SUPPLEMENTARY APPENDIX FOR THE FACTOR-LASSO AND K-STEP BOOTSTRAP APPROACH FOR INFERENCE IN HIGH-DIMENSIONAL ECONOMIC APPLICATIONS

CHRISTIAN HANSEN AND YUAN LIAO

Abstract. This supplement contains additional proofs, simulation results, and empirical results for the paper "The Factor-Lasso and K-Step Bootstrap Approach for Inference in High-Dimensional Economic Applications."

Contents

Christian Hansen, Booth School of Business, University of Chicago, Chicago, IL 60637. Christian.Hansen@chicagobooth.edu. Yuan Liao, Department of Economics, Rutgers University, New Brunswick, NJ 08901. yuan.liao@rutgers.edu.

Appendix D. Estimating Factors Using Principal Components Analysis

Let

$$\tilde{X} = \begin{pmatrix} \tilde{X}_{i1} & \cdots & \tilde{X}_{n1} \\ \vdots & & \vdots \\ \tilde{X}_{iT} & \cdots & \tilde{X}_{nT} \end{pmatrix}_{pT\times n}, \quad \tilde{\Lambda} = \begin{pmatrix} \tilde{\Lambda}_1 \\ \vdots \\ \tilde{\Lambda}_T \end{pmatrix}_{pT\times K}, \quad \tilde{F} = \begin{pmatrix} \tilde{f}_1' \\ \vdots \\ \tilde{f}_n' \end{pmatrix}_{n\times K},$$

and define $\tilde{U}$ similarly. The matrix form of the factor model is then

$$\tilde{X} = \tilde{\Lambda}\tilde{F}' + \tilde{U},$$

where the individual and time effects have already been removed. Let $\widehat{F} = (\widehat{f}_1, ..., \widehat{f}_n)'$ denote the $n\times K$ matrix of the estimated factors. The columns of $\widehat{F}/\sqrt{n}$ are the eigenvectors of the first K eigenvalues of $\tilde{X}'\tilde{X}/(npT)$. We call this the "PC estimator".[1] We now present

[1]We choose to focus on the PC estimator as a concrete example because it is relatively simple and is free of tuning parameters. One could consider other options which would also satisfy our assumed high-level conditions. For example, the weighted PC estimator (e.g. Choi (2012); Bai and Liao (2013)), can be more efficient than the standard PC estimator but requires additional tuning parameters for practical application.

regularity conditions which are sufficient to verify that the PC estimator satisfies the high-level sufficient conditions for the quality of factor estimation given in Assumption D.4 in Section D.1. These regularity conditions are standard for high-dimensional approximate factor models.

Assumption D.1 (Pervasiveness)**.** *There are* $c, C > 0$ *so that*

$$c < \lambda_{\min}\left(\frac{1}{T}\sum_{t=1}^{T}\frac{1}{p}\tilde{\Lambda}_t'\tilde{\Lambda}_t\right) \le \lambda_{\max}\left(\frac{1}{T}\sum_{t=1}^{T}\frac{1}{p}\tilde{\Lambda}_t'\tilde{\Lambda}_t\right) < C.$$

where $\lambda_{\min}(A)$ *and* $\lambda_{\max}(A)$ *respectively denote the minimum and maximum eigenvalues of* A*.*

Assumption D.2 (Second Order Weak Dependence)**.** *There is* $C > 0$*,*

$$\begin{aligned}
&\max_{m,t,i}\sum_{s=1}^{T}\sum_{v=1}^{p}\operatorname{Cov}(U_{it,m}^2, U_{is,v}^2) < C,\\
&\max_{i,m,s,t,v}\sum_{h=1}^{T}\sum_{l=1}^{p}|\operatorname{Cov}(U_{it,v}U_{is,m}, U_{ih,l}U_{is,m})| < C,\\
&\max_{i}\frac{1}{T^2p}\sum_{m,l\le p}\sum_{t,s,h,v\le T}\operatorname{Cov}(U_{it,m}U_{is,m}, U_{ih,l}U_{iv,l}) < C,\\
&\max_{i,m}\frac{1}{T^2p}\sum_{k,l\le p}\sum_{t,s,h,v\le T}|\operatorname{Cov}(U_{it,k}U_{is,m}, U_{ih,l}U_{iv,m})| < C.
\end{aligned}$$

Assumption D.3 (Uniform cross-sectional convergence)**.**

$$\begin{aligned}
\max_{i\le n}|\tfrac{1}{pT}\textstyle\sum_{t=1}^{T}\sum_{m=1}^{p}(U_{it,m}^2 - EU_{it,m}^2)| = O_P(1).\\
\max_{i\le n}\|\tfrac{1}{pT}\textstyle\sum_{t=1}^{T}\sum_{m=1}^{p}\tilde{\lambda}_{t,m}U_{it,m}\|_2 = O_P\left(\sqrt{\tfrac{\log n}{pT}}\right).
\end{aligned}$$

D.1. **High-level assumptions on the estimated factors.** More generally, the estimated factors should satisfy the following conditions in the original and the bootstrap data.

Assumption D.4 (Quality of Factor Estimation in Original Data)**.** *Suppose there is an invertible* $\dim(f_i)\times\dim(f_i)$ *matrix* H *with* $\|H\|+\|H^{-1}\| = O_P(1)$*, and non-negative sequences* Δ_F*,* Δ_{eg}*,* Δ_{ud}*,* Δ_{fum}*,* $\Delta_{fe}, \Delta_{\max}$*, so that for* $\tilde{z}_{it} \in \{\tilde{\epsilon}_{it}, \tilde{\eta}_{it}\}$*,* $\tilde{w}_{tm} \in \{\tilde{\Lambda}_t'\gamma_d, \tilde{\Lambda}_t'\gamma_y, \tilde{\delta}_{dt}, \tilde{\delta}_{yt}, \tilde{\lambda}_{tm}\}$*,* $\tilde{h}_{tk} \in \{\tilde{\delta}_{dt}, \tilde{\delta}_{yt}, \tilde{\lambda}_{tk}\}$*, and* $\gamma \in \{\gamma_d, \gamma_y\}$*,*

$$\max_{i\le n}\|\widehat{f}_i - H'\tilde{f}_i\|_2 = O_P(\Delta_{\max}), \quad \frac{1}{n}\sum_{i=1}^{n}\|\widehat{f}_i - H'\tilde{f}_i\|_2^2 = O_P(\Delta_F^2)$$

$$\frac{1}{T}\sum_{t=1}^{T}\|\frac{1}{n}\sum_{i=1}^{n}(\widehat{f}_i - H'\tilde{f}_i)\tilde{z}_{it}\|_2^2 = O_P(\Delta_{fe}^2),$$
$$\max_{m\leq p}\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(\widehat{f}_i - H'\tilde{f}_i)\tilde{z}_{it}\tilde{w}_{tm}'\|_F = O_P(\Delta_{eg}),$$
$$\max_{m,k\leq p}\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(\widehat{f}_i - H'\tilde{f}_i)\tilde{U}_{it,m}\tilde{h}_{tk}'\|_F = O_P(\Delta_{ud}),$$
$$\max_{m\leq p, t\leq T}\|\frac{1}{n}\sum_{j=1}^{n}(\widehat{f}_j - H'\tilde{f}_j)\tilde{U}_{jt,m}\|_2 = O_P(\Delta_{fum}).$$

These sequences satisfy the following restrictions:

$$\sqrt{nT}|J|_0^2\Delta_F^2 = o(1),\ \Delta_{eg} = o\left(\frac{1}{\sqrt{nT}}\right),\ \Delta_{ud} = o\left(\sqrt{\frac{\log p}{nT}}\right),\ |J|_0^2\sqrt{\log p}\Delta_{ud} = o(1),$$
$$\Delta_{fum}^2 = o\left(\frac{\log p}{T|J|_0^2\log(pT)}\right),\ \Delta_{fe}^2 = o\left(\frac{\log p}{T\log(pT)}\right),\ \Delta_{\max}^2 = O(\log(n)),\ \textit{and}$$
$$\Delta_{\max}^2|J|_0^2T\left(\lambda_n^2|J|_0 + \Delta_F^2|J|_0^2 + \frac{|J|_0}{n}\right) = o(1).$$

Assumption D.5 (Quality of Factor Estimation in Bootstrap Data)**.** *Suppose there is an invertible* $\dim(f_i)\times\dim(f_i)$ *matrix* H^* *with* $\|H^*\| + \|H^{*-1}\| = O_{P^*}(1)$*, and non-negative sequences* Δ_F^**,* Δ_{eg}^**,* Δ_{ud}^**,* Δ_{fe}^**, so that for* $\tilde{z}_{it}^* \in \{\tilde{\eta}_{it}^*, \tilde{\epsilon}_{it}^*\}$*,* $\widehat{g}_{tm} \in \{\widehat{\Lambda}_t'\widehat{\gamma}_d, \widehat{\Lambda}_t'\widehat{\gamma}_y, \widehat{\delta}_{dt}, \widehat{\delta}_{yt}, \widehat{\lambda}_{tm}\}$*, and* $\widehat{h}_{tm} \in \{\widehat{\delta}_{dt}, \widehat{\delta}_{yt}, \widehat{\lambda}_{tm}\}$*,*

$$\frac{1}{n}\sum_{i=1}^{n}\|\widehat{f}_i^* - H^{*'}\widehat{f}_i\|_2^2 = O_P(\Delta_F^{*2})$$
$$\max_{m\leq p}\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{z}_{it}^*\widehat{g}_{tm}(\widehat{f}_i^* - H^*\widehat{f}_i)'\|_F = O_{P^*}(\Delta_{eg}^*)$$
$$\max_{m,k\leq p}\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(\widehat{f}_i^* - H^{*'}\widehat{f}_i)\tilde{U}_{it,m}^*\widehat{h}_{tk}'\|_F = O_P(\Delta_{ud}^*).$$

These sequences satisfy the following restrictions:

$$\sqrt{nT}|J|_0^2\Delta_F^{*2} = o(1),\ \Delta_{eg}^* = o\left(\frac{1}{\sqrt{nT}}\right),\ \Delta_{ud}^* = o\left(\sqrt{\frac{\log p}{nT}}\right),\ |J|_0^2\sqrt{\log p}\Delta_{ud}^* = o(1),$$
$$\Delta_F^{*2} = o\left(\frac{\log p}{T\log(pT)}\right),\ \textit{and}\ \Delta_{\max}^2|J|_0^2T\Delta_F^{*2} = o(1).$$

The following proposition shows that the high-level Assumptions D.4 and D.5 are satisfied by the PC estimator.

Proposition D.1. *Assume* $|J|_0^4 = o(nT^3)$, $|J|_0^4 n = o(p^2 T)$ *and* $|J|_0^2 \log n = o(p)$. *Then Assumptions D.1 - D.3 imply Assumptions D.4 and D.5 about* $\widehat{F}$ *and* $\widehat{F}^*$.

The conditions $|J|_0^4 n = o(p^2 T)$ and $|J|_0^2 \log n = o(p)$ require lower bounds on the growth of p. These conditions differ from those used in the literature on inference in purely sparse high-dimensional models, e.g. Belloni et al. (2014), in that lower bounds on p are not required in the purely sparse setting. These lower bounds arise because accurately estimating the unknown factors using PCA requires a large number of observed series. In the special case $|J|_0 = O(1)$, these conditions require

$$T \ll n \ll p^2 T, \ \log^3 p = O(n), \ \text{and} \ \log n = o(p).$$

Technically, existing results in the literature on estimating factor models are not directly applicable to prove Proposition D.1. In Appendix I.1, we show that

$$\frac{1}{n}\sum_{i=1}^{n} \|\widehat{f}_i - H'\tilde{f}_i\|_2^2 = O_P\left(\frac{1}{pT} + \frac{1}{n^2} + \frac{1}{nT^2}\right)$$

when $\widehat{f}_i$ is estimated via PCA. This rate is fairly standard (see, e.g., Bai (2003)). On one hand, the estimation of factors depends on the number of "units" related to the factors. In the current model $\tilde{X}_{it} = \tilde{\Lambda}_t \tilde{f}_i + \tilde{U}_{it}$, this number of "units" is pT as the factors are common to the p observed covariates and the T time periods. On the other hand, the rate should also depend on the number of observations available for estimating the coefficients on the factors. This number of observations enters as $O_P(n^{-2})$. (Note that in the current notation, the rate in Bai (2003) would be $O_P(n^{-1})$ under serial correlation. We achieve a faster rate $O_P(n^{-2})$ due to the independence across i.) The term $O_P(\frac{1}{nT^2})$ is a nonstandard component, which appears due to the use of demeaned $\tilde{U}_{it}$ in the model. Specifically, we need to bound $\frac{1}{n^2}\sum_i E[\frac{1}{p}\sum_{m=1}^{p} \bar{U}_{i\cdot,m}^2]^2$, which leads to an extra term $O_P(\frac{1}{nT^2})$.

The above rate allows conditions involving Δ_F to be directly verified. However, it does not imply the uniform convergence condition $\max_{t\leq T} \|\widehat{f}_i - H'\tilde{f}_i\|_2$. Nor is this result sufficient to verify the other stated conditions because other terms, e.g. Δ_{eg}, Δ_{fum}, and Δ_{fe}, involve "weighted averages" of $\{\widehat{f}_i - H'\tilde{f}_i\}$ whose rates of convergence can be derived and shown to be faster than that of $\Delta_F^2 = \frac{1}{pT} + \frac{1}{n^2} + \frac{1}{nT^2}$. For instance, if we use a simple

Cauchy-Schwarz inequality to bound Δ_{ud}, we would have

$$\max_{m,k\leq p}\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(\widehat{f}_i-H'\tilde{f}_i)\tilde{U}_{it,m}\tilde{h}_{tk}'\|_F^2\leq\frac{1}{n}\sum_{i=1}^{n}\|\widehat{f}_i-H'\tilde{f}_i\|_2^2\max_{m,k\leq p}\frac{1}{n}\sum_{i=1}^{n}\|\frac{1}{T}\sum_{t=1}^{T}\tilde{U}_{it,m}\tilde{h}_{tk}\|_2^2.$$

It can be shown that $\max_{m,k\leq p}\frac{1}{n}\sum_{i=1}^{n}\|\frac{1}{T}\sum_{t=1}^{T}\tilde{U}_{it,m}\tilde{h}_{tk}\|_2^2=O_P(\frac{\log p}{T})$, so this crude bound gives us $\Delta_{ud}=\Delta_F\sqrt{\frac{\log p}{T}}$. Unfortunately, this bound is not sharp enough to verify that $\Delta_{ud}=o(\sqrt{\frac{\log p}{nT}})$ unless $n=o(pT)$. In the special case that T is fixed, requiring $n=o(p)$ is a restrictive condition. Rather than relying on these crude bounds, we achieve sharper bounds by directly deriving the rate of convergence for each required term in Appendix I.1, which relies on some novel technical work. Using these sharper bounds, we only require $n=o(p^2T)$ which provides much more freedom on the ratio n/p.

Appendix E. Additional Numerical Results

This section contains additional simulation results and an additional empirical example based on a cross-sectional version of the model generalized to allow for an endogenous variable. Implementing the factor-lasso estimation algorithm in this setting requires a small adaptation to the algorithm in Section 2.2 of the main paper. First, we outline the modification of the basic procedure in the panel data case with a single instrument, called Z, for clarity. We then note that, in examples with a single cross-section, we simply drop all of the individual specific effects in the obvious way as estimation with unrestricted individual-specific heterogeneity cannot proceed with a single cross-section.

Specifically, we add an additional equation to the model to obtain

$$y_{it}=\alpha d_{it}+\xi_t'f_i+U_{it}'\theta+g_i+\nu_t+\epsilon_{it} \tag{E.1}$$

$$d_{it}=\pi z_{it}+\delta_{dt}'f_i+U_{it}'\gamma_d+\zeta_i+\mu_t+\eta_{it} \tag{E.2}$$

$$z_{it}=\delta_{zt}'f_i+U_{it}'\gamma_z+a_i+b_t+\eta_{it} \tag{E.3}$$

$$X_{it}=\Lambda_t f_i+w_i+\rho_t+U_{it}, \tag{E.4}$$

and adapt the algorithm in the following manner.

Algorithm (**IV Factor-Lasso Estimation of** α.)

(1) Obtain $\{\widehat{f}_i,\widehat{U}_{it}\}_{i\leq n,t\leq T}$ by extracting factors from the model $\tilde{X}_{it}=\tilde{\Lambda}_t\tilde{f}_i+\tilde{U}_{it}$.

(2) For $\widehat{\delta}_{yt} = (\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{Y}_t$, $\widehat{\delta}_{dt} = (\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{D}_t$, and $\widehat{\delta}_{zt} = (\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{Z}_t$, run the cluster-lasso programs (2.6) and (2.7) to obtain $\tilde{\gamma}_y$ and $\tilde{\gamma}_d$ and the cluster-lasso program

$$\tilde{\gamma}_z = \arg\min_{\gamma\in\mathbb{R}^p} \frac{1}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}(\tilde{z}_{it} - \widehat{\delta}_{zt}'\widehat{f}_i - \widehat{U}_{it}'\gamma)^2 + \kappa_n\|\widehat{\Psi}^z\gamma\|_1 \tag{E.5}$$

with tuning parameters as in (2.6) and (2.7).

(3) Obtain the estimator $\widehat{\alpha}$ and corresponding estimated standard error as the coefficient on $\tilde{d}_{it} - \widehat{\delta}_{dt}'\widehat{f}_i - \widehat{U}_{it,\widehat{J}}'\widehat{\gamma}_d$ and associated clustered standard error from the instrumental variables regression of $\tilde{y}_{it} - \widehat{\delta}_{yt}'\widehat{f}_i - \widehat{U}_{it,\widehat{J}}'\widehat{\gamma}_y$ on $\tilde{d}_{it} - \widehat{\delta}_{dt}'\widehat{f}_i - \widehat{U}_{it,\widehat{J}}'\widehat{\gamma}_d$ using $\tilde{z}_{it} - \widehat{\delta}_{zt}'\widehat{f}_i - \widehat{U}_{it,\widehat{J}}'\widehat{\gamma}_z$ as the instrumental variable where $\widehat{U}_{it,\widehat{J}}$ is the subvector of $\widehat{U}_{it}$ whose elements are $\{\widehat{U}_{it,j} : j \in \widehat{J}\}$ and $\widehat{J} = \{j \le p : \tilde{\gamma}_{y,j} \neq 0\} \cup \{j \le p : \tilde{\gamma}_{d,j} \neq 0\} \cup \{j \le p : \tilde{\gamma}_{z,j} \neq 0\}$.

E.1. **Instrumental Variables Model Simulations.** We generate data from the model

$$\begin{aligned}
y_i &= \alpha d_i + (c_\xi \xi)' f_i + U_i'(c_\theta \theta) + \nu + \epsilon_i \\
d_i &= \pi z_i + (c_{\delta_d}\delta_d)' f_i + U_i'(c_{\gamma_d}\gamma_d) + \mu + \eta_i \\
z_i &= (c_{\delta_z}\delta_z)' f_i + U_i'(c_{\gamma_z}\gamma_z) + \zeta + v_i \\
X_i &= (c_\Lambda \Lambda) f_i + \rho + U_i
\end{aligned}$$

with $n = 100$, $K = 2$, and $p = 100$. Within this model, d_i is an endogenous variable with coefficient of interest α and z_i is an instrumental variable. We generate $\epsilon_i \sim N(0,1)$ and $\eta_i \sim N(0,1)$ with $\mathrm{E}[\epsilon_i\eta_i] = .8$ i.i.d. across i and independent of all other random variables. We generate i.i.d. draws for U_i as in the simulations in the main text, and $v_i \sim N(0,1)$ independently from U_i. We also generate $(\nu, \mu, \rho, \xi, \delta_d, \Lambda, c_\Lambda, c_{\gamma_z}, c_{\delta_d}, c_\xi, c_\theta)$ exactly as in the simulations in the main text. We set $\theta = \gamma_d = \gamma_z$ to be vectors with j^{th} entry given by $\theta_j = \gamma_{d,j} = \gamma_{z,j} = \frac{1}{j^2}$. To control the strength of the instrument, we choose $(c_{\delta_z}, c_{\gamma_z})$ so that the R^2 of the infeasible regression of $z_i - \zeta$ on $(c_{\delta_z}\delta_z)'f_i + U_i'(c_{\gamma_z}\gamma_z)$ is 0.7 and the factors account for 50% of the explanatory power in this regression. We set π so that the fraction of variation accounted for by z_i in the regression of d_i on z_i, f_i and U_i is 25%. Finally, we set $\alpha = 1$.

We again estimate α using six different IV procedures similar to those implemented in the main simulation results with one exception. As the number of features is equal to the sample size in these simulations, we consider an infeasible "oracle" estimator that estimates α from IV regression of $y_i - (c_\xi \xi)' f_i - U_i'(c_\theta \theta) - \nu$ on $d_i - (c_{\delta_d}\delta_d)' f_i - U_i'(c_{\gamma_d}\gamma_d) - \mu$ using $z_i - (c_{\delta_z}\delta_z)' f_i - U_i'(c_{\gamma_z}\gamma_z) - \zeta$ as instrument (Oracle). This estimator provides a type of best-case benchmark and allows us to ascertain that instruments are strong enough that the usual asymptotic approximation provides a reasonable approximation in the idealized scenario where one is able to perfectly remove the effect of confounding from all variables.

Figure 1 gives simulation RMSEs for the estimator of α resulting from applying each procedure. The RMSEs are truncated at 0.1 for readability of the figure.[2] Again, we see that the factor lasso procedure delivers good performance regardless of the relative strength of the factors and factor residuals in this simulation design. Each of the other procedures exhibits behavior that depends strongly on the exact strength of the factors in the different equations. It might be noted that the dominance of the factor-lasso estimator, in terms of RMSE, over the "Oracle" procedure is due to the definition of the oracle that we use which fully removes the variation in each variable due to factors and factor residuals even in situations in which some of these variables produce no confounding. For example, one need not remove the variation in the instruments due to the factors in cases where the factors have zero loadings in the outcome equation, but this variation is always removed due to the way we have defined the oracle model.

We report size of 5% level tests based on standard asymptotic approximations for each of the six procedures considered in Figure 2 with size truncated at 0.3 for readability of the figure. In each panel, we report the rejection frequency of the standard t-test of the null hypothesis that $\alpha = 1$ using heteroscedasticity robust standard errors. Here, we see that the only procedure that uniformly controls size is the infeasible oracle. Among the feasible procedures, the proposed factor lasso approach performs relatively well in keeping size distortions small across the majority of combinations of relative strengths of the factors. In this case, we do see that the factor-lasso procedure suffers from reasonably large size distortions when the factors account for all of the confounding in the outcome equation and a moderate amount of counfounding in the treatment equation. We also see that the pure factor model controls size well in this case, but performs very poorly once all variation in the outcome equation is not due to the factors.

[2] Theoretically, the MSE of the IV estimator does not exist in this context. We report root mean truncated squared error with a truncation point of 1.

We again conclude by looking at the performance of the k-step bootstrap in Figure 3. We see that there is a modest, but clearly visible, improvement from using the k-step bootstrap relative to the asymptotic approximation. The score based bootstrap, on the other hand, lines up reasonably well with the asymptotic approximation.

E.2. **Estimating the Effects of Institutions on Output.** We revisit the example considered in Acemoglu et al. (2001). Acemoglu et al. (2001) are interested in the parameter α in a structural model of the form

$$\log(\text{GDP per capita}_i) = \alpha(\text{Protection from Expropriation}_i) + x_i'\beta + \varepsilon_i$$

based on aggregate country level data where "Protection from Expropriation" is a measure of the strength of individual property rights that is used as a proxy for the strength of institutions and x_i is a set of variables that are meant to control for geography. Acemoglu et al. (2001) adopt an IV strategy where they instrument for institution quality using early European settler mortality to estimate α as institutions are clearly potentially endogenous. They point out that their instrument would be invalid if there were other factors that are highly persistent and related to the development of institutions within a country and to the country's GDP. A leading candidate for such a factor that they discuss is geography. To address this possibility, Acemoglu et al. (2001) control for the distance from the equator in their baseline specifications and consider different sets of geographic controls such as continent dummies within their robustness checks (e.g. Acemoglu et al. (2001) Table 4).

There are, of course, many other ways to measure geography besides distance to the equator or continent where a country is found. Rather than *ex ante* choosing a small number of variables to proxy for geography, we put a large number of variables that potentially capture geography in x_i and then use the data to reduce dimension. Specifically, we consider dummies for Africa, Asia, North America, and South America as well as longitude, renewable water, land boundary, land area, amount of coastline, territorial seas, amount of arable land, average temperature, average high temperature, average low temperature, average precipitation, elevation of highest point, elevation of lowest point, fraction of area that is low-lying, latitude, and spherical distance from London.[3]

[3]We apply the PCA to extract common factors, recognizing that it may not be the ideal method in the presence of dummy variables. Several approaches have been developed in the literature to extend the PCA to factor models with categorical data. We refer to Keller and Wansbeek (1983), Ng (2015), and references therein.

We adapt the analysis of Acemoglu et al. (2001) to the present setting by considering estimation of a partial factor instrumental variables model

$$\begin{aligned}
\log(\text{GDP per capita}_i) &= \alpha(\text{Protection from Expropriation}_i) + f_i'\xi + U_i'\theta + \varepsilon_i \\
\text{Protection from Expropriation}_i &= \pi\text{Early Settler Mortality}_i + f_i'\delta_d + U_i'\gamma_d + \eta_i \\
\text{Early Settler Mortality}_i &= f_i'\delta_z + U_i'\gamma_z + v_i \\
X_i &= \Lambda f_i + U_i
\end{aligned}$$

using our 20 geography measures as x_i and the 64 countries from the original Acemoglu et al. (2001) data. The factor-lasso approach seems quite sensible in this setting. Each of the observed geography measures could reasonably be taken as a noisy proxy for a country's geography. This relationship is likely to be complicated and uneven with the chief features leading to association between the geography proxies plausibly being only weakly related to the notions of geography that are important predictors of mortality and institutions. The factor-lasso approach, by allowing a small number of elements of U_i to enter the equation of interest in addition to any common geography factors, readily accommodates this latter possibility in a parsimonious, data-dependent way.

We report estimation results for the first stage coefficient on the instrument in Table 1. We report results from the factor-lasso approach in the row "Factor-Lasso". For comparison, we also report results from a few natural alternative models. The row labeled "Latitude" uses the single variable distance from the equator to control for geography as in the baseline results from Acemoglu et al. (2001). We report results based on using all 20 available geographic controls without dimension reduction in the row labeled "All Controls." We apply the double selection approach of Belloni et al. (2014) which would be appropriate if the relationship between geographic controls and the variables of interest were well-approximated by a sparse linear model in "Double Selection." Finally, "Factor" reduces dimension through positing a conventional factor model. All factors are estimated using PCA with number of factors selected by applying the procedure from Ahn and Horenstein (2013).

The first-stage results using only the latitude control suggest there is a fairly strong relationship between the instrument and endogenous variable if latitude is a sufficient control for geography. The first stage F-statistic using just latitude is 10.9 which many would take to indicate that the instrument is sufficiently strong to identify the effect of interest.[4]

[4]A benchmark that is commonly used in the applied literature to assess whether there is sufficient variation in the instrument to identify the effect of interest is to compare the first stage F-statistic to 10, with smaller values indicating weak identification.

The results change in a potentially substantive way after allowing for the possibility that geography is not adequately captured by latitude. For each of the remaining approaches considered, the first-stage F-statistic drops substantially below 10, with all methods besides applying the pure factor model returning first-stage coefficients that are statistically insignificant at the 5% level.

One might dismiss the lack of significance after including all controls without dimension reduction as it seems likely that a model with 20 covariates in addition to the variables of interest and only 64 observations is overfit. The next strongest result is from the pure factor model which makes use of a single extracted component and produces a first-stage F-statistic of 7.5. As evidenced in the simulation example, inference results based on a pure factor model may be misleading when elements of U_i also have explanatory power. It is then interesting that the double-selection approach and the factor-lasso approach deliver almost identical results indicating a weak association between the endogenous variable and instrument after controlling parsimoniously for geography. The double-selection procedure selects four variables[5], and the factor-lasso approach uses one factor and two additional variables.[6] One might take this to mean that the four variables selected in the double-selection procedure approximately capture the same information as the single factor and two variables used in the factor-lasso results. In either case, the results suggest that, at best, identification of the structural effect of institutions as measured by "Protection from Expropriation" using settler mortality as instrument is weak after geography is controlled for in a parsimonious, data-dependent way. Given this apparent weak identification, we do not report second stage estimates of the structural effect.[7]

Appendix F. The Lasso estimator in the presence of latent factors

F.1. The convergence of the lasso estimator on the original data.

Proposition F.1. *Under Assumptions 3.1-3.4 and Assumption D.4, for $x \in \{d, y\}$, the lasso estimator $\tilde{\gamma}_x$ satisfies: (i)*

$$\frac{1}{n}\sum_{t=1}^{T}\|\widehat{U}_t(\gamma_x - \tilde{\gamma}_x)\|_2^2 = O_P(\kappa_n^2|J|_0),$$

[5]These variables are the Africa dummy, average temperature, average high temperature, and amount of arable land.

[6]The two selected variables in addition to the factor are the Africa and Asia dummies.

[7]We note that it would be straightforward to adapt the weak-identification robust procedure of Chernozhukov and Hansen (2008) to the present setting. We do not pursue this extension for brevity.

$$\|\gamma_x - \tilde{\gamma}_x\|_1 = O_P(\kappa_n |J|_0 + \|R_x\|_1).$$

(ii) $|\widehat{J}|_0 = O_P(|J|_0)$.

Proof of Proposition F.1 (i)

In the following, we show the argument for $\tilde{\gamma}_y$. The argument for $\tilde{\gamma}_d$ is similar.

Let

$$l(v) := \frac{1}{nT}\sum_{t=1}^T \|\tilde{Y}_t - \widehat{F}\widehat{\delta}_{yt} - \widehat{U}_t v\|_2^2.$$

Recall that the solution of the lasso program is $\tilde{\gamma}_y = \arg\min_{\gamma \in \mathbb{R}^p} l(\gamma) + \kappa_n \|\widehat{\Psi}^y \gamma\|_1$. We shall use J to denote the active set. Let Ψ^y_J denote a $|J|_0 \times |J|_0$ diagonal matrix whose entries are the elements of Ψ^y in the set J. Recall that γ_y can be decomposed as

$$\gamma_y = \underbrace{\gamma_y^0}_{\text{exactly sparse}} + \underbrace{R_y}_{\text{remainder}}.$$

Recall that $\tilde{Y}_t = \tilde{F}\tilde{\delta}_{yt} + \tilde{U}_t \gamma_y + \tilde{e}_t$. Let

$$M_t = \tilde{F}\tilde{\delta}_{yt} - \widehat{F}\widehat{\delta}_{yt} + (\tilde{U}_t - \widehat{U}_t)\gamma_y, \quad \Delta_\gamma = \gamma_y^0 - \tilde{\gamma}_y.$$

Then $l(\tilde{\gamma}_y) + \kappa_n \|\widehat{\Psi}^y \tilde{\gamma}_y\|_1 \leq l(\gamma_y^0) + \kappa_n \|\widehat{\Psi}^y \gamma_y^0\|_1$ implies

$$\frac{1}{nT}\sum_{t=1}^T \left(\|\widehat{U}_t \Delta_\gamma\|_2^2 + 2(R_y' \widehat{U}_t' + \tilde{e}_t' + M_t')\widehat{U}_t \Delta_\gamma \right) + \kappa_n \|\widehat{\Psi}^y \tilde{\gamma}_y\|_1 \leq \kappa_n \|\widehat{\Psi}^y \gamma_y^0\|_1. \tag{F.1}$$

Recall that

$$\Psi^y = \operatorname{diag}\left\{ \left(\frac{1}{n}\sum_i W_{im}^2 \right)^{1/2} : m \leq p \right\}$$

where $W_{im} = \frac{1}{\sqrt{T}}\sum_{t=1}^T (U_{it,m} - \bar{U}_{i\cdot,m})(e_{it} - \bar{e}_{i\cdot})$ are independent across i. Hence, by the Bernstein inequality for independent data, for $y = M\sqrt{\frac{\log p}{n}}$ and large enough $M > 0$,

$$P\left(\max_{m \leq p} |\frac{1}{n}\sum_{i=1}^n W_{im}^2 - EW_{im}^2| > y \right) \leq p \max_m P\left(\left| \frac{1}{n}\sum_{i=1}^n W_{im}^2 - EW_{im}^2 \right| > y \right)$$
$$\leq p \exp\left(-\frac{ny^2}{4\max_{im} \operatorname{Var}(W_{im}^2)} \right) \leq \exp\left(\log p - \frac{ny^2}{4C} \right) = o(1).$$

Hence $\max_m |\frac{1}{n}\sum_i (W_{im}^2 - EW_{im}^2)| = o_P(1)$. Also, recall that $\widehat{\Psi}^y$ is defined in (2.9), which satisfies $\|\Psi^y - \widehat{\Psi}^y\|_\infty = o_P(1)$. By the assumption that $\frac{1}{n}\sum_i EW_{im}^2$ is bounded away from zero, we have, $\|(\Psi^y)^{-1}\|_\infty = O_P(1)$. In addition, by Lemma F.1,

$$\left\|(\Psi^y)^{-1}\frac{1}{nT}\sum_{t=1}^T \tilde{U}_t'\tilde{e}_t\right\|_\infty \leq \frac{1}{\sqrt{nT}}\Phi^{-1}\left(1-\frac{q_n}{2p}\right) + o_P\left(\sqrt{\frac{\log p}{nT}}\right)$$

with probability approaching one, and

$$\frac{1}{nT}\sum_{t=1}^T \left(2(R_y'\widehat{U}_t' + M_t')\widehat{U}_t\Delta_\gamma\right) = o_P\left(\sqrt{\frac{\log p}{nT}}\right)\|\Delta_\gamma\|_1.$$

Hence, for any $c_0 > 1$,

$$\begin{aligned}
&|\frac{2}{nT}\sum_{t=1}^T(R_y'\widehat{U}_t' + \tilde{e}_t' + M_t')\widehat{U}_t\Delta_\gamma| \\
&\quad\leq o_P\left(\sqrt{\frac{\log p}{nT}}\right)\|\Psi^{y-1}\|_\infty\|\Psi^y\Delta_\gamma\|_1 + \frac{2}{\sqrt{nT}}\Phi^{-1}\left(1-\frac{q_n}{2p}\right)\|\Psi^y\Delta_\gamma\|_1 \\
&\quad\leq \frac{c_0+1}{\sqrt{nT}}\Phi^{-1}\left(1-\frac{q_n}{2p}\right)\|\Psi^y\Delta_\gamma\|_1 \\
&\quad= \frac{c_0+1}{2c_0}\kappa_n\|\Psi^y\Delta_\gamma\|_1
\end{aligned}$$

with probability approaching one where the second inequality follows from the fact that there is $C > 0$ such that $\Phi^{-1}\left(1-\frac{q_n}{2p}\right) \asymp \sqrt{\log p + \log\frac{1}{q_n}} \geq \sqrt{\log p}$ and we have used the definition

$$\kappa_n = \frac{2c_0}{\sqrt{nT}}\Phi^{-1}\left(1-\frac{q_n}{2p}\right), \quad \log(q_n^{-1}) = O(\log p)$$

in the last equality. We thus also have, with probability approaching one,

$$\begin{aligned}
&|\frac{2}{nT}\sum_{t=1}^T(R_y'\widehat{U}_t' + \tilde{e}_t' + M_t')\widehat{U}_t\Delta_\gamma| \\
&\quad\leq \frac{c_0+1}{2c_0}\kappa_n\|\widehat{\Psi}^y\Delta_\gamma\|_1 + \frac{c_0+1}{2c_0}\kappa_n\|\widehat{\Psi}^y\Delta_\gamma\|_1\|\widehat{\Psi}^y - \Psi^y\|_\infty\|\widehat{\Psi}^{y-1}\|_\infty \\
&\quad\leq \frac{c_0+1}{2c_0}\kappa_n\|\widehat{\Psi}^y\Delta_\gamma\|_1\left(1+\frac{c_0-1}{4c_0}\right) \\
&\quad= \frac{3c_0+1}{4c_0}\kappa_n\|\widehat{\Psi}^y\Delta_\gamma\|_1.
\end{aligned}$$

(F.1) then implies $\frac{1}{nT}\sum_{t=1}^T\|\widehat{U}_t\Delta_\gamma\|_2^2 - \frac{3c_0+1}{4c_0}\kappa_n\|\widehat{\Psi}^y\Delta_\gamma\|_1 + \kappa_n\|\widehat{\Psi}^y\tilde{\gamma}_y\|_1 \leq \kappa_n\|\widehat{\Psi}^y\gamma_y^0\|_1$. Thus

$$\frac{1}{nT}\sum_{t=1}^T\|\widehat{U}_t\Delta_\gamma\|_2^2 + \frac{c_0-1}{4c_0}\kappa_n\|(\widehat{\Psi}^y\Delta_\gamma)_{J^c}\|_1 \leq \frac{7c_0+1}{4c_0}\kappa_n\|(\widehat{\Psi}^y\Delta_\gamma)_J\|_1. \quad \text{(F.2)}$$

Now, note that the sparse eigenvalue assumption 3.4 implies that, for some $\underline{\phi} > 0$,

$$\underline{\phi}(m) := \inf_{\delta \in \mathbb{R}^p : \|\delta_{J^c}\|_1 \leq m \|\delta_J\|_1} \mathcal{R}(\delta) \geq \underline{\phi}.$$

And for some generic $C > 0$, $\|\Delta_\gamma\|_1^2 \leq C\|\Delta_{\gamma,J}\|_1^2 \leq C\|\Delta_\gamma\|_2^2 |J|_0$. Hence $\frac{1}{nT}\sum_{t=1}^T \|\tilde{U}_t \Delta_\gamma\|_2^2 \geq \underline{\phi}\|\Delta_\gamma\|_2^2$. In addition, by Lemma F.1,

$$|J|_0 \max_{m,k\leq p} |\frac{1}{nT}\sum_{it}(\widehat{U}_{it,m}\widehat{U}_{it,k} - \tilde{U}_{it,m}\tilde{U}_{it,k})| = o_P(1).$$

Hence,

$$\begin{aligned}\frac{1}{nT}\sum_{t=1}^T \|\widehat{U}_t\Delta_\gamma\|_2^2 &\geq \frac{1}{nT}\sum_{t=1}^T \|\tilde{U}_t\Delta_\gamma\|_2^2 - \|\Delta_\gamma\|_1^2 \max_{m,k\leq p}|\frac{1}{nT}\sum_{it}(\widehat{U}_{it,m}\widehat{U}_{it,k} - \tilde{U}_{it,m}\tilde{U}_{it,k})| \\ &\geq \underline{\phi}\|\Delta_\gamma\|_2^2 - C\|\Delta_\gamma\|_2^2 |J|_0 \max_{m,k\leq p}|\frac{1}{nT}\sum_{it}(\widehat{U}_{it,m}\widehat{U}_{it,k} - \tilde{U}_{it,m}\tilde{U}_{it,k})| \\ &\geq \underline{\phi}\|\Delta_\gamma\|_2^2/2,\end{aligned}$$

with probability approaching one. Also, $\max_m |\widehat{\Psi}_m^y| = O_P(2)$. So (F.2) implies for some $C > 0$,

$$\|\Delta_\gamma\|_2^2 \leq C\kappa_n \|(\Delta_\gamma)_J\|_1 \leq \sqrt{|J|_0}\kappa_n \|(\Delta_\gamma)_J\|_2 \leq \sqrt{|J|_0}\kappa_n \|\Delta_\gamma\|_2.$$

Hence $\|\Delta_\gamma\|_2 \leq O_P(\sqrt{|J|_0}\kappa_n)$, and $\|\Delta_\gamma\|_1 \leq C\|\Delta_\gamma\|_2\sqrt{|J|_0} \leq O_P(\kappa_n |J|_0)$. Still by (F.2),

$$\frac{1}{nT}\sum_{t=1}^T \|\widehat{U}_t \Delta_\gamma\|_2^2 = O_P(\kappa_n^2 |J|_0).$$

Moreover, $\gamma_y - \tilde{\gamma}_y = \Delta_\gamma + R_y$, and by Lemma F.1, $\max_{m,k\leq p}|\frac{1}{nT}\sum_{it}\widehat{U}_{it,m}\widehat{U}_{it,k}| = O_P(1)$, so $\frac{1}{nT}\sum_{t=1}^T \|\widehat{U}_t R_y\|_2^2 \leq \|R_y\|_1^2 \max_{m,k\leq p}|\frac{1}{nT}\sum_{it}\widehat{U}_{it,m}\widehat{U}_{it,k}| = O_P(1)\|R_y\|_1^2$. We have

$$\frac{1}{nT}\sum_{t=1}^T \|\widehat{U}_t(\gamma_y - \tilde{\gamma}_y)\|_2^2 = O_P(\|R_y\|_1^2 + \kappa_n^2 |J|_0),$$

$$\|\gamma_y - \tilde{\gamma}_y\|_1 = O_P(\kappa_n |J|_0 + \|R_y\|_1).$$

where $\frac{1}{nT}\sum_{t=1}^T \|\widehat{U}_t(\gamma_y - \tilde{\gamma}_y)\|_2^2$ follows from (F.2).

Proof of Proposition F.1 (ii)

We prove the convergence for $|\tilde{\gamma}_y|_0$ only; the result for $|\tilde{\gamma}_d|_0$ follows similarly. Write $\widehat{J} = \mathrm{supp}(\tilde{\gamma}_y)$. Let $\widehat{U}_{it,\widehat{J}}$ be the $|\widehat{J}|_0 \times 1$ subvector of $\widehat{U}_{it}$ whose elements correspond to the indices in $\widehat{J}$. Also, let $\widehat{U}_{t,\widehat{J}}$ be the $n \times |\widehat{J}|_0$ matrix whose i^{th} row is $\widehat{U}'_{it,\widehat{J}}$. In addition, write

$$\widehat{U}_{\widehat{J}} = (\widehat{U}'_{1,\widehat{J}}, ..., \widehat{U}'_{T,\widehat{J}})', \quad nT \times |\widehat{J}|_0.$$

Define $\tilde{U}_{it,\widehat{J}}$, $\tilde{U}_{t,\widehat{J}}$ and $\tilde{U}_{\widehat{J}}$ similarly. Then by (F.15),

$$\begin{aligned}
\|\frac{1}{nT}\sum_{t=1}^{T}\widehat{U}'_{t,\widehat{J}}\widehat{U}_{t,\widehat{J}}\| &= \frac{1}{nT}\|\widehat{U}'_{\widehat{J}}\widehat{U}_{\widehat{J}}\| = \frac{1}{nT}\|\widehat{U}_{\widehat{J}}\|^2 \\
&\leq \frac{2}{nT}\|\tilde{U}_{\widehat{J}}\|^2 + \frac{2}{nT}\|\widehat{U}_{\widehat{J}} - \tilde{U}_{\widehat{J}}\|^2 \\
&= \|\frac{2}{nT}\sum_{t=1}^{T}\tilde{U}'_{t,\widehat{J}}\tilde{U}_{t,\widehat{J}}\| + \|\frac{2}{nT}\sum_{t=1}^{T}(\widehat{U}_{t,\widehat{J}} - \tilde{U}_{t,\widehat{J}})'(\widehat{U}_{t,\widehat{J}} - \tilde{U}_{t,\widehat{J}})\| \\
&\leq \|\frac{2}{nT}\sum_{t=1}^{T}\tilde{U}'_{t,\widehat{J}}\tilde{U}_{t,\widehat{J}}\| + \frac{2}{nT}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{m\in\widehat{J}}(\widehat{U}_{it,m} - \tilde{U}_{it,m})^2 \\
&\leq 2\phi_{\max}(|\widehat{J}|_0) + O_P(|\widehat{J}|_0\Delta_F^2 + |\widehat{J}|_0\frac{\log(pT)}{n}), \qquad \text{(F.3)}
\end{aligned}$$

where the last inequality follows from (F.15).

Let $\Psi^y(\widehat{J})$ be a $|\widehat{J}|_0 \times 1$ subvector of the diagonal entries of Ψ^y, with elements in $\widehat{J}$. Then, using the Karush-Kuhn-Tucker (KKT) condition and recalling that $\tilde{Y}_t = \tilde{F}\tilde{\delta}_{yt} + \tilde{U}_t\gamma_y + \tilde{e}_t$, and that $M_t := \tilde{F}\tilde{\delta}_{yt} - \widehat{F}\widehat{\delta}_{yt} + (\tilde{U}_t - \widehat{U}_t)\gamma_y$, we have that, with probability approaching one,

$$\begin{aligned}
\kappa_n\|\widehat{\Psi}^y(\widehat{J})\|_2 &= \|\frac{2}{nT}\sum_t \widehat{U}'_{t,\widehat{J}}(\tilde{Y}_t - \widehat{F}\widehat{\delta}_{yt} - \widehat{U}_t\tilde{\gamma}_y)\|_2 \\
&= \|\frac{2}{nT}\sum_t \widehat{U}'_{t,\widehat{J}}(M_t + \tilde{e}_t + \widehat{U}_t(\gamma_y - \tilde{\gamma}_y))\|_2 \\
&\leq \frac{2}{nT}\|\sum_t \widehat{U}'_{t,\widehat{J}}\widehat{U}_{t,\widehat{J}}\|^{1/2}(\sum_t \|\widehat{U}_t\gamma_y - \widehat{U}_t\tilde{\gamma}_y\|_2^2)^{1/2} \\
&\quad + \|\widehat{\Psi}^y(\widehat{J})\|_2\|\widehat{\Psi}^{y-1}_{\widehat{J}}\frac{2}{nT}\sum_t \widehat{U}'_{t,\widehat{J}}(M_t + \tilde{e}_t)\|_\infty \\
&\leq (2\phi_{\max}(|\widehat{J}|_0) + O_P\left(|\widehat{J}|_0\Delta_F^2 + |\widehat{J}|_0\frac{\log(pT)}{n}\right)^{1/2} O_P(\|R_y\|_1^2 + \kappa_n^2|J|_0)^{1/2} \\
&\quad + \|\widehat{\Psi}^y(\widehat{J})\|_2\frac{3c_0+1}{4c_0}\kappa_n,
\end{aligned}$$

where the last inequality follows from (F.3), Lemma F.1, and Proposition F.1. We then have

$$\kappa_n|\widehat{J}|_0^{1/2} \leq O_P\left(2\phi_{\max}(|\widehat{J}|_0) + |\widehat{J}|_0\Delta_F^2 + |\widehat{J}|_0\frac{\log(pT)}{n}\right)^{1/2}(\|R_y\|_1^2 + \kappa_n^2|J|_0)^{1/2}. \quad \text{(F.4)}$$

Case 1: $\phi_{\max}(|\widehat{J}|_0) < |\widehat{J}|_0\Delta_F^2 + |\widehat{J}|_0\frac{\log(pT)}{n}$. In this case, (F.4) implies

$$\kappa_n^2 \leq O_P\left(\Delta_F^2 + \frac{\log(pT)}{n}\right)\|R_y\|_1^2 + O_P\left(\Delta_F^2 + \frac{\log(pT)}{n}\right)\kappa_n^2|J|_0.$$

However, this result contradicts the assumptions that $(\Delta_F^2 + \frac{\log(pT)}{n})|J|_0 = o(1)$ and that $(\Delta_F^2 + \frac{\log(pT)}{n})\|R_y\|_1^2 = o(\kappa_n^2)$. Thus, we have that this case cannot happen.

Case 2: $\phi_{\max}(|\widehat{J}|_0) \geq |\widehat{J}|_0\Delta_F^2 + |\widehat{J}|_0\frac{\log(pT)}{n}$. In this case, (F.4) implies

$$|\widehat{J}|_0 \leq O_P(\|R_y\|_1^2/\kappa_n^2 + |J|_0)\phi_{\max}(|\widehat{J}|_0) = O_P(|J|_0\phi_{\max}(|\widehat{J}|_0))$$

where the equality is due to $\|R_y\|_1^2 = O(\kappa_n^2|J|_0)$. By Assumption 3.4, there is an $l_n \to \infty$ such that $\phi_{\max}(l_n|J|_0) < c_2$. Now, let $g_n = l_n|J|_0$. We now show $g_n \geq |\widehat{J}|_0$. Suppose otherwise, and let $l = |\widehat{J}|_0/g_n$. Then $l \geq 1$. Now by Lemma 3 of Belloni and Chernozhukov (2013), $\phi_{\max}([lg_n]) \leq [l]\phi_{\max}(g_n)$ for $l \geq 1$. Hence, we have, with probability arbitrarily close to one, that there is $C > 0$ such that

$$|\widehat{J}|_0 \leq C\phi_{\max}(|\widehat{J}|_0)|J|_0 \leq C\phi_{\max}(g_n)|J|_0\frac{|\widehat{J}|_0}{g_n} < Cc_2|J|_0\frac{|\widehat{J}|_0}{g_n},$$

implying $l_n|J|_0 = g_n \leq Cc_2|J|_0$. However, this result contradicts $l_n \to \infty$. Hence, $g_n \geq |\widehat{J}|_0$. Then, we have $\phi_{\max}(|\widehat{J}|_0) \leq \phi_{\max}(g_n)$ because $\phi_{\max}(.)$ is non-decreasing. Thus,

$$|\widehat{J}|_0 \leq C\phi_{\max}(|\widehat{J}|_0)|J|_0 \leq C\phi_{\max}(g_n)|J|_0 < Cc_2|J|_0.$$

Lemma F.1. *Let* $M_t = \tilde{F}\tilde{\delta}_{xt} - \widehat{F}\widehat{\delta}_{xt} + (\tilde{U}_t - \widehat{U}_t)\gamma_x$*, with* $x = y$ *or* $x = d$*. Then*
(i) $\|\Psi^{y-1}\frac{1}{nT}\sum_{t=1}^T\tilde{U}_t'\tilde{e}_t\|_\infty \leq \frac{1}{\sqrt{nT}}\Phi^{-1}(1 - \frac{q_n}{2p}) + o_P(\sqrt{\frac{\log p}{nT}})$ *with probability approaching one, provided* $\log p + \log q_n^{-1} = o(n^{1/3})$*.*
(ii) $\|\frac{1}{nT}\sum_{t=1}^T(\tilde{U}_t - \widehat{U}_t)'\tilde{e}_t\|_\infty = o_P(\sqrt{\frac{\log p}{nT}})$*.*
(iii) $\|\frac{1}{nT}\sum_{t=1}^T M_t'\widehat{U}_t\|_\infty = o_P(\sqrt{\frac{\log p}{nT}})$*,*
(iv) $|J|_0\max_{m,k\leq p}|\frac{1}{nT}\sum_{it}(\widehat{U}_{it,m}\widehat{U}_{it,k} - \tilde{U}_{it,m}\tilde{U}_{it,k})| = o_P(1)$*.*
(v) $\max_{k,m}|\frac{1}{nT}\sum_t\sum_{i=1}^n\widehat{U}_{it,k}\widehat{U}_{it,m}| = O_P(1)$*.*
(vi) $\|\frac{1}{nT}\sum_t R_x'\widehat{U}_t'\widehat{U}_t\|_\infty = o_P(\sqrt{\frac{\log p}{nT}})$*.*

Proof. (i) We have $\frac{1}{nT}\sum_{t=1}^T \tilde{U}_t'\tilde{e}_t = \frac{1}{nT}\sum_{t=1}^T\sum_{i=1}^n (U_{it}-\bar{U}_{i\cdot})(e_{it}-\bar{e}_{i\cdot}) - \frac{1}{T}\sum_{t=1}^T \bar{U}_{\cdot t}\bar{e}_{\cdot t} + \bar{\bar{U}}\bar{\bar{e}}$. Define $W_{im} = \frac{1}{\sqrt{T}}\sum_{t=1}^T (U_{it,m}-\bar{U}_{i\cdot,m})(e_{it}-\bar{e}_{i\cdot})$. Then by Assumption 3.2,

$$\max_{i\le n, m\le p}\frac{(E|W_{im}|^3)^{1/3}}{(EW_{im}^2)^{1/2}} \le C, \quad EW_{im}=0.$$

Now note that $\Psi_m^y = \sqrt{\frac{1}{n}\sum_i W_{im}^2}$. It follows from Lemma 5 of Belloni et al. (2012),

$$P\left(\max_{m\le p}\frac{|\frac{1}{n}\sum_i W_{im}|}{\sqrt{\frac{1}{n}\sum_i W_{im}^2}} < \frac{1}{\sqrt{n}}\Phi^{-1}\left(1-\frac{q_n}{2p}\right)\right) \le q_n(1+C/l_n^3)$$

for some $l_n\to\infty$, provided that $\log p + \log q_n^{-1} = o(n^{1/3}/l_n^2)$. Hence as long as $\log p + \log q_n^{-1} = o(n^{1/3})$, such $l_n\to\infty$ always exist, and thus $\max_{m\le p}\frac{|\frac{1}{n}\sum_i W_{im}|}{\Psi_m^y} < \frac{1}{\sqrt{n}}\Phi^{-1}(1-\frac{q_n}{2p})$ with probability approaching one. Hence, with probability approaching one,

$$\begin{aligned}\|(\Psi^y)^{-1}\frac{1}{nT}\sum_{t=1}^T\sum_{i=1}^n (U_{it}-\bar{U}_{i\cdot})(e_{it}-\bar{e}_{i\cdot})\|_\infty &= \frac{1}{\sqrt{T}}\max_m (\Psi_m^y)^{-1}|\frac{1}{n}\sum_i W_{im}| \\ &< \frac{1}{\sqrt{nT}}\Phi^{-1}\left(1-\frac{q_n}{2p}\right).\end{aligned}$$

In addition, we show $\|\frac{1}{T}\sum_{t=1}^T \bar{U}_{\cdot t}\bar{e}_{\cdot t}\|_\infty = o_P\left(\sqrt{\frac{\log p}{nT}}\right)$. If T does not grow, then

$$\|\frac{1}{T}\sum_t \bar{U}_{\cdot t}\bar{e}_{\cdot t}\|_\infty \le \max_t \|\bar{U}_{\cdot t}\|_\infty \max_t |\bar{e}_{\cdot t}| \le O_P(\frac{\sqrt{\log p}}{n}) = o_P\left(\sqrt{\frac{\log p}{nT}}\right)$$

where the last equality follows since $T\ge 1$ is a fixed constant.

On the other hand, if $T\to\infty$, then for any $k\le p$, let $\mathcal{X}_{tk} = \bar{U}_{\cdot t,k}\bar{e}_{\cdot t}\operatorname{Var}(\bar{U}_{\cdot t,k}\bar{e}_{\cdot t})^{-1/2}$, which satisfies the strong mixing condition across t (Assumption 3.1). To apply the Bernstein inequality for strong mixing data, let $\bar{r}=\min\{r,1\}$, $r_1=(0.5+\bar{r}^{-1})^{-1}$, $c=0.5(\gamma+1)$, then $\gamma\bar{r}\ge 2$, $r_1<1$, $cr_1>1$ and $2c\ge 1$. Because $\bar{r}\le r$, the strong mixing condition in Assumption 3.1 also holds with $\bar{r}$ in place of r. The Bernstein inequality for weakly dependent data of Merlevède et al. (2011) requires (a) the exponential tails (fulfilled by the sub-Gaussian condition in Assumption 3.1), and (b) the strong mixing condition (also assumed in Assumption 3.1). The introduction of $\bar{r}$ is to ensure that the so-defined $r_1<1$, another requirement of applying the Bernstein inequality for strong mixing sequences. Then by Theorem 1 of Merlevède et al. (2011) (proved using a "coupling argument" of Dedecker and Prieur (2004) - see also Wang et al. (2016) for a similar argument), for $y = M\frac{(\log p)^c}{\sqrt{T}}$, and sufficiently large $M>0$, we have

$$P\left(\max_{k\leq p}\left|\frac{1}{T}\sum_t \mathcal{X}_{tk}\right|>y\right)\leq p\max_{k\leq p}P\left(\left|\frac{1}{T}\sum_t \mathcal{X}_{tk}\right|>y\right)\leq A_1+A_2+A_3$$

where

$$\begin{aligned}
A_1 &= pT\exp(-C(Ty)^{r_1})=\exp(\log(pT)-CM^{r_1}T^{r_1/2}\log^{cr_1}p)=o(1), \quad (cr_1\geq 1)\\
A_2 &= p\exp\left(-C\frac{(Ty)^2}{T}\exp\left(\frac{(Ty)^{r_1(1-r_1)}}{C\log^{r_1}(Ty)}\right)\right)=o(1), \quad (2c>1, Ty\to\infty)\\
A_3 &= p\exp\left(-CTy^2\right)=\exp(\log p-CM^2\log^{2c}p)=o(1).
\end{aligned}$$

Hence $\max_{k\leq p}|\frac{1}{T}\sum_t\mathcal{X}_{tk}|=O_P\left(\frac{(\log p)^c}{\sqrt{T}}\right)$. In addition, $\max_k \mathrm{Var}(\bar U_{\cdot t,k}\bar e_{\cdot t})=O(n^{-2})$. Hence

$$\|\frac{1}{T}\sum_t \bar U_{\cdot t}\bar e_{\cdot t}\|_\infty\leq \max_{k\leq p}|\frac{1}{T}\sum_t\mathcal{X}_{tk}|\sqrt{\max_k \mathrm{Var}(\bar U_{\cdot t,k}\bar e_{\cdot t})}=O_P\left(\frac{(\log p)^c}{\sqrt{T}}\frac{1}{n}\right)=o_P\left(\sqrt{\frac{\log p}{nT}}\right)$$

where the last equality is due to $\log^{2c-1}p=\log^\gamma p=o(n)$. Also, $\|\bar{\bar U}\bar{\bar e}\|_\infty=o_P\left(\sqrt{\frac{\log p}{T}}\frac{1}{n}\right)$. Thus,

$$\|\frac{1}{T}\sum_{t=1}^T \bar U_{\cdot t}\bar e_{\cdot t}\|_\infty+\|\bar{\bar U}\bar{\bar e}\|_\infty=O_P\left(\sqrt{\frac{\log p}{T}}\frac{1}{n}\right)=o_P\left(\sqrt{\frac{\log p}{nT}}\right).$$

(ii) Throughout the paper, we denote $(\widehat\lambda_{t1},...,\widehat\lambda_{tp})$ as the columns of the $K\times p$ matrix $\widehat\Lambda_t'$. Also, denote $(\lambda_{t1},...,\lambda_{tp})$ as the columns of the $K\times p$ matrix Λ_t'. Note that for each $m\leq p$, $\widehat\lambda_{tm}$ estimates the demeaned loading $\tilde\lambda_{tm}$ up to a matrix transformation.

For $m\leq p$, $\tilde U_{it,m}-\widehat U_{it,m}=(\widehat\lambda_{tm}'-\tilde\lambda_{tm}'(H')^{-1})\widehat f_i+\tilde\lambda_{tm}'(H')^{-1}(\widehat f_i-H'\tilde f_i)$. Also, note $\widehat\lambda_{tm}-H^{-1}\tilde\lambda_{tm}=(\widehat F'\widehat F)^{-1}\widehat F'(\tilde FH-\widehat F)H^{-1}\tilde\lambda_{tm}+(\widehat F'\widehat F)^{-1}\sum_{i=1}^n\widehat f_i\tilde U_{it,m}$. Hence,

$$\begin{aligned}
\|\frac{1}{nT}\sum_{t=1}^T(\tilde U_t-\widehat U_t)'\tilde e_t\|_\infty &= \|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T\tilde e_{it}(\tilde U_{it}-\widehat U_{it})\|_\infty\\
&\leq \max_m|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T\tilde e_{it}\widehat f_i'(\widehat\lambda_{tm}-H^{-1}\tilde\lambda_{tm})|\\
&\quad +O_P(1)\max_m\|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T\tilde e_{it}(\widehat f_i-H'\tilde f_i)\tilde\lambda_{tm}'\|_F\\
&\leq \max_m\|(\widehat F'\widehat F)^{-1}\widehat F'(\tilde FH-\widehat F)H^{-1}\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T\tilde\lambda_{tm}\tilde e_{it}\widehat f_i'\|_F
\end{aligned}$$

$$
\begin{aligned}
&+ \max_m \|(\widehat{F}'\widehat{F})^{-1}\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T\sum_{j=1}^n \widehat{f}_j\tilde{U}_{jt,m}\tilde{e}_{it}\widehat{f}_i'\|_F + o_P\left(\sqrt{\frac{\log p}{nT}}\right) \\
&\leq o_P(1)\max_m \|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T \tilde{\lambda}_{tm}\tilde{e}_{it}\tilde{f}_i'\|_F \\
&+ \max_m \|\frac{1}{n^2T}\sum_{i=1}^n\sum_{t=1}^T\sum_{j=1}^n \widehat{f}_j\tilde{U}_{jt,m}\tilde{e}_{it}\widehat{f}_i'\|_F + o_P\left(\sqrt{\frac{\log p}{nT}}\right),
\end{aligned} \tag{F.5}
$$

where we used Assumption D.4 that

$$
\max_m \|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T \tilde{e}_{it}(\widehat{f}_i - H'\tilde{f}_i)\tilde{\lambda}_{tm}'\|_F = o_P\left(\sqrt{\frac{\log p}{nT}}\right).
$$

We now respectively bound the first two terms in (F.5). For the first term, we have

$$
\max_{m\leq p} \|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T \tilde{\lambda}_{tm}\tilde{e}_{it}\tilde{f}_i'\|_F = \max_m \|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T \tilde{\lambda}_{tm}\tilde{e}_{it}f_i'\|_F \leq l_1 + ... + l_4,
$$

for $l_1 - l_4$ defined in the following.

Let $W_{im} = \frac{1}{T}\sum_{t=1}^T \tilde{\lambda}_{tm}e_{it}f_i'$. For simplicity, we assume $\dim(f_i) = \dim(\tilde{\lambda}_{im}) = 1$ so W_{im} is a scalar. The multivariate case follows from an argument that investigating it element-by-element. By the strong mixing condition, $\max_{im}\operatorname{Var}(W_{im}) \leq C\frac{1}{T}$ for some $C > 0$. Using the Bernstein inequality for independent data, we can obtain, for $y = M\sqrt{\frac{\log p}{nT}}$ and large enough $M > 0$,

$$
\begin{aligned}
&P\left(\max_{m\leq p}|\frac{1}{n}\sum_{i=1}^n W_{im}| > y\right) \leq p\max_m P(|\frac{1}{n}\sum_{i=1}^n W_{im}| > y) \\
&\leq p\exp\left(-\frac{ny^2}{4\max_{im}\operatorname{Var}(W_{im})}\right) \leq \exp\left(\log p - \frac{nTy^2}{4C}\right) = o(1).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
l_1 &= \max_m \|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T \tilde{\lambda}_{tm}e_{it}f_i'\|_F = \max_m \|\frac{1}{n}\sum_{i=1}^n W_{im}\|_F = O_P\left(\sqrt{\frac{\log p}{nT}}\right) \\
l_2 &= \max_m \|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T \tilde{\lambda}_{tm}\bar{e}_{i\cdot}f_i'\|_F = O_P(1)\|\frac{1}{n}\sum_{i=1}^n \bar{e}_{i\cdot}f_i'\|_F \\
&= O_P\left(\max_{kl\leq K}(\frac{1}{n^2T^2}\sum_i\sum_{ts} Ef_{ik}f_{il}e_{it}e_{is})^{1/2}\right)
\end{aligned}
$$

$$\leq O_P\left(\max_{it}\sup_{f_i}\sum_s |E(e_{it}e_{is}|f_i)|\right)^{1/2} O_P\left(\frac{1}{\sqrt{nT}}\right)\max_{ik}(Ef_{ik}^2)^{1/2} = O_P\left(\frac{1}{\sqrt{nT}}\right),$$

in the last equality we used $\max_{it}\sup_{f_i}\sum_s |E(e_{it}e_{is}|f_i)| = O_P(1)$, because given $\{f_i\}$, the process $\{e_t = \alpha\eta_t + \epsilon_t\}_{t\leq T}$ is weakly dependent satisfying the strong mixing condition in Assumption 3.1.

$$l_3 = \max_m \|\frac{1}{T}\sum_{t=1}^T \tilde{\lambda}_{tm}\bar{e}_{\cdot t}\bar{f}'\|_F$$

$$= O_P(1)\max_m \|\frac{1}{nT}\sum_t\sum_i \tilde{\lambda}_{tm}e_{it}\|_2 = O_P\left(\sqrt{\frac{\log p}{nT}}\right)$$

$$l_4 = \max_m \|\frac{1}{T}\sum_{t=1}^T \tilde{\lambda}_{tm}\bar{\bar{e}}\bar{f}'\|_F = O_P(1)|\bar{\bar{e}}| = O_P\left(\frac{1}{\sqrt{nT}}\right).$$

Here l_3 follows from the same argument as l_1, using the Bernstein inequality for independent data. Hence, $o_P(1)\max_m \|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T \tilde{\lambda}_{tm}\tilde{e}_{it}\tilde{f}_i'\|_F = o_P\left(\sqrt{\frac{\log p}{nT}}\right)$. For the second term, we have

$$\max_m \|\frac{1}{n^2T}\sum_{i=1}^n\sum_{t=1}^T\sum_{j=1}^n \widehat{f}_j\tilde{U}_{jt,m}\tilde{e}_{it}\widehat{f}_i'\|_F \leq a_1 + ... + a_4,$$

where, up to a $\|H\| = O_P(1)$ term,

$$a_1 = \max_m \|\frac{1}{n}\sum_{j=1}^n(\widehat{f}_j - H'\tilde{f}_j)\frac{1}{nT}\sum_{t=1}^T\sum_{i=1}^n \tilde{U}_{jt,m}\tilde{e}_{it}(\widehat{f}_i - H'\tilde{f}_i)'\|_F$$

$$\leq \left(\frac{1}{T}\sum_t \|\frac{1}{n}\sum_{j=1}^n(\widehat{f}_j - H'\tilde{f}_j)\tilde{U}_{jt,m}\|_2^2\right)^{1/2}\left(\frac{1}{T}\sum_t \|\frac{1}{n}\sum_{i=1}^n \tilde{e}_{it}(\widehat{f}_i - H'\tilde{f}_i)'\|_2^2\right)^{1/2}$$

$$= O_P\left(\Delta_{fum}\Delta_{fe}\right) =^{(a)} o_P\left(\sqrt{\frac{\log(p)}{nT}}\right)$$

$$a_2 = \max_m \|\frac{1}{n^2T}\sum_{i=1}^n\sum_{t=1}^T\sum_{j=1}^n \tilde{f}_j\tilde{U}_{jt,m}\tilde{e}_{it}\tilde{f}_i'\|_F \leq^{(b)} o_P\left(\sqrt{\frac{\log(p)}{nT}}\right),$$

$$a_3 = \max_m \|\frac{1}{n^2T}\sum_{i=1}^n\sum_{t=1}^T\sum_{j=1}^n \tilde{f}_j\tilde{U}_{jt,m}\tilde{e}_{it}(\widehat{f}_i - H'\tilde{f}_i)'\|_F$$

$$\leq \max_{mt} \|\frac{1}{n}\sum_{j=1}^{n}\tilde{f}_j\tilde{U}_{jt,m}\|_2 \left(\frac{1}{T}\sum_{t=1}^{T}\|\frac{1}{n}\sum_{i=1}^{n}\tilde{e}_{it}(\widehat{f}_i - H'\tilde{f}_i)\|_2^2\right)^{1/2}$$

$$=^{(c)} O_P\left(\sqrt{\frac{\log(pT)}{n}}\right) o_P\left(\sqrt{\frac{\log p}{T\log(pT)}}\right)$$

$$a_4 = \max_{m} \|\frac{1}{n^2T}\sum_{i=1}^{n}\sum_{t=1}^{T}\sum_{j=1}^{n}(\widehat{f}_j - H'\tilde{f}_j)\tilde{U}_{jt,m}\tilde{e}_{it}\tilde{f}_i'\|_F$$

$$\leq \max_{mt} \|\frac{1}{n}\sum_{j=1}^{n}(\widehat{f}_j - H'\tilde{f}_j)\tilde{U}_{jt,m}\|_2 \left(\frac{1}{T}\sum_{t}\|\frac{1}{n}\sum_{i=1}^{n}\tilde{e}_{it}\tilde{f}_i\|_2^2\right)^{1/2}$$

$$=^{(d)} O_P\left(\frac{1}{\sqrt{n}}\right) o_P\left(\sqrt{\frac{\log p}{T}}\right),$$

where (a) follows from $\Delta_{fe}\Delta_{fum} \leq \Delta_F^2 = o\left(\sqrt{\frac{\log(p)}{nT}}\right)$ due to $\frac{1}{T}\sum_{t=1}^{T}\|\frac{1}{n}\sum_{i=1}^{n}\tilde{e}_{it}(\widehat{f}_i - H'\tilde{f}_i)\|_2^2 := O_P(\Delta_{fe}^2)$ and $\max_{mt}\|\frac{1}{n}\sum_{j=1}^{n}(\widehat{f}_j - H'\tilde{f}_j)\tilde{U}_{jt,m}\|_2 := O_P(\Delta_{fum})$; (b) follows from Lemma H.12; (c) follows using the assumption $\Delta_{fe}^2 = o\left(\frac{\log p}{T\log(pT)}\right)$; and (d) uses the assumption $\Delta_{fum} = o\left(\sqrt{\frac{\log p}{T}}\right)$.

(iii) For, $M_t = \tilde{F}\tilde{\delta}_{xt} - \widehat{F}\widehat{\delta}_{xt} + (\tilde{U}_t - \widehat{U}_t)\gamma_x$, we have

$$\|\frac{1}{nT}\sum_{t=1}^{T}M_t'\widehat{U}_t\|_\infty = \|\frac{1}{n}\sum_{i=1}^{n}\widehat{U}_{it}\frac{1}{T}\sum_{t=1}^{T}M_{it}\|_\infty$$

$$\leq \|\frac{1}{n}\sum_{i=1}^{n}\widehat{U}_{it}\frac{1}{T}\sum_{t=1}^{T}(\widehat{\delta}_{xt} - H^{-1}\tilde{\delta}_{xt})'\widehat{f}_i\|_\infty \tag{F.6}$$

$$+ \|\frac{1}{n}\sum_{i=1}^{n}\tilde{U}_{it}\frac{1}{T}\sum_{t=1}^{T}\tilde{\delta}_{xt}'(H')^{-1}(\widehat{f}_i - H'\tilde{f}_i)\|_\infty \tag{F.7}$$

$$+ \|\frac{1}{n}\sum_{i=1}^{n}(\widehat{U}_{it} - \tilde{U}_{it})\frac{1}{T}\sum_{t=1}^{T}\tilde{\delta}_{xt}'(H')^{-1}(\widehat{f}_i - H'\tilde{f}_i)\|_\infty \tag{F.8}$$

$$+ \|\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(\widehat{U}_{it} - \tilde{U}_{it})(\tilde{U}_{it} - \widehat{U}_{it})'\gamma_x\|_\infty \tag{F.9}$$

$$+ \|\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{U}_{it}(\tilde{U}_{it} - \widehat{U}_{it})'\gamma_x\|_\infty. \tag{F.10}$$

(F.6) is $\|\frac{1}{n}\widehat{U}_t'\widehat{F}\frac{1}{T}\sum_{t=1}^T(\widehat{\delta}_{xt} - H^{-1}\tilde{\delta}_{xt})\|_\infty = 0$ due to Lemma H.10. (F.7) is $o_P\left(\sqrt{\frac{\log p}{nT}}\right)$ due to the assumption $\Delta_{ud} = o\left(\sqrt{\frac{\log p}{nT}}\right)$. Terms (F.8)-(F.10) are all $o_P\left(\sqrt{\frac{\log p}{nT}}\right)$ due to Lemma H.13. Hence $\|\frac{1}{nT}\sum_{t=1}^T M_t'\widehat{U}_t\|_\infty = o_P\left(\sqrt{\frac{\log p}{nT}}\right)$.

(iv) We have

$$\max_{m,k\leq p}|\frac{1}{nT}\sum_{it}(\widehat{U}_{it,m}\widehat{U}_{it,k} - \tilde{U}_{it,m}\tilde{U}_{it,k})|$$

$$\leq \max_{m,k\leq p}|\frac{2}{nT}\sum_{it}\tilde{U}_{it,m}(\widehat{U}_{it,k} - \tilde{U}_{it,k})| \tag{F.11}$$

$$+ \max_{m,k\leq p}|\frac{1}{nT}\sum_{it}(\widehat{U}_{it,m} - \tilde{U}_{it,m})(\widehat{U}_{it,k} - \tilde{U}_{it,k})|. \tag{F.12}$$

We also have the equality

$$\tilde{U}_{it,m} - \widehat{U}_{it,m} = \widehat{f}_i'(\widehat{F}'\widehat{F})^{-1}\widehat{F}'(\tilde{F}H - \widehat{F})H^{-1}\tilde{\lambda}_{tm} + \widehat{f}_i'(\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{U}_{t,m} + \tilde{\lambda}_{tm}'H'^{-1}(\widehat{f}_i - H'\tilde{f}_i).$$

Hence, we may bound (F.11) by

$$|J|_0\max_{m,k\leq p}|\frac{1}{nT}\sum_{it}\tilde{U}_{it,m}(\widehat{U}_{it,k} - \tilde{U}_{it,k})|$$

$$\leq |J|_0\max_{m,k\leq p}|\frac{1}{nT}\sum_{it}\tilde{U}_{it,k}\widehat{f}_i'\tilde{\lambda}_{tm}|O_P(\Delta_F) + |J|_0\max_{m,k\leq p}|\frac{1}{nT}\sum_{it}\tilde{U}_{it,k}\widehat{f}_i'(\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{U}_{t,m}|$$

$$+ |J|_0\max_{m,k\leq p}|\frac{1}{nT}\sum_{it}\tilde{U}_{it,k}\tilde{\lambda}_{tm}'H'^{-1}(\widehat{f}_i - H'\tilde{f}_i)|$$

$$\leq O_P\left(\Delta_{ud} + \sqrt{\frac{\log p}{nT}}\right)|J|_0(\Delta_F) + |J|_0\max_{m\leq p}\frac{1}{T}\sum_t\|\frac{1}{n}\widehat{F}'\tilde{U}_{t,m}\|_2^2 + O_P(|J|_0\Delta_{ud})$$

$$= o_P(1) + O_P\left(\frac{\log(pT)}{n} + \Delta_{fum}^2\right)|J|_0 = o_P(1)$$

where we used Lemma H.11 (i) that $\max_{m\leq p}\left(\frac{1}{n^2T}\sum_t\|\tilde{U}_{t,m}'\widehat{F}\|_2^2\right) = O_P\left(\frac{\log(pT)}{n} + \Delta_{fum}^2\right)$, Lemma H.11 (vi) that $\max_{m,k\leq p}|\frac{1}{nT}\sum_{it}\tilde{U}_{it,k}\widehat{f}_i'\tilde{\lambda}_{tm}| = O_P\left(\Delta_{ud} + \sqrt{\frac{\log p}{nT}}\right)$, and the assumptions that $\left(\sqrt{\frac{\log p}{nT}}\right)|J|_0\Delta_F + \left(\frac{\log(pT)}{n} + \Delta_{fum}^2\right)|J|_0 = o(1)$ and that $|J|_0\Delta_{ud} \leq |J|_0\Delta_F = o(1)$. The last inequality is due to the following argument:

Proof of $\Delta_{ud} \leq \Delta_F$: By definition, Δ_{ud}, Δ_F are sequences such that

$$\max_{m,k\leq p} \|\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(\widehat{f}_i - H'\tilde{f}_i)\tilde{U}_{it,m}\tilde{h}'_{tk}\|_F = O_P(\Delta_{ud}), \quad \frac{1}{n}\sum_{i=1}^{n}\|\widehat{f}_i - H'\tilde{f}_i\|_2^2 = O_P(\Delta_F^2).$$

where $\tilde{h}_{tk} \in \{\tilde{\delta}_{dt}, \tilde{\delta}_{yt}, \tilde{\lambda}_{tk}\}$ is bounded. Using the Cauchy-Schwarz inequality, we have

$$\max_{m,k\leq p} \|\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(\widehat{f}_i - H'\tilde{f}_i)\tilde{U}_{it,m}\tilde{h}'_{tk}\|_F \leq O_P(\Delta_F)(\max_{m\leq p}\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{U}_{it,m}^2)^{1/2}. \tag{F.13}$$

In the proof of Lemma 4.1 below, we show that

$$\max_{mv}|\frac{1}{nT}\sum_i\sum_t \tilde{U}_{it,m}\tilde{U}_{it,v} - \frac{1}{nT}\sum_i\sum_t(U_{it,m}-\bar{U}_{i\cdot,m})(U_{it,v}-\bar{U}_{i\cdot,v})| = o_P(1)$$

and

$$\max_{mv}\left|\frac{1}{nT}\sum_i\sum_t(U_{it,m}-\bar{U}_{i\cdot,m})(U_{it,v}-\bar{U}_{i\cdot,v}) - E\left[\frac{1}{nT}\sum_i\sum_t(U_{it,m}-\bar{U}_{i\cdot,m})(U_{it,v}-\bar{U}_{i\cdot,v})\right]\right| = o_P(1).$$

In addition, $\max_{mv}|E\left[\frac{1}{nT}\sum_i\sum_t(U_{it,m}-\bar{U}_{i\cdot,m})(U_{it,v}-\bar{U}_{i\cdot,v})\right]| = O(1)$ due to Assumption 3.1. Hence

$$\max_{mv}|\frac{1}{nT}\sum_i\sum_t \tilde{U}_{it,m}\tilde{U}_{it,v} \leq o_P(1)+\max_{mv}\left|E\left[\frac{1}{nT}\sum_i\sum_t(U_{it,m}-\bar{U}_{i\cdot,m})(U_{it,v}-\bar{U}_{i\cdot,v})\right]\right| = O_P(1). \tag{F.14}$$

It also implies $\max_{m\leq p}\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{U}_{it,m}^2 = O_P(1)$. Hence (F.13) implies

$$\max_{m,k\leq p} \|\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(\widehat{f}_i - H'\tilde{f}_i)\tilde{U}_{it,m}\tilde{h}'_{tk}\|_F \leq O_P(\Delta_F).$$

Then by the definition of Δ_{ud}, we can take Δ_{ud} to be at least Δ_F in order to upper bound $\max_{m,k\leq p}\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(\widehat{f}_i - H'\tilde{f}_i)\tilde{U}_{it,m}\tilde{h}'_{tk}\|_F$, this means Δ_F is the least favorable choice for Δ_{ud}. Hence $\Delta_{ud} \leq \Delta_F$.

As for (F.12), we have

$$\begin{aligned}|J|_0 \max_{m,k\leq p}|\frac{1}{nT}\sum_{it}(\widehat{U}_{it,m}-\tilde{U}_{it,m})(\widehat{U}_{it,k}-\tilde{U}_{it,k})| &\leq |J|_0 O_P\left(\Delta_F^2 + \frac{\log(pT)}{n}\right) \\ &= o_P(1).\end{aligned} \tag{F.15}$$

The desired result follows.

In fact, we have just shown

$$\begin{aligned}\max_{m,k\leq p}|\frac{1}{nT}\sum_{it}\tilde{U}_{it,m}(\widehat{U}_{it,k}-\tilde{U}_{it,k})| &\leq O_P\left(\sqrt{\frac{\log p}{nT}}\Delta_F+\Delta_{ud}+\frac{\log(pT)}{n}+\Delta_{fum}^2\right)\\ \max_{m,k\leq p}|\frac{1}{nT}\sum_{it}\widehat{U}_{it,m}(\widehat{U}_{it,k}-\tilde{U}_{it,k})| &\leq O_P\left(\sqrt{\frac{\log p}{nT}}\Delta_F+\Delta_{ud}+\frac{\log(pT)}{n}+\Delta_{fum}^2\right).\end{aligned} \tag{F.16}$$

(v) By part (iv), it suffices to prove $\max_{k,m}|\frac{1}{nT}\sum_t\sum_{i=1}^n\tilde{U}_{it,k}\tilde{U}_{it,m}|=O_P(1)$.

Hence $\max_{k,m}|\frac{1}{nT}\sum_t\sum_{i=1}^n\widehat{U}_{it,k}\widehat{U}_{it,m}|\leq o_P(1)+\max_{mk}|E\left[\frac{1}{nT}\sum_i\sum_t(U_{it,m}-\bar{U}_{i\cdot,m})(U_{it,k}-\bar{U}_{i\cdot,k})\right]|$, which is $O_P(1)$.

(vi) Under the assumption that $\|R_x\|_1=o\left(\sqrt{\frac{\log p}{nT}}\right)$, by part (v), $\max_{k,m}|\frac{1}{nT}\sum_t\sum_{i=1}^n\widehat{U}_{it,k}\widehat{U}_{it,m}|=O_P(1)$, so

$$\begin{aligned}\|\frac{1}{nT}\sum_t R_x'\widehat{U}_t'\widehat{U}_t\|_\infty &= \max_m|\frac{1}{nT}\sum_t\sum_{k=1}^p\sum_{i=1}^n R_{x,k}\widehat{U}_{it,k}\widehat{U}_{it,m}|\\ &\leq \|R_x\|_1\max_{k,m}|\frac{1}{nT}\sum_t\sum_{i=1}^n\widehat{U}_{it,k}\widehat{U}_{it,m}|=o_P\left(\sqrt{\frac{\log p}{nT}}\right).\end{aligned}$$

■

F.2. **Proof of Lemma 4.1.**

Proof. We start by observing

$$\begin{aligned}&\max_{mv}|\frac{1}{nT}\sum_i\sum_t\tilde{U}_{it,m}\tilde{U}_{it,v}-\frac{1}{nT}\sum_i\sum_t(U_{it,m}-\bar{U}_{i\cdot,m})(U_{it,v}-\bar{U}_{i\cdot,v})|\\ &\leq \max_{mv}|\frac{1}{nT}\sum_i\sum_t(\tilde{U}_{it,m}-U_{it,m}+\bar{U}_{i\cdot,m})\tilde{U}_{it,v}|\\ &\quad+\max_{mv}|\frac{1}{nT}\sum_i\sum_t(U_{it,m}-\bar{U}_{i\cdot,m})(\bar{U}_{it,v}-U_{it,v}+\bar{U}_{i\cdot,v})|\\ &=\max_{mv}|\frac{1}{T}\sum_t(\bar{U}_{\cdot t,m}-\bar{\bar{U}}_m)(\bar{U}_{\cdot t,v}-\bar{\bar{U}}_v)|\\ &=\max_{mv}|\frac{1}{T}\sum_t\bar{U}_{\cdot t,m}(\bar{U}_{\cdot t,v}-\bar{\bar{U}}_v)|\end{aligned}$$

$$
\begin{aligned}
&\le \max_m \frac{1}{T}\sum_t \bar{U}^2_{\cdot t,m} + \max_m \bar{\bar{U}}^2_m \\
&= O_P\left(\frac{\log(p)}{nT}\right) + \max_m \frac{1}{T}\sum_t E\bar{U}^2_{\cdot t,m} + \max_m |\frac{1}{T}\sum_t \bar{U}^2_{\cdot t,m} - E\bar{U}^2_{\cdot t,m}| \\
&= O_P\left(\frac{\log(p)}{nT} + \frac{1}{n} + \sqrt{\frac{\log p}{n^2 T}}\right).
\end{aligned}
$$

Also, let $\mathcal{U}_{i,mv} = \frac{1}{T}\sum_t (U_{it,m} - \bar{U}_{i\cdot,m})(U_{it,v} - \bar{U}_{i\cdot,v}) = \frac{1}{T}\sum_t U_{it,m}U_{it,v} - \bar{U}_{i\cdot,m}\bar{U}_{i\cdot,v}$. Note that

$$
\begin{aligned}
\max_{mv} \frac{1}{n}\sum_i \mathrm{Var}(\frac{1}{T}\sum_t U_{it,m}U_{it,v}) &= \max_{mv}\frac{1}{n}\sum_i \frac{1}{T^2}\sum_t\sum_s \mathrm{Cov}(U_{it,m}U_{it,v}, U_{is,m}U_{is,v}) \\
&= O\left(\frac{1}{T}\right)
\end{aligned}
$$

and that

$$
\begin{aligned}
\max_{mv}\frac{1}{n}\sum_i \mathrm{Var}(\bar{U}_{i\cdot,m}\bar{U}_{i\cdot,v}) &= \max_{mv}\frac{1}{n}\sum_i \frac{1}{T^4}\sum_{tskl}\mathrm{Cov}(U_{it,m}U_{is,v}, U_{ik,m}U_{il,v}) \\
&= O\left(\frac{1}{T^2}\right).
\end{aligned}
$$

Thus, $\max_{mv}\frac{1}{n}\sum_i \mathrm{Var}(\mathcal{U}_{i,mv}) \le O(\frac{1}{T}) + O(\frac{1}{T^2}) + \sqrt{O(\frac{1}{T^2})O(\frac{1}{T})} = O(\frac{1}{T})$. Hence,

$$
\begin{aligned}
&\max_{mv}\left|\frac{1}{nT}\sum_i\sum_t (U_{it,m} - \bar{U}_{i\cdot,m})(U_{it,v} - \bar{U}_{i\cdot,v}) - E\left[\frac{1}{nT}\sum_i\sum_t (U_{it,m} - \bar{U}_{i\cdot,m})(U_{it,v} - \bar{U}_{i\cdot,v})\right]\right| \\
&\quad = \max_{mv}|\frac{1}{n}\sum_i \mathcal{U}_{i,mv} - E\mathcal{U}_{i,mv}| = O_P\left(\sqrt{\frac{\log p}{nT}}\right).
\end{aligned}
$$

Write $A_{mv} = \frac{1}{nT}\sum_i\sum_t \tilde{U}_{it,m}\tilde{U}_{it,v} - E\frac{1}{nT}\sum_i\sum_t (U_{it,m} - \bar{U}_{i\cdot,m})(U_{it,v} - \bar{U}_{i\cdot,v})$. We have proven

$$
\max_{mv}|A_{mv}| = O_P\left(\sqrt{\frac{\log p}{nT}} + \frac{\log(p)}{nT} + \frac{1}{n} + \sqrt{\frac{\log p}{n^2T}}\right) = O_P\left(\sqrt{\frac{\log p}{nT}}\right).
$$

Therefore, for any $\delta \in \mathbb{R}^p/\{0\}$,

$$
\begin{aligned}
&\delta' \frac{1}{nT}\sum_{it}[\tilde{U}_{it}\tilde{U}'_{it} - E(U_{it} - \bar{U}_{i\cdot})(U_{it} - \bar{U}_{i\cdot})']\delta \\
&\qquad \le \|\delta\|_1^2 \max_{mv}|A_{mv}| = \|\delta\|_1^2 O_P\left(\sqrt{\frac{\log p}{nT}}\right)
\end{aligned} \tag{F.17}
$$

where the $O_P\left(\sqrt{\frac{\log p}{nT}}\right)$ term is independent of δ.

We are now ready to verify Assumptions 3.4 and 4.2.

Verification of Assumption 3.4 Uniformly on the set $\{\delta : \|\delta\|_0 \leq l_T|J|_0\}$, we have $\|\delta\|_1 \leq \|\delta\|_2\sqrt{l_T|J|_0}$. Hence as long as $l_T^2|J|_0^2\log p = o(n)$ (for instance, $l_T = \log p$),

$$\begin{aligned} \mathcal{R}(\delta) &\geq c - O_P\left(\sqrt{\frac{\log p}{nT}}\right)\frac{\|\delta\|_1^2}{\|\delta\|_2^2} \geq c - O_P\left(l_T|J|_0\sqrt{\frac{\log p}{nT}}\right) \geq c/2 \\ \mathcal{R}(\delta) &\leq C + O_P\left(\sqrt{\frac{\log p}{nT}}\right)\frac{\|\delta\|_1^2}{\|\delta\|_2^2} \leq C + O_P\left(l_T|J|_0\sqrt{\frac{\log p}{nT}}\right) \leq 2C, \end{aligned}$$

where C denotes the maximum eigenvalue of $E\left[(U_{it}-\bar{U}_{i\cdot})(U_{it}-\bar{U}_{i\cdot})'\right]$.

Verification of Assumption 4.2 The result follows immediately from (F.17).

■

Appendix G. Proof of Theorem 4.1

Define $TK \times 1$ matrices $\widehat{\Xi} = (\widehat{\xi}_1', ..., \widehat{\xi}_T')'$ and $\widehat{\Delta}_d = (\widehat{\delta}_{d1}', ..., \widehat{\delta}_{dT}')'$.

Note that $\tilde{Y}^* = \tilde{D}^*\widehat{\alpha} + (I_T \otimes \widehat{F})\widehat{\Xi} + \tilde{U}^*\widehat{\theta} + \tilde{\epsilon}^*$, and that $\widehat{\eta}^* = M_{\widehat{U}^*_{\widehat{J}^*}}(I_T \otimes M_{\widehat{F}^*})\tilde{D}^*$. Hence

$$\begin{aligned} \widehat{\alpha}^* &= (\widehat{\eta}'^*\widehat{\eta}^*)^{-1}\widehat{\eta}'^* M_{\widehat{U}^*_{\widehat{J}^*}}(I_T \otimes M_{\widehat{F}^*})\tilde{Y} \\ &= \widehat{\alpha} + (\widehat{\eta}'^*\widehat{\eta}^*)^{-1}\widehat{\eta}'^* M_{\widehat{U}^*_{\widehat{J}^*}}(I_T \otimes M_{\widehat{F}^*})[(I_T \otimes \widehat{F})\widehat{\Xi} + \tilde{U}^*\widehat{\theta} + \tilde{\epsilon}^*]. \end{aligned}$$

Hence

$$\sqrt{nT}\left(\frac{1}{nT}\widehat{\eta}'^*\widehat{\eta}^*\right)(\widehat{\alpha}^* - \widehat{\alpha}) = \frac{1}{\sqrt{nT}}\tilde{\eta}^{*\prime}\tilde{\epsilon}^* + \sum_{i=1}^{6} A_i^* \tag{G.1}$$

where

$$\begin{aligned} A_1^* &= \frac{1}{\sqrt{nT}}(\widehat{\eta}^* - \tilde{\eta}^*)' M_{\widehat{U}^*_{\widehat{J}^*}}(I_T \otimes M_{\widehat{F}^*})\tilde{\epsilon}^*, \quad A_2^* = \frac{1}{\sqrt{nT}}\widehat{\eta}'^* M_{\widehat{U}^*_{\widehat{J}^*}}(I_T \otimes M_{\widehat{F}^*}\widehat{F})\widehat{\Xi} \\ A_3^* &= -\frac{1}{\sqrt{nT}}\tilde{\eta}^{*\prime}(I_T \otimes P_{\widehat{F}^*})\tilde{\epsilon}^*, \quad A_4^* = \frac{1}{\sqrt{nT}}\widehat{\eta}'^* M_{\widehat{U}^*_{\widehat{J}^*}}(I_T \otimes M_{\widehat{F}^*})\tilde{U}^*\widehat{\theta} \\ A_5^* &= -\frac{1}{\sqrt{nT}}\tilde{\eta}^{*\prime} P_{\widehat{U}^*_{\widehat{J}^*}}\tilde{\epsilon}^*, \quad A_6 = \frac{1}{\sqrt{nT}}\tilde{\eta}^{*\prime} P_{\widehat{U}^*_{\widehat{J}^*}}(I_T \otimes P_{\widehat{F}^*})\tilde{\epsilon}^* = 0. \end{aligned}$$

We shall prove that $A_i^* = o_{P^*}(1)$ for $i = 1, ..., 6$ and $\frac{1}{nT}\widehat{\eta}^{*\prime}\widehat{\eta}^* - \frac{1}{nT}\tilde{\eta}'\tilde{\eta} = o_{P^*}(1)$. Similarly to (A.5), it can be shown that

$$\begin{aligned}\frac{1}{\sqrt{nT}}(\widehat{\eta}^* - \tilde{\eta}^*) = & \frac{1}{\sqrt{nT}}M_{\widehat{U}^*_{\widehat{J}^*}}(I_T \otimes M_{\widehat{F}^*}(\widehat{F}H^* - \widehat{F}^*)H^{*-1})\widehat{\Delta}_d \\ & + \frac{1}{\sqrt{nT}}M_{\widehat{U}^*_{\widehat{J}^*}}(I_T \otimes M_{\widehat{F}^*})\tilde{U}^*\widehat{\gamma}_d \\ & - \frac{1}{\sqrt{nT}}P_{\widehat{U}^*_{\widehat{J}^*}}\tilde{\eta}^* \\ & - \frac{1}{\sqrt{nT}}M_{\widehat{U}^*_{\widehat{J}^*}}(I_T \otimes P_{\widehat{F}^*})\tilde{\eta}^*.\end{aligned} \tag{G.2}$$

Recall that a_n is defined in Assumption 4.1, which quantifies the minimization error of the k-step lasso. Write

$$\psi_n^* := \kappa_n |J|_0^{1/2} + \Delta_F^* |J|_0 + \sqrt{\frac{|J|_0}{n}} + \sqrt{a_n}. \tag{G.3}$$

Proposition G.1. *(i)* $\frac{1}{\sqrt{nT}}\|\widehat{\eta}^* - \tilde{\eta}^*\|_2 = O_P(\psi_n^*)$.
(ii) $A_l^* = o_{P^*}(1), l = 1, ..., 6$.

Proof. The proof for (i) is similar to that of Proposition A.1, based on (G.2) and Lemma H.7 below. The proof for (ii) follows the same arguments employed in Sections A.2-A.4 using on Lemmas H.7 and H.8 below. The only difference is that $A_4^* = o_{P^*}(1)$ requires $a_n\sqrt{nT} = o(1)$. We omit the details for brevity. ■

G.1. **Proof of Theorem 4.1.** It follows from equation (G.1) and Proposition G.1 that

$$\sqrt{nT}(\frac{1}{nT}\widehat{\eta}'^*\widehat{\eta}^*)(\widehat{\alpha}^* - \widehat{\alpha}) = \frac{1}{\sqrt{nT}}\tilde{\eta}^{*\prime}\tilde{\epsilon}^* + o_{P^*}(1).$$

Note that $\frac{1}{\sqrt{nT}}\tilde{\eta}^{*\prime}\tilde{\epsilon}^* = \frac{1}{\sqrt{n}}\sum_i \varpi_i w_i^Y w_i^D$, where $\varpi_i = \frac{1}{\sqrt{T}}\sum_t \widehat{\eta}_{it}\widehat{\epsilon}_{it}$. Let

$$b_n = \left[\mathrm{Var}^*(\frac{1}{\sqrt{n}}\sum_i \varpi_i w_i^Y w_i^D)\right]^{-1/2},$$

so

$$b_n\sqrt{nT}(\frac{1}{nT}\widehat{\eta}'^*\widehat{\eta}^*)(\widehat{\alpha}^* - \widehat{\alpha}) = \frac{b_n}{\sqrt{nT}}\tilde{\eta}^{*\prime}\tilde{\epsilon}^* + o_{P^*}(b_n).$$

Now note that Step 1 below shows $b_n = O_P(1)$. Step 2 shows $\frac{b_n}{\sqrt{nT}}\tilde{\eta}^{*\prime}\tilde{\epsilon}^* \to^{d^*} \mathcal{N}(0,1)$ and hence $b_n\sqrt{nT}(\widehat{\alpha}^* - \widehat{\alpha}) = O_{P^*}(1)$. Step 3 shows $|\sigma_\eta^2 - \frac{1}{nT}\widehat{\eta}'^*\widehat{\eta}^*| = o_{P^*}(1)$. Therefore,

$b_n\sqrt{nT}\frac{1}{nT}(\widehat{\eta}'^*\widehat{\eta}^* - \sigma^2_\eta)(\widehat{\alpha}^* - \widehat{\alpha}) = o_{P^*}(1)$ which yields

$$b_n\sqrt{nT}\sigma^2_\eta(\widehat{\alpha}^* - \widehat{\alpha}) = \frac{b_n}{\sqrt{nT}}\tilde{\eta}^{*'}\tilde{\epsilon}^* \to^{d^*} \mathcal{N}(0,1).$$

In addition, we also show in Step 1 below that $\sqrt{nT}(\sigma^2_\eta)(\widehat{\alpha}^* - \widehat{\alpha}) = O_{P^*}(b_n^{-1}) = O_{P^*}(1)$. Hence,

$$(b_n - \sigma_{\eta\epsilon}^{-1/2})\sqrt{nT}(\sigma^2_\eta)(\widehat{\alpha}^* - \widehat{\alpha}) = o_{P^*}(1).$$

Thus, we have both

$$\begin{aligned} \sqrt{nT}\sigma_{\eta\epsilon}^{-1/2}\sigma^2_\eta(\widehat{\alpha}^* - \widehat{\alpha}) &\to^{d^*} \mathcal{N}(0,1) \\ \sqrt{nT}\sigma_{\eta\epsilon}^{-1/2}\sigma^2_\eta(\widehat{\alpha} - \alpha) &\to^{d} \mathcal{N}(0,1). \end{aligned} \tag{G.4}$$

where the second line follows from Theorem 3.1.

Now let q^*_τ be such that $P^*(\sqrt{nT}|\widehat{\alpha}^* - \widehat{\alpha}| \leq q^*_\tau) = 1-\tau$, and let

$$\begin{aligned} a_1 &:= P(\sqrt{nT}|\widehat{\alpha} - \alpha| \leq q^*_\tau) = P(\sigma_{\eta\epsilon}^{-1/2}\sigma^2_\eta\sqrt{nT}|\widehat{\alpha} - \alpha| \leq \sigma_{\eta\epsilon}^{-1/2}\sigma^2_\eta q^*_\tau) \\ a_2 &:= P^*(\sqrt{nT}|\widehat{\alpha}^* - \widehat{\alpha}| \leq q^*_\tau) = P^*(\sigma_{\eta\epsilon}^{-1/2}\sigma^2_\eta\sqrt{nT}|\widehat{\alpha}^* - \widehat{\alpha}| \leq \sigma_{\eta\epsilon}^{-1/2}\sigma^2_\eta q^*_\tau) \end{aligned}$$

The results summarized in (G.4) then implies $a_1 \to a_2 = 1-\tau$.

Step 1: Show $b_n = O_P(1)$ and $b_n^{-1} = O_P(1)$.

Let $Z_i = \sum_t(\widehat{\eta}_{it}\widehat{\epsilon}_{it} - \tilde{\eta}_{it}\tilde{\epsilon}_{it})$. Then by Lemma H.4, $\frac{1}{nT}\sum_i Z_i^2 = o_P(1)$. Hence,

$$\begin{aligned} b_n^{-2} &= \text{Var}^*(\frac{1}{\sqrt{n}}\sum_i \varpi_i w_i^Y w_i^D) \\ &= \frac{1}{n}\sum_i \varpi_i^2 E^*(w_i^Y w_i^D)^2 = \frac{1}{n}\sum_i \varpi_i^2 \\ &= \frac{1}{n}\sum_i(\frac{1}{\sqrt{T}}\sum_t \widehat{\eta}_{it}\widehat{\epsilon}_{it})^2 = \frac{1}{nT}\sum_i(Z_i + \sum_t \tilde{\eta}_{it}\tilde{\epsilon}_{it})^2 \\ &= \frac{1}{nT}\sum_i Z_i^2 + \frac{1}{nT}\sum_i(\sum_t \tilde{\eta}_{it}\tilde{\epsilon}_{it})^2 + \frac{2}{nT}\sum_i Z_i \sum_t \tilde{\eta}_{it}\tilde{\epsilon}_{it} \\ &\leq o_P(1) + \frac{1}{nT}\sum_i(\sum_t \tilde{\eta}_{it}\tilde{\epsilon}_{it})^2 + o_P(1)\sqrt{\frac{1}{nT}\sum_i(\sum_t \tilde{\eta}_{it}\tilde{\epsilon}_{it})^2}. \end{aligned}$$

By (A.18) and (A.20), $\frac{1}{nT}\sum_{i=1}^n(\sum_{t=1}^T \tilde{\eta}_{it}\tilde{\epsilon}_{it})^2 \to^P \sigma_{\eta\epsilon}$. Hence, $b_n^{-2} \to^P \sigma_{\eta\epsilon} > c$. It follows that $b_n = O_P(1)$, $b_n \to^P \sigma_{\eta\epsilon}^{-1/2}$, and $b_n^{-1} = O_P(1)$ since $\sigma_{\eta\epsilon} < C$.

Step 2: Apply the CLT to $\frac{b_n}{\sqrt{nT}}\tilde{\eta}^{*'}\tilde{\epsilon}^*$.

Let $g_{n,i} = b_n \varpi_i w_i^Y w_i^D$ and $s_n^2 = \sum_i \mathrm{Var}^*(g_{n,i}) = n$. Note that $E^* g_{n,i} = 0$.

We now verify the Lindeberg condition for the triangular array $\{g_{n,i}\}$. For any $\varepsilon > 0$,

$$E^* \left(\frac{1}{n} \sum_i g_{n,i}^2 1\{|g_{n,i}| > \varepsilon \sqrt{n}\} \right) \leq E^* \left[\frac{1}{n} \sum_i g_{n,i}^2 \right] = 1.$$

Hence, by the dominated convergence theorem,

$$s_n^{-2} \sum_i E^* g_{n,i}^2 1\{|g_{n,i}| > \varepsilon s_n\} = E^* \left(\frac{1}{n} \sum_i g_{n,i}^2 1\{|g_{n,i}| > \varepsilon \sqrt{n}\} \right) \to 0.$$

We then have, by the Lindeberg central limit theorem,

$$\frac{1}{s_n} \sum_i g_{n,i} = \frac{b_n}{\sqrt{n}} \sum_i \varpi_i w_i^Y w_i^D = \frac{b_n}{\sqrt{nT}} \tilde{\eta}^{*\prime} \tilde{\epsilon}^* \to^{d^*} \mathcal{N}(0,1). \tag{G.5}$$

Step 3: Show $|\sigma_\eta^2 - \frac{1}{nT} \widehat{\eta}^{\prime *} \widehat{\eta}^*| = o_{P^*}(1)$.

First,

$$\begin{aligned}
|\frac{1}{nT} \widehat{\eta}^{\prime *} \widehat{\eta}^* - \frac{1}{nT} \sum_{i,t} \tilde{\eta}_{it}^2 (w_i^D)^2| &= |\frac{1}{nT} \sum_{i,t} (\widehat{\eta}_{it}^2 - \tilde{\eta}_{it}^2)(w_i^D)^2| \\
&\leq \frac{1}{nT} \sum_{i,t} |\widehat{\eta}_{it} - \tilde{\eta}_{it}||\widehat{\eta}_{it} + \tilde{\eta}_{it}| \max_i (w_i^D)^2 \\
&\leq \left(\frac{1}{nT} \sum_{i,t} (\widehat{\eta}_{it} - \tilde{\eta}_{it})^2 \right) O_{P^*}(\log n) \\
&= O_{P^*}(\psi_n \log n) = o_{P^*}(1).
\end{aligned}$$

Then, using $E^*(w_i^D)^2 = 1$ and

$$\begin{aligned}
E^* |\frac{1}{nT} \sum_{i,t} \tilde{\eta}_{it}^2 ((w_i^D)^2 - E^*(w_i^D)^2)|^2 &= \mathrm{Var}^* \left(\frac{1}{nT} \sum_{i,t} \tilde{\eta}_{it}^2 (w_i^D)^2 \right) \\
&= \frac{1}{n^2 T^2} \sum_i \mathrm{Var}^* \left((w_i^D)^2 \right) \left(\sum_t \tilde{\eta}_{it}^2 \right)^2 \\
&\leq \frac{C}{n^2} \sum_i \left(\frac{1}{T} \sum_t \tilde{\eta}_{it}^2 \right)^2 = O_P\left(\frac{1}{n} \right),
\end{aligned}$$

we have $|\frac{1}{nT}\sum_{i,t}\tilde\eta_{it}^2(w_i^D)^2-\frac{1}{nT}\sum_{i,t}\tilde\eta_{it}^2|=o_{P^*}(1)$, from which

$$|\frac{1}{nT}\widehat{\eta}'^*\widehat{\eta}^*-\frac{1}{nT}\sum_{i,t}\tilde\eta_{it}^2|=o_{P^*}(1)$$

follows. Moreover, equation (A.15) in the proof of Theorem 3.1 shows $|\frac{1}{nT}\tilde\eta'\tilde\eta-\frac{1}{nT}\sum_{i,t}E\iota_{it}|=o_P(1)$ for $\iota_{it}=(\eta_{it}-\bar\eta_{i\cdot})^2$. Hence $|\frac{1}{nT}\sum_{i,t}E\iota_{it}-\frac{1}{nT}\widehat{\eta}'^*\widehat{\eta}^*|=o_{P^*}(1)$. Finally, note that $\sigma_\eta^2:=\frac{1}{nT}\sum_{i,t}E\iota_{it}$. ■

Appendix H. Technical lemmas for the main theorems

H.1. Technical Lemmas for Theorem 3.1.

Lemma H.1. *For $\gamma\in\{\gamma_y,\gamma_d,\theta\}$,*
(i) $\|\frac{1}{\sqrt{nT}}M_{\widehat U_{\widehat J}}\tilde U\gamma\|_2^2=O_P\left(\|R_y\|_1^2+\kappa_n^2|J|_0+|J|_0^2\Delta_F^2+\frac{|J|_0}{n}\right)$.
(ii) $\|\frac{1}{\sqrt{nT}}(I_T\otimes P_{\widehat F})\tilde U\gamma\|_2^2=O_P\left(\frac{|J|_0}{n}+|J|_0^2\Delta_F^2\right)$.
(iii) $\|\frac{1}{\sqrt{nT}}(I_T\otimes P_{\widehat F})\tilde\eta\|_2^2=O_P\left(\Delta_{fe}^2+\frac{1}{n}\right)$ *and* $\|\frac{1}{\sqrt{nT}}(I_T\otimes P_{\widehat F})\tilde\epsilon\|_2^2=O_P\left(\Delta_{fe}^2+\frac{1}{n}\right)$.
(iv) $\|\frac{1}{\sqrt{nT}}P_{\widehat U_{\widehat J}}\tilde\eta\|_2^2=O_P\left(|J|_0\frac{\log p}{nT}\right)$ *and* $\|\frac{1}{\sqrt{nT}}P_{\widehat U_{\widehat J}}\tilde\epsilon\|_2^2=O_P\left(|J|_0\frac{\log p}{nT}\right)$.

Proof. (i) First, consider the following constrained problem:

$$\widehat m=\arg\min_m\|\widehat U(\gamma-m)\|_2^2,\quad m_j=0,j\notin\widehat J. \tag{H.1}$$

The solution satisfies $\widehat U\widehat m=P_{\widehat U_{\widehat J}}\widehat U\gamma$. Hence, from Proposition F.1,
$\frac{1}{nT}\sum_{t=1}^T\|\widehat U_t(\gamma_y-\tilde\gamma_y)\|_2^2=O_P(\|R_y\|_1^2+\kappa_n^2|J|_0)$, and

$$\begin{aligned}\|\frac{1}{\sqrt{nT}}M_{\widehat U_{\widehat J}}\widehat U\gamma\|_2^2&=\|\frac{1}{\sqrt{nT}}(\widehat U\gamma-P_{\widehat U_{\widehat J}}\widehat U\gamma)\|_2^2\\&=\|\frac{1}{\sqrt{nT}}\widehat U(\gamma-\widehat m)\|_2^2\le\|\frac{1}{\sqrt{nT}}\widehat U(\gamma-\tilde\gamma)\|_2^2\\&=\frac{1}{nT}\sum_{t=1}^T\|\widehat U_t(\gamma-\tilde\gamma)\|_2^2=O_P(\|R_y\|_1^2+\kappa_n^2|J|_0).\end{aligned}$$

Next, $\frac{1}{T}\sum_t\|\frac{1}{n}\widehat F'\tilde U_t\gamma\|_2^2=O_P\left(\frac{|J|_0}{n}+|J|_0^2\Delta_F^2\right)$ by Lemma H.11. Also, using the equality

$$\begin{aligned}\tilde U_{it,m}-\widehat U_{it,m}=&\widehat f_i'(\widehat F'\widehat F)^{-1}\widehat F'(\tilde FH-\widehat F)H^{-1}\tilde\lambda_{tm}+\widehat f_i'(\widehat F'\widehat F)^{-1}\widehat F'\tilde U_{t,m}\\&+\tilde\lambda_{tm}'H'^{-1}(\widehat f_i-H'\tilde f_i),\end{aligned}$$

we have

$$\begin{aligned}\left\|\frac{1}{\sqrt{nT}}M_{\widehat{U}_{\widehat{J}}}(\widehat{U}-\tilde{U})\gamma\right\|_2^2 &\leq \frac{1}{nT}\|(\widehat{U}-\tilde{U})\gamma\|_2^2 \\ &= \frac{1}{nT}\sum_{t=1}^{T}\sum_{i}\left[\sum_{m}(\widehat{U}_{it,m}-\tilde{U}_{it,m})\gamma_m\right]^2 \\ &\leq |J|_0^2 O_P(\Delta_F^2) + \frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{n}\widehat{F}'\tilde{U}_t\gamma\right\|_2^2 + |J|_0^2 O_P(\Delta_F^2) \\ &= O_P\left(|J|_0^2\Delta_F^2 + \frac{|J|_0}{n}\right).\end{aligned} \tag{H.2}$$

where we used, by Lemma H.11, $\frac{1}{T}\sum_t \|\frac{1}{n}\widehat{F}'\tilde{U}_t\gamma\|_2^2 = O_P\left(\frac{|J|_0}{n} + |J|_0^2\Delta_F^2\right)$. Combining these results yields $\left\|\frac{1}{\sqrt{nT}}M_{\widehat{U}_{\widehat{J}}}\tilde{U}\gamma\right\|_2^2 = O_P\left(\|R_y\|_1^2 + \kappa_n^2|J|_0 + |J|_0^2\Delta_F^2 + \frac{|J|_0}{n}\right)$.

(ii) By Lemma H.11,

$$\begin{aligned}\left\|\frac{1}{\sqrt{nT}}(I_T\otimes P_{\widehat{F}})\tilde{U}\gamma\right\|_2^2 &= \frac{1}{nT}\sum_t \left\|P_{\widehat{F}}\tilde{U}_t\gamma\right\|_2^2 \\ &\leq \frac{1}{n^2T}\sum_t\left\|\widehat{F}'\tilde{U}_t\gamma\right\|_2^2 \\ &= O_P\left(\frac{|J|_0}{n} + |J|_0^2\Delta_F^2\right).\end{aligned}$$

(iii) By Lemma H.11, $\frac{1}{T}\sum_t\left\|\frac{1}{n}\widehat{F}'\tilde{\eta}_t\right\|_2^2 = O_P\left(\frac{1}{n}+\Delta_{fe}^2\right)$. Thus,

$$\left\|\frac{1}{\sqrt{nT}}(I_T\otimes P_{\widehat{F}})\tilde{\eta}\right\|_2^2 = \frac{1}{nT}\sum_t\|P_{\widehat{F}}\tilde{\eta}_t\|_2^2 \leq \frac{1}{n^2T}\sum_t\left\|\widehat{F}'\tilde{\eta}_t\right\|_2^2 = O_P\left(\Delta_{fe}^2 + \frac{1}{n}\right).$$

(iv) The same argument as employed in the proof of Lemma F.1 yields

$$\max_m\left|\frac{1}{nT}\sum_t\sum_i \tilde{U}_{it,m}\tilde{\eta}_{it}\right| = O_P\left(\sqrt{\frac{\log p}{nT}}\right)$$

and

$$\max_m\left|\frac{1}{nT}\sum_t\sum_i(\tilde{U}_{it,m}-\widehat{U}_{it,m})\tilde{\eta}_{it}\right| = o_P\left(\sqrt{\frac{\log p}{nT}}\right).$$

Hence, using $|\widehat{J}|_0 = O_P(|J|_0)$ from Proposition F.1, we have

$$\begin{aligned}\left\|\frac{1}{nT}\tilde{U}'_{\widehat{J}}\tilde{\eta}\right\|_2 &= \left\|\frac{1}{nT}\sum_t \tilde{U}'_{t,\widehat{J}}\tilde{\eta}_t\right\|_2 \\ &\le \max_m \left|\frac{1}{nT}\sum_t\sum_i \tilde{U}_{it,m}\tilde{\eta}_{it}\right|\sqrt{|\widehat{J}|_0} \\ &= O_P\left(\sqrt{|J|_0\frac{\log p}{nT}}\right)\end{aligned}$$

and $\left\|\frac{1}{nT}(\tilde{U}_{\widehat{J}} - \widehat{U}_{\widehat{J}})'\tilde{\eta}\right\|_2 \le \left\|\frac{1}{nT}\sum_t(\tilde{U}'_{t,\widehat{J}} - \widehat{U}'_{t,\widehat{J}})\tilde{\eta}_t\right\|_\infty \sqrt{|\widehat{J}|_0} = o_P\left(\sqrt{|J|_0\frac{\log p}{nT}}\right)$. Hence,

$$\left\|\frac{1}{nT}\widehat{U}'_{\widehat{J}}\tilde{\eta}\right\|_2 = O_P\left(\sqrt{|J|_0\frac{\log p}{nT}}\right). \tag{H.3}$$

Next, by Lemma F.1,

$$\left\|\frac{1}{nT}\widehat{U}'_{\widehat{J}}\widehat{U}_{\widehat{J}} - \frac{1}{nT}\tilde{U}'_{\widehat{J}}\tilde{U}_{\widehat{J}}\right\| \le O_P(|J|_0)\max_{m,k\le p}\left|\frac{1}{nT}\sum_{it}(\widehat{U}_{it,m}\widehat{U}_{it,k} - \tilde{U}_{it,m}\tilde{U}_{it,k})\right| = o_P(1).$$

Therefore, $\|\frac{1}{\sqrt{nT}}\widehat{U}_{\widehat{J}}\|^2 = \|\frac{1}{nT}\widehat{U}'_{\widehat{J}}\widehat{U}_{\widehat{J}}\| \le \|\frac{1}{nT}\tilde{U}'_{\widehat{J}}\tilde{U}_{\widehat{J}}\| = O_P(\phi_{\max}(|J|_0)) + o_P(1) = O_P(1)$ and

$$\lambda_{\min}\left(\frac{1}{nT}\widehat{U}'_{\widehat{J}}\widehat{U}_{\widehat{J}}\right) \ge \lambda_{\min}\left(\frac{1}{nT}\tilde{U}'_{\widehat{J}}\tilde{U}_{\widehat{J}}\right) - o_P(1) \ge \phi_{\min}(|\widehat{J}|_0) - o_P(1), \tag{H.4}$$

which is bounded away from zero. Thus, $\left\|\left(\frac{1}{nT}\widehat{U}'_{\widehat{J}}\widehat{U}_{\widehat{J}}\right)^{-1}\right\| = O_P(1)$.

Finally, $\left\|\frac{1}{\sqrt{nT}}P_{\widehat{U}_{\widehat{J}}}\tilde{\eta}\right\|_2 \le \left\|\frac{1}{\sqrt{nT}}\widehat{U}_{\widehat{J}}(\frac{1}{nT}\widehat{U}'_{\widehat{J}}\widehat{U}_{\widehat{J}})^{-1}\frac{1}{nT}\widehat{U}'_{\widehat{J}}\tilde{\eta}\right\|_2 = O_P\left(\sqrt{|J|_0\frac{\log p}{nT}}\right)$. ■

Lemma H.2. *(i)* $\|\frac{\sqrt{T}}{n\sqrt{n}}\sum_{m=1}^p \widehat{F}'\frac{1}{T}\sum_t U_{t,m}\epsilon'_t\widehat{F}\gamma_{dm}\|_F = o_P(1)$

(ii) $\|\frac{1}{n}\sum_{j=1}^n \widehat{f}_j\bar{U}'_{j\cdot}\gamma_d\|_2 = O_P\left(\Delta_F|J|_0 + \sqrt{\frac{|J|_0}{nT}}\right)$, *and* $\|\frac{1}{n}\sum_{i=1}^n \bar{\epsilon}_{i\cdot}\widehat{f}'_i\|_2 = O_P\left(\Delta_F + \frac{1}{\sqrt{nT}}\right)$.

(iii) $\|\frac{\sqrt{nT}}{n}\sum_{m=1}^p\sum_{j=1}^n \widehat{f}_j\frac{1}{T}\sum_{t=1}^T U_{jt,m}\bar{\epsilon}_{\cdot t}\gamma_{dm}\|_2 = o_P(1)$.

(iv) $\|\frac{\sqrt{nT}}{n}\sum_{i=1}^n\frac{1}{T}\sum_{t=1}^T \bar{U}'_{\cdot t}\gamma_d\epsilon_{it}\widehat{f}'_i\|_2 = o_P(1)$ *and* $\sqrt{nT}|\sum_{m=1}^p\frac{1}{T}\sum_{t=1}^T \bar{U}_{\cdot t,m}\bar{\epsilon}_{\cdot t}\gamma_{dm}| = o_P(1)$.

(v) $\|\frac{1}{nT}\sum_t \widehat{U}'_{t,\widehat{J}}(\tilde{F}H - \widehat{F})H^{-1}\tilde{\delta}_{dt}\|_2 = O_P\left(\Delta_{ud} + \Delta_F\sqrt{\frac{\log p}{nT}}\right)\sqrt{|J|_0}$.

Proof. (i) We have

$$\left\| \frac{\sqrt{T}}{n\sqrt{n}} \sum_{m=1}^{p} \widehat{F}' \frac{1}{T} \sum_t U_{t,m} \epsilon_t' \widehat{F} \gamma_{dm} \right\|_F$$

$$\leq \left\| \frac{\sqrt{T}}{n\sqrt{n}} \sum_{m=1}^{p} (\widehat{F} - \tilde{F}H)' \frac{1}{T} \sum_t U_{t,m} \epsilon_t' (\widehat{F} - \tilde{F}H) \gamma_{dm} \right\|_F \tag{H.5}$$

$$+ O_P(1) \left\| \frac{\sqrt{T}}{n\sqrt{n}} \sum_{m=1}^{p} \tilde{F}' \frac{1}{T} \sum_t U_{t,m} \epsilon_t' \tilde{F} \gamma_{dm} \right\|_F \tag{H.6}$$

$$+ \left\| \frac{\sqrt{T}}{n\sqrt{n}} \sum_{m=1}^{p} (\widehat{F} - \tilde{F}H)' \frac{1}{T} \sum_t U_{t,m} \epsilon_t' \tilde{F} \gamma_{dm} \right\|_F O_P(1) \tag{H.7}$$

$$+ O_P(1) \left\| \frac{\sqrt{T}}{n\sqrt{n}} \sum_{m=1}^{p} \tilde{F}' \frac{1}{T} \sum_t U_{t,m} \epsilon_t' (\widehat{F} - \tilde{F}H) \gamma_{dm} \right\|_F . \tag{H.8}$$

Term (H.5) is bounded by

$$\sqrt{Tn} \left(\frac{1}{T} \sum_t \| \frac{1}{n} (\widehat{F} - \tilde{F}H)' U_t \gamma_d \|_2^2 \right)^{1/2} \left(\frac{1}{T} \sum_t \| \frac{1}{n} \epsilon_t' (\widehat{F} - \tilde{F}H) \|_2^2 \right)^{1/2} = O_P\left(|J|_0 \Delta_F^2 \sqrt{Tn} \right)$$

$$= o_P(1)$$

using the assumption that $|J|_0 \Delta_F \Delta_{fe} \sqrt{Tn} \leq |J|_0 \Delta_F^2 \sqrt{Tn} = o(1)$. Term (H.6)

$$\left\| \frac{\sqrt{T}}{n\sqrt{n}} \sum_{m=1}^{p} \tilde{F}' \frac{1}{T} \sum_t U_{t,m} \epsilon_t' \tilde{F} \gamma_{dm} \right\|_F = o_P(1).$$

For term (H.7),

$$\left\| \frac{\sqrt{T}}{n\sqrt{n}} \sum_{m=1}^{p} (\widehat{F} - \tilde{F}H)' \frac{1}{T} \sum_t U_{t,m} \epsilon_t' \tilde{F} \gamma_{dm} \right\|_F \leq \frac{\sqrt{T}}{\sqrt{n}} \| \widehat{F}' - \tilde{F}H \|_F \left\| \frac{1}{n} \sum_{m=1}^{p} \frac{1}{T} \sum_t U_{t,m} \epsilon_t' \tilde{F} \gamma_{dm} \right\|_F$$

$$\leq O_P(\sqrt{T}\Delta_F) \left(\frac{1}{T} \sum_t \| \sum_{m=1}^{p} \frac{1}{\sqrt{n}} \gamma_{dm} U_{t,m} \|_2^2 \right)^{1/2} \left(\frac{1}{T} \sum_t \| \frac{1}{\sqrt{n}} \epsilon_t' \tilde{F} \|_2^2 \right)^{1/2}$$

$$= O_P(|J|_0 \Delta_F \sqrt{T}) = o_P(1).$$

For (H.8), by Assumption D.4 that $\frac{1}{T} \sum_{t=1}^{T} \| \frac{1}{n} \sum_{i=1}^{n} (\widehat{f}_i - H' \tilde{f}_i) \epsilon_{it} \|_2^2 = O_P(\Delta_{fe}^2)$,

$$\left\| \frac{\sqrt{T}}{n\sqrt{n}} \sum_{m=1}^{p} \tilde{F}' \frac{1}{T} \sum_t U_{t,m} \epsilon_t' (\widehat{F} - \tilde{F}H) \gamma_{dm} \right\|_F$$

$$\leq (\frac{1}{T}\sum_t \|\frac{1}{n}\epsilon_t'(\widehat{F}-\tilde{F}H)\|_2^2)^{1/2}(\frac{1}{T}\sum_t \|\frac{1}{\sqrt{n}}\sum_{m=1}^p \tilde{F}'U_{t,m}\gamma_{dm}\|_2^2)^{1/2}\sqrt{T}$$
$$\leq O_P(\sqrt{T}\Delta_{fe}|J|_0) \leq O_P(\sqrt{T}\Delta_F|J|_0) = o_P(1).$$

Combining these bounds yields the result.

(ii) Because $\frac{1}{n}\sum_i \left(\frac{1}{T}\sum_t U_{it}'\gamma_d\right)^2 = O_P(|J|_0^2)$,

$$\left\|\frac{1}{nT}\sum_{i=1}^n\sum_t(\widehat{f}_i - H'\tilde{f}_i)U_{it}'\gamma_d\right\|_2^2 \leq \frac{1}{n}\sum_i\left(\frac{1}{T}\sum_t U_{it}'\gamma_d\right)^2\frac{1}{n}\sum_i\left\|\widehat{f}_i - H'\tilde{f}_i\right\|_2^2$$
$$= O_P\left(\Delta_F^2|J|_0^2\right)$$

follows by applying the Cauchy-Schwarz inequality. We also have

$$\|\frac{1}{nT}\sum_{j=1}^n\sum_t(\widehat{f}_j - H'\tilde{f}_j)\epsilon_{jt}\|_2^2 \leq \frac{1}{n}\sum_j(\frac{1}{T}\sum_t\epsilon_{jt})^2\frac{1}{n}\sum_j\|\widehat{f}_j - H'\tilde{f}_j\|_2^2 = O_P(\Delta_F)$$

due to $\frac{1}{n}\sum_j\left(\frac{1}{T}\sum_t\epsilon_{jt}\right)^2 = O_P(1)$. Hence,

$$\left\|\frac{1}{n}\sum_{j=1}^n\widehat{f}_j\bar{U}_{j.}'\gamma_d\right\|_2 \leq \left\|\frac{1}{nT}\sum_{j=1}^n\sum_t(\widehat{f}_j - H'\tilde{f}_j)U_{jt}'\gamma_d\|_2 + O_P(1)\|\frac{1}{nT}\sum_{j=1}^n\sum_t\tilde{f}_jU_{jt}'\gamma_d\right\|_2$$
$$= O_P\left(\Delta_F|J|_0\right) + O_P(1)\left\|\frac{1}{nT}\sum_{j=1}^n\sum_t f_jU_{jt}'\gamma_d\right\|_2 + O_P(1)\left|\frac{1}{nT}\sum_{j=1}^n\sum_t U_{jt}'\gamma_d\right\|_2$$
$$= O_P\left(\Delta_F|J|_0 + \sqrt{\frac{|J|_0}{nT}}\right)$$

The second bound follows similarly.

(iii) Note that

$$
\begin{aligned}
&\left\| \frac{\sqrt{nT}}{n} \sum_{m=1}^{p} \sum_{j=1}^{n} \widehat{f}_j \frac{1}{T} \sum_{t=1}^{T} U_{jt,m} \bar{\epsilon}_{\cdot t} \gamma_{dm} \right\|_2 \\
&\quad \leq \left\| \frac{\sqrt{nT}}{n} \sum_{j=1}^{n} (\widehat{f}_j - H'\tilde{f}_j) \frac{1}{T} \sum_{t=1}^{T} U_{jt}' \gamma_d \bar{\epsilon}_{\cdot t} \right\|_2 + \left\| \frac{\sqrt{nT}}{n} \sum_{j=1}^{n} H' f_j \frac{1}{T} \sum_{t=1}^{T} U_{jt}' \gamma_d \bar{\epsilon}_{\cdot t} \right\|_2 \\
&\qquad + \left\| H' \bar{f} \frac{\sqrt{nT}}{n} \sum_{j=1}^{n} \frac{1}{T} \sum_{t=1}^{T} U_{jt}' \gamma_d \bar{\epsilon}_{\cdot t} \right\|_2 \\
&\quad \leq \left\| \frac{\sqrt{nT}}{n} \sum_{j=1}^{n} (\widehat{f}_j - H'\tilde{f}_j) \frac{1}{T} \sum_{t=1}^{T} U_{jt}' \gamma_d \bar{\epsilon}_{\cdot t} \right\|_2 + O_P(1) \left\| \frac{\sqrt{nT}}{n} \sum_{j=1}^{n} f_j \frac{1}{T} \sum_{t=1}^{T} U_{jt}' \gamma_d \bar{\epsilon}_{\cdot t} \right\|_2 \\
&\qquad + O_P(1) \left| \frac{\sqrt{nT}}{n} \sum_{j=1}^{n} \frac{1}{T} \sum_{t=1}^{T} U_{jt}' \gamma_d \bar{\epsilon}_{\cdot t} \right|
\end{aligned}
\tag{H.9}
$$

We also have that

$$
\begin{aligned}
E\left[\frac{1}{n} \sum_j \left(\frac{1}{T} \sum_t U_{jt}' \gamma_d \bar{\epsilon}_{\cdot t} \right)^2 \right] &= \frac{1}{n^3 T^2} \sum_{j=1}^{n} \sum_{s,t \leq T} E\left[U_{jt}' \gamma_d \epsilon_{jt} U_{js}' \gamma_d \epsilon_{js} \right] \\
&\quad + \frac{1}{n^3 T^2} \sum_j \sum_s \sum_{i \neq j} \sum_t E\left[U_{jt}' \gamma_d U_{js}' \gamma_d \right] E\left[\epsilon_{is} \epsilon_{it} \right] \\
&= O\left(\frac{|J|_0}{n^2} \right) \max_{j,s,t} E|\epsilon_{js}\epsilon_{jt}| \max_{j,t,s,k} \sum_{m=1}^{p} \left| E\left[U_{jt,m} U_{js,k} | \epsilon_{js} \epsilon_{jt} \right] \right| \\
&\quad + O\left(\frac{|J|_0}{nT} \right) \max_{ist} \left| E\left[\epsilon_{is} \epsilon_{it} \right] \right| \max_{j,t,m} \sum_s \sum_m \left| E\left[U_{jt,m} U_{js,k} \right] \right| \\
&= O\left(\frac{|J|_0}{n^2} + \frac{|J|_0}{nT} \right),
\end{aligned}
$$

where we used the assumption that $\max_{j,t,s,k} \sum_{m=1}^{p} |E[U_{jt,m}U_{js,k}|\epsilon_{js}\epsilon_{jt}]| = O(1)$ almost surely and $\max_{j,t,m} \sum_s \sum_m |E[U_{jt,m}U_{js,k}]| = O(1)$. The first term of (H.9) is thus bounded by

$$
\sqrt{nT} \Delta_F \left(\frac{1}{n} \sum_j \left(\frac{1}{T} \sum_t U_{jt}' \gamma_d \bar{\epsilon}_{\cdot t} \right)^2 \right)^{1/2} = O_P\left(\sqrt{nT|J|_0} \Delta_F \right) \left(\frac{1}{n} + \frac{1}{\sqrt{nT}} \right) = o_P(1).
$$

For notational simplicity, suppose $\dim(f_j)=1$ in the following. Note that otherwise we can do the analysis element-by-element as $\dim(f_j)$ is fixed. We can bound the second term of (H.9) by

$$\begin{aligned}
&E\left|\frac{\sqrt{nT}}{n}\sum_{j=1}^n f_j\frac{1}{T}\sum_{t=1}^T U_{jt}'\gamma_d\bar{\epsilon}_{\cdot t}\right|^2 \\
&\quad=\frac{1}{n^3T}E\left[\sum_{j=1}^n\sum_{t=1}^T\sum_{s=1}^T f_j^2U_{jt}'\gamma_d\epsilon_{jt}U_{js}'\gamma_d\epsilon_{js}\right] \\
&\qquad+\frac{1}{n^3T}\sum_{j=1}^n\sum_{t=1}^T\sum_{i\neq j}\sum_{s=1}^T E\left[f_j^2U_{jt}'\gamma_dU_{js}'\gamma_d\right]E\left[\epsilon_{is}\epsilon_{it}\right] \\
&\quad\leq O\left(\frac{|J|_0}{n^3T}\right)\sum_{j=1}^n\sum_{t=1}^T\sum_{s=1}^T E\left|f_j^2\epsilon_{js}\epsilon_{jt}\right|\sum_m|E(U_{js,k}U_{jt,m}|\epsilon_j,f_j)| \\
&\qquad+O\left(\frac{|J|_0}{n^3T}\right)\sum_{j=1}^n\sum_{t=1}^T\sum_{i\neq j}|E\left[\epsilon_{is}\epsilon_{it}\right]|E\left[f_j^2\right]\max_{jtk}\sum_m\sum_{s=1}^T|E\left(U_{jt,m}U_{js,k}|f_j\right)| \\
&\quad=O\left(\frac{|J|_0T}{n^2}+\frac{|J|_0}{n}\right)=o(1).
\end{aligned}$$

Note that we frequently used the fact that $\{f_i,U_i,\epsilon_i\}$ are independent across i and that $E(\epsilon_i|f_i,U_i)=0$ in the above.

It follows from a similar argument to that used to bound the second term of (H.9) that the third term of (H.9) is $o_P(1)$. Hence (H.9) is $o_P(1)$.

(iv) Note that

$$\begin{aligned}
&\left\|\frac{\sqrt{nT}}{n}\sum_{i=1}^n\frac{1}{T}\sum_{t=1}^T\bar{U}_{\cdot t}'\gamma_d\epsilon_{it}\widehat{f}_i'\right\|_2 \\
&\quad\leq\left\|\frac{\sqrt{nT}}{n}\sum_{i=1}^n\frac{1}{T}\sum_{t=1}^T\bar{U}_{\cdot t}'\gamma_d\epsilon_{it}(\widehat{f}_i-H'\tilde{f}_i)'\right\|_2+\left\|\frac{\sqrt{nT}}{n}\sum_{i=1}^n\frac{1}{T}\sum_{t=1}^T\bar{U}_{\cdot t}'\gamma_d\epsilon_{it}\tilde{f}_i'\right\|_2 \\
&\quad\leq O_P\left(\sqrt{nT}\Delta_{fe}\right)\left(\frac{1}{T}\sum_t(\bar{U}_{\cdot t}'\gamma_d)^2\right)^{1/2}+\left\|\frac{\sqrt{nT}}{n}\sum_{i=1}^n\frac{1}{T}\sum_{t=1}^T\bar{U}_{\cdot t}'\gamma_d\epsilon_{it}\tilde{f}_i'\right\|_2.
\end{aligned}$$

Following similar calculations in part (iii), it can be shown that $\frac{1}{T}\sum_t(\bar{U}_{\cdot t}'\gamma_d)^2=O_P\left(\frac{|J|_0}{n}\right)$. Note that $\max_k\sum_m|EU_{it,m}U_{it,k}|=O(1)$, which enables us to remove the impact of one of the $\sum_{m\leq p}$ terms. Hence, the first term is $O_P\left(\sqrt{T|J|_0\Delta_F^2}\right) = o_P(1)$. In addition,

$$\left\|\frac{\sqrt{nT}}{n}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=1}^{T}\bar{U}_{\cdot t}'\gamma_d\epsilon_{it}\tilde{f}_i'\right\|_2 = O_P\left(\sqrt{\frac{|J|_0}{n}}\right) = o_P(1).$$

The second conclusion that

$$\sqrt{nT}|\sum_{m=1}^{p}\frac{1}{T}\sum_{t=1}^{T}\bar{U}_{\cdot t,m}\bar{\epsilon}_{\cdot t}\gamma_{dm}| = o_P(1)$$

follows from a similar calculation to that used to bound the first term given immediately above.

(v) Note that (H.32) holds when $\tilde{\delta}_{yt}$ is replaced with $\tilde{\delta}_{dt}$. Also, recall that $\Delta_F = o\left(\sqrt{\frac{\log p}{nT}}\right)$, so we have

$$\left\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(\widehat{U}_{it}-\tilde{U}_{it})\tilde{\delta}_{dt}'H'^{-1}(H'\tilde{f}_i-\widehat{f}_i)\right\|_\infty \leq O_P(\Delta_F)\left(\Delta_{ud}+\sqrt{\frac{\log p}{nT}}\right).$$

Also, by the assumption $\max_m |\frac{1}{nT}\sum_i\sum_t\tilde{U}_{it,m}(\widehat{f}_i - H'\tilde{f}_i)'H^{-1}\tilde{\delta}_{dt}| = O_P(\Delta_{ud})$, we have

$$\begin{aligned}\left\|\frac{1}{nT}\sum_t\widehat{U}_{t,\widehat{J}}'(\tilde{F}H-\widehat{F})H^{-1}\tilde{\delta}_{dt}\right\|_2 &\leq \max_m\left|\frac{1}{nT}\sum_i\sum_t\widehat{U}_{it,m}(\widehat{f}_i-H'\tilde{f}_i)'H^{-1}\tilde{\delta}_{dt}\right|O_P(\sqrt{|J|_0})\\ &\leq O_P\left(\Delta_{ud}+\Delta_F\sqrt{\frac{\log p}{nT}}\right)\sqrt{|J|_0}.\end{aligned}$$

Note that Lemma H.3 continues to hold when $\tilde{\epsilon}$ is replaced with $\tilde{\eta}$ and when γ_d is replaced with θ. ■

Lemma H.3. *(i)* $\frac{1}{\sqrt{nT}}\tilde{\epsilon}'M_{\widehat{U}_{\widehat{J}}}(I_T\otimes M_{\widehat{F}}(\tilde{F}H-\widehat{F})H^{-1})\tilde{\Delta}_d = o_P(1)$.
(ii) $\frac{1}{\sqrt{nT}}\tilde{\epsilon}'(I_T\otimes P_{\widehat{F}})\tilde{U}\gamma_d = o_P(1)$,
(iii) $\frac{1}{\sqrt{nT}}\tilde{\epsilon}'(I_T\otimes P_{\widehat{F}})\tilde{\eta} = o_P(1)$,
(iv) $\frac{1}{\sqrt{nT}}\tilde{\epsilon}'M_{\widehat{U}_{\widehat{J}}}\tilde{U}\gamma_d = o_P(1)$.
(v) $\frac{1}{\sqrt{nT}}\tilde{\epsilon}'M_{\widehat{U}_{\widehat{J}}}(I_T\otimes M_{\widehat{F}})\tilde{U}\gamma_d = o_P(1)$.
(vi) $\tilde{\Xi}'(I_T\otimes H'^{-1}(\tilde{F}H-\widehat{F})'M_{\widehat{F}})\frac{1}{\sqrt{nT}}M_{\widehat{U}_{\widehat{J}}}\tilde{U}\gamma_d = o_P(1)$.

Proof. (i) Note that

$$\left|\frac{1}{\sqrt{nT}}\sum_t \tilde{\epsilon}_t' P_{\widehat{F}}(\tilde{F}H-\widehat{F})H^{-1}\tilde{\delta}_{dt}\right|^2 = \frac{1}{nT}\operatorname{tr}^2\left(\sum_t \tilde{\delta}_{dt}\tilde{\epsilon}_t' P_{\widehat{F}}(\tilde{F}H-\widehat{F})H^{-1}\right)$$
$$\leq \frac{1}{nT}\left\|\sum_t \tilde{\delta}_{dt}\tilde{\epsilon}_t'\widehat{F}\right\|_F^2 \left\|\frac{1}{n}\widehat{F}'(\tilde{F}H-\widehat{F})\right\|_F^2 O_P(1)$$
$$= O_P(\Delta_F^2)\left\|\frac{1}{\sqrt{nT}}\sum_t \tilde{\delta}_{dt}\tilde{\epsilon}_t'(\widehat{F}-FH)\right\|_F^2$$
$$+ O_P(\Delta_F^2)\frac{1}{nT}\left\|\sum_t \tilde{\delta}_{dt}\tilde{\epsilon}_t' F\right\|_F^2 = o_P(1).$$

In addition,

$$\left\|\frac{1}{nT}\sum_t \widehat{U}_{t,\widehat{J}}'(\tilde{F}H-\widehat{F})H^{-1}\tilde{\delta}_{dt}\right\|_2 = O_P\left(\Delta_{ud}+\Delta_F\sqrt{\frac{\log p}{nT}}\right)\sqrt{|J|_0},$$

by Lemma H.2; and $\left\|\frac{1}{\sqrt{nT}}\sum_t \widehat{U}_{t,\widehat{J}}' P_{\widehat{F}}(\tilde{F}H-\widehat{F})H^{-1}\delta_{dt}\right\|_2 = 0$ due to $\widehat{U}_t'\widehat{F}=0$. It follows that

$$\left\|P_{\widehat{U}_{\widehat{J}}}\ (I_T\otimes M_{\widehat{F}}(\tilde{F}H-\widehat{F})H^{-1})\tilde{\Delta}_d\right\|_2$$
$$= \left\|\widehat{U}_{\widehat{J}}(\widehat{U}_{\widehat{J}}'\widehat{U}_{\widehat{J}})^{-1}\sum_t \widehat{U}_{t,\widehat{J}}' M_{\widehat{F}}(\tilde{F}H-\widehat{F})H^{-1}\delta_{dt}\right\|_2$$
$$\leq \left\|\frac{1}{\sqrt{nT}}\sum_t \widehat{U}_{t,\widehat{J}}' M_{\widehat{F}}(\tilde{F}H-\widehat{F})H^{-1}\delta_{dt}\right\|_2$$
$$\leq \left\|\frac{1}{\sqrt{nT}}\sum_t \widehat{U}_{t,\widehat{J}}'(\tilde{F}H-\widehat{F})H^{-1}\delta_{dt}\right\|_2 + \left\|\frac{1}{\sqrt{nT}}\sum_t \widehat{U}_{t,\widehat{J}}' P_{\widehat{F}}(\tilde{F}H-\widehat{F})H^{-1}\delta_{dt}\right\|_2$$
$$= O_P\left(\Delta_{ud}+\Delta_F\sqrt{\frac{\log p}{nT}}\right)\sqrt{|J|_0 nT}. \tag{H.10}$$

By Assumption D.4, $\|\frac{1}{\sqrt{nT}}\sum_t \tilde{\delta}_{dt}\tilde{\epsilon}_t'(\tilde{F}H-\widehat{F})\|_F = o_P(1)$. By Lemma H.1,

$$\|\frac{1}{\sqrt{nT}}P_{\widehat{U}_{\widehat{J}}}\tilde{\epsilon}\|_2^2 = O_P\left(|J|_0\frac{\log p}{nT}\right).$$

Hence, using the assumption $\left(\Delta_{ud} + \Delta_F \sqrt{\frac{\log p}{nT}}\right) \sqrt{\log p |J|_0^2} = o(1)$,

$$
\begin{aligned}
&\frac{1}{\sqrt{nT}} \tilde{\epsilon}' M_{\widehat{U}_{\hat{J}}} (I_T \otimes M_{\widehat{F}} (\tilde{F} H - \widehat{F}) H^{-1}) \tilde{\Delta}_d \\
&\quad = -\frac{1}{\sqrt{nT}} \tilde{\epsilon}' P_{\widehat{U}_{\hat{J}}} (I_T \otimes M_{\widehat{F}} (\tilde{F} H - \widehat{F}) H^{-1}) \tilde{\Delta}_d + \frac{1}{\sqrt{nT}} \tilde{\epsilon}' (I_T \otimes (\tilde{F} H - \widehat{F}) H^{-1}) \tilde{\Delta}_d \\
&\qquad - \frac{1}{\sqrt{nT}} \tilde{\epsilon}' (I_T \otimes P_{\widehat{F}} (\tilde{F} H - \widehat{F}) H^{-1}) \tilde{\Delta}_d \\
&\quad \leq \left\| \frac{1}{\sqrt{nT}} \tilde{\epsilon}' P_{\widehat{U}_{\hat{J}}} \right\|_2 \left\| P_{\widehat{U}_{\hat{J}}} (I_T \otimes M_{\widehat{F}} (\tilde{F} H - \widehat{F}) H^{-1}) \tilde{\Delta}_d \right\|_2 \\
&\qquad + \left| \frac{1}{\sqrt{nT}} \sum_t \tilde{\epsilon}_t' P_{\widehat{F}} (\tilde{F} H - \widehat{F}) H^{-1} \tilde{\delta}_{dt} \right| + \left| \frac{1}{\sqrt{nT}} \sum_t \tilde{\epsilon}_t' (\tilde{F} H - \widehat{F}) H^{-1} \tilde{\delta}_{dt} \right| \\
&\quad = o_P(1) + O_P\left(\sqrt{|J|_0 \frac{\log p}{nT}}\right) \left(\Delta_{ud} + \Delta_F \sqrt{\frac{\log p}{nT}}\right) \sqrt{|J|_0 nT} = o_P(1).
\end{aligned}
$$

(ii) Note that

$$
\begin{aligned}
\frac{1}{T} \sum_{t=1}^T \tilde{U}_{jt} \tilde{\epsilon}_{it} = \frac{1}{T} \sum_{t=1}^T U_{jt} \epsilon_{it} - \bar{U}_{j \cdot} \bar{\epsilon}_{i \cdot} - \frac{1}{T} \sum_{t=1}^T U_{jt} \bar{\epsilon}_{\cdot t} + \bar{U}_{j \cdot} \bar{\bar{\epsilon}} + \bar{\bar{U}} \bar{\epsilon}_{i \cdot} - \bar{\bar{U}} \bar{\bar{\epsilon}} \\
- \frac{1}{T} \sum_{t=1}^T \bar{U}_{\cdot t} \epsilon_{it} + \frac{1}{T} \sum_{t=1}^T \bar{U}_{\cdot t} \bar{\epsilon}_{\cdot t}.
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\frac{1}{\sqrt{nT}} \tilde{\epsilon}' (I_T \otimes P_{\widehat{F}}) \tilde{U} \gamma_d &= \frac{1}{\sqrt{nT}} \sum_t \tilde{\epsilon}_t' P_{\widehat{F}} \tilde{U}_t \gamma_d \\
&\leq \left\| \frac{\sqrt{T}}{n\sqrt{n}} \sum_{m=1}^p \widehat{F}' \frac{1}{T} \sum_t \tilde{U}_{t,m} \tilde{\epsilon}_t' \widehat{F} \gamma_{dm} \right\|_F \qquad \text{(H.11)} \\
&\leq \sum_{l=1}^8 a_l
\end{aligned}
$$

where each a_l is defined below and can be bounded employing Lemma H.2. Specifically,

$$
a_1 = \left\| \frac{\sqrt{T}}{n\sqrt{n}} \sum_{m=1}^p \widehat{F}' \frac{1}{T} \sum_t U_{t,m} \epsilon_t' \widehat{F} \gamma_{dm} \right\|_F = o_P(1).
$$

$$
\begin{aligned}
a_2 &= \sqrt{nT}\left\|\frac{1}{n}\sum_{j=1}^{n}\widehat{f}_j\bar{U}_{j\cdot}'\gamma_d\right\|_2\left\|\frac{1}{n}\sum_{i=1}^{n}\bar{\epsilon}_{i\cdot}\widehat{f}_i'\right\|_2\\
&= O_P\left(\sqrt{nT}\right)\left(\Delta_F|J|_0+\sqrt{\frac{|J|_0}{nT}}\right)\left(\Delta_F+\frac{1}{\sqrt{nT}}\right)=o_P(1).\\
a_3 &= O_P(1)\left\|\frac{\sqrt{nT}}{n}\sum_{m=1}^{p}\sum_{j=1}^{n}\widehat{f}_j\frac{1}{T}\sum_{t=1}^{T}U_{jt,m}\bar{\epsilon}_{\cdot t}\gamma_{dm}\right\|_F=o_P(1).\\
a_4 &= O_P(1)\left\|\frac{\sqrt{nT}}{n}\sum_{j=1}^{n}\widehat{f}_j\bar{U}_{j\cdot}'\gamma_d\right\|_2|\bar{\bar{\epsilon}}|\\
&= O_P\left(\sqrt{nT}\right)\left(\Delta_F|J|_0+\sqrt{\frac{|J|_0}{nT}}\right)\frac{1}{\sqrt{nT}}=o_P(1).\\
a_5 &= O_P(1)\left\|\frac{\sqrt{nT}}{n}\sum_{i=1}^{n}\bar{\epsilon}_{i\cdot}\widehat{f}_i'\right\|_2\left|\bar{\bar{U}}'\gamma_d\right|\\
&= O_P(\sqrt{nT})\left(\Delta_F+\frac{1}{\sqrt{nT}}\right)\sqrt{\frac{|J|_0}{nT}}=o_P(1).\\
a_6 &= O_P(1)\sqrt{nT}\,|\bar{\bar{\epsilon}}|\left|\bar{\bar{U}}'\gamma_d\right|=O_P\left(\sqrt{\frac{|J|_0}{nT}}\right)=o_P(1).\\
a_7 &= O_P(1)\left\|\frac{\sqrt{nT}}{n}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=1}^{T}\bar{U}_{\cdot t}'\gamma_d\epsilon_{it}\widehat{f}_i'\right\|_2\le o_P(1).\\
a_8 &= O_P(1)\sqrt{nT}\left|\sum_{m=1}^{p}\frac{1}{T}\sum_{t=1}^{T}\bar{U}_{\cdot t,m}\bar{\epsilon}_{\cdot t}\gamma_{dm}\right|\le o_P(1).
\end{aligned}
$$

where we used $\sqrt{nT}\Delta_F^2|J|_0=o(1)$.

(iii) The third conclusion of the lemma follows from an argument similar to that used to establish (ii). We have

$$
\frac{1}{\sqrt{nT}}\tilde{\epsilon}'(I_T\otimes P_{\widehat{F}})\tilde{\eta}\le\left\|\frac{\sqrt{T}}{n\sqrt{n}}\widehat{F}'\frac{1}{T}\sum_t\tilde{\eta}_t\tilde{\epsilon}_t'\widehat{F}\right\|_F\le\sum_{l=1}^{8}b_l=o_P(1)
$$

which follows from the bounds

$$
b_1=\left\|\sqrt{nT}\frac{1}{n^2}\sum_i\sum_j\widehat{f}_i\frac{1}{T}\sum_t\eta_{it}\epsilon_{jt}\widehat{f}_j'\right\|_F=O_P\left(\Delta_F^2\sqrt{Tn}+\Delta_F\sqrt{T}\right)=o_P(1)
$$

$$b_2 = \sqrt{nT}\left\|\frac{1}{n}\sum_{j=1}^{n}\widehat{f}_j\bar{\eta}_{j\cdot}\|_2\right\|\frac{1}{n}\sum_{i=1}^{n}\bar{\epsilon}_{i\cdot}\widehat{f}_i'\|_2 = O_P\left(\sqrt{nT}\right)\left(\Delta_F + \frac{1}{\sqrt{nT}}\right)^2 = o_P(1)$$

$$b_3 = O_P(1)\left\|\frac{\sqrt{nT}}{n}\sum_{j=1}^{n}\widehat{f}_j\frac{1}{T}\sum_{t=1}^{T}\eta_{jt}\bar{\epsilon}_{\cdot t}\right\|_F = o_P(1)$$

$$b_4 = O_P(1)\left\|\frac{\sqrt{nT}}{n}\sum_{j=1}^{n}\widehat{f}_j\bar{\eta}_{j\cdot}\right\|_2|\bar{\bar{\epsilon}}| = O_P\left(\sqrt{nT}\right)\left(\Delta_F + \sqrt{\frac{1}{nT}}\right)\frac{1}{\sqrt{nT}} = o_P(1)$$

$$b_5 = O_P(1)\left\|\frac{\sqrt{nT}}{n}\sum_{i=1}^{n}\bar{\epsilon}_{i\cdot}\widehat{f}_i'\right\|_2|\bar{\bar{\eta}}| = O_P\left(\sqrt{nT}\right)\left(\Delta_F + \frac{1}{\sqrt{nT}}\right)\sqrt{\frac{1}{nT}} = o_P(1)$$

$$b_6 = O_P(1)\sqrt{nT}\,|\bar{\bar{\epsilon}}|\,|\bar{\bar{\eta}}| = O_P\left(\sqrt{\frac{1}{nT}}\right) = o_P(1)$$

$$b_7 = O_P(1)\left\|\frac{\sqrt{nT}}{n}\sum_{i=1}^{n}\frac{1}{T}\sum_{t=1}^{T}\bar{\eta}_{\cdot t}\epsilon_{it}\widehat{f}_i'\right\|_2 \le o_P(1)$$

$$b_8 = O_P(1)\sqrt{nT}\left|\frac{1}{T}\sum_{t=1}^{T}\bar{\eta}_{\cdot t}\bar{\epsilon}_{\cdot t}\right| \le o_P(1)$$

where we used that $\sqrt{nT}\Delta_F^2 = o(1)$.

(iv) By (H.1), $\widehat{U}\widehat{m} = P_{\widehat{U}_{\widehat{J}}}\widehat{U}\gamma_d$. We first bound $\|\gamma_d - \widehat{m}\|_1$. By Proposition F.1,

$$\frac{1}{nT}\|\widehat{U}(\gamma_d - \widehat{m})\|_2^2 \le \frac{1}{nT}\sum_{t=1}^{T}\|\widehat{U}_t(\gamma_d - \tilde{\gamma})\|_2^2 = O_P(\|R_y\|_1^2 + \kappa_n^2|J|_0).$$

We also have $\gamma_d - \widehat{m}$ is at most $|\widehat{J}|_0 + |J|_0$-sparse. Hence, by Assumption 3.4 and using a proof similar to that used to show (H.4), we have that

$$\begin{aligned}\frac{1}{n}\left\|\widehat{U}(\gamma_d - \widehat{m})\right\|_2^2 &\ge (\phi_{\min}(|\widehat{J}|_0 + |J|_0) - o_P(1))\|\gamma_d - \widehat{m}\|_2^2\\ &\ge C(|J|_0 + |\widehat{J}|_0)^{-1}\|\gamma_d - \widehat{m}\|_1^2\end{aligned}$$

with probability approaching one. Hence,

$$\|\gamma_d - \widehat{m}\|_1^2 = O_P(\|R_y\|_1^2 + \kappa_n^2|J|_0)(|J|_0 + |\widehat{J}|_0) = O_P(\|R_y\|_1^2|J|_0 + \kappa_n^2|J|_0^2).$$

In addition, $\|\frac{1}{\sqrt{nT}}\tilde{\epsilon}'\widehat{U}\|_\infty = O_P\left(\sqrt{\log p}\right)$ by an argument similar to that in the proof of Lemma F.1. Hence,

$$\begin{aligned}\left|\frac{1}{\sqrt{nT}}\tilde{\epsilon}' M_{\widehat{U}_{\widehat{J}}}\widehat{U}\gamma_d\right| &= \left|\frac{1}{\sqrt{nT}}\tilde{\epsilon}'(I - P_{\widehat{U}_{\widehat{J}}})\widehat{U}\gamma_d\right| \\ &= \left|\frac{1}{\sqrt{nT}}\tilde{\epsilon}'\widehat{U}(\gamma_d - \widehat{m})\right| \\ &\le \left\|\frac{1}{\sqrt{nT}}\tilde{\epsilon}'\widehat{U}\right\|_\infty \|\gamma_d - \widehat{m}\|_1 \\ &= O_P\left(\|R_y\|_1 |J|_0^{1/2} + \kappa_n |J|_0\right)\sqrt{\log p} = o_P(1)\end{aligned}$$

using $\|R_y\|_1^2 |J|_0 \log p = o(1)$ and $\kappa_n |J|_0 \sqrt{\log p} = o(1)$.

Next,

$$\begin{aligned}\frac{1}{\sqrt{nT}}\tilde{\epsilon}' M_{\widehat{U}_{\widehat{J}}}(\tilde{U} - \widehat{U})\gamma_d &= \frac{1}{\sqrt{nT}}\tilde{\epsilon}'(\tilde{U} - \widehat{U})\gamma_d - \frac{1}{\sqrt{nT}}\tilde{\epsilon}' P_{\widehat{U}_{\widehat{J}}}(\tilde{U} - \widehat{U})\gamma_d \\ &= \frac{1}{\sqrt{nT}}\sum_i\sum_t\sum_m \tilde{\epsilon}_{it}(\tilde{U}_{it,m} - \widehat{U}_{it,m})\gamma_{dm} \\ &\quad - \frac{1}{\sqrt{nT}}\tilde{\epsilon}'\widehat{U}_{\widehat{J}}(\widehat{U}_{\widehat{J}}'\widehat{U}_{\widehat{J}})^{-1}\widehat{U}_{\widehat{J}}'(\tilde{U} - \widehat{U})\gamma_d \\ &\le \left\|\frac{1}{\sqrt{nT}}\sum_i\sum_m\sum_t \tilde{\epsilon}_{it}\gamma_{dm}\tilde{\lambda}_{tm}\tilde{f}_i'\right\|_F O_P(\Delta_F) \qquad \text{(H.12)} \\ &\quad + \left\|\frac{1}{\sqrt{nT}}\sum_i\sum_t\sum_m \tilde{\epsilon}_{it}\widehat{f}_i\frac{1}{n}\widehat{F}'\tilde{U}_{t,m}\gamma_{dm}\right\|_F O_P(1) \\ &\quad + \left\|\frac{1}{\sqrt{nT}}\sum_i\sum_t\sum_m \tilde{\epsilon}_{it}\gamma_{dm}\tilde{\lambda}_{tm}(\widehat{f}_i - H'\tilde{f}_i)'\right\| O_P(1) \\ &\quad + \left\|\frac{1}{\sqrt{nT}}\tilde{\epsilon}'\widehat{U}_{\widehat{J}}\|_2\|\frac{1}{nT}\widehat{U}_{\widehat{J}}'(\tilde{U} - \widehat{U})\gamma_d\right\|_2 O_P(1).\end{aligned}$$

The first term in the last inequality in (H.12) is $O_P(|J|_0)(\Delta_F) = o_P(1)$. By (H.11), the second term in the last inequality in (H.12) is $o_P(1)$. The third term in the last inequality in (H.12) is $o_P(1)$ by assumption. For the last term in the last inequality in (H.12), we have $\|\frac{1}{\sqrt{nT}}\widehat{U}_{\widehat{J}}'\tilde{\epsilon}\|_2 = O_P(\sqrt{|J|_0 \log p})$ by a proof similar to that of (H.3). Then, Lemma F.1 (v) yields

$$\frac{1}{nT}\left\|\widehat{U}_{\widehat{J}}'(\tilde{U} - \widehat{U})\gamma_d\right\|_2 \le \sqrt{|\widehat{J}|_0}\max_k\left|\frac{1}{nT}\sum_i\sum_t\sum_m \widehat{U}_{it,k}(\tilde{U}_{it,m} - \widehat{U}_{it,m})\gamma_{dm}\right|$$

$$\leq O_P\left(|J|_0^{3/2}\right) \max_{mk} \left| \frac{1}{nT} \sum_i \sum_t \sum_m \widehat{U}_{it,k}(\tilde{U}_{it,m} - \widehat{U}_{it,m}) \right|$$
$$\leq O_P\left(|J|_0^{3/2}\right) \left(\sqrt{\frac{\log p}{nT}} \Delta_F + \Delta_{ud} + \frac{\log(pT)}{n} + \Delta_{fum}^2 \right).$$

The last term of (H.12) is thus bounded by

$$\left(|J|_0^2 \sqrt{\log p}\right) O_P\left(\sqrt{\frac{\log p}{nT}} \Delta_F + \Delta_{ud} + \frac{\log(pT)}{n} + \Delta_{fum}^2 \right) = o_P(1).$$

Therefore, both $\frac{1}{\sqrt{nT}} \tilde{\epsilon}' M_{\widehat{U}_{\widehat{J}}} (\tilde{U} - \widehat{U})\gamma_d$ and $\frac{1}{\sqrt{nT}} \tilde{\epsilon}' M_{\widehat{U}_{\widehat{J}}} \widehat{U} \gamma_d$ are $o_P(1)$.

(v) Since $\frac{1}{\sqrt{nT}} \tilde{\epsilon}' P_{\widehat{U}_{\widehat{J}}} (I_T \otimes P_{\widehat{F}}) \tilde{U} \gamma_d = 0$, it follows from part (ii) and Lemma H.3(ii) and (iii) that

$$\frac{1}{\sqrt{nT}} \tilde{\epsilon}' M_{\widehat{U}_{\widehat{J}}} (I_T \otimes M_{\widehat{F}}) \tilde{U} \gamma_d = \frac{1}{\sqrt{nT}} \tilde{\epsilon}' M_{\widehat{U}_{\widehat{J}}} \tilde{U} \gamma_d - \frac{1}{\sqrt{nT}} \tilde{\epsilon}' (I_T \otimes P_{\widehat{F}}) \tilde{U} \gamma_d = o_P(1).$$

(vi) By Lemma H.1, $\left\| \frac{1}{\sqrt{nT}} M_{\widehat{U}_{\widehat{J}}} \tilde{U} \gamma \right\|_2^2 = O_P\left(\|R_y\|_1^2 + \kappa_n^2 |J|_0 + |J|_0^2 \Delta_F^2 + \frac{|J|_0}{n} \right)$. Thus, under $Tn\Delta_F^2 \left(\kappa_n^2 |J|_0 + \|R_y\|_1^2 + \Delta_F^2 |J|_0^2 + \frac{|J|_0}{n} \right) = o(1)$,

$$|\tilde{\Xi}'(I_T \otimes H^{'-1} (\tilde{F}H - \widehat{F})' M_{\widehat{F}}) \frac{1}{\sqrt{nT}} M_{\widehat{U}_{\widehat{J}}} \tilde{U} \gamma_d|^2$$
$$\leq O_P(T) \left\| \widehat{F} - \tilde{F}H \right\|_F^2 \left\| \frac{1}{\sqrt{nT}} M_{\widehat{U}_{\widehat{J}}} \tilde{U} \gamma_d \right\|_2^2$$
$$= O_P\left(Tn\Delta_F^2\right) \left(\kappa_n^2 |J|_0 + \|R_y\|_1^2 + \Delta_F^2 |J|_0^2 + \frac{|J|_0}{n} \right) = o_P(1).$$

The following technical lemma is used to prove the consistency of the asymptotic variance estimator. ■

Lemma H.4. *(i) Both* $\frac{1}{nT} \sum_{i=1}^n \left[\sum_{t=1}^T (\widehat{\eta}_{it} - \tilde{\eta}_{it}) \tilde{\epsilon}_{it} \right]^2 = o_P(1)$ *and* $\frac{1}{nT} \sum_{i=1}^n \left[\sum_{t=1}^T (\widehat{\epsilon}_{it} - \tilde{\epsilon}_{it}) \tilde{\eta}_{it} \right]^2 = o_P(1)$.
(ii) $\frac{1}{nT} \sum_{i=1}^n \left[\sum_{t=1}^T (\widehat{\eta}_{it} \widehat{\epsilon}_{it} - \tilde{\eta}_{it} \tilde{\epsilon}_{it}) \right]^2 = o_P(1)$.

Proof. (i) We first bound $\max_{i \leq n} \frac{1}{T} \sum_{t=1}^T \tilde{\epsilon}_{it}^2$. We do this by bounding $\max_{i \leq n} \left| \frac{1}{T} \sum_{t=1}^T \epsilon_{it}^2 - E\epsilon_{it}^2 \right|$ and $\max_{i \leq n} \left(\frac{1}{T} \sum_t \epsilon_{it} \right)^2$.

When T is fixed,

$$\max_{i\leq n}\left|\frac{1}{T}\sum_{t=1}^{T}\epsilon_{it}^2 - E\epsilon_{it}^2\right| \leq \max_{it}|\epsilon_{it}^2 - E\epsilon_{it}^2| = O_P(\log n) = O_P\left(\frac{\log n}{\sqrt{T}}\right). \quad (T \text{ does not grow})$$

and $\max_{i\leq n}\left(\frac{1}{T}\sum_t \epsilon_{it}\right)^2 \leq \max_{it}\epsilon_{it}^2 = O_P(\log n) = O_P(\frac{\log n}{T})$.

When $T \to \infty$, recall that by assumption, $\log^\gamma p = o(n)$ for some $\gamma > 2$ and that strong mixing (Assumption 3.1) holds for the process $\{(\eta_t, \epsilon_t)\}_{t=-\infty}^{+\infty}$ with mixing coefficient bounded by $\exp(-CT^r)$, $r > 1$. Note that we also assume $\gamma r \geq 2$. Let $\bar{r} = \min\{r, 1\}$, $r_1 = (0.5 + \bar{r}^{-1})^{-1}$, $c = 0.5(\gamma + 1)$, then $\gamma\bar{r} \geq 2$, $r_1 < 1$, $cr_1 > 1$ and $2c \geq 1$. Because $\bar{r} \leq r$, the strong mixing condition in Assumption 3.1 also holds with $\bar{r}$ in place of r. The Bernstein inequality for weakly dependent data of Merlevède et al. (2011) requires (a) exponential tails (fulfilled by the sub-Gaussian condition in Assumption 3.1), and (b) a strong mixing condition (also assumed in Assumption 3.1). The introduction of $\bar{r}$ is to ensure that the so-defined $r_1 < 1$, another requirement of applying the Bernstein inequality for strong mixing sequences. Then, by Theorem 1 of Merlevède et al. (2011) (proved using a “coupling argument” of Dedecker and Prieur (2004) - see also Wang et al. (2016) for a similar argument), for $y = M\frac{(\log n)^c}{\sqrt{T}}$, and sufficiently large $M > 0$, we have

$$P\left(\max_{i\leq n}\left|\frac{1}{T}\sum_{t=1}^{T}\epsilon_{it}^2 - E\epsilon_{it}^2\right| > y\right) \leq A_1 + A_2 + A_3$$

where

$$\begin{aligned}
A_1 &= nT\exp(-C(Ty)^{r_1}) = \exp(\log(nT) - CM^{r_1}T^{r_1/2}\log^{cr_1} n) = o(1), \quad (cr_1 \geq 1)\\
A_2 &= n\exp\left(-C\frac{(Ty)^2}{T}\exp\left(\frac{(Ty)^{r_1(1-r_1)}}{C\log^{r_1}(Ty)}\right)\right) = o(1), \quad (r_1 < 1, 2c > 1, Ty \to \infty)\\
A_3 &= n\exp\left(-CTy^2\right) = \exp(\log n - CM^2\log^{2c} n) = o(1).
\end{aligned}$$

Hence $\max_{i\leq n}\left|\frac{1}{T}\sum_{t=1}^{T}\epsilon_{it}^2 - E\epsilon_{it}^2\right| = O_P\left(\frac{(\log n)^c}{\sqrt{T}}\right) = O_P\left(\frac{(\log n)^{0.5+0.5\gamma}}{\sqrt{T}}\right)$. Similarly, $\max_{i\leq n}|\frac{1}{T}\sum_{t=1}^{T}\epsilon_{it}| = O_P\left(\frac{(\log n)^{0.5+0.5\gamma}}{\sqrt{T}}\right)$. Then combining the two cases $T \nrightarrow \infty$ and $T \to \infty$, we have

$$\begin{aligned}
\max_{i\leq n}\frac{1}{T}\sum_{t=1}^{T}\epsilon_{it}^2 &\leq \max_{i\leq n}E\epsilon_{it}^2 + \max_{i\leq n}\left|\frac{1}{T}\sum_{t=1}^{T}\epsilon_{it}^2 - E\epsilon_{it}^2\right|\\
&= O_P\left(1 + \frac{(\log n)^{0.5+0.5\gamma}}{\sqrt{T}}\right),
\end{aligned}$$

$$\max_{i\le n}\bar{\epsilon}_{i\cdot}^2 = \max_{i\le n}\left(\frac{1}{T}\sum_t \epsilon_{it}\right)^2 = O_P\left(\frac{\log^{1+\gamma} n}{T}\right),$$

$$\frac{1}{T}\sum_{t=1}^T \bar{\epsilon}_{\cdot t}^2 = \frac{1}{T}\sum_{t=1}^T\left(\frac{1}{n}\sum_i \epsilon_{it}\right)^2 = O_P\left(\frac{1}{n}\right),$$

and

$$\bar{\bar{\epsilon}}^2 = o_P(1).$$

Hence, $\max_{i\le n}\frac{1}{T}\sum_{t=1}^T\tilde{\epsilon}_{it}^2 = O_P\left(1+\frac{(\log n)^{0.5+0.5\gamma}}{\sqrt{T}}+\frac{\log^{1+\gamma} n}{T}\right)$. Thus, by the Cauchy-Schwarz inequality and with ψ_n defined in (A.6),

$$\begin{aligned}\frac{1}{nT}\sum_{i=1}^n\left[\sum_{t=1}^T(\widehat{\eta}_{it}-\tilde{\eta}_{it})\tilde{\epsilon}_{it}\right]^2 &\le \frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T(\widehat{\eta}_{it}-\tilde{\eta}_{it})^2\sum_{s=1}^T\tilde{\epsilon}_{is}^2\\ &\le \frac{T}{nT}\|\widehat{\eta}-\tilde{\eta}\|_2^2\max_{i\le n}\frac{1}{T}\sum_{t=1}^T\tilde{\epsilon}_{it}^2 = O_P(T\psi_n^2)\left(1+\frac{(\log n)^{0.5+0.5\gamma}}{\sqrt{T}}+\frac{\log^{1+\gamma} n}{T}\right)\\ &= O_P\left(\psi_n^2 T+\psi_n^2\log^{1+\gamma} n+\psi_n^2(T\log^{1+\gamma} n)^{1/2}\right) = o_P(1).\end{aligned}$$

where the last equality may be directly verified under our assumptions.

The same arguments as used in the proof of Proposition A.1 yield that

$$\frac{1}{nT}\|\widehat{e}-\tilde{e}\|_2^2 = O_P(\psi_n^2).$$

In addition, the first statement of Theorem 3.1 implies $|\alpha-\widehat{\alpha}|^2 = O_P\left(\frac{1}{nT}\right)$. Thus, from $\widehat{\epsilon}=\widehat{e}-\widehat{\alpha}\widehat{\eta}$, we have

$$\begin{aligned}\frac{1}{nT}\|\widehat{\epsilon}-\tilde{\epsilon}\|_2^2 &\le 2\frac{1}{nT}\|\widehat{e}-\tilde{e}\|_2^2+4\frac{1}{nT}\|(\alpha-\widehat{\alpha})\tilde{\eta}\|_2^2+4\frac{1}{nT}\|\widehat{\alpha}(\tilde{\eta}-\widehat{\eta})\|_2^2\\ &= O_P(\psi_n^2)\end{aligned}$$

for ψ_n defined in (A.6) and $\frac{1}{nT}=O(\psi_n^2)$. In addition, $\max_{i\le n}\frac{1}{T}\sum_{t=1}^T\tilde{\eta}_{it}^2 = O_P\left(1+\frac{\log n}{T}\right)$ follows from the same argument. Hence,

$$\begin{aligned}\frac{1}{nT}\sum_{i=1}^n\left[\sum_{t=1}^T(\widehat{\epsilon}_{it}-\tilde{\epsilon}_{it})\tilde{\eta}_{it}\right]^2 &\le \tfrac{T}{nT}\|\widehat{\epsilon}-\tilde{\epsilon}\|_2^2\max_{i\le n}\tfrac{1}{T}\textstyle\sum_{t=1}^T\tilde{\eta}_{it}^2\\ &= O_P\left(\psi_n^2 T+\psi_n^2\log n\right) = o_P(1).\end{aligned}$$

(ii) By part (i),

$$\begin{aligned}\frac{1}{nT}\sum_{i=1}^n\left[\sum_{t=1}^T(\widehat\eta_{it}\widehat\epsilon_{it}-\tilde\eta_{it}\tilde\epsilon_{it})\right]^2 &\leq \frac{2}{nT}\sum_{i=1}^n\left[\sum_{t=1}^T(\widehat\eta_{it}-\tilde\eta_{it})\tilde\epsilon_{it}\right]^2+\frac{2}{nT}\sum_{i=1}^n\left[\sum_{t=1}^T\widehat\eta_{it}(\widehat\epsilon_{it}-\tilde\epsilon_{it})\right]^2\\ &\leq o_P(1)+\frac{4T}{n}\sum_{i=1}^n\left[\frac{1}{T}\sum_{t=1}^T(\widehat\eta_{it}-\tilde\eta_{it})(\widehat\epsilon_{it}-\tilde\epsilon_{it})\right]^2\\ &\leq o_P(1)+\frac{4}{nT}\|\widehat\epsilon-\tilde\epsilon\|_2^2\max_i\sum_t(\widehat\eta_{it}-\tilde\eta_{it})^2\\ &= o_P(1)+O_P(T\psi_n^2)\max_i\frac{1}{T}\sum_t(\widehat\eta_{it}-\tilde\eta_{it})^2. \qquad \text{(H.13)}\end{aligned}$$

We now show the second term on the right of (H.13) is $o_P(1)$. Using the equalities

$$\tilde U_{it,m}-\widehat U_{it,m}=\widehat f_i'(\widehat F'\widehat F)^{-1}\widehat F'(\tilde FH-\widehat F)H^{-1}\tilde\lambda_{tm}+\widehat f_i'(\widehat F'\widehat F)^{-1}\widehat F'\tilde U_{t,m}+\tilde\lambda_{tm}'H'^{-1}(\widehat f_i-H'\tilde f_i)$$

and

$$\widehat\delta_{dt}-H^{-1}\tilde\delta_{dt}=(\widehat F'\widehat F)^{-1}\widehat F'(\tilde FH-\widehat F)H^{-1}\tilde\delta_{dt}+(\widehat F'\widehat F)^{-1}\widehat F'\tilde U_t\gamma_d+(\widehat F'\widehat F)^{-1}\widehat F'\tilde\eta_t,$$

we have

$$\begin{aligned}\tilde\eta_{it}-\widehat\eta_{it}&=\widehat f_i'(\widehat\delta_{dt}-H^{-1}\tilde\delta_{dt})+\tilde\delta_{dt}'H^{-1'}(\widehat f_i-H'\tilde f_i)\\ &\quad+\sum_m(\widehat U_{it,m}-\tilde U_{it,m})\widehat\gamma_{dm}+\tilde U_{it}'(\widehat\gamma_d-\gamma_d)\\ &=\widehat f_i'(\widehat F'\widehat F)^{-1}\widehat F'(\tilde FH-\widehat F)H^{-1}\tilde\delta_{dt}+\widehat f_i'(\widehat F'\widehat F)^{-1}\widehat F'\tilde U_t\gamma_d\\ &\quad+\widehat f_i'(\widehat F'\widehat F)^{-1}\widehat F'\tilde\eta_t+\tilde\delta_{dt}'H^{-1'}(\widehat f_i-H'\tilde f_i)\\ &\quad+\sum_m\widehat f_i'(\widehat F'\widehat F)^{-1}\widehat F'(\tilde FH-\widehat F)H^{-1}\tilde\lambda_{tm}\widehat\gamma_{dm}+\sum_m\widehat f_i'(\widehat F'\widehat F)^{-1}\widehat F'\tilde U_{t,m}\widehat\gamma_{dm}\qquad\text{(H.14)}\\ &\quad+\sum_m\tilde\lambda_{tm}'H'^{-1}(\widehat f_i-H'\tilde f_i)\widehat\gamma_{dm}+\tilde U_{it}'(\widehat\gamma_d-\gamma_d)\\ &:=\sum_{l=1}^8 C_{l,it}.\end{aligned}$$

Since f_i is sub-Gaussian, $\max_i\|\tilde f_i\|_2=O_P(\sqrt{\log n})$. Under

$$\max_i\|\widehat f_i-H'f_i\|_2=O_P(1),$$

we have that $\max_i \|\widehat{f}_i\|_2 = O_P(\sqrt{\log n})$. By Lemma H.11(iv),

$$\frac{1}{T}\sum_t \left\|\frac{1}{n}\widehat{F}'\tilde{U}_t\gamma\right\|_2^2 = O_P\left(\frac{|J|_0}{n} + |J|_0^2\Delta_F^2\right),$$

and

$$\frac{1}{T}\sum_t \left\|\frac{1}{n}\widehat{F}'\tilde{\eta}_t\right\|_2^2 = O_P\left(\frac{1}{n} + \Delta_{fe}^2\right).$$

Hence,

$$\begin{aligned}
T\psi_n^2 \max_i \frac{1}{T}\sum_t C_{1,it}^2 &= T\psi_n^2 \max_i \frac{1}{T}\sum_t (\widehat{f}_i'(\widehat{F}'\widehat{F})^{-1}\widehat{F}'(\tilde{F}H - \widehat{F})H^{-1}\tilde{\delta}_{dt})^2 \\
&= O_P(T\psi_n^2 \log n) O_P(\Delta_F^2) \\
T\psi_n^2 \max_i \frac{1}{T}\sum_t C_{2,it}^2 &= T\psi_n^2 \max_i \frac{1}{T}\sum_t (\widehat{f}_i'(\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{U}_t\gamma_d)^2 \\
&= O_P(T\psi_n^2 \log n)\frac{1}{T}\sum_t \|\frac{1}{n}\widehat{F}'\tilde{U}_t\gamma_d\|_2^2 \\
&= O_P(T\psi_n^2 \log n) O_P\left(\frac{|J|_0}{n} + |J|_0^2\Delta_F^2\right) \\
T\psi_n^2 \max_i \frac{1}{T}\sum_t C_{3,it}^2 &= T\psi_n^2 \max_i \frac{1}{T}\sum_t (\widehat{f}_i'(\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{\eta}_t)^2 \\
&= O_P(T\psi_n^2 \log n)\frac{1}{T}\sum_t \left\|\frac{1}{n}\widehat{F}'\tilde{\eta}\right\|_2^2 \\
&= O_P(T\psi_n^2 \log n) O_P\left(\frac{1}{n} + \Delta_{fe}^2\right) \\
T\psi_n^2 \max_i \frac{1}{T}\sum_t C_{4,it}^2 &= T\psi_n^2 \max_i \frac{1}{T}\sum_t (\tilde{\delta}_{dt}' H^{-1\prime}(\widehat{f}_i - H'\tilde{f}_i))^2 \\
&= O_P(T\psi_n^2) \max_i \|\widehat{f}_i - H' f_i\|_2^2 \\
T\psi_n^2 \max_i \frac{1}{T}\sum_t C_{5,it}^2 &= T\psi_n^2 \max_i \frac{1}{T}\sum_t (\sum_m \widehat{f}_i'(\widehat{F}'\widehat{F})^{-1}\widehat{F}'(\tilde{F}H - \widehat{F})H^{-1}\tilde{\lambda}_{tm}\widehat{\gamma}_{dm})^2 \\
&= O_P(T\psi_n^2 \log n)|J|_0^2 O_P(\Delta_F^2) \\
T\psi_n^2 \max_i \frac{1}{T}\sum_t C_{6,it}^2 &= T\psi_n^2 \max_i \frac{1}{T}\sum_t (\sum_m \widehat{f}_i'(\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{U}_{t,m}\widehat{\gamma}_{dm})^2 \\
&= O_P(T\psi_n^2 \log n) O_P\left(\frac{|J|_0}{n} + |J|_0^2\Delta_F^2\right)
\end{aligned}$$

$$T\psi_n^2 \max_i \frac{1}{T}\sum_t C_{7,it}^2 = T\psi_n^2 \max_i \frac{1}{T}\sum_t (\sum_m \tilde{\lambda}_{tm}' H'^{-1}(\widehat{f}_i - H'\tilde{f}_i)\widehat{\gamma}_{dm})^2$$
$$= O_P(T\psi_n^2)\max_i \|\widehat{f}_i - H' f_i\|_2^2 |J|_0^2$$
$$T\psi_n^2 \max_i \frac{1}{T}\sum_t C_{8,it}^2 = T\psi_n^2 \max_i \frac{1}{T}\sum_t (\tilde{U}_{it}'(\widehat{\gamma}_d - \gamma_d))^2$$
$$\leq T\psi_n^2 \|\widehat{\gamma}_d - \gamma_d\|_1^2 \max_{imk}\left|\frac{1}{T}\sum_t \tilde{U}_{it,m}\tilde{U}_{it,k}\right|$$
$$= O_P(T\psi_n^2)(\kappa_n^2|J|_0^2 + \|R_d\|_1^2)\left(1+\sqrt{\frac{\log(pn)}{T}}\right).$$

Therefore, $T\psi_n^2 \max_i \frac{1}{T}\sum_t (\widehat{\eta}_{it} - \tilde{\eta}_{it})^2 = O_P(T\psi_n^2 c_n)$ where

$$c_n = \log n\left(\Delta_F^2|J|_0^2 + \frac{|J|_0}{n} + |J|_0^2\Delta_F^2 + \Delta_{fe}^2\right)$$
$$+\max_i \|\widehat{f}_i - H' f_i\|_2^2|J|_0^2 + \left(\kappa_n^2|J|_0^2 + \|R_d\|_1^2\right)\left(1+\sqrt{\frac{\log(pn)}{T}}\right).$$

Under our assumptions $T\psi_n^2 c_n = o(1)$. Then (H.13) implies

$$\frac{1}{nT}\sum_{i=1}^n\left[\sum_{t=1}^T (\widehat{\eta}_{it}\widehat{\epsilon}_{it} - \tilde{\eta}_{it}\tilde{\epsilon}_{it})\right]^2 = o_P(1).$$

■

H.2. **Technical Lemmas for Theorem 4.1.** The following lemma presents a version of the maximal inequality for a U-statistic. While there are several inequalities of this type available in the literature, e.g., Nolan and Pollard (1987); Horváth and Shao (1996), we present and prove the following version, which is more directly related to the current context, and is useful for proofs in the bootstrap sampling space.

Lemma H.5. *Suppose the following conditions hold:*
(i) $\{Z_{ijm}\}$ *is a sequence of random variables in the original sampling space, satisfying*

$$\max_{m\leq p, i\leq n}\frac{1}{n}\sum_{j=1}^n Z_{ijm}^2 = O_P(a_n^2)$$

for some deterministic sequence $a_n > 0$.
(ii) $\{X_i^*, Y_i^*\}_{i\leq n}$ *is an i.i.d. sequence in the bootstrap sampling space such that* $\{X_i^*\}$ *is*

independent of $\{Y_i^*\}$, $EX_i^* = EY_i^* = 0$, *and* $\mathrm{Var}^*(X_i) < C$ *and* $\mathrm{Var}^*(Y_i) < C$ *for a constant* $C > 0$ *where* C *is non-random in both the original and bootstrap sampling space. (iii) Both* X_i^* *and* Y_i^* *are sub-exponential random variables satisfying Assumption 3.1 (iv).*

Then for any $\varepsilon_1, \varepsilon_2 > 0$, *there is a* $C_{\varepsilon_1,\varepsilon_2} > 0$ *such that*

$$P\left(P^*\left(\max_{m\le p}\left|\frac{1}{n^2}\sum_{i,j\le n} X_i^* Y_j^* Z_{ijm}\right| > \frac{2a_n C\sqrt{C_{\varepsilon_1,\varepsilon_2}\log p\log(pn)}}{n}\right) > \varepsilon_1\right) < \varepsilon_2.$$

Thus $\max_{m\le p}|\frac{1}{n^2}\sum_{i,j\le n} X_i^* Y_j^* Z_{ijm}| = O_{P^*}\left(\frac{a_n\sqrt{\log p\log(np)}}{n}\right)$.

Proof. By condition (i), for any $\delta > 0$, there is $C_\delta > 0$ such that with probability at least $1-\delta$ the event $A_\delta := \{\max_{m\le p, i\le n}\frac{1}{n}\sum_{j=1}^n Z_{ijm}^2 < a_n^2 C_\delta\}$ holds.

Let $V^* = \max_{mi}|\frac{1}{n}\sum_j Y_j^* Z_{ijm}|$. Define $W_{im}^* = X_i^*\frac{1}{n}\sum_j Y_j^* Z_{ijm}$ and $Y^* = \{Y_i^*\}_{i\le n}$. Since $\{X_i^*\}$ and $\{Y_i^*\}$ are independent, then on the event A_δ,

$$\begin{aligned}
&\max_{m,i}\frac{1}{n}\sum_j \mathrm{Var}^*(Y_j^* Z_{ijm}) = \max_{m,i}\frac{1}{n}\sum_j Z_{ijm}^2\,\mathrm{Var}(Y_j^*) < a_n^2 C C_\delta \\
&\max_m \frac{1}{n}\sum_i \mathrm{Var}^*(W_{im}^*|Y^*) = \max_m\frac{1}{n}\sum_i\left(\frac{1}{n}\sum_j Y_j^* Z_{ijm}\right)^2 \mathrm{Var}^*(X_i^*) \le C V^{*2}.
\end{aligned} \tag{H.15}$$

In the bootstrap sampling space (BSS), $\{Y_j^* Z_{ijm}\}_{j\le n}$ is independent across j and $E^* Y_j^* Z_{ijm} = 0$. By the Bernstein inequality, for $y = (2a_n^2 C C_\delta \log(pn)/n)^{1/2}$,

$$\begin{aligned}
P^*(V^* > y)1\{A_\delta\} &\le pn\max_{m,i} P^*\left(|\frac{1}{n}\sum_j Y_j^* Z_{ijm}| > y\right)1\{A_\delta\} \\
&\le \exp\left(\log(pn) - \frac{ny^2}{\max_{m,i}\frac{1}{n}\sum_j \mathrm{Var}^*(Y_j^* Z_{ijm})}\right)1\{A_\delta\} \\
&\le \exp\left(\log(pn) - \frac{ny^2}{a_n^2 C C_\delta}\right) = (pn)^{-1}.
\end{aligned} \tag{H.16}$$

In the BSS, $\{W_{im}^*\}_{i\leq n}$ is independent across i conditional on Y^*. By (H.15) and the Bernstein inequality, for $x = y\sqrt{\frac{2C\log p}{n}} = \frac{2a_nC\sqrt{C_\delta \log p \log(pn)}}{n}$,

$$
\begin{aligned}
P^*(\max_{m\leq p}|\frac{1}{n}\sum_i W_{im}^*| > x|Y^*)1\{V^* < y\}
&\leq p\max_m P^*\left(|\frac{1}{n}\sum_i W_{im}^*| > x|Y^*\right)1\{V^* < y\} \\
&\leq \exp\left(\log p - \frac{nx^2}{\max_m \frac{1}{n}\sum_i \mathrm{Var}^*(W_{im}^*|Y^*)}\right)1\{V^* < y\} \\
&\leq \exp\left(\log p - \frac{nx^2}{CV^{*2}}\right)1\{V^* < y\}\exp\left(\log p - \frac{nx^2}{Cy^2}\right) = p^{-1}.
\end{aligned}
\tag{H.17}
$$

Let E_{Y^*} denote the expectation operator with respect to the marginal distribution of Y^* in the bootstrap sampling space; i.e., E_{Y^*} is the conditional distribution of Y^* given the original data. By the law of iterated expectations, $E_{Y^*}[P^*(\cdot|Y^*)] = P^*(\cdot)$. We then have

$$
\begin{aligned}
&P^*\left(\max_{m\leq p}|\frac{1}{n^2}\sum_{i,j\leq n} X_i^*Y_j^*Z_{ijm}| > x\right) \\
&\leq P^*\left(\max_{m\leq p}|\frac{1}{n^2}\sum_{i,j\leq n} X_i^*Y_j^*Z_{ijm}| > x\right)1\{A_\delta\} + 1\{A_\delta^c\} \\
&= P^*\left(\max_{m\leq p}|\frac{1}{n}\sum_i W_{im}^*| > x\right)1\{A_\delta\} + 1\{A_\delta^c\} \\
&= E_{Y^*}P^*\left(\max_{m\leq p}|\frac{1}{n}\sum_i W_{im}^*| > x|Y^*\right)1\{A_\delta\} + 1\{A_\delta^c\} \\
&\leq^{(a)} E_{Y^*}P^*\left(\max_{m\leq p}|\frac{1}{n}\sum_i W_{im}^*| > x|Y^*\right)1\{V^* < y\}1\{A_\delta\} + E_{Y^*}1\{V^* \geq y\}1\{A_\delta\} + 1\{A_\delta^c\} \\
&\leq^{(b)} p^{-1} + P^*(V^* \geq y)1\{A_\delta\} + 1\{A_\delta^c\} \\
&\leq^{(c)} p^{-1} + (pn)^{-1} + 1\{A_\delta^c\},
\end{aligned}
$$

where we used $P^*(\cdot|Y^*) \leq P^*(\cdot|Y^*)1\{V^* < y\} + 1\{V^* \geq y\}$ in (a), (H.17) in (b), and (H.16) in (c). Because $P(A_\delta^c) \leq \delta$, taking the expectation with respect to the distribution of the original data on both sides yields

$$EP^*(\max_{m\leq p}|\frac{1}{n^2}\sum_{i,j\leq n}X_i^*Y_j^*Z_{ijm}|>x)\leq p^{-1}+(pn)^{-1}+\delta.$$

For any $\varepsilon_1,\varepsilon_2>0$, let $\delta=\varepsilon_1\varepsilon_2/2$, and call C_δ in x to be $C_{\varepsilon_1,\varepsilon_2}$. By the Markov Inequality,

$$P\left(P^*(\max_{m\leq p}|\frac{1}{n^2}\sum_{i,j\leq n}X_i^*Y_j^*Z_{ijm}|>x)>\varepsilon_1\right)\leq\frac{1}{\varepsilon_1}(p^{-1}+(pn)^{-1}+\delta)\leq\frac{\varepsilon_1\varepsilon_2/2+\delta}{\varepsilon_1}=\varepsilon_2.$$

■

Lemma H.6. *(i)* $\frac{1}{T}\sum_t\|\frac{1}{n}\widehat{F}^{*\prime}\tilde{U}_t^{*\prime}\widehat{\gamma}\|_2^2=O_{P^*}\left(\Delta_F^{*2}|J|_0^2+\frac{|J|_0}{n}\right)$.
(ii) $\frac{1}{T}\sum_t\|\frac{1}{n}\sum_{i=1}^n\tilde{e}_{it}^*\widehat{f}_i^*\|_2^2=O_{P^*}\left(\frac{1}{n}+\Delta_F^{*2}\right)$.
(iii) $\|\frac{1}{nT}\sum_t\widehat{U}'_{t,\widehat{J}}(\tilde{F}H-\widehat{F}^*)H^{*-1}\widehat{\delta}_{dt}\|_2=O_{P^*}(\Delta_{ud}^*+\Delta_F^{*2}+\Delta_F^*b_n^*)\sqrt{|J_0|}$, *where*
$b_n^*=\sqrt{\frac{\log p}{n}\left(\frac{\log(n)\log(pT)}{n}+\Delta_{\max}^2+\frac{\log(pn)}{T}\right)}$.

Proof. For notational simplicity, we assume f_i is a scalar.

(i) We have $\frac{1}{T}\sum_t\|\frac{1}{n}\widehat{F}^{*\prime}\tilde{U}_t^{*\prime}\widehat{\gamma}\|_2^2\leq O_{P^*}(|J|_0^2\Delta_F^{*2})+\frac{1}{T}\sum_t|\frac{1}{n}\sum_i\widehat{f}_iw_i^U\widehat{U}'_{it}\widehat{\gamma}|^2O_{P^*}(1)$. Hence, we need to bound $\frac{1}{T}\sum_t|\frac{1}{n}\sum_i\widehat{f}_iw_i^U\widehat{U}'_{it}\widehat{\gamma}|^2$.

$$\begin{aligned}
&E^*\left[\frac{1}{T}\sum_t|\frac{1}{n}\sum_i\widehat{f}_iw_i^U\widehat{U}'_{it}\widehat{\gamma}|^2\right]\\
&=\frac{1}{T}\sum_t\frac{1}{n^2}\sum_i(\widehat{f}_i\widehat{U}'_{it}\widehat{\gamma})^2\\
&\leq\frac{2}{n^2T}\sum_t\sum_i(\widehat{f}_i\tilde{U}'_{it}\widehat{\gamma})^2+\frac{1}{n^2T}\sum_t\sum_i\widehat{f}_i^4\left(\sum_{m=1}^p\tilde{\lambda}_{tm}\widehat{\gamma}_m\right)^2O_P(\Delta_F^2)\\
&\quad+\frac{1}{n^2T}\sum_t\sum_i\widehat{f}_i^4\left(\frac{1}{n}\widehat{F}'\tilde{U}'_t\gamma\right)^2+\frac{1}{nT}\sum_t\left(\sum_{m=1}^p\tilde{\lambda}'_{tm}\widehat{\gamma}_m\right)^2\max_i\widehat{f}_i^2O_P(\Delta_F^2)\\
&\leq\frac{4}{n^2T}\sum_t\sum_i(\widehat{f}_i\tilde{U}'_{it}(\widehat{\gamma}-\gamma))^2+\frac{4}{n^2T}\sum_t\sum_i(\widehat{f}_i\tilde{U}'_{it}\gamma)^2\\
&\quad+\max_i|\widehat{f}_i|^2\frac{1}{n}O_P(\Delta_F^2|J|_0^2)+\max_i|\widehat{f}_i|^2\frac{1}{nT}\sum_t\left(\frac{1}{n}\widehat{F}'\tilde{U}'_t\gamma\right)^2\\
&\leq\|\widehat{\gamma}-\gamma\|_1^2\max_i\widehat{f}_i^2\max_{mk}\frac{1}{n^2T}\sum_{it}|\tilde{U}_{it,m}\tilde{U}_{it,k}|+\frac{1}{n^2T}\sum_t\sum_i(\tilde{f}_i\tilde{U}'_{it}\gamma)^2
\end{aligned}$$

$$
\begin{aligned}
&+\frac{8}{n^2T}\sum_t\sum_i(\widehat{f}_i-H'\tilde{f}_i)^2(\tilde{U}_{it}'\gamma)^2+\max_i|\widehat{f}_i|^2\frac{1}{n}O_P(\Delta_F^2|J|_0^2)\\
&+\max_i|\widehat{f}_i|^2\frac{1}{nT}\sum_t\left(\frac{1}{n}\widehat{F}'\tilde{U}_t'\gamma\right)^2\\
\leq{}& O_P\left(\frac{\log n}{n}\kappa_n^2|J|_0^2+\frac{|J|_0}{n}+\frac{\log n}{n}\Delta_F^2|J|_0^2+\frac{\log n|J|_0}{n^2}+\frac{\log n|J|_0^2}{n}\Delta_F^2\right.\\
&\left.+\frac{|J|_0^2\Delta_F^2}{n}\left(\sqrt{\frac{\log n}{T}}+1\right)\right)\\
={}& O_P\left(\frac{|J|_0}{n}\right),
\end{aligned}
$$

where we used the assumption $\max_k\sum_m|EU_{it,k}U_{it,m}|<C$ to bound $\frac{1}{n^2T}\sum_t\sum_i(\tilde{f}_i\tilde{U}_{it}'\gamma)^2$, yielding its rate $O_P\left(\frac{|J|_0}{n}\right)$ instead of $O_P\left(\frac{|J|_0^2}{n}\right)$. Also,

$$
\frac{1}{T}\sum_t\left\|\frac{1}{n}\widehat{F}'\tilde{U}_t\gamma\right\|_2^2=O_P\left(\frac{|J|_0}{n}+\Delta_F^2|J|_0^2\right)
$$

follows from Lemma H.11. Finally,

$$
\frac{8}{n^2T}\sum_t\sum_i(\widehat{f}_i-H'\tilde{f}_i)^2(\tilde{U}_{it}'\gamma)^2=O_P\left(\frac{|J|_0^2\Delta_F^2}{n}\left(\sqrt{\frac{\log n}{T}}+1\right)\right)
$$

is due to

$$
\begin{aligned}
\frac{1}{n^2T}\sum_i(\widehat{f}_i-H'\tilde{f}_i)^2\sum_t(\tilde{U}_{it}'\gamma)^2 &\leq O_P\left(\frac{1}{n}\Delta_F^2\right)\max_i\frac{1}{T}\sum_t(\tilde{U}_{it}'\gamma)^2\\
&\leq O_P\left(\frac{1}{n}\Delta_F^2\right)\max_i\frac{1}{T}\sum_t(U_{it}'\gamma)^2\\
&\leq O_P\left(\frac{1}{n}\Delta_F^2\right)\max_i\left|\frac{1}{T}\sum_t(E(U_{it}'\gamma)^2-(U_{it}'\gamma)^2)\right|\\
&\quad+O_P\left(\frac{1}{n}\Delta_F^2\right)\max_i\frac{1}{T}\sum_t E(U_{it}'\gamma)^2\\
&\leq O_P\left(\frac{|J|_0^2\Delta_F^2}{n}\left(\sqrt{\frac{\log n}{T}}+1\right)\right).
\end{aligned}
$$

(ii) By Lemma H.14, $\frac{1}{T}\sum_t\|\frac{1}{n}\sum_{i=1}^n\tilde{e}_{it}^*\widehat{f}_i\|_2^2=O_{P^*}(\frac{1}{n})$, which yields the result.

(iii) By Assumption D.5 and (H.41),

$$\begin{aligned}
&\left\| \frac{1}{nT} \sum_t \widehat{U}^{*\prime}_{t,\widehat{J}^*}(\tilde{F}H - \widehat{F}^*)H^{*-1}\widehat{\delta}_{dt} \right\|_2 \\
&\quad \leq O_{P^*}(\sqrt{|J|_0}) \max_m \left| \frac{1}{n}\sum_{i=1}^n \frac{1}{T}\sum_{t=1}^T \tilde{U}^*_{it,m}\widehat{\delta}_{yt}' H^{*'-1}(\widehat{f}_i^* - H^{*'}\widehat{f}_i) \right| \\
&\qquad + O_{P^*}(\sqrt{|J|_0}) \max_m \left| \frac{1}{n}\sum_{i=1}^n \frac{1}{T}\sum_{t=1}^T (\widehat{U}^*_{it,m} - \tilde{U}^*_{it,m})\widehat{\delta}_{yt}' H^{*'-1}(\widehat{f}_i^* - H^{*'}\widehat{f}_i) \right| \\
&\quad \leq O_{P^*}(\Delta^*_{ud} + \Delta_F^{*2} + \Delta_F^* b_n^*)\sqrt{|J_0|}
\end{aligned}$$

where $b_n^* = \sqrt{\frac{\log p}{n}\left(\frac{\log(n)\log(pT)}{n} + \Delta^2_{\max} + \frac{\log(pn)}{T}\right)}$. ■

In the following lemma, recall that $\tilde{\gamma}^*_y$ denotes the k-step lasso estimator, and $\widehat{\gamma}_y$ denotes the lasso estimator using the original data.

Lemma H.7. *For* $\widehat{\gamma} \in \{\widehat{\gamma}_y, \widehat{\gamma}_d, \theta\}$,

(i) $\left\| \frac{1}{\sqrt{nT}} M_{\widehat{U}^*_{\widehat{J}^*}} \tilde{U}^*\widehat{\gamma} \right\|_2^2 = O_{P^*}\left(a_n + \kappa_n^2|J|_0 + |J|_0^2\Delta_F^{*2} + \frac{|J|_0}{n}\right)$.

(ii) $\left\| \frac{1}{\sqrt{nT}} (I_T \otimes P_{\widehat{F}^*})\tilde{U}^*\widehat{\gamma} \right\|_2^2 = O_{P^*}\left(\frac{|J|_0}{n} + |J|_0^2\Delta_F^{*2}\right)$.

(iii) $\left\| \frac{1}{\sqrt{nT}} (I_T \otimes P_{\widehat{F}^*})\tilde{\eta}^* \right\|_2^2 = O_{P^*}\left(\Delta_F^{*2} + \frac{1}{n}\right)$ *and* $\left\| \frac{1}{\sqrt{nT}} (I_T \otimes P_{\widehat{F}^*})\tilde{\epsilon}^* \right\|_2^2 = O_{P^*}\left(\Delta_F^{*2} + \frac{1}{n}\right)$.

(iv) $\left\| \frac{1}{\sqrt{nT}} P_{\widehat{U}^*_{\widehat{J}^*}} \tilde{\eta}^* \right\|_2^2 = O_{P^*}\left(|J|_0 \frac{\log p}{nT}\right)$ *and* $\left\| \frac{1}{\sqrt{nT}} P_{\widehat{U}^*_{\widehat{J}^*}} \tilde{\epsilon}^* \right\|_2^2 = O_{P^*}\left(|J|_0 \frac{\log p}{nT}\right)$.

(v) $\left\| \frac{1}{\sqrt{nT}} M_{\widehat{U}^*_{\widehat{J}^*}} (I_T \otimes M_{\widehat{F}^*})\tilde{U}^*\widehat{\gamma}_d \right\| = O_{P^*}\left(a_n + \kappa_n^2|J|_0 + |J|_0^2\Delta_F^{*2} + \frac{|J|_0}{n}\right)$.

Proof. (i) First, consider the following constrained problem:

$$\widehat{m} = \arg\min_m \|\widehat{U}^*(\widehat{\gamma} - m)\|_2^2, \quad m_j = 0, j \notin \widehat{J}^*,$$

where $\widehat{J}^*$ is the support of the k-step lasso (instead of the support of the complete bootstrap lasso estimator $\tilde{\gamma}^*_{y,lasso}$). The solution satisfies $\widehat{U}^*\widehat{m} = P_{\widehat{U}^*_{\widehat{J}^*}}\widehat{U}^*\widehat{\gamma}$. For $\tilde{\gamma}^*$ being the k-step lasso (either $\tilde{\gamma}^*_y$ or $\tilde{\gamma}^*_d$), $\frac{1}{nT}\sum_{t=1}^T \|\widehat{U}^*_t(\widehat{\gamma} - \tilde{\gamma}^*)\|_2^2 = O_{P^*}(\kappa_n^2|J|_0 + a_n)$ by Proposition B.2. Hence

$$\left\| \frac{1}{\sqrt{nT}} M_{\widehat{U}^*_{\widehat{J}^*}} \widehat{U}^*\widehat{\gamma} \right\|_2^2 = \left\| \frac{1}{\sqrt{nT}} \widehat{U}^*(\widehat{\gamma} - \widehat{m}) \right\|_2^2 \leq \left\| \frac{1}{\sqrt{nT}} \widehat{U}^*(\widehat{\gamma} - \tilde{\gamma}^*) \right\|_2^2 = O_{P^*}(a_n + \kappa_n^2|J|_0).$$

Also, using $\frac{1}{T}\sum_t \|\frac{1}{n}\widehat{F}^{*\prime}\tilde{U}_t^{*\prime}\widehat{\gamma}\|_2^2 = O_{P^*}\left(|J|_0^2\Delta_F^{*2} + \frac{|J|_0}{n}\right)$ from Lemma H.6 and

$$\tilde{U}_{it,m}^* - \widehat{U}_{it,m}^* = \widehat{f}_i^{*\prime}(\widehat{F}^{*\prime}\widehat{F}^*)^{-1}\widehat{F}^{*\prime}(\widehat{F}H^* - \widehat{F}^*)H^{*-1}\widehat{\lambda}_{tm} + \widehat{f}_i^{*\prime}(\widehat{F}^{*\prime}\widehat{F}^*)^{-1}\widehat{F}^{*\prime}\tilde{U}_{t,m}^* + \widehat{\lambda}_{tm}^{\prime}H^{*\prime-1}(\widehat{f}_i^* - H^{*\prime}\widehat{f}_i),$$

we have

$$\begin{aligned}\left\|\frac{1}{\sqrt{nT}}M_{\widehat{U}^*_{\widehat{J}^*}}(\widehat{U}^* - \tilde{U}^*)\widehat{\gamma}\right\|_2^2 &\leq \frac{1}{nT}\left\|(\widehat{U}^* - \tilde{U}^*)\widehat{\gamma}\right\|_2^2 \\ &= \frac{1}{nT}\sum_{t=1}^T\sum_i[\sum_m(\widehat{U}_{it,m}^* - \tilde{U}_{it,m}^*)\widehat{\gamma}_m]^2 \\ &\leq \frac{1}{T}\sum_{t=1}^T\left[\sum_m\frac{1}{n}\widehat{F}^{*\prime}\tilde{U}_{t,m}^*\widehat{\gamma}_m\right]^2 O_{P^*}(1) + \frac{1}{T}\sum_{t=1}^T[\sum_m\widehat{\lambda}_{tm}^{\prime}\widehat{\gamma}_m]^2 O_{P^*}(\Delta_F^{*2}) \\ &= O_{P^*}\left(|J|_0^2\Delta_F^{*2} + \frac{|J|_0}{n}\right).\end{aligned}$$

Combining these yields $\left\|\frac{1}{\sqrt{nT}}M_{\widehat{U}^*_{\widehat{J}^*}}\tilde{U}^*\widehat{\gamma}\right\|_2^2 = O_{P^*}\left(a_n + \kappa_n^2|J|_0 + |J|_0^2\Delta_F^{*2} + \frac{|J|_0}{n}\right)$.

(ii) Still by by Lemma H.6,

$$\begin{aligned}\left\|\frac{1}{\sqrt{nT}}(I_T\otimes P_{\widehat{F}^*})\tilde{U}^*\widehat{\gamma}\right\|_2^2 &= \frac{1}{nT}\sum_t\left\|P_{\widehat{F}^*}\tilde{U}_t^*\widehat{\gamma}\right\|_2^2 \\ &\leq \frac{1}{n^2T}\sum_t\left\|\widehat{F}^{*\prime}\tilde{U}_t^{*\prime}\widehat{\gamma}\right\|_2^2 \\ &= O_{P^*}\left(\frac{|J|_0}{n} + |J|_0^2\Delta_F^{*2}\right).\end{aligned}$$

(iii) By Lemma H.6, $\frac{1}{T}\sum_t\|\frac{1}{n}\sum_{i=1}^n\tilde{e}_{it}^*\widehat{f}_i\|_2^2 = O_{P^*}(\frac{1}{n})$. Hence,

$$\|\frac{1}{\sqrt{nT}}(I_T\otimes P_{\widehat{F}^*})\tilde{\eta}^*\|_2^2 = \frac{1}{nT}\sum_t\|P_{\widehat{F}^*}\tilde{\eta}_t^*\|_2^2 \leq \frac{1}{n^2T}\sum_t\|\widehat{F}^{*\prime}\tilde{\eta}_t^*\|_2^2 = O_{P^*}\left(\Delta_F^{*2} + \frac{1}{n}\right).$$

(iv) The same argument as used in the proof of Lemma H.17 yields

$$\max_m\left|\frac{1}{nT}\sum_t\sum_i\tilde{U}_{it,m}^*\tilde{\eta}_{it}^*\right| = O_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right) \text{ and}$$

$$\max_m\left|\frac{1}{nT}\sum_t\sum_i(\tilde{U}_{it,m}^* - \widehat{U}_{it,m}^*)\tilde{\eta}_{it}^*\right| = o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right).$$

Also, by Assumption 4.1, $|\widehat{J}^*| = O_{P^*}(|J|_0)$. Hence,

$$\left\|\frac{1}{nT}\tilde{U}^{*\prime}_{\widehat{J}^*}\tilde{\eta}^*\right\|_2 = \left\|\frac{1}{nT}\sum_t \tilde{U}^{*\prime}_{t,\widehat{J}^*}\tilde{\eta}^*_t\right\|_2 \leq \max_m \left|\frac{1}{nT}\sum_t\sum_i \tilde{U}^*_{it,m}\tilde{\eta}^*_{it}\right|\sqrt{|\widehat{J}^*|_0} = O_{P^*}\left(\sqrt{|J|_0\frac{\log p}{nT}}\right)$$

and $\|\frac{1}{nT}(\tilde{U}^*_{\widehat{J}^*} - \widehat{U}^*_{\widehat{J}^*})'\tilde{\eta}^*\|_2 \leq \|\frac{1}{nT}\sum_t(\tilde{U}^{*\prime}_{t,\widehat{J}^*} - \widehat{U}^{*\prime}_{t,\widehat{J}^*})\tilde{\eta}^*_t\|_\infty\sqrt{|\widehat{J}|_0} = o_{P^*}\left(\sqrt{|J|_0\frac{\log p}{nT}}\right)$. We then have

$$\left\|\frac{1}{nT}\widehat{U}^{*\prime}_{\widehat{J}^*}\tilde{\eta}^*\right\|_2 = O_{P^*}\left(\sqrt{|J|_0\frac{\log p}{nT}}\right). \tag{H.18}$$

Next, $\phi_{\min}(|\widehat{J}^*|) \leq \lambda_{\min}\left(\frac{1}{nT}\tilde{U}'_{\widehat{J}^*}\tilde{U}_{\widehat{J}^*}\right) \leq \lambda_{\max}\left(\frac{1}{nT}\tilde{U}'_{\widehat{J}^*}\tilde{U}_{\widehat{J}^*}\right) \leq \phi_{\max}(|\widehat{J}^*|)$. Also by (H.49),

$$\left\|\frac{1}{nT}\widehat{U}^{*\prime}_{\widehat{J}^*}\widehat{U}^*_{\widehat{J}^*} - \frac{1}{nT}\tilde{U}'_{\widehat{J}^*}\tilde{U}_{\widehat{J}^*}\right\| \leq O_{P^*}(|J|_0)\max_{m,k\leq p}\left|\frac{1}{nT}\sum_{it}(\widehat{U}^*_{it,m}\widehat{U}^*_{it,k} - \tilde{U}_{it,m}\tilde{U}_{it,k})\right| = o_{P^*}(1).$$

This result implies

$$\phi_{\min}(|\widehat{J}^*|) - o_{P^*}(1) \leq \lambda_{\min}\left(\frac{1}{nT}\widehat{U}^{*\prime}_{\widehat{J}^*}\widehat{U}^*_{\widehat{J}^*}\right) \leq \lambda_{\max}\left(\frac{1}{nT}\widehat{U}^{*\prime}_{\widehat{J}^*}\widehat{U}^*_{\widehat{J}^*}\right) \leq \phi_{\max}(|\widehat{J}^*|) + o_{P^*}(1). \tag{H.19}$$

Hence $\left\|\frac{1}{\sqrt{nT}}\widehat{U}^*_{\widehat{J}^*}\right\|^2 = \lambda_{\max}\left(\frac{1}{nT}\widehat{U}^{*\prime}_{\widehat{J}^*}\widehat{U}^*_{\widehat{J}^*}\right) = O_{P^*}(1)$ and $\left\|(\frac{1}{nT}\widehat{U}^{*\prime}_{\widehat{J}^*}\widehat{U}^*_{\widehat{J}^*})^{-1}\right\| = O_{P^*}(1)$. Finally,

$$\left\|\frac{1}{\sqrt{nT}}P_{\widehat{U}^*_{\widehat{J}^*}}\tilde{\eta}^*\right\|_2 \leq \left\|\frac{1}{\sqrt{nT}}\widehat{U}^*_{\widehat{J}^*}(\frac{1}{nT}\widehat{U}^{*\prime}_{\widehat{J}^*}\widehat{U}^*_{\widehat{J}^*})^{-1}\frac{1}{nT}\widehat{U}^{*\prime}_{\widehat{J}^*}\tilde{\eta}^*\right\|_2 = O_{P^*}\left(\sqrt{|J|_0\frac{\log p}{nT}}\right).$$

(v) The result follows immediately from parts (i) and (ii) and the following inequality:

$$\left\|\frac{1}{\sqrt{nT}}M_{\widehat{U}^*_{\widehat{J}^*}}(I_T\otimes M_{\widehat{F}^*})\tilde{U}^*\widehat{\gamma}_d\right\|_2 \leq \left\|\frac{1}{\sqrt{nT}}M_{\widehat{U}^*_{\widehat{J}^*}}\tilde{U}^*\widehat{\gamma}_d\right\|_2 + \left\|\frac{1}{\sqrt{nT}}(I_T\otimes P_{\widehat{F}^*})\tilde{U}^*\widehat{\gamma}_d\right\|_2.$$

■

Lemma H.8. *(i)* $\frac{1}{\sqrt{nT}}\tilde{\epsilon}^{*\prime}M_{\widehat{U}^*_{\widehat{J}^*}}(I_T\otimes M_{\widehat{F}^*}(\widehat{F}H^* - \widehat{F}^*)H^{*-1})\widehat{\Delta}_d = o_{P^*}(1)$.
(ii) $\frac{1}{\sqrt{nT}}\tilde{\epsilon}^{*\prime}(I_T\otimes P_{\widehat{F}^*})\tilde{U}^*\widehat{\gamma}_d = o_{P^*}(1)$,
(iii) $\frac{1}{\sqrt{nT}}\tilde{\epsilon}^{*\prime}(I_T\otimes P_{\widehat{F}^*})\tilde{\eta}^* = o_{P^*}(1)$,
(iv) $\frac{1}{\sqrt{nT}}\tilde{\epsilon}^{*\prime}M_{\widehat{U}^*_{\widehat{J}^*}}\tilde{U}^*\widehat{\gamma}_d = o_{P^*}(1)$.
(v) $\frac{1}{\sqrt{nT}}\tilde{\epsilon}^{*\prime}M_{\widehat{U}^*_{\widehat{J}^*}}(I_T\otimes M_{\widehat{F}^*})\tilde{U}^*\widehat{\gamma}_d = o_{P^*}(1)$.
(vi) $\widehat{\Xi}'(I_T\otimes H^{*\prime-1}(\widehat{F}H^* - \widehat{F}^*)'M_{\widehat{F}^*})\frac{1}{\sqrt{nT}}M_{\widehat{U}^*_{\widehat{J}^*}}\tilde{U}^*\widehat{\gamma}_d = o_{P^*}(1)$.

Proof. (i) First, by Lemma H.14, $\frac{1}{T}\sum_t \|\frac{1}{n}\sum_{i=1}^n \tilde{\epsilon}_{it}^*\widehat{f}_i\|_2^2 = O_{P^*}\left(\frac{1}{n}\right)$. Thus,

$$\begin{aligned}
&\left|\frac{1}{\sqrt{nT}}\ \sum_t \tilde{\epsilon}_t^{*\prime} P_{\widehat{F}^*}(\widehat{F}H^* - \widehat{F}^*)H^{*-1}\widehat{\delta}_{dt}\right|^2 \\
&\qquad = \frac{1}{nT}\operatorname{tr}^2\left(\sum_t \widehat{\delta}_{dt}\tilde{\epsilon}_t^{*\prime} P_{\widehat{F}^*}(\widehat{F}H^* - \widehat{F}^*)H^{*-1}\right) \\
&\qquad \leq \frac{1}{nT}\left\|\sum_t \widehat{\delta}_{dt}\tilde{\epsilon}_t^{*\prime}\widehat{F}^*\right\|_F^2 \left\|\frac{1}{n}\widehat{F}^{*\prime}(\widehat{F}H^* - \widehat{F}^*)\right\|_F^2 O_{P^*}(1) \\
&\qquad = O_{P^*}(\Delta_F^{*2})\left\|\frac{1}{\sqrt{nT}}\sum_t \widehat{\delta}_{dt}\tilde{\epsilon}_t^{*\prime}(\widehat{F}^* - \widehat{F}H^*)\right\|_F^2 + O_{P^*}(\Delta_F^{*2})\frac{1}{nT}\left\|\sum_t \widehat{\delta}_{dt}\tilde{\epsilon}_t^{*\prime}\widehat{F}\right\|_F^2 \\
&\qquad = O_{P^*}(\Delta_F^{*4}nT) + O_{P^*}(\Delta_F^{*2})nT\frac{1}{T}\sum_t \|\frac{1}{n}\tilde{\epsilon}_t^{*\prime}\widehat{F}\|_F^2 \\
&\qquad = o_{P^*}(1).
\end{aligned}$$

By Lemma H.6,

$$\left\|\frac{1}{nT}\sum_t \widehat{U}_{t,\widehat{J}}'(\tilde{F}H - \widehat{F}^*)H^{*-1}\widehat{\delta}_{dt}\right\|_2 = O_{P^*}\left(\Delta_{ud}^* + \Delta_F^{*2} + \Delta_F^* b_n^*\right)\sqrt{|J_0|},$$

and

$$\left\|\frac{1}{\sqrt{nT}}\sum_t \widehat{U}_{t,\widehat{J}^*}^{*\prime} P_{\widehat{F}^*}(\widehat{F}H^* - \widehat{F}^*)H^{*-1}\widehat{\delta}_{dt}\right\|_2 = 0$$

due to $\widehat{U}_t'\widehat{F}^* = 0$. Therefore,

$$\begin{aligned}
&\left\|P_{\widehat{U}^*_{\widehat{J}^*}}\ (I_T \otimes M_{\widehat{F}^*}(\widehat{F}H^* - \widehat{F}^*)H^{*-1})\widehat{\Delta}_d\right\|_2 \\
&\qquad = \left\|\widehat{U}^*_{\widehat{J}^*}(\widehat{U}^{*\prime}_{\widehat{J}^*}\widehat{U}^*_{\widehat{J}^*})^{-1}\sum_t \widehat{U}^{*\prime}_{t,\widehat{J}^*} M_{\widehat{F}^*}(\widehat{F}H^* - \widehat{F}^*)H^{*-1}\widehat{\delta}_{dt}\right\|_2 \\
&\qquad \leq \left\|\frac{1}{\sqrt{nT}}\sum_t \widehat{U}^{*\prime}_{t,\widehat{J}^*} M_{\widehat{F}^*}(\widehat{F}H^* - \widehat{F}^*)H^{*-1}\widehat{\delta}_{dt}\right\|_2 \\
&\qquad \leq \left\|\frac{1}{\sqrt{nT}}\sum_t \widehat{U}^{*\prime}_{t,\widehat{J}^*}(\widehat{F}H^* - \widehat{F}^*)H^{*-1}\widehat{\delta}_{dt}\right\|_2 \\
&\qquad = O_{P^*}\left(\Delta_{ud}^* + \Delta_F^{*2} + \Delta_F^* b_n^*\right)\sqrt{|J|_0 nT}.
\end{aligned} \tag{H.20}$$

By Assumption D.5, $\|\frac{1}{\sqrt{nT}}\sum_t \widehat{\delta}_{dt}\tilde{\epsilon}_t^{*\prime}(\widehat{F}H^*-\widehat{F}^*)\|_F = o_{P^*}(1)$, and $\|\frac{1}{\sqrt{nT}}P_{\widehat{U}^*_{\widehat{J}^*}}\tilde{\epsilon}\|_2^2 = O_{P^*}(|J|_0\frac{\log p}{nT})$ by Lemma H.7. Hence,

$$\begin{aligned}
&\frac{1}{\sqrt{nT}}\tilde{\epsilon}^{*\prime}M_{\widehat{U}^*_{\widehat{J}^*}}(I_T\otimes M_{\widehat{F}^*}(\widehat{F}H^*-\widehat{F}^*)H^{*-1})\widehat{\Delta}_d \\
&\quad= -\frac{1}{\sqrt{nT}}\tilde{\epsilon}^{*\prime}P_{\widehat{U}^*_{\widehat{J}^*}}(I_T\otimes M_{\widehat{F}^*}(\widehat{F}H^*-\widehat{F}^*)H^{*-1})\widehat{\Delta}_d \\
&\qquad+\frac{1}{\sqrt{nT}}\tilde{\epsilon}^{*\prime}(I_T\otimes(\widehat{F}H^*-\widehat{F}^*)H^{*-1})\widehat{\Delta}_d \\
&\qquad-\frac{1}{\sqrt{nT}}\tilde{\epsilon}^{*\prime}(I_T\otimes P_{\widehat{F}^*}(\widehat{F}H^*-\widehat{F}^*)H^{*-1})\widehat{\Delta}_d \\
&\quad\le \left\|\frac{1}{\sqrt{nT}}\tilde{\epsilon}^{*\prime}P_{\widehat{U}^*_{\widehat{J}^*}}\right\|_2\left\|P_{\widehat{U}^*_{\widehat{J}^*}}(I_T\otimes M_{\widehat{F}^*}(\widehat{F}H^*-\widehat{F}^*)H^{*-1})\widehat{\Delta}_d\right\|_2 \\
&\qquad+\left|\frac{1}{\sqrt{nT}}\sum_t\tilde{\epsilon}_t^{*\prime}P_{\widehat{F}^*}(\widehat{F}H^*-\widehat{F}^*)H^{-1}\tilde{\delta}_{dt}\right| \\
&\qquad+\left|\frac{1}{\sqrt{nT}}\sum_t\tilde{\epsilon}_t^{*\prime}(\widehat{F}H^*-\widehat{F}^*)H^{-1}\widehat{\delta}_{dt}\right| \\
&\quad= o_{P^*}(1)+O_{P^*}\left(\sqrt{|J|_0\frac{\log p}{nT}}\right)\left(\Delta^*_{ud}+\Delta^{*2}_F+\Delta^*_F b^*_n\right)\sqrt{|J|_0 nT} \\
&\quad= o_{P^*}(1)
\end{aligned}$$

where we have used the assumption that $(\Delta^*_{ud}+\Delta^{*2}_F+\Delta^*_F b^*_n)\sqrt{\log p|J|_0^2} = o(1)$.

(ii) By the Cauchy-Schwarz inequality and Lemma H.6,

$$\begin{aligned}
\frac{1}{\sqrt{nT}}\tilde{\epsilon}^{*\prime}(I_T\otimes P_{\widehat{F}^*})\tilde{U}^*\widehat{\gamma}_d &= \frac{1}{\sqrt{nT}}\sum_t\tilde{\epsilon}_t^{*\prime}P_{\widehat{F}^*}\tilde{U}_t^*\widehat{\gamma}_d \\
&\le \left\|\frac{\sqrt{nT}}{n^2}\sum_{m=1}^p\frac{1}{T}\sum_t\widehat{F}^{*\prime}\tilde{U}^*_{t,m}\tilde{\epsilon}_t^{*\prime}\widehat{F}^*\widehat{\gamma}_{dm}\right\|_F \\
&\le \sqrt{nT}\left(\frac{1}{T}\sum_t(\frac{1}{n}\widehat{F}^{*\prime}\tilde{U}_t^{*\prime}\widehat{\gamma}_d)^2\right)^{1/2}\left(\frac{1}{T}\sum_t\left(\frac{1}{n}\tilde{\epsilon}_t^{*\prime}\widehat{F}^*\right)^2\right)^{1/2} \\
&= O_{P^*}\left(\sqrt{nT}\left(\Delta^*_F|J|_0+\sqrt{\frac{|J|_0}{n}}\right)\right)\left(\frac{1}{\sqrt{n}}+\Delta^*_F\right) \\
&= o_{P^*}(1).
\end{aligned}$$

(iii) It can be shown that

$$\frac{1}{\sqrt{nT}}\tilde{\epsilon}^{*\prime}(I_T\otimes P_{\widehat{F}^*})\tilde{\eta}^* \leq \left\|\frac{\sqrt{T}}{n\sqrt{n}}\widehat{F}^{*\prime}\frac{1}{T}\sum_t \tilde{\eta}_t^*\tilde{\epsilon}_t^{*\prime}\widehat{F}^*\right\|_F \leq O_{P^*}\left(\sqrt{nT}\left(\frac{1}{\sqrt{n}}+\Delta_F^*\right)^2\right) = o_{P^*}(1).$$

(iv) Recall that there is $\widehat{m}$ such that $\widehat{U}^*\widehat{m} = P_{\widehat{U}^*_{\widehat{J}^*}}\widehat{U}^*\widehat{\gamma}_d$. We first bound $\|\widehat{\gamma}_d - \widehat{m}\|_1$. By Proposition B.2,

$$\frac{1}{nT}\|\widehat{U}^*(\widehat{\gamma}_d - \widehat{m})\|_2^2 \leq \frac{1}{nT}\sum_{t=1}^T \|\widehat{U}_t^*(\widehat{\gamma}_d - \tilde{\gamma}^*)\|_2^2 = O_{P^*}(a_n + \kappa_n^2|J|_0).$$

Also, $\widehat{\gamma}_d - \widehat{m}$ is at most $|\widehat{J}|_0 + |\widehat{J}^*|_0$-sparse. Hence, by Assumption 3.4 and using arguments similar to those used in the proof of (H.19),

$$\frac{1}{nT}\|\widehat{U}^*(\widehat{\gamma}_d - \widehat{m})\|_2^2 \geq (\phi_{\min}(|\widehat{J}|_0 + |\widehat{J}^*|_0) - o_{P^*}(1))\|\widehat{\gamma}_d - \widehat{m}\|_2^2 \geq C|J|_0^{-1}\|\widehat{\gamma}_d - \widehat{m}\|_1^2.$$

Hence, $\|\widehat{\gamma}_d - \widehat{m}\|_1^2 = O_{P^*}(a_n + \kappa_n^2|J|_0)|J|_0$.

In addition, similar to the proof of Lemma H.17, $\left\|\frac{1}{\sqrt{nT}}\tilde{\epsilon}^{*\prime}\widehat{U}^*\right\|_\infty = O_{P^*}(\sqrt{\log p})$. Hence,

$$\begin{aligned}\left|\frac{1}{\sqrt{nT}}\tilde{\epsilon}^{*\prime}M_{\widehat{U}^*_{\widehat{J}^*}}\widehat{U}^*\widehat{\gamma}_d\right| &= \left|\frac{1}{\sqrt{nT}}\tilde{\epsilon}^{*\prime}\widehat{U}^*(\widehat{\gamma}_d - \widehat{m})\right| \\ &\leq \left\|\frac{1}{\sqrt{nT}}\tilde{\epsilon}^{*\prime}\widehat{U}^*\|_\infty\|\widehat{\gamma}_d - \widehat{m}\right\|_1 \\ &= O_{P^*}\left(a_n^{1/2}|J|_0^{1/2} + \kappa_n|J|_0\right)\sqrt{\log p} = o_{P^*}(1)\end{aligned}$$

under the assumption $a_n|J|_0\log p = o(1)$.

In addition,

$$\begin{aligned}\frac{1}{\sqrt{nT}}\tilde{\epsilon}^{*\prime}M_{\widehat{U}^*_{\widehat{J}^*}}(\tilde{U}^* - \widehat{U}^*)\widehat{\gamma}_d &= \frac{1}{\sqrt{nT}}\tilde{\epsilon}^{*\prime}(\tilde{U}^* - \widehat{U}^*)\widehat{\gamma}_d - \frac{1}{\sqrt{nT}}\tilde{\epsilon}^{*\prime}P_{\widehat{U}^*_{\widehat{J}^*}}(\tilde{U}^* - \widehat{U}^*)\widehat{\gamma}_d \\ &= \frac{1}{\sqrt{nT}}\sum_i\sum_t\sum_m \tilde{\epsilon}_{it}^*(\tilde{U}_{it,m}^* - \widehat{U}_{it,m}^*)\widehat{\gamma}_{dm} \\ &\quad - \frac{1}{\sqrt{nT}}\tilde{\epsilon}^{*\prime}\widehat{U}^*_{\widehat{J}^*}(\widehat{U}^{*\prime}_{\widehat{J}^*}\widehat{U}^*_{\widehat{J}^*})^{-1}\widehat{U}^{*\prime}_{\widehat{J}^*}(\tilde{U}^* - \widehat{U}^*)\widehat{\gamma}_d\end{aligned}$$

$$\leq \left\| \frac{1}{\sqrt{nT}} \sum_i \sum_m \sum_t \tilde{\epsilon}^*_{it} \widehat{\gamma}_{dm} \widehat{\lambda}_{tm} \widehat{f}_i' \right\|_F O_{P^*}(\Delta^*_F)$$
$$+ \left\| \frac{1}{\sqrt{nT}} \sum_i \sum_t \sum_m \tilde{\epsilon}^*_{it} \widehat{f}_i \frac{1}{n} \widehat{F}^{*\prime} \tilde{U}^*_{t,m} \widehat{\gamma}_{dm} \right\|_F O_{P^*}(1)$$
$$+ \left\| \frac{1}{\sqrt{nT}} \sum_i \sum_t \sum_m \tilde{\epsilon}^*_{it} \widehat{\gamma}_{dm} \widehat{\lambda}_{tm} (\widehat{f}^*_i - H^* \widehat{f}_i)' \right\| O_{P^*}(1)$$
$$+ \left\| \frac{1}{\sqrt{nT}} \tilde{\epsilon}^{*\prime} \widehat{U}^*_{\widehat{J}^*} \|_2 \| \frac{1}{nT} \widehat{U}^{*\prime}_{\widehat{J}^*} (\tilde{U}^* - \widehat{U}^*) \widehat{\gamma}_d \right\|_2 O_{P^*}(1). \tag{H.21}$$

The first term on the right of the inequality in (H.21) is $O_{P^*}(|J|_0)\Delta^*_F = o_{P^*}(1)$. By part (ii), the second term on the right of the inequality in (H.21) is $o_{P^*}(1)$, and the third term is $O_{P^*}(\Delta^*_{eg}\sqrt{nT}) = o_{P^*}(1)$ by Assumption D.5. For the last term on the right of the inequality in (H.21), we have $\left\| \frac{1}{\sqrt{nT}} \widehat{U}^{*\prime}_{\widehat{J}^*} \tilde{\epsilon}^* \right\|_2 = O_{P^*}(\sqrt{|J|_0 \log p})$. Also, by (H.46)

$$\frac{1}{nT} \| \widehat{U}^{*\prime}_{\widehat{J}^*} (\tilde{U}^* - \widehat{U}^*) \widehat{\gamma}_d \|_2 \leq \sqrt{|\widehat{J}|_0} \max_k \left| \frac{1}{nT} \sum_i \sum_t \sum_m \widehat{U}^*_{it,k} (\tilde{U}^*_{it,m} - \widehat{U}^*_{it,m}) \widehat{\gamma}_{dm} \right|$$
$$\leq O_{P^*}(|J|_0^{1/2}) \left(\Delta^{*2}_F |J|_0 + \Delta^*_{ud} |J|_0 + |J|_0 \Delta^*_F b_n + |J|_0 \frac{\log(pT)}{n} \right) := O_{P^*}(c_n),$$

where $b_n = \left(\sqrt{\frac{\log p}{n}} \right) \sqrt{\frac{\log(n)\log(pT)}{n} + \Delta^2_{\max} + \frac{\log(pn)}{T}}$. Therefore, the last term of (H.21) is bounded by $O_{P^*}(\sqrt{|J|_0 \log p}) c_n = o_{P^*}(1)$. Thus, both $\frac{1}{\sqrt{nT}} \tilde{\epsilon}^{*\prime} M_{\widehat{U}^*_{\widehat{J}^*}} (\tilde{U}^* - \widehat{U}^*) \widehat{\gamma}_d$ and $\frac{1}{\sqrt{nT}} \tilde{\epsilon}^{*\prime} M_{\widehat{U}^*_{\widehat{J}^*}} \widehat{U}^* \widehat{\gamma}_d$ are $o_{P^*}(1)$.

(v) Since $\frac{1}{\sqrt{nT}} \tilde{\epsilon}^{*\prime} P_{\widehat{U}^*_{\widehat{J}^*}} (I_T \otimes P_{\widehat{F}^*}) \tilde{U}^* \widehat{\gamma}_d = 0$, it follows from parts (ii) and (iv) that

$$\frac{1}{\sqrt{nT}} \tilde{\epsilon}^{*\prime} M_{\widehat{U}^*_{\widehat{J}^*}} (I_T \otimes M_{\widehat{F}^*}) \tilde{U}^* \widehat{\gamma}_d = \frac{1}{\sqrt{nT}} \tilde{\epsilon}^{*\prime} M_{\widehat{U}^*_{\widehat{J}^*}} \tilde{U}^* \widehat{\gamma}_d - \frac{1}{\sqrt{nT}} \tilde{\epsilon}^{*\prime} (I_T \otimes P_{\widehat{F}^*}) \tilde{U}^* \widehat{\gamma}_d = o_{P^*}(1).$$

(vi) By Lemma H.7, $\left\| \frac{1}{\sqrt{nT}} M_{\widehat{U}^*_{\widehat{J}^*}} \tilde{U}^* \widehat{\gamma} \right\|_2^2 = O_{P^*} \left(a_n + \kappa_n^2 |J|_0 + |J|_0^2 \Delta^{*2}_F + \frac{|J|_0}{n} \right)$. Thus,

$$\left| \widehat{\Xi}' \, (I_T \otimes H^{*\prime -1} (\widehat{F} H^* - \widehat{F}^*)' M_{\widehat{F}^*}) \frac{1}{\sqrt{nT}} M_{\widehat{U}^*_{\widehat{J}^*}} \tilde{U}^* \widehat{\gamma}_d \right|^2$$
$$\leq O_{P^*}(T) \left\| \widehat{F}^* - \widehat{F} H^* \right\|_F^2 \left\| \frac{1}{\sqrt{nT}} M_{\widehat{U}^*_{\widehat{J}^*}} \tilde{U}^* \widehat{\gamma}_d \right\|_2^2$$

$$= O_{P^*}\left(nT\Delta_F^{*2}\right)\left(\kappa_n^2|J|_0 + a_n + \Delta_F^{*2}|J|_0^2 + \frac{|J|_0}{n}\right) = o_{P^*}(1).$$

■

H.3. **Technical lemmas for Proposition F.1.**

Lemma H.9. $\kappa_n^2|J|_0\sqrt{nT} = o(1)$.

Proof. Note that $\kappa_n^2 = O(\log(pq_n^{-1})/(nT))$. To prove $\kappa_n^2|J|_0\sqrt{nT} = o(1)$, it thus suffices to show $|J|_0^2\log^2 p = o(nT)$ and $|J|_0^2\log^2\left(\frac{1}{q_n}\right) = o(nT)$, which is guaranteed by our assumptions. ■

Lemma H.10. *For all t, the $K\times p$ matrix $\widehat{F}'\widehat{U}_t$ satisfies $\widehat{F}'\widehat{U}_t = 0$. Also, $P_{\widehat{U}_{\widehat{J}}}(I_T\otimes P_{\widehat{F}}) = 0$, and $\widehat{U}_{\widehat{J}}'(I_T\otimes P_{\widehat{F}}) = 0$.*

Proof. Note that the $n\times p$ matrix $\widehat{U}_t = \tilde{X}_t - \widehat{F}(\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{X}_t$. Thus, $\widehat{F}'\widehat{U}_t = \widehat{F}'M_{\widehat{F}}\tilde{X}_t = 0$. Straightforward calculations yield $\widehat{U}_{\widehat{J}}'(I_T\otimes P_{\widehat{F}}) = \sum_t \widehat{U}_{t,\widehat{J}}'\widehat{F}(\widehat{F}'\widehat{F})^{-1}\widehat{F}' = 0$ since $\widehat{U}_t'\widehat{F} = 0$. It also follows that $P_{\widehat{U}_{\widehat{J}}}(I_T\otimes P_{\widehat{F}}) = 0$. ■

Lemma H.11. *For $\gamma = \gamma_y$ or $\gamma = \gamma_d$,*
(i) Both

$$\max_{t\leq T, m\leq p}\|\frac{1}{n}\sum_{i=1}^n \widehat{f}_i\tilde{U}_{it,m}\|_2 = O_P\left(\sqrt{\frac{\log(pT)}{n}} + \Delta_{fum}\right)$$

and

$$\max_{m\leq p}\left(\frac{1}{n^2T}\sum_t\|\tilde{U}_{t,m}'\widehat{F}\|_2^2\right)^{1/2} = O_P\left(\sqrt{\frac{\log(pT)}{n}} + \Delta_{fum}\right).$$

(ii) $\max_{m\leq p}\|\frac{1}{nT}\sum_{t=1}^T\widehat{F}'\tilde{U}_t\gamma\tilde{\lambda}_{tm}'\|_F = O_P\left(\sqrt{\frac{|J|_0\log p}{nT}} + \Delta_F|J|_0\right)$.
(iii) Both

$$\max_{m\leq p}\|\frac{1}{nT}\sum_{t=1}^T\widehat{F}'\tilde{e}_t\tilde{\lambda}_{tm}'\|_F = O_P\left(\Delta_{eg} + \sqrt{\frac{\log p}{nT}}\right)$$

and

$$\max_m\|\frac{1}{T}\sum_{t=1}^T\frac{1}{n}\widehat{F}'\tilde{U}_{t,m}\tilde{\delta}_{yt}'\|_F = O_P\left(\Delta_{ud} + \sqrt{\frac{\log p}{nT}}\right).$$

(iv) We have

$$\frac{1}{T}\sum_t\|\frac{1}{n}\widehat{F}'\tilde{e}_t\|_2^2 = O_P\left(\frac{1}{n} + \Delta_{fe}^2\right),$$

$$\frac{1}{T}\sum_t \|\frac{1}{n}\widehat{F}'\tilde{\eta}_t\|_2^2 = O_P\left(\frac{1}{n}+\Delta_{fe}^2\right),$$

and

$$\frac{1}{T}\sum_t \|\frac{1}{n}\widehat{F}'\tilde{U}_t\gamma\|_2^2 = O_P\left(\frac{|J|_0}{n}+|J|_0^2\Delta_F^2\right).$$

(v) $\max_{m\le p}\|\frac{1}{nT}\sum_{t=1}^T \widehat{F}'\tilde{U}_{t,m}\gamma'\tilde{\Lambda}_t\|_F = O_P\left(\Delta_F|J|_0+|J|_0\sqrt{\frac{\log p}{nT}}\right)$.

(vi) $\max_{m,l\le p}|\frac{1}{nT}\sum_{it}\tilde{U}_{it,l}\widehat{f}_i'\tilde{\lambda}_{tm}| = O_P\left(\Delta_{ud}+\sqrt{\frac{\log p}{nT}}\right)$.

Proof. For notational simplicity, we assume f_i to be a scalar without loss of generality.

(i) First, $\frac{1}{n}\sum_{i=1}^n \tilde{f}_i\tilde{U}_{it,m} = \frac{1}{n}\sum_{i=1}^n f_iU_{it,m} - \bar{f}\bar{U}_{\cdot t,m} - \frac{1}{n}\sum_i f_i\bar{U}_{i\cdot,m} + \bar{f}\bar{\bar{U}}_m$. Then

$$\max_{t\le T,m\le p}\|\frac{1}{n}\sum_{i=1}^n \tilde{f}_i\tilde{U}_{it,m}\|_2 = O_P\left(\sqrt{\frac{\log(pT)}{n}}\right).$$

Thus, $\max_{t\le T,m\le p}\|\frac{1}{n}\sum_{i=1}^n \widehat{f}_i\tilde{U}_{it,m}\|_2 = O_P\left(\sqrt{\frac{\log(pT)}{n}}+\Delta_{fum}\right)$. Also,

$$\max_m \frac{1}{n^2T}\sum_t \|\tilde{U}_{t,m}'\widehat{F}\|_2^2 = \max_m \frac{1}{T}\sum_t \|\frac{1}{n}\sum_i \tilde{U}_{it,m}\widehat{f}_i\|_2^2 = O_P\left(\frac{\log(pT)}{n}+\Delta_{fum}^2\right).$$

(ii) We first bound $\max_m\|\frac{1}{nT}\sum_{t=1}^T \tilde{F}'\tilde{U}_t\gamma\tilde{\lambda}_{tm}'\|_F$.

$$\begin{aligned}\max_m\left\|\frac{1}{nT}\sum_{t=1}^T \tilde{F}'\tilde{U}_t\gamma\tilde{\lambda}_{tm}'\right\|_F &= \max_m\left\|\frac{1}{nT}\sum_{t=1}^T\sum_i f_i\tilde{U}_{it}'\gamma\tilde{\lambda}_{tm}'\right\|_F \\ &\le |\bar{\bar{U}}'\gamma|O_P(1) &\text{(H.22)}\\ &\quad +\max_m\left\|\frac{1}{nT}\sum_{t=1}^T\sum_i f_iU_{it}'\gamma\tilde{\lambda}_{tm}'\right\|_F &\text{(H.23)}\\ &\quad +O_P(1)\max_m\left\|\frac{1}{T}\sum_{t=1}^T \bar{U}_{\cdot t}'\gamma\tilde{\lambda}_{tm}'\right\|_2 &\text{(H.24)}\\ &\quad +O_P(1)\left\|\frac{1}{n}\sum_i f_i\bar{U}_{i\cdot}'\gamma\right\|_2. &\text{(H.25)}\end{aligned}$$

For term (H.22), note

$$E|\bar{\bar{U}}'\gamma|^2 \le \frac{1}{nT}\|\gamma\|_1\max_{i,t,m}\sum_{s=1}^T\sum_{k=1}^p |EU_{it,m}U_{is,k}| = O\left(\frac{|J|_0}{nT}\right).$$

Hence, $|\bar{\bar{U}}'\gamma| = O_P\left(\sqrt{\frac{|J|_0}{nT}}\right)$.

For term (H.23), let $W_{im} = \frac{1}{T}\sum_t f_i U_{it}'\gamma\tilde{\lambda}_{tm}$. Note that the W_{im}'s are independent across i and $E(W_{im}) = 0$. By assumption, $\max_{l,t}\sum_{k=1}^p\sum_s |E(U_{it,k}\ U_{is,l}|F)| = O(1)$; so

$$\max_{im} \mathrm{Var}(W_{im}) \leq \max_{im} \frac{1}{T}\|\gamma\|_1 E f_i^2 \max_{l,t}\sum_{k=1}^p\sum_s |E(U_{it,k}\ U_{is,l}|F)| = O\left(\frac{|J|_0}{T}\right).$$

Hence $\max_m |\frac{1}{n}\sum_i W_{im}| = O_P\left(\sqrt{\frac{|J|_0 \log p}{nT}}\right)$.

For term (H.24), let $Z_{im} = \frac{1}{T}\sum_t U_{it}'\gamma\tilde{\lambda}_{tm}$. Then $\max_{im}\mathrm{Var}(Z_{im}) = O\left(\frac{|J|_0}{T}\right)$, and

$$\max_m \|\frac{1}{T}\sum_{t=1}^T \bar{U}_{\cdot t}'\gamma\tilde{\lambda}_{tm}'\|_2 = \max_m \|\frac{1}{n}\sum_i Z_{im}\|_2 = O_P(\sqrt{\frac{|J|_0 \log p}{nT}}).$$

Finally,

$$E|\frac{1}{n}\sum_i f_i \bar{U}_{i\cdot}'\gamma|^2 \leq \frac{1}{nT}\|\gamma\|_1 E f_i^2 \max_{imt}\sum_s\sum_{k=1}^p |E(U_{is,k}U_{it,m}|F)|.$$

Hence, $\|\frac{1}{n}\sum_i f_i\bar{U}_{i\cdot}'\gamma\|_2 = O_P\left(\sqrt{\frac{|J|_0}{nT}}\right)$. Combining the four terms above, we have

$$\max_m \|\frac{1}{nT}\sum_{t=1}^T \tilde{F}'\tilde{U}_t\gamma\tilde{\lambda}_{tm}'\|_F = O_P\left(\sqrt{\frac{|J|_0 \log p}{nT}}\right).$$

Next, $\frac{1}{n}\sum_i |\frac{1}{T}\sum_t \tilde{U}_{it}'\gamma|^2 \max_m \|\tilde{\lambda}_m\|_2^2 = O_P(|J|_0^2)$, so

$$\begin{aligned}\max_m \|\frac{1}{nT}\sum_{t=1}^T(\tilde{F}H - \widehat{F})'\tilde{U}_t\gamma\tilde{\lambda}_{tm}'\|_F^2 &\leq \frac{1}{n}\sum_i \|\widehat{f}_i - H'\tilde{f}_i\|_2^2 \frac{1}{n}\sum_i |\frac{1}{T}\sum_t \tilde{U}_{it}'\gamma|^2 \max_m \|\tilde{\lambda}_m\|_2^2 \\ &= O_P(\Delta_F^2|J|_0^2).\end{aligned}$$

Hence, we have $\max_m \|\frac{1}{nT}\sum_{t=1}^T \widehat{F}'\tilde{U}_t\gamma\tilde{\lambda}_{tm}'\|_F = O_P\left(\sqrt{\frac{|J|_0 \log p}{nT}} + \Delta_F|J|_0\right)$.

(iii) First of all, note that $\frac{1}{nT}\sum_{t=1}^T\sum_{i=1}^n \tilde{e}_{it}\tilde{\lambda}_{tm}' = 0$ and $\frac{1}{T}\sum_t \tilde{\lambda}_{tm} = 0$, so

$$\max_{m\leq p}\|\frac{1}{nT}\sum_{t=1}^T \tilde{F}'\tilde{e}_t\tilde{\lambda}_{tm}'\|_F = \max_{m\leq p}\|\frac{1}{nT}\sum_{t=1}^T\sum_{i=1}^n \tilde{f}_i\tilde{e}_{it}\tilde{\lambda}_{tm}'\|_F = \max_{m\leq p}\|\frac{1}{nT}\sum_{t=1}^T\sum_{i=1}^n f_i\tilde{e}_{it}\tilde{\lambda}_{tm}'\|_F$$

$$\leq \max_{m\leq p}\|\frac{1}{nT}\sum_{t=1}^T\sum_{i=1}^n f_ie_{it}\tilde\lambda_{tm}'\|_F+\max_{m\leq p}\|\frac{1}{nT}\sum_{t=1}^T\sum_{i=1}^n f_i\bar e_{i\cdot}\tilde\lambda_{tm}'\|_F$$
$$+\max_{m\leq p}\|\frac{1}{nT}\sum_{t=1}^T\sum_{i=1}^n f_i\bar e_{\cdot t}\tilde\lambda_{tm}'\|_F+\max_{m\leq p}\|\frac{1}{nT}\sum_{t=1}^T\sum_{i=1}^n f_i\bar{\bar e}\tilde\lambda_{tm}'\|_F$$
$$\leq \max_{m\leq p}\|\frac{1}{nT}\sum_{t=1}^T\sum_{i=1}^n f_ie_{it}\tilde\lambda_{tm}'\|_F+\max_{m\leq p}\|\frac{1}{T}\sum_{t=1}^T\bar e_{\cdot t}\tilde\lambda_{tm}'\|_2O_P(1).$$

Fix an element k of $\tilde\lambda_{tm}$, let $W_{im,k}=\frac{1}{T}\sum_{t=1}^T f_ie_{it}\tilde\lambda_{tm,k}$. Then $EW_{im,k}=0$, and $\{W_{im,k}\}$'s are independent across $i\leq n$; $\max_{imk}\mathrm{Var}(W_{im,k})=O(\frac{1}{T})$. In addition, by Assumption 3.1, $W_{im,k}$ satisfies the exponential-tail condition. Also, $K=O(1)$, hence we can ignore $\max_{k\leq K}$ in the follows by bounding it for each fixed k. By the Bernstein inequality for independent data, (e.g., Bühlmann and van de Geer (2011)) we reach

$$\max_{m\leq p}\|\frac{1}{nT}\sum_{t=1}^T\sum_{i=1}^n f_ie_{it}\tilde\lambda_{tm}'\|_F\leq K\max_{k\leq K}\max_{m\leq p}|\frac{1}{n}\sum_i W_{im,k}|=O_P(\sqrt{\frac{\log p}{nT}}).$$

More specifically, by Assumption 3.1 (iv), $W_{im,k}$ has exponential tails, then we can apply the Bernstein inequality for independent data to reach ($K=O(1)$), for $y=M\sqrt{\frac{\log p}{nT}}$ and sufficiently large $M>0$,

$$P(\max_{k\leq K}\max_{m\leq p}|\frac{1}{n}\sum_i W_{im,k}|>y)\leq Kp\max_{k\leq K}\max_{m\leq p}P(|\frac{1}{n}\sum_i W_{im,k}|>y)$$
$$\leq \exp\left(\log(Kp)-\frac{ny^2}{\max_{imk}\mathrm{Var}(W_{im,k})}\right)\leq\exp\left(\log(Kp)-CTny^2\right)=o(1).$$

In addition, let $Z_{im,k}=\frac{1}{T}\sum_{t=1}^T e_{it}\tilde\lambda_{tm,k}$. Then $\max_{ikm}\mathrm{Var}(Z_{imk})=O(\frac{1}{T})$. Hence

$$\max_{m\leq p}\|\frac{1}{T}\sum_{t=1}^T\bar e_{\cdot t}\tilde\lambda_{tm}'\|_2=\max_{m\leq p}\|\frac{1}{Tn}\sum_{t=1}^T\sum_{i=1}^n e_{it}\tilde\lambda_{tm}'\|_2\leq\sqrt{K}\max_k\max_{m\leq p}|\frac{1}{n}\sum_{i=1}^n Z_{im,k}|=O_P(\sqrt{\frac{\log p}{nT}}).$$

This implies

$$\max_{m\leq p}\|\frac{1}{nT}\sum_{t=1}^T\tilde F'\tilde e_t\tilde\lambda_{tm}'\|_F=O_P(\sqrt{\frac{\log p}{nT}}).$$

Also, by Assumption D.4, $\max_{m\leq p}\|\frac{1}{nT}\sum_{t=1}^T(\widehat F-\tilde FH)'\tilde e_t\tilde\lambda_{tm}'\|_F=O_P(\Delta_{eg})$, implying

$$\max_{m\leq p}\|\frac{1}{nT}\sum_{t=1}^T\widehat F'\tilde e_t\tilde\lambda_{tm}'\|_F=O_P\left(\Delta_{eg}+\sqrt{\frac{\log p}{nT}}\right).$$

On the other hand, similar calculations yield

$$\max_m \|\frac{1}{T}\sum_{t=1}^T \frac{1}{n}\tilde F'\tilde U_{t,m}\tilde\delta_{yt}'\|_F \leq \max_{m\leq p}\|\frac{1}{nT}\sum_{t=1}^T\sum_{i=1}^n f_iU_{it,m}\tilde\delta_{yt}'\|_F + \max_{m\leq p}\|\frac{1}{T}\sum_{t=1}^T \bar U_{\cdot t,m}\tilde\delta_{yt}'\|_2 O_P(1)$$
$$= O_P(\sqrt{\frac{\log p}{nT}}).$$

Also, by Assumption D.4, $\max_m \|\frac{1}{T}\sum_{t=1}^T\frac{1}{n}(\widehat F-\tilde F H)'\tilde U_{t,m}\tilde\delta_{yt}'\|_F = O_P(\Delta_{ud})$. Hence

$$\max_m \|\frac{1}{T}\sum_{t=1}^T \frac{1}{n}\widehat F'\tilde U_{t,m}\tilde\delta_{yt}'\|_F = O_P(\Delta_{ud}+\sqrt{\frac{\log p}{nT}}).$$

(iv) Note that $\tilde e_{it} = e_{it}-\bar e_{\cdot t}-\bar e_{i\cdot}+\bar{\bar e}$. Also, $\frac{1}{n}\sum_i \tilde e_{it}=0$. Hence,

$$\frac{1}{n}\sum_i \tilde f_i\tilde e_{it} = \frac{1}{n}\sum_i f_ie_{it} - \frac{1}{n}\sum_i f_i\bar e_{i\cdot} - \bar f\bar e_{\cdot t} + \bar f\bar{\bar e}.$$

Therefore,

$$\frac{1}{T}\sum_t \|\frac{1}{n}\widehat F'\tilde e_t\|_2^2 \leq \frac{2}{T}\sum_t\|\frac{1}{n}\sum_i(\widehat f_i - H'\tilde f_i)\tilde e_{it}\|_2^2 + O_P(1)\frac{1}{T}\sum_t\|\frac{1}{n}\sum_i\tilde f_i\tilde e_{it}\|_2^2$$
$$= O_P\left(\frac{1}{n}+\Delta_{fe}^2\right)$$
$$\frac{1}{T}\sum_t \|\frac{1}{n}\widehat F'\tilde U_t\gamma\|_2^2 \leq \frac{2}{T}\sum_t\|\frac{1}{n}\sum_i(\widehat f_i - H'\tilde f_i)\tilde U_{it}'\gamma\|_2^2 + O_P(1)\frac{1}{T}\sum_t\|\frac{1}{n}\sum_i\tilde f_i\tilde U_{it}'\gamma\|_2^2$$
$$= O_P\left(\frac{|J|_0}{n}+|J|_0^2\Delta_F^2\right).$$

(v) The proof is very similar to that of (ii). First,

$$\max_{m\leq p}\left\|\frac{1}{nT}\sum_{t=1}^T \tilde F'\tilde U_{t,m}\gamma'\tilde\Lambda_t\right\|_F \leq \max_m\left\|\frac{1}{nT}\sum_{t=1}^T\sum_i f_iU_{it,m}\gamma'\tilde\Lambda_t\right\|_F \tag{H.26}$$
$$+ O_P(1)\max_m\left\|\frac{1}{T}\sum_t \bar U_{\cdot t,m}\gamma'\tilde\Lambda_t\right\|_2. \tag{H.27}$$

For the term (H.26), let $Z_{im} = \frac{1}{T}\sum_t f_iU_{it,m}\gamma'\tilde\Lambda_t$. Then

$$\max_{im}\mathrm{Var}(Z_{im}) = O\left(\frac{|J|_0^2}{T}\right),$$

and

$$\max_m \|\frac{1}{n}\sum_i Z_{im}\|_2 = O_P\left(|J|_0\sqrt{\frac{\log p}{nT}}\right),$$

which gives the rate of convergence of (H.26). For term (H.27), let $W_{im} = \frac{1}{T}\sum_t U_{it,m}\gamma'\tilde{\Lambda}_t$. The W_{im}'s are independent across i, have $E(W_{im}) = 0$, and have $\max_{im}\mathrm{Var}(W_{im}) = O(|J|_0^2/T)$. Hence, $\max_m |\frac{1}{n}\sum_i W_{im}| = O_P\left(|J|_0\sqrt{\frac{\log p}{nT}}\right)$, which gives the rate of convergence for (H.27). Combining then yields

$$\max_{m\leq p}\|\frac{1}{nT}\sum_{t=1}^T \tilde{F}'\tilde{U}_{t,m}\gamma'\tilde{\Lambda}_t\|_F = O_P\left(|J|_0\sqrt{\frac{\log p}{nT}}\right).$$

Finally, using $\max_m \frac{1}{nT}\sum_{it}\tilde{U}_{it,m}^2\frac{1}{T}\sum_t\|\gamma_y'\tilde{\Lambda}_t\|_2^2 = O_P(|J|_0^2)$, we have

$$\begin{aligned}&\max_m\left\|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T(\widehat{f}_i - H'\tilde{f}_i)\tilde{U}_{it,m}\gamma_y'\tilde{\Lambda}_t\right\|_F^2\\ &\leq \frac{1}{n}\sum_i\|\widehat{f}_i - H'\tilde{f}_i\|_2^2\max_m\frac{1}{nT}\sum_{it}\tilde{U}_{it,m}^2\frac{1}{T}\sum_t\|\gamma_y'\tilde{\Lambda}_t\|_2^2 = O_P\left(\Delta_F^2|J|_0^2\right).\end{aligned}$$

Hence, we obtain the desired result.

(vi) A proof similar to that of part (iii) yields

$$\max_{m,l\leq p}|\frac{1}{nT}\sum_{it}\tilde{U}_{it,l}\tilde{f}_i'\tilde{\lambda}_{tm}| = O_P\left(\sqrt{\frac{\log p}{nT}}\right).$$

Then applying Assumption D.4 yields $\max_{m,l\leq p}|\frac{1}{nT}\sum_{it}\tilde{U}_{it,l}(\widehat{f}_i - H'\tilde{f}_i)'\tilde{\lambda}_{tm}| = O_P(\Delta_{ud})$. Hence

$$\max_{m,l\leq p}|\frac{1}{nT}\sum_{it}\tilde{U}_{it,l}\widehat{f}_i'\tilde{\lambda}_{tm}| = O_P\left(\Delta_{ud} + \sqrt{\frac{\log p}{nT}}\right).$$

■

Lemma H.12. *(i)* $\max_m\|\frac{1}{n^2T}\sum_{i=1}^n\sum_{t=1}^T\sum_{j=1}^n\tilde{f}_j\tilde{U}_{jt,m}\tilde{e}_{it}\tilde{f}_i'\|_F = o_P(\sqrt{\frac{\log p}{nT}})$.
(ii) $\max_{m\leq p}\|\frac{1}{n^2}\widehat{F}'\frac{1}{T}\sum_t\tilde{U}_t\gamma_y\tilde{U}_{t,m}'\widehat{F}\|_F = o_P(\sqrt{\frac{\log p}{nT}})$.
(iii) $\|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T(\widehat{U}_{it} - \tilde{U}_{it})\tilde{f}_i'H(\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{U}_t\gamma_y\|_\infty = o_P(\sqrt{\frac{\log p}{nT}})$.
(iv) $\|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T(\widehat{U}_{it} - \tilde{U}_{it})\tilde{f}_i'H(\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{e}_t\|_\infty = o_P(\sqrt{\frac{\log p}{nT}})$.

Proof. For notational simplicity, we take f_i to be a scalar without loss of generality.

(i) We have

$$\frac{1}{n^2T}\sum_{i=1}^n\sum_{t=1}^T\sum_{j=1}^n \tilde f_j\tilde U_{jt,m}\tilde e_{it}\tilde f_i' = \frac{1}{n^2T}\sum_{i=1}^n\sum_{t=1}^T\sum_{j=1}^n \tilde f_j U_{jt,m} e_{it}\tilde f_i' - \frac{1}{n^2T}\sum_{i=1}^n\sum_{t=1}^T\sum_{j=1}^n \tilde f_j \bar U_{j\cdot,m}\bar e_{i\cdot}\tilde f_i'.$$

Hence, as $n\to\infty$,

$$\max_m\left\|\frac{1}{n^2T}\sum_{i=1}^n\sum_{t=1}^T\sum_{j=1}^n \tilde f_j\tilde U_{jt,m}\tilde e_{it}\tilde f_i'\right\|_F \leq \max_m\left\|\frac{1}{n^2T}\sum_{i=1}^n\sum_{t=1}^T\sum_{j=1}^n \tilde f_j U_{jt,m} e_{it}\tilde f_i'\right\|_F + o_P\left(\sqrt{\frac{\log p}{nT}}\right).$$

Define $W_{tm} = \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n U_{jt,m}e_{it}\tilde f_j\tilde f_i'$. We aim to bound $\max_m\|\frac{1}{T}\sum_t W_{tm}\|_F$. We assume $\tilde f_j$ is one dimensional for simplicity, and the multivariate case follows from a similar argument. By assumption, $E(U_{jt}e_{it}|F)=0$ almost surely. If $T=O(1)$, then because $\{(U_{it},e_{it})\}_{i\leq n}$ are conditionally independent given F,

$$\begin{aligned}\max_m|\frac{1}{T}\sum_t W_{tm}| &\leq \max_{mt}|\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n U_{jt,m}e_{it}\tilde f_j\tilde f_i| \\ &\leq \max_{mt}|\frac{1}{n}\sum_{i=1}^n U_{jt,m}\tilde f_j|\max_t|\frac{1}{n}\sum_{i=1}^n e_{it}\tilde f_i| \\ &\leq O_P(\frac{\sqrt{\log p}}{n}) = O_P(\sqrt{\frac{\log p}{Tn^2}}). \quad (T=O(1))\end{aligned}$$

If $T\to\infty$, using the assumption that $\{(U_{it},e_{it})\}_{i\leq n}$ are conditionally independent given F, for $E_f(\cdot)=E(\cdot|F)$,

$$\begin{aligned}\max_{tm}\mathrm{Var}(W_{tm}|F) &\leq \frac{1}{n^4}\sum_{l=1}^n\tilde f_l^4E_fU_{lt,m}^2e_{lt}^2 + \frac{1}{n^3}\sum_{l=1}^n\tilde f_l^2E_fU_{lt,m}^2\frac{1}{n}\sum_{k=1}^n\tilde f_k^2E_fe_{kt}^2 \\ &\leq \frac{1}{n^3}\left(\frac{1}{n}\sum_{l=1}^n\tilde f_l^8\right)^{1/2}\left(\frac{1}{n}\sum_{l=1}^nE_fU_{lt,m}^4e_{lt}^4\right)^{1/2} \\ &\quad + \frac{1}{n^2}\left(\frac{1}{n}\sum_{l=1}^n\tilde f_l^4\right)\left(\frac{1}{n}\sum_l E_fU_{lt,m}^4\right)^{1/2}\left(\frac{1}{n}\sum_{k=1}^nE_fe_{kt}^4\right)^{1/2}.\end{aligned}$$

There is $C > 0$ such that $\max_m \frac{1}{n}\sum_l E_f U_{lt,m}^8$ and $\sum_{k=1}^n E_f e_{kt}^8$ are upper bounded by C almost surely (a.s.) in F. Thus a.s. in F, $\max_{tm} \mathrm{Var}(W_{tm}|F) \leq \frac{C}{n^2}(\frac{1}{n}\sum_{l=1}^n \tilde{f}_l^8)^{1/2}$. Hence, for any $x > 0$ and a generic constant $C > 0$,

$$\begin{aligned} P\left(\max_m |\frac{1}{T}\sum_t W_{tm}| > x\right) &\leq p \max_m EP\left(|\frac{1}{T}\sum_t W_{tm}| > x|F\right) \\ &\leq p \max_m EP\left(|\frac{1}{T}\sum_t W_{tm}| > x|F\right) 1\{\frac{1}{n}\sum_{l=1}^n \tilde{f}_l^8 < C\} \\ &\quad + pP\left(|\frac{1}{n}\sum_{l=1}^n (f_l^{16} - Ef_l^{16})| > C\right) + pP\left(|\bar{f} - E\bar{f}| > C\right) \\ &\leq^{(a)} p \max_m EP\left(|\frac{1}{T}\sum_t W_{tm}| > x|F\right) 1\{\max_{tm} \mathrm{Var}(W_{tm}|F) \leq \frac{C}{n^2}\} + o(1) \end{aligned}$$

where (a) follows from the Bernstein inequality for independent data and $\log p = o(n)$, applied to $pP\left(|\frac{1}{n}\sum_{l=1}^n (f_l^{16} - Ef_l^{16})| > C\right)$ and $pP\left(|\bar{f} - E\bar{f}| > C\right)$.

We now bound $P\left(|\frac{1}{T}\sum_t W_{tm}| > x|F\right)$ conditioning on the event $\max_{tm} \mathrm{Var}(W_{tm}|F) \leq \frac{C}{n^2}$. We have $\min_{tm}[\mathrm{Var}(W_{tm}|F)^{-1/2}] \geq C_0 n$ for a generic $C_0 > 0$. Conditional on F, $\{W_{tm}\}$ is a strong mixing sequence across t with uniform mixing condition (uniform over $m \leq p$) that is bounded by $\exp(-CT^r)$ (Assumption 3.1). Write $\mathcal{X}_{tm} = W_{tm} \mathrm{Var}(W_{tm}|F)^{-1/2}$. Recall that by assumption, $\log^\gamma p = o(n)$ for some $\gamma > 2$ and the strong mixing (Assumption 3.1) holds for the process $\{(\eta_t, \epsilon_t)\}_{t=-\infty}^{+\infty}$, with mixing coefficient bounded by $\exp(-CT^r)$, $r > 1$, and also we assume $\gamma r \geq 2$. Let $\bar{r} = \min\{r, 1\}$, $r_1 = (0.5 + \bar{r}^{-1})^{-1}$, $c = 0.5(\gamma + 1)$, then $r_1 < 1$, $cr_1 > 1$ and $2c \geq 1$. Because $\bar{r} \leq r$, the strong mixing condition in Assumption 3.1 also holds with $\bar{r}$ in place of r. All these constants are independent of F. Then by the Bernstein inequality for strong mixing sequences Merlevède et al. (2011), for $y = M\frac{(\log p)^c}{\sqrt{T}}$, and sufficiently large $M > 0$, for $C_0 n x = y$,

$$\begin{aligned} p \max_m P\left(|\frac{1}{T}\sum_t W_{tm}| > x|F\right) &\leq p \max_m P\left(|\frac{1}{T}\sum_t \mathcal{X}_{tm}| > x \min_{tm}[\mathrm{Var}(W_{tm}|F)^{-1/2}]|F\right) \\ &\leq p \max_m P\left(|\frac{1}{T}\sum_t \mathcal{X}_{tm}| > C_0 n x|F\right) \leq A_1 + A_2 + A_3 \end{aligned}$$

where for a generic constant C that is independent of F,

$$A_1 = pT\exp(-C(Ty)^{r_1}) = \exp(\log(pT) - CM^{r_1}T^{r_1/2}\log^{cr_1} p) = o(1), \quad (cr_1 \geq 1)$$

$$
\begin{aligned}
A_2 &= p\exp\left(-C\frac{(Ty)^2}{T}\exp\left(\frac{(Ty)^{r_1(1-r_1)}}{C\log^{r_1}(Ty)}\right)\right) = o(1), \quad (r_1<1, 2c>1, Ty\to\infty)\\
A_3 &= p\exp\left(-CTy^2\right) = \exp(\log p - CM^2\log^{2c}p) = o(1),
\end{aligned}
$$

and $o(1)$ in the above are non-stochastic converging sequences. (This is because the involved generic constant C in Merlevède et al. (2011) is independent of F, given that the conditional strong mixing and sub-Gaussian conditions (given F) in Assumption 3.1 hold almost surely in F.) Hence

$$
p\max_m EP\left(|\frac{1}{T}\sum_t W_{tm}|>x|F\right)1\{\max_{tm}\mathrm{Var}(W_{tm}|F)\leq\frac{C}{n^2}\} = o(1).
$$

Consequently, $\max_m|\frac{1}{T}\sum_t W_{tm}| = O_P(x) = O_P(\frac{\log^{0.5\gamma+0.5}p}{\sqrt{Tn^2}})$ when $T\to\infty$. Combining with the case $T=O(1)$, we have $\max_m|\frac{1}{T}\sum_t W_{tm}| = O_P(\frac{\log^{0.5\gamma+0.5}p}{\sqrt{Tn^2}})$. Hence,

$$
\max_m\|\frac{1}{n^2T}\sum_{i=1}^n\sum_{t=1}^T\sum_{j=1}^n\tilde f_j U_{jt,m}e_{it}\tilde f_i'\|_F = \max_m|\frac{1}{T}\sum_t W_{tm}| = O_P(\frac{\log^{0.5\gamma+0.5}p}{\sqrt{Tn^2}}) = o_P(\sqrt{\frac{\log p}{nT}}).
$$

where the last equality is due to $\log^\gamma p = o(n)$.

(ii) By Assumption D.4, $\max_{mt}\|\frac{1}{n}\sum_{j=1}^n(\widehat f_j - H'\tilde f_j)\tilde U_{jt,m}\|_2 = o_P\left(\sqrt{\frac{\log p}{T}}\right)$. Hence,

$$
\begin{aligned}
&\max_{m\leq p}\|\frac{1}{n^2}\widehat F'\frac{1}{T}\sum_t\tilde U_t\gamma_y\tilde U_{t,m}'\widehat F\|_F\\
&\quad\leq O_P(1)\max_{m\leq p}\|\frac{1}{n^2}\tilde F'\frac{1}{T}\sum_t\tilde U_t\gamma_y\tilde U_{t,m}'\tilde F\|_F\\
&\qquad + \|\gamma_y\|_1\max_{mt}|\frac{1}{n}\sum_i(\widehat f_i' - H'\tilde f_i)\tilde U_{it,m}|\max_{m\leq p}\left(\frac{1}{n^2T}\sum_t\|\tilde U_{t,m}'\widehat F\|_2^2\right)^{1/2}.
\end{aligned}
\tag{H.28}
$$

By Lemma H.11, $\max_{m\leq p}\left(\frac{1}{n^2T}\sum_t\|\tilde U_{t,m}'\widehat F\|_2^2\right)^{1/2} = O_P\left(\sqrt{\frac{\log(pT)}{n}}+\Delta_{fum}\right)$. Hence, the second term on the right-hand-side of the inequality in (H.28) is $O_P\left(|J|_0\Delta_{fum}\right)\left(\sqrt{\frac{\log(pT)}{n}}+\Delta_{fum}\right)$.

We now bound the first term on the right-hand-side of the inequality in (H.28). We have

$$\begin{aligned}
&\max_{m\le p}|\frac{1}{n^2}\tilde{F}'\frac{1}{T}\sum_t \tilde{U}_t\gamma_y\tilde{U}_{t,m}'\tilde{F}| \\
&= \max_{m\le p}|\frac{1}{n^2T}\sum_t\sum_{j=1}^n\sum_{i=1}^n\sum_{l=1}^p \tilde{f}_i\tilde{f}_j\tilde{U}_{it,l}\tilde{U}_{jt,m}\gamma_{yl}| \\
&\le O_P\left(\frac{\log p}{nT}|J|_0\right) \\
&\quad + \max_{m\le p}|\frac{1}{n^2T}\sum_l\gamma_{yl}\sum_t\sum_{j=1}^n\sum_{i=1}^n \tilde{f}_i\tilde{f}_j(U_{it,l}-\bar{U}_{\cdot t,l})(U_{jt,m}-\bar{U}_{\cdot t,m})|
\end{aligned} \tag{H.29}$$

where we obtain the $O_P\left(|J|_0\frac{\log p}{nT}\right)$ term in the second inequality by bounding

$$\begin{aligned}
\max_m|\frac{1}{T}\sum_i f_i\bar{U}_{i\cdot m}|^2 &= O_P\left(\frac{\log p}{n}\max_{im}\mathrm{Var}(f_i\bar{U}_{i,\cdot,m})\right) \\
&= O_P\left(\frac{\log p}{n}\right)\max_{im}\frac{1}{T^2}\sum_t\sum_s Ef_i^2U_{it,m}U_{is,m} \\
&= O_P\left(\frac{\log p}{n}\right)\max_{im,t}\frac{1}{T}Ef_i^2|\sum_s(EU_{it,m}U_{is,m}|f_i)| \\
&= O_P\left(\frac{\log p}{nT}\right),
\end{aligned}$$

using Assumption 3.2 that $\max_{im,t}\sum_s|E(U_{it,m}U_{is,m}|f_i)|<C$ almost surely in f_i.

It remains to bound the remaining term on the right-hand-side of the inequality in (H.29). This term is less than

$$\begin{aligned}
&\underbrace{\max_{m\le p}|\frac{1}{n^2T}\sum_l\gamma_{yl}\sum_t\sum_{j=1}^n\sum_{i=1}^n\tilde{f}_i\tilde{f}_jU_{it,l}U_{jt,m}|}_{z_1} + \underbrace{\max_{m\le p}|\frac{1}{n^2T}\sum_l\gamma_{yl}\sum_t\sum_{j=1}^n\sum_{i=1}^n\tilde{f}_i\tilde{f}_jU_{it,l}\bar{U}_{\cdot t,m}|}_{z_2} \\
&\quad + \underbrace{\max_{m\le p}|\frac{1}{n^2T}\sum_l\gamma_{yl}\sum_t\sum_{j=1}^n\sum_{i=1}^n\tilde{f}_i\tilde{f}_j\bar{U}_{\cdot t,l}U_{jt,m}|}_{z_3} + \underbrace{\max_{m\le p}|\frac{1}{n^2T}\sum_l\gamma_{yl}\sum_t\sum_{j=1}^n\sum_{i=1}^n\tilde{f}_i\tilde{f}_j\bar{U}_{\cdot t,l}\bar{U}_{\cdot t,m}|}_{z_4}.
\end{aligned}$$

We now bound each of z_1-z_4.

z_1: Define

$$Z_{1t,lm}=\frac{1}{n}\sum_{i=1}^n\tilde f_i^2(U_{it,l}U_{it,m}-E(U_{it,l}U_{it,m}|F))$$

and

$$Z_{2t,lm}=\frac{1}{n^2}\sum_{i=1}^n\sum_{j\neq i}\tilde f_i\tilde f_jU_{it,l}U_{jt,m}.$$

By the assumption that $E(U_{it,l}U_{jt,m}|F)=0$ for $i\neq j$, we have $E(Z_{1t,lm}|F)=0$. Also note that $\|\gamma_y\|_1\leq O(|J|_0)+\|R_y\|_1=O(|J|_0)+o(1)=O(|J|_0)$.

Now, consider

$$\begin{aligned}
z_1 &\leq \max_{m\leq p}|\frac{1}{n^2T}\sum_l\gamma_{yl}\sum_t\sum_{i=1}^n\tilde f_i^2U_{it,l}U_{it,m}|+\max_{m\leq p}|\frac{1}{n^2T}\sum_l\gamma_{yl}\sum_t\sum_{i=1}^n\sum_{j\neq i}\tilde f_i\tilde f_jU_{it,l}U_{jt,m}|\\
&\leq \max_{m\leq p}\left|\frac{1}{n^2T}\sum_l\gamma_{yl}\sum_t\sum_{i=1}^n\tilde f_i^2(U_{it,l}U_{it,m}-E(U_{it,l}U_{it,m}|F))\right|\\
&\quad+\max_{m\leq p}|\frac{1}{n^2T}\sum_l\gamma_{yl}\sum_t\sum_{i=1}^n\tilde f_i^2E(U_{it,l}U_{it,m}|F)|\\
&\quad+\max_{m\leq p}|\frac{1}{n^2T}\sum_l\gamma_{yl}\sum_t\sum_{i=1}^n\sum_{j\neq i}\tilde f_i\tilde f_jU_{it,l}U_{jt,m}|\\
&=\max_{m\leq p}\left|\frac{1}{n}\sum_l\gamma_{yl}\frac{1}{T}\sum_t Z_{1t,lm}\right|+\max_{m\leq p}|\frac{1}{n^2T}\sum_l\gamma_{yl}\sum_t\sum_{i=1}^n\tilde f_i^2E(U_{it,l}U_{it,m}|F)|\\
&\quad+\max_{m\leq p}|\frac{1}{T}\sum_l\gamma_{yl}\sum_t Z_{2t,lm}|\\
&\leq \|\gamma_y\|_1\frac{1}{n}\max_{m,l\leq p}\left|\frac{1}{T}\sum_t Z_{1t,lm}\right|+\|\gamma_y\|_1\max_{m,l\leq p}|\frac{1}{T}\sum_t Z_{2t,lm}|\\
&\quad+\max_{l\leq p}|\gamma_{yl}|\max_{m\leq p}\frac{1}{n^2T}\sum_t\sum_{i=1}^n\tilde f_i^2\sum_{l=1}^p|E(U_{it,l}U_{it,m}|F)|\\
&\leq O(\frac{|J|_0}{n})\max_{m,l\leq p}\left|\frac{1}{T}\sum_t Z_{1t,lm}\right|+O(|J|_0)\max_{l,m\leq p}\left|\frac{1}{T}\sum_t Z_{2t,lm}\right|\\
&\quad+O_P(\frac{1}{n})\max_{m\leq p,i\leq n,t\leq T}\sum_{l=1}^p|E(U_{it,l}U_{it,m}|F)|
\end{aligned}$$

$$= O\left(\frac{|J|_0}{n}\right)\max_{m\le p}\left|\frac{1}{T}\sum_t Z_{1t,lm}\right| + O(|J|_0)\max_{l,m\le p}\left|\frac{1}{T}\sum_t Z_{2t,lm}\right| + O\left(\frac{1}{n}\right),$$

where in the last equality we used the assumption that $\max_{m\le p,i\le n,t\le T}\sum_{l=1}^p |E(U_{it,l}U_{it,m}|F)| = O(1)$ almost surely in F (Assumption 3.1 (iii): weak cross-sectional correlations).

To bound $\max_{m\le p}|\frac{1}{T}\sum_t Z_{1t,lm}|$, first note that

$$\begin{aligned}\mathrm{Var}(Z_{1t,lm}|F) &= \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^n \tilde f_i^2 U_{it,l}U_{it,m}|F\right)\\ &= \frac{1}{n^2}\sum_i \tilde f_i^4\,\mathrm{Var}(U_{it,l}U_{it,m}|F)\\ &\le \frac{C}{n}\left(\frac{1}{n}\sum_i \tilde f_i^8\right)^{1/2}\end{aligned}$$

since $\max_{m,t}\frac{1}{n}\sum_i E_f U_{it,m}^8 < C$ a.s. in F.

Conditional on F, $\{Z_{1t,lm}\}_{t\le T}$ is a strong mixing sequence. Hence, for $x = C\sqrt{\frac{\log p}{Tn}}$,

$$\begin{aligned}&P\left(\max_{lm}\left|\frac{1}{T}\sum_t Z_{1t,lm}\right| > x\right)\\ &\quad\le p^2\max_{lm}E\left[P\left(\left|\frac{1}{T}\sum_t Z_{1t,lm}\right| > x|F\right)\right]\\ &\quad\le p^2\max_{lm}E\left[P\left(\left|\frac{1}{T}\sum_t Z_{1t,lm}\right| > x|F\right)1\{\frac{1}{n}\sum_{l=1}^n \tilde f_l^8 < C\}\right]\\ &\qquad + p^2 P\left(\left|\frac{1}{n}\sum_{l=1}^n (f_l^{16} - Ef_l^{16})\right| > C\right) + p^2 P(|\bar f - E\bar f| > C)\\ &\quad\le p^2\max_{lm}E\left[P\left(\left|\frac{1}{T}\sum_t Z_{1t,lm}\right| > x|F\right)1\{\max_{tm}\mathrm{Var}(Z_{1t,lm}|F)\le\frac{C}{n}\}\right] + o(1)\\ &\quad\le E\left[\exp\left(2\log p - \frac{Tx^2}{\max_{tm}\mathrm{Var}(Z_{1t,lm}|F)}\right)1\{\max_{tm}\mathrm{Var}(Z_{1t,lm}|F)\le\frac{C}{n}\}\right] + o(1)\\ &\quad\le \exp(2\log p - CTnx^2) + o(1) = o(p^{-1}),\ \text{for some large } C.\end{aligned}$$

This bound then implies $O(\frac{|J|_0}{n})\max_{m\le p}\left|\frac{1}{T}\sum_t Z_{1t,lm}\right| = O_P\left(\frac{|J|_0}{n}\sqrt{\frac{\log p}{Tn}}\right)$.

To bound $\max_{m\leq p}|\frac{1}{T}\sum_t Z_{2t,lm}|$, uniformly in t,l,m, (recall $E_f(.)=E(.|F)$) we note that

$$\begin{aligned}\mathrm{Var}(Z_{2t,lm}|F) &= \frac{1}{n^4}\sum_{i=1}^n\sum_{j\neq i}\sum_{k=1}^n\sum_{v\neq k}\tilde f_i\tilde f_j\tilde f_k\tilde f_v E_f U_{it,l}U_{jt,m}U_{kt,l}U_{vt,m}\\ &= \frac{1}{n^4}\sum_{i=1}^n\sum_{j\neq i}\tilde f_i^2\tilde f_j^2 E_f U_{it,l}^2 E_f U_{jt,m}^2 + \frac{1}{n^4}\sum_{i=1}^n\sum_{j\neq i}\tilde f_i^2\tilde f_j^2 E_f U_{it,l}U_{it,m}E_f U_{jt,m}U_{jt,l}\\ &\leq \frac{C}{n^3}\sum_{i=1}^n \tilde f_i^4.\end{aligned}$$

Hence, following an argument similar to that used to bound $\max_{m\leq p}\left|\frac{1}{T}\sum_t Z_{1t,lm}\right|$, it can be shown that $\max_{m\leq p}|\frac{1}{T}\sum_t Z_{2t,lm}| = O_P\left(\sqrt{\frac{\log p}{Tn^2}}\right)$. Therefore,

$$z_1 = O_P\left(\frac{1}{n} + |J|_0\sqrt{\frac{\log p}{Tn^2}}\right).$$

For z_2, z_3, note that

$$\begin{aligned} z_2 &\leq \max_{m\leq p}\left|\frac{1}{n^3T}\sum_l \gamma_{yl}\sum_t\sum_{j=1}^n\sum_{i=1}^n\sum_{k=1}^n \tilde f_i\tilde f_j U_{it,l}U_{kt,m}\right|\\ &\leq \max_{m\leq p}\left|\frac{1}{n^2T}\sum_l\gamma_{yl}\sum_t\sum_{i=1}^n\sum_{k=1}^n\tilde f_i U_{it,l}U_{kt,m}\right| O_P(1)\end{aligned}$$

Following the same argument as we used to bound z_1, we reach $z_2 = O_P\left(\frac{1}{n} + |J|_0\sqrt{\frac{\log p}{Tn^2}}\right)$. Similarly, we have $z_3 = O_P\left(\frac{1}{n} + |J|_0\sqrt{\frac{\log p}{Tn^2}}\right)$. Finally, as $\frac{1}{n}\sum_{j=1}^n \tilde f_i = \frac{1}{n}\sum_{j=1}^n f_i - \bar f = 0$, we have

$$z_4 = \max_{m\leq p}|\frac{1}{n^2T}\sum_l\gamma_{yl}\sum_t\sum_{j=1}^n\sum_{i=1}^n \tilde f_i\tilde f_j \bar U_{\cdot t,l}\bar U_{\cdot t,m}| = 0$$

Combining the above, we have that the second term of (H.29) is $O_P\left(\frac{1}{n} + |J|_0\sqrt{\frac{\log p}{Tn^2}}\right)$. Therefore, $\max_{m\leq p}|\frac{1}{n^2}\tilde F'\frac{1}{T}\sum_t \tilde U_t\gamma_y\tilde U_{t,m}'\tilde F| = O_P\left(\frac{\log p}{nT}|J|_0 + |J|_0\sqrt{\frac{\log p}{Tn^2}} + \frac{1}{n}\right)$. Thus by (H.28), we have

$$\max_{m\leq p}\left\|\frac{1}{n^2}\widehat F'\frac{1}{T}\sum_t\tilde U_t\gamma_y\tilde U_{t,m}'\widehat F\right\|_F$$

$$= O_P\left(|J|_0\Delta_{fum}\right)\left(\sqrt{\frac{\log(pT)}{n}}+\Delta_{fum}\right)+O_P\left(\frac{\log p}{nT}|J|_0+|J|_0\sqrt{\frac{\log p}{Tn^2}}+\frac{1}{n}\right),$$

which is $o_P\left(\sqrt{\frac{\log p}{nT}}\right)$ given that $\Delta^2_{fum}=o\left(\frac{\log p}{T|J|^2\log(pT)}\right)$, $\Delta^4_{fum}=o\left(\frac{\log p}{|J|^2nT}\right)$, $|J|_0^2\log p = o(nT)$, and $|J|_0^2T=o(n)$. $\Delta^2_{fum}=o\left(\frac{\log p}{T|J|^2\log(pT)}\right)$ is guaranteed in Assumption D.4. To establish $\Delta^4_{fum}=o\left(\frac{\log p}{|J|^2nT}\right)$, note that

$$\max_{mt}\|\frac{1}{n}\sum_{j=1}^n(\widehat{f}_j-H'\tilde{f}_j)\tilde{U}_{jt,m}\|_2^2\le\frac{1}{n}\sum_i\|\widehat{f}_i-H'\tilde{f}_i\|_2^2\max_{mt}\frac{1}{n}\sum_i\tilde{U}^2_{jt,m}.$$

Because $\max_{mt}\frac{1}{n}\sum_i\tilde{U}^2_{it,m}\le C\max_{imt}(EU^2_{it,m}+E\bar{U}^2_{i\cdot,m})+o_P(1)=O_P(1)$, we have $\Delta_{fum}\le\Delta_F$. We then have $\Delta^4_{fum}\le\Delta^4_F=o(\frac{1}{nT|J|_0^4})=o(\frac{\log p}{|J|_0^2nT})$.

(iii) Note that $\tilde{U}_{it,m}-\widehat{U}_{it,m}=(\widehat{\lambda}'_{tm}-\tilde{\lambda}'_{tm}H'^{-1})\widehat{f}_i+\tilde{\lambda}'_{tm}H'^{-1}(\widehat{f}_i-H'\tilde{f}_i)$. Also note $\widehat{\lambda}_{tm}-H^{-1}\tilde{\lambda}_{tm}=(\widehat{F}'\widehat{F})^{-1}\widehat{F}'(\tilde{F}H-\widehat{F})H^{-1}\tilde{\lambda}_{tm}+(\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{U}_{t,m}$. Then up to an $O_P(1)$ term, the object of interest is

$$\max_{m\le p}\left|\frac{1}{n}\sum_{i=1}^n\tilde{f}_i'H(\widehat{F}'\widehat{F})^{-1}\frac{1}{T}\sum_{t=1}^T\widehat{F}'\tilde{U}_t\gamma_y(\widehat{U}_{it,m}-\tilde{U}_{it,m})\right|$$
$$\le\max_{m\le p}\left\|\frac{1}{nT}\sum_{t=1}^T\widehat{F}'\tilde{U}_t\gamma_y\tilde{\lambda}'_{tm}\right\|_F O_P(\Delta_F)+\max_{m\le p}\left\|\frac{1}{n^2}\widehat{F}'\frac{1}{T}\sum_t\tilde{U}_t\gamma_y\tilde{U}'_{t,m}\widehat{F}\right\|_F.$$

By part (ii),

$$\max_{m\le p}\left\|\frac{1}{n^2}\widehat{F}'\frac{1}{T}\sum_t\tilde{U}_t\gamma_y\tilde{U}'_{t,m}\widehat{F}\right\|_F=o_P\left(\sqrt{\frac{\log p}{nT}}\right).$$

By Lemma H.11,

$$\Delta_F\max_{m\le p}\left\|\frac{1}{nT}\sum_{t=1}^T\widehat{F}'\tilde{U}_t\gamma_y\tilde{\lambda}'_{tm}\right\|_F=O_P\left(\Delta_F|J|_0+\sqrt{\frac{|J|_0\log p}{nT}}\right)\Delta_F$$
$$=o_P\left(\sqrt{\frac{\log p}{nT}}\right),$$

given the assumptions that $\sqrt{nT}\Delta_F^2|J|_0=o(1)$. The result follows.

(iv) The object of interest is bounded by

$$\max_{m\le p}\left\|\frac{1}{nT}\sum_{t=1}^T\widehat{F}'\tilde{e}_t\tilde{\lambda}'_{tm}\right\|_F O_P(\Delta_F)+\max_{m\le p}\left\|\frac{1}{n^2}\widehat{F}'\frac{1}{T}\sum_t\tilde{e}_t\tilde{U}'_{t,m}\widehat{F}\right\|_F. \tag{H.30}$$

The first term in equation (H.30) can be bounded similarly to the term

$$\max_{m\leq p}\|\frac{1}{nT}\sum_{t=1}^{T}\widehat{F}'\tilde{U}_t\gamma_y\tilde{\lambda}_{tm}'\|_F$$

in Lemma H.11; and we have, using from Lemma H.11 that

$$\max_{m\leq p}\|\frac{1}{nT}\sum_{t=1}^{T}\widehat{F}'\tilde{e}_t\tilde{\lambda}_{tm}'\|_F = O_P\left(\Delta_{eg}+\sqrt{\frac{\log p}{nT}}\right),$$

that

$$O_P(\Delta_F)\max_{m\leq p}\left\|\frac{1}{nT}\sum_{t=1}^{T}\widehat{F}'\tilde{e}_t\tilde{\lambda}_{tm}'\right\|_F = O_P(\Delta_F)\left(\Delta_{eg}+\sqrt{\frac{\log p}{nT}}\right) = o_P\left(\sqrt{\frac{\log p}{nT}}\right),$$

given Assumption D.4 that $\Delta_{eg} = o\left(\frac{1}{\sqrt{nT}}\right)$.

The second term in (H.30) is bounded similarly to the term $\max_{m\leq p}\left\|\frac{1}{n^2}\widehat{F}'\frac{1}{T}\sum_t\tilde{U}_t\gamma_y\tilde{U}_{t,m}'\widehat{F}\right\|_F$ in part (ii) using that $\max_m\left\|\frac{1}{n^2T}\sum_{i=1}^n\sum_{t=1}^T\sum_{j=1}^n\tilde{f}_jU_{jt,m}e_{it}\tilde{f}_i'\right\|_F = O_P\left(\sqrt{\frac{\log p}{Tn^2}}\right)$ from part (i). Specifically, we have

$$\begin{aligned}\max_{m\leq p}\left\|\frac{1}{n^2}\widehat{F}'\frac{1}{T}\sum_t\tilde{e}_t\tilde{U}_{t,m}'\widehat{F}\right\|_F &= O_P(\Delta_{fe})\left(\sqrt{\frac{\log(pT)}{n}}+\Delta_{fum}\right)+O_P\left(\frac{\log p}{nT}+\sqrt{\frac{\log p}{Tn^2}}\right)\\ &= o_P\left(\sqrt{\frac{\log p}{nT}}\right),\end{aligned} \tag{H.31}$$

due to $\Delta_{fe}\Delta_{fum}\leq\Delta_F^2 = o\left(\sqrt{\frac{\log p}{nT}}\right)$ and the assumption that $\Delta_{fe}^2 = o\left(\frac{\log p}{T\log(pT)}\right)$. We omit the details for brevity. ∎

Lemma H.13. *(i)* $\|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T(\widehat{U}_{it}-\tilde{U}_{it})\tilde{\delta}_{yt}'H'^{-1}(H'\tilde{f}_i-\widehat{f}_i)\|_\infty = o_P(\sqrt{\frac{\log p}{nT}})$.
(ii) $\|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T(\widehat{U}_{it}-\tilde{U}_{it})(\widehat{U}_{it}-\tilde{U}_{it})'\gamma_y\|_\infty = o_P(\sqrt{\frac{\log p}{nT}})$.
(iii) $\|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T\tilde{U}_{it}(\tilde{U}_{it}-\widehat{U}_{it})'\gamma_y\|_\infty = o_P(\sqrt{\frac{\log p}{nT}})$.

Proof. (i) By Lemma H.11, $\max_m \|\frac{1}{T}\sum_{t=1}^T \frac{1}{n}\widehat{F}'\tilde{U}_{t,m}\tilde{\delta}_{yt}'\|_F = O_P\left(\Delta_{ud} + \sqrt{\frac{\log p}{nT}}\right)$. Hence,

$$\begin{aligned}
&\|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T(\widehat{U}_{it}-\tilde{U}_{it})\tilde{\delta}_{yt}'H'^{-1}(H'\tilde{f}_i-\widehat{f}_i)\|_\infty \\
&\leq \max_m \left\|\frac{1}{T}\sum_{t=1}^T(\widehat{\lambda}_{tm}-H^{-1}\tilde{\lambda}_{tm})\tilde{\delta}_{yt}'\right\|_F O_P(\Delta_F)+O_P(\Delta_F^2) \\
&\leq O_P(\Delta_F^2)+\max_m \left\|\frac{1}{T}\sum_{t=1}^T\frac{1}{n}\widehat{F}'\tilde{U}_{t,m}\tilde{\delta}_{yt}'\right\|_F O_P(\Delta_F) \qquad \text{(H.32)}\\
&= O_P(\Delta_F^2)+O_P(\Delta_F)\left(\Delta_{ud}+\sqrt{\frac{\log p}{nT}}\right) \\
&= o_P\left(\sqrt{\frac{\log p}{nT}}\right)
\end{aligned}$$

under the assumption $\Delta_F^2 + \Delta_F\Delta_{ud} = o\left(\sqrt{\frac{\log p}{nT}}\right)$.

(ii) Using the equalities

$$\begin{aligned}
\tilde{U}_{it,m}-\widehat{U}_{it,m} &= (\widehat{\lambda}_{tm}'-\tilde{\lambda}_{tm}'H'^{-1})\widehat{f}_i+\tilde{\lambda}_{tm}'H'^{-1}(\widehat{f}_i-H'\tilde{f}_i), \\
\widehat{\lambda}_{tm}-H^{-1}\tilde{\lambda}_{tm} &= (\widehat{F}'\widehat{F})^{-1}\widehat{F}'(\tilde{F}H-\widehat{F})H^{-1}\tilde{\lambda}_{tm}+(\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{U}_{t,m}, \text{and} \\
\widehat{\delta}_{yt}-H^{-1}\tilde{\delta}_{yt} &= (\widehat{F}'\widehat{F})^{-1}\widehat{F}'(\tilde{F}H-\widehat{F})H^{-1}\tilde{\delta}_{yt}+(\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{U}_t\gamma_y+(\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{e}_t,
\end{aligned}$$

we have

$$\begin{aligned}
&\|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T(\widehat{U}_{it}-\tilde{U}_{it})(\widehat{U}_{it}-\tilde{U}_{it})'\gamma_y\|_\infty \\
&= \max_m |\frac{1}{nT}\sum_{i=1}^n\sum_{k=1}^p\sum_{t=1}^T(\widehat{U}_{it,m}-\tilde{U}_{it,m})(\widehat{U}_{it,k}-\tilde{U}_{it,k})\gamma_{yk}| \\
&\leq \max_m \|\frac{1}{T}\sum_{k=1}^p\sum_{t=1}^T\tilde{\lambda}_{tm}(\widehat{\lambda}_{tk}-H^{-1}\tilde{\lambda}_{tk})'\gamma_{yk}\|_F O_P(\Delta_F) \\
&\quad+\max_m \|\frac{1}{T}\sum_{k=1}^p\sum_{t=1}^T(\widehat{\lambda}_{tm}-H^{-1}\tilde{\lambda}_{tm})\tilde{\lambda}_{tk}'\gamma_{yk}\|_F O_P(\Delta_F) \\
&\quad+\max_m \|\frac{1}{T}\sum_{k=1}^p\sum_{t=1}^T(\widehat{\lambda}_{tm}-H^{-1}\tilde{\lambda}_{tm})(\widehat{\lambda}_{tk}'-\tilde{\lambda}_{tk}'H'^{-1})\gamma_{yk}\|_F+O_P(\Delta_F^2|J|_0)
\end{aligned}$$

$$
\begin{aligned}
&\leq \left(\max_m \|\frac{1}{T}\sum_{k=1}^p\sum_{t=1}^T \tilde{\lambda}_{tm}\frac{1}{n}\widehat{F}'\tilde{U}_{t,k}\gamma_{yk}\|_F + \max_m \|\frac{1}{T}\sum_{k=1}^p\sum_{t=1}^T \frac{1}{n}\widehat{F}'\tilde{U}_{t,m}\tilde{\lambda}_{tk}'\gamma_{yk}\|_F\right) O_P(\Delta_F) \\
&\quad + \max_m \|\frac{1}{T}\sum_{k=1}^p\sum_{t=1}^T \frac{1}{n}\widehat{F}'\tilde{U}_{t,m}\frac{1}{n}\widehat{F}'\tilde{U}_{t,k}\gamma_{yk}\|_F + O_P(\Delta_F^2|J|_0) \\
&= \left(|J|_0\sqrt{\frac{\log p}{nT}} + \Delta_F|J|_0\right) O_P(\Delta_F) + o_P\left(\sqrt{\frac{\log p}{nT}}\right) + O_P(\Delta_F^2|J|_0) \\
&= o_P\left(\sqrt{\frac{\log p}{nT}}\right),
\end{aligned}
$$

where the bound on the first term in the second to last equality follows from Lemma H.11 (ii)(v) and the bound on the second term in the second to last equality follows from Lemma H.12. Finally, the last equality follows from the assumption that $|J|_0\Delta_F^2 = o\left(\sqrt{\frac{\log p}{nT}}\right)$.

(iii) From the equality

$$
\begin{aligned}
\tilde{U}_{it,m} - \widehat{U}_{it,m} = \widehat{f}_i'(\widehat{F}'\widehat{F})^{-1}\widehat{F}'(\tilde{F}H - \widehat{F})H^{-1}\tilde{\lambda}_{tm} + \widehat{f}_i'(\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{U}_{t,m} \\
+ \tilde{\lambda}_{tm}'H'^{-1}(\widehat{f}_i - H'\tilde{f}_i),
\end{aligned}
$$

we have

$$
\begin{aligned}
&\|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T \tilde{U}_{it}(\tilde{U}_{it} - \widehat{U}_{it})'\gamma_y\|_\infty \\
&\quad \leq \max_m \left|\frac{1}{nT}\sum_{i=1}^n\sum_{k=1}^p\sum_{t=1}^T \tilde{U}_{it,m}\tilde{\lambda}_{tk}'H'^{-1}(\widehat{f}_i - H'\tilde{f}_i)\gamma_{yk}\right| \\
&\quad\quad + \max_m \left|\frac{1}{nT}\sum_{i=1}^n\sum_{k=1}^p\sum_{t=1}^T \tilde{U}_{it,m}\tilde{\lambda}_{tk}\widehat{f}_i\gamma_{yk}\right| O_P(\Delta_F) \\
&\quad\quad + \max_m \left|\frac{1}{nT}\sum_{i=1}^n\sum_{k=1}^p\sum_{t=1}^T \tilde{U}_{it,m}\frac{1}{n}\widehat{F}'\tilde{U}_{t,k}\widehat{f}_i\gamma_{yk}\right| \\
&\quad = o_P\left(\sqrt{\frac{\log p}{nT}}\right).
\end{aligned}
$$

■

H.4. **Technical lemmas for Appendix B.**

Lemma H.14. *(i)* $\max_m \|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T \tilde{e}^*_{it}\widehat{f}_i\widehat{\lambda}'_{tm}\|_F = O_{P^*}\left(\sqrt{\frac{\log(p)\log(np)}{nT}}\right)$.
(ii) $\max_{mj}\frac{1}{n}\sum_i(\frac{1}{T}\sum_t\widehat{\eta}_{it}\widehat{f}_j\widehat{U}_{jt,m}\widehat{f}_i)^2 = O_P\left(\frac{\log^2(n)\log(pn)}{T}\right)$.
(iii) $\max_m \|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T \tilde{e}^*_{it}\frac{1}{n}\widehat{F}'\tilde{U}^*_{t,m}\widehat{f}_i\|_F = o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right)$.
(iv) $\max_{mt}\|\frac{1}{n}\widehat{F}'\tilde{U}^*_{t,m}\|_2 = O_{P^*}\left(\sqrt{\frac{\log(pT)}{n}}\right)$, $\max_m\frac{1}{T}\sum_t\|\frac{1}{n}\widehat{F}'\tilde{U}^*_{t,m}\|_2^2 = O_{P^*}\left(\sqrt{\frac{\log(pT)}{n}}\right)$, *and* $\frac{1}{T}\sum_t\|\frac{1}{n}\sum_{i=1}^n\tilde{e}^*_{it}\widehat{f}_i\|_2^2 = O_{P^*}\left(\frac{1}{n}\right)$.
(v) $\max_m \|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T \tilde{e}^*_{it}\frac{1}{n}\widehat{F}^{*\prime}\tilde{U}^*_{t,m}\widehat{f}^{*\prime}_i\|_F = o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right)$.

Proof. For notational simplicity, we assume f_i is one dimensional without loss of generality.

(i) Note $\tilde{e}^*_{it} = \tilde{\epsilon}^*_{it} + \widehat{\alpha}\tilde{\eta}^*_{it} = w_i^Y\widehat{\epsilon}_{it} + \widehat{\alpha}w_i^D\widehat{\eta}_{it}$. We bound $\max_m|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T w_i^D\widehat{\eta}_{it}\widehat{f}_i\widehat{\lambda}_{tm}|$. Let $\Gamma_{im} = \frac{1}{T}\sum_{t=1}^T w_i^D\widehat{\eta}_{it}\widehat{f}_i\widehat{\lambda}_{tm}$. Then, since $\mathrm{Var}(w_i^D) = 1$, uniformly in $m \leq p$,

$$
\begin{aligned}
\frac{1}{n}\sum_i \mathrm{Var}^*(\Gamma_{im}) &= \frac{1}{n}\sum_i\left(\frac{1}{T}\sum_t\widehat{\eta}_{it}\widehat{f}_i\widehat{\lambda}_{tm}\right)^2 \\
&\leq O_P(1)\frac{1}{n}\sum_i\left(\frac{1}{T}\sum_t\widehat{\eta}_{it}\widehat{\lambda}_{tm}\right)^2 \\
&\leq O_P\left(\frac{1}{n}\right)\sum_i\left(\frac{1}{T}\sum_t\tilde{\eta}_{it}\tilde{\lambda}_{tm}\right)^2 \\
&\quad + O_P(1)\frac{1}{n}\sum_i\left(\frac{1}{T}\sum_t(\widehat{\eta}_{it}-\tilde{\eta}_{it})\widehat{\lambda}_{tm}\right)^2 \\
&\quad + O_P(1)\frac{1}{n}\sum_i\left(\frac{1}{T}\sum_t\tilde{\eta}_{it}(\widehat{\lambda}_{tm}-H^{-1}\tilde{\lambda}_{tm})\right)^2 \\
&\leq^{(a)} O_P(\psi_n^2)\max_m\frac{1}{T}\sum_t\widehat{\lambda}^2_{tm} \\
&\quad + O_P(1)\max_m\frac{1}{n}\sum_i\left(\frac{1}{T}\sum_t\tilde{\eta}_{it}(\widehat{F}'\widehat{F})^{-1}\widehat{F}'(\tilde{F}H-\widehat{F})H^{-1}\tilde{\lambda}_{tm}\right)^2 \\
&\quad + O_P(1)\max_m\frac{1}{n}\sum_i\left(\frac{1}{T}\sum_t\tilde{\eta}_{it}(\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{U}_{t,m}\right)^2
\end{aligned}
$$

$$
\begin{aligned}
&+ O_P(1) \max_m \frac{1}{n} \sum_i \left(\frac{1}{T} \sum_t \tilde{\eta}_{it} \tilde{\lambda}_{tm} \right)^2 \\
\leq^{(b)}\ & O_P(\psi_n^2) + \max_m \frac{1}{n} \sum_i \left(\frac{1}{T} \sum_t \tilde{\eta}_{it} \tilde{\lambda}_{tm} \right)^2 O_P(\Delta_F^2 + 1) \\
&+ O_P(1) \max_m \frac{1}{n} \sum_i \left(\frac{1}{Tn} \sum_t \tilde{\eta}_{it} \widehat{F}' \tilde{U}_{t,m} \right)^2 \\
\leq\ & O_P(\psi_n^2) + \max_{im} \left(\frac{1}{T} \sum_t \tilde{\eta}_{it} \tilde{\lambda}_{tm} \right)^2 O_P(1) + \frac{1}{n^2} \|\widehat{F}\|_F^2 \max_{im} \left\| \frac{1}{T} \sum_t \tilde{\eta}_{it} \tilde{U}_{t,m} \right\|_2^2 \\
=\ & O_P\left(\psi_n^2 + \frac{\log(np)}{T} \right) \\
=^{(c)}\ & O_P\left(\frac{\log(np)}{T} \right)
\end{aligned}
$$

where ψ_n is defined in (A.6). In (a), we used the equality

$$
\widehat{\lambda}_{tm} - H^{-1}\tilde{\lambda}_{tm} = (\widehat{F}'\widehat{F})^{-1}\widehat{F}'(\tilde{F}H - \widehat{F})H^{-1}\tilde{\lambda}_{tm} + (\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{U}_{t,m}
$$

and Proposition A.1. In (b), we used

$$
\begin{aligned}
\max_{mt} |\widehat{\lambda}_{tm}| &\leq O_P(1) \max_{mt} |\tilde{\lambda}_{tm}| + \max_{mt} |\widehat{\lambda}_{tm} - H^{-1}\tilde{\lambda}_{tm}| \\
&\leq O_P(1) + \max_{mt} |\frac{1}{n}\widehat{F}'\tilde{U}_{t,m}| = O_P(1).
\end{aligned}
\tag{H.33}
$$

which follows from Lemma H.11. Equality (c) follows from $\psi_n^2 = O\left(\frac{1}{T}\right)$, given that $\Delta_F^2 |J|_0^2 = o\left(\frac{1}{\sqrt{nT}}\right) = o\left(\frac{1}{T}\right)$.

Recall that $\Gamma_{im} = \frac{1}{T}\sum_{t=1}^T w_i^D \widehat{\eta}_{it} \widehat{f}_i \widehat{\lambda}_{tm}$. It then follows that, for any arbitrarily small $\varepsilon > 0$, there is a $C_\varepsilon > 0$ which depends only on the true data generating process and not on any realized data such that the event $A = \left\{ \max_m \frac{1}{n}\sum_i \mathrm{Var}^*(\Gamma_{im}) \leq C_\varepsilon \frac{\log(np)}{T} \right\}$ holds with probability at least $1-\varepsilon$. On this event, by the Bernstein inequality for independent data, for $x = \sqrt{\frac{2C_\varepsilon \log(p)\log(np)}{nT}}$,

$$
P^*\left(\max_m |\frac{1}{n}\sum_i \Gamma_{im}| > x \right) 1\{A\} \leq \exp\left(\log p - \frac{nTx^2}{C_\varepsilon \log(np)} \right) = p^{-1}.
$$

Thus, $P^*(\max_m |\frac{1}{n}\sum_i \Gamma_{im}| > x) \leq p^{-1} + 1\{A^c\}$ where P^* is with respect to the bootstrap sampling space conditional on the original data. Taking expectations on both sides, we reach $P\left(\max_m \left|\frac{1}{n}\sum_i \Gamma_{im}\right| > x\right) \leq p^{-1} + \varepsilon$ where P is with respect to the probability space of the joint distribution of $\{w_i^D, \text{original data}\}$. Now for any $\varepsilon_1, \varepsilon_2 > 0$, take $\varepsilon = \varepsilon_1\varepsilon_2/2$. Then, as $p \to \infty$,

$$P\left(P^*\left(\max_m |\frac{1}{n}\sum_i \Gamma_{im}| > x\right) > \varepsilon_1\right) < (p^{-1}+\varepsilon)/\varepsilon_1 < \varepsilon_2.$$

Since C_ε is not random in either the bootstrap or the original sampling space, we have

$$\max_m \left|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T w_i^D \widehat{\eta}_{it}\widehat{f}_i\widehat{\lambda}_{tm}\right| = O_{P^*}\left(\sqrt{\frac{\log(p)\log(np)}{nT}}\right).$$

The same argument applies to

$$\max_m \left|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T w_i^Y \widehat{\epsilon}_{it}\widehat{f}_i\widehat{\lambda}_{tm}\right|$$

and thus

$$\max_m \left|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T \tilde{e}_{it}^*\widehat{f}_i\widehat{\lambda}_{tm}\right|$$

as well. The conclusion follows.

(ii) First, by Lemma H.11 and Assumption D.4,

$$\max_{m\le p}\left(\frac{1}{n^2T}\sum_t \|\tilde{U}_{t,m}'\widehat{F}\|_2^2\right) = O_P\left(\frac{\log(pT)}{n} + \Delta_{fum}^2\right) = O_P\left(\Delta_F^2 + \frac{\log(pT)}{n}\right).$$

Hence,

$$\begin{aligned}
\max_{mi}\frac{1}{T}\sum_t(\tilde{U}_{it,m}-\widehat{U}_{it,m})^2 &\le \max_{mi}\frac{1}{T}\sum_t(\widehat{f}_i'(\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{U}_{t,m})^2 \\
&\quad + \max_{mi}\frac{1}{T}\sum_t(\widehat{f}_i'(\widehat{F}'\widehat{F})^{-1}\widehat{F}'(\tilde{F}H-\widehat{F})H^{-1}\tilde{\lambda}_{tm})^2 \\
&\quad + \max_{mi}\frac{1}{T}\sum_t(\tilde{\lambda}_{tm}'H'^{-1}(\widehat{f}_i - H'\tilde{f}_i))^2 \\
&\le O_P(\log n)\max_m\frac{1}{T}\sum_t\left(\frac{1}{n}\widehat{F}'\tilde{U}_{t,m}\right)^2 + O_P(\log(n)\Delta_F^2+\Delta_{\max}^2) \\
&= O_P\left(\log(n)\frac{\log(pT)}{n} + \log(n)\Delta_F^2 + \Delta_{\max}^2\right) = O_P(\log n)
\end{aligned} \tag{H.34}$$

under the assumption $\Delta_{\max} = O\left(\sqrt{\log n}\right)$. Then we have

$$
\begin{aligned}
&\max_{mj} \frac{1}{n}\sum_i \left(\frac{1}{T}\sum_t \widehat{\eta}_{it}\widehat{f}_j\widehat{U}_{jt,m}\widehat{f}_i\right)^2 \\
&\quad\leq \max_{mj} \frac{1}{n}\sum_i \left(\frac{1}{T}\sum_t \widehat{\eta}_{it}\widehat{U}_{jt,m}\right)^2 \widehat{f}_i^2 \max_j \widehat{f}_j^2 \\
&\quad\leq O_P(\log n) \max_{mj} \frac{1}{n}\sum_i \left(\frac{1}{T}\sum_t \widehat{\eta}_{it}\widehat{U}_{jt,m}\right)^2 \widehat{f}_i^2 \\
&\quad\leq O_P(\log n) \max_{mj} \frac{1}{n}\sum_i \left(\frac{1}{T}\sum_t (\widehat{\eta}_{it} - \tilde{\eta}_{it})\widehat{U}_{jt,m}\right)^2 \widehat{f}_i^2 \\
&\qquad + O_P(\log n)\frac{1}{n}\sum_i \widehat{f}_i^2 \max_{mj} \left(\frac{1}{T}\sum_t \tilde{\eta}_{it}\tilde{U}_{jt,m}\right)^2 \\
&\qquad + O_P(\log n) \max_{mj} \frac{1}{n}\sum_i \left(\frac{1}{T}\sum_t \tilde{\eta}_{it}(\widehat{U}_{jt,m} - \tilde{U}_{jt,m})\right)^2 \widehat{f}_i^2 \\
&\quad\leq O_P(\log^2 n)\frac{1}{nT}\sum_{it}(\widehat{\eta}_{it} - \tilde{\eta}_{it})^2 \max_{mj} \frac{1}{T}\sum_s \widehat{U}_{js,m}^2 + O_P(\log n) \max_{mij} \left(\frac{1}{T}\sum_t \tilde{\eta}_{it}\tilde{U}_{jt,m}\right)^2 \\
&\qquad + O_P(\log n) \max_{mj} \frac{1}{n}\sum_i \left(\frac{1}{T}\sum_t \tilde{\eta}_{it}\widehat{f}_j'(\widehat{F}'\widehat{F})^{-1}\widehat{F}'(\tilde{F}H - \widehat{F})H^{-1}\tilde{\lambda}_{tm}\right)^2 \widehat{f}_i^2 \\
&\qquad + O_P(\log n) \max_{mj} \frac{1}{n}\sum_i \left(\frac{1}{T}\sum_t \tilde{\eta}_{it}\widehat{f}_j'(\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{U}_{t,m}\right)^2 \widehat{f}_i^2 \\
&\qquad + O_P(\log n) \max_{mj} \frac{1}{n}\sum_i \left(\frac{1}{T}\sum_t \tilde{\eta}_{it}\tilde{\lambda}_{tm}'H'^{-1}(\widehat{f}_j - H'\tilde{f}_j)\right)^2 \widehat{f}_i^2 \\
&\quad\leq O_P(\log^2(n)\psi_n^2)\left(\max_{mj} \frac{1}{T}\sum_s \tilde{U}_{js,m}^2 + \log n\right) + O_P\left(\log(n)\frac{\log(pn)}{T}\right) \\
&\qquad + O_P(\log^2(n)\Delta_F^2) \max_{mi} \left(\frac{1}{T}\sum_t \tilde{\eta}_{it}\tilde{\lambda}_{tm}\right)^2 \\
&\qquad + O_P(\log^2 n) \max_i \frac{1}{T}\sum_t |\tilde{\eta}_{it}|^2 \max_{tm} \left|\frac{1}{n}\widehat{F}'\tilde{U}_{t,m}\right|^2
\end{aligned}
$$

$$+ O_P(\log(n)\Delta_{\max}^2) \max_{mi} \left(\frac{1}{T} \sum_t \tilde{\eta}_{it} \tilde{\lambda}_{tm} \right)^2$$
$$\leq O_P \left(\frac{\log^2(n)\log(pn)}{T} \right),$$

where we used

$$\max_{t \leq T, m \leq p} \left\| \frac{1}{n} \sum_{i=1}^n \widehat{f}_i \tilde{U}_{it,m} \right\|_2 = O_P \left(\sqrt{\frac{\log(pT)}{n}} + \Delta_{fum} \right)$$

and

$$\max_i \frac{1}{T} \sum_t \tilde{\eta}_{it}^2 = O_P \left(\max_{it} E\eta_{it}^2 + \max_i \frac{1}{T} \sum_t (\eta_{it}^2 - E\eta_{it}^2) \right)$$

from Lemma H.11, (A.6) for the definition of ψ_n, $|J|_0^2 \Delta_F^2 = o\left((nT)^{-1/2}\right)$ from Assumption D.4, and $T = o(n)$.

(iii) We have

$$\frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T \tilde{e}_{it}^* \frac{1}{n} \widehat{F}' \tilde{U}_{t,m}^* \widehat{f}_i' = \frac{1}{n^2 T} \sum_{i,j \leq n} \sum_{t=1}^T w_i^U (\tilde{\epsilon}_{it}^* + \widehat{\alpha} \tilde{\eta}_{it}^*) \widehat{f}_j \widehat{U}_{jt,m} \widehat{f}_i'.$$

We bound

$$\max_m \left| \frac{1}{n^2 T} \sum_{i,j \leq n} \sum_{t=1}^T w_j^U \tilde{\eta}_{it}^* \widehat{f}_j \widehat{U}_{jt,m} \widehat{f}_i \right| = \max_m \left| \frac{1}{n^2 T} \sum_{i,j \leq n} \sum_{t=1}^T w_j^U w_i^D \widehat{\eta}_{it} \widehat{f}_j \widehat{U}_{jt,m} \widehat{f}_i \right|$$

using Lemma H.5.

We first rewrite

$$\max_m \left| \frac{1}{n^2 T} \sum_{i,j \leq n} \sum_{t=1}^T w_j^U w_i^D \widehat{\eta}_{it} \widehat{f}_j \widehat{U}_{jt,m} \widehat{f}_i \right|$$

as

$$\max_m \left| \frac{1}{n^2 T} \sum_{i,j \leq n} \sum_{t=1}^T w_i^U w_j^D \widehat{\eta}_{jt} \widehat{f}_i \widehat{U}_{it,m} \widehat{f}_j \right|$$

so that it looks more friendly for application of Lemma H.5. Now define $Z_{ijm} = \frac{1}{T} \sum_t \widehat{\eta}_{jt} \widehat{f}_i \widehat{U}_{it,m} \widehat{f}_j$. Then by part (ii) of the present lemma,

$$\max_{m \leq p, i \leq n} \frac{1}{n} \sum_{j=1}^n Z_{ijm}^2 = \max_{mi} \frac{1}{n} \sum_j \left(\frac{1}{T} \sum_t \widehat{\eta}_{jt} \widehat{f}_i \widehat{U}_{it,m} \widehat{f}_j \right)^2$$
$$= O_P \left(\frac{\log^2(n)\log(pn)}{T} \right) := O_P(a_n^2).$$

Now define $X_i^* = w_i^U$ and $Y_j^* = w_j^D$. Apply Lemma H.5,

$$\max_m \left| \frac{1}{n^2 T} \sum_{i,j \le n} \sum_{t=1}^T w_i^U w_j^D \widehat{\eta}_{jt} \widehat{f}_i \widehat{U}_{it,m} \widehat{f}_j \right| = \max_m \left| \frac{1}{n^2} \sum_{ij} X_i^* Y_j^* Z_{ijm} \right| = O_{P^*}\left(\frac{\log(n)\log(pn)}{\sqrt{n}} \sqrt{\frac{\log p}{nT}} \right).$$

Then, using the assumption that $\log^3 p = O(n)$, we have

$$\max_m |\frac{1}{n^2 T} \sum_{i,j \le n} \sum_{t=1}^T w_i^U w_j^D \widehat{\eta}_{jt} \widehat{f}_i \widehat{U}_{it,m} \widehat{f}_j| = o_{P^*}\left(\sqrt{\frac{\log p}{nT}} \right).$$

Similarly,

$$\max_m \left| \frac{1}{n^2 T} \sum_{i,j \le n} \sum_{t=1}^T w_j^U \tilde{\epsilon}_{it}^* \widehat{f}_j \widehat{U}_{jt,m} \widehat{f}_i \right| = o_{P^*}\left(\sqrt{\frac{\log p}{nT}} \right).$$

Therefore,

$$\max_m \left| \frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T \tilde{e}_{it}^* \frac{1}{n} \widehat{F}' \tilde{U}_{t,m}^* \widehat{f}_i' \right| = o_{P^*}\left(\sqrt{\frac{\log p}{nT}} \right).$$

(iv) First, we bound $\max_{tm} \frac{1}{n} \sum_i |\widehat{U}_{it,m} - \tilde{U}_{it,m}|^g$, for some $g \ge 2$. Again, we use the following equality

$$\tilde{U}_{it,m} - \widehat{U}_{it,m} = \widehat{f}_i'(\widehat{F}'\widehat{F})^{-1}\widehat{F}'(\tilde{F}H - \widehat{F})H^{-1}\tilde{\lambda}_{tm} + \widehat{f}_i'(\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{U}_{t,m} + \tilde{\lambda}_{tm}' H'^{-1}(\widehat{f}_i - H'\tilde{f}_i).$$

Note that

$$\frac{1}{n} \sum_i |\widehat{f}_i|^g \le \max_i |\widehat{f}_i|^{g-2} \frac{1}{n} \sum_i |\widehat{f}_i|^2 = O_P((\log n)^{g-2}),$$

$$\frac{1}{n} \sum_i |\widehat{f}_i - H'\tilde{f}_i|^g \le \max_i |\widehat{f}_i - H'\tilde{f}_i|^{g-2} \frac{1}{n} \sum_i |\widehat{f}_i - H'\tilde{f}_i|^2 = O_P(\Delta_F^2 \Delta_{\max}^{g-2}),$$

and

$$\max_{tm} |\frac{1}{n} \widehat{F}' \tilde{U}_{t,m}|^g = (\max_{tm} |\frac{1}{n} \widehat{F}' \tilde{U}_{t,m}|^2)^{g/2} = O_P((\frac{\log(pT)}{n})^{g/2} + \Delta_F^g).$$

Thus,

$$\begin{aligned}
\max_{tm}\frac{1}{n}\sum_i(\widehat{U}_{it,m}-\tilde{U}_{it,m})^g &\leq \max_{tm}|\tilde{\lambda}_{tm}|^g\frac{1}{n}\sum_i|\widehat{f}_i|^g O_P(\Delta_F^g)\\
&\quad + O_P(1)\frac{1}{n}\sum_i|\widehat{f}_i|^g\max_{tm}|\frac{1}{n}\widehat{F}'\tilde{U}_{t,m}|^g\\
&\quad + \max_{tm}|\tilde{\lambda}_{tm}|^g\frac{1}{n}\sum_i|\widehat{f}_i-H'\tilde{f}_i|^g O_P(1)\\
&\leq O_P\left(\Delta_F^g(\log n)^{g-2}+(\frac{\log(pT)}{n})^{g/2}(\log n)^{g-2}+\Delta_F^2\Delta_{\max}^{g-2}\right).
\end{aligned} \tag{H.35}$$

When $g=2$, $\max_{tm}\frac{1}{n}\sum_i(\widehat{U}_{it,m}-\tilde{U}_{it,m})^2=o_P(1)$, yielding $\max_{tm}\frac{1}{n}\sum_i\widehat{U}_{it,m}^2\leq O_P(1)$. Hence, uniformly in $m\leq p, t\leq T$,

$$\begin{aligned}
\frac{1}{n}\sum_i\mathrm{Var}^*(w_i^U\widehat{f}_i\widehat{U}_{it,m}) &= \frac{1}{n}\sum_i\widehat{f}_i^2\widehat{U}_{it,m}^2\\
&\leq \frac{2}{n}\sum_i\widehat{f}_i^2\tilde{U}_{it,m}^2+\frac{2}{n}\sum_i\widehat{f}_i^2(\widehat{U}_{it,m}-\tilde{U}_{it,m})^2\\
&\leq \frac{4}{n}\sum_i(\widehat{f}_i-H'\tilde{f}_i)^2\tilde{U}_{it,m}^2+O_P(1)\frac{1}{n}\sum_i\tilde{f}_i^2\tilde{U}_{it,m}^2\\
&\quad +\frac{1}{n}\sum_i\widehat{f}_i^4\tilde{\lambda}_{tm}^2O_P(\Delta_F^2)+\frac{1}{n}\sum_i\widehat{f}_i^4\left(\frac{1}{n}\widehat{F}'\tilde{U}_{t,m}\right)^2\\
&\quad +\frac{1}{n}\sum_i\widehat{f}_i^2(\tilde{\lambda}_{tm}'H'^{-1}(\widehat{f}_i-H'\tilde{f}_i))^2\\
&\leq O_P(\Delta_F^2)\max_{itm}\tilde{U}_{it,m}^2+O_P(1)+O_P(\Delta_F^2\log n)\\
&\quad +O_P(\log n)\left(\Delta_{fum}^2+\frac{\log(pT)}{n}\right)\\
&= O_P\left(\Delta_F^2\log(nTp)+1\right)=O_P(1).
\end{aligned}$$

It then follows from the union bound and Bernstein inequality for independent data that

$$\begin{aligned}
\max_{mt}\|\frac{1}{n}\widehat{F}'\tilde{U}_{t,m}^*\|_2 &= \max_{mt}\|\frac{1}{n}\sum_i w_i^U\widehat{f}_i\widehat{U}_{it,m}\|_2\\
&= O_{P^*}\left(\sqrt{\max_{tm}\frac{1}{n}\sum_i\mathrm{Var}^*(w_i^U\widehat{f}_i\widehat{U}_{it,m})\frac{\log(pT)}{n}}\right)
\end{aligned}$$

$$= O_{P^*}\left(\sqrt{\frac{\log(pT)}{n}}\right).$$

Next, note

$$\frac{1}{T}\sum_t \|\frac{1}{n}\sum_{i=1}^n \tilde{e}_{it}^* \widehat{f}_i\|_2^2 = \frac{1}{T}\sum_t \|\frac{1}{n}\sum_{i=1}^n (\tilde{\epsilon}_{it}^* + \widehat{\alpha}\tilde{\eta}_{it}^*)\widehat{f}_i\|_2^2.$$

Here we just bound $\frac{1}{T}\sum_t \|\frac{1}{n}\sum_{i=1}^n \tilde{\eta}_{it}^* \widehat{f}_i\|_2^2$. Note that bounding $\frac{1}{T}\sum_t \|\frac{1}{n}\sum_{i=1}^n \tilde{\epsilon}_{it}^* \widehat{f}_i\|_2^2$ follows from the same argument. Recall that we have taken f_i to be a scalar for simplicity. We then have

$$\begin{aligned}
E^*\frac{1}{T}\sum_t \left(\frac{1}{n}\sum_{i=1}^n \tilde{\eta}_{it}^* \widehat{f}_i\right)^2 &= \frac{1}{T}\sum_t \mathrm{Var}^*\left(\frac{1}{n}\sum_{i=1}^n w_i^D \widehat{\eta}_{it}\widehat{f}_i\right) \\
&= \frac{1}{Tn^2}\sum_{it}(\widehat{\eta}_{it}\widehat{f}_i)^2 \\
&\leq \frac{2}{Tn^2}\sum_{it}(\tilde{\eta}_{it}\widehat{f}_i)^2 + \frac{2}{Tn^2}\sum_{it}(\widehat{\eta}_{it}-\tilde{\eta}_{it})^2 \max_i \widehat{f}_i^2 \\
&\leq \frac{2}{Tn^2}\sum_{it}(\tilde{\eta}_{it}\tilde{f}_i)^2 O_P(1) + O_P\left(\frac{1}{n}\Delta_F^2\right)\max_i \frac{1}{T}\sum_t \tilde{\eta}_{it}^2 + O_P\left(\frac{\log n}{n}\psi_n^2\right) \\
&= O_P\left(\frac{1}{n} + \frac{\Delta_F^2}{n}\left(1+\sqrt{\frac{\log n}{T}}\right) + \frac{\log n}{n}\psi_n^2\right) \\
&= O_P\left(\frac{1}{n}\right),
\end{aligned}$$

where we used $\log n\psi_n^2 = o(1)$ from (A.6).

(v) We have $\max_m \|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T \tilde{e}_{it}^* \frac{1}{n}\widehat{F}^{*\prime}\tilde{U}_{t,m}^*\widehat{f}_i^{*\prime}\|_F \leq \sum_{l=1}^4 A_l$. Each A_l is defined and bounded below.

$$\begin{aligned}
A_1 &= \max_m \left\|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T \tilde{e}_{it}^*\frac{1}{n}(\widehat{F}^* - \widehat{F}H^*)'\tilde{U}_{t,m}^*(\widehat{f}_i^* - H^{*\prime}\widehat{f}_i)'\right\|_F \\
&\leq \frac{1}{T}\sum_{t=1}^T \left\|\frac{1}{n}\sum_{i=1}^n \tilde{e}_{it}^*(\widehat{f}_i^* - H^{*\prime}\widehat{f}_i)\right\|_2 \max_{tm}\left\|\frac{1}{n}(\widehat{F}^* - \widehat{F}H^*)'\tilde{U}_{t,m}^*\right\|_2 \\
&= O_{P^*}(\Delta_{fe}^*\Delta_{fum}^*) =^{(a)} o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right).
\end{aligned}$$

$$
\begin{aligned}
A_2 &= \max_m \left\| \frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T \tilde{e}^*_{it} \frac{1}{n} \widehat{F}' \tilde{U}^*_{t,m} \widehat{f}_i' \right\|_F \\
&= o_{P^*}\left(\sqrt{\frac{\log p}{nT}} \right) \text{ by part (iii).} \\
A_3 &= \max_m \left\| \frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T \tilde{e}^*_{it} \frac{1}{n} \widehat{F}' \tilde{U}^*_{t,m} (\widehat{f}^*_i - H^{*\prime} \widehat{f}_i)' \right\|_F \\
&\leq \max_{mt} \left\| \frac{1}{n} \widehat{F}' \tilde{U}^*_{t,m} \right\|_2 \frac{1}{T} \sum_{t=1}^T \left\| \frac{1}{n} \sum_{i=1}^n \tilde{e}^*_{it} (\widehat{f}^*_i - H^{*\prime} \widehat{f}_i)' \right\|_F \\
&\leq \max_{mt} \left\| \frac{1}{n} \widehat{F}' \tilde{U}^*_{t,m} \right\|_2 O_{P^*}\left(\Delta^*_{fe} \right) \\
&= O_{P^*}\left(\sqrt{\frac{\log(pT)}{n}} \Delta^*_{fe} \right) =^{(b)} o_{P^*}\left(\sqrt{\frac{\log p}{nT}} \right) \\
A_4 &= \max_m \left\| \frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T \frac{1}{n} (\widehat{F} H^* - \widehat{F}^*)' \tilde{U}^*_{t,m} \widehat{f}_i' \tilde{e}^*_{it} \right\|_F \\
&\leq O_{P^*}(\Delta^*_{fum}) \left(\frac{1}{T} \sum_t \| \frac{1}{n} \sum_{i=1}^n \tilde{e}^*_{it} \widehat{f}_i \|_2^2 \right)^{1/2} \\
&\leq^{(c)} O_{P^*}\left(\Delta^*_{fum} \sqrt{\frac{1}{n}} \right) =^{(d)} o_{P^*}\left(\sqrt{\frac{\log p}{nT}} \right).
\end{aligned}
$$

In the above derivations, Δ^*_{fe} is taken such that

$$
\frac{1}{T} \sum_{t=1}^T \left\| \frac{1}{n} \sum_{i=1}^n \tilde{e}^*_{it} (\widehat{f}^*_i - H^{*\prime} \widehat{f}_i) \right\|_2^2 = O_{P^*}(\Delta^{*2}_{fe}).
$$

Also, the sequence Δ^*_{fum} is taken such that

$$
\max_{tm} \left\| \frac{1}{n} (\widehat{F}^* - \widehat{F} H^*)' \tilde{U}^*_{t,m} \right\|_2^2 = O_{P^*}(\Delta^{*2}_{fum}).
$$

Then (c) follows from part (iv) of the present lemma. We now prove (b) by showing that $\Delta^*_{fe} \leq \Delta^*_F$, then (b) follows from Assumption D.5 that $\Delta^{*2}_F = o\left(\frac{\log p}{T(\log pT)}\right)$. In addition, we prove (d) that $\Delta^{*2}_{fum} \leq \Delta^{*2}_F = o\left(\frac{\log p}{T}\right)$ and (a) that $\Delta^*_{fe}\Delta^*_{fum} = o\left(\sqrt{\frac{\log p}{nT}}\right)$ in the following.

Show $\mathbf{\Delta^*_{fe}} \leq \mathbf{\Delta^*_F}$: By Cauchy-Schwarz,

$$\frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{n}\sum_{i=1}^{n}\tilde{e}^*_{it}(\widehat{f}^*_i - H^{*'}\widehat{f}_i)\right\|_2^2 \leq \frac{1}{nT}\sum_{it}\tilde{e}^{*2}_{it}\frac{1}{n}\sum_{i=1}^{n}\left\|\widehat{f}^*_i - H^{*'}\widehat{f}_i\right\|_2^2 = O_{P^*}(\Delta^{*2}_F).$$

Hence, we can always take Δ^*_{fe} so that $\Delta^*_{fe} \leq \Delta^*_F$.

Show $\mathbf{\Delta^*_{fum}} \leq \mathbf{\Delta^*_F}$: Let $X^*_{i,tm} = \widehat{U}^2_{it,m}(w^U_i)^2$. For a generic $C > 0$, by (H.35) with $g = 4$, we have

$$\begin{aligned}
\max_{tm}\frac{1}{n}\sum_i \mathrm{Var}^*(X^*_{i,tm}) &\leq \max_{tm} C\frac{1}{n}\sum_i \widehat{U}^4_{it,m} \\
&\leq C\max_{tm}\frac{1}{n}\sum_i(\widehat{U}_{it,m} - \tilde{U}_{it,m})^4 + C\max_{tm}\frac{1}{n}\sum_i \tilde{U}^4_{it,m} \\
&\leq O_P(1) + O_P\left(\Delta^4_F(\log n)^2 + (\frac{\log(pT)}{n})^2(\log n)^2 + \Delta^2_F\Delta^2_{\max}\right) = O_P(1).
\end{aligned}$$

Applying a Bernstein inequality for independent data and union bounds and noting that $E^*X^*_{i,tm} = \widehat{U}^2_{it,m}$, we have

$$\begin{aligned}
\max_{tm}\frac{1}{n}\sum_i X^*_{i,tm} &\leq \max_{tm}\left|\frac{1}{n}\sum_i(X^*_{i,tm} - E^*X^*_{i,tm})\right| + \max_{tm}\frac{1}{n}\sum_i \widehat{U}^2_{it,m} \\
&= O_{P^*}\left(\sqrt{\frac{\log(pT)}{n}} + 1\right) = O_{P^*}(1).
\end{aligned}$$

Now recall that Δ^*_{fum} is taken to satisfy

$$\max_{tm}\|\frac{1}{n}\sum_i(\widehat{f}^*_i - H^{*'}\widehat{f}_i)\tilde{U}^*_{it,m}\|_2^2 = O_{P^*}(\Delta^{*2}_{fum}).$$

By Cauchy-Schwarz, we have

$$\max_{tm}\left\|\frac{1}{n}\sum_i(\widehat{f}^*_i - H^{*'}\widehat{f}_i)\tilde{U}^*_{it,m}\right\|_2^2 \leq O_{P^*}(\Delta^{*2}_F)\max_{tm}\frac{1}{n}\sum_i \tilde{U}^{*2}_{it,m} \tag{H.36}$$

$$= O_{P^*}(\Delta^{*2}_F)\max_{tm}\frac{1}{n}\sum_i X^*_{i,tm} = O_{P^*}(\Delta^{*2}_F). \tag{H.37}$$

Hence, we can take Δ^*_{fum} so that $\Delta^*_{fum} \leq \Delta^*_F$.

Therefore, $\Delta^*_{fe}\Delta^*_{fum} \leq \Delta^{*2}_F = o\left(\frac{\log p}{nT}\right)$. ■

Lemma H.15. *Recall that $\widehat{\gamma}_y$ is the post-lasso estimator based on the original data, used as the coefficient to generate the bootstrap data.*

(i) $\max_m \left| \frac{1}{T}\sum_{t=1}^T \frac{1}{n}\widehat{F}^{*\prime}\tilde{U}^*_{t,m}\widehat{\delta}'_{yt}\right| = O_{P^*}\left(\sqrt{\frac{\log p}{n}\left(\frac{\log(n)\log(pT)}{n} + \Delta^2_{\max} + \frac{\log(pn)}{T}\right)}\right)$.

(ii) $\max_{m,v\le p}\left|\frac{1}{T}\sum_{k=1}^p\sum_{t=1}^T \widehat{\lambda}_{tm}\frac{1}{n}\widehat{F}^{*\prime}\tilde{U}^*_{t,v}\widehat{\gamma}_y\right| = O_{P^*}\left(\sqrt{\frac{|J|_0^2\log(p)}{n}\left(\frac{\log(n)\log(pT)}{n}+\Delta^2_{\max}+\frac{\log(pn)}{T}\right)}\right)$.

(iii) $\max_m\left|\frac{1}{T}\sum_{k=1}^p\sum_{t=1}^T \frac{1}{n}\widehat{F}^{*\prime}\tilde{U}^*_{t,m}\frac{1}{n}\widehat{F}^{*\prime}\tilde{U}^*_{t,k}\widehat{\gamma}_y\right| = o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right)$.

(iv) $\max_{m,k\le p}\left|\frac{1}{nT}\sum_{it}\tilde{U}^*_{it,k}\widehat{f}_i\widehat{\lambda}_{tm}\right| = O_{P^*}\left(\sqrt{\frac{\log p}{n}}\right)\sqrt{\frac{\log(n)\log(pT)}{n}+\Delta^2_{\max}+\frac{\log(pn)}{T}}$.

Proof. (i) **Bound** $\mathbf{\frac{1}{T}\sum_t(\widehat{\delta}_{yt} - H^{-1}\tilde{\delta}_{yt})^2}$: We have the following equality:

$$\widehat{\delta}_{yt} - H^{-1}\tilde{\delta}_{yt} = (\widehat{F}'\widehat{F})^{-1}\widehat{F}'(\tilde{F}H - \widehat{F})H^{-1}\tilde{\delta}_{yt} + (\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{U}_t\gamma_y + (\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{e}_t.$$

Hence, using Lemma H.11, we have

$$\frac{1}{T}\sum_t(\widehat{\delta}_{yt} - H^{-1}\tilde{\delta}_{yt})^2 = O_P(\Delta_F^2 + \frac{|J|_0}{n} + \Delta^2_{fe}) = O_P(|J|_0^2\Delta_F^2 + \frac{|J|_0}{n})$$

where we used Cauchy-Schwarz to bound $\Delta^2_{fe}\le\Delta^2_F$.

Bound $\max_{\mathbf{m}}|\mathbf{\frac{1}{T}\sum_{t=1}^T\frac{1}{n}\widehat{F}\tilde{U}^*_{t,m}\widehat{\delta}'_{yt}}|$: Let $\mathcal{X}_{im} = w_i^U\frac{1}{T}\sum_{t=1}^T\widehat{f}_i\widehat{U}_{it,m}\widehat{\delta}_{yt}$.

$$\begin{aligned}
\frac{1}{n}\sum_i \mathrm{Var}^*(\mathcal{X}_{im}) &= \frac{1}{n}\sum_i\left(\frac{1}{T}\sum_{t=1}^T\widehat{f}_i\widehat{U}_{it,m}\widehat{\delta}_{yt}\right)^2\\
&= \frac{1}{n}\sum_i\widehat{f}_i^2\left(\frac{1}{T}\sum_{t=1}^T\widehat{U}_{it,m}\widehat{\delta}_{yt}\right)^2\\
&\le O_P(1)\left(\frac{1}{T}\sum_{t=1}^T\widehat{U}_{it,m}\widehat{\delta}_{yt}\right)^2\\
&\le O_P(1)\left(\frac{1}{T}\sum_{t=1}^T(\widehat{U}_{it,m}-\tilde{U}_{it,m})\widehat{\delta}_{yt}\right)^2 + O_P(1)\left(\frac{1}{T}\sum_{t=1}^T\tilde{U}_{it,m}(\widehat{\delta}_{yt}-H^{-1}\tilde{\delta}_{yt})\right)^2\\
&\quad + O_P(1)\left(\frac{1}{T}\sum_{t=1}^T\tilde{U}_{it,m}\tilde{\delta}_{yt}\right)^2\\
&\le O_P(1)\max_{im}\frac{1}{T}\sum_t(\widehat{U}_{it,m}-\tilde{U}_{it,m})^2\\
&\quad + O_P\left(1+\sqrt{\frac{\log(pn)}{T}}\right)\frac{1}{T}\sum_t(\widehat{\delta}_{yt}-H^{-1}\tilde{\delta}_{yt})^2 + O_P\left(\frac{\log(pn)}{T}\right)
\end{aligned}$$

$$
\begin{aligned}
&=^{(a)} O_P\left(\log(n)\frac{\log(pT)}{n} + \log(n)\Delta_F^2 + \Delta_{\max}^2 + \frac{\log(pn)}{T}\right) \\
&\quad + O_P\left(1+\sqrt{\frac{\log(pn)}{T}}\right)\left(|J|_0^2\Delta_F^2 + \frac{|J|_0}{n}\right) \\
&= O_P\left(\log(n)\frac{\log(pT)}{n} + \Delta_{\max}^2 + \frac{\log(pn)}{T}\right)
\end{aligned}
$$

where (a) follows from (H.34) and the bound is uniform in $m \le p$. Therefore,

$$
\begin{aligned}
\max_m \left|\frac{1}{T}\sum_{t=1}^T \frac{1}{n}\widehat{F}\tilde{U}_{t,m}^*\widehat{\delta}_{yt}'\right| &= \max_m \left|\frac{1}{n}\sum_i \mathcal{X}_{im}\right| \\
&\le O_{P^*}\left(\sqrt{\frac{\log p}{n}(\frac{\log(n)\log(pT)}{n} + \Delta_{\max}^2 + \frac{\log(pn)}{T})}\right).
\end{aligned}
$$

Bound $\max_{\mathbf{m}} |\frac{\mathbf{1}}{\mathbf{T}}\sum_{\mathbf{t=1}}^{\mathbf{T}} \frac{\mathbf{1}}{\mathbf{n}}\widehat{\mathbf{F}}^{*\prime}\tilde{\mathbf{U}}_{\mathbf{t,m}}^*\widehat{\delta}_{\mathbf{yt}}'|$:

$$
\begin{aligned}
&\max_m \left|\frac{1}{T}\sum_{t=1}^T \frac{1}{n}\widehat{F}^{*\prime}\tilde{U}_{t,m}^*\widehat{\delta}_{yt}'\right| \\
&\quad \le \max_m \left|\frac{1}{T}\sum_{t=1}^T \frac{1}{n}(\widehat{F}^* - \widehat{F}H^*)'\tilde{U}_{t,m}^*\widehat{\delta}_{yt}'\right| + \max_m \left|\frac{1}{T}\sum_{t=1}^T \frac{1}{n}\widehat{F}\tilde{U}_{t,m}^*\widehat{\delta}_{yt}'\right| O_{P^*}(1).
\end{aligned}
$$

The second term is bounded from above. The first term is $O_{P^*}(\Delta_{ud}^*) = o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right)$.

(ii) First, $\gamma_y = \gamma_y^0 + R_y$ is the decomposition of exactly sparse and remainder terms for γ_y. We know that $\|\gamma_y^0\|_1 = O(|J|_0)$ and $\|R_y\|_1 = o(1)$, so $\|\gamma_y\|_1 = O(|J|_0)$.

$$
\|\widehat{\gamma}_y\|_1 \le \|\widehat{\gamma}_y - \gamma_y\|_1 + \|\gamma_y\|_1 = o_P(1) + O(|J|_0) = O_P(|J|_0).
$$

Also, $\max_{tm}|\widehat{\lambda}_{tm}| = O_P(1)$ as shown in (H.33). So by Cauchy-Schwarz,

$$
\begin{aligned}
\max_{m,v\le p}\left|\frac{1}{T}\sum_{k=1}^p\sum_{t=1}^T \widehat{\lambda}_{tm}\frac{1}{n}(\widehat{F}^* - \widehat{F}H^*)'\tilde{U}_{t,v}^*\widehat{\gamma}_y\right| &\le \max_{m,v\le p}\left(\frac{1}{nT}\sum_{it}\tilde{U}_{it,v}^{*2}\right)^{1/2} O_{P^*}(\Delta_F^*|J|_0) \\
&= O_{P^*}(\Delta_F^*|J|_0).
\end{aligned}
$$

Secondly,

$$
\max_{m,v\le p}\left|\frac{1}{T}\sum_{k=1}^p\sum_{t=1}^T \widehat{\lambda}_{tm}\frac{1}{n}\widehat{F}'\tilde{U}_{t,v}^*\widehat{\gamma}_y\right| = \max_{m,v\le p}\left|\frac{1}{Tn}\sum_{k=1}^p\sum_{t=1}^T\sum_i \widehat{\lambda}_{tm}\widehat{f}_i\widehat{U}_{it,v}w_i^U\widehat{\gamma}_y\right|.
$$

Let $\mathcal{X}_{i,mv} = \frac{1}{T}\sum_{k=1}^{p}\sum_{t=1}^{T}\widehat{\lambda}_{tm}\widehat{f}_i\widehat{U}_{it,v}w_i^U\widehat{\gamma}_y$. Then

$$\begin{aligned}
\max_{mv}\frac{1}{n}\sum_i \mathrm{Var}^*(\mathcal{X}_{i,mv}) &= \max_{mv}\frac{1}{n}\sum_i\left(\frac{1}{T}\sum_{k=1}^{p}\sum_{t=1}^{T}\widehat{\lambda}_{tm}\widehat{f}_i\widehat{U}_{it,v}\widehat{\gamma}_y\right)^2\\
&\le \max_{mv}\frac{1}{n}\sum_i \widehat{f}_i^2\left(\frac{1}{T}\sum_{t=1}^{T}\widehat{\lambda}_{tm}\widehat{U}_{it,v}\right)^2 O_P(|J|_0^2)\\
&\le \max_{imv}\left(\frac{1}{T}\sum_{t=1}^{T}\widehat{\lambda}_{tm}\widehat{U}_{it,v}\right)^2 O_P(|J|_0^2)\\
&\le O_P(|J|_0^2)\max_{imv}\left(\frac{1}{T}\sum_{t=1}^{T}(\widehat{U}_{it,v}-\tilde{U}_{it,v})\widehat{\lambda}_{tm}\right)^2\\
&\quad + O_P(|J|_0^2)\max_{imv}\left(\frac{1}{T}\sum_{t=1}^{T}\tilde{U}_{it,v}(\widehat{\lambda}_{tm}-H^{-1}\tilde{\lambda}_{tm})\right)^2\\
&\quad + O_P(|J|_0^2)\max_{imv}\left(\frac{1}{T}\sum_{t=1}^{T}\tilde{U}_{it,v}\tilde{\lambda}_{tm}\right)^2\\
&\le O_P\left(\frac{|J|_0^2\log(pn)}{T}\right) + O_P(|J|_0^2)\max_{iv}\frac{1}{T}\sum_t(\widehat{U}_{it,v}-\tilde{U}_{it,v})^2\\
&\quad + O_P\left(1+\sqrt{\frac{\log(pn)}{T}}\right)\max_m\frac{|J|_0^2}{T}\sum_t(\widehat{\lambda}_{tm}-H^{-1}\tilde{\lambda}_{tm})^2\\
&=^{(a)} O_P\left(\log(n)\frac{\log(pT)}{n}+\Delta_{\max}^2+\frac{\log(pn)}{T}\right)|J|_0^2\\
&\quad + O_P\left(1+\sqrt{\frac{\log(pn)}{T}}\right)|J|_0^2\left(\Delta_F^2+\frac{\log(pT)}{n}\right)\\
&= O_P\left(\log(n)\frac{\log(pT)}{n}+\Delta_{\max}^2+\frac{\log(pn)}{T}\right)|J|_0^2 \qquad \text{(H.38)}
\end{aligned}$$

where in (a) we used

$$\widehat{\lambda}_{tm}-H^{-1}\tilde{\lambda}_{tm} = (\widehat{F}'\widehat{F})^{-1}\widehat{F}'(\tilde{F}H-\widehat{F})H^{-1}\tilde{\lambda}_{tm} + (\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{U}_{t,m}$$

and thus

$$\max_{tm}|\widehat{\lambda}_{tm}-H^{-1}\tilde{\lambda}_{tm}| = O_P(\Delta_F+\sqrt{\frac{\log(pT)}{n}}).$$

Therefore,

$$\max_{mv}\left|\frac{1}{T}\sum_{k=1}^{p}\sum_{t=1}^{T}\widehat{\lambda}_{tm}\frac{1}{n}\widehat{F}'\tilde{U}^*_{t,v}\widehat{\gamma}_y\right| = \max_{mv}\left|\frac{1}{n}\sum_i \mathcal{X}_{i,mv}\right|$$
$$= O_{P^*}\left(\sqrt{\frac{|J|_0^2\log(p)}{n}(\frac{\log(n)\log(pT)}{n}+\Delta^2_{\max}+\frac{\log(pn)}{T})}\right).$$

(iii) Note that

$$\max_m|\frac{1}{T}\sum_{k=1}^{p}\sum_{t=1}^{T}\frac{1}{n}\widehat{F}^{*\prime}\tilde{U}^*_{t,m}\frac{1}{n}\widehat{F}^{*\prime}\tilde{U}^*_{t,k}\widehat{\gamma}_y| \leq \max_m\frac{1}{T}\sum_{t=1}^{T}|\frac{1}{n}\widehat{F}^{*\prime}\tilde{U}^*_{t,m}|^2 O_P(|J|_0).$$

By repeatedly using the triangle inequality, we have

$$\max_m\frac{1}{T}\sum_{t=1}^{T}\left|\frac{1}{n}\widehat{F}^{*\prime}\tilde{U}^*_{t,m}\right|^2 \leq \max_m\frac{1}{T}\sum_{t=1}^{T}\left|\frac{1}{n}(\widehat{F}^*-\widehat{F}H^*)'\tilde{U}^*_{t,m}\right|^2 \quad (A_1)$$
$$+O_{P^*}(1)\max_m\frac{1}{T}\sum_{t=1}^{T}\left|\frac{1}{n}\sum_i w_i^U(\widehat{f}_i-H'\tilde{f}_i)\widehat{U}_{it,m}\right|^2 \quad (A_2)$$
$$+O_{P^*}(1)\max_m\frac{1}{T}\sum_{t=1}^{T}\left|\frac{1}{n}\sum_i w_i^U\tilde{f}_i(\widehat{U}_{it,m}-\tilde{U}_{it,m})\right|^2 \quad (A_3)$$
$$+O_{P^*}(1)\left|\frac{1}{n}\sum_i w_i^U\tilde{f}_i\right|^2\max_m\frac{1}{T}\sum_{t=1}^{T}\left|\bar{U}_{\cdot t,m}\right|^2 \quad (A_4)$$
$$+O_{P^*}(1)\max_m\left|\frac{1}{n}\sum_i w_i^U\tilde{f}_i\bar{U}_{i\cdot,m}\right|^2 \quad (A_5)$$
$$+O_{P^*}(1)\left|\frac{1}{n}\sum_i w_i^U\tilde{f}_i\right|^2\max_m\left|\bar{\bar{U}}_m\right|^2 \quad (A_6)$$
$$+O_{P^*}(1)\frac{1}{n}\sum_i\frac{1}{n}(w_i^U)^2\tilde{f}_i^2\max_m\frac{1}{T}\sum_{t=1}^{T}EU^2_{it,m} \quad (A_7)$$
$$+O_{P^*}(1)\max_m\frac{1}{T}\sum_{t=1}^{T}\frac{1}{n}\sum_i\frac{1}{n}(w_i^U)^2\tilde{f}_i^2(U^2_{it,m}-EU^2_{it,m}) \quad (A_8)$$
$$+O_{P^*}(1)\max_m\frac{1}{T}\sum_{t=1}^{T}\frac{1}{n}\sum_i\frac{1}{n}\sum_{j\neq i}w_i^U\tilde{f}_jw_j^U\tilde{f}_iU_{it,m}U_{jt,m} \quad (A_9)$$

$$:= \sum_{l=1}^{9} A_l.$$

We shall show each $A_l = o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right)$. Note that we shall show that the last two terms are $o_{PP^*}\left(\sqrt{\frac{\log p}{nT}}\right)$, where $o_{PP^*}(1)$ means convergence in probability with respect to the joint distribution of the original data and the bootstrap weights $\{w_i^U\}$. This convergence then implies by the Markov inequality that they are also $o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right)$, i.e., that they converge in probability with respect to the bootstrap sampling distribution (see Lemma 3 of Cheng and Huang (2010)).

Term $\mathbf{A_1}$. It is clear that term (A_1) is $O_{P^*}(\Delta_{fum}^{*2})$. By (H.36), $\Delta_{fum}^* \leq \Delta_F^*$. Hence $|J|_0\Delta_{fum}^{*2} \leq |J|_0\Delta_F^{*2} = o\left(\sqrt{\frac{\log p}{nT}}\right)$.

Term $\mathbf{A_2}$. Term (A.2) is bounded by $\frac{1}{n}\sum_i (w_i^U(\widehat{f}_i - H'\tilde{f}_i))^2 \max_m \frac{1}{T}\sum_{t=1}^T \frac{1}{n}\sum_j \widehat{U}_{jt,m}^2$. Note that $E^*\left[\frac{1}{n}\sum_i (w_i^U(\widehat{f}_i - H'\tilde{f}_i))^2\right] = O_P(\Delta_F^2)$, and $\max_m \frac{1}{T}\sum_{t=1}^T \frac{1}{n}\sum_j \widehat{U}_{jt,m}^2 = O_P(1)$. Hence, $|J|_0 A_2 = O_{P^*}(|J|_0\Delta_F^2)O_P(1) = O_{P^*}(|J|_0\Delta_F^2) = o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right)$, where we used $O_{P^*} \times O_P = O_{P^*}$ using Lemma 3 of Cheng and Huang (2010).

Term $\mathbf{A_3 \sim A_6}$. These terms can be shown to be $o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right)$ similarly to the other terms with provided details. We omit the details for brevity.

Term $\mathbf{A_7}$. Term (A_7) is bounded by $\frac{1}{n}\sum_i \frac{1}{n}\tilde{f}_i^2(w_i^U)^2 \max_{itm}|(EU_{it,m}^2)|$. Also,

$$E^*\left(\frac{1}{n}\sum_i \frac{1}{n}\tilde{f}_i^2(w_i^U)^2\right) = \frac{1}{n}\sum_i \frac{1}{n}\tilde{f}_i^2 = O_P\left(\frac{1}{n}\right).$$

Hence, $|J|_0 A_7 = O_{P^*}(\frac{|J|_0}{n}) = o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right)$.

Term $\mathbf{A_8}$. Let $W^U = \{w_i^*\}_{i\leq n}$. Now let P_{U,F,W^U}, $P_{U|F,W^U}$ and P_{F,W^U} respectively be the probability measures with respect to the joint distribution of $\{U, F, W^U\}$, the conditional distribution $U|F, W^U$, and the joint distribution of (F, W^U). Let E_{F,W^U} be the expectation operator with respect to the distribution of (F, W^U).

If $T = O(1)$, then for $Z_{itm} := \tilde{f}_i^2(U_{it,m}^2 - EU_{it,m}^2)(w_i^U)^2$, $A_8 \leq \max_{it}|\frac{1}{n}\sum_i \frac{1}{n}Z_{itm}|$. Then conditional on F, W^U,

$$\max_{tm} \mathrm{Var}(Z_{itm}|F, W^U) \leq \tilde{f}_i^4(w_i^U)^4 \max_{tm} \mathrm{Var}(U_{it,m}^2|F, W^U)$$

$$\leq C\tilde{f}_i^4(w_i^U)^4,$$

almost surely in F, W^U for a generic constant C that does not depend on F, W^U. $\forall x > 0$, and $l \leq 8$,

$$\begin{aligned}
&P_{U,F,W^U}(A_8 > x) \\
&\quad \leq P_{F,W^U}(\frac{1}{n}\sum_i(|w_i^U|^l - E|w_i^U|^l) > C)P_{F,W^U}(\frac{1}{n}\sum_i(|f_i|^l - E|f_i|^l) > C) \\
&\qquad + E_{F,W^U}P_{U|F,W^U}(A_8 > x)\,1\left\{\frac{1}{n}\sum_i \tilde{f}_i^4(w_i^U)^4 < C\right\} \\
&\quad = o(1) \\
&\qquad + E_{F,W^U}P_{U|F,W^U}\left(\max_{it}|\frac{1}{n}\sum_i Z_{itm}| > nx\right)1\{\max_{tm}\mathrm{Var}(Z_{itm}|F, W^U) < C\}.
\end{aligned}$$

Because $\{U_i\}_{i\leq n}$ are independent conditional on $\{w_i^U, F\}$, we can apply the Bernstein inequality for independent data to reach, for $x = M\frac{1}{n}\sqrt{\frac{\log(pT)}{n}}$, for sufficiently large $M > 0$,

$$\begin{aligned}
&E_{F,W^U}P_{U|F,W^U}\left(\max_{it}|\frac{1}{n}\sum_i Z_{itm}| > nx\right)1\{\max_{tm}\mathrm{Var}(Z_{itm}|F, W^U) < C\} \\
&\leq E_{F,W^U}\exp\left(\log(pT) - \frac{M^2\log(pT)}{\max_{tm}\mathrm{Var}(Z_{itm}|F, W^U)}\right)1\left\{\max_{tm}\mathrm{Var}(Z_{itm}|F, W^U) < C\right\} = o(1).
\end{aligned}$$

Hence $A_8 = O_P(x) = O_P(\sqrt{\frac{\log(p)}{n^3}})$ when $T = O(1)$.

When $T \to \infty$, define $\mathcal{X}_{t,m} = \frac{1}{n}\sum_i \frac{1}{n}\tilde{f}_i^2(U_{it,m}^2 - EU_{it,m}^2)(w_i^U)^2$, so $A_8 = \max_m|\frac{1}{T}\sum_{t=1}^T \mathcal{X}_{t,m}|$. Then conditional on F, W^U,

$$\begin{aligned}
\max_{tm}\mathrm{Var}(\mathcal{X}_{t,m}|F, W^U) &\leq \frac{1}{n^4}\sum_i \tilde{f}_i^4(w_i^U)^4\max_{tm}\mathrm{Var}(U_{it,m}^2|F, W^U) \\
&\leq C\frac{1}{n^4}\sum_i \tilde{f}_i^4(w_i^U)^4.
\end{aligned}$$

almost surely in F, W^U for a generic constant C that does not depend on F, W^U. So

$$\begin{aligned}
&P_{U,F,W^U}\left(\max_m|\frac{1}{T}\sum_{t=1}^T \mathcal{X}_{t,m}| > x\right) \\
&\quad \leq P_{F,W^U}(\frac{1}{n}\sum_i(|w_i^U|^l - E|w_i^U|^l) > C)P_{F,W^U}(\frac{1}{n}\sum_i(|f_i|^l - E|f_i|^l) > C)
\end{aligned}$$

$$
\begin{aligned}
&+ E_{F,W^U} P_{U|F,W^U}\left(\max_m |\frac{1}{T}\sum_{t=1}^T \mathcal{X}_{t,m}| > x\right) 1\left\{\frac{1}{n}\sum_i \tilde{f}_i^4 (w_i^U)^4 < C\right\} \\
= o(1)& \\
&+ E_{F,W^U} P_{U|F,W^U}\left(\max_m |\frac{1}{T}\sum_{t=1}^T \mathcal{X}_{t,m}| > x\right) 1\{\max_{tm} \mathrm{Var}(\mathcal{X}_{t,m}|F, W^U) < \frac{C_0}{n^3}\}.
\end{aligned}
$$

We now bound $P_{U|F,W^U}\left(\max_m |\frac{1}{T}\sum_{t=1}^T \mathcal{X}_{t,m}| > x\right)$ conditional on $\max_{tm} \mathrm{Var}(\mathcal{X}_{t,m}|F, W^U) < \frac{C_0}{n^{3/2}}$. We have $\min_{tm}[\mathrm{Var}(\mathcal{X}_{t,m}|F)^{-1/2}] \geq C_0 n^{3/2}$ for a generic $C_0 > 0$. Conditional on F, $\{\mathcal{X}_{t,m}\}$ is a strong mixing sequence across t with uniform mixing condition (uniform over $m \leq p$) that is bounded by $\exp(-CT^r)$ (Assumption 3.1). Write $\mathcal{Z}_{t,m} = \mathcal{X}_{t,m} \mathrm{Var}(\mathcal{X}_{t,m}|F)^{-1/2}$, whose conditional variance is one. Recall that by assumption, $\log^\gamma p = o(n)$ for some $\gamma > 2$ and the conditional strong mixing (Assumption 3.1) holds for the process $\{(U_t)\}_{t=-\infty}^{+\infty}$ given F, W^U, with mixing coefficient bounded by $\exp(-CT^r)$, $r > 1$, and also we assume $\gamma r \geq 2$. Let $\bar{r} = \min\{r, 1\}$, $r_1 = (0.5 + \bar{r}^{-1})^{-1}$, $c = 0.5(\gamma + 1)$, then $r_1 < 1$, $cr_1 > 1$ and $2c \geq 1$. Because $\bar{r} \leq r$, the strong mixing condition in Assumption 3.1 also holds with $\bar{r}$ in place of r. All these constants are independent of F. Then apply the Bernstein inequality for strong mixing sequences Merlevède et al. (2011) on $|\frac{1}{T}\sum_t \mathcal{Z}_{t,m}|$, for $y = M\frac{(\log p)^c}{\sqrt{T}}$, and sufficiently large $M > 0$, for $x = y/(C_0 n^{3/2})$,

$$
\begin{aligned}
P_{U|F,W^U}\left(\max_m |\frac{1}{T}\sum_{t=1}^T \mathcal{X}_{t,m}| > x\right) &\leq p \max_m P_{U|F,W^U}\left(|\frac{1}{T}\sum_t \mathcal{Z}_{t,m}| > x \min_{tm}[\mathrm{Var}(\mathcal{X}_{t,m}|F)^{-1/2}]\right) \\
&\leq p \max_m P_{U|F,W^U}\left(|\frac{1}{T}\sum_t \mathcal{Z}_{t,m}| > C_0 n^{3/2} x\right) \leq A_1 + A_2 + A_3
\end{aligned}
$$

where for a generic constant C that is independent of F,

$$
\begin{aligned}
A_1 &= pT\exp(-C(Ty)^{r_1}) = \exp(\log(pT) - CM^{r_1}T^{r_1/2}\log^{cr_1} p) = o(1), \quad (cr_1 \geq 1) \\
A_2 &= p\exp\left(-C\frac{(Ty)^2}{T}\exp\left(\frac{(Ty)^{r_1(1-r_1)}}{C\log^{r_1}(Ty)}\right)\right) = o(1), \quad (r_1 < 1, 2c > 1, Ty \to \infty) \\
A_3 &= p\exp\left(-CTy^2\right) = \exp(\log p - CM^2\log^{2c} p) = o(1),
\end{aligned}
$$

and $o(1)$ in the above are non-stochastic converging sequences. (This is because the involved generic constant C in Merlevède et al. (2011) is independent of F, W^U. Hence

$$
E_{F,W^U} P_{U|F,W^U}\left(\max_m |\frac{1}{T}\sum_{t=1}^T \mathcal{X}_{t,m}| > x\right) 1\{\max_{tm} \mathrm{Var}(\mathcal{X}_{t,m}|F, W^U) < \frac{C_0}{n^3}\} = o(1).
$$

Consequently, $\max_m |\frac{1}{T}\sum_t \mathcal{X}_{t,m}| = O_P(x) = O_P(\frac{\log^{0.5\gamma+0.5} p}{\sqrt{Tn^3}})$ when $T\to\infty$.

Combining with the case $T = O(1)$, we have $A_8 \leq \max_m |\frac{1}{T}\sum_t \mathcal{X}_{t,m}| = O_P(\frac{\log^{0.5\gamma+0.5} p}{\sqrt{Tn^3}})$. Because the convergence is with respect to the joint distribution of (U, F, W^U), the above result implies, by Lemma 3 of Cheng and Huang (2010), $|J|_0 A_8 = O_{P^*}\left(|J|_0 \frac{\log^{0.5\gamma+0.5} p}{\sqrt{Tn^3}}\right)$. Thus, $|J|_0 A_8 = o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right)$, given that $|J|_0^2 \log^\gamma p = o(n^2)$.

Term $\mathbf{A_9}$. If $T = O(1)$, then

$$A_9 \leq \max_{mt} \| \frac{1}{n}\sum_i w_i^U \tilde f_i U_{it,m}\| \max_{mt} \|\frac{1}{n}\sum_j \tilde f_j w_j^U U_{jt,m}\|.$$

Using the same argument when we bound A_9 when $T = O(1)$, we conclude that $A_9 = O_P(\frac{\log p}{n})$.

When $T\to\infty$, define $G_{t,m} = \frac{1}{n}\sum_i \frac{1}{n}\sum_{j\neq i} \tilde f_j w_i^U \tilde f_i U_{it,m} U_{jt,m} w_j^U$. Then $E(G_{t,m}|F, W^U) = 0$, and almost surely in F, W^U, and $A_9 \leq \max_m \|\frac{1}{T}\sum_t G_{t,m}\|_F$. We only focus on the case $\dim(f_i) = 1$.

$$\begin{aligned}
&\max_{tm} \mathrm{Var}(G_{t,m}|F, W^U)\\
&\quad\leq \max_{tm}\sum_i\sum_{j\neq i}\frac{1}{n^4}\tilde f_j^2 w_i^U \tilde f_i^2 w_j^U w_i^U w_j^U E(U_{it,m}^2 U_{jt,m}^2|F, W^U)\\
&\qquad + \max_{tm}\sum_i\sum_{j\neq i}\frac{1}{n^4}\tilde f_j w_i^U \tilde f_i w_j^U \tilde f_i w_j^U \tilde f_j w_i^U E(U_{jt,m}^2 U_{it,m}^2|F, W^U)\\
&\quad\leq \frac{C}{n^2}\frac{1}{n}\sum_i (w_i^U)^2 \tilde f_i^2 \frac{1}{n}\sum_{j\neq i}\tilde f_j^2 (w_j^U)^2.
\end{aligned}$$

Hence, for $l \leq 4$ and any $x > 0$,

$$\begin{aligned}
&P_{U,F,W^U}\left(\max_m \left|\frac{1}{T}\sum_{t=1}^T G_{t,m}\right| > x\right)\\
&\quad\leq P_{F,W^U}\left(\frac{1}{n}\sum_i (|w_i^U|^l - E|w_i^U|^l) > C\right) + P_{F,W^U}\left(\frac{1}{n}\sum_i (|f_i|^l - E|f_i|^l) > C\right)\\
&\qquad + E_{F,W^U} P_{U|F,W^U}\left(\max_m |\frac{1}{T}\sum_{t=1}^T G_{t,m}| > x\right) 1\left\{\frac{1}{n}\sum_i (w_i^U)^2 \tilde f_i^2 \frac{1}{n}\sum_{j\neq i}\tilde f_j^2 (w_j^U)^2 < C\right\}\\
&\quad\leq o(1) + E_{F,W^U} P_{U|F,W^U}\left(\max_m |\frac{1}{T}\sum_{t=1}^T G_{t,m}| > x\right) 1\left\{\max_{tm}\mathrm{Var}(G_{t,m}|F, W^U) < \frac{C_0}{n^2}\right\}.
\end{aligned}$$

We note that $\{G_{t,m}\}_{t\leq T}$ is conditionally strong mixing, given F, W^U, whose mixing coefficient is independent of $m \leq p, F, W^U$. Thus, we can use the same argument as in bounding term A_8 for the case of $T\to\infty$, to conclude that when $x = O_P\left(\frac{\log^{0.5\gamma+0.5} p}{\sqrt{Tn^2}}\right)$

$$E_{F,W^U} P_{U|F,W^U}\left(\max_m |\frac{1}{T}\sum_{t=1}^T G_{t,m}| > x\right) 1\left\{\max_{tm} \mathrm{Var}(G_{t,m}|F, W^U) < \frac{C_0}{n^2}\right\} \leq o(1)$$

Combining with the case of $T = O(1)$, $A_9 = O_P(\frac{\log^{0.5\gamma+0.5} p}{\sqrt{Tn^2}}) = O_{P^*}(\frac{\log^{0.5\gamma+0.5} p}{\sqrt{Tn^2}})$, where the second equality is due to Lemma 3 of Cheng and Huang (2010). Thus $|J|_0 A_9 = O_{P^*}\left(|J|_0 \frac{\log^{0.5\gamma+0.5} p}{\sqrt{Tn^2}}\right) = o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right)$ given that $J^2\log^\gamma p = o(n)$. Hence, we have proven that

$$\begin{aligned}\max_m \left|\frac{1}{T}\sum_{k=1}^p\sum_{t=1}^T \frac{1}{n}\widehat{F}^{*\prime}\tilde{U}^*_{t,m}\frac{1}{n}\widehat{F}^{*\prime}\tilde{U}^*_{t,k}\widehat{\gamma}_y\right| &\leq \max_m \frac{1}{T}\sum_{t=1}^T\left|\frac{1}{n}\widehat{F}^{*\prime}\tilde{U}^*_{t,m}\right|^2 O_P(|J|_0) \\ &= o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right).\end{aligned} \tag{H.39}$$

(iv) Let $G_{i,mk} = \frac{1}{T}\sum_t \widehat{U}_{it,k} w_i^U \widehat{f}_i \widehat{\lambda}_{tm}$. Then by (H.38),

$$\begin{aligned}\max_{mk}\frac{1}{n}\sum_i \mathrm{Var}^*(G_{i,mk}) &\leq \max_{mk}\frac{1}{n}\sum_i \widehat{f}_i^2\left(\frac{1}{T}\sum_{t=1}^T \widehat{\lambda}_{tm}\widehat{U}_{it,k}\right)^2 \\ &= O_P\left(\frac{\log(n)\log(pT)}{n} + \Delta^2_{\max} + \frac{\log(pn)}{T}\right).\end{aligned}$$

Hence,

$$\begin{aligned}\max_{m,k\leq p}\left|\frac{1}{nT}\sum_{it}\tilde{U}^*_{it,k}\widehat{f}_i\widehat{\lambda}_{tm}\right| &= \max_{mk}\left|\frac{1}{n}\sum_i G_{i,mk}\right| \\ &= O_{P^*}\left(\sqrt{\frac{\log p}{n}}\right)\sqrt{\frac{\log(n)\log(pT)}{n} + \Delta^2_{\max} + \frac{\log(pn)}{T}}.\end{aligned}$$

■

Lemma H.16. *(i)* $\left\|\frac{1}{n}\sum_{i=1}^n \frac{1}{T}\sum_{t=1}^T(\widehat{U}^*_{it} - \tilde{U}^*_{it})\widehat{\delta}'_{yt}H^{*\prime -1}(\widehat{f}^*_i - H^{*\prime}\widehat{f}_i)\right\|_\infty = o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right)$.

(ii) $\left\|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T(\widehat{U}^*_{it} - \tilde{U}^*_{it})(\tilde{U}^*_{it} - \widehat{U}^*_{it})'\widehat{\gamma}_y\right\|_\infty = o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right)$

(iii) $\left\|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T \tilde{U}^*_{it}(\tilde{U}^*_{it} - \widehat{U}^*_{it})'\widehat{\gamma}_y\right\|_\infty = o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right)$.

Proof. We shall repeatedly use the following equality:

$$\tilde{U}^*_{it,m} - \widehat{U}^*_{it,m} = \widehat{f}^{*\prime}_i(\widehat{F}^{*\prime}\widehat{F}^*)^{-1}\widehat{F}^{*\prime}(\widehat{F}H^* - \widehat{F}^*)H^{*-1}\widehat{\lambda}_{tm} + \widehat{f}^{*\prime}_i(\widehat{F}^{*\prime}\widehat{F}^*)^{-1}\widehat{F}^{*\prime}\tilde{U}^*_{t,m} + \widehat{\lambda}'_{tm}H^{*\prime-1}(\widehat{f}^*_i - H^{*\prime}\widehat{f}_i), \tag{H.40}$$

where $\widehat{F}^*$ and $\widehat{U}^*$ represent the estimators using bootstrapped data.

(i) First, by (H.33), $\max_{tm}|\widehat{\lambda}_{tm}| = O_P(1)$. Hence,

$$\max_m\left|\frac{1}{T}\sum_{t=1}^T\widehat{\lambda}_{tm}\widehat{\delta}_{yt}\right| = O_P(1)\left(\frac{1}{T}\sum_t\widehat{\delta}^2_{yt}\right)^{1/2} = O_P(1).$$

Also,

$$\begin{aligned}
&\max_m\left|\frac{1}{n}\sum_{i=1}^n\frac{1}{T}\sum_{t=1}^T(\widehat{U}^*_{it,m} - \tilde{U}^*_{it,m})\widehat{\delta}'_{yt}H^{*\prime-1}(\widehat{f}^*_i - H^{*\prime}\widehat{f}_i)\right| \\
&= \max_m\left|\frac{1}{n}\sum_{i=1}^n\frac{1}{T}\sum_{t=1}^T\widehat{f}^{*\prime}_i(\widehat{F}^{*\prime}\widehat{F}^*)^{-1}\widehat{F}^{*\prime}(\widehat{F}H^* - \widehat{F}^*)H^{*-1}\widehat{\lambda}_{tm}\widehat{\delta}'_{yt}H^{*\prime-1}(\widehat{f}^*_i - H^{*\prime}\widehat{f}_i)\right| \\
&\quad + \max_m\left|\frac{1}{n}\sum_{i=1}^n\frac{1}{T}\sum_{t=1}^T\widehat{f}^{*\prime}_i(\widehat{F}^{*\prime}\widehat{F}^*)^{-1}\widehat{F}^{*\prime}\tilde{U}^*_{t,m}\widehat{\delta}'_{yt}H^{*\prime-1}(\widehat{f}^*_i - H^{*\prime}\widehat{f}_i)\right| \\
&\quad + \max_m\left|\frac{1}{n}\sum_{i=1}^n\frac{1}{T}\sum_{t=1}^T\widehat{\lambda}'_{tm}H^{*\prime-1}(\widehat{f}^*_i - H^{*\prime}\widehat{f}_i)\widehat{\delta}'_{yt}H^{*\prime-1}(\widehat{f}^*_i - H^{*\prime}\widehat{f}_i)\right| \\
&\leq \max_m\left|\frac{1}{T}\sum_{t=1}^T\widehat{\lambda}_{tm}\widehat{\delta}_{yt}\right|O_{P^*}(\Delta^{*2}_F) + \max_m\left|\frac{1}{T}\sum_{t=1}^T\frac{1}{n}\widehat{F}^{*\prime}\tilde{U}^*_{t,m}\widehat{\delta}'_{yt}\right|O_{P^*}(\Delta^*_F) \\
&=^{(a)} O_{P^*}(\Delta^{*2}_F) + O_{P^*}\left(\sqrt{\frac{\log p}{n}\left(\frac{\log(n)\log(pT)}{n} + \Delta^2_{\max} + \frac{\log(pn)}{T}\right)}\right)\Delta^*_F \\
&=^{(b)} o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right)
\end{aligned} \tag{H.41}$$

where (a) follows from Lemma H.15(i) and (b) follows from Assumption D.4.

(ii) We have

$$\begin{aligned}
&\max_m\left|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T(\widehat{U}^*_{it,m} - \tilde{U}^*_{it,m})(\tilde{U}^*_{it} - \widehat{U}^*_{it})'\widehat{\gamma}_y\right| \\
&\leq \max_m\left|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T\widehat{f}^{*\prime}_i\widehat{\lambda}_{tm}(\tilde{U}^*_{it} - \widehat{U}^*_{it})'\widehat{\gamma}_y\right|O_{P^*}(\Delta^*_F)
\end{aligned}$$

$$
\begin{aligned}
&+\max_m \left| \frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T \widehat{f}_i^{*\prime}\frac{1}{n}\widehat{F}^{*\prime}\tilde{U}_{t,m}^*(\tilde{U}_{it}^*-\widehat{U}_{it}^*)'\widehat{\gamma}_y\right| O_{P^*}(1)\\
&+\max_m \left| \frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T \widehat{\lambda}_{tm}'(\widehat{f}_i^*-H^{*\prime}\widehat{f}_i)(\tilde{U}_{it}^*-\widehat{U}_{it}^*)'\widehat{\gamma}_y\right| O_{P^*}(1)\\
\le&\max_m \left| \frac{1}{T}\sum_{k=1}^p\sum_{t=1}^T \widehat{\lambda}_{tm}\widehat{\lambda}_{tk}\widehat{\gamma}_y\right| O_{P^*}(\Delta_F^{*2})\\
&+\max_m \left| \frac{1}{T}\sum_{k=1}^p\sum_{t=1}^T \widehat{\lambda}_{tm}\frac{1}{n}\widehat{F}^{*\prime}\tilde{U}_{t,k}^*\widehat{\gamma}_y\right| O_{P^*}(\Delta_F^{*})\\
&+\max_m \left| \frac{1}{T}\sum_{k=1}^p\sum_{t=1}^T \frac{1}{n}\widehat{F}^{*\prime}\tilde{U}_{t,m}^*\widehat{\lambda}_{tk}\widehat{\gamma}_y\right| O_{P^*}(\Delta_F^{*})\\
&+\max_m \left| \frac{1}{T}\sum_{k=1}^p\sum_{t=1}^T \frac{1}{n}\widehat{F}^{*\prime}\tilde{U}_{t,m}^*\frac{1}{n}\widehat{F}^{*\prime}\tilde{U}_{t,k}^*\widehat{\gamma}_y\right| O_{P^*}(1)\\
\le^{(a)}& O_{P^*}(|J|_0\Delta_F^{*2})+O_{P^*}\left(\sqrt{\frac{|J|_0^2\log(p)}{n}\left(\frac{\log(n)\log(pT)}{n}+\Delta_{\max}^2+\frac{\log(pn)}{T}\right)}\right)\Delta_F^*\\
&+o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right)\\
=&\, o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right),
\end{aligned}
$$

where we used Lemma H.15 in (a) and the assumption that $\Delta_{\max}^2|J|_0^2T\Delta_F^{*2}=o(1)$.

(iii) We have

$$
\begin{aligned}
\left\|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T\tilde{U}_{it}^*(\tilde{U}_{it}^*-\widehat{U}_{it}^*)'\widehat{\gamma}_y\right\|_\infty &= \max_{k\le p}\left|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T\tilde{U}_{it,k}^*\sum_{m=1}^p\widehat{\gamma}_{ym}(\tilde{U}_{it,m}^*-\widehat{U}_{it,m}^*)\right|\\
&\le \max_{k\le p}\left|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T\tilde{U}_{it,k}^*\widehat{f}_i^*\widehat{\gamma}_y'\widehat{\Lambda}_t\right| O_{P^*}(\Delta_F^*)\\
&\quad+\max_{k\le p}\left|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T\tilde{U}_{it,k}^*\widehat{f}_i^*\sum_{m=1}^p\widehat{\gamma}_{ym}\frac{1}{n}\widehat{F}^{*\prime}\tilde{U}_{t,m}^*\right| O_{P^*}(1)\\
&\quad+\max_{k\le p}\left|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T\tilde{U}_{it,k}^*(\widehat{f}_i^*-H^{*\prime}\widehat{f}_i)\widehat{\gamma}_y'\widehat{\Lambda}_t\right| O_{P^*}(1)
\end{aligned}
$$

$$= o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right).$$

The first term following the inequality is $O_{P^*}\left(\sqrt{\frac{|J|_0^2\log(p)}{n}\left(\frac{\log(n)\log(pT)}{n}+\Delta_{\max}^2+\frac{\log(pn)}{T}\right)}\right)\Delta_F^* = o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right)$ follows from Lemma H.15. The second term following the inequality is bounded using Lemma H.15(iii). By Assumption D.5, the third term is $o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right)$.

■

Lemma H.17. *(i)* $\|\widehat{\Psi}^{y-1}\frac{1}{nT}\sum_{t=1}^T \tilde{U}_t^{*\prime}\tilde{e}_t^*\|_\infty \leq \frac{1}{\sqrt{nT}}\Phi^{-1}(1-\frac{q_n}{2p})(1+o_{P^*}(1))$.
(ii) $\|\frac{1}{nT}\sum_{t=1}^T(\tilde{U}_t^*-\widehat{U}_t^*)'\tilde{e}_t^*\|_\infty = o_{P^*}(\sqrt{\frac{\log p}{nT}})$.
(iii) $\|\frac{1}{nT}\sum_{t=1}^T M_t^{*\prime}\widehat{U}_t^*\|_\infty = o_P(\sqrt{\frac{\log p}{nT}})$.
(iv) $|J|_0\max_{m,k\leq p}|\frac{1}{nT}\sum_{it}(\widehat{U}_{it,m}^*\widehat{U}_{it,k}^*-\tilde{U}_{it,m}^*\tilde{U}_{it,k}^*)| = o_{P^*}(1)$.
(v) $\frac{1}{nT}\sum_{t=1}^T\|\widehat{U}_t^*\beta\|_2^2 \geq \frac{1}{nT}\sum_{t=1}^T\|\tilde{U}_t\beta\|_2^2 - \|\beta\|_2^2 o_{P^*}(1)$*, for any vector* $p\times 1$ *vector* $\|\beta\|_1^2\leq C\|\beta\|_2^2|J|_0$.

Proof. (i) Note that $\tilde{e}_{it}^* = w_i^Y\widehat{\epsilon}_{it}+\widehat{\alpha}w_i^D\widehat{\eta}_{it}$. Define $a_{im}=\frac{1}{\sqrt{T}}\sum_t\widehat{U}_{it,m}\widehat{\epsilon}_{it}$, $b_{im}=\frac{1}{\sqrt{T}}\sum_t\widehat{U}_{it,m}\widehat{\alpha}\widehat{\eta}_{it}$, and $W_{im}=\frac{1}{\sqrt{T}}\sum_t\widehat{U}_{it,m}^*\tilde{e}_{it}^* = w_i^Uw_i^Ya_{im}+w_i^Uw_i^Db_{im}$. Note that

$$\begin{aligned}
E^*\left|w_i^Uw_i^Ya_{im}\right|^3 &\leq |a_{im}|^3E^*|w_i^U|^3E^*|w_i^Y|^3\\
&\leq C\left|\frac{1}{\sqrt{T}}\sum_t\widehat{U}_{it,m}\widehat{\epsilon}_{it}\right|^3\\
&= \left|\frac{C}{\sqrt{T}}\sum_t\tilde{U}_{it,m}\tilde{\epsilon}_{it}\right|^3+o_P(1)=O_P(1).\\
E^*|w_i^Uw_i^Db_{im}|^3 &\leq |b_{im}|^3E^*|w_i^U|^3E^*|w_i^D|^3\\
&\leq C\left|\frac{1}{\sqrt{T}}\sum_t\widehat{U}_{it,m}\widehat{\alpha}\widehat{\eta}_{it}\right|^3\\
&= \left|\frac{C}{\sqrt{T}}\sum_t\tilde{U}_{it,m}\alpha\tilde{\eta}_{it}\right|^3+o_P(1)=O_P(1).\\
E^*|w_i^Uw_i^Db_{im}|^2|w_i^Uw_i^Ya_{im}| &\leq C|b_{im}|^2|a_{im}|=O_P(1).\\
E^*|w_i^Uw_i^Ya_{im}|^2|w_i^Uw_i^Db_{im}| &\leq C|a_{im}|^2|b_{im}|=O_P(1).
\end{aligned}$$

Therefore, $E^*|W_{im}|^3 = O_P(1)$. Also,

$$\begin{aligned} E^*W_{im}^2 &= E^*(w_i^U w_i^Y)^2 a_{im}^2 + (w_i^U w_i^D)^2 b_{im}^2 + 2(w_i^U)^2 w_i^Y w_i^D a_{im} b_{im} \\ &= a_{im}^2 + b_{im}^2 \\ &\geq \left(\frac{1}{\sqrt{T}}\sum_t \tilde{U}_{it,m}\tilde{\epsilon}_{it}\right)^2 + \left(\frac{1}{\sqrt{T}}\sum_t \tilde{U}_{it,m}\tilde{\alpha}\tilde{\eta}_{it}\right)^2 - o_P(1). \end{aligned}$$

Hence, $E^*W_{im}^2 \geq C - o_P(1)$ uniformly in i which implies

$$\max_{i\leq n, m\leq p} \frac{(E^*|W_{im}|^3)^{1/3}}{(E^*W_{im}^2)^{1/2}} \leq C + o_P(1).$$

Also note that $E^*W_{im} = 0$. It follows from Lemma 5 of Belloni et al. (2012),

$$\max_{m\leq p} \frac{|\frac{1}{n}\sum_i W_{im}|}{(\frac{1}{n}\sum_i W_{im}^2)^{1/2}} < \frac{1}{\sqrt{n}}\Phi^{-1}\left(1 - \frac{q_n}{2p}\right)$$

with probability approaching one.

Since

$$\frac{1}{nT}\sum_i \sum_{t=1}^T \tilde{U}_{it,m}^* \tilde{e}_{it}^* = \frac{1}{nT}\sum_i \sum_{t=1}^T \widehat{U}_{it,m} w_i^U (w_i^Y \widehat{\epsilon}_{it} + \widehat{\alpha} w_i^D \widehat{\eta}_{it}) = \frac{1}{n\sqrt{T}}\sum_i W_{im},$$

by the triangle inequality, for

$$G = \max_m |[\frac{1}{n}\sum_i (\frac{1}{\sqrt{T}}\sum_t \widehat{U}_{it,m}\widehat{e}_{it})^2]^{-1/2} - [\frac{1}{n}\sum_i W_{im}^2]^{-1/2}|,$$

we have

$$
\begin{aligned}
&\left\|\widehat{\Psi}^{y-1} \ \frac{1}{nT}\sum_{t=1}^{T}\tilde{U}_t^{*\prime}\tilde{e}_t^*\right\|_\infty \\
&\quad\le G\max_m\left|\frac{1}{nT}\sum_{it}\tilde{U}^*_{it,m}\tilde{e}^*_{it}\right| + \max_m\left|\left[\frac{1}{n}\sum_i W_{im}^2\right]^{-1/2}\frac{1}{nT}\sum_{it}\tilde{U}^*_{it,m}\tilde{e}^*_{it}\right| \\
&\quad\le G\max_m\left[\frac{1}{n}\sum_i W_{im}^2\right]^{1/2}\max_m\left|\left[\frac{1}{n}\sum_i W_{im}^2\right]^{-1/2}\frac{1}{\sqrt{T}}\frac{1}{n}\sum_i W_{im}\right| \qquad \text{(H.42)} \\
&\qquad + \max_m\left|\left[\frac{1}{n}\sum_i W_{im}^2\right]^{-1/2}\frac{1}{\sqrt{T}}\frac{1}{n}\sum_i W_{im}\right| \\
&\quad\le\left(G\max_m[\frac{1}{n}\sum_i W_{im}^2]^{1/2}+1\right)\frac{1}{\sqrt{nT}}\Phi^{-1}\left(1-\frac{q_n}{2p}\right).
\end{aligned}
$$

We now show $G = o_{P^*}(1)$ for which it suffices to show $\max_m|\frac{1}{n}\sum_i(\frac{1}{\sqrt{T}}\sum_t\widehat{U}_{it,m}\widehat{e}_{it})^2 - \frac{1}{n}\sum_i W_{im}^2| = o_{P^*}(1)$. In the bootstrap sampling space, $\{W_{im}\}_{i\le n}$ is i.i.d. with exponential tails. Hence, by the Bernstein inequality for i.i.d. data and union bounds, it can be shown that

$$\max_m|\frac{1}{n}\sum_i W_{im}^2 - E^*\frac{1}{n}\sum_i W_{im}^2| = o_{P^*}(1)$$

where $E^*\frac{1}{n}\sum_i W_{im}^2 = \frac{1}{n}\sum_i(a_{im}^2+b_{im}^2)$. Also since $\widehat{e}_{it}=\widehat{\epsilon}_{it}+\widehat{\alpha}\widehat{\eta}_{it}$, we have

$$\frac{1}{n}\sum_i(\frac{1}{\sqrt{T}}\sum_t\widehat{U}_{it,m}\widehat{e}_{it})^2 = \frac{1}{n}\sum_i a_{im}^2+b_{im}^2+2a_{im}b_{im}.$$

Thus,

$$\max_m\left|\frac{1}{n}\sum_i(\frac{1}{\sqrt{T}}\sum_t\widehat{U}_{it,m}\widehat{e}_{it})^2-\frac{1}{n}\sum_i W_{im}^2\right| \le o_{P^*}(1)+\max_m\left|\frac{2}{n}\sum_i a_{im}b_{im}\right| \qquad \text{(H.43)}$$

It remains to show the second term is $o_P(1)$. It follows from the Cauchy-Schwarz inequality,

$$\max_m\frac{1}{n}\sum_i(\frac{1}{\sqrt{T}}\sum_t\widehat{U}_{it,m}\widehat{\epsilon}_{it}-\tilde{U}_{it,m}\tilde{\epsilon}_{it})^2 = o_P(1),$$

and

$$\max_m\frac{1}{n}\sum_i(\frac{1}{\sqrt{T}}\sum_t\widehat{U}_{it,m}\widehat{\alpha}\widehat{\eta}_{it}-\tilde{U}_{it,m}\alpha\tilde{\eta}_{it})^2 = o_P(1)$$

that

$$\max_m \left| \frac{1}{n}\sum_i a_{im}b_{im} - \frac{1}{n}\sum_i (\frac{1}{\sqrt{T}}\sum_t \tilde{U}_{it,m}\tilde{\epsilon}_{it})(\frac{1}{\sqrt{T}}\sum_t \tilde{U}_{it,m}\alpha\tilde{\eta}_{it}) \right| = o_P(1).$$

Also,

$$\max_m \left| \frac{1}{n}\sum_i (\frac{1}{\sqrt{T}}\sum_t \tilde{U}_{it,m}\tilde{\epsilon}_{it})(\frac{1}{\sqrt{T}}\sum_t \tilde{U}_{it,m}\alpha\tilde{\eta}_{it}) \right| = o_P(1) + \max_m \left| \frac{1}{n}\sum_i X_{im} \right|,$$

where $X_{im} = \left(\frac{1}{\sqrt{T}}\sum_t U_{it,m}\epsilon_{it}\right)\left(\frac{1}{\sqrt{T}}\sum_t U_{it,m}\alpha\eta_{it}\right)$. Now $EX_{im} = 0$ since $E(\epsilon|U,\eta) = 0$, and $\max_{im} \mathrm{Var}(X_{im}) < C$. Thus by the Bernstein inequality for independent data and union bounds, it follows that

$$\max_m \left| \frac{1}{n}\sum_i X_{im} \right| = o_P(1).$$

This proves that $\max_i \frac{1}{n}\sum_i a_{im}b_{im} = o_P(1)$, and therefore the left hand side of (H.43)) is $o_{P^*}(1)$, which also implies $G = o_{P^*}(1)$.

In view of (H.42), we have, $\|\widehat{\Psi}^{y-1}\frac{1}{nT}\sum_{t=1}^T \tilde{U}_t^{*\prime}\tilde{e}_t^*\|_\infty \leq (1+o_{P^*}(1))\frac{1}{\sqrt{nT}}\Phi^{-1}\left(1-\frac{q_n}{2p}\right)$.

(ii) We make use of equality (H.40). For any $w_{it,k}$,

$$\begin{aligned}
&\max_{m,k}\left|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T w_{it,k}(\tilde{U}_{it,m}^* - \widehat{U}_{it,m}^*)\right| \\
&\leq \max_{m,k}\left|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T w_{it,k}\widehat{f}_i^{*\prime}(\widehat{F}^{*\prime}\widehat{F}^*)^{-1}\widehat{F}^{*\prime}(\widehat{F}H^* - \widehat{F}^*)H^{*-1}\widehat{\lambda}_{tm}\right| \\
&\quad + \max_{m,k}\left|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T w_{it,k}\widehat{f}_i^{*\prime}(\widehat{F}^{*\prime}\widehat{F}^*)^{-1}\widehat{F}^{*\prime}\tilde{U}_{t,m}^*\right| \\
&\quad + \max_{m,k}\left|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T w_{it,k}\widehat{\lambda}_{tm}'H^{*\prime-1}(\widehat{f}_i^* - H^{*\prime}\widehat{f}_i)\right| \qquad \text{(H.44)} \\
&\leq \max_{m,k}\left\|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T w_{it,k}\widehat{f}_i^*\widehat{\lambda}_{tm}'\right\|_F O_{P^*}(\Delta_F^*) + \max_{m,k}\left\|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T w_{it,k}\frac{1}{n}\widehat{F}^{*\prime}\tilde{U}_{t,m}^*\widehat{f}_i^{*\prime}\right\|_F \\
&\quad + \max_{m,k}\left|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T w_{it,k}\widehat{\lambda}_{tm}'H^{*\prime-1}(\widehat{f}_i^* - H^{*\prime}\widehat{f}_i)\right|. \qquad \text{(H.45)}
\end{aligned}$$

Let $w_{it,k} = \tilde{e}^*_{it}$. Then (H.45) implies $\max_m |\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T \tilde{e}^*_{it}(\tilde{U}^*_{it,m} - \widehat{U}^*_{it,m})|$ is bounded by

$$\max_m \|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T \tilde{e}^*_{it}\widehat{f}_i\widehat{\lambda}'_{tm}\|_F O_{P^*}(\Delta^*_F) + \max_m \|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T \tilde{e}^*_{it}\frac{1}{n}\widehat{F}^{*\prime}\tilde{U}^*_{t,m}\widehat{f}^{*\prime}_i\|_F + O_{P^*}(\Delta^*_{eg}).$$

By Assumption D.5, $\Delta^*_{eg} = o(\sqrt{\frac{\log p}{nT}})$. By Lemma H.14, the first term is $O_{P^*}\left(\sqrt{\frac{\log(p)\log(np)}{nT}}\Delta^*_F\right)$, which is $o_{P^*}\left(\sqrt{\frac{\log(p)}{nT}}\right)$ given that $\log(np)\Delta^{*2}_F = o(1)$. By Lemma H.14, the second term satisfies

$$\max_m \left\|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T \tilde{e}^*_{it}\frac{1}{n}\widehat{F}^{*\prime}\tilde{U}^*_{t,m}\widehat{f}^{*\prime}_i\right\|_F = o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right).$$

Hence

$$\max_m \left|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T \tilde{e}^*_{it}(\tilde{U}^*_{it,m} - \widehat{U}^*_{it,m})\right| = o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right).$$

(iii) Note that $M^*_t = \widehat{F}\widehat{\delta}_{yt} - \widehat{F}^*\widehat{\delta}^*_{yt} + (\tilde{U}^*_t - \widehat{U}^*_t)\widehat{\gamma}_y$. Thus,

$$\begin{aligned}
&\left\|\frac{1}{nT}\sum_{t=1}^T M^{*\prime}_t\widehat{U}^*_t\right\|_\infty \\
&\leq \left\|\frac{1}{n}\sum_{i=1}^n\frac{1}{T}\sum_{t=1}^T \widehat{U}^*_{it}(\widehat{\delta}^*_{yt} - H^{*-1}\widehat{\delta}_{yt})'\widehat{f}^*_i\right\|_\infty \\
&\quad + \left\|\frac{1}{n}\sum_{i=1}^n\frac{1}{T}\sum_{t=1}^T \widehat{\delta}'_{yt}H^{*\prime-1}(\widehat{f}^*_i - H^{*\prime}\widehat{f}_i)\tilde{U}^*_{it}\right\|_\infty \\
&\quad + \left\|\frac{1}{n}\sum_{i=1}^n\frac{1}{T}\sum_{t=1}^T(\widehat{U}^*_{it} - \tilde{U}^*_{it})\widehat{\delta}'_{yt}H^{*\prime-1}(\widehat{f}^*_i - H^{*\prime}\widehat{f}_i)\right\|_\infty \\
&\quad + \left\|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T(\widehat{U}^*_{it} - \tilde{U}^*_{it})(\tilde{U}^*_{it} - \widehat{U}^*_{it})'\widehat{\gamma}_y\right\|_\infty \\
&\quad + \left\|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T\tilde{U}^*_{it}(\tilde{U}^*_{it} - \widehat{U}^*_{it})'\widehat{\gamma}_y\right\|_\infty .
\end{aligned}$$

The first term after the inequality in the preceding display is zero since $\widehat{F}^{*\prime}\widehat{U}^*_t = 0$. The second term is $O_{P^*}(\Delta^*_{ud}) = o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right)$ due to Assumption D.5. That the other three terms following the inequality are also $o_{P^*}\left(\sqrt{\frac{\log p}{nT}}\right)$ follows from Lemma H.16.

(iv) By Lemma H.15,

$$\max_{m,k\le p}|\frac{1}{nT}\sum_{it}\tilde U^*_{it,k}\widehat f_i\widehat\lambda_{tm}| = O_{P^*}\left(\sqrt{\frac{\log p}{n}}\right)\sqrt{\frac{\log(n)\log(pT)}{n}+\Delta^2_{\max}+\frac{\log(pn)}{T}}.$$

By Lemma H.14, $\max_m \frac{1}{T}\sum_t\|\frac{1}{n}\widehat F'\tilde U^*_{t,m}\|_2^2 = O_{P^*}\left(\frac{\log(pT)}{n}\right)$. Thus, we have

$$|J|_0\max_{m,k\le p}|\frac{1}{nT}\sum_{it}(\widehat U^*_{it,m}\widehat U^*_{it,k}-\tilde U^*_{it,m}\tilde U^*_{it,k})|\le A+B,$$

where

$$\begin{aligned}
A &:= |J|_0\max_{m,k\le p}\left|\frac{2}{nT}\sum_{it}\tilde U^*_{it,m}(\widehat U^*_{it,k}-\tilde U^*_{it,k})\right| \\
&\le |J|_0\max_{m,k\le p}\left|\frac{2}{nT}\sum_{it}\tilde U^*_{it,k}\widehat f_i\widehat\lambda_{tm}\right| O_{P^*}(\Delta^*_F) \\
&\quad + O_{P^*}|J|_0\max_m\frac{1}{T}\sum_t\left|\frac{1}{n}\widehat F^{*\prime}\tilde U^*_{t,m}\right|^2 + O_{P^*}(\Delta^*_{ud}|J|_0) \\
&= o_{P^*}(1) \qquad \text{(H.46)}\\
B &:= |J|_0\max_{m,k\le p}\left|\frac{1}{nT}\sum_{it}(\widehat U^*_{it,m}-\tilde U^*_{it,m})(\widehat U^*_{it,k}-\tilde U^*_{it,k})\right| \\
&\le |J|_0 O_{P^*}(\Delta^{*2}_F) + \max_m\frac{1}{T}\sum_{t=1}^T\left|\frac{1}{n}\widehat F'\tilde U^*_{t,m}\right|^2 O_{P^*}(|J|_0) \\
&= O_{P^*}\left(|J|_0\Delta^{*2}_F + |J|_0\frac{\log(pT)}{n}\right) = o_{P^*}(1).
\end{aligned}$$

(v) First note that for any β, by (H.46),

$$\begin{aligned}
\frac{1}{nT}\sum_{t=1}^T\|(\widehat U^*_t-\tilde U^*_t)\beta\|_2^2 &= \frac{1}{nT}\sum_{it}((\widehat U^*_{it}-\tilde U^*_{it})'\beta)^2 \\
&= \beta'\frac{1}{nT}\sum_{it}(\widehat U^*_{it}-\tilde U^*_{it})(\widehat U^*_{it}-\tilde U^*_{it})'\beta \qquad \text{(H.47)}\\
&\le \|\beta\|_1^2\max_{mk}\left|\frac{1}{nT}\sum_{it}(\widehat U^*_{it,m}-\tilde U^*_{it,m})(\widehat U^*_{it,k}-\tilde U^*_{it,k})\right| \\
&\le \|\beta\|_1^2 o_{P^*}(|J|_0^{-1}) \le o_{P^*}(1)\|\beta\|_2^2.
\end{aligned}$$

Note that the condition $\|\beta\|_1^2\le C\|\beta\|_2^2|J|_0$ is only used in the last inequality.

Secondly, define $X_{i,mk} = \frac{1}{T}\sum_t \widehat{U}_{it,m}\widehat{U}_{it,k}((w_i^U)^2 - E^*(w_i^U)^2)$. Then

$$\max_{mk} \frac{1}{n}\sum_i \mathrm{Var}^*(X_{i,mk}) \leq C \max_{mk} \frac{1}{n}\sum_i (\frac{1}{T}\sum_t \widehat{U}_{it,m}\widehat{U}_{it,k})^2 = O_P(1).$$

Hence $\max_{mk} |\frac{1}{n}\sum_i X_{i,mk}| = O_{P^*}(\sqrt{\frac{\log(p)}{n}})$ and

$$\begin{aligned}
&\left|\beta' \ \frac{1}{nT}\sum_{it} \widehat{U}_{it}\widehat{U}_{it}'((w_i^U)^2 - E^*(w_i^U)^2)\beta\right| \\
&\leq \|\beta\|_1^2 \max_{mk} \left|\frac{1}{nT}\sum_{it} \widehat{U}_{it,m}\widehat{U}_{it,k}((w_i^U)^2 - E^*(w_i^U)^2)\right| \\
&= \|\beta\|_1^2 \max_{mk} \left|\frac{1}{n}\sum_i X_{i,mk}\right| \\
&= \|\beta\|_1^2 O_{P^*}\left(\sqrt{\frac{\log(p)}{n}}\right).
\end{aligned}$$

Combining the above result with (H.47) and noting that

$$\frac{1}{nT}\sum_{t=1}^T \|\tilde{U}_t^*\beta\|_2^2 = \beta' \frac{1}{nT}\sum_{it} \widehat{U}_{it}\widehat{U}_{it}'(w_i^U)^2\beta,$$

we obtain, for any β,

$$\begin{aligned}
\frac{1}{nT}\sum_{t=1}^T \|\widehat{U}_t^*\beta\|_2^2 &\geq \frac{1}{nT}\sum_{t=1}^T \|\tilde{U}_t^*\beta\|_2^2 - o_{P^*}(|J|_0^{-1})\|\beta\|_1^2 \\
&\geq \beta' \frac{1}{nT}\sum_{it} \widehat{U}_{it}\widehat{U}_{it}'\beta - \|\beta\|_1^2 \left(o_{P^*}(|J|_0^{-1}) + O_{P^*}\left(\sqrt{\frac{\log(p)}{n}}\right)\right) \\
&\geq^{(a.1)} \frac{1}{nT}\sum_{t=1}^T \|\tilde{U}_t\beta\|_2^2 - \|\beta\|_1^2 \left(o_{P^*}(|J|_0^{-1}) + o_P(|J|_0^{-1})\right) \\
&\geq^{(a.2)} \frac{1}{nT}\sum_{t=1}^T \|\tilde{U}_t\beta\|_2^2 - \|\beta\|_2^2 o_{P^*}(1),
\end{aligned} \tag{H.48}$$

where (a.1) follows from Lemma F.1 (v) that

$$\max_{m,k\leq p} |\frac{1}{nT}\sum_{it}(\widehat{U}_{it,m}\widehat{U}_{it,k} - \tilde{U}_{it,m}\tilde{U}_{it,k})| = o_P(|J|_0^{-1})$$

and (a.2) follows from $\|\beta\|_1^2 \leq C\|\beta\|_2^2|J|_0$. Note that we have shown

$$|J|_0 \max_{m,k\leq p}\left|\frac{1}{nT}\sum_{it}(\widehat{U}^*_{it,m}\widehat{U}^*_{it,k} - \tilde{U}_{it,m}\tilde{U}_{it,k})\right| = o_{P^*}(1). \tag{H.49}$$

■

Appendix I. Technical lemmas for the estimated factors

I.1. **Proof of Proposition D.1 (for $\widehat{F}$ using the original data).** This section verifies Assumption D.4 when factors are estimated using PCA.

(i) By Assumption D.1, it can be shown that $\|H\| = O_P(1) = \|V^{-1}\|$. In addition, we have the following identity:

$$\widehat{f}_i - H'\tilde{f}_i = V^{-1}\sum_{l=1}^{5} A_{il},$$

where

$$\begin{aligned}
A_{i1} &= \frac{1}{pTn}\sum_{j=1}^{n}\widehat{f}_j\sum_{t=1}^{T}(\tilde{U}_{it}'\tilde{U}_{jt} - U_{it}'U_{jt}),\\
A_{i2} &= \frac{1}{pTn}\sum_{j=1}^{n}\widehat{f}_j\sum_{t=1}^{T}(U_{it}'U_{jt} - EU_{it}'U_{jt}),\\
A_{i3} &= \frac{1}{pTn}\sum_{j=1}^{n}\widehat{f}_j\sum_{t=1}^{T}EU_{it}'U_{jt},\\
A_{i4} &= \frac{1}{pTn}\sum_{j=1}^{n}\widehat{f}_j\sum_{t=1}^{T}\tilde{f}_j'\tilde{\Lambda}_t'\tilde{U}_{it},\\
A_{i5} &= \frac{1}{pTn}\sum_{j=1}^{n}\widehat{f}_j\sum_{t=1}^{T}\tilde{f}_i'\tilde{\Lambda}_t'\tilde{U}_{jt}.
\end{aligned} \tag{I.1}$$

Each term can be written in the form $A_{il} = \frac{1}{pTn}\sum_{j=1}^{n}\widehat{f}_j\sum_{t=1}^{T}B_{ijt,l}$. By Cauchy-Schwarz,

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n}\|\widehat{f}_i - H'\tilde{f}_i\|_2^2 &= O_P(1)\sum_{l=1}^{5}\frac{1}{n}\sum_{i=1}^{n}\|A_{il}\|_2^2\\
&\leq O_P(1)\sum_{l=1}^{5}\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{1}{pT}\sum_{t=1}^{T}B_{ijt,l}\right)^2.
\end{aligned} \tag{I.2}$$

Note that we used $\frac{1}{n}\sum_{i=1}^n\|\widehat{f}_i\|_2^2=K=O(1)$ because the columns of $n^{-1/2}(\widehat{f}_1,...,\widehat{f}_K)'$ are defined as the eigenvectors of $\tilde{X}'\tilde{X}$. We bound the terms in (I.2) in Lemmas I.6 and I.7 below. Then, applying the bounds in Lemmas I.6 and I.7 and using $T=o(n)$, we have

$$\frac{1}{n}\sum_{i=1}^n\|\widehat{f}_i-H'\tilde{f}_i\|_2^2=O_P(\Delta_F^2),\quad \Delta_F^2=\frac{1}{n^2}+\frac{1}{nT^2}+\frac{1}{pT}.$$

It is then straightforward to verify that $|J|_0^2\Delta_F^2=o\left(\sqrt{\frac{1}{nT}}\right)$ holds when $|J|_0^4=o(nT^3)$, $|J|_0^4n=o(p^2T)$, $|J|_0^2\sqrt{\log p}\log(pT)=o(n)$, and $|J|_0^2T=o(n)$. For example, to show $|J|_0^2\sqrt{\log p}\log(T)=o(n)$, note that $|J|_0=o(\sqrt{n/(\log p)})$ and $|J|_0=o(\sqrt{n/T})$ implies

$$|J|_0^2\sqrt{\log p}\log(T)=o(n\log T\sqrt{\log p}/\sqrt{T\log p})=o(n).$$

(ii) We now verify that we can produce sequences Δ_{eg} so that $\Delta_{eg}=o\left(\frac{1}{\sqrt{nT}}\right)$. First, note that we can set $g_{tm}\in\{\gamma_d'\tilde{\Lambda}_t,\tilde{\lambda}_{tm},\tilde{\delta}_t\}$ in applying Lemma I.2, each of which yields $\omega_n=O(|J|_0^2)$ for ω_n defined in Lemma I.2. It then follows from Lemma I.2 that we can take $\Delta_{eg}=(\frac{1}{\sqrt{npT}}+\frac{1}{n})|J|_0$ so that $\Delta_{eg}=o\left(\sqrt{\frac{1}{nT}}\right)$, given $T|J|_0^2=o(n)$ and $|J|_0^2=o(p)$. Note that $|J|_0^2=o(p)$ is implied by the assumption that $|J|_0^4n=o(p^2T)$.

(iii) By Lemma I.10, we take $\Delta_{fum}=\frac{1}{n}+\frac{1}{T\sqrt{n}}+\frac{1}{T\sqrt{p}}+\sqrt{\frac{\log(pT)}{npT}}$, and $\Delta_{fe}^2=\frac{1}{n^2}+\frac{1}{Tpn}$. We then have

$$\max_{s\le T,m\le p}|\frac{1}{n}\sum_{i=1}^n(H'\tilde{f}_i-\widehat{f}_i)\tilde{U}_{is,m}|=O_P(\Delta_{fum})$$

and

$$\frac{1}{T}\sum_{t=1}^T\|\frac{1}{n}\sum_{i=1}^n\tilde{e}_{it}(\widehat{f}_i-H'\tilde{f}_i)\|_2^2=O_P(\Delta_{fe}^2).$$

Then, it is straightforward to check $\Delta_{fum}^2=o\left(\frac{\log p}{T|J|^2\log(pT)}\right)$ and $\Delta_{fe}^2=o\left(\frac{\log p}{T\log(pT)}\right)$.

(iv) By Lemma I.3, we can define $\Delta_{ud}=\frac{1}{\sqrt{n}}\sqrt{\frac{\log(pT)}{nT}}+\sqrt{\frac{\log p}{nT}}(\frac{1}{\sqrt{pT}}+\frac{1}{T\sqrt{n}})+\frac{1}{pT}$. Given $|J|_0^4n=o(p^2T)$ and $|J|_0^4=o(nT^3)$, it is straightforward to verify that $\Delta_{ud}=o\left(\sqrt{\frac{\log p}{nT}}\right)$ and $|J|_0^2\sqrt{\log p}\Delta_{ud}=o(1)$. This result follows by verifying $\sqrt{\frac{\log p}{nT}}\frac{1}{\sqrt{pT}}|J|_0^2\sqrt{\log p}=o(1)$ which can be shown by noting that $|J|_0^4n=o(p^2T)$ implies $|J_0|^2=o(p\sqrt{T/n})$ and that $|J|_0^2\log(p)=o(n)$. Thus, because $\log^2p=o(n)$,

$$\left(\sqrt{\frac{\log p}{nT}}\frac{|J|_0^2\sqrt{\log p}}{\sqrt{pT}}\right)^2=\frac{\log^2p|J|_0^4}{npT^2}=o\left(\frac{\log(p)pn\sqrt{T}}{npT^2\sqrt{n}}\right)=o\left(\frac{\log(p)\sqrt{T}}{T^2\sqrt{n}}\right)=o(1).$$

(v) First, by Lemma I.1, we can take $\Delta_{\max} = \frac{1}{\sqrt{n}} + \sqrt{\frac{\log n}{Tp}}$. Also $\Delta_F^2 = \frac{1}{n^2} + \frac{1}{nT^2} + \frac{1}{pT}$. This implies $\Delta_F^2 |J|_0^2 + \frac{|J|_0}{n} = O\left(\frac{|J|_0}{n} + \frac{|J|_0^2}{nT^2} + \frac{|J|_0^2}{pT}\right)$. In addition, $\kappa_n^2 |J|_0 \sqrt{nT} = o(1)$ and $\|R_y\|_1^2 = o\left(\frac{\log p}{nT}\right)$ imply $\lambda_n^2 |J|_0 + \|R_y\|_1^2 = O\left(\frac{1}{\sqrt{nT}}\right)$. Hence with the conditions $|J|_0^4 = o(nT^3)$, $|J|_0^4 n = o(p^2 T)$, we have

$$\lambda_n^2 |J|_0 + \|R_y\|_1^2 + \Delta_F^2 |J|_0^2 + \frac{|J|_0}{n} = O\left(\frac{1}{\sqrt{nT}}\right).$$

Thus, in order to verify $\Delta_{\max}^2 |J|_0^2 T(\lambda_n^2 |J|_0 + \|R_y\|_1^2 + \Delta_F^2 |J|_0^2 + \frac{|J|_0}{n}) = o(1)$, it suffices to verify

$$\left(\frac{1}{n} + \frac{\log n}{Tp}\right) |J|_0^2 T \frac{1}{\sqrt{nT}} = o(1),$$

which holds given the conditions $|J|_0^2 T = o(n)$ and $|J|_0^2 = o(p)$. Note that $|J|_0^2 = o(p)$ is implied by $|J|_0^4 n = o(p^2 T)$. ■

Lemma I.1.

$$\max_i \|\widehat{f}_i - H' \tilde{f}_i\|_2 = O_P\left(\frac{1}{\sqrt{n}} + \sqrt{\frac{\log n}{Tp}}\right) := O_P(\Delta_{\max}^2).$$

Proof. By Assumption D.3, $\max_i \frac{1}{pT} \sum_{tm} (U_{it,m}^2 - EU_{it,m}^2) = O_P(1)$. Also, $\max_{it,m} EU_{it,m}^2 = O(1)$. Hence $\max_i \frac{1}{pT} \sum_{tm} U_{it,m}^2 = O_P(1)$. So $\max_i \|\widehat{f}_i - H' \tilde{f}_i\|_2 \leq \sum_{l=1}^3 G_l$ where each G_l is defined and bounded below. Specifically,

$$\begin{aligned}
G_1 &= \max_i \|\frac{1}{pTn} \sum_{t=1}^T \tilde{U}_{it}' \sum_{j=1}^n \tilde{U}_{jt} \widehat{f}_j\|_2 \\
&\leq \max_i \|\frac{1}{pTn} \sum_{t=1}^T U_{it}' \sum_{j=1}^n \tilde{U}_{jt} \widehat{f}_j\|_2 + \|\frac{1}{pTn} \sum_{t=1}^T \bar{U}_{\cdot t}' \sum_{j=1}^n \tilde{U}_{jt} \widehat{f}_j\|_2 \\
&\leq \left(2 \max_i \frac{1}{pT} \sum_{tm} U_{it,m}^2 + 2 \frac{1}{pT} \sum_{tm} \bar{U}_{\cdot t,m}^2\right)^{1/2} \left(\frac{1}{pT} \sum_{tm} \|\frac{1}{n} \sum_j \tilde{U}_{jt,m} \widehat{f}_j\|_2^2\right)^{1/2} \\
&= O_P\left(O_P(1)\left(\frac{1}{T\sqrt{p}} + \frac{1}{\sqrt{n}}\right)\right)
\end{aligned}$$

by equation (I.5) given below in Lemma I.2.

$$G_2 = \max_i \|\frac{1}{pTn} \sum_{j=1}^n \widehat{f}_j \sum_{t=1}^T \tilde{f}_j' \tilde{\Lambda}_t' \tilde{U}_{it}\|_2$$

$$
\begin{aligned}
&\leq O_P(1)\max_i \|\frac{1}{pT}\sum_{t=1}^{T}\tilde{\Lambda}_t' U_{it}\|_2 + O_P(1)\|\frac{1}{pT}\sum_{t=1}^{T}\tilde{\Lambda}_t'\bar{U}_{\cdot t}\|_2 \\
&= O_P\left(\sqrt{\frac{\log n}{Tp}}\right)
\end{aligned}
$$

by Assumption D.3 that $\max_{i\leq n}\|\frac{1}{pT}\sum_{t=1}^{T}\sum_{m=1}^{p}\tilde{\lambda}_{t,m}U_{it,m}\|_2 = O_P(\sqrt{\frac{\log n}{pT}})$. Finally, recall that $\|A\|_F^2 = \mathrm{tr}(A'A)$ and $\|A\|_2 \leq \|A\|_F$ for any matrix A,

$$
\begin{aligned}
G_3 &= \max_i \|\frac{1}{pTn}\sum_{t=1}^{T}\tilde{f}_i'\tilde{\Lambda}_t'\sum_{j=1}^{n}\tilde{U}_{jt}\widehat{f}_j'\|_2 \\
&\leq \max_i \|\tilde{f}_i\|_2 \|\frac{1}{pTn}\sum_{t=1}^{T}\tilde{\Lambda}_t'\sum_{j=1}^{n}\tilde{U}_{jt}\widehat{f}_j'\|_F \\
&\leq O_P(\sqrt{\log n})\|\frac{1}{pTn}\sum_{t=1}^{T}\tilde{\Lambda}_t'\sum_{j=1}^{n}\tilde{U}_{jt}\tilde{f}_j'\|_F \\
&\quad + O_P(\sqrt{\log n})\|\frac{1}{pTn}\sum_{t=1}^{T}\tilde{\Lambda}_t'\sum_{j=1}^{n}\tilde{U}_{jt}(\widehat{f}_j - H'\tilde{f}_j)'\|_F \\
&= O_P\left(\sqrt{\frac{\log n}{pTn}}\right) + O_P(\sqrt{\log n}\Delta_F)\left(\frac{1}{n}\sum_j \|\frac{1}{pT}\sum_t \tilde{\Lambda}_t'\tilde{U}_{jt}\|_2^2\right)^{1/2} \\
&= O_P\left(\sqrt{\frac{\log n}{pTn}} + \Delta_F\sqrt{\frac{\log n}{pT}}\right).
\end{aligned}
$$

Hence, $\max_i \|\widehat{f}_i - H'\tilde{f}_i\|_2 = O_P\left(\frac{1}{\sqrt{n}} + \sqrt{\frac{\log n}{Tp}}\right)$. ■

Lemma I.2. *Let $\{z_{it}\}$ be a random sequence with $E(z_{it}|f_t, U_t) = 0$ and $\mathrm{Var}(z_{it}) > 0$. In addition, let $\{g_{tm}\}$ be a deterministic sequence of vectors with a fixed dimension, $m \leq p$. Then for $\tilde{z}_{it} = z_{it} - \bar{z}_{i\cdot} - \bar{z}_{\cdot t} + \bar{\bar{z}}$, and $\omega_n = \max_{m\leq p}\frac{1}{T}\sum_{t=1}^{T}\|g_{tm}\|_2^2$,*

$$
\max_{m\leq p}\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}g_{tm}\tilde{z}_{it}(\widehat{f}_i - H'\tilde{f}_i)'\|_F = O_P\left(\frac{1}{\sqrt{npT}} + \frac{1}{n}\right)\omega_n^{1/2}.
$$

Proof. It follows from equation (I.1) that

$$
\max_{m\leq p}\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}g_{tm}\tilde{z}_{it}(\widehat{f}_i - H'\tilde{f}_i)'\|_F \leq \sum_{l=1}^{3}\bar{C}_l O_P(1),
$$

where each term $\bar{C}_l$ is defined and bounded in below.

$$\begin{aligned}
\bar{C}_1 &= \max_{m\leq p}\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{1}{pTn}\sum_{j=1}^{n}\widehat{f}_j\sum_{s=1}^{T}\tilde{U}_{is}'\tilde{U}_{js}\tilde{z}_{it}g_{tm}'\|_F \\
&\leq \max_{m\leq p}\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{1}{pTn}\sum_{j=1}^{n}(\widehat{f}_j - H'\tilde{f}_j)\sum_{s=1}^{T}\tilde{U}_{is}'\tilde{U}_{js}\tilde{z}_{it}g_{tm}'\|_F \\
&\quad + O_P(1)\max_{m\leq p}\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{1}{pTn}\sum_{j=1}^{n}\tilde{f}_j\sum_{s=1}^{T}\tilde{U}_{is}'\tilde{U}_{js}\tilde{z}_{it}g_{tm}'\|_F \\
&\leq (x_n\omega_n)^{1/2}\left(\left(\frac{1}{npT}\|\tilde{U}\|_F^2\right)^{1/2}O_P(\Delta_F) + O_P(1)(\frac{1}{pT}\sum_{s}\|\frac{1}{n}\sum_{j=1}^{n}\tilde{f}_j\tilde{U}_{js}'\|_F^2)^{1/2}\right) \\
&= (x_n\omega_n)^{1/2}O_P(\Delta_F + \frac{1}{\sqrt{n}}) \\
&= \omega_n^{1/2}O_P\left(\frac{1}{\sqrt{npT}} + \frac{1}{n}\right).
\end{aligned}$$

where $x_n = \frac{1}{pT^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\|\frac{1}{n}\sum_{i=1}^{n}\tilde{z}_{it}\tilde{U}_{is}\|_2^2 = O_P(\frac{1}{n})$ and $\omega_n = \max_{m\leq p}\frac{1}{T}\sum_{t=1}^{T}\|g_{tm}\|_2^2$. Next,

$$\begin{aligned}
\bar{C}_2 &= \max_{m\leq p}\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{1}{pTn}\sum_{j=1}^{n}\widehat{f}_j\sum_{s=1}^{T}\tilde{f}_j'\tilde{\Lambda}_s'\tilde{U}_{is}\tilde{z}_{it}g_{tm}'\|_F \\
&\leq O_P(\omega_n^{1/2})\left(\frac{1}{T}\sum_{t=1}^{T}\|\frac{1}{pT}\frac{1}{n}\sum_{i=1}^{n}\sum_{s=1}^{T}\tilde{\Lambda}_s'\tilde{U}_{is}\tilde{z}_{it}\|_2^2\right)^{1/2} \\
&= O_P\left(\omega_n^{1/2}\frac{1}{\sqrt{npT}}\right)
\end{aligned}$$

by Lemma I.8). Finally,

$$\begin{aligned}
\bar{C}_3 &= \max_{m\leq p}\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{1}{pTn}\sum_{j=1}^{n}\widehat{f}_j\sum_{s=1}^{T}\tilde{f}_i'\tilde{\Lambda}_s'\tilde{U}_{js}\tilde{z}_{it}g_{tm}'\|_F \\
&\leq \max_{m\leq p}\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{1}{pTn}\sum_{j=1}^{n}(\widehat{f}_j - H'\tilde{f}_j)\sum_{s=1}^{T}\tilde{f}_i'\tilde{\Lambda}_s'\tilde{U}_{js}\tilde{z}_{it}g_{tm}'\|_F \\
&\quad + O_P(1)\max_{m\leq p}\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\frac{1}{pTn}\sum_{j=1}^{n}\tilde{f}_j\sum_{s=1}^{T}\tilde{f}_i'\tilde{\Lambda}_s'\tilde{U}_{js}\tilde{z}_{it}g_{tm}'\|_F
\end{aligned}$$

$$\leq (\omega_n c_n)^{1/2}\left(O_P(\Delta_F)(\frac{1}{n}\sum_j \|\frac{1}{pT}\sum_{s=1}^T \tilde\Lambda_s'\tilde U_{js}\|_2^2)^{1/2} + O_P(1)\max_{m\leq p}\|\frac{1}{pTn}\sum_{j=1}^n\sum_{s=1}^T \tilde\Lambda_s'\tilde U_{js}\tilde f_j'\|_F\right)$$
$$= \omega_n^{1/2} O_P\left(\frac{1}{\sqrt{npT}} + \frac{1}{n}\right)$$

where $c_n = \frac{1}{T}\sum_{t=1}^T \|\frac{1}{n}\sum_{i=1}^n \tilde f_i \tilde z_{it}\|_2^2 = O_P(\frac{1}{n})$. Therefore,

$$\max_{m\leq p}\|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T g_{tm}\tilde z_{it}(\widehat f_i - H'\tilde f_i)'\|_F = O_P\left(\frac{1}{\sqrt{npT}} + \frac{1}{n}\right)\omega_n^{1/2}.$$

■

Lemma I.3. *Let* $\{g_{tk} : k \leq p\}$ *be a deterministic sequence of vectors of fixed dimension with* $\max_{tk}\|g_{tk}\|_2 = O(1)$*, and let* $\tilde g_{tk} = g_{tk} - \frac{1}{T}\sum_t g_{tk}$*. Then*

$$\max_{m,k\leq p}\|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T(\widehat f_i - H'\tilde f_i)\tilde U_{it,m}\tilde g_{tk}'\|_F = O_P\left(\frac{1}{\sqrt n}\sqrt{\frac{\log(pT)}{nT}} + \sqrt{\frac{\log p}{nT}}(\frac{1}{\sqrt{pT}} + \frac{1}{T\sqrt n}) + \frac{1}{pT}\right)$$
$$:= O_P(\Delta_{ud}).$$

Proof. First, note that $\sum_t \tilde g_{tk} = 0$. Hence,

$$\max_{m,k\leq p}\|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T(\widehat f_i - H'\tilde f_i)\tilde U_{it,m}\tilde g_{tk}'\|_F$$
$$\leq \max_{m,k\leq p}\|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T(\widehat f_i - H'\tilde f_i)U_{it,m}\tilde g_{tk}'\|_F \tag{I.3}$$
$$+ \|\frac{1}{n}\sum_{i=1}^n(\widehat f_i - H'\tilde f_i)\|_2 \max_{m,k\leq p}\|\frac{1}{T}\sum_{t=1}^T \bar U_{\cdot t,m}\tilde g_{tk}'\|_F. \tag{I.4}$$

Term (I.4) is $O_P\left(\Delta_F\sqrt{\frac{\log p}{nT}}\right)$. Term (I.3) is bounded by $\sum_{l=1}^7 C_l$, where each C_l is defined and bounded blow.

First, note that applying Lemma I.10 gives

$$\begin{aligned}
\frac{1}{\sqrt{p}}\left(\frac{1}{T}\sum_t \|\frac{1}{n}\sum_j \widehat{f}_j \tilde{U}_{jt}'\|_F^2\right)^{1/2} &\leq \frac{2}{\sqrt{p}}\left(\frac{1}{T}\sum_t \|\frac{1}{n}\sum_j (\widehat{f}_j - H'\tilde{f}_j)\tilde{U}_{jt}'\|_F^2\right)^{1/2} \\
&\quad + \frac{2}{\sqrt{p}}\left(\frac{1}{T}\sum_t \|\frac{1}{n}\sum_j H'\tilde{f}_j \tilde{U}_{jt}'\|_F^2\right)^{1/2} \\
&= O_P\left(\frac{1}{n} + \frac{1}{T\sqrt{n}} + \frac{1}{T\sqrt{p}} + \sqrt{\frac{\log(pT)}{npT}} + \frac{1}{\sqrt{n}}\right) \qquad \text{(I.5)} \\
&= O_P(\frac{1}{T\sqrt{p}} + \frac{1}{\sqrt{n}}) \text{ and} \\
\max_{tl}\|\frac{1}{n}\sum_{j=1}^n \widehat{f}_j \tilde{U}_{jt,l}\|_2 &= O_P\left(\sqrt{\frac{\log(pT)}{n}} + \frac{1}{T\sqrt{p}}\right).
\end{aligned}$$

We then have, up to an $\|V^{-1}\| = O_P(1)$ term,

$$\begin{aligned}
C_1 &= \max_{m,k\leq p}\|\frac{1}{nT}\sum_{i=1}^n\sum_{s=1}^T \frac{1}{pTn}\sum_{j=1}^n \widehat{f}_j \sum_{t=1}^T \tilde{U}_{jt}'(\tilde{U}_{it} - U_{it})U_{is,m}\tilde{g}_{sk}'\|_F \\
&= \max_{m,k\leq p}\|\frac{1}{nT}\sum_{i=1}^n\sum_{s=1}^T \frac{1}{pTn}\sum_{j=1}^n \widehat{f}_j \sum_{t=1}^T \tilde{U}_{jt}'\bar{U}_{\cdot t}U_{is,m}\tilde{g}_{sk}'\|_F \text{ (because } \frac{1}{T}\sum_t \tilde{U}_{jt} = 0) \\
&\leq \left(\frac{1}{T}\sum_t \|\frac{1}{n}\sum_j \widehat{f}_j \tilde{U}_{jt}'\|_F^2\right)^{1/2}\left(\frac{1}{T}\sum_t \|\bar{U}_{\cdot t}\|_2^2\right)^{1/2}\max_{mk}\|\frac{1}{nTp}\sum_{i=1}^n\sum_{s=1}^T U_{is,m}\tilde{g}_{sk}'\|_2 \\
&= O_P\left(\frac{1}{T\sqrt{p}} + \frac{1}{\sqrt{n}}\right)\frac{1}{\sqrt{n}}\sqrt{\frac{\log p}{nT}} \\
&= O_P\left(\frac{1}{T\sqrt{pn}} + \frac{1}{n}\right)\sqrt{\frac{\log p}{nT}}.
\end{aligned}$$

$$
\begin{aligned}
C_2 &= \max_{m,k\le p}\left\|\frac{1}{nT}\sum_{i=1}^n\sum_{s=1}^T\frac{1}{pTn}\sum_{j=1}^n\widehat{f}_j\sum_{t=1}^T\tilde{U}_{jt}'(U_{it}U_{is,m}-EU_{it}U_{is,m})\tilde{g}_{sk}'\right\|_F \\
&\le\left(\frac{1}{T}\sum_t\|\frac{1}{n}\sum_j\widehat{f}_j\tilde{U}_{jt}'\|_F^2\right)^{1/2}\max_{mkt}\|\frac{1}{nTp}\sum_{i=1}^n\sum_{s=1}^T(U_{it}U_{is,m}-EU_{it}U_{is,m})\tilde{g}_{sk}'\|_F \qquad \text{(I.6)}\\
&= O_P\left(\frac{1}{T\sqrt{p}}+\frac{1}{\sqrt{n}}\right)\left(\sqrt{\frac{\log(pT)}{nT}}\right)
\end{aligned}
$$

using Lemma I.9.

$$
\begin{aligned}
C_3 &= \max_{m,k\le p}\left\|\frac{1}{nT}\sum_{i=1}^n\sum_{s=1}^T\frac{1}{pTn}\sum_{j=1}^n\widehat{f}_j\sum_{t=1}^T\tilde{U}_{jt}'(EU_{it}U_{is,m})\tilde{g}_{sk}'\right\|_F \\
&\le\max_{itm}\sum_{s=1}^T\sum_{l=1}^p|(EU_{it,l}U_{is,m})|\max_{sk}\|\tilde{g}_{sk}\|_2\frac{1}{pT}\max_{tl}\|\frac{1}{n}\sum_{j=1}^n\widehat{f}_j\tilde{U}_{jt,l}\|_2 \\
&= O_P\left(\frac{1}{pT}\right)O_P\left(\sqrt{\frac{\log(pT)}{n}}+\frac{1}{T\sqrt{p}}\right).
\end{aligned}
$$

$$
\begin{aligned}
C_4 &= \max_{m,k\le p}\left\|\frac{1}{nT}\sum_{i=1}^n\sum_{s=1}^T\frac{1}{pTn}\sum_{j=1}^n\widehat{f}_j\sum_{t=1}^T\tilde{f}_j'\tilde{\Lambda}_t'(\tilde{U}_{it}-U_{it})U_{is,m}\tilde{g}_{sk}'\right\|_F \\
&= \max_{m,k\le p}\left\|\frac{1}{nT}\sum_{i=1}^n\sum_{s=1}^T\frac{1}{pTn}\sum_{j=1}^n\widehat{f}_j\sum_{t=1}^T\tilde{f}_j'\tilde{\Lambda}_t'\bar{U}_{\cdot t}U_{is,m}\tilde{g}_{sk}'\right\|_F \quad \left(\text{because }\sum_t\tilde{\Lambda}_t=0\right) \\
&\le O_P(1)\left\|\frac{1}{pT}\sum_{t=1}^T\tilde{\Lambda}_t'\bar{U}_{\cdot t}\right\|_2\max_{m,k\le p}\left\|\frac{1}{nT}\sum_{i=1}^n\sum_{s=1}^T U_{is,m}\tilde{g}_{sk}'\right\|_2 \\
&= O_P\left(\frac{1}{\sqrt{npT}}\right)\sqrt{\frac{\log p}{nT}}.
\end{aligned}
$$

$$
\begin{aligned}
C_5 &= \max_{m,k\le p}\left\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}\frac{1}{pTn}\sum_{j=1}^{n}\widehat{f}_j\sum_{t=1}^{T}\tilde{f}_j'\tilde{\Lambda}_t'(U_{it}U_{is,m}-EU_{it}U_{is,m})\tilde{g}_{sk}'\right\|_F \\
&\le O_P(1)\max_{m,k\le p}\left\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}\frac{1}{pT}\sum_{t=1}^{T}\tilde{\Lambda}_t'(U_{it}U_{is,m}-EU_{it}U_{is,m})\tilde{g}_{sk}'\right\|_F \qquad \text{(I.7)}\\
&= O_P\left(\sqrt{\frac{\log p}{npT^2}}\right)
\end{aligned}
$$

using the same proof as that of Lemma I.9 (ii).

$$
\begin{aligned}
C_6 &= \max_{m,k\le p}\left\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}\frac{1}{pTn}\sum_{j=1}^{n}\widehat{f}_j\sum_{t=1}^{T}\tilde{f}_j'\tilde{\Lambda}_t'(EU_{it}U_{is,m})\tilde{g}_{sk}'\right\|_F \\
&\le O_P(1)\max_{m,k\le p}\left\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}\frac{1}{pT}\sum_{t=1}^{T}\tilde{\Lambda}_t'(EU_{it}U_{is,m})\tilde{g}_{sk}'\right\|_F \\
&= O_P\left(\frac{1}{pT}\right).
\end{aligned}
$$

$$
\begin{aligned}
C_7 &= \max_{m,k\le p}\left\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}\frac{1}{pTn}\sum_{j=1}^{n}\widehat{f}_j\sum_{t=1}^{T}\tilde{f}_i'\tilde{\Lambda}_t'\tilde{U}_{jt}U_{is,m}\tilde{g}_{sk}'\right\|_F \\
&\le \left\|\frac{1}{pTn}\sum_{j=1}^{n}\sum_{t=1}^{T}\widehat{f}_j\tilde{U}_{jt}'\tilde{\Lambda}_t\right\|_2 \max_{m,k\le p}\left\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}\tilde{f}_iU_{is,m}\tilde{g}_{sk}'\right\|_F \\
&\le O_P\left(\sqrt{\frac{\log p}{nT}}\right)\left(\|\frac{1}{pTn}\sum_{j=1}^{n}\sum_{t=1}^{T}\tilde{f}_j\tilde{U}_{jt}'\tilde{\Lambda}_t\|_2+\|\frac{1}{n}\sum_{j=1}^{n}(\widehat{f}_j-H'\tilde{f}_j)\tilde{U}_{jt,m}\|_2\right) \\
&= O_P\left(\sqrt{\frac{\log p}{nT}}\right)\left(\frac{1}{n}+\frac{1}{T\sqrt{n}}+\frac{1}{T\sqrt{p}}+\sqrt{\frac{\log(pT)}{npT}}\right).
\end{aligned}
$$

Combining the above, we reach

$$
\begin{aligned}
&\max_{m,k\le p}\left\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(\widehat{f}_i-H'\tilde{f}_i)\tilde{U}_{it,m}\tilde{g}_{tk}'\right\|_F \\
&\qquad = O_P\left(\frac{1}{\sqrt{n}}\sqrt{\frac{\log(pT)}{nT}}+\sqrt{\frac{\log p}{nT}}(\frac{1}{\sqrt{pT}}+\frac{1}{T\sqrt{n}}+\Delta_F)+\frac{1}{pT}\right)
\end{aligned}
$$

$$= O_P\left(\frac{1}{\sqrt{n}}\sqrt{\frac{\log(pT)}{nT}} + \sqrt{\frac{\log p}{nT}}(\frac{1}{\sqrt{pT}} + \frac{1}{T\sqrt{n}}) + \frac{1}{pT}\right).$$

■

I.2. **Proof of Proposition D.1 (for $\widehat{F}^*$ using the bootstrap data).** This section verifies Assumption D.5 when factors are estimated using PCA.

(i) Similar to (I.1), it can be proven that there is $\|H^*\| = O_{P^*}(1)$ and $\|V^{*-1}\| = O_{P^*}(1)$such that

$$\widehat{f}_i^* - H^{*'}\widehat{f}_i = V^{*-1}\sum_{l=1}^{4} A_{il}^*,$$

where

$$A_{i1}^* = \frac{1}{pTn}\sum_{j=1}^{n}\widehat{f}_j^*\sum_{t=1}^{T}(\tilde{U}_{it}^{*'}\tilde{U}_{jt}^* - E^*\tilde{U}_{it}^{*'}\tilde{U}_{jt}^*), \quad A_{i2}^* = \frac{1}{pTn}\sum_{j=1}^{n}\widehat{f}_j^*\sum_{t=1}^{T}E^*\tilde{U}_{it}^{*'}\tilde{U}_{jt}^*,$$
$$A_{i3}^* = \frac{1}{pTn}\sum_{j=1}^{n}\widehat{f}_j^*\sum_{t=1}^{T}\widehat{f}_j'\widehat{\Lambda}_t'\tilde{U}_{it}^*, \qquad A_{i4}^* = \frac{1}{pTn}\sum_{j=1}^{n}\widehat{f}_j^*\sum_{t=1}^{T}\widehat{f}_i'\widehat{\Lambda}_t'\tilde{U}_{jt}^*. \tag{I.8}$$

We first treat $A_{i2}^* - A_{i4}^*$. Because $\tilde{U}_{it}^*$ and $\tilde{U}_{jt}^*$ are independent if $i \neq j$, we have

$$\begin{aligned}\frac{1}{n}\sum_i \|A_{i2}^*\|_2^2 &= \frac{1}{n}\sum_i \|\frac{1}{pTn}\sum_{j=1}^{n}\widehat{f}_j^*\sum_{t=1}^{T}E^*\tilde{U}_{it}^{*'}\tilde{U}_{jt}^*\|_2^2 \\ &= \frac{1}{n}\sum_i \|\frac{1}{pTn}\widehat{f}_i^*\sum_{t=1}^{T}E^*\tilde{U}_{it}^{*'}\tilde{U}_{it}^*\|_2^2 \\ &= O_{P^*}\left(\frac{1}{n^2}\right).\end{aligned}$$

By Lemma I.12, $\frac{1}{n}\sum_i \|A_{i3}^*\|_2^2 + \frac{1}{n}\sum_i \|A_{i4}^*\|_2^2 = O_{P^*}(\Delta_F^2)$. Hence

$$\sum_{l=2}^{4}\frac{1}{n}\sum_i \|A_{il}^*\|_2^2 = O_{P^*}(\Delta_F^2). \tag{I.9}$$

Now we bound $\frac{1}{n}\sum_{i=1}^n \|A_{i1}^*\|_2^2$. A preliminary rate is provided by Lemma I.11 where we have that

$$\frac{1}{n}\sum_{i=1}^{n}\|A_{i1}^*\|_2^2 = O_{P^*}\left(\Delta_F^2 + \frac{1}{n}\right).$$

However, this rate is not sharp due to the $O_{P^*}(n^{-1})$ term and can be improved. Specifically, the proof of Lemma I.11 (iii) uses a Cauchy-Schwarz inequality and is not sharp for terms involving $E\left[U_{it}'U_{jt}\right]$. To see this intuitively, consider a simple example where we bound

$$\frac{1}{n}\sum_{i=1}^{n}\|\frac{1}{nTp}\sum_{j=1}^{n}\sum_{t=1}^{T}f_jEU_{it}'U_{jt}\|_2^2.$$

Since U_{it} and U_{jt} are independent when $i \neq j$, this term may be simplified to

$$\frac{1}{n}\sum_{i=1}^{n}\|\frac{1}{nTp}\sum_{t=1}^{T}f_iEU_{it}'U_{it}\|_2^2 = O_P\left(\frac{1}{n^2}\right).$$

In contrast, using the Cauchy-Schwarz inequality gives

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\|\frac{1}{nTp}\sum_{j=1}^{n}\sum_{t=1}^{T}f_jEU_{it}'U_{jt}\|_2^2 &\leq \frac{1}{n}\sum_{j}\|\widehat{f}_j\|_2^2\frac{1}{n^2}\sum_{j}\sum_{i=1}^{n}\left(\frac{1}{Tp}\sum_{t=1}^{T}EU_{it}'U_{jt}\right)^2\\ &= O_P\left(\frac{1}{n}\right).\end{aligned}$$

Lemma I.11 (iii) does provide a useful preliminary rate to build upon. Applying Lemma I.11 (iii) and (I.9), we obtain a preliminary rate

$$\frac{1}{n}\sum_{i=1}^{n}\|\widehat{f}_i^* - H^{*'}\widehat{f}_i\|_2^2 = O_{P^*}\left(\Delta_F^2 + \frac{1}{n}\right).$$

Our goal is to remove the term $\frac{1}{n}$ through improving the bound for $\frac{1}{n}\sum_{i=1}^{n}\|A_{i1}^*\|_2^2$. By the triangle inequality,

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\|A_{i1}^*\|_2^2 \leq \frac{2}{n}\sum_{i=1}^{n}&\left\|\frac{1}{pTn}\sum_{j=1}^{n}H^{*'}\widehat{f}_j\sum_{t=1}^{T}(\tilde{U}_{it}^{*'}\tilde{U}_{jt}^* - E^*\tilde{U}_{it}^{*'}\tilde{U}_{jt}^*)\right\|_2^2\\ &+\frac{2}{n}\sum_{i=1}^{n}\left\|\frac{1}{n}\sum_{j=1}^{n}(\widehat{f}_j^* - H^{*'}\widehat{f}_j)\frac{1}{pT}\sum_{t=1}^{T}(\tilde{U}_{it}^{*'}\tilde{U}_{jt}^* - E^*\tilde{U}_{it}^{*'}\tilde{U}_{jt}^*)\right\|_2^2\\ &\leq^{(a)} O_{P^*}(\Delta_F^2) + O_{P^*}(\Delta_F^2)\frac{1}{n^2}\sum_{ij}\left(\frac{1}{pT}\sum_{t=1}^{T}(\tilde{U}_{it}^{*'}\tilde{U}_{jt}^* - E^*\tilde{U}_{it}^{*'}\tilde{U}_{jt}^*)\right)^2\\ &= O_{P^*}(\Delta_F^2)\end{aligned}$$

where in (a) we used Lemma I.13 and the last equality follows from (I.15). Hence combining with (I.9), we have $\frac{1}{n}\sum_{i=1}^{n}\|\widehat{f}_i^* - H^{*'}\widehat{f}_i\|_2^2 = O_{P^*}(\Delta_F^2)$. Thus, we have $\Delta_F^* = \Delta_F$.

(ii) We now verify the conditions in Assumption D.5: $\sqrt{nT}|J|_0^2\Delta_F^{*2} = o(1)$, $\Delta_{eg}^* = o(\frac{1}{\sqrt{nT}})$, $\Delta_{ud}^* = o(\sqrt{\frac{\log p}{nT}})$, $|J|_0^2\sqrt{\log p}\Delta_{ud}^* = o(1)$, $\Delta_F^{*2} = o(\frac{\log p}{T\log(pT)})$, and $\Delta_{\max}^2|J|_0^2 T\Delta_F^{*2} = o(1)$.

Δ_F^*: We have previously proven that $\Delta_F^2 = \frac{1}{n^2} + \frac{1}{nT^2} + \frac{1}{pT}$. In addition, Lemma I.1 gives $\Delta_{\max} = \frac{1}{\sqrt{n}} + \sqrt{\frac{\log n}{Tp}}$. Hence it is straightforward to verify the required conditions involving Δ_F^*, given the assumption that $|J|_0^2\log n = o(p)$.

Δ_{ud}^*: Δ_{ud}^* is defined in Lemma I.5 which gives $\Delta_{ud}^* = b_n + \Delta_{ud}$ for

$$b_n = \frac{1}{T}\sqrt{\frac{\log p\log(np)}{np}} + \sqrt{\frac{\log n}{npT}} + \frac{\log p}{n\sqrt{T}} + \frac{1}{pT} + \sqrt{\frac{\log p}{npT}} + \frac{\sqrt{\log p}}{n} + \frac{\sqrt{|J|_0}}{nT} + \sqrt{\frac{|J|_0}{npT}}.$$

In the proof of Proposition D.1, we verified $\Delta_{ud} = o\left(\sqrt{\frac{\log p}{nT}}\right)$ and $|J|_0^2\sqrt{\log p}\Delta_{ud} = o(1)$. It is also straightforward to verify that $b_n = o\left(\sqrt{\frac{\log p}{nT}}\right)$ and that $|J|_0^2\sqrt{\log p}b_n = o(1)$ given that $\log n = o(p)$, $|J|_0^4 n = o(p^2T)$, $|J|_0^2\log^3 p = o(n)$, and $|J|_0^4 = o(nT^3)$. In particular, we need to verify $|J|_0^5\log p\left(\frac{1}{n^2T^2} + \frac{1}{npT}\right) = o(1)$. To verify this condition, we use $|J|_0^4 n = o(p^2T)$ and $|J|_0^4 = o(nT^3)$ to show

$$\begin{aligned}
|J|_0^5\log p\left(\frac{1}{n^2T^2} + \frac{1}{npT}\right) &= |J|_0^3\log p\left(\frac{|J|_0^2}{n^2T^2} + \frac{|J|_0^2 n^{1/2}}{n^{3/2}pT}\right) \\
&= o(1)|J|_0^3\log p\left(\frac{n^{1/2}T^{3/2}}{n^2T^2} + \frac{pT^{1/2}}{n^{3/2}pT}\right) \\
&= o(1)\frac{|J|_0^3\log p}{n^{3/2}T^{1/2}} \\
&= o(1)(\frac{|J|_0\log p}{n^{1/2}})^3 = o(1).
\end{aligned}$$

Δ_{eg}^*: Note that for $\widehat{g}_{tm} \in \{\widehat{\Lambda}_t'\widehat{\gamma}_d, \widehat{\Lambda}_t'\widehat{\gamma}_y, \widehat{\delta}_{dt}, \widehat{\delta}_{yt}, \widehat{\lambda}_{tm}\}$, we have

$$\omega_n^* = \max_{m\le p}\frac{1}{T}\sum_{t=1}^T\|\widehat{g}_{tm}\|_2^2 = O_P(|J|_0^2).$$

Hence by Lemma I.4, $\Delta_{eg}^{*2} = \left(\frac{1}{n^2} + \frac{\log n}{npT} + \frac{\log n}{n^2T^2} + \frac{\log^{1/2} n}{n^2T^{1/2}}\right)|J|_0^2$. Given $|J|_0^2\log n = o(p)$ and $|J|_0^2 T = o(n)$, it is then straightforward to verify $\Delta_{eg}^{*2} = o\left(\frac{1}{nT}\right)$ which follows by verifying

$\frac{|J|_0^2 \log n}{nT} = o(1)$. To see $\frac{|J|_0^2 \log n}{nT} = o(1)$, note that we have, by $|J|_0^{4/3} = o(n^{1/3}T)$,

$$\frac{|J|_0^2 \log n}{nT} = \frac{|J|_0^{2/3}|J|_0^{4/3} \log n}{nT} = o(1)\frac{|J|_0^{2/3} n^{1/3} T \log n}{nT} = o(1)\frac{|J|_0^{2/3} \log n}{n^{2/3}} = o(1).$$

■

Lemma I.4. *In the bootstrap sampling space, let* $\tilde{z}_{it}^* = \widehat{z}_{it} w_i^Z$ *where* $\{w_i^Z\}_{i=1}^n$ *are i.i.d. with mean zero and bounded variance and independent of* $\{w_i^U\}$ *and* $\widehat{z}_{it} = \widehat{\eta}_{it}$ *or* $\widehat{z}_{it} = \widehat{\epsilon}_{it}$*. In addition, let* $\{\widehat{g}_{tm}\}$ *be a deterministic sequence (in the bootstrap sampling space) of vectors with a fixed dimension,* $m \leq p$*. Then for* $\omega_n^* = \max_{m \leq p} \frac{1}{T} \sum_{t=1}^T \|\widehat{g}_{tm}\|_2^2$,

$$\max_{m \leq p} \left\| \frac{1}{nT} \sum_{i=1}^n \sum_{t=1}^T \widehat{g}_{tm} \tilde{z}_{it}^* (\widehat{f}_i^* - H^{*'} \widehat{f}_i)' \right\|_F^2 = O_{P^*}\left(\frac{1}{n^2} + \frac{\log n}{npT} + \frac{\log n}{n^2 T^2} + \frac{\log^{1/2} n}{n^2 T^{1/2}} \right) \omega_n^*$$

where the term $O_{P^*}\left(\frac{1}{n^2} + \frac{\log n}{npT} + \frac{\log n}{n^2 T^2} + \frac{\log^{1/2} n}{n^2 T^{1/2}} \right) \omega_n^*$ *defines* Δ_{eg}^{*2}.

Proof. It follows from (I.8) that

$$\max_{m \leq p} \| \frac{1}{nT} \sum_{i=1}^n \sum_{s=1}^T \tilde{z}_{is}^* (\widehat{f}_i^* - H^{*'} \widehat{f}_i) \widehat{g}_{sm}' \|_F \leq \sum_{l=1}^3 \bar{C}_l O_P(1),$$

where each term $\bar{C}_l$ is defined and bounded below.

First, we have

$$\begin{aligned}
\bar{C}_1 &= \max_{m \leq p} \left\| \frac{1}{nT} \sum_{i=1}^n \sum_{s=1}^T \tilde{z}_{is}^* \frac{1}{pTn} \sum_{j=1}^n \widehat{f}_j^* \sum_{t=1}^T \tilde{U}_{it}^{*'} \tilde{U}_{jt}^* \widehat{g}_{sm}' \right\|_F \\
&\leq \left(\frac{1}{pT} \sum_t \| \frac{1}{n} \sum_{j=1}^n \widehat{f}_j^* \tilde{U}_{jt}^{*'} \|_F^2 \right)^{1/2} \left(\frac{1}{T^2 p} \sum_{s,t} \| \frac{1}{n} \sum_{i=1}^n \tilde{z}_{is}^* \tilde{U}_{it}^* \|_2^2 \right)^{1/2} \omega_n^{*1/2} \\
&= O_{P^*}\left(\Delta_F + \frac{1}{\sqrt{n}} \right) \frac{1}{\sqrt{n}} \omega_n^{*1/2} \\
&= \omega_n^{1/2} O_{P^*}\left(\frac{1}{\sqrt{npT}} + \frac{1}{n} \right),
\end{aligned}$$

where we used

$$
\begin{aligned}
\frac{1}{pT}\sum_t \|\frac{1}{n}\sum_{j=1}^n \widehat{f}_j^* \tilde{U}_{jt}^{*\prime}\|_F^2 &\leq O_{P^*}(\Delta_F^2) + O_{P^*}(1)\frac{1}{pT}\sum_t \|\frac{1}{n}\sum_{j=1}^n \widehat{f}_j \tilde{U}_{jt}^{*\prime}\|_F^2 \\
&= O_{P^*}\left(\Delta_F^2 + \frac{1}{n}\right), \\
\frac{1}{T^2p}\sum_{s,t}\|\frac{1}{n}\sum_{i=1}^n \tilde{z}_{is}^* \tilde{U}_{it}^*\|_2^2 &= \frac{1}{T^2p}\sum_{s,t}\|\frac{1}{n}\sum_{i=1}^n \widehat{z}_{is} w_i^Z w_i^U \widehat{U}_{it}\|_2^2 \\
&= O_{P^*}\left(\frac{1}{n}\right),
\end{aligned}
\tag{I.10}
$$

and $\Delta_F = \Delta_F^*$.

For the second term, we have by Lemma I.13 that

$$
\begin{aligned}
\bar{C}_2 &= \max_{m\leq p}\left\|\frac{1}{nT}\sum_{i=1}^n\sum_{s=1}^T \tilde{z}_{is}^* \frac{1}{pTn}\sum_{j=1}^n \widehat{f}_j^* \sum_{t=1}^T \widehat{f}_j' \widehat{\Lambda}_t' \tilde{U}_{it}^* \widehat{g}_{sm}'\right\|_F \\
&\leq O_{P^*}(\omega_n^{*1/2})\left(\frac{1}{T}\sum_{s=1}^T \|\frac{1}{pTn}\sum_{i=1}^n\sum_{t=1}^T \widehat{\Lambda}_t' \tilde{U}_{it}^* \tilde{z}_{is}^*\|_2^2\right)^{1/2} \\
&= O_{P^*}\left(\sqrt{\frac{\log n}{npT}} + \frac{\sqrt{\log n}}{nT} + \frac{1}{n} + \frac{1}{n}(\frac{\log n}{T})^{1/4}\right)\omega_n^{*1/2}.
\end{aligned}
$$

Finally,

$$
\begin{aligned}
\bar{C}_3 &= \max_{m\leq p}\left\|\frac{1}{nT}\sum_{i=1}^n\sum_{s=1}^T \tilde{z}_{is}^* \frac{1}{pTn}\sum_{j=1}^n \widehat{f}_j^* \sum_{t=1}^T \widehat{f}_i' \widehat{\Lambda}_t' \tilde{U}_{jt}^* \widehat{g}_{sm}'\right\|_F \\
&\leq \left(\frac{1}{T}\sum_{s=1}^T (\frac{1}{n}\sum_{i=1}^n \tilde{z}_{is}^* \widehat{f}_i')^2\right)^{1/2} \omega_n^{*1/2}\left\|\frac{1}{nTp}\sum_{j=1}^n\sum_{t=1}^T \widehat{f}_j^* \widehat{\Lambda}_t' \tilde{U}_{jt}^*\right\|_2 \\
&= O_{P^*}\left(\omega_n^{*1/2}\frac{1}{\sqrt{n}}\right)\left(\Delta_F^2 + \sqrt{\frac{\Delta_F^2 \log n}{n}}\right)
\end{aligned}
$$

where we used Lemma H.14 to obtain $\frac{1}{T}\sum_t \|\frac{1}{n}\sum_{i=1}^n \tilde{z}_{it}^* \widehat{f}_i\|_2^2 = O_{P^*}\left(\frac{1}{n}\right)$.

Combining the above, we have

$$
\max_{m\leq p}\|\frac{1}{nT}\sum_{i=1}^n\sum_{t=1}^T \widehat{g}_{tm}\tilde{z}_{it}^*(\widehat{f}_i^* - H^{*\prime}\widehat{f}_i)'\|_F^2 = O_{P^*}\left(\frac{1}{n^2} + \frac{\log n}{npT} + \frac{\log n}{n^2T^2} + \frac{\log^{1/2} n}{n^2T^{1/2}}\right)\omega_n^*.
$$

■

Lemma I.5. *For* $\widehat{h}_{tk} \in \{\widehat{\delta}_{yt}, \widehat{\delta}_{dt}, \widehat{\lambda}_{tk}\}$, *we have*

$$\max_{m,k\le p}\left\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(\widehat{f}_i^* - H^{*'}\widehat{f}_i)\tilde{U}^*_{it,m}\widehat{h}'_{tk}\right\|_F \le O_{P^*}\left(\frac{1}{T}\sqrt{\frac{\log p\log(np)}{np}} + \sqrt{\frac{\log n}{npT}} + \frac{\log p}{n\sqrt{T}} + \frac{1}{pT} + \sqrt{\frac{\log p}{npT}} + \frac{\sqrt{\log p}}{n} + \frac{\sqrt{|J|_0}}{nT} + \sqrt{\frac{|J|_0}{npT}}\right) + O_{P^*}(\Delta_{ud})$$

for Δ_{ud} *defined as in Lemma I.3. The term on the right-hand-side of the inequality defines* Δ^*_{ud}.

Proof. We have $\max_{m,k\le p}\|\frac{1}{nT}\sum_{i=1}^n\sum_{s=1}^T(\widehat{f}_i^* - H^{*'}\widehat{f}_i)\tilde{U}^*_{is,m}\widehat{h}'_{sk}\|_F \le \sum_{l=1}^5 D_l$, where each D_l is defined and bounded below.

$$\begin{aligned} D_1 &= \max_{m,k\le p}\left\|\frac{1}{T}\sum_{t=1}^{T}\frac{1}{n}\sum_{j=1}^{n}\widehat{f}_j^*\tilde{U}^{*'}_{jt}\frac{1}{nTp}\sum_{i=1}^{n}\sum_{s=1}^{T}(\tilde{U}^*_{it}\tilde{U}^*_{is,m} - E^*\tilde{U}^*_{it}\tilde{U}^*_{is,m})\widehat{h}'_{sk}\right\|_F \\ &\le \left(\frac{1}{T}\sum_{t=1}^{T}\|\frac{1}{n}\sum_{j=1}^{n}\widehat{f}_j^*\tilde{U}^{*'}_{jt}\|_2^2\right)^{1/2} \\ &\quad\times \max_{m,k\le p}\left(\frac{1}{T}\sum_{t=1}^{T}\|\frac{1}{nTp}\sum_{i=1}^{n}\sum_{s=1}^{T}(\tilde{U}^*_{it}\tilde{U}^*_{is,m} - E^*\tilde{U}^*_{it}\tilde{U}^*_{is,m})\widehat{h}'_{sk}\|_F^2\right)^{1/2} \\ &= O_{P^*}\left(\sqrt{\frac{\log(pT)}{n}(\frac{\log n\log(pT)}{n} + \frac{\log(np)}{T})}\right)\left(\Delta_F + \frac{1}{\sqrt{n}}\right) \end{aligned}$$

where we used Lemma I.13 and equation (I.10) from the proof of Lemma I.4 which gives

$$\frac{1}{pT}\sum_t\|\frac{1}{n}\sum_{j=1}^n\widehat{f}_j^*\tilde{U}^{*'}_{jt}\|_F^2 \le O_{P^*}\left(\Delta_F^2 + \frac{1}{n}\right).$$

Next,

$$\begin{aligned} D_2 &= \max_{m,k\le p}\left\|\sum_{l=1}^{p}\frac{1}{T}\sum_{t=1}^{T}\frac{1}{pn}\sum_{j=1}^{n}\widehat{f}_j^*\tilde{U}^{*'}_{jt,l}\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}E^*(\tilde{U}^*_{it,l}\tilde{U}^*_{is,m})\widehat{h}'_{sk}\right\|_F \\ &\le \max_{tl}\left\|\frac{1}{n}\sum_{j=1}^{n}\widehat{f}_j^*\tilde{U}^{*'}_{jt,l}\right\|_2 \max_{m,k\le p}\frac{1}{pT}\sum_{l=1}^{p}\sum_{t=1}^{T}\left\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}E^*(\tilde{U}^*_{it,l}\tilde{U}^*_{is,m})\widehat{h}'_{sk}\right\|_F \end{aligned}$$

$$= O_{P^*}\left(\sqrt{\frac{\log(pT)}{n}}+\Delta_F\right)\left(\Delta_F+\frac{\sqrt{\log(pT)\log p}}{n}+\sqrt{\frac{\log(pT)}{nT}}\right)$$

where we used Lemma I.13 and $\max_{mt}\|\frac{1}{n}\widehat{F}^{*\prime}\tilde{U}^*_{t,m}\|_2 = O_{P^*}\left(\sqrt{\frac{\log(pT)}{n}}+\Delta_F\right)$ due to Lemma H.14. We then have

$$\begin{aligned} D_3 &= \max_{m,k\leq p}\|\frac{1}{nT}\sum_{i=1}^n\sum_{s=1}^T\frac{1}{pTn}\sum_{j=1}^n\widehat{f}^*_j\sum_{t=1}^T\widehat{f}_j'\widehat{\Lambda}_t'(\tilde{U}^*_{it}\tilde{U}^*_{is,m}-E^*\tilde{U}^*_{it}\tilde{U}^*_{is,m})\widehat{h}'_{sk}\|_F \\ &\leq O_{P^*}(1)\max_{m,k\leq p}\|\frac{1}{nT}\sum_{i=1}^n\sum_{s=1}^T\frac{1}{pT}\sum_{t=1}^T\widehat{\Lambda}_t'(\tilde{U}^*_{it}\tilde{U}^*_{is,m}-E^*\tilde{U}^*_{it}\tilde{U}^*_{is,m})\widehat{h}'_{sk}\|_F \\ &= O_{P^*}\left(\sqrt{\frac{\log p}{n}(\frac{\log n\log(pT)}{n}+\frac{\log(np)}{T})\Delta_F^2}\right) \end{aligned}$$

by Lemma I.13. We also have

$$\begin{aligned} D_4 &= \max_{m,k\leq p}\|\frac{1}{nT}\sum_{i=1}^n\sum_{s=1}^T\frac{1}{pTn}\sum_{j=1}^n\widehat{f}^*_j\sum_{t=1}^T\widehat{f}_j'\widehat{\Lambda}_t'E^*\tilde{U}^*_{it}\tilde{U}^*_{is,m}\widehat{h}'_{sk}\|_F \\ &\leq O_{P^*}(1)\max_{m,k\leq p}\|\frac{1}{nT}\sum_{i=1}^n\sum_{s=1}^T\frac{1}{pT}\sum_{t=1}^T\widehat{\Lambda}_t'E^*\tilde{U}^*_{it}\tilde{U}^*_{is,m}\widehat{h}'_{sk}\|_F \\ &= O_{P^*}\left(\frac{1}{pT}+\sqrt{\frac{\log p}{npT}}+\frac{\sqrt{\log p}}{n}+\frac{\sqrt{|J|_0}}{nT}+\sqrt{\frac{|J|_0}{npT}}\right)+O_{P^*}(\Delta_{ud}) \end{aligned}$$

where the inequality follows from Lemma I.14 (iii). Finally,

$$\begin{aligned} D_5 &= \max_{m,k\leq p}\left\|\frac{1}{pTn}\sum_{j=1}^n\widehat{f}^*_j\sum_{t=1}^T\widehat{\Lambda}_t'\tilde{U}^*_{jt}\frac{1}{nT}\sum_{s=1}^T\sum_{i=1}^n\widehat{f}_i\tilde{U}^*_{is,m}\widehat{h}'_{sk}\right\|_F \\ &\leq \left\|\frac{1}{pTn}\sum_{j=1}^n\widehat{f}^*_j\sum_{t=1}^T\widehat{\Lambda}_t'\tilde{U}^*_{jt}\right\|_2\max_{m,k\leq p}\left\|\frac{1}{nT}\sum_{s=1}^T\sum_{i=1}^n\widehat{f}_i\tilde{U}^*_{is,m}\widehat{h}'_{sk}\right\|_F \\ &=^{(a)} O_{P^*}\left(\Delta_F^2+\sqrt{\frac{\Delta_F^2\log n}{n}}\right)\max_{m,k\leq p}\left\|\frac{1}{nT}\sum_{s=1}^T\sum_{i=1}^n\widehat{f}_i\tilde{U}^*_{is,m}\widehat{h}'_{sk}\right\|_F \\ &= O_{P^*}\left(\Delta_F^2+\sqrt{\frac{\Delta_F^2\log n}{n}}\right) \end{aligned}$$

where equality (a) results by applying Lemma I.13 (iii). Note that the upper bound achieved in the last equality is not sharp but is sufficient to verify Assumptions about Δ^*_{ud}.

Combining the above terms, we reach

$$\max_{m,k\le p}\left\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}(\widehat{f}_i^{*}-H^{*'}\widehat{f}_i)\tilde{U}_{is,m}^{*}\widehat{h}_{sk}'\right\|_F$$

$$= O_{P^*}\left(\sqrt{\frac{\log(pT)}{n}(\frac{\log n\log(pT)}{n}+\frac{\log(np)}{T})}\right)\left(\Delta_F+\frac{1}{\sqrt{n}}\right)$$

$$+ O_{P^*}\left(\sqrt{\frac{\log(pT)}{n}}+\Delta_F\right)\left(\Delta_F+\frac{\sqrt{\log(pT)\log p}}{n}+\sqrt{\frac{\log(pT)}{nT}}\right)$$

$$+ O_{P^*}(\Delta_{ud}+\Delta_F\sqrt{\frac{\log n}{n}})+O_{P^*}\left(\Delta_F\sqrt{\frac{\log p}{n}(\frac{\log n\log(pT)}{n}+\frac{\log(np)}{T})}\right)$$

$$+ O_{P^*}\left(\frac{1}{pT}+\sqrt{\frac{\log p}{npT}}+\frac{\sqrt{\log p}}{n}+\frac{\sqrt{|J|_0}}{nT}+\sqrt{\frac{|J|_0}{npT}}\right)$$

$$= O_{P^*}\left(\frac{1}{T}\sqrt{\frac{\log p\log(np)}{np}}+\sqrt{\frac{\log n}{npT}}+\frac{\log p}{n\sqrt{T}}+\frac{1}{pT}+\sqrt{\frac{\log p}{npT}}+\frac{\sqrt{\log p}}{n}+\frac{\sqrt{|J|_0}}{nT}+\sqrt{\frac{|J|_0}{npT}}\right)$$

$$+ O_{P^*}(\Delta_{ud})$$

where Δ_{ud} is defined as in Lemma I.3. ■

I.3. **Technical lemmas for Proposition D.1: the original data.**

Lemma I.6. *Let A_{i1} be as defined in (I.1). Then*

$$\frac{1}{n}\sum_{i=1}^{n}\|A_{i1}\|_2^2 = O_P\left(\frac{1}{n^2}+\frac{1}{npT}+\frac{1}{nT^2}+\frac{1}{pT^2}\right).$$

Proof. First, we have the following equality:

$$\frac{1}{T}\sum_{t=1}^{T}\tilde{U}_{it}'\tilde{U}_{jt} = \frac{1}{T}\sum_{t=1}^{T}U_{it}'U_{jt}-\bar{U}_{i\cdot}'\bar{U}_{j\cdot}-\frac{1}{T}\sum_{t=1}^{T}U_{it}'\bar{U}_{\cdot t}+\bar{U}_{i\cdot}'\bar{\bar{U}}+\bar{\bar{U}}'\bar{U}_{j\cdot}-\bar{\bar{U}}'\bar{\bar{U}}-\frac{1}{T}\sum_{t=1}^{T}\bar{U}_{\cdot t}'U_{jt}+\frac{1}{T}\sum_{t=1}^{T}\bar{U}_{\cdot t}'\bar{U}_{\cdot t}.$$

Therefore, $\frac{1}{n}\sum_{i=1}^{n}\|A_{i1}\|_2^2 \le \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(\frac{1}{pT}\sum_{t=1}^{T}\tilde{U}_{it}'\tilde{U}_{jt}-U_{it}'U_{jt})^2O_P(1)$, and

$$E\left[\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(\frac{1}{pT}\sum_{t=1}^{T}(\tilde{U}_{it}'\tilde{U}_{jt}-U_{it}'U_{jt}))^2\right]$$

$$\le CE\left[\frac{1}{n}\sum_{i=1}^{n}(\frac{1}{pT}\sum_{t=1}^{T}\bar{U}_{\cdot t}'U_{it})^2\right]+CE\left[(\frac{1}{pT}\sum_{t=1}^{T}\bar{U}_{\cdot t}'\bar{U}_{\cdot t})^2\right]$$

$$+ CE\left[\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n(\frac{1}{p}\bar{U}_{i\cdot}'\bar{U}_{j\cdot})^2\right] + CE\left[\frac{1}{n}\sum_{i=1}^n(\frac{1}{p}\bar{U}_{i\cdot}'\bar{\bar{U}})^2\right] + CE\left[(\frac{1}{p}\bar{\bar{U}}'\bar{\bar{U}})^2\right]$$

$$:= \sum_{l=1}^5 \Delta_l.$$

We now separately bound $\Delta_1 - \Delta_5$.

$$\begin{aligned}
\Delta_1 =& E\left[\frac{1}{n}\sum_{i=1}^n(\frac{1}{pT}\sum_{t=1}^T \bar{U}_{\cdot t}'U_{it})^2\right] \\
&= \frac{1}{n^3}\sum_{i=1}^n\sum_{j=1}^n\sum_{l=1}^n\frac{1}{p^2T^2}\sum_{t=1}^T\sum_{s=1}^T E\left[U_{is}'U_{js}U_{it}'U_{lt}\right] \\
&= \frac{1}{n^3}\sum_{i=1}^n\frac{1}{p^2T^2}\sum_{t=1}^T\sum_{s=1}^T E\left[U_{is}'U_{is}U_{it}'U_{it}\right] + \frac{1}{n^3}\sum_{i=1}^n\sum_{j\neq i}\frac{1}{p^2T^2}\sum_{t=1}^T\sum_{s=1}^T E\left[U_{is}'U_{js}U_{it}'U_{jt}\right] \\
&= \frac{1}{n^3}\sum_{i=1}^n\sum_{j=1}^n\frac{1}{p^2T^2}E\left[\left(\sum_{s=1}^T U_{is}'U_{js}\right)^2\right] \\
&= \frac{1}{n^3}\sum_{i=1}^n\sum_{j=1}^n\frac{1}{p^2T^2}\operatorname{Var}\left(\sum_{s=1}^T U_{is}'U_{js}\right) + \frac{1}{n^3}\sum_{i=1}^n\sum_{j=1}^n\frac{1}{p^2T^2}\left(E\left[\sum_{s=1}^T U_{is}'U_{js}\right]\right)^2 \\
&= \frac{1}{n^3}\sum_{i=1}^n\sum_{j=1}^n\frac{1}{p^2T^2}\sum_{t=1}^T\sum_{s=1}^T \operatorname{Cov}(U_{is}'U_{js}, U_{it}'U_{jt}) + \frac{1}{n^3}\sum_{i=1}^n\frac{1}{p^2T^2}\left(\sum_{s=1}^T E\left[U_{is}'U_{is}\right]\right)^2 \\
&= \frac{1}{n^3}\sum_{i=1}^n\frac{1}{p^2T^2}\sum_{t=1}^T\sum_{s=1}^T \operatorname{Cov}(U_{is}'U_{is}, U_{it}'U_{it}) \\
&\quad + \frac{1}{n^3}\sum_{i=1}^n\sum_{j\neq i}\frac{1}{p^2T^2}\sum_{t=1}^T\sum_{s=1}^T\sum_{m=1}^p\sum_{v=1}^p E\left[U_{is,m}U_{it,v}EU_{js,m}U_{jt,v}\right] \\
&\quad + \frac{1}{n^3}\sum_{i=1}^n\frac{1}{p^2T^2}\left(\sum_{s=1}^T\sum_{m=1}^p E\left[U_{is,m}^2\right]\right)^2 \\
&\leq \frac{1}{n^2T}\max_{i\leq n, t\leq T}\sum_{s=1}^T\left|\operatorname{Cov}\left(\frac{1}{p}U_{is}'U_{is}, \frac{1}{p}U_{it}'U_{it}\right)\right| \\
&\quad + \frac{1}{n^3}\sum_{i=1}^n\sum_{j\neq i}\frac{1}{p^2T^2}\sum_{t=1}^T\sum_{m=1}^p\max_{istmv}\left|E\left[U_{is,m}U_{it,v}\right]\right|\sum_{v=1}^p\sum_{s=1}^T\left|E\left[U_{js,m}U_{jt,v}\right]\right|
\end{aligned}$$

$$+\frac{1}{n^2}\max_{i,s,m}\mathrm{Var}^2(U_{is,m})$$
$$=O\left(\frac{1}{n^2}+\frac{1}{npT}\right),$$

where we used the condition that $\max_{i\le n}\max_{m\le p,t\le T}\sum_{s=1}^{T}\sum_{v=1}^{p}|\mathrm{Cov}(U_{it,m},U_{is,v})| = O(1)$.

$$\Delta_2 = E\left[\left(\frac{1}{pT}\sum_{t=1}^{T}\bar{U}_{\cdot t}'\bar{U}_{\cdot t}\right)^2\right]$$
$$= E\left[\frac{1}{p^2T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\bar{U}_{\cdot t}'\bar{U}_{\cdot t}\bar{U}_{\cdot s}'\bar{U}_{\cdot s}\right]$$
$$=\frac{1}{p^2T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{1}{n^4}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{h=1}^{n}E\left[U_{it}'U_{jt}U_{ks}'U_{hs}\right]$$
$$=\frac{1}{p^2T^2}\frac{1}{n^4}\sum_{i=1}^{n}E\left[\left(\sum_{t=1}^{T}U_{it}'U_{it}\right)^2\right]+\frac{1}{p^2T^2}\frac{1}{n^4}\sum_{i=1}^{n}\sum_{j\neq i}E\left[\left(\sum_{t=1}^{T}U_{it}'U_{jt}\right)^2\right]$$
$$+\frac{1}{p^2T^2}\frac{1}{n^4}\sum_{i=1}^{n}\sum_{k\neq i}\sum_{t=1}^{T}E\left[U_{it}'U_{it}\right]\sum_{s=1}^{T}E\left[U_{ks}'U_{ks}\right]+\frac{1}{p^2T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{1}{n^4}\sum_{i=1}^{n}\sum_{j\neq i}E\left[U_{it}'U_{jt}U_{js}'U_{is}\right]$$
$$=O\left(\frac{1}{n^2}\right).$$

$$\Delta_3=\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}E\left[\left(\frac{1}{p}\bar{U}_{i\cdot}'\bar{U}_{j\cdot}\right)^2\right]$$
$$=\frac{1}{n^2}\sum_{i=1}^{n}E\left[\left(\frac{1}{p}\sum_{m=1}^{p}\bar{U}_{i\cdot,m}^2\right)^2\right]$$
$$+\frac{1}{n^2p^2T^4}\sum_{i=1}^{n}\sum_{j\neq i}\sum_{k=1}^{T}\sum_{v=1}^{T}\sum_{s=1}^{T}\sum_{t=1}^{T}\sum_{m=1}^{p}\sum_{l=1}^{p}E\left[U_{it,m}U_{iv,l}EU_{jk,l}U_{js,m}\right]$$
$$\le\frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}\left(\frac{1}{p}\sum_{m=1}^{p}\bar{U}_{i\cdot,m}^2\right)+\frac{1}{n^2}\sum_{i=1}^{n}\left(\frac{1}{p}\sum_{m=1}^{p}E\left[\bar{U}_{i\cdot,m}^2\right]\right)^2$$

$$+\frac{1}{n^2p^2T^4}\sum_{i=1}^{n}\sum_{j\neq i}\sum_{s=1}^{T}\sum_{t=1}^{T}\sum_{m=1}^{p}\max_{l\leq p}\sum_{k=1}^{T}|E[U_{jk,l}U_{js,m}]|\max_{i\leq n,m\leq p,t\leq T}\sum_{l=1}^{p}\sum_{v=1}^{T}|E[U_{it,m}U_{iv,l}]|$$

$$\leq\frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}\left(\frac{1}{p}\sum_{m=1}^{p}\bar{U}_{i\cdot,m}^2\right)+\frac{1}{n}\left[\max_{i\leq n,m\leq p}\mathrm{Var}\left(\bar{U}_{i\cdot,m}\right)\right]^2+\frac{C}{pT^2}$$

$$=O\left(\frac{1}{npT^2}\right)+O\left(\frac{1}{nT^2}\right)+O\left(\frac{1}{pT^2}\right)$$

where we assumed $\frac{1}{n}\sum_i\mathrm{Var}\left(\frac{1}{p}\sum_{m=1}^{p}\bar{U}_{i\cdot,m}^2\right)=O\left(\frac{1}{pT^2}\right)$.

$$\Delta_4=E\left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{p}\bar{U}_{i\cdot}'\bar{\bar{U}}\right)^2\right]$$

$$=\frac{1}{np^2}\frac{1}{n^2T^4}\sum_{i=1}^{n}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{k=1}^{T}\sum_{v=1}^{T}\sum_{m=1}^{p}\sum_{l=1}^{p}E[U_{it,m}U_{is,m}U_{ik,l}U_{iv,l}]$$

$$+\frac{1}{np^2}\frac{1}{n^2T^4}\sum_{i=1}^{n}\sum_{j\neq i}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{k=1}^{T}\sum_{m=1}^{p}\sum_{l=1}^{p}\sum_{v=1}^{T}E[U_{it,m}U_{ik,l}EU_{jv,l}U_{js,m}]$$

$$\leq\frac{1}{n^3}\sum_{i=1}^{n}E\left[\left(\frac{1}{p}\sum_{m=1}^{p}\bar{U}_{i\cdot,m}^2\right)^2\right]$$

$$+\frac{1}{n^3T^4p^2}\sum_{i=1}^{n}\sum_{j\neq i}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{m=1}^{p}\max_{itml}\sum_{k=1}^{T}|E[U_{it,m}U_{ik,l}]|\max_{jsm}\sum_{l=1}^{p}\sum_{v=1}^{T}|E[U_{jv,l}U_{js,m}]|$$

$$=O\left(\frac{1}{n^2p}+\frac{1}{n^2T^2}\right)+O\left(\frac{1}{npT^2}\right).$$

$$\Delta_5=E\left[\left(\frac{1}{p}\bar{\bar{U}}'\bar{\bar{U}}\right)^2\right]$$

$$=\frac{1}{p^2}E\left[\left(\sum_{m=1}^{p}\left(\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}U_{it,m}\right)^2\right)^2\right]$$

$$=\frac{1}{n^2}\frac{1}{n^2}\sum_{i=1}^{n}E\left[\left(\frac{1}{p}\sum_{m=1}^{p}\bar{U}_{i\cdot,m}^2\right)^2\right]$$

$$+\frac{1}{p^2}\frac{1}{n^2T^2}\frac{1}{n^2T^2}\sum_{i=1}^{n}\sum_{t=1}^{T}\sum_{j\neq i}\sum_{s=1}^{T}\sum_{l=1}^{p}\sum_{m=1}^{p}\sum_{v=1}^{T}E[U_{iv,l}U_{it,m}]\sum_{b=1}^{T}E[U_{jb,l}U_{js,m}]$$

$$+\frac{1}{p^2}\frac{1}{n^2T^2}\frac{1}{n^2T^2}\sum_{i=1}^{n}\sum_{j\neq i}\sum_{v=1}^{T}\sum_{b=1}^{T}\sum_{l=1}^{p}E\left[U_{jv,l}U_{jb,l}\right]\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{m=1}^{p}E\left[U_{it,m}U_{is,m}\right]$$

$$+\frac{1}{p^2}\frac{1}{n^2T^2}\frac{1}{n^2T^2}\sum_{i=1}^{n}\sum_{m=1}^{p}\sum_{t=1}^{T}\sum_{j\neq i}\sum_{s=1}^{T}\sum_{l=1}^{p}\sum_{v=1}^{T}E\left[U_{jv,l}U_{js,m}\right]\sum_{b=1}^{T}E\left[U_{ib,l}U_{it,m}\right]$$

$$\leq O\left(\frac{1}{n^3p}+\frac{1}{n^3T^2}\right)+O\left(\frac{1}{n^2pT^2}\right)+O\left(\frac{1}{n^2T^2}\right)+O\left(\frac{1}{n^2pT^2}\right).$$

Hence, $\frac{1}{n}\sum_{i=1}^{n}\|A_{i1}\|_2^2=O_P(1)\sum_{l=1}^{5}\Delta_l=O_P\left(\frac{1}{n^2}+\frac{1}{npT}+\frac{1}{nT^2}+\frac{1}{pT^2}\right)$. ■

Lemma I.7. *Let $A_{i2},...,A_{i5}$ be as defined in (I.1). Then*

$$\frac{1}{n}\sum_{i=1}^{n}\|A_{i2}+...+A_{i5}\|_2^2=O_P\left(\frac{1}{pT}+\frac{1}{n^2}\right).$$

Proof. A_{i2}: We prove $\frac{1}{n}\sum_{i=1}^{n}\|A_{i2}\|_2^2=O_P\left(\frac{1}{Tp}\right)$. By (I.2),

$$\begin{aligned}
&\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}E\left[\left(\frac{1}{pT}\sum_{t=1}^{T}U_{it}'U_{jt}-EU_{it}'U_{jt}\right)^2\right]\\
&\quad=\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{Var}\left(\frac{1}{pT}\sum_{t=1}^{T}U_{it}'U_{jt}\right)\\
&\quad=\frac{1}{n^2}\sum_{i=1}^{n}\frac{1}{p^2T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\sum_{m=1}^{p}\sum_{v=1}^{p}\mathrm{Cov}(U_{it,m}^2,U_{is,v}^2)\\
&\qquad+\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j\neq i}\frac{1}{p^2T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\sum_{m=1}^{p}\sum_{v=1}^{p}E\left[U_{it,m}U_{is,v}EU_{js,v}U_{jt,m}\right]\\
&\quad\leq\frac{1}{np}\max_{i\leq n,m\leq p}\max_{t,s\leq T}\sum_{v=1}^{p}\left|\mathrm{Cov}(U_{it,m}^2,U_{is,v}^2)\right|\\
&\qquad+\max_{j,s,v,t,m}|EU_{js,v}U_{jt,m}|\frac{1}{pT}\max_{i\leq n}\max_{m\leq p,t\leq T}\sum_{s=1}^{T}\sum_{v=1}^{p}|\mathrm{Cov}(U_{it,m},U_{is,v})| \qquad (I.11)\\
&\quad=O\left(\frac{1}{pT}\right),
\end{aligned}$$

where we used the conditions

$$\max_{i\leq n}\max_{m\leq p,t\leq T}\sum_{s=1}^{T}\sum_{v=1}^{p}|\,\mathrm{Cov}(U_{it,m},U_{is,v})|=O(1),$$

$$\max_{i\leq n, m\leq p}\max_{t,s\leq T}\sum_{v=1}^{p}|\operatorname{Cov}(U_{it,m}^2, U_{is,v}^2)| = O(1),$$

and $T = o(n)$.

A_{i3} : We prove $\frac{1}{n}\sum_{i=1}^n \|A_{i3}\|_2^2 = O_P\left(\frac{1}{n^2}\right)$.

$$\frac{1}{n}\sum_{i=1}^n \|A_{i3}\|_2^2 = \frac{1}{n}\sum_{i=1}^n \left\|\frac{1}{pTn}\widehat{f}_i\sum_{t=1}^T EU_{it}'U_{it}\right\|_2^2 = O\left(\frac{1}{n^2}\right).$$

A_{i4} : We prove $\frac{1}{n}\sum_{i=1}^n \|A_{i4}\|_2^2 = O_P\left(\frac{1}{pT}\right)$. Note that

$$\frac{1}{T}\sum_{t=1}^T \tilde{\Lambda}_t'\bar{U}_{i\cdot} = \frac{1}{pT}\sum_{t=1}^T \tilde{\Lambda}_t'\bar{\bar{U}} = 0.$$

Then

$$\begin{aligned}
&\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n\left(\frac{1}{pT}\sum_{t=1}^T \tilde{f}_j'\tilde{\Lambda}_t'\tilde{U}_{it}\right)^2 \\
&\qquad\leq \frac{2}{n^2}\sum_{i=1}^n\sum_{j=1}^n(\tilde{f}_j'\frac{1}{pT}\sum_{t=1}^T\tilde{\Lambda}_t'U_{it})^2 + \frac{2}{n^2}\sum_{i=1}^n\sum_{j=1}^n\left(\frac{1}{pT}\sum_{t=1}^T\tilde{f}_j'\tilde{\Lambda}_t'\bar{U}_{\cdot t}\right)^2 \\
&\qquad\leq \frac{2}{n}\sum_{j=1}^n\left\|\tilde{f}_j\right\|_2^2\frac{1}{n}\sum_{i=1}^n\left\|\frac{1}{pT}\sum_{t=1}^T\tilde{\Lambda}_t'U_{it}\right\|_2^2 + \frac{2}{n}\sum_{j=1}^n\left\|\tilde{f}_j\right\|_2^2\left\|\frac{1}{pT}\sum_{t=1}^T\tilde{\Lambda}_t'\bar{U}_{\cdot t}\right\|_2^2.
\end{aligned}$$

The first term following the inequality is $O_P\left(\frac{1}{pT}\right)$ given the condition

$$\max_{i\leq n}\max_{m\leq p, t\leq T}\sum_{s=1}^T\sum_{v=1}^p|\operatorname{Cov}(U_{it,m}, U_{is,v})| = O(1).$$

The second term is $O_P\left(\frac{1}{np}\right)$ and is negligible since $T = o(n)$.

Finally, $\frac{1}{n}\sum_{i=1}^n \|A_{i5}\|_2^2 = O_P\left(\frac{1}{pT}\right)$ is bounded similarly to $\frac{1}{n}\sum_{i=1}^n \|A_{i4}\|_2^2$. ■

Lemma I.8. *Let $\{z_{it}\}$ be a random sequence with $E(z_{it}|f_t, U_t) = 0$, $E(z_{it}^2) = O(1)$, and*

$$\max_{ism}\sum_{h=1}^T\sum_{k=1}^p|E[U_{is,m}U_{ih,k}|z_{i1}, ..., z_{iT}]| < C$$

almost surely in $\{z_{it}\}$*. Then for* $\tilde{z}_{it} = z_{it} - \bar{z}_{i\cdot} - \bar{z}_{\cdot t} + \bar{\bar{z}}$,

$$\frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{pT}\frac{1}{n}\sum_{i=1}^{n}\sum_{s=1}^{T}\tilde{\Lambda}_s'\tilde{U}_{is}\tilde{z}_{it}\right\|_2^2 = O\left(\frac{1}{pTn}\right).$$

Proof. Since $\tilde{\Lambda}_s'$ has K number of rows with fixed K, for notational simplicity, we assume that $\tilde{\Lambda}_s'$ is a row vector without loss of generality as we can always look at each row of $\tilde{\Lambda}_s'$ separately. Then

$$E\left[\frac{1}{T}\sum_{t=1}^{T}\left|\frac{1}{pT}\frac{1}{n}\sum_{i=1}^{n}\sum_{s=1}^{T}\tilde{\Lambda}_s'\tilde{U}_{is}\tilde{z}_{it}\right|^2\right] \leq C\sum_{l=1}^{7}\bar{A}_l$$

where each $\bar{A}_l$ is defined and bounded below.

$$\begin{aligned}
\bar{A}_1 &= E\left[\frac{1}{T}\sum_{t=1}^{T}\left|\frac{1}{pT}\frac{1}{n}\sum_{i=1}^{n}\sum_{s=1}^{T}\tilde{\Lambda}_s'U_{is}z_{it}\right|^2\right]\\
&= \frac{1}{T}\sum_{t=1}^{T}\frac{1}{p^2T^2n^2}\sum_{m=1}^{p}\sum_{k=1}^{p}\sum_{h=1}^{T}\sum_{i=1}^{n}\sum_{s=1}^{T}\tilde{\lambda}_{sm}\tilde{\lambda}_{hk}E\left[z_{it}^2U_{is,m}U_{ih,k}\right]\\
&\leq O(1)\frac{1}{T}\sum_{t=1}^{T}\frac{1}{p^2T^2n^2}\sum_{m=1}^{p}\sum_{i=1}^{n}\sum_{s=1}^{T}E\left[z_{it}^2\sum_{h=1}^{T}\sum_{k=1}^{p}\left|E\left[U_{is,m}U_{ih,k}|z_{it}\right]\right|\right] = O\left(\frac{1}{pTn}\right).\\
\bar{A}_2 &= E\left[\frac{1}{T}\sum_{t=1}^{T}\left|\frac{1}{pT}\frac{1}{n}\sum_{i=1}^{n}\sum_{s=1}^{T}\tilde{\Lambda}_s'U_{is}\bar{z}_{i\cdot}\right|^2\right]\\
&\leq O(1)\frac{1}{T}\sum_{t=1}^{T}\frac{1}{pT}\frac{1}{n^2}\frac{1}{T^2}\sum_{h=1}^{T}\frac{1}{pT}\sum_{i=1}^{n}\sum_{v=1}^{T}\sum_{s=1}^{T}\sum_{m=1}^{p}E\left[|z_{ih}z_{iv}|\sum_{w=1}^{T}\sum_{k=1}^{p}\left|E\left[U_{is,m}U_{iw,k}|z_i\right]\right|\right]\\
&= O\left(\frac{1}{pTn}\right).\\
\bar{A}_3 &= E\left[\frac{1}{T}\sum_{t=1}^{T}\left|\frac{1}{pT}\frac{1}{n}\sum_{i=1}^{n}\sum_{s=1}^{T}\tilde{\Lambda}_s'U_{is}\bar{z}_{\cdot t}\right|^2\right]\\
&\leq O(1)\frac{1}{T}\sum_{t=1}^{T}\frac{1}{pT}\frac{1}{n^4}\sum_{i=1}^{n}\sum_{s=1}^{T}\frac{1}{pT}\sum_{m=1}^{p}E\left[z_{it}^2\sum_{k=1}^{p}\sum_{v=1}^{T}\left|E\left[U_{is,m}U_{iv,k}|z_i\right]\right|\right]\\
&\quad + O(1)\frac{1}{T}\sum_{t=1}^{T}\frac{1}{pT}\frac{1}{n^4}\sum_{i=1}^{n}\sum_{s=1}^{T}\sum_{j\neq i}\frac{1}{pT}\sum_{m=1}^{p}E\left[z_{jt}^2\right]\sum_{v=1}^{T}\sum_{k=1}^{p}\left|E\left[U_{is,m}U_{iv,k}\right]\right|
\end{aligned}$$

$$= O\left(\frac{1}{Tpn^2}\right).$$

$$\bar{A}_4 = E\left[\frac{1}{T}\sum_{t=1}^{T}|\frac{1}{pT}\frac{1}{n}\sum_{i=1}^{n}\sum_{s=1}^{T}\tilde{\Lambda}_s' U_{is}\bar{\bar{z}}|^2\right] = O\left(\frac{1}{n^2pT}\right).$$

$$\bar{A}_5 = E\left[\frac{1}{T}\sum_{t=1}^{T}|\frac{1}{pT}\frac{1}{n}\sum_{i=1}^{n}\sum_{s=1}^{T}\tilde{\Lambda}_s' \bar{U}_{i\cdot}\tilde{z}_{it}|^2\right] = 0 \quad (\text{because } \frac{1}{T}\sum_s \tilde{\Lambda}_s = 0).$$

$$\bar{A}_6 = E\left[\frac{1}{T}\sum_{t=1}^{T}|\frac{1}{pT}\frac{1}{n}\sum_{i=1}^{n}\sum_{s=1}^{T}\tilde{\Lambda}_s' \bar{U}_{\cdot s}\tilde{z}_{it}|^2\right] = 0 \quad (\text{because } \frac{1}{n}\sum_i \tilde{z}_{it} = 0).$$

$$\bar{A}_7 = E\left[\frac{1}{T}\sum_{t=1}^{T}|\frac{1}{pT}\frac{1}{n}\sum_{i=1}^{n}\sum_{s=1}^{T}\tilde{\Lambda}_s' \bar{\bar{U}}\tilde{z}_{it}|^2\right] = 0.$$

■

The following lemmas are used to bound $\frac{1}{T}\sum_{t=1}^{T}\|\frac{1}{n}\sum_{i=1}^{n}\tilde{e}_{it}(\widehat{f}_i - H'\tilde{f}_i)\|_2^2$ and $\max_{s\leq T, m\leq p}|\frac{1}{n}\sum_{i=1}^{n}(H'\tilde{f}_i - \widehat{f}_i)\tilde{U}_{is,m}|$.

Lemma I.9. *(i)* $\max_{m,s}\left\|\frac{1}{pTn^2}\sum_{i,j}\widehat{f}_j\sum_{t=1}^{T}(U_{it}'U_{jt} - EU_{it}'U_{jt})\tilde{U}_{is,m}\right\|_2 = O_P\left(\frac{1}{n} + \frac{1}{pT}\right).$
(ii) $\max_m\left\|\frac{1}{pT}\sum_{t=1}^{T}\tilde{\Lambda}_t'\frac{1}{n}\sum_{i=1}^{n}U_{it}\bar{U}_{i\cdot,m}\right\|_2 = O_P\left(\sqrt{\frac{\log p}{nT^2p}} + \frac{1}{pT}\right).$
(iii) $\max_{m,s}\left\|\frac{1}{pTn^2}\sum_{i,j}\widehat{f}_j\sum_{t=1}^{T}\tilde{f}_j'\tilde{\Lambda}_t'(\tilde{U}_{it}\tilde{U}_{is,m} - U_{it}U_{is,m})\right\|_2 = O_P\left(\sqrt{\frac{\log p}{nT^2p}} + \frac{1}{pT} + \sqrt{\frac{\log(pT)}{n^2pT}}\right).$
(iv) $\max_{mktl}\left\|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}(U_{it,l}U_{is,m} - EU_{it,l}U_{is,m})\tilde{g}_{sk}'\right\|_F = O_P\left(\sqrt{\frac{\log(pT)}{nT}}\right)$*, Here* $\{g_{sk} : k\leq p\}$*, as in Lemma I.3, is a deterministic sequence of vectors of fixed dimension with* $\max_{sk}\|g_{sk}\|_2 = O(1)$*, and* $\tilde{g}_{sk} = g_{sk} - \frac{1}{T}\sum_s g_{sk}$*.*

Proof. (i) Let $\tilde{W}_i = \frac{1}{pTn}\sum_{j=1}^{n}\widehat{f}_j\sum_{t=1}^{T}(U_{it}'U_{jt} - EU_{it}'U_{jt})$. We first bound $\frac{1}{n}\sum_i\|\tilde{W}_i\|_2^2$.

$$\frac{1}{n}\sum_i\|\tilde{W}_i\|_2^2 \leq \frac{2}{n}\sum_i\left\|\frac{1}{pTn}\sum_{j=1}^{n}(\widehat{f}_j - H'\tilde{f}_j)\sum_{t=1}^{T}(U_{it}'U_{jt} - EU_{it}'U_{jt})\right\|_2^2$$
$$+ \frac{2}{n}\sum_i\left\|\frac{1}{pTn}\sum_{j=1}^{n}H'\tilde{f}_j\sum_{t=1}^{T}(U_{it}'U_{jt} - EU_{it}'U_{jt})\right\|_2^2$$
$$\leq \frac{2}{p^2T^2n^3}\sum_{l=1}^{n}\|\widehat{f}_l - H'\tilde{f}_l\|_2^2\sum_i\sum_{j=1}^{n}(\sum_{t=1}^{T}U_{it}'U_{jt} - EU_{it}'U_{jt})^2$$

$$
\begin{aligned}
&+ O_P\left(\frac{1}{p^2T^2n^3}\right)\sum_i \left\|\sum_{j=1}^n f_j \sum_{t=1}^T (U_{it}'U_{jt} - EU_{it}'U_{jt})\right\|_2^2 \\
&+ O_P\left(\frac{\|\bar f\|_2^2}{p^2T^2n^3}\right)\sum_i \left|\sum_{j=1}^n \sum_{t=1}^T (U_{it}'U_{jt} - EU_{it}'U_{jt})\right|^2 \\
= {}& O_P\left(\frac{\Delta_F^2}{p^2T^2n^2}\right)\sum_i \sum_{j=1}^n \mathrm{Var}\left(\sum_{t=1}^T U_{it}'U_{jt}\right) + O_P\left(\frac{1}{p^2T^2n^3}\right)\sum_i \mathrm{Var}\left(\sum_{j=1}^n \sum_{t=1}^T U_{it}'U_{jt}\right) \\
&+ O_P\left(\frac{1}{p^2T^2n^3}\right)\sum_i E\left\|\sum_{j=1}^n \sum_{t=1}^T f_j(U_{it}'U_{jt} - EU_{it}'U_{jt})\right\|_2^2 .
\end{aligned}
$$

Below we bound each of the three terms on the right of the equal sign in the preceding display. First,

$$
\begin{aligned}
\sum_i \sum_{j=1}^n \mathrm{Var}\left(\sum_{t=1}^T U_{it}'U_{jt}\right) &= \sum_i \sum_{j=1}^n \sum_{t=1}^T \sum_{m=1}^p \sum_{s=1}^T \sum_{v=1}^p \mathrm{Cov}(U_{it,m}U_{jt,m}, U_{is,v}U_{js,v}) \\
&\le pnT \max_{mti} \sum_{s=1}^T \sum_{v=1}^p \mathrm{Cov}(U_{it,m}^2, U_{is,v}^2) + O(n^2Tp) \max_{tm} \sum_{s=1}^T \sum_{v=1}^p |EU_{it,m}U_{is,v}| \\
&= O(n^2Tp).
\end{aligned}
$$

Hence, $O_P\left(\frac{\Delta_F^2}{p^2T^2n^2}\right)\sum_i \sum_{j=1}^n \mathrm{Var}\left(\sum_{t=1}^T U_{it}'U_{jt}\right) = O_P\left(\frac{\Delta_F^2}{pT}\right)$.

Second,

$$
\begin{aligned}
\sum_i \mathrm{Var}\left(\sum_{j=1}^n \sum_{t=1}^T U_{it}'U_{jt}\right) &= \sum_i \sum_{j=1}^n \sum_{m=1}^p \sum_{t=1}^T \sum_{l=1}^n \sum_{v=1}^p \sum_{s=1}^T \mathrm{Cov}(U_{it,m}U_{jt,m}, U_{is,v}U_{ls,v}) \\
&= \sum_i \sum_{m=1}^p \sum_{t=1}^T \sum_{v=1}^p \sum_{s=1}^T \mathrm{Cov}(U_{it,m}^2, U_{is,v}^2) \\
&\quad + \sum_i \sum_{j\neq i} \sum_{m=1}^p \sum_{t=1}^T \sum_{v=1}^p \sum_{s=1}^T EU_{it,m}U_{is,v}EU_{js,v}U_{jt,m} \\
&\le O(Tpn) \max_{tmi} \sum_{v=1}^p \sum_{s=1}^T \mathrm{Cov}(U_{it,m}^2, U_{is,v}^2)
\end{aligned}
$$

$$+ O(n^2Tp)\max_{tmi}\sum_{v=1}^{p}\sum_{s=1}^{T}|EU_{it,m}U_{is,v}|$$
$$= O(n^2Tp).$$

Hence, $O_P\left(\frac{1}{p^2T^2n^3}\right)\sum_i \mathrm{Var}\left(\sum_{j=1}^n\sum_{t=1}^T U_{it}'U_{jt}\right) = O\left(\frac{1}{pTn}\right)$.

Third, for each $k \leq K$,

$$\sum_i E\left|\sum_{j=1}^{n}\sum_{t=1}^{T} f_{jk}(U_{it}'U_{jt} - EU_{it}'U_{jt})\right|^2$$
$$\leq 2\sum_i E\left|\sum_{t=1}^{T} f_{ik}(\|U_{it}\|_2^2 - E\|U_{it}\|_2^2)\right|^2 + 2\sum_i \mathrm{Var}\left(\sum_{j\neq i}\sum_{t=1}^{T} f_{jk}U_{it}'U_{jt}\right)$$
$$= O(np^2T^2) + 2\sum_i\sum_{l\neq i}\sum_{s=1}^{T}\sum_{v=1}^{p}\sum_{t=1}^{T}\sum_{m=1}^{p} E\left[U_{it,m}U_{is,v}\right]E\left[f_{lk}^2U_{lt,m}U_{ls,v}\right]$$
$$= O(np^2T^2) + O(n^2Tp)\max_{tmi}\sum_{s=1}^{T}\sum_{v=1}^{p}|EU_{it,m}U_{is,v}|$$
$$= O(np^2T^2 + n^2Tp).$$

Hence, $O_P\left(\frac{1}{p^2T^2n^3}\right)\sum_i E\left\|\sum_{j=1}^n\sum_{t=1}^T f_j(U_{it}'U_{jt} - EU_{it}'U_{jt})\right\|_2^2 = O_P\left(\frac{1}{n^2} + \frac{1}{pTn}\right)$. Therefore,

$$\frac{1}{n}\sum_i\left\|\frac{1}{pTn}\sum_{j=1}^{n}\tilde{f}_j\sum_{t=1}^{T}(U_{it}'U_{jt} - EU_{it}'U_{jt})\right\|_2^2 \leq O_P\left(\frac{1}{pTn} + \frac{1}{n^2}\right). \tag{I.12}$$

By $\Delta_F^2 = \frac{1}{n^2} + \frac{1}{nT^2} + \frac{1}{pT}$, $\frac{1}{n}\sum_i\|\tilde{W}_i\|_2^2 = O_P\left(\frac{\Delta_F^2}{pT} + \frac{1}{pTn} + \frac{1}{n^2}\right) = O_P\left(\frac{1}{n^2} + \frac{1}{p^2T^2}\right)$. Hence, the object we aim to bound is $\max_{ms}\left\|\frac{1}{n}\sum_{i=1}^n\tilde{W}_i\tilde{U}_{is,m}\right\|_2 \leq \left(\frac{1}{n}\sum_i\|\tilde{W}_i\|_2^2\right)^{1/2}\max_{ms}\left(\frac{1}{n}\sum_i\tilde{U}_{is,m}^2\right)^{1/2}$. We now show $\max_{ms}\left(\frac{1}{n}\sum_i\tilde{U}_{is,m}^2\right)^{1/2} = O_P(1)$. Once this is done, then $\max_{ms}\left\|\frac{1}{n}\sum_{i=1}^n\tilde{W}_i\tilde{U}_{is,m}\right\|_2 = O_P\left(\frac{1}{n} + \frac{1}{pT}\right)$ as desired. Note that

$$\max_{ms}\frac{1}{n}\sum_i\tilde{U}_{is,m}^2 \leq \max_{ms}\frac{2}{n}\sum_i(U_{is,m} - \bar{U}_{i\cdot,m})^2 + 2\max_{ms}(\bar{U}_{\cdot s,m} - \bar{\bar{U}}_m)^2$$

Let $\mathcal{U}_{i,ms} = 2(U_{ms,i} - \bar{U}_{i\cdot,m})^2 - 2E(U_{ms,i} - \bar{U}_{i\cdot,m})^2$. Then $E\mathcal{U}_{i,ms} = 0$ and $\max_{i,ms} \operatorname{Var}(\mathcal{U}_{i,ms}) < C$. Hence by the Bernstein inequality for independent data, we have

$$\max_{ms} \frac{1}{n} \sum_i \mathcal{U}_{i,ms} = O_P(\sqrt{\frac{\log(Tp)}{n}}) = o_P(1).$$

This implies $\max_{ms} \frac{2}{n} \sum_i (U_{is,m} - \bar{U}_{i\cdot,m})^2 \leq o_P(1) + \max_{ms} \frac{2}{n} \sum_i E(U_{is,m} - \bar{U}_{i\cdot,m})^2 = O_P(1)$. On the other hand,

$$\begin{aligned}
\max_{ms} (\bar{U}_{\cdot s,m} - \bar{\bar{U}}_m)^2 &\leq 2(\max_{ms} |\bar{U}_{\cdot s,m}|)^2 + 2(\max_{ms} |\bar{\bar{U}}_m|)^2 \\
&\leq 2(\max_{ms} |\bar{U}_{\cdot s,m}|)^2 + 2(\max_{ms} |\frac{1}{T} \sum_s \bar{U}_{\cdot s,m}|)^2 \leq 4(\max_{ms} |\bar{U}_{\cdot s,m}|)^2 \\
&\leq 4(\max_{ms} |\frac{1}{n} \sum_i U_{is,m}|)^2 = O_P\left(\frac{\log(Tp)}{n}\right).
\end{aligned}$$

This shows $\max_{ms} \frac{1}{n} \sum_i \tilde{U}_{is,m}^2 = O_P(1)$.

(ii) Define $\Gamma_{im} = \frac{1}{pT^2} \sum_{t=1}^T \sum_{s=1}^T \tilde{\Lambda}_t'(U_{it}U_{is,m} - EU_{it}U_{is,m})$. Without loss of generality, we assume $\tilde{\Lambda}_t'$ has a single row, as we can always conduct the analysis for each fixed row of $\tilde{\Lambda}_t'$. Then $EX_{im} = 0$, and

$$\begin{aligned}
\max_{im} \operatorname{Var}(X_{im}) &\leq \max_{im} \frac{1}{T^4 p^2} \sum_{t=1}^T \sum_{k=1}^p \sum_{h=1}^T \sum_{l=1}^p \sum_{v=1}^T \sum_{s=1}^T |\operatorname{Cov}(U_{it,k}U_{is,m}, U_{ih,l}U_{iv,m})| \\
&= O\left(\frac{1}{T^2 p}\right).
\end{aligned}$$

By the Bernstein inequality for independent data, $\max_m \|\frac{1}{n} \sum_{i=1}^n X_{im}\|_2 = O_P\left(\sqrt{\frac{\log p}{nT^2 p}}\right)$. Thus,

$$\begin{aligned}
\max_m \left\| \frac{1}{pT} \sum_{t=1}^T \tilde{\Lambda}_t' \frac{1}{n} \sum_{i=1}^n U_{it}\bar{U}_{i\cdot,m} \right\|_2 &\leq \max_m \left\| \frac{1}{n} \sum_{i=1}^n X_{im} \right\|_2 + \max_m \left\| \frac{1}{pT^2} \sum_{t=1}^T \frac{1}{n} \sum_{s=1}^T \sum_{i=1}^n \tilde{\Lambda}_t' EU_{it}U_{is,m} \right\|_2 \\
&\leq O_P\left(\sqrt{\frac{\log p}{nT^2 p}} + \frac{1}{pT}\right).
\end{aligned}$$

(iii) With $\left\| \frac{1}{n} \sum_i \widehat{f}_j \tilde{f}_j' \right\| = O_P(1)$ and the equalities $\frac{1}{T} \sum_t \tilde{\Lambda}_t = 0$ and

$$\frac{1}{n} \sum_{i=1}^n (\tilde{U}_{it}\tilde{U}_{is,m} - U_{it}U_{is,m}) = -\frac{1}{n} \sum_{i=1}^n U_{it}\bar{U}_{i\cdot,m} - \bar{U}_{\cdot t}\bar{U}_{\cdot s,m} + \bar{U}_{\cdot t}\bar{\bar{U}}_m - \frac{1}{n} \sum_{i=1}^n \bar{U}_{i\cdot}U_{is,m}$$

$$+\bar{\bar{U}}\bar{U}_{\cdot s,m}+\frac{1}{n}\sum_{i=1}^{n}\bar{U}_{i\cdot}\bar{U}_{i\cdot,m}-\bar{\bar{U}}\bar{\bar{U}}_{m},$$

we have that $\max_{m,s}|\frac{1}{n}\sum_{i=1}^{n}\frac{1}{pTn}\sum_{j=1}^{n}\widehat{f}_j\sum_{t=1}^{T}\tilde{f}_j'\tilde{\Lambda}_t'(\tilde{U}_{it}\tilde{U}_{is,m}-U_{it}U_{is,m})|$ is bounded by

$$\begin{aligned}
&O_P(1)\max_{ms}\left\|\frac{1}{pT}\sum_{t=1}^{T}\tilde{\Lambda}_t'\frac{1}{n}\sum_{i=1}^{n}(\tilde{U}_{it}\tilde{U}_{is,m}-U_{it}U_{is,m})\right\|_2\\
&\quad\le O_P(1)\max_{m}\left\|\frac{1}{pT}\sum_{t=1}^{T}\tilde{\Lambda}_t'\frac{1}{n}\sum_{i=1}^{n}U_{it}\bar{U}_{i\cdot,m}\right\|_2+O_P(1)\max_{ms}\left\|\frac{1}{pT}\sum_{t=1}^{T}\tilde{\Lambda}_t'\bar{U}_{\cdot t}\bar{U}_{\cdot s,m}\right\|_2\\
&\qquad+O_P(1)\max_{m}\left\|\frac{1}{pT}\sum_{t=1}^{T}\tilde{\Lambda}_t'\bar{U}_{\cdot t}\bar{\bar{U}}_m\right\|_2\\
&\quad\le O_P(1)\max_{m}\left\|\frac{1}{pT}\sum_{t=1}^{T}\tilde{\Lambda}_t'\frac{1}{n}\sum_{i=1}^{n}U_{it}\bar{U}_{i\cdot,m}\right\|_2+O_P(1)\left\|\frac{1}{pT}\sum_{t=1}^{T}\tilde{\Lambda}_t'\bar{U}_{\cdot t}\right\|_2\max_{ms}\left(|\bar{U}_{\cdot s,m}|+|\bar{\bar{U}}_m|\right)\\
&\quad=O_P\left(\sqrt{\frac{\log p}{nT^2p}}+\frac{1}{pT}+\sqrt{\frac{\log(pT)}{n^2pT}}\right),
\end{aligned}$$

where the last equality follows from part (ii) that

$$\max_{m}\left\|\frac{1}{pT}\sum_{t=1}^{T}\tilde{\Lambda}_t'\frac{1}{n}\sum_{i=1}^{n}U_{it}\bar{U}_{i\cdot,m}\right\|_2=O_P\left(\sqrt{\frac{\log p}{nT^2p}}+\frac{1}{pT}\right),$$

that $\max_{ms}(|\bar{U}_{\cdot s,m}|+|\bar{\bar{U}}_m|)=O_P\left(\sqrt{\frac{\log pT}{n}}\right)$, and that $\left\|\frac{1}{pT}\sum_{t=1}^{T}\tilde{\Lambda}_t'\bar{U}_{\cdot t}\right\|_2=O\left(\frac{1}{\sqrt{pTn}}\right)$.

(iv) As the dimension of g_{sk} does not grow, without loss of generality, we assume it is a scalar. Let $X_{i,mktl}=\frac{1}{T}\sum_{s=1}^{T}(U_{it,l}U_{is,m}-EU_{it,l}U_{is,m})\tilde{g}_{sk}$. Then $EX_{i,mktl}=0$ and we are calculating the rate of $\max_{mktl}|\frac{1}{n}\sum_{i=1}^{n}X_{i,mktl}|$. Note that by Assumption D.2,

$$\max_{imktl}\operatorname{Var}(X_{i,mktl})\le\max_{imt}\frac{1}{T^2}\sum_{s=1}^{T}\sum_{v=1}^{T}|\operatorname{Cov}(U_{it,l}U_{is,m},U_{it,l}U_{iv,m})|=O\left(\frac{1}{T}\right).$$

Hence by the Bernstein inequality with independent data,

$$\max_{mktl}|\frac{1}{n}\sum_{i=1}^{n}X_{i,mktl}|=O_P\left(\sqrt{\frac{\log(pT)}{nT}}\right).$$

■

Lemma I.10. *(i)* $\max_{s\leq T,m\leq p}\left|\frac{1}{n}\sum_{i=1}^n(H'\tilde{f}_i-\widehat{f}_i)\tilde{U}_{is,m}\right|=O_P\left(\frac{1}{n}+\frac{1}{T\sqrt{n}}+\frac{1}{T\sqrt{p}}+\sqrt{\frac{\log(pT)}{npT}}\right)$.
(ii) $\frac{1}{T}\sum_{t=1}^T\left\|\frac{1}{n}\sum_{i=1}^n\tilde{e}_{it}(\widehat{f}_i-H'\tilde{f}_i)\right\|_2^2=O_P\left(\frac{1}{n^2}+\frac{1}{Tpn}\right)$.

Note that the rate of convergence in (i) is better than the rate obtained from the simple inequality

$$\max_{s\leq T,m\leq p}\left|\frac{1}{n}\sum_{i=1}^n(H'\tilde{f}_i-\widehat{f}_i)\tilde{U}_{is,m}\right|\leq O_P(\Delta_F)\max_{ms}\sqrt{\frac{1}{n}\sum_i\tilde{U}_{is,m}^2}=O_P\left(\frac{1}{n}+\frac{1}{T\sqrt{n}}+\frac{1}{\sqrt{pT}}\right)$$

in some regimes such as when $p/n\ll T\ll n^2/p$.

Proof. (i) We have $\frac{1}{n}\sum_{i=1}^n(\widehat{f}_i-H'\tilde{f}_i)\tilde{U}_{is,m}=V^{-1}\sum_{l=1}^7 D_{l,ms}$ where

$$D_{1,ms}=\frac{1}{n}\sum_{i=1}^n\frac{1}{pTn}\sum_{j=1}^n\widehat{f}_j\sum_{t=1}^T(U_{it}'U_{jt}-EU_{it}'U_{jt})\tilde{U}_{is,m},$$

$$D_{2,ms}=\frac{1}{n}\sum_{i=1}^n\frac{1}{pTn}\sum_{j=1}^n\widehat{f}_j\sum_{t=1}^T\tilde{f}_j'\tilde{\Lambda}_t'(\tilde{U}_{it}\tilde{U}_{is,m}-U_{it}U_{is,m})$$

$$D_{3,ms}=\frac{1}{n}\sum_{i=1}^n\frac{1}{pTn}\sum_{j=1}^n\widehat{f}_j\sum_{t=1}^T(EU_{it}'U_{jt})\tilde{U}_{is,m},$$

$$D_{4,ms}=\frac{1}{n}\sum_{i=1}^n\frac{1}{pTn}\sum_{j=1}^n\widehat{f}_j\sum_{t=1}^T\tilde{f}_j'\tilde{\Lambda}_t'(EU_{it}U_{is,m}),$$

$$D_{5,ms}=\frac{1}{n}\sum_{i=1}^n\frac{1}{pTn}\sum_{j=1}^n\widehat{f}_j\sum_{t=1}^T(\tilde{U}_{it}'\tilde{U}_{jt}-U_{it}'U_{jt})\tilde{U}_{is,m},$$

$$D_{6,ms}=\frac{1}{n}\sum_{i=1}^n\frac{1}{pTn}\sum_{j=1}^n\widehat{f}_j\sum_{t=1}^T\tilde{f}_i'\tilde{\Lambda}_t'\tilde{U}_{jt}\tilde{U}_{is,m},$$

$$D_{7,ms}=\frac{1}{n}\sum_{i=1}^n\frac{1}{pTn}\sum_{j=1}^n\widehat{f}_j\sum_{t=1}^T\tilde{f}_j'\tilde{\Lambda}_t'(U_{it}U_{is,m}-EU_{it}U_{is,m}).$$

We respectively bound these terms.

Bounding $\mathbf{D_{1,ms}}$ and $\mathbf{D_{2,ms}}$: The bounds

$$\max_{ms}\|D_{1,ms}\|_2=O_P\left(\frac{1}{n}+\frac{1}{pT}\right)$$

and

$$\max_{ms}\|D_{2,ms}\|_2 = O_P\left(\sqrt{\frac{\log p}{nT^2p}} + \frac{1}{pT} + \sqrt{\frac{\log(pT)}{n^2pT}}\right)$$

directly follow from Lemma I.9.

Bounding $\mathbf{D_{3,ms}}$ and $\mathbf{D_{4,ms}}$:

$$\max_{ms}\|D_{3,ms}\|_2 = \max_{ms}\left\|\frac{1}{n}\sum_{i=1}^{n}\frac{1}{n}\widehat{f}_i\tilde{U}_{is,m}\frac{1}{Tp}\sum_{t=1}^{T}E\|U_{it}\|_2^2\right\|_2 = O_P\left(\frac{1}{n}\right).$$

$$\max_{ms}\|D_{4,ms}\|_2 = O_P(1)\max_{ms}\frac{1}{pTn}\sum_{i=1}^{n}\sum_{t=1}^{T}\sum_{v=1}^{p}|EU_{it,v}U_{is,m}| = O_P\left(\frac{1}{pT}\right).$$

Bounding $\mathbf{D_{5,ms}}$: By the Cauchy-Schwarz inequality and Lemma I.6, we have

$$\max_{ms}\|D_{5,ms}\|_2 \leq \max_{ms}\left(\frac{1}{n}\sum_{i}\tilde{U}_{is,m}^2\right)^{1/2}\left(\frac{1}{n}\sum_{i=1}^{n}\|A_{i1}\|_2^2\right)^{1/2}$$
$$= O_P\left(\frac{1}{n} + \frac{1}{\sqrt{npT}} + \frac{1}{T\sqrt{n}} + \frac{1}{T\sqrt{p}}\right).$$

Bounding $\mathbf{D_{6,ms}}$: First note that $\left\|\frac{1}{npT}\sum_{j=1}^{n}\sum_{t=1}^{T}\tilde{\Lambda}_t'\tilde{U}_{jt}\widehat{f}_j'\right\|_F \leq O_P\left(\frac{\Delta_F}{\sqrt{pT}} + \frac{1}{\sqrt{pTn}}\right)$. Then

$$\max_{ms}\|D_{6,ms}\|_2 \leq O(1)\max_{ms}\left\|\frac{1}{n}\sum_{i=1}^{n}\tilde{U}_{is,m}\tilde{f}_i\right\|_2\left\|\frac{1}{npT}\sum_{j=1}^{n}\sum_{t=1}^{T}\tilde{\Lambda}_t'\tilde{U}_{jt}\widehat{f}_j'\right\|_F$$
$$= O_P\left(\frac{\Delta_F}{\sqrt{pT}} + \frac{1}{\sqrt{pTn}}\right)\sqrt{\frac{\log(pT)}{n}}.$$

Bounding $\mathbf{D_{7,ms}}$: Let $X_{i,ms} = \frac{1}{pT}\sum_{t=1}^{T}\tilde{\Lambda}_t'(U_{it}U_{is,m} - EU_{it}U_{is,m})$. Then $EX_{i,ms} = 0$; and for each $k \leq K$ where K denotes the number of factors,

$$\max_{ims}\operatorname{Var}(X_{i,ms}) = \max_{ims}\frac{1}{p^2T^2}\operatorname{Var}\left(\sum_{t=1}^{T}\sum_{v=1}^{p}\lambda_{tk,v}U_{it,v}U_{is,m}\right)$$
$$= \max_{ims}\frac{1}{p^2T^2}\sum_{j=1}^{T}\sum_{l=1}^{p}\sum_{t=1}^{T}\sum_{v=1}^{p}\lambda_{jk,l}\lambda_{tk,v}\operatorname{Cov}(U_{it,v}U_{is,m}, U_{ij,l}U_{is,m})$$
$$\leq O\left(\frac{1}{pT}\right)\max_{imstv}\sum_{k=1}^{T}\sum_{l=1}^{p}|\operatorname{Cov}(U_{it,v}U_{is,m}, U_{ik,l}U_{is,m})| = O\left(\frac{1}{pT}\right).$$

Hence, by the Bernstein inequality for independent data,

$$\max_{ms}\|D_{7,ms}\|_2 \quad \le \quad O_P(1)\max_{ms}\|\frac{1}{n}\sum_{i=1}^{n}X_{i,ms}\|_2 = O_P\left(\sqrt{\frac{\log(pT)}{npT}}\right).$$

Combining the above results, we reach

$$\max_{s\le T, m\le p}|\frac{1}{n}\sum_{i=1}^{n}(H'\tilde f_i-\widehat f_i)\tilde U_{is,m}| = O_P\left(\frac{1}{n}+\frac{1}{T\sqrt{n}}+\frac{1}{T\sqrt{p}}+\sqrt{\frac{\log(pT)}{npT}}\right).$$

(ii) $\frac{1}{T}\sum_{t=1}^{T}\|\frac{1}{n}\sum_{i=1}^{n}\tilde e_{it}(\widehat f_i - H'\tilde f_i)\|_2^2 \le \sum_{l=1}^{3}\bar B_l$ where each $\bar B_l$ is defined and bounded below making use of (I.1).

$$\begin{aligned}
\bar B_1 &= \frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{n}\sum_{i=1}^{n}\tilde e_{is}\frac{1}{pTn}\sum_{j=1}^{n}\widehat f_j\sum_{t=1}^{T}\tilde U_{it}'\tilde U_{jt}\right\|_2^2\\
&\le \frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{n}\sum_{i=1}^{n}\tilde e_{is}\frac{1}{pTn}\sum_{j=1}^{n}\tilde f_j\sum_{t=1}^{T}\tilde U_{it}'\tilde U_{jt}\right\|_2^2 O_P(1)\\
&\quad + \frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{n}\sum_{j=1}^{n}(\widehat f_j - H'\tilde f_j)\tilde U_{jt}'\right\|_F^2\frac{1}{T^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\left\|\frac{1}{pn}\sum_{i=1}^{n}\tilde U_{it}\tilde e_{is}\right\|_2^2\\
&\le \frac{1}{T}\sum_{s=1}^{T}\frac{1}{T}\sum_{t}\left\|\frac{1}{n}\sum_{i=1}^{n}\tilde e_{is}\tilde U_{it}\right\|_2^2\frac{1}{T}\sum_{t}\left\|\frac{1}{pn}\sum_{j=1}^{n}\tilde f_j\tilde U_{jt}\right\|^2 O_P(1)\\
&\quad + \frac{1}{pT^2}\sum_{s=1}^{T}\sum_{t=1}^{T}\left\|\frac{1}{n}\sum_{i=1}^{n}\tilde U_{it}\tilde e_{is}\right\|_2^2\max_{tm}\left\|\frac{1}{n}\sum_{j=1}^{n}(\widehat f_j - H'\tilde f_j)\tilde U_{jt,m}\right\|_2^2\\
&\le O_P\left(\frac{1}{n^2}\right)+\frac{1}{n}O_P\left(\frac{1}{n^2}+\frac{1}{T^2n}+\frac{1}{T^2p}+\frac{\log(pT)}{npT}\right)\\
&= O_P\left(\frac{1}{n^2}+\frac{1}{T^2pn}\right). \text{ (Lemma I.10)}\\
\bar B_2 &= \frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{n}\sum_{i=1}^{n}\tilde e_{is}\frac{1}{pTn}\sum_{j=1}^{n}\widehat f_j\sum_{t=1}^{T}\tilde f_j'\tilde\Lambda_t'\tilde U_{it}\right\|_2^2\\
&\le O_P(1)\frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{n}\sum_{i=1}^{n}\tilde e_{is}\frac{1}{pT}\sum_{t=1}^{T}\tilde\Lambda_t'\tilde U_{it}\right\|_2^2
\end{aligned}$$

$$= O\left(\frac{1}{pTn}\right). \text{ (Lemma I.8)}$$

$$\bar{B}_3 = \frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{n}\sum_{i=1}^{n}\tilde{e}_{is}\frac{1}{pTn}\sum_{j=1}^{n}\widehat{f}_j\sum_{t=1}^{T}\tilde{f}_i'\tilde{\Lambda}_t'\tilde{U}_{jt}\right\|_2^2$$

$$\leq \frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{n}\sum_{i=1}^{n}\tilde{e}_{is}\frac{1}{pTn}\sum_{j=1}^{n}\tilde{f}_j\sum_{t=1}^{T}\tilde{f}_i'\tilde{\Lambda}_t'\tilde{U}_{jt}\right\|_2^2 O_P(1)$$

$$+ \frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{n}\sum_{j=1}^{n}(\widehat{f}_j - H'\tilde{f}_j)\tilde{U}_{jt}'\right\|_F^2 \frac{1}{T}\sum_{s=1}^{T}\frac{1}{p}\left\|\frac{1}{n}\sum_{i=1}^{n}\tilde{f}_i'\tilde{e}_{is}\right\|_2^2$$

$$= O_P\left(\frac{1}{n^2pT}\right) + \frac{1}{n}O_P\left(\frac{1}{n^2} + \frac{1}{T^2n} + \frac{1}{T^2p} + \frac{\log(pT)}{npT}\right)$$

$$= O_P\left(\frac{1}{n^3} + \frac{1}{T^2n^2} + \frac{1}{T^2pn} + \frac{\log(pT)}{n^2pT}\right).$$

Hence,

$$\frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{n}\sum_{i=1}^{n}\tilde{e}_{it}(\widehat{f}_i - H'\tilde{f}_i)\right\|_2^2 = O_P\left(\frac{1}{n^2} + \frac{1}{Tpn}\right).$$

■

I.4. **Technical lemmas for Proposition D.1: the bootstrap data.**

Lemma I.11. *(i)* $\frac{1}{n}\sum_{j=1}^{n}\left(\frac{1}{pT}\sum_{m=1}^{p}\sum_{t=1}^{T}\frac{1}{n}\tilde{F}'\tilde{U}_{t,m}\tilde{U}_{jt,m}\right)^2 = O_P\left(\frac{1}{n^2} + \frac{1}{nT^2} + \frac{1}{pTn}\right)$.
(ii) $\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{pT}\sum_{t=1}^{T}\sum_{m=1}^{p}\frac{1}{n}\tilde{F}'\tilde{U}_{t,m}\widehat{U}_{it,m}\right)^2 = O_P\left(\frac{1}{n^2} + \frac{1}{nT^2} + \frac{1}{pTn}\right)$.
(iii) $\frac{1}{n}\sum_{i=1}^{n}\|A_{i1}^*\|_2^2 = O_{P^*}\left(\Delta_F^2 + \frac{1}{n}\right)$ *for* A_{i1}^* *defined in (I.8).*

Proof. (i) Note that $\sum \tilde{f}_i = \sum_i \tilde{U}_{it} = \sum_t \tilde{U}_{it} = 0$. We may then bound

$$\frac{1}{n}\sum_{j=1}^{n}\left(\frac{1}{pT}\sum_{m=1}^{p}\sum_{t=1}^{T}\frac{1}{n}\tilde{F}'\tilde{U}_{t,m}\tilde{U}_{jt,m}\right)^2$$

from above with

$$\frac{C}{n}\sum_{j=1}^{n}(\frac{1}{pT}\sum_{m=1}^{p}\sum_{t=1}^{T}\frac{1}{n}\sum_{i=1}^{n}\tilde{f}_i\tilde{U}_{it,m}U_{jt,m})^2 + \frac{C}{n}\sum_{j=1}^{n}(\frac{1}{pT}\sum_{m=1}^{p}\sum_{t=1}^{T}\frac{1}{n}\sum_{i=1}^{n}\tilde{f}_i\tilde{U}_{it,m}\bar{U}_{\cdot t,m})^2$$

$$\le O_P\left(\frac{1}{n^2}\right) + \frac{C}{n}\sum_{j=1}^n\left(\frac{1}{pT}\sum_{m=1}^p\sum_{t=1}^T\frac{1}{n}\sum_{i=1}^n \tilde{f}_i U_{it,m}U_{jt,m}\right)^2$$
$$+ \frac{C}{n}\sum_{j=1}^n\left(\frac{1}{pT}\sum_{m=1}^p\sum_{t=1}^T\frac{1}{n}\sum_{i=1}^n \tilde{f}_i \bar{U}_{i\cdot,m}U_{jt,m}\right)^2$$
$$\le O_P\left(\frac{1}{n^2}+\frac{1}{nT^2}\right) + \frac{C}{n}\sum_{j=1}^n\left(\frac{1}{pT}\sum_{m=1}^p\sum_{t=1}^T\frac{1}{n}\sum_{i=1}^n \tilde{f}_i (U_{it,m}U_{jt,m} - EU_{it,m}U_{jt,m})\right)^2$$
$$+ \frac{C}{n}\sum_{j=1}^n\left(\frac{1}{pT}\sum_{m=1}^p\sum_{t=1}^T\frac{1}{n}\sum_{i=1}^n \tilde{f}_i EU_{it,m}U_{jt,m}\right)^2$$
$$= O_P\left(\frac{1}{n^2}+\frac{1}{nT^2}+\frac{1}{pTn}\right),$$

where we used (I.12) that

$$\frac{1}{n}\sum_i\left\|\frac{1}{pTn}\sum_{j=1}^n\tilde{f}_j\sum_{t=1}^T(U_{it}'U_{jt} - EU_{it}'U_{jt})\right\|_2^2 \le O_P\left(\frac{1}{pTn}+\frac{1}{n^2}\right).$$

(ii) The conclusion follows from part (i) and the following inequality:

$$\frac{1}{n}\sum_{i=1}^n\left(\frac{1}{pT}\sum_{t=1}^T\sum_{m=1}^p\frac{1}{n}\tilde{F}'\tilde{U}_{t,m}(\widehat{U}_{it,m}-\tilde{U}_{it,m})\right)^2$$
$$\le \left(\frac{1}{pT}\sum_{t=1}^T\sum_{m=1}^p\frac{1}{n}\tilde{F}'\tilde{U}_{t,m}\tilde{\lambda}_{tm}\right)^2 O_P(\Delta_F^2) + \left(\frac{1}{pT}\sum_{t=1}^T\sum_{m=1}^p\left\|\frac{1}{n}\tilde{F}'\tilde{U}_{t,m}\right\|^2\right)^2$$
$$= O_P\left(\frac{\Delta_F^2}{nTp}+\frac{1}{n^2}\right).$$

(iii) Note that $\max_{ij}\mathrm{Var}^*(w_i^U w_j^U) < C$.

$$\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n E^*\left(\frac{1}{pT}\sum_{t=1}^T\tilde{U}_{it}^{*\prime}\tilde{U}_{jt}^* - E^*\tilde{U}_{it}^{*\prime}\tilde{U}_{jt}^*\right)^2$$
$$= \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n \mathrm{Var}^*\left(\frac{1}{pT}\sum_{t=1}^T\tilde{U}_{it}^{*\prime}\tilde{U}_{jt}^*\right)$$

$$
\begin{aligned}
&= \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n\left(\frac{1}{pT}\sum_{t=1}^T \widehat{U}_{it}'\widehat{U}_{jt}\right)^2 \mathrm{Var}^*(w_i^U w_j^U) \\
&\le \underbrace{\frac{8}{n^2}\sum_{i=1}^n\sum_{j=1}^n\left(\frac{1}{pT}\sum_{t=1}^T(\tilde{U}_{it}'\tilde{U}_{jt} - EU_{it}'U_{jt})\right)^2 C}_{a_1} + \underbrace{\frac{8}{n^2}\sum_{i=1}^n\sum_{j=1}^n\left(\frac{1}{pT}\sum_{t=1}^T EU_{it}'U_{jt}\right)^2 C}_{a_2} \\
&\quad + \underbrace{\frac{8}{n^2}\sum_{i=1}^n\sum_{j=1}^n\left(\frac{1}{pT}\sum_{t=1}^T \widehat{U}_{it}'\widehat{U}_{jt} - \tilde{U}_{it}'\tilde{U}_{jt}\right)^2 C}_{a_3}.
\end{aligned}
$$

Note that

$$
a_1 \le \frac{C}{n^2}\sum_{i=1}^n\sum_{j=1}^n(\frac{1}{pT}\sum_{t=1}^T(\tilde{U}_{it}'\tilde{U}_{jt} - U_{it}'U_{jt}))^2 + \frac{C}{n^2}\sum_{i=1}^n\sum_{j=1}^n(\frac{1}{pT}\sum_{t=1}^T(U_{it}'U_{jt} - EU_{it}'U_{jt}))^2.
$$

The first term in the upper bound for a_1 is $O_P\left(\frac{1}{n^2} + \frac{1}{npT} + \frac{1}{nT^2} + \frac{1}{pT^2}\right)$ by Lemma I.6. The second term in the bound for a_1 is $O_P\left(\frac{1}{pT}\right)$ by (I.11). Therefore,

$$
a_1 = O_P\left(\frac{1}{n^2} + \frac{1}{nT^2} + \frac{1}{pT}\right) = O_P(\Delta_F^2). \tag{I.13}
$$

Next we have $a_2 = \frac{C}{n^2}\sum_{i=1}^n\left(\frac{1}{pT}\sum_{t=1}^T EU_{it}'U_{it}\right)^2 = O\left(\frac{1}{n}\right)$. Finally,

$$
\begin{aligned}
a_3 &= \frac{C}{n^2}\sum_{i=1}^n\sum_{j=1}^n\left(\frac{1}{pT}\sum_{m=1}^p\sum_{t=1}^T \widehat{U}_{it,m}\widehat{U}_{jt,m} - \tilde{U}_{it,m}\tilde{U}_{jt,m}\right)^2 \\
&\le \frac{C}{n^2}\sum_{i=1}^n\sum_{j=1}^n\left(\frac{1}{pT}\sum_{m=1}^p\sum_{t=1}^T(\widehat{U}_{it,m} - \tilde{U}_{it,m})\widehat{U}_{jt,m}\right)^2 \\
&\quad + \frac{C}{n^2}\sum_{i=1}^n\sum_{j=1}^n\left(\frac{1}{pT}\sum_{m=1}^p\sum_{t=1}^T(\widehat{U}_{it,m} - \tilde{U}_{it,m})\tilde{U}_{jt,m}\right)^2 \\
&\le \frac{C}{n}\sum_{j=1}^n\left(\frac{1}{pT}\sum_{m=1}^p\sum_{t=1}^T \tilde{\lambda}_{tm}\widehat{U}_{jt,m}\right)^2 O_P(\Delta_F^2) + \frac{C}{n}\sum_{j=1}^n\left(\frac{1}{pT}\sum_{m=1}^p\sum_{t=1}^T\frac{1}{n}\widehat{F}'\tilde{U}_{t,m}\widehat{U}_{jt,m}\right)^2 \\
&\quad + \frac{C}{n}\sum_{j=1}^n\left(\frac{1}{pT}\sum_{m=1}^p\sum_{t=1}^T \tilde{\lambda}_{tm}\tilde{U}_{jt,m}\right)^2 O_P(\Delta_F^2) + \frac{C}{n}\sum_{j=1}^n\left(\frac{1}{pT}\sum_{m=1}^p\sum_{t=1}^T\frac{1}{n}\widehat{F}'\tilde{U}_{t,m}\tilde{U}_{jt,m}\right)^2
\end{aligned}
$$

$$
\begin{aligned}
&= O_P(\Delta_F^2) + \frac{C}{n}\sum_{j=1}^{n}\left(\frac{1}{pT}\sum_{m=1}^{p}\sum_{t=1}^{T}\frac{1}{n}\tilde{F}'\tilde{U}_{t,m}\tilde{U}_{jt,m}\right)^2 \\
&\quad + \frac{C}{n}\sum_{i=1}^{n}\left(\frac{1}{pT}\sum_{t=1}^{T}\sum_{m=1}^{p}\frac{1}{n}\tilde{F}'\tilde{U}_{t,m}(\widehat{U}_{it,m} - \tilde{U}_{it,m})\right)^2 \qquad \text{(I.14)} \\
&= O_P(\Delta_F^2).
\end{aligned}
$$

where $\frac{1}{n}\sum_{j=1}^{n}\left(\frac{1}{pT}\sum_{m=1}^{p}\sum_{t=1}^{T}\frac{1}{n}\tilde{F}'\tilde{U}_{t,m}\tilde{U}_{jt,m}\right)^2 = O_P\left(\frac{1}{n^2} + \frac{1}{nT^2} + \frac{1}{pTn}\right)$ follows from part (i). Hence, we have proved

$$
\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}E^*\left(\frac{1}{pT}\sum_{t=1}^{T}\tilde{U}_{it}^{*\prime}\tilde{U}_{jt}^{*} - E^*\tilde{U}_{it}^{*\prime}\tilde{U}_{jt}^{*}\right)^2 = O_{P^*}\left(\Delta_F^2 + \frac{1}{n}\right). \qquad \text{(I.15)}
$$

The Cauchy-Schwarz inequality then implies

$$
\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n}\|A_{i1}^*\|_2^2 &\le O_{P^*}(1)\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{1}{pT}\sum_{t=1}^{T}\tilde{U}_{it}^{*\prime}\tilde{U}_{jt}^{*} - E^*\tilde{U}_{it}^{*\prime}\tilde{U}_{jt}^{*}\right)^2 \\
&= O_{P^*}\left(\Delta_F^2 + \frac{1}{n}\right).
\end{aligned}
$$

■

Lemma I.12. *$\frac{1}{n}\sum_i\|A_{i3}^*\|_2^2 = O_{P^*}(\Delta_F^2)$ and $\frac{1}{n}\sum_i\|A_{i4}^*\|_2^2 = O_{P^*}(\Delta_F^2)$ for A_{i3}^* and A_{i4}^* defined in (I.8).*

Proof. By the Cauchy-Schwarz inequality, $\frac{1}{n}\sum_{i=1}^{n}\|A_{i3}^*\|_2^2 \le O_{P^*}(1)\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{pT}\sum_{t=1}^{T}\widehat{\Lambda}_t'\tilde{U}_{it}^*\right)^2$. Now

$$
\begin{aligned}
E^*\left[\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{pT}\sum_{t=1}^{T}\widehat{\Lambda}_t'\tilde{U}_{it}^*\right)^2\right] &= \frac{1}{n}\sum_{i=1}^{n}\mathrm{Var}^*\left(\frac{1}{pT}\sum_{t=1}^{T}\widehat{\Lambda}_t'\widehat{U}_{it}w_i^U\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{pT}\sum_{t=1}^{T}\widehat{\Lambda}_t'\widehat{U}_{it}\right)^2 \\
&\le O_P(1)\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{pT}\sum_{t=1}^{T}\tilde{\Lambda}_t'\widehat{U}_{it}\right)^2 \\
&\quad + \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{pT}\sum_{t=1}^{T}(\widehat{\Lambda}_t - \tilde{\Lambda}_t H'^{-1})'\widehat{U}_{it}\right)^2
\end{aligned}
$$

$$
\begin{aligned}
&\le O_P(1)\frac{1}{n}\sum_{i=1}^n\left(\frac{1}{pT}\sum_{t=1}^T\tilde\Lambda_t'\widehat U_{it}\right)^2+O_P(\Delta_F^2)\\
&\quad+\frac{1}{n}\sum_{i=1}^n\left(\frac{1}{pT}\sum_{t=1}^T\sum_{m=1}^p\frac{1}{n}\tilde F'\tilde U_{t,m}\widehat U_{it,m}\right)^2O_P(1)\\
&\le O_P(\Delta_F^2)+O_P(1)\frac{1}{n}\sum_{i=1}^n\left(\frac{1}{pT}\sum_{t=1}^T\tilde\Lambda_t'\tilde U_{it}\right)^2\\
&\quad+O_P(1)\frac{1}{n}\sum_{i=1}^n\left(\frac{1}{pT}\sum_{t=1}^T\sum_{m=1}^p\tilde\lambda_{tm}(\widehat U_{it,m}-\tilde U_{it,m})\right)^2\\
&\le O_P(\Delta_F^2)+O_P(1)\left(\frac{1}{pT}\sum_{t=1}^T\sum_{m=1}^p\frac{1}{n}\tilde F'\tilde U_{t,m}\tilde\lambda_{tm}\right)^2\\
&=O_P\left(\Delta_F^2+\frac{1}{nTp}\right)=O_P(\Delta_F^2). \qquad\text{(I.16)}
\end{aligned}
$$

where we used Lemma I.11 (ii) to bound $\frac{1}{n}\sum_{i=1}^n(\frac{1}{pT}\sum_{t=1}^T\sum_{m=1}^p\frac{1}{n}\tilde F'\tilde U_{t,m}\widehat U_{it,m})^2=O_P(\Delta_F^2)$ and the equalities $\widehat\lambda_{tm}-H^{-1}\tilde\lambda_{tm}=(\widehat F'\widehat F)^{-1}\widehat F'(\tilde FH-\widehat F)H^{-1}\tilde\lambda_{tm}+(\widehat F'\widehat F)^{-1}\widehat F'\tilde U_{t,m}$,
$\tilde U_{it,m}-\widehat U_{it,m}=\widehat f_i'(\widehat F'\widehat F)^{-1}\widehat F'(\tilde FH-\widehat F)H^{-1}\tilde\lambda_{tm}+\widehat f_i'(\widehat F'\widehat F)^{-1}\widehat F'\tilde U_{t,m}+\tilde\lambda_{tm}'H'^{-1}(\widehat f_i-H'\tilde f_i)$.

Hence, $\frac{1}{n}\sum_i\|A_{i3}^*\|_2^2=O_{P^*}(\Delta_F^2)$. $\frac{1}{n}\sum_i\|A_{i4}^*\|_2^2$ is bounded similarly. ■

Lemma I.13. *(i)* $\frac{1}{n}\sum_i\|\frac{1}{pTn}\sum_{j=1}^n\widehat f_j\sum_{t=1}^T(\tilde U_{it}^{*\prime}\tilde U_{jt}^*-E^*\tilde U_{it}^{*\prime}\tilde U_{jt}^*)\|_2^2=O_{P^*}(\Delta_F^2)$.
(ii) $\frac{1}{T}\sum_s\left\|\frac{1}{pTn}\sum_{i=1}^n\sum_t\widehat\Lambda_t'\tilde U_{it}^*\tilde z_{is}^*\right\|_2^2=O_{P^*}\left(\frac{\log n}{npT}+\frac{\log n}{n^2T^2}+\frac{1}{n^2}(1+\sqrt{\frac{\log n}{T}})\right)$ *where* $\tilde z_{is}^*\in\{\tilde\eta_{is}^*,\tilde\epsilon_{is}^*\}$.
(iii) $\left\|\frac{1}{nTp}\sum_{j=1}^n\sum_{t=1}^T\widehat f_j^*\widehat\Lambda_t'\tilde U_{jt}^*\right\|_2=O_{P^*}\left(\Delta_F^2+\sqrt{\frac{\Delta_F^2\log n}{n}}\right)$.
(iv)

$$
\max_{m,k,l,t}\left\|\frac{1}{nT}\sum_{i=1}^n\sum_{s=1}^T(\tilde U_{it,l}^*\tilde U_{is,m}^*-E^*\tilde U_{it,l}^*\tilde U_{is,m}^*)\widehat h_{sk}'\right\|_F
=O_{P^*}\left(\sqrt{\frac{\log(pT)}{n}\left(\frac{\log n\log(pT)}{n}+\frac{\log(np)}{T}\right)}\right)
$$

(v)

$$\max_{m,k\le p}\frac{1}{pT}\sum_{l=1}^{p}\sum_{t=1}^{T}\left\|\frac{1}{nT}\sum_{is}E^*(\tilde{U}^*_{it,l}\tilde{U}^*_{is,m})\widehat{h}'_{sk}\right\|_F = O_P\left(\Delta_F+\frac{\sqrt{\log(pT)\log p}}{n}+\sqrt{\frac{\log(pT)}{nT}}\right),$$

(vi)

$$\max_{m,k\le p}\left\|\frac{1}{nT}\sum_{i=1}^{n}\frac{1}{pT}\sum_{s,t}\widehat{\Lambda}'_t(\tilde{U}^*_{it}\tilde{U}^*_{is,m}-E^*\tilde{U}^*_{it}\tilde{U}^*_{is,m})\widehat{h}'_{sk}\right\|_F = O_{P^*}\left(\sqrt{\frac{\log p}{n}\left(\frac{\log n\log(pT)}{n}+\frac{\log(np)}{T}\right)\Delta_F^2}\right),$$

where $\widehat{h}_{sk}\in\{\widehat{\delta}_{ys},\widehat{\delta}_{ds},\widehat{\lambda}_{sk}\}$, $s\le T, k\le p$.

Proof. We shall assume f_j to be a scalar without loss of generality.

(i)

$$\begin{aligned}
&\frac{1}{n}\sum_i E^*\left|\frac{1}{pTn}\sum_{j=1}^{n}\widehat{f}_j\sum_{t=1}^{T}(\tilde{U}^{*\prime}_{it}\tilde{U}^*_{jt}-E^*\tilde{U}^{*\prime}_{it}\tilde{U}^*_{jt})\right|^2\\
&=\frac{1}{n}\sum_i \mathrm{Var}^*\left(\frac{1}{pTn}\sum_{j=1}^{n}\widehat{f}_j\sum_{t=1}^{T}\tilde{U}^{*\prime}_{it}\tilde{U}^*_{jt}\right)\\
&=\frac{1}{n}\sum_i\left(\frac{1}{pTn}\sum_{j=1}^{n}\widehat{f}_j\sum_{t=1}^{T}\widehat{U}'_{it}\widehat{U}_{jt})^2\,\mathrm{Var}^*(w_i^Uw_j^U\right)\\
&\le\frac{C}{n}\sum_i\left(\frac{1}{pTn}\sum_{j=1}^{n}\widehat{f}_j\sum_{t=1}^{T}(\widehat{U}'_{it}\widehat{U}_{jt}-\tilde{U}'_{it}\tilde{U}_{jt})\right)^2\\
&\quad+\frac{C}{n}\sum_i\left(\frac{1}{pTn}\sum_{j=1}^{n}\widehat{f}_j\sum_{t=1}^{T}(\tilde{U}'_{it}\tilde{U}_{jt}-EU'_{it}U_{jt})\right)^2+\frac{C}{n}\sum_i\left(\frac{1}{pTn}\sum_{j=1}^{n}\widehat{f}_j\sum_{t=1}^{T}EU'_{it}U_{jt}\right)^2\\
&\le\underbrace{\frac{C}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{1}{pT}\sum_{t=1}^{T}(\widehat{U}'_{it}\widehat{U}_{jt}-\tilde{U}'_{it}\tilde{U}_{jt})\right)^2}_{a_3}+\underbrace{\frac{C}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{1}{pT}\sum_{t=1}^{T}(\tilde{U}'_{it}\tilde{U}_{jt}-EU'_{it}U_{jt})\right)^2}_{a_1}
\end{aligned}$$

$$
\begin{aligned}
&+ \frac{C}{n}\sum_i \left(\frac{1}{pTn}\widehat{f}_j\sum_{t=1}^T EU_{it}^{'}U_{it}\right)^2 \\
&= a_1 + a_3 + O_P\left(\frac{1}{n^2}\right) = O_P(\Delta_F^2)
\end{aligned}
$$

where a_1 and a_3 are as defined in Lemma I.11. We also have $a_1 = O_P(\Delta_F^2)$ and $a_3 = O_P(\Delta_F^2)$ by (I.13) and (I.14).

(ii) We prove the result for $\tilde{z}_{is}^* = \tilde{\eta}_{is}^*$ only. The other results follow similarly. Let $a_n := \frac{\log n}{npT} + \frac{\Delta_F^2 \log n}{n} + \frac{\Delta_{\max}^2}{n}$.

$$
\begin{aligned}
&E^*\left[\frac{1}{T}\sum_{s=1}^T\left|\frac{1}{pTn}\sum_{i=1}^n\sum_{t=1}^T \widehat{\Lambda}_t'\tilde{U}_{it}^*\tilde{z}_{is}^*\right|^2\right] \\
&= \frac{1}{T}\sum_{s=1}^T \mathrm{Var}^*\left(\frac{1}{pTn}\sum_{i=1}^n\sum_{t=1}^T \widehat{\Lambda}_t'\widehat{U}_{it}\widehat{\eta}_{is}w_i^U w_i^Z\right) \\
&= \frac{1}{T}\sum_{s=1}^T \frac{1}{n^2}\sum_{i=1}^n \left(\frac{1}{pT}\sum_{t=1}^T \widehat{\Lambda}_t'\widehat{U}_{it}\widehat{\eta}_{is}\right)^2 \mathrm{Var}^*(w_i^U w_i^Z) \\
&\le \frac{1}{T}\sum_{s=1}^T \frac{1}{n^2}\sum_{i=1}^n \widehat{\eta}_{is}^2 \left(\frac{1}{pT}\sum_{t=1}^T \widehat{\Lambda}_t'\widehat{U}_{it}\right)^2 C \\
&\le O_P\left(\frac{1}{n}\right)\max_i \left(\frac{1}{pT}\sum_{m=1}^p\sum_{t=1}^T \widehat{\lambda}_{tm}(\widehat{U}_{it,m} - \tilde{U}_{it,m})\right)^2 \\
&\quad + O_P\left(\frac{1}{n}\right)\max_i \left(\frac{1}{pT}\sum_{m=1}^p\sum_{t=1}^T (\widehat{\lambda}_{tm} - H^{'-1}\tilde{\lambda}_{tm})\tilde{U}_{it,m}\right)^2 \\
&\quad + O_P\left(\frac{1}{n}\right)\max_i \left(\frac{1}{pT}\sum_{m=1}^p\sum_{t=1}^T \tilde{\lambda}_{tm}\tilde{U}_{it,m}\right)^2 \\
&\le O_P\left(\frac{\log n}{npT} + \frac{\Delta_F^2\log n}{n} + \frac{\Delta_{\max}^2}{n}\right) + O_P\left(\frac{\log n}{n}\right)\left(\frac{1}{pT}\sum_{m=1}^p\sum_{t=1}^T \widehat{\lambda}_{tm}\frac{1}{n}\widehat{F}'\tilde{U}_{t,m}\right)^2 \\
&\quad + O_P\left(\frac{1}{n}\right)\max_i \left(\frac{1}{pT}\sum_{m=1}^p\sum_{t=1}^T (\widehat{\lambda}_{tm} - H^{'-1}\tilde{\lambda}_{tm})\tilde{U}_{it,m}\right)^2 \\
&\le O_P(a_n) + O_P\left(\frac{\log n}{n}\right)\left(\frac{1}{pT}\sum_{m=1}^p\sum_{t=1}^T \widehat{\lambda}_{tm}\frac{1}{n}\widehat{F}'\tilde{U}_{t,m}\right)^2
\end{aligned}
$$

$$
\begin{aligned}
&+ O_P\left(\frac{1}{n}\right)\max_i\left(\frac{1}{pT}\sum_{m=1}^{p}\sum_{t=1}^{T}\frac{1}{n}\widehat{F}'\tilde{U}_{t,m}\tilde{U}_{it,m}\right)^2 + O_P\left(\frac{\Delta_F^2}{n}\right)\max_i\left(\frac{1}{pT}\sum_{m=1}^{p}\sum_{t=1}^{T}\tilde{\lambda}_{tm}\tilde{U}_{it,m}\right)^2 \\
&\leq O_P(a_n) + O_P\left(\frac{1}{n}\right)\max_i\left(\frac{1}{pT}\sum_{m=1}^{p}\sum_{t=1}^{T}\frac{1}{n}\tilde{F}'\tilde{U}_{t,m}\tilde{U}_{it,m}\right)^2 \\
&+ O_P\left(\frac{\log n}{n}\right)\left(\frac{1}{pT}\sum_{m=1}^{p}\sum_{t=1}^{T}\widehat{\lambda}_{tm}\frac{1}{n}\tilde{F}'\tilde{U}_{t,m}\right)^2 \\
&\leq O_P(a_n) + O_P\left(\frac{1}{n^2}\right)\max_i\frac{1}{pT}\sum_{m=1}^{p}\sum_{t=1}^{T}\tilde{U}_{it,m}^2 + O_P\left(\frac{\log n}{n}\right)\left(\frac{1}{pT}\sum_{m=1}^{p}\sum_{t=1}^{T}\tilde{\lambda}_{tm}\frac{1}{n}\tilde{F}'\tilde{U}_{t,m}\right)^2 \\
&+ O_P\left(\frac{\log n}{n}\right)\left(\frac{1}{pT}\sum_{m=1}^{p}\sum_{t=1}^{T}(\widehat{\lambda}_{tm} - H'^{-1}\tilde{\lambda}_{tm})\frac{1}{n}\tilde{F}'\tilde{U}_{t,m}\right)^2 \\
&\leq O_P\left(a_n + \frac{1}{n^2}\left(1+\sqrt{\frac{\log n}{T}}\right)\right) + O_P\left(\frac{\log n}{n}\right)\left(\frac{1}{pT}\sum_{m=1}^{p}\sum_{t=1}^{T}(\frac{1}{n}\tilde{F}'\tilde{U}_{t,m})^2\right)^2 \\
&= O_P\left(\frac{\log n}{npT} + \frac{\Delta_F^2\log n}{n} + \frac{\Delta_{\max}^2}{n} + \frac{1}{n^2} + \frac{1}{n^2}\sqrt{\frac{\log n}{T}}\right) \\
&= O_P\left(\frac{\log n}{npT} + \frac{\log n}{n^2T^2} + \frac{1}{n^2} + \frac{1}{n^2}\sqrt{\frac{\log n}{T}}\right)
\end{aligned}
$$

where we use $\max_i \|\widehat{f}_i - H'\tilde{f}_i\|_2 = O_P(\Delta_{\max}) = O_P\left(\frac{1}{\sqrt{n}} + \sqrt{\frac{\log n}{Tp}}\right)$ from Lemma I.1.

(iii) By (I.16), $\frac{1}{n}\sum_{i=1}^{n}(\frac{1}{pT}\sum_{t=1}^{T}\widehat{\Lambda}_t'\tilde{U}_{it}^*)^2 = O_{P^*}(\Delta_F^2)$ and $\frac{1}{n}\sum_{i=1}^{n}(\frac{1}{pT}\sum_{t=1}^{T}\widehat{\Lambda}_t'\widehat{U}_{it})^2 = O_P(\Delta_F^2)$. Therefore,

$$
\begin{aligned}
E^*\left|\frac{1}{nTp}\sum_{j=1}^{n}\sum_{t=1}^{T}\widehat{f}_j\widehat{\Lambda}_t'\tilde{U}_{jt}^*\right|^2 &= \frac{1}{n^2}\sum_{j=1}^{n}\left(\frac{1}{Tp}\sum_{t=1}^{T}\widehat{f}_j\widehat{\Lambda}_t'\widehat{U}_{jt}\right)^2 \mathrm{Var}^*(w_j^U) \\
&= \frac{1}{n^2}\sum_{j=1}^{n}\left(\frac{1}{Tp}\sum_{t=1}^{T}\widehat{\Lambda}_t'\widehat{U}_{jt}\right)^2 \widehat{f}_j^2 \\
&\leq \frac{1}{n^2}\sum_{j=1}^{n}\left(\frac{1}{Tp}\sum_{t=1}^{T}\widehat{\Lambda}_t'\widehat{U}_{jt}\right)^2 \max_j \widehat{f}_j^2 \\
&= O_P\left(\frac{\Delta_F^2\log n}{n}\right).
\end{aligned}
$$

Hence,

$$\left\|\frac{1}{nTp}\sum_{j=1}^{n}\sum_{t=1}^{T}\widehat{f}_j^*\widehat{\Lambda}_t'\tilde{U}_{jt}^*\right\|_2^2 \leq O_{P^*}(\Delta_F^2)\frac{1}{n}\sum_{j=1}^{n}\left(\frac{1}{pT}\sum_{t=1}^{T}\widehat{\Lambda}_t'\tilde{U}_{jt}^*\right)^2 + \left|\frac{1}{nTp}\sum_{j=1}^{n}\sum_{t=1}^{T}\widehat{f}_j\widehat{\Lambda}_t'\tilde{U}_{jt}^*\right|^2 O_{P^*}(1)$$
$$\leq O_{P^*}\left(\Delta_F^4 + \frac{\Delta_F^2\log n}{n}\right).$$

(iv) For notational simplicity, we assume $\widehat{h}_{sk}$ to be a scalar. Then

$$\max_{mklt}|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}(\tilde{U}_{it,l}^*\tilde{U}_{is,m}^* - E^*\tilde{U}_{it,l}^*\tilde{U}_{is,m}^*)\widehat{h}_{sk}| = \max_{mklt}|\frac{1}{n}\sum_{i=1}^{n}\Gamma_{i,mklt}|$$

where $\Gamma_{i,mklt} = \frac{1}{T}\sum_{s=1}^{T}\widehat{U}_{it,l}\widehat{U}_{is,m}\widehat{h}_{sk}\left[(w_i^U)^2 - E^*(w_i^U)^2\right]$. We still use the Bernstein inequality for independent data to achieve

$$\max_{mklt}|\frac{1}{n}\sum_{i=1}^{n}\Gamma_{i,mklt}| = O_{P^*}(\sqrt{\max_{mklt}\frac{1}{n}\sum_{i=1}^{n}\mathrm{Var}^*(\Gamma_{i,mklt})\frac{\log(pT)}{n}}).$$

So we now bound $\max_{mklt}\frac{1}{n}\sum_{i=1}^{n}\mathrm{Var}^*(\Gamma_{i,mklt})$ below. Recall the equality:

$\tilde{U}_{it,m} - \widehat{U}_{it,m} = \widehat{f}_i'(\widehat{F}'\widehat{F})^{-1}\widehat{F}'(\tilde{F}H - \widehat{F})H^{-1}\tilde{\lambda}_{tm} + \widehat{f}_i'(\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{U}_{t,m} + \tilde{\lambda}_{tm}'H'^{-1}(\widehat{f}_i - H'\tilde{f}_i)$,

$$\max_{mklt}\frac{1}{n}\sum_{i=1}^{n}\mathrm{Var}^*(\Gamma_{i,mklt}) = \max_{mklt}\frac{1}{n}\sum_{i=1}^{n}\widehat{U}_{it,l}^2\left(\frac{1}{T}\sum_{s=1}^{T}\widehat{U}_{is,m}\widehat{h}_{sk}\right)^2\mathrm{Var}^*\left((w_i^U)^2\right)$$
$$\leq O_P(1)\max_{imk}\left(\frac{1}{T}\sum_{s=1}^{T}\widehat{U}_{is,m}\widehat{h}_{sk}\right)^2$$
$$\leq O_P(1)\max_{imk}\left|\frac{1}{T}\sum_{s=1}^{T}\tilde{U}_{is,m}\widehat{h}_{sk}\right|^2$$
$$+\max_{mk}\frac{1}{T}\sum_{t=1}^{T}\left|\tilde{\lambda}_{tm}\widehat{h}_{tk}\right|^2 O_P\left(\Delta_F^2\log n + \Delta_{\max}^2\right)$$
$$+O_P(\log n)\max_{m}\frac{1}{T}\sum_{t=1}^{T}\left|\frac{1}{n}\widehat{F}'\tilde{U}_{t,m}\right|^2$$

$$\leq O_P\left(\Delta_F^2 \log n + \Delta_{\max}^2 + \frac{\log n \log(pT)}{n} + \frac{\log(np)}{T}\right) + \max_{imk}\left|\frac{1}{T}\sum_{t=1}^T \tilde{U}_{it,m}(\widehat{h}_{tk} - H^{-1}\tilde{h}_{tk})\right|^2. \tag{I.17}$$

In the derivation above, we used

$$\max_{m\leq p}\frac{1}{n^2T}\sum_t \|\tilde{U}_{t,m}'\widehat{F}\|_2^2 = O_P\left(\frac{\log(pT)}{n} + \Delta_{fum}^2\right) = O_P\left(\frac{\log(pT)}{n} + \Delta_F^2\right)$$

from Lemma H.11 and $\max_{imk}\frac{1}{T}\sum_{t=1}^T|\tilde{\lambda}_{tm}\widehat{h}_{tk}| = O_P(1)$ which follows from $\max_{tm}|\widehat{\lambda}_{tm}| = O_P(1)$ and $\max_t|\widehat{h}_{tk}| = O_P(1)$ for $\widehat{h}_{tk}\in\{\widehat{\lambda}_{tm}, \widehat{\delta}_{yt}\}$. We now bound the second term on the right in (I.17) in two cases.

Case 1: $\widehat{h}_{tk} = \widehat{\delta}_{yt}$ for all $k\leq p$. Using the equality

$$\widehat{\delta}_{yt} - H^{-1}\tilde{\delta}_{yt} = (\widehat{F}'\widehat{F})^{-1}\widehat{F}'(\tilde{F}H - \widehat{F})H^{-1}\tilde{\delta}_{yt} + (\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{U}_t\gamma_y + (\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{e}_t,$$

we have

$$\begin{aligned}\max_{imk}\left|\frac{1}{T}\sum_{t=1}^T \tilde{U}_{it,m}(\widehat{h}_{tk} - H^{-1}\tilde{h}_{tk})\right|^2 &\leq \max_{im}\left|\frac{1}{T}\sum_{t=1}^T \tilde{U}_{it,m}\tilde{\delta}_{yt}\right|^2 O_P(\Delta_F^2) + \max_{im}\left|\frac{1}{T}\sum_{t=1}^T \tilde{U}_{it,m}\frac{1}{n}\widehat{F}'\tilde{U}_t\gamma_y\right|^2 \\ &\quad + \max_{im}\left|\frac{1}{T}\sum_{t=1}^T \tilde{U}_{it,m}\frac{1}{n}\widehat{F}'\tilde{e}_t\right|^2 \\ &\leq \frac{\log(np)}{T}O_P(\Delta_F^2) + \max_{im}\frac{1}{T}\sum_t \tilde{U}_{it,m}^2 O_P\left(\frac{|J|_0}{n} + |J|_0^2\Delta_F^2\right) \\ &= \frac{\log(np)}{T}O_P(\Delta_F^2) + \left(1 + \sqrt{\frac{\log(np)}{T}}\right)O_P\left(\frac{|J|_0}{n} + |J|_0^2\Delta_F^2\right),\end{aligned}$$

where we used Lemma H.11 (iv) that $\frac{1}{T}\sum_t\|\frac{1}{n}\widehat{F}'\tilde{e}_t\|_2^2 + \frac{1}{T}\sum_t\|\frac{1}{n}\widehat{F}'\tilde{U}_t\gamma\|_2^2 = O_P\left(\frac{|J|_0}{n} + |J|_0^2\Delta_F^2\right)$.

Case 2: $\widehat{h}_{tk} = \widehat{\lambda}_{tk}$. Using the equality

$$\widehat{\lambda}_{tk} - H^{-1}\tilde{\lambda}_{tk} = (\widehat{F}'\widehat{F})^{-1}\widehat{F}'(\tilde{F}H - \widehat{F})H^{-1}\tilde{\lambda}_{tk} + (\widehat{F}'\widehat{F})^{-1}\widehat{F}'\tilde{U}_{t,k},$$

we have

$$\max_{imk}\left|\frac{1}{T}\sum_{t=1}^T \tilde{U}_{it,m}(\widehat{h}_{tk} - H^{-1}\tilde{h}_{tk})\right|^2 \leq \max_{imk}\left|\frac{1}{T}\sum_{t=1}^T \tilde{U}_{it,m}\tilde{\lambda}_{tk}\right|^2 O_P(\Delta_F^2)$$

$$
\begin{aligned}
&+\max_{imk}\left|\frac{1}{T}\sum_{t=1}^{T}\tilde{U}_{it,m}\frac{1}{n}\widehat{F}'\tilde{U}_{t,k}\right|^2\\
&\leq O_P\left(\frac{\Delta_F^2\log(np)}{T}\right)+\max_{imk}\frac{1}{T}\sum_{t=1}^{T}\tilde{U}_{it,m}^2\frac{1}{T}\sum_t\left(\frac{1}{n}\widehat{F}'\tilde{U}_{t,k}\right)^2\\
&=O_P\left(\frac{\Delta_F^2\log(np)}{T}\right)+\left(1+\sqrt{\frac{\log(pn)}{T}}\right)\left(\Delta_F^2+\frac{\log(pT)}{n}\right).
\end{aligned}
$$

Hence $\max_{imk}|\frac{1}{T}\sum_{t=1}^{T}\tilde{U}_{it,m}(\widehat{h}_{tk}-H^{-1}\tilde{h}_{tk})|^2=O_P(a_n)$, where

$$
a_n:=\frac{\log(np)}{T}(\Delta_F^2)+\left(1+\sqrt{\frac{\log(np)}{T}}\right)\left(\frac{|J|_0}{n}+|J|_0^2\Delta_F^2+\frac{\log(pT)}{n}\right).
$$

Substituting this expression into (I.17), we obtain

$$
\begin{aligned}
\max_{mklt}\frac{1}{n}\sum_{i=1}^{n}\mathrm{Var}^*(\Gamma_{i,mklt})&=O_P(1)\max_{imk}\left(\frac{1}{T}\sum_{s=1}^{T}\widehat{U}_{is,m}\widehat{h}_{sk}\right)^2\\
&=O_P\left(\frac{\log n\log(pT)}{n}+\frac{\log(np)}{T}\right).
\end{aligned}
\tag{I.18}
$$

Hence, $\max_{mklt}\left|\frac{1}{n}\sum_{i=1}^{n}\Gamma_{i,mklt}\right|=O_{P^*}\left(\sqrt{\frac{\log(pT)}{n}\left(\frac{\log n\log(pT)}{n}+\frac{\log(np)}{T}\right)}\right)$.

(v) We assume $\widehat{h}_{sk}$ to be a scalar for notational simplicity.

$$
\begin{aligned}
&\max_{m,k\leq p}\frac{1}{pT}\sum_{l=1}^{p}\sum_{t=1}^{T}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}E^*(\tilde{U}_{it,l}^*\tilde{U}_{is,m}^*)\widehat{h}_{sk}\right|\\
&\quad=\max_{m,k\leq p}\frac{1}{pT}\sum_{l=1}^{p}\sum_{t=1}^{T}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}\widehat{U}_{it,l}\widehat{U}_{is,m}\widehat{h}_{sk}\right|\\
&\quad\leq\max_{m,k\leq p}\frac{1}{pT}\sum_{l=1}^{p}\sum_{t=1}^{T}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}(\tilde{U}_{it,l}\tilde{U}_{is,m})\widehat{h}_{sk}\right|\\
&\qquad+\max_{m,k\leq p}\frac{1}{pT}\sum_{l=1}^{p}\sum_{t=1}^{T}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}\tilde{U}_{is,m}(\widehat{U}_{it,l}-\tilde{U}_{it,l})\widehat{h}_{sk}\right|\\
&\qquad+\max_{m,k\leq p}\frac{1}{pT}\sum_{l=1}^{p}\sum_{t=1}^{T}\left|\frac{1}{n}\sum_{i=1}^{n}\widehat{U}_{it,l}\widehat{f}_i'\right|O_P(\Delta_F)
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{1}{pT}\sum_{l=1}^{p}\sum_{t=1}^{T}\left|\frac{1}{n}\sum_{i=1}^{n}\widehat{U}_{it,l}\widehat{f}_i'\right|\max_{m,k\leq p}\left|\frac{1}{Tn}\sum_{s=1}^{T}\tilde{F}'\tilde{U}_{s,m}\widehat{h}_{sk}\right| \\
&+\frac{1}{pT}\sum_{l=1}^{p}\sum_{t=1}^{T}\left|\frac{1}{n}\sum_{i=1}^{n}\tilde{U}_{it,l}(\widehat{f}_i-H'\tilde{f}_i)\right|O_P(1) \\
\leq\ & O_P(\Delta_F)+\max_{m,k\leq p}\frac{1}{pT}\sum_{l=1}^{p}\sum_{t=1}^{T}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}(\tilde{U}_{it,l}\tilde{U}_{is,m})\widehat{h}_{sk}\right| \\
&+\frac{1}{pT}\sum_{l=1}^{p}\sum_{t=1}^{T}\left|\frac{1}{n}\sum_{i=1}^{n}\widehat{U}_{it,l}\widehat{f}_i'\right|\max_{m,k\leq p}\left|\frac{1}{Tn}\sum_{s=1}^{T}\tilde{F}'\tilde{U}_{s,m}\widehat{h}_{sk}\right| \\
\leq\ & O_P(\Delta_F)+\max_{m,k\leq p}\frac{1}{pT}\sum_{l=1}^{p}\sum_{t=1}^{T}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}(\tilde{U}_{it,l}\tilde{U}_{is,m})\widehat{h}_{sk}\right| \\
&+\max_{m,k\leq p}\left|\frac{1}{Tn}\sum_{s=1}^{T}\tilde{F}'\tilde{U}_{s,m}\widehat{h}_{sk}\right|O_P\left(\frac{1}{\sqrt{n}}\right) \\
\leq\ & O_P\left(\Delta_F+\frac{\sqrt{\log p}}{n}\right)+\max_{m,k\leq p}\frac{1}{pT}\sum_{l=1}^{p}\sum_{t=1}^{T}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}\tilde{U}_{it,l}\tilde{U}_{is,m}\widehat{h}_{sk}\right| \\
\leq\ & O_P\left(\Delta_F+\frac{\sqrt{\log p}}{n}\right)+O_P(1)\max_{m\leq p}\frac{1}{p}\sum_{l=1}^{p}\left|\frac{1}{n}\sum_{i=1}^{n}\bar{U}_{i\cdot,l}\bar{U}_{i\cdot,m}\right| \\
&+O_P(1)\max_{m\leq p}\frac{1}{pT}\sum_{l=1}^{p}\sum_{t=1}^{T}\left|\frac{1}{n}\sum_{i=1}^{n}U_{it,l}\bar{U}_{i\cdot,m}\right|+\max_{m,k\leq p}\frac{1}{pT}\sum_{l=1}^{p}\sum_{t=1}^{T}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}U_{it,l}U_{is,m}\widehat{h}_{sk}\right| \\
&+\max_{m,k\leq p}\frac{1}{p}\sum_{l=1}^{p}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}\bar{U}_{i\cdot,l}U_{is,m}\widehat{h}_{sk}\right| \\
:=\ & O_P\left(\Delta_F+\frac{\sqrt{\log p}}{n}\right)+\sum_{l=1}^{4}D_lO_P(1).
\end{aligned}
$$

We now bound each of D_1-D_4.

$$
\begin{aligned}
D_1&=\max_{m\leq p}\frac{1}{p}\sum_{l=1}^{p}\left|\frac{1}{n}\sum_{i=1}^{n}\bar{U}_{i\cdot,l}\bar{U}_{i\cdot,m}\right| \\
&\leq\frac{1}{pT}\max_{ms}\sum_{t}\sum_{l=1}^{p}|EU_{is,l}U_{it,m}|
\end{aligned}
$$

$$+\max_{m,l\le p}\left|\frac{1}{n}\sum_{i=1}^{n}\frac{1}{T^2}\sum_{s,t}(U_{is,l}U_{it,m}-EU_{is,l}U_{it,m})\right|$$

$$=O\left(\frac{1}{pT}\right)+\max_{m,l\le p}\left|\frac{1}{n}\sum_{i=1}^{n}X_{i,ml}\right|$$

where $X_{i,ml}=\frac{1}{T^2}\sum_{s,t}(U_{is,l}U_{it,m}-EU_{is,l}U_{it,m})$ with $\max_{ml}\frac{1}{n}\sum_i \mathrm{Var}(X_{i,ml})=O(\frac{1}{T^2})$. Thus, $D_1=O_P\left(\frac{1}{T}\sqrt{\frac{\log p}{n}}+\frac{1}{pT}\right)$.

$$D_2=\max_{m\le p}\frac{1}{pT}\sum_{l=1}^{p}\sum_{t=1}^{T}\left|\frac{1}{n}\sum_{i=1}^{n}U_{it,l}\bar{U}_{i\cdot,m}\right|$$

$$\le\max_{m\le p}\frac{1}{pT}\sum_{l=1}^{p}\sum_{t=1}^{T}\left|\frac{1}{n}\sum_{i=1}^{n}\frac{1}{T}\sum_{s}EU_{it,l}U_{is,m}\right|$$

$$+\max_{mlt}\left|\frac{1}{n}\sum_{i=1}^{n}\frac{1}{T}\sum_{s}(U_{it,l}U_{is,m}-EU_{it,l}U_{is,m})\right|$$

$$=O\left(\frac{1}{pT}\right)+\max_{mlt}\left|\frac{1}{n}\sum_{i=1}^{n}X_{i,mtl}\right|$$

where $X_{i,mtl}=\frac{1}{T}\sum_{s}(U_{it,l}U_{is,m}-EU_{it,l}U_{is,m})$ with $\max_{mtl}\frac{1}{n}\sum_i \mathrm{Var}(X_{i,mtl})=O\left(\frac{1}{T}\right)$. Hence, $D_2=O_P\left(\sqrt{\frac{\log(pT)}{nT}}+\frac{1}{pT}\right)$.

$$D_3+D_4=\max_{m,k\le p}\frac{1}{pT}\sum_{l=1}^{p}\sum_{t=1}^{T}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}U_{it,l}U_{is,m}\widehat{h}_{sk}\right|$$

$$+\max_{m,k\le p}\frac{1}{p}\sum_{l=1}^{p}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}\bar{U}_{i\cdot,l}U_{is,m}\widehat{h}_{sk}\right|$$

$$\le O\left(\frac{1}{pT}+\sqrt{\frac{\log(pT)}{nT}}\right)+\max_{m,k\le p}\frac{1}{pT}\sum_{l=1}^{p}\sum_{t=1}^{T}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}U_{it,l}U_{is,m}(\widehat{h}_{sk}-H^{-1}\tilde{h}_{sk})\right|$$

$$+\max_{m,k\le p}\frac{1}{p}\sum_{l=1}^{p}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}\bar{U}_{i\cdot,l}U_{is,m}(\widehat{h}_{sk}-H^{-1}\tilde{h}_{sk})\right|.$$

To bound the last two terms on the right-hand-side of the last inequality, we discuss two cases.

Case 1: $\widehat{h}_{tk} = \widehat{\delta}_{yt}$ for all $k \le p$. We have

$$\max_{m,k\le p} \frac{1}{pT}\sum_{l=1}^{p}\sum_{t=1}^{T}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T} U_{it,l}U_{is,m}(\widehat{h}_{sk} - H^{-1}\tilde{h}_{sk})\right|$$
$$+\frac{1}{p}\sum_{l=1}^{p}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}\bar{U}_{i\cdot,l}U_{is,m}(\widehat{h}_{sk} - H^{-1}\tilde{h}_{sk})\right| \le O_P\left(\Delta_F + \frac{\sqrt{|J|_0\log(pT)}}{n}\right).$$

Case 2: $\widehat{h}_{tk} = \widehat{\lambda}_{tk}$. We have

$$\max_{m,k\le p} \frac{1}{pT}\sum_{l=1}^{p}\sum_{t=1}^{T}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T} U_{it,l}U_{is,m}(\widehat{h}_{sk} - H^{-1}\tilde{h}_{sk})\right|$$
$$+\frac{1}{p}\sum_{l=1}^{p}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}\bar{U}_{i\cdot,l}U_{is,m}(\widehat{h}_{sk} - H^{-1}\tilde{h}_{sk})\right| \le O_P\left(\Delta_F + \frac{\sqrt{\log(pT)\log p}}{n}\right).$$

Combining all of the above bounds yields

$$\max_{m,k\le p}\frac{1}{pT}\sum_{l=1}^{p}\sum_{t=1}^{T}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}E^*(\tilde{U}^*_{it,l}\tilde{U}^*_{is,m})\widehat{h}_{sk}\right| = O_P\left(\Delta_F + \frac{\sqrt{\log(pT)\log p}}{n} + \sqrt{\frac{\log(pT)}{nT}}\right).$$

(vi) We assume $\widehat{\lambda}'_t$ is a row vector and $\widehat{h}_{sk}$ is a scalar. Now let

$$\max_{m,k\le p}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}\frac{1}{pT}\sum_{t=1}^{T}\widehat{\Lambda}'_t(\tilde{U}^*_{it}\tilde{U}^*_{is,m} - E^*\tilde{U}^*_{it}\tilde{U}^*_{is,m})\widehat{h}'_{sk}\right| = \max_{m,k\le p}\left|\frac{1}{n}\sum_i X_{i,mk}\right|$$

where $X_{i,mk} = \frac{1}{T}\sum_{s=1}^{T}\frac{1}{pT}\sum_{t=1}^{T}\widehat{\Lambda}'_t\widehat{U}_{it}\widehat{U}_{is,m}\widehat{h}_{sk}(w_i^{U2} - E^*w_i^{U2})$. Note that $X_{i,mk}$ is mean zero and satisfies

$$\max_{mk}\frac{1}{n}\sum_{i=1}^{n}\mathrm{Var}^*(X_{i,mk}) = \max_{mk}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{pT}\sum_{t=1}^{T}\widehat{\Lambda}'_t\widehat{U}_{it}\right)^2\left(\frac{1}{T}\sum_{s=1}^{T}\widehat{U}_{is,m}\widehat{h}_{sk}\right)^2\mathrm{Var}^*(w_i^{U2})$$
$$\le^{(I.18)}\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{pT}\sum_{t=1}^{T}\widehat{\Lambda}'_t\widehat{U}_{it}\right)^2 O_P\left(\frac{\log n\log(pT)}{n} + \frac{\log(np)}{T}\right)$$
$$=^{(I.16)} O_P\left(\frac{\log n\log(pT)}{n} + \frac{\log(np)}{T}\right)\Delta_F^2.$$

Hence, $\max_{m,k\le p}|\frac{1}{n}\sum_i X_{i,mk}| = O_{P^*}\left(\sqrt{\frac{\log p}{n}\left(\frac{\log n\log(pT)}{n} + \frac{\log(np)}{T}\right)\Delta_F^2}\right)$. ■

Lemma I.14. *For $\widehat{h}_{sk}\in\{\widehat{\delta}_{ys},\widehat{\delta}_{ds},\widehat{\lambda}_{sk}\}$, $s\leq T, k\leq p$:*
(i)

$$\max_{m,k\leq p}\left|\frac{1}{nT}\frac{1}{pT}\sum_{ist}\tilde{\Lambda}_t'\tilde{U}_{it}\tilde{U}_{is,m}\widehat{h}_{sk}\right|$$
$$=O_P\left(\frac{1}{pT}+\sqrt{\frac{\log p}{npT^2}}+\left(\sqrt{\frac{\log(pT)}{npT}}\right)\left(\sqrt{\frac{|J|_0}{n}}+\sqrt{\frac{\log(pT)}{n}}\right)\right)$$

(ii) $\max_{m,k\leq p}\left|\frac{1}{nT}\sum_{v=1}^p\sum_{i=1}^n\frac{1}{pnT}\sum_{st}\widehat{F}'\tilde{U}_{t,v}\tilde{U}_{it,v}\tilde{U}_{is,m}\widehat{h}_{sk}\right|=O_P\left(\sqrt{\frac{\log p}{npT}}+\frac{\sqrt{\log p}}{n}+\frac{1}{pT}\right)$
(iii)

$$\max_{m,k\leq p}\left\|\frac{1}{nT}\sum_{i=1}^n\sum_{s=1}^T\frac{1}{pT}\sum_{t=1}^T\widehat{\Lambda}_t'E^*\tilde{U}_{it}^*\tilde{U}_{is,m}^*\widehat{h}_{sk}'\right\|_F$$
$$=O_P\left(\frac{1}{pT}+\sqrt{\frac{\log p}{npT}}+\frac{\sqrt{\log p}}{n}+\frac{\sqrt{|J|_0}}{nT}+\sqrt{\frac{|J|_0}{npT}}\right)+O_P(\Delta_{ud})$$

Proof. (i) $\max_{m,k\leq p}|\frac{1}{nT}\sum_{i=1}^n\sum_{s=1}^T\frac{1}{pT}\sum_{t=1}^T\tilde{\Lambda}_t'\tilde{U}_{it}\tilde{U}_{is,m}\widehat{h}_{sk}|$ is bounded by

$$\max_{m,k\leq p}\left|\frac{1}{nT}\sum_{i=1}^n\sum_{s=1}^T\frac{1}{pT}\sum_{t=1}^T\tilde{\Lambda}_t'(EU_{it}U_{is,m})\widehat{h}_{sk}\right|+\max_{m,k\leq p}\left|\frac{1}{nT}\sum_{i=1}^n\sum_{s=1}^T\frac{1}{pT}\sum_{t=1}^T\tilde{\Lambda}_t'U_{it}\bar{U}_{i\cdot,m}\widehat{h}_{sk}\right|$$
$$+\max_{m,k\leq p}\left|\frac{1}{nT}\sum_{i=1}^n\sum_{s=1}^T\frac{1}{pT}\sum_{t=1}^T\tilde{\Lambda}_t'(U_{it}U_{is,m}-EU_{it}U_{is,m})\widehat{h}_{sk}\right|+O_P\left(\frac{1}{\sqrt{npT}}\sqrt{\frac{\log p}{n}}\right)\quad\text{(I.19)}$$

The first term in (I.19) is $O_P\left(\frac{1}{pT}\right)$. The second term in (I.19) is stochastically bounded by $O_P\left(\frac{1}{pT}\right)$ plus the third term. Note that

$$\max_{m,k\leq p}\left|\frac{1}{nT}\sum_{i=1}^n\sum_{s=1}^T\frac{1}{pT}\sum_{t=1}^T\tilde{\Lambda}_t'(U_{it}U_{is,m}-EU_{it}U_{is,m})\tilde{h}_{sk}\right|=O_P\left(\sqrt{\frac{\log p}{npT^2}}\right)$$

by (I.7). Hence, the third term in (I.19) is bounded by

$$O_P\left(\sqrt{\frac{\log p}{npT^2}}\right)+\max_{m,k\leq p}\left|\frac{1}{nT}\sum_{i=1}^n\sum_{s=1}^T\frac{1}{pT}\sum_{t=1}^T\tilde{\Lambda}_t'(U_{it}U_{is,m}-EU_{it}U_{is,m})(\widehat{h}_{sk}-H^{-1}\tilde{h}_{sk})\right|$$
$$\leq O_P\left(\sqrt{\frac{\log p}{npT^2}}\right)+\left[\max_{m,s}\left|\frac{1}{n}\sum_{i=1}^n\frac{1}{pT}\sum_{t=1}^T\tilde{\Lambda}_t'(U_{it}U_{is,m}-EU_{it}U_{is,m})\right|\times\right.$$

$$\max_{k\le p}\frac{1}{T}\sum_s\left(\left|\frac{1}{n}\widehat{F}'\tilde{U}_{s,k}\right|+\left|\frac{1}{n}\widehat{F}'\tilde{U}_t\gamma\right|+\left|\frac{1}{n}\widehat{F}'\tilde{e}_s\right|+\Delta_F\right)\Bigg]$$

$$\le O_P\left(\sqrt{\frac{\log p}{npT^2}}\right)+O_P\left(\sqrt{\frac{\log(pT)}{npT}}\right)\left(\sqrt{\frac{|J|_0}{n}}+\sqrt{\frac{\log(pT)}{n}}\right)$$

regardless of whether $h_{sk}=\delta_t$ or λ_{tk}. Hence,

$$\max_{m,k\le p}\left|\frac{1}{nT}\sum_{i=1}^n\sum_{s=1}^T\frac{1}{pT}\sum_{t=1}^T\tilde{\Lambda}_t'\tilde{U}_{it}\tilde{U}_{is,m}\widehat{h}_{sk}\right|$$

$$=O_P\left(\frac{1}{pT}+\sqrt{\frac{\log p}{npT^2}}+\left(\sqrt{\frac{\log(pT)}{npT}}\right)\left(\sqrt{\frac{|J|_0}{n}}+\sqrt{\frac{\log(pT)}{n}}\right)\right).$$

(ii) Note that $\sum_s\tilde{h}_{sk}=0$ for $h\in\{\delta,\lambda\}$. We then have that

$$\max_{m,k\le p}\left|\frac{1}{nT}\sum_{v=1}^p\sum_{i=1}^n\sum_{s=1}^T\frac{1}{pT}\sum_{t=1}^T\frac{1}{n}\sum_{j=1}^n\widehat{f}_j\tilde{U}_{jt,v}\tilde{U}_{it,v}\tilde{U}_{is,m}\widehat{h}_{sk}\right|$$

$$\le\max_{m,k\le p}\left|\frac{1}{nT}\sum_{v=1}^p\sum_{i=1}^n\sum_{s=1}^T\frac{1}{pT}\sum_{t=1}^T\frac{1}{n}\sum_{j=1}^n\widehat{f}_j\tilde{U}_{jt,v}U_{it,v}\tilde{U}_{is,m}\widehat{h}_{sk}\right|$$

$$+\max_{m,k\le p}\left|\frac{1}{nT}\sum_{v=1}^p\sum_{i=1}^n\sum_{s=1}^T\frac{1}{pT}\sum_{t=1}^T\frac{1}{n}\sum_{j=1}^n\widehat{f}_j\tilde{U}_{jt,v}\bar{U}_{t,v}\tilde{U}_{is,m}\widehat{h}_{sk}\right|$$

$$\le O_P\left(\Delta_F\frac{\sqrt{\log p}}{\sqrt{n}}+\frac{\sqrt{\log p}}{n}\right)$$

$$+\max_{m,k\le p}\left|\frac{1}{nT}\sum_{v=1}^p\sum_{i=1}^n\sum_{s=1}^T\frac{1}{pT}\sum_{t=1}^T\frac{1}{n}\sum_{j=1}^n\widehat{f}_j\tilde{U}_{jt,v}(EU_{it,v}U_{is,m})\widehat{h}_{sk}\right|$$

$$+\max_{m\le p}\left|\frac{1}{n}\sum_{v=1}^p\sum_{i=1}^n\frac{1}{pT}\sum_{t=1}^T\frac{1}{n}\sum_{j=1}^n\widehat{f}_j\tilde{U}_{jt,v}U_{it,v}\bar{U}_{i,m}\right|$$

$$+\max_{m,k\le p}\left|\frac{1}{nT}\sum_{v=1}^p\sum_{i=1}^n\sum_{s=1}^T\frac{1}{pT}\sum_{t=1}^T\frac{1}{n}\sum_{j=1}^n\widehat{f}_j\tilde{U}_{jt,v}(U_{it,v}U_{is,m}-EU_{it,v}U_{is,m})\widehat{h}_{sk}\right|$$

$$\le^{(a)}O_P\left(\Delta_F\frac{\sqrt{\log p}}{\sqrt{n}}+\frac{\sqrt{\log p}}{n}+\frac{1}{pT}\right)+O_P\left(\frac{1}{T\sqrt{p}}+\frac{1}{\sqrt{n}}\right)\left(\sqrt{\frac{\log(pT)}{nT}}\right)$$

$$+\max_{m,k\leq p}\left|\frac{1}{nT}\sum_{v=1}^{p}\sum_{i=1}^{n}\sum_{s=1}^{T}\frac{1}{pT}\sum_{t=1}^{T}\frac{1}{n}\sum_{j=1}^{n}\widehat{f}_j\tilde{U}_{jt,v}(U_{it,v}U_{is,m}-EU_{it,v}U_{is,m})(\widehat{h}_{sk}-H^{-1}\tilde{h}_{sk})\right|$$

$$\leq^{(b)} O_P\left(\Delta_F\frac{\sqrt{\log p}}{\sqrt{n}}+\frac{\sqrt{\log p}}{n}+\frac{1}{pT}\right)+O_P\left(\frac{1}{T\sqrt{p}}+\frac{1}{\sqrt{n}}\right)\left(\sqrt{\frac{\log(pT)}{nT}}\right)$$

$$+\max_{m,k\leq p}\left|\frac{1}{nT}\sum_{v=1}^{p}\sum_{i=1}^{n}\sum_{s=1}^{T}\frac{1}{pT}\sum_{t=1}^{T}\frac{1}{n}\sum_{j=1}^{n}\widehat{f}_j\tilde{U}_{jt,v}(U_{it,v}U_{is,m}-EU_{it,v}U_{is,m})\tilde{h}_{sk}\right|O_P(\Delta_F)$$

$$+\max_{m,k\leq p}\left|\frac{1}{nT}\sum_{v=1}^{p}\sum_{i=1}^{n}\sum_{s=1}^{T}\frac{1}{pT}\sum_{t=1}^{T}\frac{1}{n}\sum_{j=1}^{n}\widehat{f}_j\tilde{U}_{jt,v}(U_{it,v}U_{is,m}-EU_{it,v}U_{is,m})\frac{1}{n}\widehat{F}'\tilde{U}_{s,k}\right|$$

$$+\max_{m,k\leq p}\left|\frac{1}{nT}\sum_{v=1}^{p}\sum_{i=1}^{n}\sum_{s=1}^{T}\frac{1}{pT}\sum_{t=1}^{T}\frac{1}{n}\sum_{j=1}^{n}\widehat{f}_j\tilde{U}_{jt,v}(U_{it,v}U_{is,m}-EU_{it,v}U_{is,m})\frac{1}{n}\widehat{F}'\tilde{U}_{s}\gamma\right|$$

$$+\max_{m,k\leq p}\left|\frac{1}{nT}\sum_{v=1}^{p}\sum_{i=1}^{n}\sum_{s=1}^{T}\frac{1}{pT}\sum_{t=1}^{T}\frac{1}{n}\sum_{j=1}^{n}\widehat{f}_j\tilde{U}_{jt,v}(U_{it,v}U_{is,m}-EU_{it,v}U_{is,m})\frac{1}{n}\widehat{F}'\tilde{e}_{s}\right|$$

$$=O_P\left(\Delta_F\frac{\sqrt{\log p}}{\sqrt{n}}+\frac{\sqrt{\log p}}{n}+\frac{1}{pT}\right)+O_P\left(\frac{1}{T\sqrt{p}}+\frac{1}{\sqrt{n}}\right)\left(\sqrt{\frac{\log(pT)}{nT}}\right)$$

$$+O_P\left(\Delta_F+\frac{1}{\sqrt{n}}\right)\left(\Delta_F+\sqrt{\frac{\log(pT)}{n}}+\sqrt{\frac{|J|_0}{n}}\right)\sqrt{\frac{\log(pT)}{n}}$$

$$=O_P\left(\sqrt{\frac{\log p}{npT}}+\frac{\sqrt{\log p}}{n}+\frac{1}{pT}\right).$$

In (a) above, we used the bound in (I.6) for C_2. Also, we did not formally treat the term

$$\max_{m\leq p}\left|\frac{1}{n}\sum_{v=1}^{p}\sum_{i=1}^{n}\frac{1}{pT}\sum_{t=1}^{T}\frac{1}{n}\sum_{j=1}^{n}\widehat{f}_j\tilde{U}_{jt,v}U_{it,v}\bar{U}_{i,m}\right|$$

as can be bounded similarly to the term C_2 in (I.6). The inequality (b) holds for $h\in\{\delta,\lambda\}$.

(iii)

$$\max_{m,k\leq p}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}\frac{1}{pT}\sum_{t=1}^{T}\widehat{\Lambda}_t'E^*\tilde{U}_{it}^*\tilde{U}_{is,m}^*\widehat{h}_{sk}\right|$$

$$=\max_{m,k\leq p}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}\frac{1}{pT}\sum_{t=1}^{T}\widehat{\Lambda}_t'\widehat{U}_{it}\widehat{U}_{is,m}\widehat{h}_{sk}\right|$$

$$
\begin{aligned}
&\leq \max_{m,k\leq p}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}\frac{1}{pT}\sum_{t=1}^{T}\widehat{\Lambda}_t'\widehat{U}_{it}\tilde{U}_{is,m}\widehat{h}_{sk}\right| + \left|\frac{1}{n}\sum_{i=1}^{n}\frac{1}{pT}\sum_{t=1}^{T}\widehat{\Lambda}_t'\widehat{U}_{it}\widehat{f}_i\right| O_P(\Delta_F) \\
&\quad + \left|\frac{1}{n}\sum_{i=1}^{n}\frac{1}{pT}\sum_{t=1}^{T}\widehat{\Lambda}_t'\widehat{U}_{it}\widehat{f}_i\right| \max_{m,k\leq p}\left|\frac{1}{T}\sum_{s=1}^{T}\frac{1}{n}\tilde{F}'\tilde{U}_{s,m}\widehat{h}_{sk}\right| \\
&\quad + \left(\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{pT}\sum_{t=1}^{T}\widehat{\Lambda}_t'\widehat{U}_{it}\right)^2\right)^{1/2} O_P(\Delta_F).
\end{aligned}
$$

Note that $\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{pT}\sum_{t=1}^{T}\widehat{\Lambda}_t'\widehat{U}_{it}\right)^2 = O_P(\Delta_F)$ by (I.16) and $\max_{m,k\leq p}\left|\frac{1}{T}\sum_{s=1}^{T}\frac{1}{n}\tilde{F}'\tilde{U}_{s,m}\widehat{h}_{sk}\right| = O_P\left(\sqrt{\frac{\log p}{n}}\right)$, so $\left|\frac{1}{n}\sum_{i=1}^{n}\frac{1}{pT}\sum_{t=1}^{T}\widehat{\Lambda}_t'\widehat{U}_{it}\widehat{f}_i\right| = O_P(\Delta_F)$. We may then complete the bound as

$$
\begin{aligned}
&\max_{m,k\leq p}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}\frac{1}{pT}\sum_{t=1}^{T}\widehat{\Lambda}_t'E^*\tilde{U}_{it}^*\tilde{U}_{is,m}^*\widehat{h}_{sk}\right| \\
&\qquad \leq \max_{m,k\leq p}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}\frac{1}{pT}\sum_{t=1}^{T}\widehat{\Lambda}_t'\widehat{U}_{it}\tilde{U}_{is,m}\widehat{h}_{sk}\right| + O_P\left(\Delta_F + \sqrt{\frac{\log p}{n}}\right)\Delta_F \\
&\qquad \leq O_P\left(\Delta_F + \sqrt{\frac{\log p}{n}}\right)\Delta_F + \max_{m,k\leq p}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}\frac{1}{pT}\sum_{t=1}^{T}\widehat{\Lambda}_t'\tilde{U}_{it}\tilde{U}_{is,m}\widehat{h}_{sk}\right| \\
&\qquad\quad + \max_{m,k\leq p}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}(\widehat{f}_i - H'\tilde{f}_i)\tilde{U}_{is,m}\tilde{h}_{sk}\right| O_P(1) \\
&\qquad\quad + \max_{m,k\leq p}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}(\widehat{f}_i - H'\tilde{f}_i)\tilde{U}_{is,m}(\widehat{h}_{sk} - H^{-1}\tilde{h}_{sk})\right| \\
&\qquad\quad + \left(\frac{1}{p}\sum_{v=1}^{p}\frac{1}{T}\sum_{t=1}^{T}\left|\frac{1}{n}\tilde{F}'\tilde{U}_{t,v}\right|^2 + \left|\frac{1}{p}\sum_{v=1}^{p}\frac{1}{T}\sum_{t=1}^{T}\frac{1}{n}\tilde{F}'\tilde{U}_{t,v}\tilde{\lambda}_{tv}\right|\right)\max_{m,k\leq p}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}\tilde{f}_i'\tilde{U}_{is,m}\widehat{h}_{sk}\right| \\
&\qquad \leq O_P\left(\Delta_F + \sqrt{\frac{\log p}{n}} + \sqrt{\frac{|J|_0}{n}}\right)\Delta_F + O_P\left(\frac{1}{n} + \frac{1}{\sqrt{npT}}\right)\sqrt{\frac{\log p}{n}} + O_P(\Delta_{ud}) \\
&\qquad\quad + \max_{m,k\leq p}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}\frac{1}{pT}\sum_{t=1}^{T}\tilde{\Lambda}_t'\tilde{U}_{it}\tilde{U}_{is,m}\widehat{h}_{sk}\right| \\
&\qquad\quad + \max_{m,k\leq p}\left|\frac{1}{nT}\sum_{v=1}^{p}\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{\lambda}_{tv}\tilde{U}_{it,v}\sum_{s=1}^{T}\frac{1}{pT}\tilde{U}_{is,m}\widehat{h}_{sk}\right| O_P(\Delta_F)
\end{aligned}
$$

$$+\max_{m,k\leq p}\left|\frac{1}{nT}\sum_{v=1}^{p}\sum_{i=1}^{n}\sum_{s=1}^{T}\frac{1}{pT}\sum_{t=1}^{T}\frac{1}{n}\widehat{F}'\tilde{U}_{t,v}\tilde{U}_{it,v}\tilde{U}_{is,m}\widehat{h}_{sk}\right|$$

$$\leq O_P\left(\Delta_F+\sqrt{\frac{\log p}{n}}+\sqrt{\frac{|J|_0}{n}}+\frac{1}{T\sqrt{p}}\right)\Delta_F+O_P\left(\frac{1}{n}+\frac{1}{\sqrt{npT}}\right)\sqrt{\frac{\log p}{n}}+O_P(\Delta_{ud})$$

$$+\max_{m,k\leq p}\left|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}\frac{1}{pT}\sum_{t=1}^{T}\tilde{\Lambda}_t'\tilde{U}_{it}\tilde{U}_{is,m}\widehat{h}_{sk}\right|$$

$$+\max_{m,k\leq p}\left|\frac{1}{nT}\sum_{v=1}^{p}\sum_{i=1}^{n}\sum_{s=1}^{T}\frac{1}{pT}\sum_{t=1}^{T}\frac{1}{n}\widehat{F}'\tilde{U}_{t,v}\tilde{U}_{it,v}\tilde{U}_{is,m}\widehat{h}_{sk}\right|$$

$$\leq^{(i)(ii)}O_P\left(\Delta_F+\sqrt{\frac{\log p}{n}}+\sqrt{\frac{|J|_0}{n}}+\frac{1}{T\sqrt{p}}\right)\Delta_F+O_P\left(\frac{1}{n}+\frac{1}{\sqrt{npT}}\right)\sqrt{\frac{\log p}{n}}+O_P(\Delta_{ud})$$

$$+O_P\left(\frac{1}{pT}+\sqrt{\frac{\log p}{npT^2}}+\left(\sqrt{\frac{\log(pT)}{npT}}\right)\left(\sqrt{\frac{|J|_0}{n}}+\sqrt{\frac{\log(pT)}{n}}\right)+\sqrt{\frac{\log p}{npT}}+\frac{\sqrt{\log p}}{n}\right)$$

$$=O_P\left(\frac{1}{pT}+\sqrt{\frac{\log p}{npT}}+\frac{\sqrt{\log p}}{n}+\frac{\sqrt{|J|_0}}{nT}+\sqrt{\frac{|J|_0}{npT}}\right)+O_P(\Delta_{ud})$$

where $\max_{m,k\leq p}|\frac{1}{nT}\sum_{i=1}^{n}\sum_{s=1}^{T}(\widehat{f}_i-H'\tilde{f}_i)\tilde{U}_{is,m}\tilde{h}_{sk}|=O_P(\Delta_{ud})$ is given in Lemma I.3.

■

References

Acemoglu, D., Johnson, S. and Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review* **91** 1369–1401.

Ahn, S. C. and Horenstein, A. R. (2013). Eigenvalue ratio test for the number of factors. *Econometrica* **81** 1203–1227.

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica* **71** 135–171.

Bai, J. and Liao, Y. (2013). Statistical inferences using large estimated covariances for panel data and factor models. Tech. rep., University of Maryland.

Belloni, A., Chen, D., Chernozhukov, V. and Hansen, C. (2012). Sparse models and methods for optimal instruments with an application to eminent domain. *Econometrica* **80** 2369–2429.

Belloni, A. and Chernozhukov, V. (2013). Least squares after model selection in high-dimensional sparse models. *Bernoulli* **19** 521–547.

Belloni, A., Chernozhukov, V. and Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies* **81** 608–650.

Bühlmann, P. and van de Geer, S. (2011). *Statistics for high-dimensional data, methods, theory and applications*. The first edition ed. Springer, New York.

Cheng, G. and Huang, J. Z. (2010). Bootstrap consistency for general semiparametric m-estimation. *The Annals of Statistics* **38** 2884–2915.

Chernozhukov, V. and Hansen, C. (2008). The reduced form: A simple approach to inference with weak instruments. *Economics Letters* **100** 68–71.

Choi, I. (2012). Efficient estimation of factor models. *Econometric Theory* **28** 274–308.

Dedecker, J. and Prieur, C. (2004). Coupling for τ-dependent sequences and applications. *Journal of Theoretical Probability* **17** 861–885.

Horváth, L. and Shao, Q.-M. (1996). Limit theorem for maximum of standardized u-statistics with an application. *The Annals of Statistics* **24** 2266–2279.

Keller, W. J. and Wansbeek, T. (1983). Multivariate methods for quantitative and qualitative data. *Journal of Econometrics* **22** 91–111.

Merlevède, F., Peligrad, M. and Rio, E. (2011). A bernstein type inequality and moderate deviations for weakly dependent sequences. *Probability Theory and Related Fields* **151** 435–474.

Ng, S. (2015). Constructing common factors from continuous and categorical data. *Econometric Reviews* **34** 1141–1171.

Nolan, D. and Pollard, D. (1987). U-processes: rates of convergence. *The Annals of Statistics* 780–799.

Wang, W., Phillips, P. C. and Su, L. (2016). Homogeneity pursuit in panel data models: theory and applications .

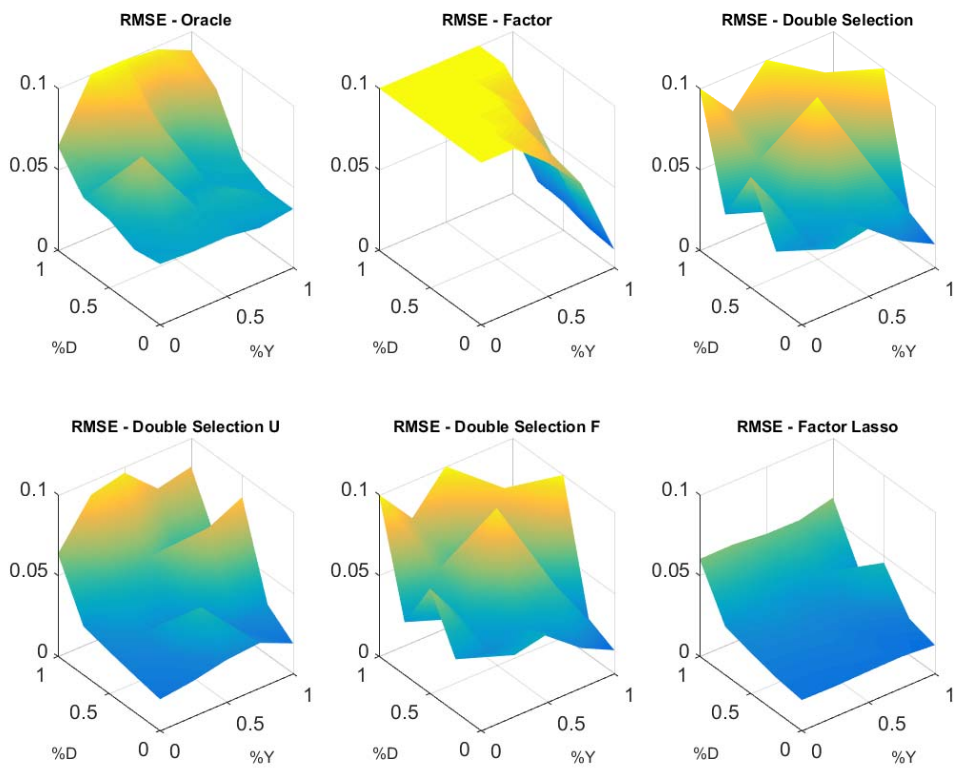

FIGURE 1. This figure shows the simulation RMSE of each of the estimators described in the text for estimating the coefficient of interest in an IV partial factor model. RMSE (truncated at 0.1) is shown in the vertical axis. The horizontal axes give the fraction of the explanatory power in an infeasible regression of Y on factors and factor residuals, "%Y," and the fraction of the explanatory power in an infeasible regression of D on factors and factor residuals, "%D," where the infeasible regressions are described in the text.

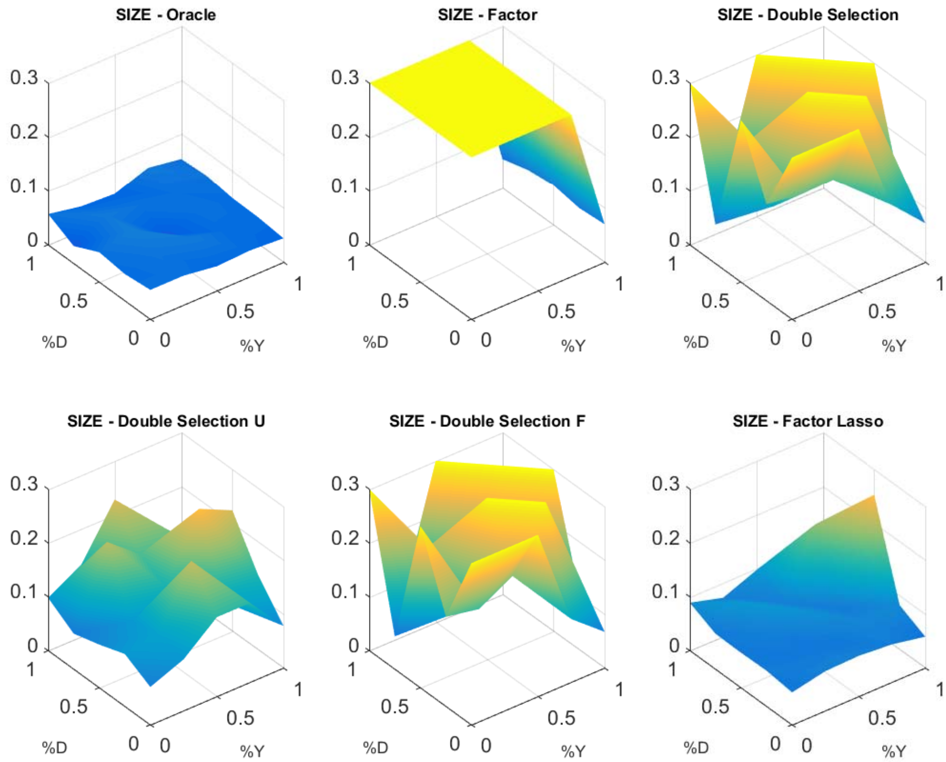

FIGURE 2. This figure shows the simulation size of 5% level tests based on each of the estimators described in the text for the IV partial factor model. Size (truncated at 0.3) is shown in the vertical axis. The horizontal axes give the fraction of the explanatory power in an infeasible regression of Y on factors and factor residuals, "%Y," and the fraction of the explanatory power in an infeasible regression of D on factors and factor residuals, "%D," where the infeasible regressions are described in the text.

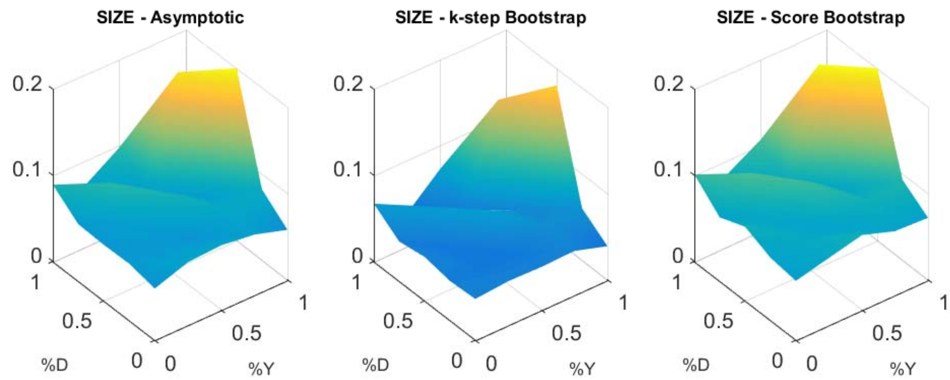

FIGURE 3. This figure shows the simulation size of 5% level tests based on the factor-lasso estimator in the IV partial factor model and the asymptotic Gaussian approximation, the k-step bootstrap, and a score based bootstrap. Size is shown in the vertical axis. The horizontal axes give the fraction of the explanatory power in an infeasible regression of Y on factors and factor residuals, "%Y," and the fraction of the explanatory power in an infeasible regression of D on factors and factor residuals, "%D," where the infeasible regressions are described in the text.

TABLE 1. Estimates of the First-Stage Relationship between Settler Mortality and Protection from Expropriation

	$\widehat{\pi}$	Estimated s.e.	Bootstrap C.I.
Latitude	-0.549	(0.166)	[-0.851,-0.246]
All Controls	-0.218	(0.168)	[-0.778,0.341]
Double Selection	-0.364	(0.178)	[-0.885,0.158]
Factor	-0.475	(0.173)	[-0.880,-0.070]
Factor-Lasso	-0.353	(0.183)	[-0.708.0.002]

This table presents estimates of the coefficient on the instrument (Settler Mortality) in the first-stage regression of the endogenous variable from the Acemoglu et al. (2001) example (Protection from Expropriation) on the instrument and geographic controls using different methods. The row labeled "Latitude" uses the single variable distance from the equator to control for geography. "All Controls" uses all 20 geographic controls without dimension reduction. "Double Selection" uses the approach of Belloni et al. (2014) to select important controls from among the 20 potential geography measures. "Factor" reduces dimension through positing a conventional factor model. "Factor-Lasso" makes use of the approach developed in this paper. Point estimates from each method are provided in the column "$\widehat{\pi}$" and the associated estimated asymptotic standard errors are given in "Estimated s.e.". The k-step bootstrap 95% confidence interval is reported in "Bootstrap C.I.".

BOOTH SCHOOL OF BUSINESS, UNIVERSITY OF CHICAGO, CHICAGO, IL 60637, CHRISTIAN.HANSEN@CHICAGOBOOT

E-mail address: Christian.Hansen@chicagobooth.edu

DEPARTMENT OF ECONOMICS, RUTGERS UNIVERSITY, NEW BRUNSWICK, NJ 08901

E-mail address: yuan.liao@rutgers.edu