bmatrix} G_1 & G_2 \end{bmatrix}$ is an orthogonal matrix, we have

$$\begin{aligned}
\widehat{G}_1' \widehat{G}_1 &= (I_{Kp} + R'R)^{-1/2} (G_1' + R'G_2') (G_1 + G_2R) (I_{Kp} + R'R)^{-1/2} \\
&= (I_{Kp} + R'R)^{-1/2} (G_1'G_1 + R'G_2'G_1 + G_1'G_2R + R'G_2'G_2R) (I_{Kp} + R'R)^{-1/2} \\
&= (I_{Kp} + R'R)^{-1/2} (I_{Kp} + R'R) (I_{Kp} + R'R)^{-1/2} \\
&= I_{Kp}
\end{aligned}$$

which, in turn, implies that $\widehat{G}_1'(A + E)\widehat{G}_1 = \widehat{G}_1'([\Gamma M_{FF}\Gamma'/N_1] + E)\widehat{G}_1 = \widehat{G}_1'\widehat{G}_1L = L$, so that L must be symmetric since, in our case here, $A + E = (\Gamma M_{FF}\Gamma'/N_1) + \widehat{\Sigma}(\widehat{H}^c) - (\Gamma M_{FF}\Gamma'/N_1) = \widehat{\Sigma}(\widehat{H}^c) = Z(\widehat{H}^c)'Z(\widehat{H}^c)/(N_1T_0)$ is a symmetric matrix. Now, let $\widehat{\Lambda} = \text{diag}(\widehat{\lambda}_1, \dots, \widehat{\lambda}_{Kp})$ and $\widehat{V} = (\widehat{v}_1 \ \widehat{v}_2 \ \dots \ \widehat{v}_{Kp})$ be as defined in expressions (5) and (6). The fact that L is symmetric implies that \widehat{V} is an orthogonal matrix. In addition, further application of Theorem 3.9 of Stewart and Sun (1990) shows that $(\widehat{G}_1\widehat{v}_g, \widehat{\lambda}_g)$ is an eigenpair for the matrix $A + E$ for $g = 1, \dots, Kp$.

Finally, to show part (d), let $G = \begin{pmatrix} G_1 & G_2 \end{pmatrix}$, and note that, by assumption,

$$G'AG = \begin{pmatrix} G_1'AG_1 & G_1'AG_2 \\ G_2'AG_1 & G_2'AG_2 \end{pmatrix} = \begin{pmatrix} \Lambda_1 & 0 \\ 0 & 0 \end{pmatrix} = \Lambda,$$

where $\Lambda_1 = \text{diag}(\lambda_{1,1}, \dots, \lambda_{1,Kp})$ contains the Kp largest eigenvalues of A . Without loss of generality, we can further assume that $\lambda_{1,1}, \dots, \lambda_{1,Kp}$ are ordered, so that $\lambda_{1,j} = \lambda_{(j)}(A)$,

i.e., $\lambda_{1,j}$ is the j^{th} largest eigenvalue of A .⁵ Given that, $G'G = GG' = I_N$, we have $\begin{pmatrix} AG_1 & AG_2 \end{pmatrix} = AG = G\Lambda = \begin{pmatrix} G_1\Lambda_1 & 0 \end{pmatrix}$, from which it follows that

$$AG_1G'_1\widehat{G}_1\widehat{v}_\ell = G_1\Lambda_1G'_1\widehat{G}_1\widehat{v}_\ell, \text{ for } \ell \in \{1, \dots, Kp\}. \quad (11)$$

Now, the result of part (c) above shows $(\widehat{G}_1\widehat{v}_\ell, \widehat{\lambda}_\ell)$ to be an eigenpair of the matrix $A + E$ for $\ell \in \{1, \dots, Kp\}$, so that

$$(A + E)\widehat{G}_1\widehat{v}_\ell = \widehat{\lambda}_\ell\widehat{G}_1\widehat{v}_\ell \text{ for } \ell \in \{1, \dots, Kp\} \quad (12)$$

where $\widehat{G}_1 = (G_1 + G_2R)(I_{Kp} + R'R)^{-1/2}$. Multiplying both sides of expression (12) by $\widehat{v}'_\ell\widehat{G}'_1G_1G'_1$, we get

$$\begin{aligned} \widehat{\lambda}_\ell\widehat{v}'_\ell\widehat{G}'_1G_1G'_1\widehat{G}_1\widehat{v}_\ell &= \widehat{v}'_\ell\widehat{G}'_1G_1G'_1(A + E)\widehat{G}_1\widehat{v}_\ell \\ &= \widehat{v}'_\ell\widehat{G}'_1G_1G'_1A\widehat{G}_1\widehat{v}_\ell + \widehat{v}'_\ell\widehat{G}'_1G_1G'_1E\widehat{G}_1\widehat{v}_\ell \end{aligned} \quad (13)$$

Since $A = \Gamma M_{FF}\Gamma'/N_1$ is symmetric, it further follows by expression (11) that

$$\widehat{v}'_\ell\widehat{G}'_1G_1G'_1A = \widehat{v}'_\ell\widehat{G}'_1G_1G'_1A' = \widehat{v}'_\ell\widehat{G}'_1G_1\Lambda_1G'_1 \quad (14)$$

Moreover, note that

$$\begin{aligned} 0 &\leq \left(\widehat{v}'_\ell\widehat{G}'_1G_1G'_1E\widehat{G}_1\widehat{v}_\ell\right)^2 \\ &\leq \left(\widehat{v}'_\ell\widehat{G}'_1G_1G'_1\widehat{G}_1\widehat{v}_\ell\right)\left(\widehat{v}'_\ell\widehat{G}'_1E'E\widehat{G}_1\widehat{v}_\ell\right) \quad (\text{by CS inequality and the fact that } G'_1G_1 = I_{Kp}) \\ &= \left[\widehat{v}'_\ell(I_{Kp} + R'R)^{-\frac{1}{2}}(G'_1 + R'G'_2)G_1G'_1(G_1 + G_2R)(I_{Kp} + R'R)^{-\frac{1}{2}}\widehat{v}_\ell\right]\left(\widehat{v}'_\ell\widehat{G}'_1E'E\widehat{G}_1\widehat{v}_\ell\right) \\ &\leq \left[\widehat{v}'_\ell(I_{Kp} + R'R)^{-1}\widehat{v}_\ell\right]\lambda_{\max}(E'E) \end{aligned}$$

from which it follows that

$$\begin{aligned} -\sqrt{\widehat{v}'_\ell(I_{Kp} + R'R)^{-1}\widehat{v}_\ell}\|E\|_2 &= -\sqrt{\widehat{v}'_\ell(I_{Kp} + R'R)^{-1}\widehat{v}_\ell}\sqrt{\lambda_{\max}(E'E)} \\ &\leq -\sqrt{\left(\widehat{v}'_\ell\widehat{G}'_1G_1G'_1E\widehat{G}_1\widehat{v}_\ell\right)^2} \\ &\leq -\left|\widehat{v}'_\ell\widehat{G}'_1G_1G'_1E\widehat{G}_1\widehat{v}_\ell\right| \end{aligned}$$

⁵If this is not the case; then, we can always define a permutation matrix \mathcal{P} such that $\Lambda^* = \mathcal{P}'\Lambda\mathcal{P}$ results in a diagonal matrix whose diagonal elements are repermuted in such a way, so that the required ordering of the eigenvalues is satisfied. Moreover, since \mathcal{P} is an orthogonal matrix, it further follows that $A = G\mathcal{P}\mathcal{P}'\Lambda\mathcal{P}\mathcal{P}'G' = G\mathcal{P}\Lambda^*\mathcal{P}'G'$. Now, define $\widetilde{G} = G\mathcal{P}$, and note that \widetilde{G} is an orthogonal matrix whose columns are just the columns of G repermuted. Hence, we can simply proceed with our analysis using \widetilde{G} in lieu of G , and the associated eigenvalues will be in the order which we have assumed.

$$\leq \tilde{v}'_\ell \widehat{G}'_1 G_1 G'_1 E \widehat{G}_1 \widehat{v}_\ell. \quad (15)$$

Combining expressions (13), (14), and (15), we see that

$$\begin{aligned} \widehat{\lambda}_\ell \tilde{v}'_\ell \widehat{G}'_1 G_1 G'_1 \widehat{G}_1 \widehat{v}_\ell &= \tilde{v}'_\ell \widehat{G}'_1 G_1 G'_1 A \widehat{G}_1 \widehat{v}_\ell + \tilde{v}'_\ell \widehat{G}'_1 G_1 G'_1 E \widehat{G}_1 \widehat{v}_\ell \\ &\geq \tilde{v}'_\ell \widehat{G}'_1 G_1 \Lambda_1 G'_1 \widehat{G}_1 \widehat{v}_\ell - \sqrt{\tilde{v}'_\ell (I_{Kp} + R'R)^{-1} \widehat{v}_\ell} \|E\|_2 \end{aligned} \quad (16)$$

for $\ell \in \{1, \dots, Kp\}$. In addition, note that

$$\begin{aligned} \tilde{v}'_\ell \widehat{G}'_1 G_1 G'_1 \widehat{G}_1 \widehat{v}_\ell &= \tilde{v}'_\ell \widehat{G}'_1 G_1 G'_1 (G_1 + G_2 R) (I_{Kp} + R'R)^{-1/2} \widehat{v}_\ell \\ &= \tilde{v}'_\ell \widehat{G}'_1 G_1 (I_{Kp} + R'R)^{-1/2} \widehat{v}_\ell \\ &= \tilde{v}'_\ell (I_{Kp} + R'R)^{-1/2} (G'_1 + R'G'_2) G_1 (I_{Kp} + R'R)^{-1/2} \widehat{v}_\ell \\ &= \tilde{v}'_\ell (I_{Kp} + R'R)^{-1} \widehat{v}_\ell \\ &> 0 \end{aligned}$$

Hence, dividing both sides of expression (16) by $\tilde{v}'_\ell \widehat{G}'_1 G_1 G'_1 \widehat{G}_1 \widehat{v}_\ell$, we obtain

$$\begin{aligned} \widehat{\lambda}_\ell &\geq \frac{\tilde{v}'_\ell \widehat{G}'_1 G_1 \Lambda_1 G'_1 \widehat{G}_1 \widehat{v}_\ell}{\tilde{v}'_\ell \widehat{G}'_1 G_1 G'_1 \widehat{G}_1 \widehat{v}_\ell} - \frac{\sqrt{\tilde{v}'_\ell (I_{Kp} + R'R)^{-1} \widehat{v}_\ell} \|E\|_2}{\tilde{v}'_\ell \widehat{G}'_1 G_1 G'_1 \widehat{G}_1 \widehat{v}_\ell} \\ &= \tilde{v}'_\ell \Lambda_1 \tilde{v}_\ell - \frac{\sqrt{\tilde{v}'_\ell (I_{Kp} + R'R)^{-1} \widehat{v}_\ell} \|E\|_2}{\tilde{v}'_\ell (I_{Kp} + R'R)^{-1} \widehat{v}_\ell} \\ &= \tilde{v}'_\ell \Lambda_1 \tilde{v}_\ell - \frac{\|E\|_2}{\sqrt{\tilde{v}'_\ell (I_{Kp} + R'R)^{-1} \widehat{v}_\ell}} \\ &= \sum_{j=1}^{Kp} \tilde{v}_{\ell,j}^2 \lambda_{1,j} - \frac{\|E\|_2}{\sqrt{\tilde{v}'_\ell (I_{Kp} + R'R)^{-1} \widehat{v}_\ell}} \end{aligned}$$

where $\tilde{v}_\ell = G'_1 \widehat{G}_1 \widehat{v}_\ell / (\tilde{v}'_\ell \widehat{G}'_1 G_1 G'_1 \widehat{G}_1 \widehat{v}_\ell)$ so that $\|\tilde{v}_\ell\|_2^2 = \sum_{\ell=1}^{Kp} \tilde{v}_{\ell,j}^2 = 1$. Note also that, by applying Theorem 8.1.10 of Golub and van Loan (1996) (or GvL for short), we have

$$\begin{aligned} \tilde{v}'_\ell (I_{Kp} + R'R)^{-1} \widehat{v}_\ell &\geq \lambda_{\min} \left\{ (I_{Kp} + R'R)^{-1} \right\} \tilde{v}'_\ell \widehat{v}_\ell \\ &= \frac{1}{\lambda_{\max} (I_{Kp} + R'R)} \quad (\text{given that } \|\widehat{v}_\ell\|_2^2 = 1) \\ &\geq \frac{1}{1 + \lambda_{\max} (R'R)} = \frac{1}{1 + \|R\|_2^2} \end{aligned}$$

$$\begin{aligned}
&\geq \left[1 + \frac{16 \|E_{21}\|_2^2}{(\text{sep}(\Lambda_1, \Lambda_2))^2} \right]^{-1} \quad (\text{by Theorem 8.1.10 of GvL}) \\
&\geq \left[1 + \frac{16 \|E\|_2^2}{(\text{sep}(\Lambda_1, \Lambda_2))^2} \right]^{-1} \quad (\text{since } \|E_{21}\|_2^2 \leq \|E\|_2^2) \\
&\geq \left[1 + \frac{16 (\text{sep}(\Lambda_1, \Lambda_2))^2 / 25}{(\text{sep}(\Lambda_1, \Lambda_2))^2} \right]^{-1} = \frac{25}{41}.
\end{aligned}$$

Making use of this lower bound, we obtain $\hat{\lambda}_\ell \geq \sum_{j=1}^{Kp} \tilde{v}_{\ell,j}^2 \lambda_{1,j} - \left(\|E\|_2 / \sqrt{\hat{v}_\ell (I_{Kp} + R'R)^{-1} \hat{v}_\ell} \right)$
 $= \sum_{j=1}^{Kp} \tilde{v}_{\ell,j}^2 \lambda_{1,j} - \frac{25}{41} \|E\|_2$. Next, the ordering of the eigenvalues of the matrices $A + E$ and A be given by $\lambda_{(1)}(A + E) \geq \dots \geq \lambda_{(Kp)}(A + E) \geq \lambda_{(Kp+1)}(A + E) \geq \dots \geq \lambda_{(N)}(A + E)$ and $\lambda_{(1)}(A) \geq \dots \geq \lambda_{(Kp)}(A) \geq \lambda_{(Kp+1)}(A) \geq \dots \geq \lambda_{(N)}(A)$. Since $A = \Gamma M_{FF} \Gamma' / N_1$ and since $\text{Rank}(A) = Kp$ for all N_1, N_2 , and T sufficiently large⁶; it follows that $\lambda_{(Kp+1)}(A) = \dots = \lambda_{(N)}(A) = 0$. In addition, by Corollary 8.1.6 of Golub and van Loan (1996), we have the inequality $\lambda_{(Kp+1)}(A + E) \leq \lambda_{(Kp+1)}(A) + \|E\|_2$. Making use of these results, we see that, for any $\ell \in \{1, \dots, Kp\}$,

$$\begin{aligned}
\hat{\lambda}_\ell - \lambda_{(Kp+1)}(A + E) &\geq \sum_{j=1}^{Kp} \tilde{v}_{\ell,j}^2 \lambda_{1,j} - \frac{25}{41} \|E\|_2 - \{ \lambda_{(Kp+1)}(A) + \|E\|_2 \} \\
&= \sum_{j=1}^{Kp} \tilde{v}_{\ell,j}^2 \lambda_{(j)}(A) - \frac{66}{41} \|E\|_2 \quad (\text{since } \lambda_{(Kp+1)}(A) = 0 \text{ and } \lambda_{1,j} = \lambda_{(j)}(A)) \\
&\geq \sum_{j=1}^{Kp} \tilde{v}_{\ell,j}^2 \lambda_{(j)}(A) - \frac{66 \text{sep}(\Lambda_1, \Lambda_2)}{41 \cdot 5} \quad (\text{by Theorem 8.1.10 of GvL}) \\
&= \sum_{j=1}^{Kp} \tilde{v}_{\ell,j}^2 \lambda_{(j)}(A) - \frac{66}{205} \text{sep}(\Lambda_1, 0) \quad (\text{since } \Lambda_2 = 0 \text{ here}) \\
&\geq \lambda_{\min}(\Lambda_1) - \frac{66}{205} \text{sep}(\Lambda_1, 0) \quad (\text{since } \Lambda_1 = \text{diag}(\lambda_{(1)}(A), \dots, \lambda_{(Kp)}(A))) \\
&= \frac{139}{205} \text{sep}(\Lambda_1, 0) \\
&\geq \frac{139}{205} c > 0 \quad (\text{by expression (9)}).
\end{aligned}$$

where the last equality above follows from the fact that $\text{sep}(\Lambda_1, 0) = \lambda_{\min}(\Lambda_1)$ by Theorem

⁶An explicit proof that $\text{Rank}(A) = Kp$ for all N_1, N_2 , and T sufficiently large is given in the Technical Appendix of an earlier version of this paper, Chao and Swanson (2022b). In particular, this result is shown in part (a) of Lemma D-11 in Appendix D of Chao and Swanson (2022b).

3.1 on page 247 of Stewart and Sun (1990) since A is symmetric in this case. This shows that the set $\{\hat{\lambda}_1, \dots, \hat{\lambda}_{Kp}\}$ contains the Kp largest eigenvalues of the matrix $A + E$. It further follows from the result given in part (c) that the columns of the matrix $\hat{G}_1 \hat{V} = \begin{pmatrix} \hat{G}_1 \hat{v}_1 & \hat{G}_1 \hat{v}_2 & \dots & \hat{G}_1 \hat{v}_{Kp} \end{pmatrix}$ are the eigenvectors associated with the Kp largest eigenvalues of the matrix $A + E$. \square

Lemma OA-2: Suppose that Assumptions 2-1, 2-2, 2-3, 2-4, 2-5, 2-6, 2-7, 2-8, 2-9, OA-3, and OA-4 hold. Then, the following statements are true.

- (a) $(\hat{N}_1 - N_1) / N_1 \xrightarrow{p} 0$
- (b) $\left\| \left[\Gamma \left(\widehat{H}^c \right) - \Gamma \right] / \sqrt{N_1} \right\|_2 \xrightarrow{p} 0$
- (c) Let $\hat{G}_1 = (G_1 + G_2 R) (I_{Kp} + R' R)^{-1/2}$, where G_1 , G_2 , and R are as defined in Lemma OA-1 above. Also, let \hat{V} be the $Kp \times Kp$ orthogonal matrix given in expression (6) of Lemma OA-1. Then, there exists some positive constant \bar{C} such that $\left\| \hat{V}' \hat{G}_1' \Gamma / \sqrt{N_1} \right\|_2 \leq \sqrt{\lambda_{\max}(\Gamma' \Gamma / N_1)} \leq \bar{C} < \infty$ for N_1, N_2 , and T sufficiently large. In addition, $\left\| \hat{V}' \hat{G}_1' \Gamma / \sqrt{N_1} - Q \right\|_2 \xrightarrow{p} 0$, where $Q = (\Gamma' \Gamma / N_1)^{\frac{1}{2}} \Xi \hat{V}$, with Ξ being the $Kp \times Kp$ orthogonal matrix whose columns are the eigenvectors of the matrix $M_{FF}^* = (\Gamma' \Gamma / N_1)^{1/2} \frac{1}{T-p+1} \sum_{t=p}^T E [\underline{E}_t \underline{E}_t'] (\Gamma' \Gamma / N_1)^{1/2}$.
- (d) For all fixed index t , $\left\| G_1' U_{t,N} \left(\widehat{H}^c \right) / \sqrt{N_1} \right\|_2 = o_p(1)$.
- (e) For all fixed index t , $\left\| U_{t,N} \left(\widehat{H}^c \right) / \sqrt{N_1} \right\|_2^2 = O_p(1)$.
- (f) For all fixed index t , $\left\| G_2' U_{t,N} \left(\widehat{H}^c \right) / \sqrt{N_1} \right\|_2 = O_p(1)$.
- (g) Let $\hat{G}_1 = (G_1 + G_2 R) (I_{Kp} + R' R)^{-1/2}$, where G_1 , G_2 , and R are as defined in Lemma OA-1 above. Also, let \hat{V} be the $Kp \times Kp$ orthogonal matrix given in expression (6) of Lemma OA-1. Then, for all fixed index t , $\left\| \hat{V}' \hat{G}_1' U_{t,N} \left(\widehat{H}^c \right) / \sqrt{\hat{N}_1} \right\|_2 \xrightarrow{p} 0$ as N_1, N_2 , and $T \rightarrow \infty$.
- (h) $\|\underline{E}_t\|_2 = O_p(1)$ for all t .

Proof of Lemma OA-2:

To show part (a), note first that, for any $\epsilon > 0$,

$$\begin{aligned}
\left\{ \left| \frac{\widehat{N}_1 - N_1}{N_1} \right| \geq \epsilon \right\} &= \left\{ \left| \frac{1}{N_1} \sum_{i=1}^N \mathbb{I}\{i \in \widehat{H}^c\} - 1 \right| \geq \epsilon \right\} \\
&= \left\{ \left| \frac{1}{N_1} \sum_{i \in H^c} (\mathbb{I}\{i \in \widehat{H}^c\} - 1) + \frac{1}{N_1} \sum_{i \in H} \mathbb{I}\{i \in \widehat{H}^c\} \right| \geq \epsilon \right\} \\
&\subseteq \left\{ \left| \frac{1}{N_1} \sum_{i \in H^c} (\mathbb{I}\{i \in \widehat{H}^c\} - 1) \right| + \left| \frac{1}{N_1} \sum_{i \in H} \mathbb{I}\{i \in \widehat{H}^c\} \right| \geq \epsilon \right\} \\
&\subseteq \left\{ \left| \frac{1}{N_1} \sum_{i \in H^c} (\mathbb{I}\{i \in \widehat{H}^c\} - 1) \right| \geq \frac{\epsilon}{2} \right\} \cup \left\{ \left| \frac{1}{N_1} \sum_{i \in H} \mathbb{I}\{i \in \widehat{H}^c\} \right| \geq \frac{\epsilon}{2} \right\}
\end{aligned}$$

By Markov's inequality, we have

$$\begin{aligned}
&\Pr \left(\left| \frac{1}{N_1} \sum_{i \in H^c} (\mathbb{I}\{i \in \widehat{H}^c\} - 1) \right| \geq \frac{\epsilon}{2} \right) \\
&= \Pr \left(\left| \frac{1}{N_1} \sum_{i \in H^c} (\mathbb{I}\{i \in \widehat{H}^c\} - 1) \right|^2 \geq \frac{\epsilon^2}{4} \right) \\
&\leq \frac{4}{\epsilon^2} E \left\{ \left| \frac{1}{N_1} \sum_{i \in H^c} (\mathbb{I}\{i \in \widehat{H}^c\} - 1) \right|^2 \right\} \\
&= \frac{4}{\epsilon^2} \frac{1}{N_1^2} \sum_{i \in H^c} \sum_{k \in H^c} E \left[(\mathbb{I}\{i \in \widehat{H}^c\} - 1) (\mathbb{I}\{k \in \widehat{H}^c\} - 1) \right] \\
&= \frac{4}{\epsilon^2} \frac{1}{N_1^2} \sum_{i \in H^c} \sum_{k \in H^c} \left\{ \Pr \left(\{i \in \widehat{H}^c\} \cap \{k \in \widehat{H}^c\} \right) - \Pr \left(k \in \widehat{H}^c \right) \right\} \\
&\quad + \frac{4}{\epsilon^2} \frac{1}{N_1^2} \sum_{i \in H^c} \sum_{k \in H^c} \left\{ 1 - \Pr \left(i \in \widehat{H}^c \right) \right\} \\
&\leq \frac{4}{\epsilon^2} \frac{1}{N_1^2} \sum_{i \in H^c} \sum_{k \in H^c} \left\{ \Pr \left(k \in \widehat{H}^c \right) - \Pr \left(k \in \widehat{H}^c \right) \right\} + \frac{4}{\epsilon^2} \frac{1}{N_1} \sum_{i \in H^c} \left\{ 1 - \Pr \left(i \in \widehat{H}^c \right) \right\} \\
&\leq \frac{4}{\epsilon^2} \frac{1}{N_1} \sum_{i \in H^c} \left\{ 1 - \min_{i \in H^c} \Pr \left(i \in \widehat{H}^c \right) \right\} \rightarrow 0 \text{ as } N_1, N_2, \text{ and } T \rightarrow \infty.
\end{aligned}$$

where the last line above follows from the fact that, for $i \in H^c$ and for either the case where $\mathbb{S}_{i,T}^+ = \sum_{\ell=1}^d \varpi_\ell |S_{i,\ell,T}|$ or the case where $\mathbb{S}_{i,T}^+ = \max_{1 \leq \ell \leq d} |S_{i,\ell,T}|$, we can apply the results of

Theorem 2 of the main paper to obtain $\min_{i \in H^c} \Pr \left(i \in \widehat{H}^c \right) \geq \Pr \left(\bigcap_{i \in H^c} \{ \mathbb{S}_{i,T}^+ \geq \Phi^{-1} (1 - \frac{\varphi}{2N}) \} \right)$

$= P\left(\min_{i \in H^c} \mathbb{S}_{i,T}^+ \geq \Phi^{-1}\left(1 - \frac{\varphi}{2N}\right)\right) \rightarrow 1$. Also, making use of Markov's inequality, we obtain, for either the case where $\mathbb{S}_{i,T}^+ = \sum_{\ell=1}^d \varpi_{\ell} |S_{i,\ell,T}|$ or the case where $\mathbb{S}_{i,T}^+ = \max_{1 \leq \ell \leq d} |S_{i,\ell,T}|$,

$$\begin{aligned}
\Pr\left(\left|\frac{1}{N_1} \sum_{i \in H} \mathbb{I}\{i \in \widehat{H}^c\}\right| \geq \frac{\epsilon}{2}\right) &= \Pr\left(\frac{1}{N_1} \sum_{i \in H} \mathbb{I}\{i \in \widehat{H}^c\} \geq \frac{\epsilon}{2}\right) \\
&\leq \frac{2}{\epsilon} E\left[\frac{1}{N_1} \sum_{i \in H} \mathbb{I}\{i \in \widehat{H}^c\}\right] \\
&= \frac{2}{\epsilon} \frac{1}{N_1} \sum_{i \in H} \Pr(i \in \widehat{H}^c) \\
&= \frac{2}{\epsilon} \frac{1}{N_1} \sum_{i \in H} \Pr\left(\mathbb{S}_{i,T}^+ \geq \Phi^{-1}\left(1 - \frac{\varphi}{2N}\right)\right) \\
&\leq \frac{2}{\epsilon} \frac{dN_2\varphi}{NN_1} [1 + o(1)] \\
&\rightarrow 0 \quad \left(\text{since } \frac{\varphi}{N_1} \rightarrow 0 \text{ and } \frac{N_2}{N} = O(1)\right)
\end{aligned}$$

where the last inequality above follows by an argument similar to that given in the proof of Theorem 1 of the main paper. Combining these results, we have that

$$\begin{aligned}
&\Pr\left(\left|\frac{\widehat{N}_1 - N_1}{N_1}\right| \geq \epsilon\right) \\
&\leq \Pr\left(\left\{\left|\frac{1}{N_1} \sum_{i \in H^c} (\mathbb{I}\{i \in \widehat{H}^c\} - 1)\right| \geq \frac{\epsilon}{2}\right\} \cup \left\{\left|\frac{1}{N_1} \sum_{i \in H} \mathbb{I}\{i \in \widehat{H}^c\}\right| \geq \frac{\epsilon}{2}\right\}\right) \\
&\leq \Pr\left(\left|\frac{1}{N_1} \sum_{i \in H^c} (\mathbb{I}\{i \in \widehat{H}^c\} - 1)\right| \geq \frac{\epsilon}{2}\right) + \Pr\left(\left|\frac{1}{N_1} \sum_{i \in H} \mathbb{I}\{i \in \widehat{H}^c\}\right| \geq \frac{\epsilon}{2}\right) \\
&\rightarrow 0
\end{aligned}$$

For part (b), note that

$$\begin{aligned}
\left\|\frac{\Gamma(\widehat{H}^c) - \Gamma}{\sqrt{N_1}}\right\|_F^2 &= \frac{1}{N_1} \text{tr} \left\{ \left(\Gamma(\widehat{H}^c) - \Gamma\right)' \left(\Gamma(\widehat{H}^c) - \Gamma\right) \right\} \\
&= \frac{1}{N_1} \sum_{i=1}^N \text{tr} \left\{ \left(\mathbb{I}\{i \in \widehat{H}^c\} \gamma_i - \gamma_i\right) \left(\mathbb{I}\{i \in \widehat{H}^c\} \gamma_i - \gamma_i\right)' \right\} \\
&= \frac{1}{N_1} \sum_{i=1}^N \left(\mathbb{I}\{i \in \widehat{H}^c\} \gamma_i - \gamma_i\right)' \left(\mathbb{I}\{i \in \widehat{H}^c\} \gamma_i - \gamma_i\right)
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{N_1} \sum_{i=1}^N \gamma'_i \gamma_i \left[1 - \mathbb{I} \left\{ i \in \widehat{H}^c \right\} \right] \\
&= \frac{1}{N_1} \sum_{i \in H^c} \gamma'_i \gamma_i \left[1 - \mathbb{I} \left\{ i \in \widehat{H}^c \right\} \right] \quad (\text{since } \gamma_i = 0 \text{ for } i \in H)
\end{aligned}$$

Applying Markov's inequality, we have, for any $\epsilon > 0$,

$$\begin{aligned}
\Pr \left(\left\| \frac{\Gamma(\widehat{H}^c) - \Gamma}{\sqrt{N_1}} \right\|_F^2 \geq \epsilon \right) &\leq \frac{1}{\epsilon} E \left\{ \frac{1}{N_1} \sum_{i \in H^c} \gamma'_i \gamma_i \left[1 - \mathbb{I} \left\{ i \in \widehat{H}^c \right\} \right] \right\} \\
&= \frac{1}{\epsilon} \frac{1}{N_1} \sum_{i \in H^c} \gamma'_i \gamma_i \left[1 - \Pr \left(i \in \widehat{H}^c \right) \right] \\
&\leq \frac{1}{\epsilon} \left[1 - \min_{i \in H^c} \Pr \left(i \in \widehat{H}^c \right) \right] \frac{1}{N_1} \sum_{i \in H^c} \gamma'_i \gamma_i \\
&\leq \frac{1}{\epsilon} \left[1 - \min_{i \in H^c} \Pr \left(i \in \widehat{H}^c \right) \right] \left(\sup_{i \in H^c} \|\gamma_i\|_2 \right)^2 \\
&\leq \frac{1}{\epsilon} \left[1 - \min_{i \in H^c} \Pr \left(i \in \widehat{H}^c \right) \right] \overline{C}^2 \quad (\text{by Assumption 2-5}) \\
&\rightarrow 0
\end{aligned}$$

where the last line follows from the fact that $\min_{i \in H^c} \Pr \left(i \in \widehat{H}^c \right) \rightarrow 1$ by Theorem 2 of the main paper. Using this result, we further deduce that $\left\| \left(\Gamma(\widehat{H}^c) - \Gamma \right) / \sqrt{N_1} \right\|_2 \leq \left\| \left(\Gamma(\widehat{H}^c) - \Gamma \right) / \sqrt{N_1} \right\|_F \xrightarrow{p} 0$.

Turning our attention to part (c), note that since, by definition, $\widehat{G}_1 = (G_1 + G_2 R) (I_{Kp} + R' R)^{-1/2}$, where $G'_1 G_1 = I_{Kp}$, $G'_2 G_2 = I_{N-Kp}$, and $G'_1 G_2 = 0$; it follows that $\widehat{G}'_1 \widehat{G}_1 = (I_{Kp} + R' R)^{-1/2} (G'_1 + R' G'_2) (G_1 + G_2 R) (I_{Kp} + R' R)^{-1/2} = (I_{Kp} + R' R)^{-1/2} (I_{Kp} + R' R) (I_{Kp} + R' R)^{-1/2} = I_{Kp}$. Hence, by Assumption OA-3,

$$\begin{aligned}
\left\| \frac{\widehat{V}' \widehat{G}'_1 \Gamma}{\sqrt{N_1}} \right\|_2 &\leq \left\| \widehat{V}' \widehat{G}'_1 \right\|_2 \left\| \frac{\Gamma}{\sqrt{N_1}} \right\|_2 = \sqrt{\lambda_{\max} \left(\widehat{G}_1 \widehat{V} \widehat{V}' \widehat{G}'_1 \right)} \sqrt{\lambda_{\max} \left(\frac{\Gamma' \Gamma}{N_1} \right)} \\
&= \sqrt{\lambda_{\max} \left(\widehat{V}' \widehat{G}'_1 \widehat{G}_1 \widehat{V} \right)} \sqrt{\lambda_{\max} \left(\frac{\Gamma' \Gamma}{N_1} \right)} \\
&= \sqrt{\lambda_{\max} (I_{Kp})} \sqrt{\lambda_{\max} \left(\frac{\Gamma' \Gamma}{N_1} \right)} \quad (\text{since } \widehat{V} \text{ is an orthogonal matrix})
\end{aligned}$$

$$= \sqrt{\lambda_{\max} \left(\frac{\Gamma' \Gamma}{N_1} \right)} \leq \bar{C} < \infty \text{ for } N_1, N_2 \text{ sufficiently large}$$

Now, to show the second result in part (c), note that, since $Q = (\Gamma' \Gamma / N_1)^{\frac{1}{2}} \Xi \hat{V}$ and $G_1 = (\Gamma / \sqrt{N_1}) (\Gamma' \Gamma / N_1)^{-1/2} \Xi = \Gamma (\Gamma' \Gamma)^{-1/2} \Xi$, we can write

$$\begin{aligned} \frac{\hat{V}' \hat{G}'_1 \Gamma}{\sqrt{N_1}} - Q' &= \frac{\hat{V}' \hat{G}'_1 \Gamma}{\sqrt{N_1}} - \hat{V}' \Xi' \left(\frac{\Gamma' \Gamma}{N_1} \right)^{\frac{1}{2}} = \frac{\hat{V}' \hat{G}'_1 \Gamma}{\sqrt{N_1}} - \hat{V}' \Xi' \left(\frac{\Gamma' \Gamma}{N_1} \right)^{-1/2} \frac{\Gamma' \Gamma}{N_1} \\ &= \frac{\hat{V}' \hat{G}'_1 \Gamma}{\sqrt{N_1}} - \frac{\hat{V}' G'_1 \Gamma}{\sqrt{N_1}} = \hat{V}' (\hat{G}_1 - G_1)' \frac{\Gamma}{\sqrt{N_1}} \end{aligned}$$

from which it follows that

$$\begin{aligned} \left\| \frac{\hat{V}' \hat{G}'_1 \Gamma}{\sqrt{N_1}} - Q' \right\|_2 &\leq \left\| \hat{V}' \right\|_2 \left\| (\hat{G}_1 - G_1)' \right\|_2 \left\| \frac{\Gamma}{\sqrt{N_1}} \right\|_2 \\ &= \sqrt{\lambda_{\max} (\hat{V} \hat{V}')} \sqrt{\lambda_{\max} \left\{ (\hat{G}_1 - G_1) (\hat{G}_1 - G_1)' \right\}} \sqrt{\lambda_{\max} \left(\frac{\Gamma' \Gamma}{N_1} \right)} \\ &= \sqrt{\lambda_{\max} (I_{Kp})} \sqrt{\lambda_{\max} \left\{ (\hat{G}_1 - G_1)' (\hat{G}_1 - G_1) \right\}} \sqrt{\lambda_{\max} \left(\frac{\Gamma' \Gamma}{N_1} \right)} \\ &\leq \sqrt{\bar{C}} \left\| \hat{G}_1 - G_1 \right\|_2 \quad (\text{by Assumption OA-3}) \\ &= o_p(1) \text{ as } N_1, N_2, \text{ and } T \rightarrow \infty \quad (\text{by part (b) of Lemma OA-1}). \end{aligned}$$

Next, to show part (d), we first write

$$\begin{aligned} &\left\| \frac{G'_1 U_{t,N}(\hat{H}^c)}{\sqrt{N_1}} \right\|_2^2 \\ &= \sum_{k=1}^{Kp} \left(\sum_{i \in H^c} \mathbb{I}\{i \in \hat{H}^c\} \frac{g_{1,ik} u_{i,t}}{\sqrt{N_1}} + \sum_{i \in H} \mathbb{I}\{i \in \hat{H}^c\} \frac{g_{1,ik} u_{i,t}}{\sqrt{N_1}} \right)^2 \\ &\leq 2 \sum_{k=1}^{Kp} \left(\sum_{i \in H^c} \mathbb{I}\{i \in \hat{H}^c\} \frac{g_{1,ik} u_{i,t}}{\sqrt{N_1}} \right)^2 + 2 \sum_{k=1}^{Kp} \left(\sum_{i \in H} \mathbb{I}\{i \in \hat{H}^c\} \frac{g_{1,ik} u_{i,t}}{\sqrt{N_1}} \right)^2 \quad (17) \end{aligned}$$

where $g_{1,ik}$ denotes the $(i, k)^{th}$ element of G_1 . Now, consider the first term on the right-hand side of expression (17). Write

$$2 \sum_{k=1}^{Kp} \left(\sum_{i \in H^c} \mathbb{I}\{i \in \hat{H}^c\} \frac{g_{1,ik} u_{i,t}}{\sqrt{N_1}} \right)^2$$

$$\begin{aligned}
&= \frac{2}{N_1} \sum_{k=1}^{Kp} \left(\sum_{i \in H^c} \left[\mathbb{I}\{i \in \widehat{H}^c\} - 1 + 1 \right] g_{1,ik} u_{i,t} \right)^2 \\
&= \frac{2}{N_1} \sum_{k=1}^{Kp} \left(\sum_{i \in H^c} \left(\mathbb{I}\{i \in \widehat{H}^c\} - 1 \right) g_{1,ik} u_{i,t} \right)^2 + \frac{2}{N_1} \sum_{k=1}^{Kp} \left(\sum_{i \in H^c} g_{1,ik} u_{i,t} \right)^2 \\
&\quad + \frac{2}{N_1} \sum_{k=1}^{Kp} \sum_{i \in H^c} \left(\mathbb{I}\{i \in \widehat{H}^c\} - 1 \right) g_{1,ik} u_{i,t} \sum_{j \in H^c} g_{1,jk} u_{j,t} \\
&\quad + \frac{2}{N_1} \sum_{k=1}^{Kp} \sum_{i \in H^c} g_{1,ik} u_{i,t} \sum_{j \in H^c} \left(\mathbb{I}\{j \in \widehat{H}^c\} - 1 \right) g_{1,jk} u_{j,t} \\
&= \mathcal{E}_{1,1,t} + \mathcal{E}_{1,2,t} + \mathcal{E}_{1,3,t} + \mathcal{E}_{1,4,t}, \quad (\text{say}).
\end{aligned}$$

Focusing first on the term $\mathcal{E}_{1,1,t}$, we have

$$\begin{aligned}
\mathcal{E}_{1,1,t} &= \frac{2}{N_1} \sum_{k=1}^{Kp} \left(\sum_{i \in H^c} \left(\mathbb{I}\{i \in \widehat{H}^c\} - 1 \right) g_{1,ik} u_{i,t} \right)^2 \\
&= \frac{2}{N_1} \sum_{k=1}^{Kp} \left(\left| \sum_{i \in H^c} \left(\mathbb{I}\{i \in \widehat{H}^c\} - 1 \right) g_{1,ik} u_{i,t} \right| \right)^2 \\
&\leq 2 \sum_{k=1}^{Kp} \left(\frac{1}{N_1} \sum_{i \in H^c} \left(\mathbb{I}\{i \in \widehat{H}^c\} - 1 \right)^2 \right) \left(\sum_{i \in H^c} g_{1,ik}^2 u_{i,t}^2 \right) \quad (\text{by CS inequality}) \\
&= 2 \sum_{k=1}^{Kp} \left[\frac{1}{N_1} \sum_{i \in H^c} \left(1 - \mathbb{I}\{i \in \widehat{H}^c\} \right) \right] \left(\sum_{i \in H^c} g_{1,ik}^2 u_{i,t}^2 \right)
\end{aligned}$$

Now, for either the case where $\mathbb{S}_{i,T}^+ = \sum_{\ell=1}^d \varpi_\ell |S_{i,\ell,T}|$ or the case where $\mathbb{S}_{i,T}^+ = \max_{1 \leq \ell \leq d} |S_{i,\ell,T}|$, we have

$$\begin{aligned}
0 &\leq E \left[\frac{1}{N_1} \sum_{i \in H^c} \left(1 - \mathbb{I}\{i \in \widehat{H}^c\} \right) \right] = \frac{1}{N_1} \sum_{i \in H^c} \left[1 - \Pr(i \in \widehat{H}^c) \right] \\
&= \frac{1}{N_1} \sum_{i \in H^c} \left[1 - P \left(\mathbb{S}_{i,T}^+ \geq \Phi^{-1} \left(1 - \frac{\varphi}{2N} \right) \right) \right] \\
&\leq 1 - P \left(\min_{i \in H^c} \mathbb{S}_{i,T}^+ \geq \Phi^{-1} \left(1 - \frac{\varphi}{2N} \right) \right) \quad (\text{given that } N_1 = \#\{H^c\}) \\
&\rightarrow 0 \quad \left(\text{since } P \left(\min_{i \in H^c} \mathbb{S}_{i,T}^+ \geq \Phi^{-1} \left(1 - \frac{\varphi}{2N} \right) \right) \rightarrow 1 \text{ by Theorem 2 of the main paper} \right).
\end{aligned}$$

Moreover, by part (b) of Assumption 2-3, we have $E \left[\sum_{i \in H^c} g_{1,ik}^2 u_{i,t}^2 \right] = \sum_{i \in H^c} g_{1,ik}^2 E[u_{i,t}^2] \leq$

$C \sum_{i=1}^N g_{1,ik}^2 \leq C$. It follows by Markov's inequality that $\frac{1}{N_1} \sum_{i \in H^c} \left(1 - \mathbb{I}\{i \in \widehat{H}^c\}\right) = o_p(1)$ and $\sum_{i \in H^c} g_{1,ik}^2 u_{i,t}^2 = O_p(1)$ from which we deduce that $\mathcal{E}_{1,1,t} = 2N_1^{-1} \sum_{k=1}^{Kp} \left(\sum_{i \in H^c} \left(\mathbb{I}\{i \in \widehat{H}^c\} - 1 \right) g_{1,ik} u_{i,t} \right)^2 \leq 2 \sum_{k=1}^{Kp} \left[N_1^{-1} \sum_{i \in H^c} \left(1 - \mathbb{I}\{i \in \widehat{H}^c\}\right) \right] \left(\sum_{i \in H^c} g_{1,ik}^2 u_{i,t}^2 \right) = o_p(1)$. Consider next the term $\mathcal{E}_{1,2,t}$. To proceed, let $U_{t,N}(H^c)$ denote an $N \times 1$ vector whose i^{th} component $U_{i,t,N}(H^c)$ is given by $U_{i,t,N}(H^c) = u_{i,t} \mathbb{I}\{i \in H^c\}$, and we can write

$$\begin{aligned} \mathcal{E}_{1,2,t} &= \frac{2}{N_1} \sum_{k=1}^{Kp} \left(\sum_{i \in H^c} g_{1,ik} u_{i,t} \right)^2 = 2 \left\| \frac{G_1' U_{t,N}(H^c)}{\sqrt{N_1}} \right\|_2^2 \\ &\leq 2tr \left\{ \frac{G_1' U_{t,N}(H^c) U_{t,N}(H^c)' G_1}{N_1} \right\} \\ &= 2tr \left\{ \Xi' \left(\frac{\Gamma' \Gamma}{N_1} \right)^{-1/2} \frac{\Gamma'}{\sqrt{N_1}} \frac{U_{t,N}(H^c) U_{t,N}(H^c)'}{N_1} \frac{\Gamma}{\sqrt{N_1}} \left(\frac{\Gamma' \Gamma}{N_1} \right)^{-1/2} \Xi \right\} \\ &= 2tr \left\{ \left(\frac{\Gamma' \Gamma}{N_1} \right)^{-1/2} \frac{\Gamma'}{\sqrt{N_1}} \frac{U_{t,N}(H^c) U_{t,N}(H^c)'}{N_1} \frac{\Gamma}{\sqrt{N_1}} \left(\frac{\Gamma' \Gamma}{N_1} \right)^{-1/2} \right\} \\ &= 2tr \left\{ \frac{\Gamma_*' U_{t,N}(H^c) U_{t,N}(H^c)' \Gamma_*}{N_1^2} \right\} \left(\text{letting } \Gamma_* = \Gamma \left(\frac{\Gamma' \Gamma}{N_1} \right)^{-1/2} \right) \\ &= \frac{2}{N_1^2} U_{t,N}(H^c)' \Gamma_* \Gamma_*' U_{t,N}(H^c) = \frac{2}{N_1^2} \sum_{i \in H^c} \sum_{j \in H^c} \gamma'_{*,i} \gamma_{*,j} u_{i,t} u_{j,t} \end{aligned}$$

where $\gamma'_{*,i}$ denotes the i^{th} row of $\Gamma_* = \Gamma (\Gamma' \Gamma / N_1)^{-1/2}$. Hence, by Assumptions 2-5, OA-2, and OA-3, there exist positive constants \bar{c} , \underline{C} , and \bar{C} such that

$$\begin{aligned} 0 &\leq E[\mathcal{E}_{1,2,t}] \\ &= \frac{2}{N_1} \sum_{k=1}^{Kp} \sum_{i \in H^c} \sum_{j \in H^c} g_{1,ik} g_{1,jk} E[u_{i,t} u_{j,t}] \\ &= \frac{2}{N_1^2} \sum_{i \in H^c} \sum_{j \in H^c} \gamma'_{*,i} \gamma_{*,j} E[u_{i,t} u_{j,t}] = \frac{2}{N_1^2} \sum_{i \in H^c} \sum_{j \in H^c} \gamma'_i \left(\frac{\Gamma' \Gamma}{N_1} \right)^{-1/2} \left(\frac{\Gamma' \Gamma}{N_1} \right)^{-1/2} \gamma_j E[u_{i,t} u_{j,t}] \\ &\leq \frac{2}{N_1^2} \sum_{i \in H^c} \sum_{j \in H^c} \left| \gamma'_i \left(\frac{\Gamma' \Gamma}{N_1} \right)^{-1/2} \left(\frac{\Gamma' \Gamma}{N_1} \right)^{-1/2} \gamma_j \right| |E[u_{i,t} u_{j,t}]| \\ &\leq \frac{2}{N_1^2} \sum_{i \in H^c} \sum_{j \in H^c} \sqrt{\gamma'_i \left(\frac{\Gamma' \Gamma}{N_1} \right)^{-1} \gamma_i} \sqrt{\gamma'_j \left(\frac{\Gamma' \Gamma}{N_1} \right)^{-1} \gamma_j} |E[u_{i,t} u_{j,t}]| \end{aligned}$$

$$\begin{aligned}
&\leq 2\bar{C} \frac{1}{N_1^2} \sum_{i \in H^c} \sum_{j \in H^c} |E[u_{i,t} u_{j,t}]| \quad (\text{by Assumption 2-5 and OA-3}) \\
&\leq 2\bar{C} \frac{C}{N_1} \rightarrow 0 \text{ as } N_1 \rightarrow \infty \quad (\text{by Assumption OA-2}).
\end{aligned}$$

It follows by Markov's inequality that $\mathcal{E}_{1,2,t} = o_p(1)$. Now, for $\mathcal{E}_{1,3,t}$, note that, by applying the triangle and CS inequalities, we have

$$\begin{aligned}
|\mathcal{E}_{1,3,t}| &\leq 2N_1^{-1} \sum_{k=1}^{Kp} \left| \sum_{i \in H^c} \left(\mathbb{I}\{i \in \widehat{H}^c\} - 1 \right) g_{1,ik} u_{i,t} \sum_{j \in H^c} g_{1,jk} u_{j,t} \right| \\
&\leq \sqrt{2N_1^{-1} \sum_{k=1}^{Kp} \left[\sum_{i \in H^c} \left(\mathbb{I}\{i \in \widehat{H}^c\} - 1 \right) g_{1,ik} u_{i,t} \right]^2} \sqrt{2N_1^{-1} \sum_{k=1}^{Kp} \left[\sum_{j \in H^c} g_{1,jk} u_{j,t} \right]^2} \\
&= \sqrt{\mathcal{E}_{1,1,t}} \sqrt{\mathcal{E}_{1,2,t}} = o_p(1). \text{ Similarly, we also have} \\
|\mathcal{E}_{1,4,t}| &\leq 2N_1^{-1} \sum_{k=1}^{Kp} \left| \sum_{i \in H^c} g_{1,ik} u_{i,t} \sum_{j \in H^c} \left(\mathbb{I}\{j \in \widehat{H}^c\} - 1 \right) g_{1,jk} u_{j,t} \right| \\
&\leq \sqrt{\mathcal{E}_{1,2,t}} \sqrt{\mathcal{E}_{1,1,t}} = o_p(1).
\end{aligned}$$

Application of the Slutsky's theorem then allows us to deduce that

$$\frac{2}{N_1} \sum_{k=1}^{Kp} \sum_{i \in H^c} \sum_{j \in H^c} \mathbb{I}\{i \in \widehat{H}^c\} \mathbb{I}\{j \in \widehat{H}^c\} g_{1,ik} g_{1,jk} u_{i,t} u_{j,t} = \mathcal{E}_{1,1,t} + \mathcal{E}_{1,2,t} + \mathcal{E}_{1,3,t} + \mathcal{E}_{1,4,t} = o_p(1).$$

Next, consider the second term on the right-hand side of expression (17). In this case, write

$$\begin{aligned}
&\frac{2}{N_1} \sum_{k=1}^{Kp} \sum_{i \in H} \sum_{j \in H} \mathbb{I}\{i \in \widehat{H}^c\} \mathbb{I}\{j \in \widehat{H}^c\} g_{1,ik} g_{1,jk} u_{i,t} u_{j,t} \\
&= \frac{2}{N_1} \sum_{k=1}^{Kp} \left(\sum_{i \in H} \mathbb{I}\{i \in \widehat{H}^c\} g_{1,ik} u_{i,t} \right)^2 = \frac{2}{N_1} \sum_{k=1}^{Kp} \left(\left| \sum_{i \in H} \mathbb{I}\{i \in \widehat{H}^c\} g_{1,ik} u_{i,t} \right| \right)^2 \\
&\leq 2 \sum_{k=1}^{Kp} \left[\frac{1}{N_1} \sum_{i \in H} \mathbb{I}\{i \in \widehat{H}^c\} \right] \left[\sum_{i \in H} g_{1,ik}^2 u_{i,t}^2 \right]
\end{aligned}$$

Note that, for either the case where $\mathbb{S}_{i,T}^+ = \sum_{\ell=1}^d \varpi_\ell |S_{i,\ell,T}|$ or the case where $\mathbb{S}_{i,T}^+ = \max_{1 \leq \ell \leq d} |S_{i,\ell,T}|$, we have, by applying an argument similar to that given in the proof of Theorem 1 of the main paper,

$$\begin{aligned}
0 &\leq E \left[\frac{1}{N_1} \sum_{i \in H} \mathbb{I}\{i \in \widehat{H}^c\} \right] = \frac{1}{N_1} \sum_{i \in H} \Pr(i \in \widehat{H}^c) = \frac{1}{N_1} \sum_{i \in H} P(\mathbb{S}_{i,T}^+ \geq \Phi^{-1}(1 - \frac{\varphi}{2N})) \\
&\leq \frac{dN_2\varphi}{NN_1} \left\{ 1 + 2^2 A \left[1 + \left(\Phi^{-1}\left(1 - \frac{\varphi}{2N}\right) \right)^3 \right] T^{-(1-\alpha_1)\frac{1}{2}} \right\} = \frac{dN_2\varphi}{N_1(N_1 + N_2)} [1 + o(1)] \rightarrow 0,
\end{aligned}$$

as $N_1, N_2, T \rightarrow \infty$. Moreover, making use of part (b) of Assumption 2-3, we have

$E \left[\sum_{i \in H} g_{1,ik}^2 u_{i,t}^2 \right] = \sum_{i \in H} g_{1,ik}^2 E[u_{i,t}^2] \leq C \sum_{i=1}^N g_{1,ik}^2 \leq C$. It follows by Markov's inequality that

$$\begin{aligned} N_1^{-1} \sum_{i \in H} \mathbb{I} \{i \in \widehat{H}^c\} &= o_p(1) \text{ and } \sum_{i \in H} g_{1,ik}^2 u_{i,t}^2 = O_p(1), \text{ from which we deduce that} \\ 2N^{-1} \sum_{k=1}^{Kp} \sum_{i \in H} \sum_{j \in H} \mathbb{I} \{i \in \widehat{H}^c\} \mathbb{I} \{j \in \widehat{H}^c\} g_{1,ik} g_{1,jk} u_{i,t} u_{j,t} \\ &\leq 2 \sum_{k=1}^{Kp} \left[N^{-1} \sum_{i \in H} \mathbb{I} \{i \in \widehat{H}^c\} \right] \left[\sum_{i \in H} g_{1,ik}^2 u_{i,t}^2 \right] = o_p(1). \end{aligned}$$

Combining these results and using the inequality $\sqrt{a_1 + a_2} \leq \sqrt{a_1} + \sqrt{a_2}$ for $a_1 \geq 0$ and $a_2 \geq 0$, we further obtain, for all t ,

$$\begin{aligned} \left\| \frac{G'_1 U_{t,N}(\widehat{H}^c)}{\sqrt{N_1}} \right\|_2 &\leq \sqrt{\frac{2}{N_1} \sum_{k=1}^K \sum_{i \in H^c} \sum_{j \in H^c} \mathbb{I} \{i \in \widehat{H}^c\} \mathbb{I} \{j \in \widehat{H}^c\} g_{1,ik} g_{1,jk} u_{i,t} u_{j,t}} \\ &\quad + \sqrt{\frac{2}{N_1} \sum_{k=1}^K \sum_{i \in H} \sum_{j \in H} \mathbb{I} \{i \in \widehat{H}^c\} \mathbb{I} \{j \in \widehat{H}^c\} g_{1,ik} g_{1,jk} u_{i,t} u_{j,t}} \\ &= o_p(1). \end{aligned}$$

For part (e), write

$$\begin{aligned} \left\| U_{t,N}(\widehat{H}^c) / \sqrt{N_1} \right\|_2^2 &= N_1^{-1} \sum_{i=1}^N \mathbb{I} \{i \in \widehat{H}^c\} u_{i,t}^2 \\ &= N_1^{-1} \sum_{i \in H^c} \mathbb{I} \{i \in \widehat{H}^c\} u_{i,t}^2 + N_1^{-1} \sum_{i \in H} \mathbb{I} \{i \in \widehat{H}^c\} u_{i,t}^2 \\ &\leq N_1^{-1} \sum_{i \in H^c} u_{i,t}^2 + N_1^{-1} \sum_{i \in H} \mathbb{I} \{i \in \widehat{H}^c\} u_{i,t}^2 \end{aligned}$$

Note that, by Assumption 2-3(b), $E \left[N_1^{-1} \sum_{i \in H^c} u_{i,t}^2 \right] = N_1^{-1} \sum_{i \in H^c} E[u_{i,t}^2] \leq C$

(since $N_1 = \# \{H^c\}$), so that, by applying Markov's inequality, we obtain $N_1^{-1} \sum_{i \in H^c} u_{i,t}^2 =$

$O_p(1)$. Moreover, note that, for any $\epsilon > 0$, $\bigcap_{i \in H} \{i \notin \widehat{H}^c\} \subseteq \left\{ N_1^{-1} \sum_{i \in H} \mathbb{I} \{i \in \widehat{H}^c\} u_{i,t}^2 < \epsilon \right\}$,

so that by DeMorgan's law $\left\{ N_1^{-1} \sum_{i \in H} \mathbb{I} \{i \in \widehat{H}^c\} u_{i,t}^2 \geq \epsilon \right\} \subseteq \left\{ \bigcap_{i \in H} \{i \notin \widehat{H}^c\} \right\}^c$

$= \bigcup_{i \in H} \{i \in \widehat{H}^c\}$. Hence, for either the case where $\mathbb{S}_{i,T}^+ = \sum_{\ell=1}^d \varpi_\ell |S_{i,\ell,T}|$ or the case where $\mathbb{S}_{i,T}^+ = \max_{1 \leq \ell \leq d} |S_{i,\ell,T}|$, we have, by applying an argument similar to that given in the proof of Theorem 1 of the main paper,

$$\Pr \left\{ \frac{1}{N_1} \sum_{i \in H} \mathbb{I} \{i \in \widehat{H}^c\} u_{i,t}^2 \geq \epsilon \right\} \leq \Pr \left\{ \bigcup_{i \in H} \{i \in \widehat{H}^c\} \right\}$$

$$\begin{aligned}
&\leq \sum_{i \in H} \Pr \left\{ i \in \widehat{H}^c \right\} = \sum_{i \in H} P \left(\mathbb{S}_{i,T}^+ \geq \Phi^{-1} \left(1 - \frac{\varphi}{2N} \right) \right) \\
&\leq \frac{dN_2\varphi}{N} \left\{ 1 + 2^2 A \left[1 + \left(\Phi^{-1} \left(1 - \frac{\varphi}{2N} \right) \right)^3 \right] T^{-(1-\alpha_1)\frac{1}{2}} \right\} \\
&= \frac{dN_2\varphi}{N_1 + N_2} [1 + o(1)] \rightarrow 0 \text{ as } N_1, N_2, T \rightarrow \infty
\end{aligned}$$

Hence, $N_1^{-1} \sum_{i \in H} \mathbb{I} \left\{ i \in \widehat{H}^c \right\} u_{i,t}^2 = o_p(1)$ from which it further follows that

$$\left\| U_{t,N} \left(\widehat{H}^c \right) / \sqrt{N_1} \right\|_2^2 \leq \frac{1}{N_1} \sum_{i \in H^c} u_{i,t}^2 + \frac{1}{N_1} \sum_{i \in H} \mathbb{I} \left\{ i \in \widehat{H}^c \right\} u_{i,t}^2 = O_p(1) + o_p(1) = O_p(1).$$

Turning our attention to part (f), note first that since $G = \begin{bmatrix} G_1 & G_2 \end{bmatrix}$ is an orthogonal matrix, we have $I_N = GG' = G_1G_1' + G_2G_2'$ or $G_2G_2' = I_N - G_1G_1'$. Hence, we can write $\left\| G_2' U_{t,N} \left(\widehat{H}^c \right) / \sqrt{N_1} \right\|_2^2 = U_{t,N} \left(\widehat{H}^c \right)' U_{t,N} \left(\widehat{H}^c \right) / N_1 - U_{t,N} \left(\widehat{H}^c \right)' G_1 G_1' U_{t,N} \left(\widehat{H}^c \right) / N_1 \leq \left\| U_{t,N} \left(\widehat{H}^c \right) / \sqrt{N_1} \right\|_2^2 + \left\| G_1' U_{t,N} \left(\widehat{H}^c \right) / \sqrt{N_1} \right\|_2^2$. Applying the results from parts (d) and (e) above, we then obtain $\left\| G_2' U_{t,N} \left(\widehat{H}^c \right) / \sqrt{N_1} \right\|_2^2 \leq \left\| U_{t,N} \left(\widehat{H}^c \right) / \sqrt{N_1} \right\|_2^2 + \left\| G_1' U_{t,N} \left(\widehat{H}^c \right) / \sqrt{N_1} \right\|_2^2 = O_p(1) + o_p(1) = O_p(1)$, from which it follows by Slutsky's theorem that $\left\| G_2' U_{t,N} \left(\widehat{H}^c \right) / \sqrt{N_1} \right\|_2 = O_p(1)$.

Now, to show part (g), first write $\widehat{V}' \widehat{G}_1' U_{t,N} \left(\widehat{H}^c \right) / \sqrt{\widehat{N}_1}$
 $= \left(1 + \left[\widehat{N}_1 - N_1 \right] / N_1 \right)^{-\frac{1}{2}} \widehat{V}' \widehat{G}_1' U_{t,N} \left(\widehat{H}^c \right) / \sqrt{N_1}$. Using the fact that $\widehat{V}' \widehat{V} = I_{Kp}$ so that $\left\| \widehat{V} \right\|_2 = 1$, we have

$$\begin{aligned}
\left\| \frac{\widehat{V}' \widehat{G}_1' U_{t,N} \left(\widehat{H}^c \right)}{\sqrt{\widehat{N}_1}} \right\|_2 &= \left\| \frac{\widehat{V}' (I_{Kp} + R'R)^{-1/2} \left[G_1' U_{t,N} \left(\widehat{H}^c \right) + R' G_2' U_{t,N} \left(\widehat{H}^c \right) \right]}{\sqrt{N_1}} \right\|_2 \\
&\leq \left\| \widehat{V} \right\|_2 \left\| (I_{Kp} + R'R)^{-1/2} \right\|_2 \left\{ \left\| \frac{G_1' U_{t,N} \left(\widehat{H}^c \right)}{\sqrt{N_1}} \right\|_2 + \|R\|_2 \left\| \frac{G_2' U_{t,N} \left(\widehat{H}^c \right)}{\sqrt{N_1}} \right\|_2 \right\} \\
&\leq \frac{1}{\sqrt{1 + \lambda_{\min}(R'R)}} \left\{ \left\| \frac{G_1' U_{t,N} \left(\widehat{H}^c \right)}{\sqrt{N_1}} \right\|_2 + \|R\|_2 \left\| \frac{G_2' U_{t,N} \left(\widehat{H}^c \right)}{\sqrt{N_1}} \right\|_2 \right\}.
\end{aligned}$$

It follows that

$$\left\| \frac{\widehat{V}' \widehat{G}_1' U_{t,N} \left(\widehat{H}^c \right)}{\sqrt{\widehat{N}_1}} \right\|_2 = \left| \left(1 + \frac{\widehat{N}_1 - N_1}{N_1} \right)^{-\frac{1}{2}} \right| \left\| \frac{\widehat{V}' \widehat{G}_1' U_{t,N} \left(\widehat{H}^c \right)}{\sqrt{N_1}} \right\|_2$$

$$\begin{aligned}
&\leq \left| \left(1 + \frac{\hat{N}_1 - N_1}{N_1} \right)^{-\frac{1}{2}} \right| \frac{1}{\sqrt{1 + \lambda_{\min}(R'R)}} \left\{ \left\| \frac{G'_1 U_{t,N}(\widehat{H}^c)}{\sqrt{N_1}} \right\|_2 + \|R\|_2 \left\| \frac{G'_2 U_{t,N}(\widehat{H}^c)}{\sqrt{N_1}} \right\|_2 \right\} \\
&\leq \left| \left(1 + \frac{\hat{N}_1 - N_1}{N_1} \right)^{-\frac{1}{2}} \right| \left\{ \left\| \frac{G'_1 U_{t,N}(\widehat{H}^c)}{\sqrt{N_1}} \right\|_2 + \|R\|_2 \left\| \frac{G'_2 U_{t,N}(\widehat{H}^c)}{\sqrt{N_1}} \right\|_2 \right\} = o_p(1)
\end{aligned}$$

where the last equality follows from the fact that $\|R\|_2 \xrightarrow{p} 0$, $\left| \left(1 + \frac{\hat{N}_1 - N_1}{N_1} \right)^{-\frac{1}{2}} \right| \xrightarrow{p} 1$, $\left\| G'_1 U_{t,N}(\widehat{H}^c) / \sqrt{N_1} \right\|_2 \xrightarrow{p} 0$, and $\left\| G'_2 U_{t,N}(\widehat{H}^c) / \sqrt{N_1} \right\|_2 = O_p(1)$, given results previously provided in part (a) in Lemma OA-1 and in parts (a), (d), and (f) of this lemma.

To show part (h), note first that, under Assumptions 2-1, 2-2(a)-(b), 2-5, and OA-3; there exists a positive constant \overline{C} such that $E \|\underline{E}_t\|_2^6 \leq \overline{C} < \infty$ ⁷. Hence, for any $\epsilon > 0$, let $C_\epsilon = \overline{C}^{\frac{1}{6}} / \sqrt{\epsilon}$; and we can apply Markov's inequality and then Liapunov's inequality to obtain $\Pr(\|\underline{E}_t\|_2 \geq C_\epsilon) \leq \Pr(\|\underline{E}_t\|_2^2 \geq C_\epsilon^2) \leq C_\epsilon^{-2} E \|\underline{E}_t\|_2^2 \leq C_\epsilon^{-2} (E \|\underline{E}_t\|_2^6)^{\frac{1}{3}} \leq \epsilon \overline{C}^{-\frac{1}{3}} \overline{C}^{\frac{1}{3}} \leq \epsilon$, from which it follows that $\|\underline{E}_t\|_2 = O_p(1)$ for all t . \square

References

- [1] Bai, J. and S. Ng (2021): "Approximate Factor Models with Weaker Loading," Working Paper, Columbia University.
- [2] Chao, J. C. and N. R. Swanson (2022a): "Consistent Estimation, Variable Selection, and Forecasting in Factor-Augmented VAR Models," Working Paper, Rutgers University and University of Maryland.
- [3] Chao, J. C. and N. R. Swanson (2022b): Technical Appendix to "Consistent Estimation, Variable Selection, and Forecasting in Factor-Augmented VAR Models," Working Paper, Rutgers University and University of Maryland.
- [4] Golub, G. H. and C. F. van Loan (1996): *Matrix Computations*, 3rd Edition. Baltimore: The Johns Hopkins University Press.
- [5] Stewart, G.W. and J. Sun (1990): *Matrix Perturbation Theory*. Boston: Academic Press.
- [6] Stock, J. H. and M. W. Watson (2002): "Forecasting Using Principal Components from a Large Number of Predictors," *Journal of the American Statistical Association*, 97, 1167-1179.

⁷An explicit proof that, under Assumptions 2-1, 2-2(a)-(b), 2-5, and 2-6; there exists some positive constant \overline{C} such that $E \|\underline{E}_t\|_2^6 \leq \overline{C}$ is given in the Technical Appendix of an earlier version of this paper, Chao and Swanson (2022b). In particular, this result is shown in Lemma C-5 in Appendix C of Chao and Swanson (2022b).