# A Bayesian framework for sparse estimation in high-dimensional mixed frequency Vector Autoregressive models

Nilanjana Chakraborty, Kshitij Khare, George Michailidis

*University of Florida*

**Supplementary Material**

## S1 Observation-driven VAR model in Ghysels (2016)

For $k_1$ monthly and $k_2$ quarterly variables, the mixed frequency Bayesian VAR model with lag $d$ proposed by Ghysels (2016) is given by

$$\begin{bmatrix} x_H(\tau_L,1)_{k_1\times 1} \\ x_H(\tau_L,2)_{k_1\times 1} \\ x_H(\tau_L,3)_{k_1\times 1} \\ x_L(\tau_L)_{k_2\times 1} \end{bmatrix} = \sum_{u=1}^{d} \begin{bmatrix} \mathbf{W}_u^{1,1} & \mathbf{W}_u^{1,2} & \mathbf{W}_u^{1,3} & \mathbf{W}_u^{1,4} \\ \mathbf{W}_u^{2,1} & \mathbf{W}_u^{2,2} & \mathbf{W}_u^{2,3} & \mathbf{W}_u^{2,4} \\ \mathbf{W}_u^{3,1} & \mathbf{W}_u^{3,2} & \mathbf{W}_u^{3,3} & \mathbf{W}_u^{3,4} \\ \mathbf{W}_u^{4,1} & \mathbf{W}_u^{4,2} & \mathbf{W}_u^{4,3} & \mathbf{W}_u^{4,4} \end{bmatrix} \begin{bmatrix} x_H(\tau_L-u,1) \\ x_H(\tau_L-u,2) \\ x_H(\tau_L-u,3) \\ x_L(\tau_L-u) \end{bmatrix} + \boldsymbol{\varepsilon}(\tau_L) \tag{S1.1}$$

where $x_L(\tau_L)$ is a $k_2\times 1$ quarterly process and $x_H(\tau_L,m)$ is a $k_1\times 1$ monthly process which is observed at the $m^{th}$ month, $m=1,2,3$, during a particular quarter. Here $\dim(\mathbf{W}_u^{4,4}) = k_2 \times k_2$, $\dim(\mathbf{W}_u^{i,4}) = k_1 \times k_2$ for $i = 1,2,3$, $\dim(\mathbf{W}_u^{4,j}) = k_2 \times k_1$ for $j = 1,2,3$ and $\dim(\mathbf{W}_u^{a,b}, a, b = 1,2,3) = k_1 \times k_1$.

As mentioned in Remark S1, Ghysels (2016) assumed the following structure in order to deal with parameter proliferation

$$\left[\mathbf{W}_u^{4,1}\ \mathbf{W}_u^{4,2}\ \mathbf{W}_u^{4,3}\right] = (\boldsymbol{\omega}(\boldsymbol{\gamma})_1\ \ \boldsymbol{\omega}(\boldsymbol{\gamma})_2\ \ \boldsymbol{\omega}(\boldsymbol{\gamma})_3) \otimes \mathbf{B}_u$$

which reduces the number of parameters in that specific block from $3k_1k_2$ to $(k_1k_2+3)$ in terms of a $k_2 \times k_1$ dimensional matrix $\mathbf{B}_u$ and a low dimensional vector $\boldsymbol{\omega}(\boldsymbol{\gamma})$.

# S2 Additional methodological details for the Bayesian MF model

## S2.1 Expressions for $y_{i,\mathbf{H}}^{t-1/3}$ and $y_{i,\mathbf{H}}^{t}$

The forecasted value of the $i^{th}$ monthly variable at time $t-1/3$ and time $t$ are given by the following equations (S2.1) and (S2.2) respectively.

$$a_{i,1}\theta^3 y_{1,\mathrm{H}}^{t-5/3} + a_{i,1}\theta^2 y_{1,\mathrm{H}}^{t-4/3} + a_{i,1}\theta y_{1,\mathrm{H}}^{t-1} + \cdots + a_{i,k_1}\theta^3 y_{k_1,\mathrm{H}}^{t-5/3} + a_{i,k_1}\theta^2 y_{k_1,\mathrm{H}}^{t-4/3} + a_{i,k_1}\theta y_{k_1,\mathrm{H}}^{t-1} + a_{i,k_1+1}\theta y_{1,\mathrm{L}}^{t-1} + \cdots + a_{i,k_1+k_2}\theta y_{k_2,\mathrm{L}}^{t-1} + \varepsilon_{i,\mathrm{H}}^{t-1/3} \tag{S2.1}$$

and,

$$a_{i,1}\theta^4 y_{1,\mathrm{H}}^{t-5/3} + a_{i,1}\theta^3 y_{1,\mathrm{H}}^{t-4/3} + a_{i,1}\theta^2 y_{1,\mathrm{H}}^{t-1} + \cdots + a_{i,k_1}\theta^4 y_{k_1,\mathrm{H}}^{t-5/3} + a_{i,k_1}\theta^3 y_{k_1,\mathrm{H}}^{t-4/3} + a_{i,k_1}\theta^2 y_{k_1,\mathrm{H}}^{t-1} +$$

$$a_{i,k_1+1}\theta^2 y_{1,\mathrm{L}}^{t-1} + \cdots + a_{i,k_1+k_2}\theta^2 y_{k_2,\mathrm{L}}^{t-1} + \varepsilon_{i,\mathrm{H}}^{t} \qquad \text{(S2.2)}$$

## S2.2 Diagrammatic representation of the proposed VAR model

Equations (2.4) and (2.5) in Section 2 of the main paper are visualized in Figure 1 for a model with $k_1 = 1$ and $k_2 = 2$ variables, wherein only the temporal dependence of all the lagged values on one monthly ($y_{1,\mathrm{H}}^{t-2/3}$) and one low-frequency variable ($y_{2,\mathrm{L}}^{t}$) are depicted (blue directed edges). The contemporaneous dependence of the model is captured by the red bi-directed edges.

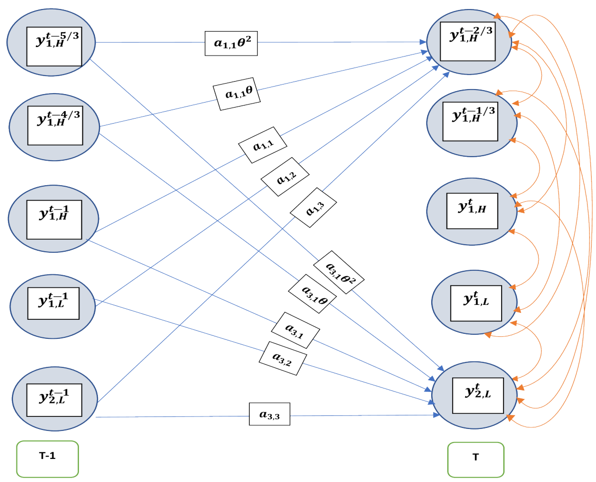


Figure 1: Temporal dependence structure of the VAR model in (2.1), (2.2) and (2.3) of the main paper with $d = 1$ and one monthly and two quarterly variables

## S2.3 A regression interpretation of the pseudo-likelihood function

It is well known that a VAR model can be expressed in the form of a multivariate regression model (Lütkepohl, 2005)). Specfically, based on the data $\{\boldsymbol{y}^0, \boldsymbol{y}^1, \cdots, \boldsymbol{y}^T\}$ (Note that, the likelihood and pseudo-likelihood functions in (3.1) and (3.3) of the main paper are implicitly conditional on the initial observation $\boldsymbol{y}^0$.) , define the response matrix $\mathbf{Y}$ and design matrix $\mathbf{X}$ as follows,

$$\mathbf{Y} = \begin{bmatrix} (\boldsymbol{y}^T)' \\ \vdots \\ (\boldsymbol{y}^d)' \end{bmatrix}_{n\times(3k_1+k_2)} \qquad \mathbf{X} = \begin{bmatrix} (\boldsymbol{y}^{T-1})' & \cdots & (\boldsymbol{y}^{T-d})' \\ \vdots & \ddots & \vdots \\ (\boldsymbol{y}^{d-1})' & \cdots & (\boldsymbol{y}^0)' \end{bmatrix}_{n\times d(3k_1+k_2)}.$$

The VAR model in (2.6) of the main paper can be expressed as a multivariate linear regression of the form

$$\mathbf{Y} = \mathbf{X}\mathbf{W}^{\perp} + \mathbf{E}, \tag{S2.3}$$

where, $\mathbf{W}^{\perp} = [\mathbf{W}_1, \mathbf{W}_2, \ldots, \mathbf{W}_d]'$ and $\mathbf{E} = \left[\boldsymbol{\varepsilon}^T, \ldots, \boldsymbol{\varepsilon}^d\right]'$. In the above formulation, the number of samples is, $n = T - d + 1$. Note that the autocovariance function of a $p$-dimensional centered covariance-stationary time series $\{\boldsymbol{y}^t\}$ is defined as $\Gamma(h) = \mathrm{Cov}(\boldsymbol{y}^t, \boldsymbol{y}^{t+h}) \quad t, h \in \mathbb{Z}$ and it is invariant in $t$ for a stable VAR process. Since each $\boldsymbol{\varepsilon}^t$ is independent and identically distributed (i.i.d,) according to a $\mathcal{N}_{(3k_1+k_2)}(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}})$, each row of $\mathbf{X}$ is distributed as $\mathcal{N}_{d(3k_1+k_2)}(\mathbf{0}, \mathbf{C}_X)$, where the structure of the covariance matrix is given by,

$$\mathbf{C}_X = \begin{bmatrix} \boldsymbol{\Gamma}(0) & \boldsymbol{\Gamma}(1) & \cdots & \boldsymbol{\Gamma}(d-1) \\ \boldsymbol{\Gamma}(1)' & \boldsymbol{\Gamma}(0) & \cdots & \boldsymbol{\Gamma}(d-2) \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Gamma}(d-1)' & \boldsymbol{\Gamma}(d-2)' & \cdots & \boldsymbol{\Gamma}(0) \end{bmatrix}_{d(3k_1+k_2)\times d(3k_1+k_2)} \tag{S2.4}$$

The joint matrix formulation in (S2.3) can be equivalently expressed as a collection of individual regression equations of $k_1$ high frequency and $k_2$ low frequency variables. In particular, for $3k_1 + 1 \leq j \leq 3k_1 + k_2$, focusing on the $j^{th}$ column of $\mathbf{Y}$ in (S2.3) gives us the linear regression equation

$$\boldsymbol{y}_{j,\mathrm{L}} = \sum_{u=1}^{d} \mathbf{Z}_u \mathbf{A}'_{u_{j.}}, + \boldsymbol{\xi}_{j,\mathrm{L}} \tag{S2.5}$$

where the $n \times 1$ response vector $\boldsymbol{y}_{j,\mathrm{L}} = (y^t_{j,\mathrm{L}})^T_{t=d}$ contains all observations of the $j^{th}$ low frequency variable, the design matrix $\mathbf{Z}_u$ is a $n \times (k_1 + k_2)$ matrix whose $t^{th}$ row is given by

$$\boldsymbol{z}_u^{t-1\prime} = \left( y^{t-1}_{1,\mathrm{H}} + \theta^2 y^{t-5/3}_{1,\mathrm{H}} + \theta y^{t-4/3}_{1,\mathrm{H}}, \ldots, y^{t-1}_{k_1,\mathrm{H}} + \theta^2 y^{t-5/3}_{k_1,\mathrm{H}} + \theta y^{t-4/3}_{k_1,\mathrm{H}}, y^{t-1}_{1,\mathrm{L}}, \ldots, y^{t-1}_{k_2,\mathrm{L}} \right), \tag{S2.6}$$

$\mathbf{A}'_{u_{j.}}$ is a $(k_1 + k_2) \times 1$ coefficient vector corresponding to the $(k_1 + j)^{th}$ row of $\mathbf{A}_u$, and the error vector $\boldsymbol{\xi}_{j,\mathrm{L}}$ corresponds to the $(3k_1 + j)^{th}$ column of $\mathbf{E}$.

Similarly, for $1 \leq i \leq k_1$, focusing on the $3i - 2, 3i - 1, 3i$ columns of $\mathbf{Y}$ in (S2.3) gives us the 3-variate regression equation

$$\mathbf{Y}_{i,\mathrm{H}} = \sum_{u=1}^{d} \mathbf{Z}_u \left( (\theta^2 \ \theta \ 1) \otimes \mathbf{A}'_{u_{i.}} \right) + \mathbf{E}_{i,\mathrm{H}}, \tag{S2.7}$$

where the $n \times 3$ response matrix $\mathbf{Y}_{i,\mathrm{H}}$ has three columns containing the observations of the third, first, and second months in each quarter for the $i^{th}$ high frequency variable, respectively, the design matrix $\mathbf{Z}$ is as defined in (S2.6), the $(k_1 + k_2)$ coefficient matrix is obtained by taking the Kronecker product of the vector $\mathbf{A}'_{u_{i.}}$ ($i^{th}$ row of $\mathbf{A}_u$) and $(\theta^2 \ \theta \ 1)$, and the $n \times 3$

error matrix $\mathbf{E}_{i,\mathrm{H}}$ is obtained from the $3i-2, 3i-1, 3i$ columns of $\mathbf{E}$. Note that each row of $\mathbf{Z}$ is distributed as $\mathcal{N}_{d(k_1+k_2)}(\mathbf{0}, \mathrm{V}(\boldsymbol{z}^t))$. Let us denote $\mathbf{C}_Z \equiv \mathrm{V}(\boldsymbol{z}^t)$.

Simple calculations show that the pseudo-likelihood function $L_{\text{pseudo}}$ in (3.3) can be alternatively expressed as

$$
\begin{aligned}
& L_{\text{pseudo}}\left(\{\mathbf{A}_u\}_{u=1}^{d}, \theta, \{\Sigma_{i,\mathrm{H}}\}_{i=1}^{k_1}, \left\{\sigma_j^2\right\}_{j=1}^{k_2} \mid \boldsymbol{y}^t, t=d,\cdots,n+d-1\right) \\
\propto\ & \prod_{i=1}^{k_1}\left\{\exp\left[-\frac{1}{2}tr\left(\left(\mathbf{Y}_{i,\mathrm{H}}-\sum_{u=1}^{d}\mathbf{Z}_u\left((\theta^2\ 1\ \theta)\otimes\mathbf{A}'_{u_{i.}}\right)\right)\boldsymbol{\Sigma}_{i,\mathrm{H}}^{-1}\left(\mathbf{Y}_{i,\mathrm{H}}-\sum_{u=1}^{d}\mathbf{Z}_u\left((\theta^2\ 1\ \theta)\otimes\mathbf{A}'_{u_{i.}}\right)\right)'\right)\right]\right. \\
& \left.\times|\boldsymbol{\Sigma}_{i,\mathrm{H}}|^{-n/2}\right\}\prod_{j=1}^{k_2}\left\{\exp\left[-\frac{1}{2\sigma_j^2}\left(\boldsymbol{y}_{j,\mathrm{L}}-\sum_{u=1}^{d}\mathbf{Z}_u\mathbf{A}'_{u_{j.}}\right)'\left(\boldsymbol{y}_{j,\mathrm{L}}-\sum_{u=1}^{d}\mathbf{Z}_u\mathbf{A}'_{u_{j.}}\right)\right]\left(\frac{1}{\sigma_j^2}\right)^{n/2}\right\}. \qquad \text{(S2.8)}
\end{aligned}
$$

Hence, $L_{pseudo}$ can be algebraically obtained by the following process: (a) Compute the "likelihood" for each of the $k_1+k_2$ regressions outlined in (S2.5) and (S2.7) by treating the corresponding design matrix as fixed and uncorrelated with the errors. (b) Take the product of all these $k_1+k_2$ "likelihood" functions. This algebraic representation/interpretation of $L_{\text{pseudo}}$ will be quite useful both for computational and theoretical purposes in subsequent analysis.

## S2.4 Gibbs Sampling Algorithm to generate samples from the joint posterior of the parameters of the Bayesian MF model

Step 1. **Input**: data $\{\boldsymbol{y}^t, t=1,2,\cdots,T\}, \tau, \omega, q, \mathbf{V}, \alpha, \beta$

Step 2. **Initialize**: all the components of the parameter space to $\mathbf{A}^{(0)}, \theta^{(0)}, \left\{\mathbf{\Sigma}_{i,\mathrm{H}}^{(0)}\right\}_{i=1}^{k_1}$, $\left\{(\sigma_j^2)^{(0)}\right\}_{j=1}^{k_2}$. Set $r = 0$.

Step 3. Sample sequentially from the full conditional posterior distribution of each entry of $\mathbf{A}$ (mixture of point mass at zero and an appropriate normal), with all other parameters fixed at their current values.

Step 4. Sample $\{\mathbf{\Sigma}_{i,\mathrm{H}}\}_{i=1}^{k_1}$ from their respective full conditional posterior distributions (independent Inverse-Wisharts), with all the other parameters fixed at their current values

Step 5. Sample $\left(\sigma_j^2\right)_{j=1}^{k_2}$ from their respective full conditional posterior distributions (independent Inverse-gammas), with all the other parameters fixed at their current values

Step 6. Sample $\theta$ from its full conditional posterior distribution.

Step 7. Denote the current values of the parameters by $\mathbf{A}^{(r)}, \theta^{(r)}, \left\{\mathbf{\Sigma}_{i,\mathrm{H}}^{(r)}\right\}_{i=1}^{k_1}, \left\{(\sigma_j^2)^{(r)}\right\}_{j=1}^{k_2}$.

Step 8. If convergence criterion (based on number of iterations, deviation in the cumulative average of an appropriate function etc.) is satisfied, stop. Otherwise $r = r + 1$, go to Step 2.

Step 9. **Output**: $\left(\mathbf{A}^{(r)}, \theta^{(r)}, \left\{\mathbf{\Sigma}_{i,\mathrm{H}}^{(r)}\right\}_{i=1}^{k_1}, \left\{(\sigma_j^2)^{(r)}\right\}_{j=1}^{k_2}\right)_{r=1}^{K}$

### S2.5 Posterior probabilities for activity graphs for the Bayesian MF model

Here the goal is to derive a closed form expression (up to an unknown multiplicative constant) for the posterior probability distribution for the activity graph $\mathcal{G}$. Recall that the activity graph $\mathcal{G}$ captures the sparsity pattern in $\mathbf{A}$. We first define some quantities related to the individual regressions in (S2.5) and (S2.7), that will appear in the aforementioned closed form expression for $\pi(\mathcal{G} \mid \mathcal{Y})$.

For every $1 \leq i \leq k_1$, we have from (S2.7) that

$$\mathbf{Y}_{i,\mathrm{H}} = \mathbf{Z}\left((\theta^2\ \theta\ 1) \otimes \mathbf{A}'_{i\cdot}\right) + \mathbf{E}_{i,\mathrm{H}}.$$

Post-multiplying by the orthogonal matrix $\mathbf{C}$ where $\mathbf{C} = [\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\delta}]$, $\boldsymbol{a}, \boldsymbol{b}$ being arbitrary (see the material after (3.4)), gives

$$\tilde{\mathbf{Y}}_{i,\mathrm{H}} = \mathbf{Y}_{i,\mathrm{H}}\mathbf{C} = \mathbf{Z}\left((0\ 0\ \ \|\boldsymbol{\delta}\|^2) \otimes \mathbf{A}'_{i\cdot}\right) + \mathbf{E}_{i,\mathrm{H}}\mathbf{C} \tag{S2.9}$$

Let $\tilde{\mathbf{y}}_{i1} = \mathbf{Y}_{i,\mathrm{H}}\boldsymbol{a}, \tilde{\mathbf{y}}_{i2} = \mathbf{Y}_{i,\mathrm{H}}\boldsymbol{b}, \tilde{\mathbf{y}}_{i3} = \mathbf{Y}_{i,\mathrm{H}}\boldsymbol{\delta}$ be the three columns of $\tilde{\mathbf{Y}}_{i,\mathrm{H}}$. Then (S2.9) can be alternatively expressed as

$$\begin{aligned}
\tilde{\mathbf{y}}_{i1} &= \mathbf{E}_{i,\mathrm{H}}\boldsymbol{a} =: \tilde{\boldsymbol{\xi}}_{i1} \\
\tilde{\mathbf{y}}_{i2} &= \mathbf{E}_{i,\mathrm{H}}\boldsymbol{b} =: \tilde{\boldsymbol{\xi}}_{i2} \\
\tilde{\mathbf{y}}_{i3} &= \|\boldsymbol{\delta}\|^2\, \mathbf{Z}\mathbf{A}'_{i\cdot} + \mathbf{E}_{i,\mathrm{H}}\boldsymbol{\delta} \qquad \text{(S2.10)}
\end{aligned}$$

Thus, the first two regressions in (S2.10) have no predictors. The matrix $\mathbf{Z}$ and the coeffcient vector $\mathbf{A}_{i\cdot}$ play a role only in the last regression.

For $1 \leq i \leq k_1 + k_2$, let $\nu_i = \nu_i(\mathcal{G})$ denote the number of non-zero/active entries in the $i^{th}$ row of $\mathbf{A}$ under the activity graph $\mathcal{G}$. Let $\mathbf{Z}_{\mathcal{G},i}$ and $\mathbf{Z}_{\mathcal{G},j}$ denote the sub-matrix of $\mathbf{Z}$ consisting only of columns of $\mathbf{Z}$ corresponding to the active entries in the $i^{th}$ and $(k_1 + j)^{th}$ row of $\mathbf{A}$ respectively. Let $\mathbf{A}_{\mathcal{G},i\cdot}$ denote the $\nu_i$ dimensional sub-vector of $\mathbf{A}_{i\cdot}$ consisting only the active coefficients in the $i^{th}$ row of $\mathbf{A}$ corresponding to the $i^{th}$ high frequency variable. Similarly, let $\mathbf{A}_{\mathcal{G},j\cdot}$ denote the $\nu_{k_1+j}$ dimensional sub-vector of $\mathbf{A}_{j\cdot}$ consisting only the active coefficients in the $(k_1+j)^{th}$ row of $\mathbf{A}$ corresponding to the $j^{th}$ low frequency variable. Then, the transformed high-frequency regression equations in (S2.10) can be written as,

$$\begin{aligned}
\tilde{\boldsymbol{y}}_{i1} &= \tilde{\boldsymbol{\xi}}_{i1} \\
\tilde{\boldsymbol{y}}_{i2} &= \tilde{\boldsymbol{\xi}}_{i2} \\
\tilde{\boldsymbol{y}}_{i3} &= \|\boldsymbol{\delta}\|^2 \mathbf{Z}_{\mathcal{G},i} \mathbf{A}'_{\mathcal{G},i\cdot} + \mathbf{E}_{i,\mathrm{H}} \boldsymbol{\delta}
\end{aligned} \tag{S2.11}$$

for $1 \leq i \leq k_1$, and the low-frequency regression equations in (S2.5) can be written as

$$\boldsymbol{y}_{j,\mathrm{L}} = \mathbf{Z}_{\mathcal{G},j} \mathbf{A}'_{\mathcal{G},j\cdot} + \boldsymbol{\xi}_{j,\mathrm{L}} \tag{S2.12}$$

for $1 \leq j \leq k_2$. Let

$$\mathbf{R}_i = \frac{1}{n}\tilde{\mathbf{Y}}'_{i,\mathrm{H}}\tilde{\mathbf{Y}}_{i,\mathrm{H}} - \left(\frac{1}{n}\tilde{\mathbf{Y}}'_{i,\mathrm{H}}\mathbf{Z}_{\mathcal{G},i}\right)\left(\frac{1}{n}\mathbf{Z}'_{\mathcal{G},i}\mathbf{Z}_{\mathcal{G},i} + \frac{\mathbf{I}}{n\tau^2}\right)^{-1}\left(\frac{1}{n}\mathbf{Z}'_{\mathcal{G},i}\tilde{\mathbf{Y}}_{i,\mathrm{H}}\right) + \frac{\mathbf{C}^{-1}\mathbf{V}\mathbf{C}'^{-1}}{n}$$

for $1 \leq i \leq k_1$, and

$$R_j = \frac{\boldsymbol{y}'_{j,\mathrm{L}}\boldsymbol{y}_{j,\mathrm{L}}}{n} - \frac{\boldsymbol{y}'_{j,\mathrm{L}}\mathbf{Z}_{\mathcal{G},j}}{n}\left(\frac{\mathbf{Z}'_{\mathcal{G},j}\mathbf{Z}_{\mathcal{G},j}}{n} + \frac{\mathbf{I}}{n\tau^2}\right)^{-1}\frac{\mathbf{Z}'_{\mathcal{G},j}\boldsymbol{y}_{j,\mathrm{L}}}{n} + \frac{\beta}{n}$$

for $1 \leq j \leq k_2$. Let us write $\mathbf{R}_i$ as

$$\mathbf{R}_i = \begin{bmatrix} \mathbf{R}_{i,11_{2\times 2}} & \mathbf{R}_{i,12_{2\times 1}} \\ \mathbf{R}_{i,21_{1\times 2}} & \mathbf{R}_{i,22_{1\times 1}} \end{bmatrix}$$

The $3 \times 3$ matrices $\{\mathbf{R}_i\}_{i=1}^{k_1}$ and the scalars $\{R_j\}_{j=1}^{k_2}$ represent (up to a small modification of the $\mathbf{I}/n\tau^2$ term) the sample error covariance matrices/scalars for the regressions outlined in (S2.11) and (S2.12) respectively. Using (S2.8) of main paper, (S2.11) and (S2.12), it can be shown by straightforward computations that for fixed $\theta$,

$$\begin{aligned}
& \pi_{\text{pseudo}}(\mathcal{G} \mid \mathcal{Y}) \\
\propto\ & \pi(\mathcal{G}) \int\int L_{\text{pseudo}}(\mathbf{A}, \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} \mid \boldsymbol{y}^t, t = 1, \cdots, T)\pi(\mathbf{A}|\mathcal{G}, \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}})\pi(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}})\mathbf{d}(\mathbf{A})\,\mathbf{d}(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}) \\
\propto\ & \prod_{j=1}^{k_2}\left(\frac{1}{\tau\sqrt{n}}\right)^{\nu_{k_1+j}(\mathcal{G})}\left|\frac{\mathbf{Z}'_{\mathcal{G},j}\mathbf{Z}_{\mathcal{G},j}}{n} + \frac{\mathbf{I}}{n\tau^2}\right|^{-1/2}(R_j)^{-(\frac{n}{2}+\alpha)}\, q^{\nu_{k_1+j}(\mathcal{G})}\,(1-q)^{(k_1+k_2)-\nu_{k_1+j}(\mathcal{G})} \times \\
& \prod_{i=1}^{k_1}\left(\frac{1}{\tau\sqrt{n}}\right)^{\nu_i(\mathcal{G})}\left|\left(\frac{\mathbf{Z}'_{\mathcal{G},i}\mathbf{Z}_{\mathcal{G},i}}{n} + \frac{\mathbf{I}}{n\tau^2}\right)\mathbf{R}_{i,11}\right|^{-1/2}\left(\mathbf{R}_{i,22} - \mathbf{R}_{i,12}{}'\mathbf{R}_{i,11}{}^{-1}\mathbf{R}_{i,12}\right)^{-(\frac{n+\nu}{2})} \times \\
& q^{\nu_i(\mathcal{G})}\,(1-q)^{(k_1+k_2)-\nu_i(\mathcal{G})} \qquad \text{(S2.13)}
\end{aligned}$$

In practice, we can always obtain a quick estimate of $\theta$ by running just a few hundred iterations of the Gibbs sampler algorithm described in Section

S2.4. Such estimates of $\theta$ are reasonably accurate based on our numerical work. Thus the posterior graph selection probability can be factored into terms regarding to $(k_1+k_2)$ individual regressions of high and low frequency variables as given in (S2.5) and (S2.7). The proof is provided in Section S4.

## S2.6 Incorporating contemporaneous relationship for posterior predictive distribution estimation

As described in Section 6.1 of the main paper, we compute the forecasts of the quarterly variables using the median of their posterior predictive distributions. Note that, generating a sample from the posterior predictive distribution at a future time point requires generating samples both from the posterior distribution of the $\mathbf{W}_u$s and the error distribution. Recall that, in our pseudo-likelihood approach we assume a block-diagonal structure of the nuisance parameter $\mathbf{\Sigma}_{\boldsymbol{\varepsilon}}$ to obtain the posterior estimate of $\mathbf{W}$. Hence, again ignoring the contemporaneous correlations while evaluating the posterior predictive distribution, by imposing a block diagonal structure on $\mathbf{\Sigma}_{\boldsymbol{\varepsilon}}$, might lead to a further loss of efficiency. To alleviate this concern, we use the following covariance estimation method for the purposes of evaluating the posterior predictive distributions.

- Use the posterior means $\{\hat{\mathbf{W}}_{u,PM}\}_{u=1}^{d}$ to compute the estimated errors

$\hat{\boldsymbol{\varepsilon}}^t = \boldsymbol{y}^t - \sum_{u=1}^{d} \hat{\mathbf{W}}_{u,PM} \boldsymbol{y}^{t-u}$.

- Note that the unobserved errors $\{\boldsymbol{\varepsilon}^t\}_{t=1}^T$ ate i.i.d. multivariate normal with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$. Use the estimated errors as proxies for the unobserved errors to obtain a sparse estimate of $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$ using the graphical lasso algorithm.

Now the samples from the posterior distribution of $\mathbf{W}$ along with the estimated $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$ can be used to generate samples from the posterior predictive distribution at a future time point.

### S2.7 Comparison with existing approaches

**Remark S1.** *(Comparison with the observational approach in Ghysels (2016)): As previously mentioned, we start with the observation-driven approach of Ghysels (2016) by expanding each high-frequency time series into a corresponding number of low-frequency series. However, there are crucial differences in the two approaches in terms of how parameter reduction is achieved. Consider the quarterly-monthly setting for illustration. In the Bayesian approach discussed in (Ghysels, 2016, Section 5.2), the parameter proliferation is controlled only in the sub-matrices* $\boldsymbol{W}_u^{4,1}$, $\boldsymbol{W}_u^{4,2}$, $\boldsymbol{W}_u^{4,3}$ *of the VAR transition matrix* $\boldsymbol{W}_u$ *(details given in Supplemental Section S1) by expressing these* 3 *block matrices in terms of a single* $k_2 \times k_1$ *matrix*

*$\boldsymbol{B}_u$ and a low dimensional vector $\boldsymbol{\omega}(\boldsymbol{\gamma})_{3\times 1}$. The other blocks of the VAR transition matrix $\boldsymbol{W}_u$ are kept intact, and zero mean Minnesota prior distributions with appropriate prior covariance matrices are used. In all, there are $d*(9k_1^2+k_2^2+4k_1k_2+3)$ parameters, which does not resolve the issue of over-parametrization when $k_1, k_2$ are both comparable to or larger than the number of observations in the data and it makes it worse whenever $k_1$ is large. McCracken et al. (2015) also use the observational approach of Ghysels (2016) for forecasting purposes using a data set with 1 quarterly variable (GDP) and 12 monthly variables. A conjugate multivariate normal/Inverse Wishart prior structure is used for the regression coefficients and the error covariance matrix, with hyperparameters introduced to control the tightness of the prior. However, direct approaches for parameter reduction (sparsity or otherwise) are not considered, and this can lead to scalability issues in moderate/high-dimensinal settings. On the other hand, the structure imposed in (2.2) and (2.3) allows us to express the VAR transition matrices $\boldsymbol{W}_u$ as in (2.7) (with $p^2=(3k_1+k_2)^2$ entries) in terms of $\boldsymbol{A}_u$ and $\theta$ (a total of $(k_1+k_2)^2+1$ parameters). A further parameter reduction is performed for meaningful estimation when $(k_1+k_2)$ (total number of variables in the data set) is comparable to or larger than the sample size $T$ by using sparsity inducing prior distributions on the elements of $\boldsymbol{A}$*

*(see Section 3.1). To summarize, in a high dimensional setting, the proposed approach deals with the over-parametrization issue in a more efficient way, compared to Ghysels (2016) by (a) reducing the dimensionality of the parameter space of the model and (b) allowing for an arbitrary sparsity pattern. Another crucial issue in a high-dimensional setting is to deal with the* $(3k_1 + k_2)$*-dimensional error covariance matrix which we tackle by developing a pseudo-likelihood based approach for estimating* $\{W_u\}_{u=1}^{d}$ *(based only on diagonal entries/submatrices of* $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$*) first, and then obtaining a sparse estimate of the error precision matrix* $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}$*.*

**Remark S2.** *(Comparison with the state-space based approach in Schorfheide and Song (2015)) In this model, all series are assumed to evolve at a monthly level and the quarterly ones are also treated as monthly series with missing observations. It leads to introducing* $2k_2T$ *missing observations on top of the* $d(k_1 + k_2)^2$ *parameters for the transition matrix, as well as the parameters for the error covariance. The dampening effect adopted in our approach leads to* $d(k_1 + k_2)^2 + 1$ *parameters. Additionally, we impose further sparsity constraints in the model parameters through spike and slab prior distributions that lead to additional parameter reduction.*

### S2.8 Stability of VAR process

The VAR($d$) process in (2.6) of the main paper is stable if the matrix-valued polynomials $\mathcal{A}(z) = \mathbf{I}_p - \sum_{u=1}^{d} \mathbf{W}_u \boldsymbol{z}^u$ satisfies $\det(\mathcal{A}(z)) \neq 0$ on the unit circle of the complex plane $\{z \in \mathcal{C}, |z| \leq 1\}$. Note that $\mathbf{W}_u$ in (2.7) is a function of $\mathbf{A}_u$ and $\theta$. The next lemma derives the spectrum (set of all eigenvalues) of $\mathbf{W}_u$, denoted as $\mathrm{sp}(\mathbf{W}_u)$, in terms of $\mathbf{A}_u$ and $\theta$.

**Lemma 1.** *Let $\lambda$ be an eigenvalue of $\boldsymbol{W}_u$. It can be shown that $\lambda \in sp(\boldsymbol{W}_u)$ iff*

$$\lambda \in sp\left(\boldsymbol{A}_\theta^u + 2\begin{bmatrix} \boldsymbol{A}_{11}^u\theta^2 & 0 \\ \boldsymbol{A}_{21}^u & 0 \end{bmatrix}\right) \ \textit{where, } \boldsymbol{A}_\theta^u = \begin{bmatrix} \theta^2 \boldsymbol{A}^u_{11_{k_1\times k_1}} & \theta^2 \boldsymbol{A}^u_{12_{k_1\times k_2}} \\ \boldsymbol{A}^u_{21_{k_2\times k_1}} & \boldsymbol{A}^u_{22_{k_2\times k_2}} \end{bmatrix}$$

Our theoretical analysis of high-dimensional time series uses the quantity $\mu_{\max}(\mathcal{A})$ as defined below

$$\mu_{\max}(\mathcal{A}) := \max_{\theta\in[-\pi,\pi]} \lambda_{\max}\left(\left(\mathbf{I}_p - \sum_{u=1}^{d} \mathbf{W}_u{}'\mathbf{e}^{iu\theta}\right)\left(\mathbf{I}_p - \sum_{u=1}^{d} \mathbf{W}_u\mathbf{e}^{-iu\theta}\right)\right)$$

which provides a measure of stability of the process Basu and Michailidis (2015). Processes with larger $\mu_{\max}(\mathcal{A})$ will be considered less stable. On the other hand, the inter-dependence between time series in the model is captured by $\mu_{\min}(\mathcal{A})$, where

$$\mu_{\min}(\mathcal{A}) := \min_{\theta\in[-\pi,\pi]} \lambda_{\min}\left(\left(\mathbf{I}_p - \sum_{u=1}^{d} \mathbf{W}_u{}'\mathbf{e}^{iu\theta}\right)\left(\mathbf{I}_p - \sum_{u=1}^{d} \mathbf{W}_u\mathbf{e}^{-iu\theta}\right)\right).$$

The smaller the value of $\mu_{\min}(\mathcal{A})$ the more dependent the components of the time series are. Hence, for a stable VAR($d$) model we desire $\frac{\mu_{\max}(\mathcal{A})}{\mu_{\min}(\mathcal{A})}$ to be small. The quantities $\mu_{\min}(\mathcal{A})$ and $\mu_{\max}(\mathcal{A})$ provide a useful bound for the eigenvalues of $\mathbf{C}_X$ where $\mathbf{C}_X$ is as given in (S2.4). Proposition 2.3 and equation (2.6) of Basu and Michailidis (2015) gives the following chain of inequalities,

$$\frac{\lambda_{\min}(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}})}{\mu_{\max}(\mathcal{A})} \leq 2\pi\mathfrak{m}(f_X) \leq \lambda_{\min}(\mathbf{C}_X) \leq \lambda_{\max}(\mathbf{C}_X) \leq 2\pi\mathcal{M}(f_X) \leq \frac{\lambda_{\max}(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}})}{\mu_{\min}(\mathcal{A})}. \tag{S2.14}$$

### S2.9 Regularity assumptions for high-dimensional consistency

We impose the following regularity assumptions on the true model parameters.

**Assumption A1.** $\mathcal{M}_n\sqrt{\dfrac{b_n \log p}{n}} \to 0$, *where* $\mathcal{M}_n = 4\pi\lambda_{max}\left(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}\right)\dfrac{1+\mu_{\max}(\mathcal{A})}{\mu_{\min}(\mathcal{A})}$.

**Assumption A2** (Bounded eigenvalues)**.** *The minimum and the maximum eigenvalues of* $\boldsymbol{C}_X$ *and* $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon},0}$ *are both bounded above and below by constants not depending on* $n$*. Hence, there exist* $0 < \lambda_1 < \lambda_2 < \infty$ *and* $0 < \sigma_{min} < \sigma_{max} < \infty$ *not depending on* $n$ *such that* $\lambda_1 < \lambda_{min}\left(\boldsymbol{C_X}\right) < \lambda_{max}\left(\boldsymbol{C_X}\right) < \lambda_2$ *and* $\sigma_{min} < \lambda_{min}\left(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}\right) < \lambda_{max}\left(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}\right) < \sigma_{max}$.

**Assumption A3** (Rate of decay of edge probability)**.** *The edge inclusion probabilities satisfy* $q_n \leq (p^2)^{-\frac{8\mathcal{M}_n^2 b_n}{\sigma_{min}\lambda_1}}$ *for large enough* $n$.

**Assumption A4** (Minimum signal strength). $\frac{\mathcal{M}_n^2 b_n \log p + \log n}{n s_n^2} \to 0$

Assumption A1 states that $p$ can increase at an appropriate sub-exponential rate as a function of the sample size $n$. It ensures that the concentration inequalities on the quantities $\mathbf{X}'\mathbf{X}/n$ and $\mathbf{X}'\mathbf{E}/n$, as mentioned in the Supplement (Sec. S5.2), are applicable. Assumption A2 allows us to bound the minimum and maximum eigenvalues of the matrix $\mathbf{X}'\mathbf{X}/n$ away from zero and infinity, respectively, with high probability. Assumption A3 provides the rate at which the prior edge probability decreases to zero, as the dimension of the activity graph space increases. This assumption plays an important role in reducing the effective dimension of the true parameters. Assumption A4 provides a lower bound for the minimum signal strength, i.e., the smallest entry (in absolute value) of $\mathbf{A}_0$. This assumption is only needed to ensure that the posterior distribution is able to distinguish asymptotically the true model and models where the set of non-zero parameters is a strict subset of the true model.

# S3 Derivation of full conditional posterior Distributions for the Model Parameters of the Bayesian MF model

We derive the full conditional posterior distributions of all parameters of the Bayesian MF model as described in Section 3.2 of the main paper.

Let us define

$$\mathbf{S}_1 = \sum_{t=1}^{T} \mathbf{y}^{t-1}\mathbf{y}^{t-1\prime}, \mathbf{S}_2 = \sum_{t=1}^{T} \mathbf{y}^{t}\mathbf{y}^{t-1\prime}, \mathbf{S}_3 = \sum_{t=1}^{T} \mathbf{y}^{t}\mathbf{y}^{t\prime}$$

For simplicity of exposition, we will derive the conditionals for $d = 1$. For a general $d$, the derivations are very similar to what is provided below; only $\mathbf{S}_1, \mathbf{S}_2$ and $\mathbf{S}_3$ need to be computed by replacing $\boldsymbol{y}^t$ with

$$\boldsymbol{y}_u^t = \left( \boldsymbol{y}^t - \sum_{u' \neq u} \mathbf{W}_{u'} \boldsymbol{y}^{t-u'} \right)$$

while deriving conditionals for $\mathbf{A}_u$. Let us write $\mathbf{S}_1$ and $\mathbf{S}_2$ as

$$\mathbf{S}_1 = \begin{bmatrix} \mathbf{S}_{1,11_{3k_1 \times 3k_1}} & \mathbf{S}_{1,12_{3k_1 \times k_2}} \\ \mathbf{S}_{1,21_{k_2 \times 3k_1}} & \mathbf{S}_{1,22_{k_2 \times k_2}} \end{bmatrix}, \ \mathbf{S}_2 = \begin{bmatrix} \mathbf{S}_{2,11_{3k_1 \times 3k_1}} & \mathbf{S}_{2,12_{3k_1 \times k_2}} \\ \mathbf{S}_{2,21_{k_2 \times 3k_1}} & \mathbf{S}_{2,22_{k_2 \times k_2}} \end{bmatrix}$$

and $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}$ as

$$\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1} = \begin{bmatrix} {\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}}_{11_{3k_1 \times 3k_1}} & {\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}}_{12_{3k_1 \times k_2}} \\ {\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}}_{21_{k_2 \times 3k_1}} & {\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}}_{22_{k_2 \times k_2}} \end{bmatrix}$$

Let $\mathbf{M} = \begin{bmatrix} \theta^2 & \theta^3 & \theta^4 \\ \theta & \theta^2 & \theta^3 \\ 1 & \theta & \theta^2 \end{bmatrix} = ((m_{ij}))_{i,j=1,2,3}$ and,

$$\boldsymbol{u} = \left(\theta^2 \; \theta \; 1\right) = ((u_i))_{i=1,2,3}, \quad \boldsymbol{v} = \left(1 \; \theta \; \theta^2\right) = ((v_j))_{j=1,2,3}$$

Additionally we denote, $\boldsymbol{a_1} = \text{vec}(\mathbf{A}_{11})$, $\boldsymbol{a_2} = \text{vec}(\mathbf{A}_{12})$, $\boldsymbol{a_3} = \text{vec}(\mathbf{A}_{21})$ and $\boldsymbol{a_4} = \text{vec}(\mathbf{A}_{22})$.

**Full conditional posterior distribution of $\boldsymbol{a_1}$:** The conditional likelihood of $\boldsymbol{a_1}$ can be expressed as:

$$\exp\left[-\frac{1}{2}\left(\boldsymbol{a_1} - \boldsymbol{\Gamma}_{11}^{-1} d_{11}\right)' \boldsymbol{\Gamma}_{11}\left(\boldsymbol{a_1} - \boldsymbol{\Gamma}_{11}^{-1} d_{11}\right)\right] + \text{terms not involving } \boldsymbol{a_1}$$

where

$$\boldsymbol{\Gamma}_{11} = \sum_{i=1}^{3}\sum_{j=1}^{3}\left[\left(\sum_{k=1}^{3} m_{jk}\mathbf{S}_{ik}\right)' \otimes \left(\sum_{k=1}^{3} m_{ki}(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}{}_{11})_{jk}\right)\right]$$

$$\boldsymbol{d}_{11} = \text{vec}\left(\sum_{i=1}^{3}\sum_{j=1}^{3} m_{ji}(\mathbf{P}_{11})_{ji}\right)$$

$$\mathbf{P}_{11} = \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}{}_{11}\tilde{\mathbf{M}}_{11}$$

$$\tilde{\mathbf{M}}_{11} = \mathbf{S}_{2,11} - \mathbf{W}_{12}\mathbf{S}_{1,12}{}'$$

Here $\mathbf{S}_{ab}$, $(\mathbf{P}_{11})_{ab}$ and $(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}{}_{11})_{ab}$ denote $ab-$th $3 \times 3$ block of $\mathbf{S}_{1,11}$, $\mathbf{P}_{11}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}{}_{11}$ respectively, $a, b = 1, \cdots, k_1$. Combining the above likelihood of $\boldsymbol{a_1}$ with the prior specified in (3.4) of the main paper, the full conditional posterior of $a_{1i}$ (the elements of $\boldsymbol{a_1}$), $i = 1, \cdots, {k_1}^2$ turns out to be a spike and slab distribution;

$$a_{1i}|\boldsymbol{y}, -a_{1i} \sim q_{11} 1_{\{0\}} + (1 - q_{11}) N(\mu_{11}, v_{11})\ ;\ i = 1, \cdots, {k_1}^2$$

where

$$q_{11} \;=\; \frac{q}{q + (1-q)\, g\left(\boldsymbol{a_1}, \boldsymbol{\Gamma}_{11}, \boldsymbol{d}_{11}\right)}; \quad \mu_{11} = \tilde{g}\left(\boldsymbol{a_1}, \boldsymbol{\Gamma}_{11}, \boldsymbol{d}_{11}\right); \quad v_{11} = \tilde{\tilde{g}}\left(\boldsymbol{\Gamma}_{11}\right)$$

and the functions $g(.)$, $\tilde{g}(.)$ and $\tilde{\tilde{g}}(.)$ are defined as

$$\begin{aligned}
g\left(\boldsymbol{a}, \boldsymbol{\Gamma}, \boldsymbol{d}\right) &= \frac{1}{\tau\sqrt{\boldsymbol{\Gamma}_{ii} + \frac{1}{\tau^2}}} \exp\left(\frac{\left(\sum_{j \neq i} \boldsymbol{a}_j \boldsymbol{\Gamma}_{ji} - \boldsymbol{d}_i\right)^2}{2\left(\boldsymbol{\Gamma}_{ii} + \frac{1}{\tau^2}\right)}\right) \\
\tilde{g}\left(\boldsymbol{a}, \boldsymbol{\Gamma}, \boldsymbol{d}\right) &= -\frac{\left(\sum_{j \neq i} \boldsymbol{a}_j \boldsymbol{\Gamma}_{ji} - \boldsymbol{d}_i\right)}{\boldsymbol{\Gamma}_{ii} + \frac{1}{\tau^2}} \\
\tilde{\tilde{g}}\left(\boldsymbol{\Gamma}\right) &= \frac{1}{\sqrt{\boldsymbol{\Gamma}_{ii} + \frac{1}{\tau^2}}}
\end{aligned}$$

and $-a_{1i}$ denotes all the model parameters except the i-th element of $\boldsymbol{a_1}$.

**Full conditional posterior distribution of $\boldsymbol{a_2}$:** The conditional likelihood of $\boldsymbol{a_2}$ can be expressed as:

$$\exp\left[-\frac{1}{2}\left(\boldsymbol{a_2} - \boldsymbol{\Gamma}_{12}^{-1}\boldsymbol{d}_{12}\right)' \boldsymbol{\Gamma}_{12}\left(\boldsymbol{a_2} - \boldsymbol{\Gamma}_{12}^{-1}\boldsymbol{d}_{12}\right)\right] + \text{terms not involving } \boldsymbol{a_2}$$

where

$$\boldsymbol{\Gamma}_{12} = (\mathbf{S}_{1,22})' \otimes (\sum_{i=1}^{3}\sum_{j=1}^{3} u_i u_j (\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}{}_{11})_{ij})$$

$$\boldsymbol{d}_{12} = \text{vec}\left(\sum_{i=1}^{3} u_i (\mathbf{P}_{12})_i\right)$$

$$\mathbf{P}_{12} = \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}{}_{11}\tilde{\mathbf{M}}_{12}$$

$$\tilde{\mathbf{M}}_{12} = \mathbf{S}_{2,12} - \mathbf{W}_{11}\mathbf{S}_{1,12}$$

Here $(\mathbf{P}_{12})_i$ denotes $i-$th $k_1 \times k_2$ block of $\mathbf{P}_{12}, i = 1, 2, 3.$.

Combining the above likelihood of $\boldsymbol{a_2}$ with the prior specified in (3.4), the full conditional posterior of $a_{2i}$ (the elements of $\boldsymbol{a_2}$) turns out to be another spike and slab distribution;

$$a_{2i}|\boldsymbol{y}, -a_{2i} \sim q_{12} 1_{\{0\}} + (1 - q_{12}) N(\mu_{12}, v_{12});\ i = 1, \cdots, k_1 k_2$$

where

$$q_{12} : \frac{q}{q + (1-q)\, g\left(\boldsymbol{a_2}, \boldsymbol{\Gamma}_{12}, \boldsymbol{d}_{12}\right)}; \quad \mu_{12} = \tilde{g}\left(\boldsymbol{a_2}, \boldsymbol{\Gamma}_{12}, \boldsymbol{d}_{12}\right); \quad v_{12} = \tilde{\tilde{g}}\left(\boldsymbol{\Gamma}_{12}\right)$$

**Full conditional posterior distribution of $\boldsymbol{a_3}$:** The conditional likelihood of $\boldsymbol{a_3}$ can be expressed as:

$$\exp\left[-\frac{1}{2}\left(\boldsymbol{a_3} - \boldsymbol{\Gamma}_{21}^{-1}\boldsymbol{d}_{21}\right)' \boldsymbol{\Gamma}_{21}\left(\boldsymbol{a_3} - \boldsymbol{\Gamma}_{21}^{-1}\boldsymbol{d}_{21}\right)\right] + \text{terms not involving } \boldsymbol{a_3}$$

where

$$\mathbf{\Gamma}_{21} = \sum_{i=1}^{3}\sum_{j=1}^{3} v_i v_j \left(\mathbf{S}'_{ij} \otimes (\mathbf{\Sigma}_{\boldsymbol{\varepsilon}}^{-1})_{22}\right)$$

$$\boldsymbol{d}_{21} = \text{vec}\left(\sum_{i=1}^{3} v_i (\mathbf{P}_{21})_i\right)$$

$$\mathbf{P}_{21} = \mathbf{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}{}_{22}\tilde{\mathbf{M}}_{21}$$

$$\tilde{\mathbf{M}}_{21} = \mathbf{S}_{2,12}{}' - \mathbf{W}_{22}\mathbf{S}_{1,12}{}'$$

Here $(\mathbf{P}_{21})_i$ denotes i-th $k_2 \times k_1$ block of $\mathbf{P}_{21}$, i=1,2,3.

Combining the above likelihood of $\boldsymbol{a_3}$ with the prior specified in (3.4) the full conditional posterior of $a_{3i}$ (the elements of $\boldsymbol{a_3}$) turns out to be another spike and slab distribution;

$$a_{3i}|\boldsymbol{y}, -a_{3i} \sim q_{21}1_{\{0\}} + (1 - q_{21})N(\mu_{21}, v_{21});\ i = 1, \cdots, k_1 k_2$$

where

$$q_{21} : \frac{q}{q + (1-q)\, g\left(\boldsymbol{a_3}, \mathbf{\Gamma}_{21}, \boldsymbol{d}_{21}\right)}; \quad \mu_{21} = \tilde{g}\left(\boldsymbol{a_3}, \mathbf{\Gamma}_{21}, \boldsymbol{d}_{21}\right); \quad v_{21} = \tilde{\tilde{g}}\left(\mathbf{\Gamma}_{21}\right)$$

**Full conditional posterior distribution of $\boldsymbol{a_4}$:** The conditional likelihood of $\boldsymbol{a_4}$ can be expressed as:

$$\exp\left[-\frac{1}{2}\left(\boldsymbol{a_4} - \mathbf{\Gamma}_{22}^{-1}\boldsymbol{d}_{22}\right)'\mathbf{\Gamma}_{22}\left(\boldsymbol{a_4} - \mathbf{\Gamma}_{22}^{-1}\boldsymbol{d}_{22}\right)\right] + \text{terms not involving } \boldsymbol{a_4}$$

where,

$$\boldsymbol{\Gamma}_{22} = (\mathbf{S}_{1,22})' \otimes (\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}{}_{22})$$

$$\boldsymbol{d}_{22} = \text{vec}\left((\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}{}_{22})\,(\mathbf{S}_{2,22} - \mathbf{W}_{21}\mathbf{S}_{1,12})\right)$$

Combining the above likelihood of $\boldsymbol{a_4}$ with the prior specified in (3.4) the full conditional posterior of $a_{4i}$ (the elements of $\boldsymbol{a_4}$) turns out to be another spike and slab distribution;

$$a_{4i}|\boldsymbol{y}, -a_{4i} \sim q_{22}1_{\{0\}} + (1-q_{22})N(\mu_{22}, v_{22}); i = 1, \cdots, {k_2}^2$$

where

$$q_{22} : \frac{q}{q + (1-q)\, g\,(\boldsymbol{a_4}, \boldsymbol{\Gamma}_{22}, \boldsymbol{d}_{22})}; \quad \mu_{22} = \tilde{g}\,(\boldsymbol{a_4}, \boldsymbol{\Gamma}_{22}, \boldsymbol{d}_{22}))\,; \quad v_{22} = \tilde{\tilde{g}}\,(\boldsymbol{\Gamma}_{22}))$$

**Full conditional posterior distribution of $\theta$:** The conditional likelihood of $\theta$ can be expressed as,

$$\begin{aligned} &exp\left[-\frac{1}{2}tr\left(\mathbf{W}'\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{W}\mathbf{S}_1 - 2\mathbf{W}'\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{S}_2\right)\right] \\ &\propto exp\left(-\frac{1}{2}\sum_{j=-4}^{4} d_j\theta^j\right), \text{ for some } \boldsymbol{d} \text{ that needs to be computed.} \end{aligned}$$

Under the prior of $\theta$ specified in Section 3.1 of main paper, the full conditional posterior distribution of $\theta$ is then proportional to

$$exp\left(-\frac{1}{2}\sum_{j=-4}^{4} d_j\theta^j\right)$$

We estimate $\boldsymbol{d}$ using ordinary least squares where the response $\boldsymbol{y}$ and the design matrix $\mathbf{X}$ are as follows:

$$\boldsymbol{y} = f(\theta) = \exp\left[-\frac{1}{2}tr\left(\mathbf{W}'\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{W}\mathbf{S}_1 - 2\mathbf{W}'\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}^{-1}\mathbf{S}_2\right)\right] \text{ where } \mathbf{W} \text{ involves } \theta.$$

$$\mathbf{X} = ((\theta^j)), j = -4, -3, \cdots, 3, 4$$

The values of $\boldsymbol{y}$ and $\mathbf{X}$ are computed for, say, 20 values of $\theta$ and $\boldsymbol{d}$ is obtained as,

$$\hat{\boldsymbol{d}} = \underset{d \in \mathcal{R}^9}{\operatorname{argmin}} \sum_{i=1}^{20} \left( f(\theta_i) - \sum_{j=-4}^{4} d_j {\theta_i}^j \right)$$

Once $\boldsymbol{d}$ is estimated, finally we use discretization method to sample $\theta$ from its posterior.

**Full conditional posterior distribution of $\boldsymbol{\Sigma}_{i,\mathbf{H}}, i = 1, \cdots, k_1$:** The conditional likelihood of $\boldsymbol{\Omega}_{i,\mathrm{H}} = {\boldsymbol{\Sigma}_{i,\mathrm{H}}}^{-1}$ is given by,

$$Q\left(\boldsymbol{\Omega}_{i,\mathrm{H}} \mid \boldsymbol{y}, \mathbf{A}, \mathcal{G}\right) \propto \exp\left(-\frac{tr\left(\boldsymbol{\Omega}_{i,\mathrm{H}}\mathbf{L}_i\right)}{2}\right) |\boldsymbol{\Omega}_{i,\mathrm{H}}|^{\frac{n}{2}}$$

where, $\mathbf{L}_i = \left(\boldsymbol{y}_{i,\mathrm{H}}^t - \boldsymbol{\delta}\boldsymbol{z}_{\mathcal{G},i}^{t-1}\mathbf{A}'_{\mathcal{G},i\cdot}\right)\left(\boldsymbol{y}_{i,\mathrm{H}}^t - \boldsymbol{\delta}\boldsymbol{z}_{\mathcal{G},i}^{t-1}\mathbf{A}'_{\mathcal{G},i\cdot}\right)'$

Let us consider a transformation:

$$\boldsymbol{\Omega}_{i,\mathrm{H}} \to \mathbf{F}_i = \mathbf{C}'\boldsymbol{\Omega}_{i,\mathrm{H}}\mathbf{C} \tag{S3.1}$$

where $\mathbf{C}$ is as defined in Section 3.1 of the main paper. Thus,

$$Q\left(\mathbf{F}_i \mid \boldsymbol{y}, \mathbf{A}, \mathcal{G}\right) \propto \exp\left(-\frac{tr\left(\mathbf{F}_i\tilde{\mathbf{L}}_i\right)}{2}\right) |\mathbf{F}_i|^{\frac{n}{2}}$$

where, $\tilde{\mathbf{L}}_i = \mathbf{C}^{-1}\mathbf{L}_i(\mathbf{C}')^{-1}$. The prior on $\boldsymbol{\Sigma}_{i,\mathrm{H}}$ as given in Section 3.1 induces the following prior on $\mathbf{F}_i$

$$\mathbf{F}_i \sim \text{ Wishart } \left(\mathbf{C}'\mathbf{V}^{-1}\mathbf{C}, \nu\right)$$

Also,

$$\Pi(\mathbf{A}_{\mathcal{G},i\cdot}|\mathbf{F}_i) \propto \left(\tilde{\sigma}_i^2\right)^{-\frac{\nu_i(\mathcal{G})}{2}} \exp\left(-\frac{\mathbf{A}_{\mathcal{G},i\cdot}\mathbf{A}'_{\mathcal{G},i\cdot}}{2\tau^2\tilde{\sigma}_i^2}\right)$$

where $\tilde{\sigma}_i^2$ is as defined in section 3.1. Then the full conditional posterior of $\mathbf{F}_i$ can be written as,

$$\Pi\left(\mathbf{F}_i \mid \boldsymbol{y}, \mathbf{A}, \mathcal{G}\right) \propto Q\left(\mathbf{F}_i \mid \boldsymbol{y}, \mathbf{A}, \mathcal{G}\right)\ \Pi(\mathbf{A}_{\mathcal{G},i\cdot}|\mathbf{F}_i)\, \Pi(\mathbf{F}_i)$$

which does not have a standard form. Consider the following block Cholesky decomposition of $\mathbf{F}_i$:

$$\mathbf{F}_i = \begin{bmatrix} \mathbf{F}_{i,11} & \mathbf{F}_{i,12} \\ \mathbf{F}_{i,21} & \mathbf{F}_{i,22} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{2\times 2} & \frac{\mathbf{F}_{i,12}}{\mathbf{F}_{i,22}} \\ \mathbf{O} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{F}_{i,11} - \mathbf{F}_{i,12}\mathbf{F}_{i,22}^{-1}\mathbf{F}_{i,21} & \mathbf{O} \\ \mathbf{O} & \mathbf{F}_{i,22} \end{bmatrix} \begin{bmatrix} \mathbf{I}_{2\times 2} & \mathbf{O} \\ \frac{\mathbf{F}_{i,21}}{\mathbf{F}_{i,22}} & 1 \end{bmatrix} \tag{S3.2}$$

Denote, $\mathbf{e} = \frac{\mathbf{F}_{i,12}}{\mathbf{F}_{i,22}}, f = \mathbf{F}_{i,22}, \mathbf{B}_1 = \mathbf{F}_{i,11} - \mathbf{F}_{i,12}\mathbf{F}_{i,22}^{-1}\mathbf{F}_{i,21}$

Then, $\mathbf{F}_i$ takes the form,

$$\mathbf{F}_i = \begin{bmatrix} \mathbf{B}_1 + \mathbf{e}\mathbf{e}'f & \mathbf{e}f \\ \mathbf{e}'f & f \end{bmatrix} \tag{S3.3}$$

Using the transformation : $\mathbf{F}_i \to (\mathbf{B}_1, \mathbf{e}, f)$, the conditional posterior of

$(\mathbf{B}_1, \mathbf{e}, f)$ is given by,

$$\begin{aligned}
& \Pi\left(\mathbf{B}_1, \mathbf{e}, f \mid \boldsymbol{y}, \mathbf{A}, \mathcal{G}\right) \\
= \; & \exp\left[-\frac{1}{2} tr\left(\left(\mathbf{B}_1 + \mathbf{e}\mathbf{e}'f\right)\mathbf{D}_{11} + \mathbf{e}f\mathbf{D}_{21} + \mathbf{e}'f\mathbf{D}_{12} + f\mathbf{D}_{22}\right)\right] |\mathbf{B}_1|^{\frac{n+\nu-4}{2}} \\
& \qquad \times f^{\frac{n+\nu+\nu_i(\mathcal{G})}{2}} \exp\left(-\frac{\mathbf{A}_{\mathcal{G},i\cdot}\mathbf{A}'_{\mathcal{G},i\cdot}f}{2\tau^2}\right)
\end{aligned}$$

where, $\mathbf{D} = \mathbf{C}^{-1}\left(\mathbf{V} + \mathbf{L}_i\right)(\mathbf{C}')^{-1}$. Then by straightforward calculation, it can be shown that,

$$\begin{aligned}
\Pi\left(\mathbf{B}_1 \mid \boldsymbol{y}, \mathbf{A}, \mathcal{G}\right) &\sim \text{Wishart}\left(\mathbf{D}_{11}, n + \nu - 1\right) \\
\Pi\left(\mathbf{e} \mid \boldsymbol{y}, f, \mathbf{A}, \mathcal{G}\right) &\sim \mathcal{N}\left(-\mathbf{D}_{11}^{-1}\mathbf{D}_{12}, \frac{1}{f}\mathbf{D}_{11}^{-1}\right) \\
\Pi\left(f \mid \boldsymbol{y}, \mathbf{A}, \mathcal{G}\right) &\sim \text{Gamma}\left(\frac{n + \nu + \nu_i(\mathcal{G})}{2}, tr\left(\frac{\mathbf{A}_{\mathcal{G},i\cdot}\mathbf{A}'_{\mathcal{G},i\cdot}}{\tau^2} + \mathbf{D}_{22} - \mathbf{D}'_{12}\mathbf{D}_{11}^{-1}\mathbf{D}_{12}\right)\right)
\end{aligned}$$

and thus $\boldsymbol{\Omega}_{i,\mathrm{H}}$ (and hence $\boldsymbol{\Sigma}_{i,\mathrm{H}}$) can be estimated once $\mathbf{B}_1, \mathbf{e}$ and $f$ are sampled from their respective full conditional posterior distributions.

**Full conditional posterior distribution of $\sigma_j^2, j = 1 \cdots k_2$:** The conditional likelihood of $\sigma_j^2$ is given by,

$$Q(\sigma_j^2 \mid \boldsymbol{y}, \mathbf{A}, \mathcal{G}) \propto \exp\left[-\frac{\left(y_{j,\mathrm{L}}^t - \mathbf{A}_{\mathcal{G},j\cdot}\boldsymbol{z}_{\mathcal{G},j}^{t-1}\right)^2}{2\sigma_j^2}\right]\left(\frac{1}{\sigma_j^2}\right)^{\frac{n}{2}}$$

Also,

$$\Pi(\mathbf{A}_{\mathcal{G},j\cdot}|\sigma_j^2) \propto \left(\sigma_j^2\right)^{-\frac{\nu_{k_1+j}(\mathcal{G})}{2}} \exp\left(-\frac{\mathbf{A}_{\mathcal{G},j\cdot}\mathbf{A}'_{\mathcal{G},j\cdot}}{2\tau^2\sigma_j^2}\right)$$

and, $\sigma_j^2 \sim \mathcal{IG}\left(\alpha, \beta\right)$. Then by straightforward calculation, full conditional posterior of $\sigma_j^2$ can be shown as,

$$\sigma_j^2 | \mathbf{A}, \boldsymbol{y}, \mathcal{G} \sim \mathcal{IG}\left( \left( \frac{n}{2} + \alpha + \frac{\nu_{k_1+j}(\mathcal{G})}{2} - 1 \right), \left( \frac{\left(y_{j,\mathrm{L}}^t - \mathbf{A}_{\mathcal{G},j\cdot} \boldsymbol{z}_{\mathcal{G},j}^{t-1}\right)^2}{2} + \beta + \frac{\mathbf{A}_{\mathcal{G},j\cdot} \mathbf{A}'_{\mathcal{G},j\cdot}}{2\tau^2} \right) \right)$$

## S4 Derivation of $\Pi(\mathcal{G} \mid \mathcal{Y})$ for the Bayesian MF model

Given the graph $\mathcal{G}$ let us denote the $i-$th row of $\mathbf{A}$ with active coefficients by $\mathbf{A}_{\mathcal{G},i}$. Let $\mathbf{A}^c$ denote the inactive coefficients of $\mathbf{A}$. We have,

$$\pi\left(\mathbf{A}_{\mathcal{G},i\cdot} \mid \mathcal{G}, \boldsymbol{\Omega}_i^{\mathrm{H}}\right) = \left( \frac{(\boldsymbol{\delta}' \boldsymbol{\Omega}_i^{\mathrm{H}} \boldsymbol{\delta})}{2\pi\tau^2} \right)^{\frac{\nu_i(\mathcal{G})}{2}} \exp\left( -\frac{1}{2\tau^2} (\boldsymbol{\delta}' \boldsymbol{\Omega}_i^{\mathrm{H}} \boldsymbol{\delta}) \mathbf{A}_{\mathcal{G},i\cdot} \mathbf{A}'_{\mathcal{G},i\cdot} \right) \text{ for } i = 1, \cdots, k_1$$

$$\pi\left(\mathbf{A}_{\mathcal{G},j\cdot} \mid \mathcal{G}, \sigma_j^2\right) = \frac{1}{\left(2\pi\tau^2\sigma_j^2\right)^{\frac{\nu_{k_1+j}(\mathcal{G})}{2}}} \exp\left( -\frac{1}{2} \frac{\mathbf{A}_{\mathcal{G},j\cdot} \mathbf{A}'_{\mathcal{G},j\cdot}}{\tau^2 \sigma_j^2} \right) \text{ for } j = 1, \cdots, k_2$$

$$\pi\left(\mathbf{A}^C{}_{ij} \mid \mathcal{G}, \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}\right) = 1 \quad \text{for all } 1 \le i \le (k_1 + k_2), 1 \le j \le (k_1 + k_2) \text{ with } \gamma_{ij} = 0$$

$$\pi(\mathcal{G}) = \prod_{i=1}^{k_1} \left\{ q^{\nu_i(\mathcal{G})} (1-q)^{(k_1+k_2) - \nu_i(\mathcal{G})} \right\} \prod_{j=1}^{k_2} \left\{ q^{\nu_{k_1+j}(\mathcal{G})} (1-q)^{(k_1+k_2) - \nu_{k_1+j}(\mathcal{G})} \right\}$$

Note that

$$\begin{aligned}
& L_{\text{pseudo}}\left(\mathbf{A}, \{\boldsymbol{\Omega}_{i,\mathrm{H}}\}_{i=1}^{k_1}, \{\sigma_j^2\}_{j=1}^{k_2} \mid \boldsymbol{y}^t, t = 1, \cdots, T\right) \\
\propto \; & \prod_{i=1}^{k_1} \left\{ \exp\left[ -\frac{1}{2} \sum_{t=1}^{T} \left( \boldsymbol{y}_{i,\mathrm{H}}^t - \left(\mathbf{W}\boldsymbol{y}^{t-1}\right)_{i,\mathrm{H}} \right)' \boldsymbol{\Omega}_{i,\mathrm{H}} \left( \boldsymbol{y}_{i,\mathrm{H}}^t - \left(\mathbf{W}\boldsymbol{y}^{t-1}\right)_{i,\mathrm{H}} \right) \right] |\boldsymbol{\Omega}_{i,\mathrm{H}}|^{n/2} \right\} \times \\
& \prod_{j=1}^{k_2} \left\{ \exp\left[ -\frac{1}{2} \sum_{t=1}^{T} \frac{\left( y_{j,\mathrm{L}}^t - (\mathbf{W}\boldsymbol{y}^{t-1})_{3k_1+j,\mathrm{L}} \right)^2}{\sigma_j^2} \right] \left( \frac{1}{\sigma_j^2} \right)^{n/2} \right\}
\end{aligned}$$

Then,

$$
\begin{aligned}
& \pi_{\text{pseudo}}\big(\mathcal{G} \mid \mathcal{Y}\big) \\
\propto\ & \pi(\mathcal{G}) \int\int L_{\text{pseudo}}\Big(\mathbf{A}, \{\boldsymbol{\Omega}_{i,\mathrm{H}}\}_{i=1}^{k_1}, \big\{\sigma_j^2\big\}_{j=1}^{k_2} \mid \boldsymbol{y}^t, t=1,\cdots,T\Big)\pi(\mathbf{A}|\mathcal{G},\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}})\pi(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}})\mathbf{dA}\ \mathbf{d}\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} \\
=\ & (const.)\Bigg\{\prod_{i=1}^{k_1}\Big(q^{\nu_i(\mathcal{G})}\,(1-q)^{(k_1+k_2)-\nu_i(\mathcal{G})}\Big) \times \\
& \int\int \prod_{i=1}^{k_1} |\boldsymbol{\Omega}_{i,\mathrm{H}}|^{n/2} \exp\left[-\frac{1}{2}\sum_{i=1}^{k_1}\sum_{t=1}^{T}\Big(\boldsymbol{y}_{i,\mathrm{H}}^t - \big(\mathbf{W}\boldsymbol{y}^{t-1}\big)_{i,\mathrm{H}}\Big)' \boldsymbol{\Omega}_{i,\mathrm{H}} \Big(\boldsymbol{y}_{i,\mathrm{H}}^t - \big(\mathbf{W}\boldsymbol{y}^{t-1}\big)_{i,\mathrm{H}}\Big)\right] \times \\
& \pi\left(\mathbf{A}_{\mathcal{G},i\cdot} \mid \mathcal{G}, \boldsymbol{\Omega}_{i,\mathrm{H}}\right) d\mathbf{A} \prod_{i=1}^{k_1} \pi\left(\boldsymbol{\Omega}_{i,\mathrm{H}}\right) d\boldsymbol{\Omega}_{i,\mathrm{H}}\Bigg\} \times \\
& \Bigg\{\prod_{j=1}^{k_2}\Big(q^{\nu_{k_1+j}(\mathcal{G})}\,(1-q)^{(k_1+k_2)-\nu_{k_1+j}(\mathcal{G})}\Big) \int\int \frac{1}{\prod_{j=1}^{k_2}\big(\sigma_j^2\big)^{n/2}} \times \\
& \exp\left[-\frac{1}{2}\sum_{j=1}^{k_2}\sum_{t=1}^{T}\frac{\Big(y_{j,\mathrm{L}}^t - (\mathbf{W}\boldsymbol{y}^{t-1})_{3k_1+j,\mathrm{L}}\Big)^2}{\sigma_j^2}\right] \pi\left(\mathbf{A}_{\mathcal{G},j\cdot} \mid \mathcal{G}, \sigma_j^2\right) d(\mathbf{A}) \prod_{j=1}^{k_2} \pi\left(\sigma_j^2\right) d\sigma_j^2\Bigg\}
\end{aligned}
\tag{S4.1}
$$

We will first simplify the second product of $k_2$ terms in (S4.1).

$$\begin{aligned}
&\prod_{j=1}^{k_2}\left(q^{\nu_{k_1+j}(\mathcal{G})}\,(1-q)^{(k_1+k_2)-\nu_{k_1+j}(\mathcal{G})}\right)\times\\
&\int\int\frac{1}{\prod_{j=1}^{k_2}\left(\sigma_j^2\right)^{n/2}}\exp\left[-\frac{1}{2}\sum_{j=1}^{k_2}\sum_{t=1}^{T}\frac{\left(y_{j,\mathrm{L}}^t-(\mathbf{W}\boldsymbol{y}^{t-1})_{3k_1+j,\mathrm{L}}\right)^2}{\sigma_j^2}\right]\times\\
&\qquad\qquad\qquad\pi\left(\mathbf{A}_{\mathcal{G},j\cdot}\mid\mathcal{G},\sigma_j^2\right)d(\mathbf{A})\prod_{j=1}^{k_2}\pi\left(\sigma_j^2\right)d\sigma_j^2\\
=\;&\prod_{j=1}^{k_2}\left(q^{\nu_{k_1+j}(\mathcal{G})}\,(1-q)^{(k_1+k_2)-\nu_{k_1+j}(\mathcal{G})}\right)\times\\
&\int\int\frac{1}{\prod_{j=1}^{k_2}\left(\sigma_j^2\right)^{n/2}}\exp\left[-\frac{1}{2}\sum_{j=1}^{k_2}\sum_{t=1}^{T}\frac{\left(y_{j,\mathrm{L}}^t-\boldsymbol{z}_{\mathcal{G},j}^{t-1\prime}\mathbf{A}_{\mathcal{G},j\cdot}'\right)^2}{\sigma_j^2}\right]\times\\
&\frac{1}{\prod_{j=1}^{k_2}\left(2\pi\tau^2\sigma_j^2\right)^{\frac{\nu_{k_1+j}(\mathcal{G})}{2}}}\exp\left(-\frac{1}{2}\sum_{j=1}^{k_2}\frac{\mathbf{A}_{\mathcal{G},j\cdot}\mathbf{A}_{\mathcal{G},j\cdot}'}{\tau^2\sigma_j^2}\right)\prod_{j=1}^{k_2}\pi\left(\sigma_j^2\right)d\sigma_j^2
\end{aligned}\tag{S4.2}$$

where $\boldsymbol{z}_{\mathcal{G},j}^{t-1\prime}$ is the $t^{th}$ row of $\mathbf{Z}_{\mathcal{G},j}$. Since $\prod_{j=1}^{k_2}\left(q^{\nu_{k_1+j}(\mathcal{G})}\,(1-q)^{(k_1+k_2)-\nu_{k_1+j}(\mathcal{G})}\right)$ doesn't depend on $\mathbf{A}$ or $\sigma_j^2$, for the ease of writing we denote it by $\mathcal{P}_2$. Now collecting all the terms involving $\mathbf{A}$, (S4.2) can be written as

$$\begin{aligned}
&\mathcal{P}_2\prod_{j=1}^{k_2}\int\frac{1}{\left(\sigma_j^2\right)^{n/2}\left(2\pi\tau^2\sigma_j^2\right)^{\frac{\nu_{k_1+j}(\mathcal{G})}{2}}}\exp\left[-\frac{1}{2}\frac{\sum_{t=1}^{T}\left(y_{j,\mathrm{L}}^t\right)^2}{\sigma_j^2}\right]\times\\
&\left\{\int\exp\left[-\frac{1}{2\sigma_j^2}\left\{\mathbf{A}_{\mathcal{G},j\cdot}\left(\sum_{t=1}^{T}\boldsymbol{z}_{\mathcal{G},j}^{t-1}\boldsymbol{z}_{\mathcal{G},j}^{t-1\prime}+\frac{\mathbf{I}}{\tau^2}\right)\mathbf{A}_{\mathcal{G},j\cdot}'+2\mathbf{A}_{\mathcal{G},j\cdot}\left(\sum_{t=1}^{T}y_{j,\mathrm{L}}^t\boldsymbol{z}_{\mathcal{G},j}^{t-1}\right)\right\}\right]d\mathbf{A}\right\}\times\\
&\qquad\qquad\qquad\pi\left(\sigma_j^2\right)d\sigma_j^2
\end{aligned}\tag{S4.3}$$

Now,

$$\int \exp\left[-\frac{1}{2\sigma_j^2}\left\{\mathbf{A}_{\mathcal{G},j\cdot}\left(\sum_{t=1}^{T}\boldsymbol{z}_{\mathcal{G},j}^{t-1}\boldsymbol{z}_{\mathcal{G},j}^{t-1\prime}+\frac{\mathbf{I}}{\tau^2}\right)\mathbf{A}_{\mathcal{G},j\cdot}^{\prime}+2\mathbf{A}_{\mathcal{G},j\cdot}\left(\sum_{t=1}^{T}y_{j,\mathrm{L}}^{t}\boldsymbol{z}_{\mathcal{G},j}^{t-1}\right)\right\}\right]d\mathbf{A}$$

$$= \left(2\pi\sigma_j^2\right)^{\frac{\nu_{k_1+j}(\mathcal{G})}{2}}\left|\sum_{t=1}^{T}\boldsymbol{z}_{\mathcal{G},j}^{t-1\prime}\boldsymbol{z}_{\mathcal{G},j}^{t-1}+\frac{\mathbf{I}}{\tau^2}\right|^{-1/2}\times$$

$$\exp\left[\frac{\left(\sum_{t=1}^{T}y_{j,\mathrm{L}}^{t}\boldsymbol{z}_{\mathcal{G},j}^{t-1}\right)^{\prime}\left(\sum_{t=1}^{T}\boldsymbol{z}_{\mathcal{G},j}^{t-1}\boldsymbol{z}_{\mathcal{G},j}^{t-1\prime}+\frac{\mathbf{I}}{\tau^2}\right)^{-1}\left(\sum_{t=1}^{T}y_{j,\mathrm{L}}^{t}\boldsymbol{z}_{\mathcal{G},j}^{t-1}\right)}{2\sigma_j^2}\right]$$

Then, (S4.3) becomes

$$\mathcal{P}_2\prod_{j=1}^{k_2}\int\left\{\frac{1}{\left(\tau^2\right)^{\frac{\nu_{k_1+j}(\mathcal{G})}{2}}}\left|\sum_{t=1}^{T}\boldsymbol{z}_{\mathcal{G},j}^{t-1}\boldsymbol{z}_{\mathcal{G},j}^{t-1\prime}+\frac{\mathbf{I}}{\tau^2}\right|^{-1/2}\times\right.$$

$$\exp\left[-\frac{\sum_{t=1}^{T}\left(y_{j,\mathrm{L}}^{t}\right)^2-\left(\sum_{t=1}^{T}y_{j,\mathrm{L}}^{t}\boldsymbol{z}_{\mathcal{G},j}^{t-1}\right)^{\prime}\left(\sum_{t=1}^{T}\boldsymbol{z}_{\mathcal{G},j}^{t-1}\boldsymbol{z}_{\mathcal{G},j}^{t-1\prime}+\frac{\mathbf{I}}{\tau^2}\right)^{-1}\left(\sum_{t=1}^{T}y_{j,\mathrm{L}}^{t}\boldsymbol{z}_{\mathcal{G},j}^{t-1}\right)}{2\sigma_j^2}\right]\times$$

$$\left.\frac{1}{\left(\sigma_j^2\right)^{n/2}}\pi\left(\sigma_j^2\right)d\sigma_j^2\right\}$$

$$= \prod_{j=1}^{k_2}\int\left\{q^{\nu_{k_1+j}(\mathcal{G})}\left(1-q\right)^{(k_1+k_2)-\nu_{k_1+j}(\mathcal{G})}\frac{1}{\left(\tau\sqrt{n}\right)^{\nu_{k_1+j}(\mathcal{G})}}\left|\frac{\mathbf{Z}_{\mathcal{G},j}^{\prime}\mathbf{Z}_{\mathcal{G},j}}{n}+\frac{\mathbf{I}}{n\tau^2}\right|^{-1/2}\right.$$

$$\left.\exp\left(-\frac{\boldsymbol{y}_{j,\mathrm{L}}^{\prime}\boldsymbol{y}_{j,\mathrm{L}}-\boldsymbol{y}_{j,\mathrm{L}}^{\prime}\mathbf{Z}_{\mathcal{G},j}\left(\mathbf{Z}_{\mathcal{G},j}^{\prime}\mathbf{Z}_{\mathcal{G},j}+\frac{\mathbf{I}}{\tau^2}\right)^{-1}\mathbf{Z}_{\mathcal{G},j}^{\prime}\boldsymbol{y}_{j,\mathrm{L}}}{2\sigma_j^2}\right)\frac{1}{\left(\sigma_j^2\right)^{n/2}}\pi\left(\sigma_j^2\right)d\sigma_j^2\right\}$$

$$= \prod_{j=1}^{k_2}q^{\nu_{k_1+j}(\mathcal{G})}\left(1-q\right)^{(k_1+k_2)-\nu_{k_1+j}(\mathcal{G})}\left(\frac{1}{\tau\sqrt{n}}\right)^{\nu_{k_1+j}(\mathcal{G})}\left|\frac{\mathbf{Z}_{\mathcal{G},j}^{\prime}\mathbf{Z}_{\mathcal{G},j}}{n}+\frac{\mathbf{I}}{n\tau^2}\right|^{-1/2}\left(R_j\right)^{-\left(\frac{n}{2}+\alpha\right)} \tag{S4.4}$$

where $R_j = \frac{\boldsymbol{y}_{j,\mathrm{L}}^{\prime}\boldsymbol{y}_{j,\mathrm{L}}}{n} - \frac{\boldsymbol{y}_{j,\mathrm{L}}^{\prime}\mathbf{Z}_{\mathcal{G},j}}{n}\left(\frac{\mathbf{Z}_{\mathcal{G},j}^{\prime}\mathbf{Z}_{\mathcal{G},j}}{n}+\frac{\mathbf{I}}{n\tau^2}\right)^{-1}\frac{\mathbf{Z}_{\mathcal{G},j}^{\prime}\boldsymbol{y}_{j,\mathrm{L}}}{n}+\frac{\beta}{n}$

The last equality follows by writing the density of $\sigma_j^2 - \pi\left(\sigma_j^2\right) \propto \left(\sigma_j^2\right)^{-(\alpha+1)}\exp\left(-\frac{\beta}{2\sigma_j^2}\right)$ and integrating the above with respect to $\sigma_j^2$'s. Now we will simplify the first product of $k_1$ terms in (S4.1). Note that

$$\begin{aligned}
& \exp\left[-\frac{1}{2}\sum_{t=1}^{T}\left(\boldsymbol{y}_{i,\mathrm{H}}^{t} - \left(\mathbf{W}\boldsymbol{y}^{t-1}\right)_{i,\mathrm{H}}\right)'\boldsymbol{\Omega}_{i,\mathrm{H}}\left(\boldsymbol{y}_{i,\mathrm{H}}^{t} - \left(\mathbf{W}\boldsymbol{y}^{t-1}\right)_{i,\mathrm{H}}\right)\right] \\
= \;& \exp\left[-\frac{1}{2}\sum_{t=1}^{T} tr\left(\boldsymbol{y}_{i,\mathrm{H}}^{t} - \boldsymbol{\delta}\boldsymbol{z}_{\mathcal{G},i}^{t-1\prime}\mathbf{A}'_{\mathcal{G},i\cdot}\right)'\boldsymbol{\Omega}_{i,\mathrm{H}}\left(\boldsymbol{y}_{i,\mathrm{H}}^{t} - \boldsymbol{\delta}\boldsymbol{z}_{\mathcal{G},i}^{t-1\prime}\mathbf{A}'_{\mathcal{G},i\cdot}\right)\right] \\
= \;& \exp\left[-\frac{1}{2}\sum_{t=1}^{T} tr\left(\boldsymbol{y}_{i,\mathrm{H}}^{t}{}'\boldsymbol{\Omega}_{i,\mathrm{H}}\boldsymbol{y}_{i,\mathrm{H}}^{t} - 2\boldsymbol{y}_{i,\mathrm{H}}^{t}{}'\boldsymbol{\Omega}_{i,\mathrm{H}}\boldsymbol{\delta}\boldsymbol{z}_{\mathcal{G},i}^{t-1\prime}\mathbf{A}'_{\mathcal{G},i\cdot} + \mathbf{A}_{\mathcal{G},i\cdot}\boldsymbol{z}_{\mathcal{G},i}^{t-1}\boldsymbol{\delta}'\boldsymbol{\Omega}_{i,\mathrm{H}}\boldsymbol{\delta}\boldsymbol{z}_{\mathcal{G},i}^{t-1\prime}\mathbf{A}'_{\mathcal{G},i\cdot}\right)\right] \\
\overset{(i)}{=} \;& \exp\left[-\frac{1}{2}\sum_{t=1}^{T} tr\left(\tilde{\mathbf{Y}}'_{t\cdot}\tilde{\mathbf{Y}}_{t\cdot}\mathbf{F}_i - 2\mathbf{A}'_{\mathcal{G},i\cdot}\boldsymbol{z}_{\mathcal{G},i}^{t-1\prime}\tilde{\mathbf{Y}}_{t\cdot}\mathbf{F}_{i\cdot 3} + \mathbf{F}_{i33}\mathbf{A}_{\mathcal{G},i\cdot}\boldsymbol{z}_{\mathcal{G},i}^{t-1\prime}\boldsymbol{z}_{\mathcal{G},i}^{t-1}\mathbf{A}'_{\mathcal{G},i\cdot}\right)\right] \qquad \text{(S4.5)}
\end{aligned}$$

where $\tilde{\mathbf{Y}}_{t\cdot}$ is the $t^{th}$ row of $\tilde{\mathbf{Y}}_{i,\mathrm{H}}$ (as defined in Section S2.5) and (i) follows by applying the transformation: $\boldsymbol{\Omega}_{i,\mathrm{H}} \to \mathbf{F}_i = \mathbf{C}'\boldsymbol{\Omega}_{i,\mathrm{H}}\mathbf{C}$ where $\mathbf{C} = [\boldsymbol{a} \quad \boldsymbol{b} \quad \boldsymbol{\delta}]$ is as defined in Section 3.1 of the main paper. Let us denote $\prod_{i=1}^{k_1}\left(q^{\nu_i(\mathcal{G})}(1-q)^{(k_1+k_2)-\nu_i(\mathcal{G})}\right)$ by $\mathcal{P}_1$. Also, as $\boldsymbol{\Sigma}_{i,\mathrm{H}} \overset{iid}{\sim}$ Inverse-Wishart $(\omega, \mathbf{V})$, for $i = 1, 2, \cdots, k_1$, the prior on $\mathbf{F}_i$ is given by,

$$\mathbf{F}_i \overset{iid}{\sim} \text{Wishart}(\omega, \mathbf{C}'\mathbf{V}^{-1}\mathbf{C}) \text{ for } i = 1, 2, \cdots, k_1$$

Then,

$$
\begin{aligned}
& \prod_{i=1}^{k_1} \left( q^{\nu_i(\mathcal{G})} (1-q)^{(k_1+k_2)-\nu_i(\mathcal{G})} \right) \times \\
& \int\int \prod_{i=1}^{k_1} |\boldsymbol{\Omega}_{i,\mathrm{H}}|^{n/2} \exp\left[ -\frac{1}{2} \sum_{i=1}^{k_1} \sum_{t=1}^{T} \left( \boldsymbol{y}_{i,\mathrm{H}}^t - \left(\mathbf{W}\boldsymbol{y}^{t-1}\right)_{i,\mathrm{H}} \right)' \boldsymbol{\Omega}_{i,\mathrm{H}} \left( \boldsymbol{y}_{i,\mathrm{H}}^t - \left(\mathbf{W}\boldsymbol{y}^{t-1}\right)_{i,\mathrm{H}} \right) \right] \times \\
& \qquad\qquad\qquad\qquad \pi\left(\mathbf{A}_{\mathcal{G},i\cdot} \mid \mathcal{G}, \boldsymbol{\Omega}_{\mathrm{H}}\right) d\mathbf{A} \prod_{i=1}^{k_1} \pi\left(\boldsymbol{\Omega}_{i,\mathrm{H}}\right) d\boldsymbol{\Omega}_{i,\mathrm{H}} \\
= & \ \mathcal{P}_1 \prod_{i=1}^{k_1} \int\int \Bigg\{ |\mathbf{F}_i|^{\frac{n}{2}} \exp\left[ -\frac{1}{2} \sum_{t=1}^{T} tr \left( \tilde{\mathbf{Y}}_{t\cdot}' \tilde{\mathbf{Y}}_{t\cdot} \mathbf{F}_i - 2\mathbf{A}_{\mathcal{G},i\cdot} \boldsymbol{z}_{\mathcal{G},i}^{t-1} \tilde{\mathbf{Y}}_{t\cdot} \mathbf{F}_{i\cdot 3} + \frac{\mathbf{A}_{\mathcal{G},i\cdot} \boldsymbol{z}_{\mathcal{G},i}^{t-1} \boldsymbol{z}_{\mathcal{G},i}^{t-1\prime} \mathbf{A}_{\mathcal{G},i\cdot}'}{\tilde{\sigma}_i^2} \right) \right] \times \\
& \left( \frac{1}{2\pi \tilde{\sigma}_i^2 \tau^2} \right)^{\frac{\nu_i(\mathcal{G})}{2}} \exp\left[ -\frac{\mathbf{A}_{\mathcal{G},i\cdot} \mathbf{A}_{\mathcal{G},i\cdot}'}{2\tilde{\sigma}_i^2 \tau^2} \right] \exp\left[ -\frac{tr(\mathbf{F}_i \tilde{\mathbf{S}})}{2} \right] |\mathbf{F}_i|^{\frac{\omega-4}{2}} \Bigg\} d\mathbf{A}_{\mathcal{G},i\cdot} d\mathbf{F}_i
\end{aligned}
\tag{S4.6}
$$

where $\tilde{\mathbf{S}} = \mathbf{C}^{-1}\mathbf{V}(\mathbf{C}')^{-1}$. Now collecting all terms involving $\mathbf{A}$, we have

$$
\begin{aligned}
& \int \exp\left[ -\frac{1}{2} tr \left( \mathbf{A}_{\mathcal{G},i\cdot} \left( \sum_{t=1}^{T} \frac{\boldsymbol{z}_{\mathcal{G},i}^{t-1} \boldsymbol{z}_{\mathcal{G},i}^{t-1\prime}}{\tilde{\sigma}_i^2} + \frac{\mathbf{I}}{\tilde{\sigma}_i^2 \tau^2} \right) \mathbf{A}_{\mathcal{G},i\cdot}' - 2\mathbf{A}_{\mathcal{G},i\cdot} \sum_{t=1}^{T} \boldsymbol{z}_{\mathcal{G},i}^{t-1} \tilde{\mathbf{Y}}_{t\cdot} \mathbf{F}_{i\cdot 3} \right) \right] d\mathbf{A} \\
= & \left(2\pi\tilde{\sigma}_i^2\right)^{\frac{\nu_i(\mathcal{G})}{2}} \left| \sum_{t=1}^{T} \boldsymbol{z}_{\mathcal{G},i}^{t-1} \boldsymbol{z}_{\mathcal{G},i}^{t-1\prime} + \frac{\mathbf{I}}{\tau^2} \right|^{-\frac{1}{2}} \times \\
& \exp\left( \frac{\mathbf{F}_{i\cdot 3}{}' \left( \sum_{t=1}^{T} \tilde{\mathbf{Y}}_{t\cdot}' \boldsymbol{z}_{\mathcal{G},i}^{t-1\prime} \right) \left( \sum_{t=1}^{T} \boldsymbol{z}_{\mathcal{G},i}^{t-1} \boldsymbol{z}_{\mathcal{G},i}^{t-1\prime} + \frac{\mathbf{I}}{\tau^2} \right)^{-1} \sum_{t=1}^{T} \left( \boldsymbol{z}_{\mathcal{G},i}^{t-1} \tilde{\mathbf{Y}}_{t\cdot} \right) \mathbf{F}_{i\cdot 3}}{2\mathbf{F}_{i33}} \right)
\end{aligned}
$$

Denote by

$$
\begin{aligned}
\mathbf{P}_i &= \left( \sum_{t=1}^{T} \tilde{\mathbf{Y}}_{t\cdot}' \tilde{\mathbf{Y}}_{t\cdot} + \tilde{\mathbf{S}} \right) \\
\mathbf{Q}_i &= \left( \sum_{t=1}^{T} \tilde{\mathbf{Y}}_{t\cdot}' \boldsymbol{z}_{\mathcal{G},i}^{t-1} \right) \left( \sum_{t=1}^{T} \boldsymbol{z}_{\mathcal{G},i}^{t-1} \boldsymbol{z}_{\mathcal{G},i}^{t-1\prime} + \frac{\mathbf{I}}{\tau^2} \right)^{-1} \sum_{t=1}^{T} \left( \boldsymbol{z}_{\mathcal{G},i}^{t-1} \tilde{\mathbf{Y}}_{t\cdot} \right)
\end{aligned}
$$

Then, (S4.6) becomes

$$\mathcal{P}_1 \prod_{i=1}^{k_1} \int |\mathbf{F}_i|^{\frac{n+\omega-4}{2}} \frac{1}{(\tau^2)^{\frac{\nu_i(\mathcal{G})}{2}}} \exp\left[-\frac{1}{2} tr\left(\mathbf{P}_i \mathbf{F}_i\right)\right] \left| \sum_{t=1}^{T} \boldsymbol{z}_{\mathcal{G},i}^{t-1} \boldsymbol{z}_{\mathcal{G},i}^{t-1\prime} + \frac{\mathbf{I}}{\tau^2} \right|^{-\frac{1}{2}} \times$$
$$\exp\left(tr\left(\mathbf{Q}_i \frac{\mathbf{F}_{i\cdot3}\mathbf{F}_{i\cdot3}{}'}{2\mathbf{F}_{i33}}\right)\right) d\mathbf{F}_i \tag{S4.7}$$

Next, consider the block Cholesky decomposition of $\mathbf{F}_i$ in terms of $\mathbf{B}_1$, $\mathbf{e}$ and $f$ as in (S3.2), $\mathbf{F}_i$ can be written as,

$$\mathbf{F}_i = \begin{bmatrix} \mathbf{B}_1 + \mathbf{e}\mathbf{e}' f & \mathbf{e} f \\ \mathbf{e}' f & f \end{bmatrix} \tag{S4.8}$$

Consider the transformation: $\mathbf{F}_i \to (\mathbf{B}_1, \boldsymbol{e}, f)$. The Jacobian of the transformation is given by: $|\boldsymbol{J}| = f^2$. Then (S4.7) reduces to

$$\mathcal{P}_1 \prod_{i=1}^{k_1} \frac{\left| \sum_{t=1}^{T} \boldsymbol{z}_{\mathcal{G},i}^{t-1} \boldsymbol{z}_{\mathcal{G},i}^{t-1\prime} + \frac{\mathbf{I}}{\tau^2} \right|^{-\frac{1}{2}}}{\tau^{2\frac{\nu_i(\mathcal{G})}{2}}} \int\int\int |\mathbf{B}_1|^{\frac{n+\omega-4}{2}} f^{\frac{n+\omega}{2}} \times$$
$$\exp\left[ -\frac{1}{2} tr \left\{ \mathbf{P}_i \begin{bmatrix} \mathbf{B}_1 & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix} + (\mathbf{P}_i - \mathbf{Q}_i) \begin{bmatrix} \boldsymbol{e}\boldsymbol{e}' f & \boldsymbol{e} f \\ \boldsymbol{e}' f & f \end{bmatrix} \right\} \right] d\mathbf{B}_1 \, d\boldsymbol{e} \, df$$
$$= \mathcal{P}_1 \prod_{i=1}^{k_1} \frac{\left| \sum_{t=1}^{T} \boldsymbol{z}_{\mathcal{G},i}^{t-1} \boldsymbol{z}_{\mathcal{G},i}^{t-1\prime} + \frac{\mathbf{I}}{\tau^2} \right|^{-\frac{1}{2}}}{\tau^{2\frac{\nu_i(\mathcal{G})}{2}}} \int\int\int |\mathbf{B}_1|^{\frac{n+\omega-4}{2}} f^{\frac{n+\omega}{2}} \exp\left[ -\frac{1}{2} tr\left( \mathbf{P}_{i,11} \mathbf{B}_1 \right) \right] \times$$
$$\exp\left[ -\frac{f}{2} tr \left( \boldsymbol{e}' \mathbf{R}_{i,11}^{\perp} \boldsymbol{e} + 2 \mathbf{R}_{i,12}^{\perp}{}' \boldsymbol{e} + \mathbf{R}_{i,22}^{\perp} \right) \right] d\mathbf{B}_1 \, d\boldsymbol{e} \, df \text{ , where } \mathbf{R}_i^{\perp} = (\mathbf{P}_i - \mathbf{Q}_i) = n\mathbf{R}_i$$
$$= (const.) \mathcal{P}_1 \prod_{i=1}^{k_1} \frac{\left| \sum_{t=1}^{T} \boldsymbol{z}_{\mathcal{G},i}^{t-1} \boldsymbol{z}_{\mathcal{G},i}^{t-1\prime} + \frac{\mathbf{I}}{\tau^2} \right|^{-\frac{1}{2}}}{\tau^{2\frac{\nu_i(\mathcal{G})}{2}}} \int \left\{ \int \exp\left[ -\frac{f}{2} tr \left( \boldsymbol{e}' \mathbf{R}_{i,11}^{\perp} \boldsymbol{e} + 2 \mathbf{R}_{i,12}^{\perp}{}' \boldsymbol{e} \right) \right] d\boldsymbol{e} \right\} \times$$
$$f^{\frac{n+\omega}{2}} \exp\left[ -\frac{\mathbf{R}_{i,22}^{\perp} f}{2} \right] df \left( \substack{\text{ignoring the terms that are constant} \\ \text{over all the graphs}} \right)$$
$$= (const.) \mathcal{P}_1 \prod_{i=1}^{k_1} \frac{\left| \frac{\mathbf{Z}_{\mathcal{G},i}{}' \mathbf{Z}_{\mathcal{G},i}}{n} + \frac{\mathbf{I}}{n\tau^2} \right|^{-\frac{1}{2}} |\mathbf{R}_{i,11}|^{-\frac{1}{2}}}{(\tau\sqrt{n})^{\nu_i(\mathcal{G})}} \int \exp\left[ -\frac{f}{2} \left( \mathbf{R}_{i,22}^{\perp} - \mathbf{R}_{i,12}^{\perp}{}' \mathbf{R}_{i,11}^{\perp}{}^{-1} \mathbf{R}_{i,12}^{\perp} \right) \right] \times$$
$$f^{\frac{n+\omega}{2}-1} \, df$$
$$= (const.) \prod_{i=1}^{k_1} q^{\nu_i(\mathcal{G})} (1-q)^{(k_1+k_2)-\nu_i(\mathcal{G})} \left( \frac{1}{\tau\sqrt{n}} \right)^{\nu_i(\mathcal{G})} \left| \left( \frac{\mathbf{Z}_{\mathcal{G},i}' \mathbf{Z}_{\mathcal{G},i}}{n} + \frac{\mathbf{I}}{n\tau^2} \right) \mathbf{R}_{i,11} \right|^{-1/2} \times$$
$$\left( \mathbf{R}_{i,22} - \mathbf{R}_{i,12}{}' \mathbf{R}_{i,11}{}^{-1} \mathbf{R}_{i,12} \right)^{-(\frac{n+\nu}{2})}$$

# S5 Proofs of Theoretical Results

## S5.1 Proof of Lemma 1:

Let $\tilde{\mathbf{W}} = \mathbf{P}'\mathbf{W}\mathbf{P}$ be a permuted version of $\mathbf{W}$ where $\mathbf{P}$ is a permutation matrix where

$$
\begin{aligned}
\tilde{\mathbf{W}} &= \begin{bmatrix} \mathbf{W}_{1_{(k_1+k_2)\times(k_1+k_2)}} & \mathbf{W}_{2_{(k_1+k_2)\times 2k_1}} \\ \mathbf{W}_{3_{2k_1\times(k_1+k_2)}} & \mathbf{W}_{4_{2k_1\times 2k_1}} \end{bmatrix} \text{ and} \\
\mathbf{W}_1 &= \begin{bmatrix} \mathbf{A}_{11_{k_1\times k_1}}\theta^2 & \mathbf{A}_{12_{k_1\times k_2}}\theta^2 \\ \mathbf{A}_{21_{k_2\times k_1}} & \mathbf{A}_{22_{k_2\times k_2}} \end{bmatrix} \otimes 1 = \mathbf{A}_\theta \otimes 1, \text{ say} \\
\mathbf{W}_2 &= \begin{bmatrix} \mathbf{A}_{11}\theta^2 \\ \mathbf{A}_{21} \end{bmatrix} \otimes \underbrace{\begin{bmatrix} \theta^2 & \theta \end{bmatrix}}_{\boldsymbol{v}'} \\
\mathbf{W}_3 &= \begin{bmatrix} \mathbf{A}_{11}\theta^2 & \mathbf{A}_{12}\theta^2 \end{bmatrix} \otimes \underbrace{\begin{bmatrix} 1/\theta^2 \\ 1/\theta \end{bmatrix}}_{\boldsymbol{u}} \\
\mathbf{W}_4 &= \mathbf{A}_{11}\theta^2 \otimes \boldsymbol{u}\boldsymbol{v}'
\end{aligned}
$$

Now,

$$
\begin{aligned}
\mathbf{W}_3\mathbf{W}_1^{-1}\mathbf{W}_2 &= \left(\begin{bmatrix}\mathbf{A}_{11}\theta^2 & \mathbf{A}_{12}\theta^2\end{bmatrix}\otimes \boldsymbol{u}\right)\left(\mathbf{A}_\theta^{-1}\otimes 1\right)\left(\begin{bmatrix}\mathbf{A}_{11}\theta^2 \\ \mathbf{A}_{21}\end{bmatrix}\otimes \boldsymbol{v}'\right) \\
&= \begin{bmatrix}\mathbf{A}_{11}\theta^2 & \mathbf{A}_{12}\theta^2\end{bmatrix}\mathbf{A}_\theta^{-1}\begin{bmatrix}\mathbf{A}_{11}\theta^2 \\ \mathbf{A}_{21}\end{bmatrix}\otimes \boldsymbol{u}\boldsymbol{v}' \\
&= \begin{bmatrix}\mathbf{I}_{k_1\times k_1} \mid 0_{k_1\times k_2}\end{bmatrix}\mathbf{A}_\theta\mathbf{A}_\theta^{-1}\mathbf{A}_\theta\begin{bmatrix}\mathbf{I}_{k_1\times k_1} \\ \hline 0_{k_2\times k_1}\end{bmatrix}\otimes \boldsymbol{u}\boldsymbol{v}' \\
&= \left(\mathbf{A}_{11}\theta^2\right)\otimes \boldsymbol{u}\boldsymbol{v}' \\
&= \mathbf{W}_4
\end{aligned}
$$

Next, we establish $\lambda \in Sp\begin{bmatrix}\mathbf{W}_1 & \mathbf{W}_2 \\ \mathbf{W}_3 & \mathbf{W}_4\end{bmatrix}$ iff $\lambda \in Sp\left(\mathbf{W}_1 + \mathbf{W}_2\mathbf{W}_3\mathbf{W}_1^{-1}\right)$

Note that

$$
\begin{aligned}
\lambda \in Sp\begin{bmatrix}\mathbf{W}_1 & \mathbf{W}_2 \\ \mathbf{W}_3 & \mathbf{W}_4\end{bmatrix} &\iff \begin{bmatrix}\mathbf{W}_1 & \mathbf{W}_2 \\ \mathbf{W}_3 & \mathbf{W}_3\mathbf{W}_1^{-1}\mathbf{W}_2\end{bmatrix}\begin{bmatrix}x \\ y\end{bmatrix} = \lambda\begin{bmatrix}x \\ y\end{bmatrix} \\
&\iff \mathbf{W}_1x + \mathbf{W}_2y = \lambda x \quad \text{and} \quad \mathbf{W}_3\mathbf{W}_1^{-1}(\mathbf{W}_1x + \mathbf{W}_2y) = \lambda y \\
&\iff y = \mathbf{W}_3\mathbf{W}_1^{-1}x \quad \text{and} \quad \mathbf{W}_1x + \mathbf{W}_2y = \lambda x \\
&\iff y = \mathbf{W}_3\mathbf{W}_1^{-1}x \quad \text{and} \quad \left(\mathbf{W}_1 + \mathbf{W}_2\mathbf{W}_3\mathbf{W}_1^{-1}\right)x = \lambda x \\
&\iff \lambda \in Sp\left(\mathbf{W}_1 + \mathbf{W}_2\mathbf{W}_3\mathbf{W}_1^{-1}\right)
\end{aligned}
$$

Then,

$$
\begin{aligned}
& \left(\mathbf{W}_1 + \mathbf{W}_2\mathbf{W}_3\mathbf{W}_1^{-1}\right) \\
= \;& \mathbf{W}_1 + \left( \begin{bmatrix} \mathbf{A}_{11}\theta^2 \\ \mathbf{A}_{21} \end{bmatrix} \otimes \begin{bmatrix} \theta^2 & \theta \end{bmatrix} \right) \left( \begin{bmatrix} \mathbf{A}_{11}\theta^2 & \mathbf{A}_{12}\theta^2 \end{bmatrix} \otimes \begin{bmatrix} 1/\theta^2 \\ 1/\theta \end{bmatrix} \right) \left(\mathbf{A}_\theta^{-1} \otimes 1\right) \\
= \;& \mathbf{W}_1 + 2 \begin{bmatrix} \mathbf{A}_{11}\theta^2 \\ \mathbf{A}_{21} \end{bmatrix} \begin{bmatrix} \mathbf{A}_{11}\theta^2 & \mathbf{A}_{12}\theta^2 \end{bmatrix} \mathbf{A}_\theta^{-1} \\
= \;& \mathbf{W}_1 + 2 \left( \begin{bmatrix} \mathbf{A}_{11}\theta^2 \\ \mathbf{A}_{21} \end{bmatrix} \begin{bmatrix} \mathbf{I}_{k_1 \times k_1} \mid 0_{k_1 \times k_2} \end{bmatrix} \mathbf{A}_\theta \mathbf{A}_\theta^{-1} \right) \\
= \;& \mathbf{W}_1 + 2 \begin{bmatrix} \mathbf{A}_{11}\theta^2 & 0 \\ \mathbf{A}_{21} & 0 \end{bmatrix} \\
= \;& \mathbf{A}_\theta + 2 \begin{bmatrix} \mathbf{A}_{11}\theta^2 & 0 \\ \mathbf{A}_{21} & 0 \end{bmatrix}
\end{aligned}
$$

Thus, we have

$$
Sp(\mathbf{W}) = Sp(\tilde{\mathbf{W}}) = Sp\left(\mathbf{W}_1 + \mathbf{W}_2\mathbf{W}_3\mathbf{W}_1^{-1}\right) = Sp\left( \mathbf{A}_\theta + 2 \begin{bmatrix} \mathbf{A}_{11}\theta^2 & 0 \\ \mathbf{A}_{21} & 0 \end{bmatrix} \right)
$$

which provides the following upper bound for the maximum eigen value of $\mathbf{W}$ (denoted by $\rho(\mathbf{W})$),

$$
\rho(\mathbf{W}) \leq \rho(\mathbf{A}_\theta) + 2\theta^2 \rho\left(\mathbf{A}_{11}\right).
$$

### S5.2 Proof of Theorem 1:

Throughout the proof, for ease of presentation, we denote $\pi_{\text{pseudo}}(. \mid \mathcal{Y})$, $\mathbf{Z}_{\mathcal{G},i}$ and $\mathbf{Z}_{\mathcal{G},j}$ by $\pi(. \mid \mathcal{Y})$, $\mathbf{Z}_i$ and $\mathbf{Z}_j$ respectively. In the entire proof, $p = 3k_1 + k_2$. For notational convenience, we will use $\nu_j$ to denote $\nu_{k_1+j}, j = 1, \cdots, k_2$. Before presenting the proof of the theorems, we introduce next notations needed in subsequent developments. Let $\boldsymbol{t}_i = \{j_1, \cdots, j_{\nu_{t_i}}\} \subset \{1, \cdots, (k_1 + k_2)\}$ be the set of column indices from $\mathbf{Z}$ corresponding to the non-zero positions of $\mathbf{A}_{0n,i.}$ in the true model for $i = 1, \cdots, k_1$ and it represents the neighbors of $i$ in the true activity graph $\mathcal{G}_0$. Similarly let $\boldsymbol{t}_j = \{s_1, \cdots, s_{\nu_{t_j}}\} \subset \{1, \cdots, (k_1 + k_2)\}$ be the set of column indices from $\mathbf{Z}$ corresponding to the non-zero positions of $\mathbf{A}_{0n,j.}$ (i.e. the $(k_1 + j)^{th}$ row of $\mathbf{A}_{0n}$) in the true model for $j = 1, \cdots, k_2$. Also $\boldsymbol{m}_i = \{i_1, \cdots, i_{\nu_{m_i}}\}$ represents the same for any candidate model (distinct from the true model) for $i = 1, \cdots, k_1$. Similarly $\boldsymbol{m}_j$ can be defined for $j = 1, \cdots, k_2$. Given two activity graphs $\mathcal{G}_0$ and $\mathcal{G}_m$, $\boldsymbol{t}_i = \boldsymbol{m}_i$ implies the neighbors of $i$ are identical in both of the graphs (i.e. $\mathcal{G}_0$ and $\mathcal{G}_m$ have the same $i$-th column). Let us define $d(\boldsymbol{m}_i, \boldsymbol{t}_i) :=$ Card $(\{\boldsymbol{t}_i \cup \boldsymbol{m}_i\} \setminus \{\boldsymbol{t}_i \cap \boldsymbol{m}_i\})$ - number of disagreements in the $i$-th column between $\mathcal{G}_m$ and $\mathcal{G}_0$. Similarly $d(\boldsymbol{m}_j, \boldsymbol{t}_j)$ can be defined. Thus total number of disagreements is denoted by $D(m, t)$ and is equal to

$\left(\sum_{i=1}^{k_1} d(\boldsymbol{m}_i, \boldsymbol{t}_i) + \sum_{j=1}^{k_2} d(\boldsymbol{m}_j, \boldsymbol{t}_j)\right)$.

$$\mathcal{G}_0 = \begin{array}{c} \quad\quad \boldsymbol{t}_i \quad\quad \\ \left[\begin{array}{ccc} \cdots & \begin{pmatrix} \vdots \\ 1 \\ \vdots \\ 0 \\ \vdots \end{pmatrix} & \cdots \end{array}\right]_{(k_1+k_2)\times(k_1+k_2)} \end{array} \qquad \mathcal{G}_m = \begin{array}{c} \quad\quad \boldsymbol{m}_i \quad\quad \\ \left[\begin{array}{ccc} \cdots & \begin{pmatrix} \vdots \\ 0 \\ \vdots \\ 1 \\ \vdots \end{pmatrix} & \cdots \end{array}\right]_{(k_1+k_2)\times(k_1+k_2)} \end{array}$$

Note that given the true activity graph $\mathcal{G}_0$ and another arbitrary graph $\mathcal{G}_m$

$$\begin{aligned} & \frac{\pi(\mathcal{G}_m \mid \mathcal{Y})}{\pi(\mathcal{G}_0 \mid \mathcal{Y})} \\ = \quad & \prod_{i=1}^{k_1} \left(\frac{1}{\tau\sqrt{n}}\right)^{(\nu_{m_i}-\nu_{t_i})} \frac{q^{\nu_{m_i}}(1-q)^{(k_1+k_2)-\nu_{m_i}}}{q^{\nu_{t_i}}(1-q)^{(k_1+k_2)-\nu_{t_i}}} \frac{\left|\frac{\mathbf{Z}'_{m_i}\mathbf{Z}_{m_i}}{n} + \frac{\mathbf{I}_{\nu_{m_i}}}{n\tau^2}\right|^{-1/2}}{\left|\frac{\mathbf{Z}'_{t_i}\mathbf{Z}_{t_i}}{n} + \frac{\mathbf{I}_{\nu_{t_i}}}{n\tau^2}\right|^{-1/2}} \frac{\left|\mathbf{R}_{m_i,11}\right|^{-1/2}}{\left|\mathbf{R}_{t_i,11}\right|^{-1/2}} \left(\frac{\mathbf{U}_{\mathbf{R}_{m_i}}}{\mathbf{U}_{\mathbf{R}_{t_i}}}\right)^{-(\frac{n+\nu}{2})} \\ & \prod_{j=1}^{k_2} \left(\frac{1}{\tau\sqrt{n}}\right)^{(\nu_{m_j}-\nu_{t_j})} \frac{q^{\nu_{m_j}}(1-q)^{(k_1+k_2)-\nu_{m_j}}}{q^{\nu_{t_j}}(1-q)^{(k_1+k_2)-\nu_{t_j}}} \frac{\left|\frac{\mathbf{Z}'_{m_j}\mathbf{Z}_{m_j}}{n} + \frac{\mathbf{I}_{\nu_{m_j}}}{n\tau^2}\right|^{-1/2}}{\left|\frac{\mathbf{Z}'_{t_j}\mathbf{Z}_{t_j}}{n} + \frac{\mathbf{I}_{\nu_{t_j}}}{n\tau^2}\right|^{-1/2}} \left(\frac{R_{m_j}}{R_{t_j}}\right)^{-(\frac{n}{2}+\alpha)} \end{aligned} \tag{S5.1}$$

where, $\mathbf{U}_{\mathbf{R}_i} = \mathbf{R}_{i,22} - \mathbf{R}_{i,12}{}'(\mathbf{R}_{i,11})^{-1}\mathbf{R}_{i,12}$ is the schur complement of $\mathbf{R}_{i,11}$. Using the same argument as in Ghosh et al. (2021), it can be shown that

$$
\begin{aligned}
& \frac{1}{2}\frac{\pi(\mathcal{G}_m \mid \mathcal{Y})}{\pi(\mathcal{G}_0 \mid \mathcal{Y})} \\
\leq \quad & \prod_{i=1}^{k_1}\left(\frac{2q}{\tau\sqrt{n}}\right)^{(\nu_{m_i}-\nu_{t_i})} \frac{\left|\frac{\mathbf{Z}_{m_i}'\mathbf{Z}_{m_i}}{n} + \frac{\mathbf{I}_{\nu_{m_i}}}{n\tau^2}\right|^{-1/2}\left|\mathbf{R}_{m_i,11}\right|^{-1/2}}{\left|\frac{\mathbf{Z}_{t_i}'\mathbf{Z}_{t_i}}{n} + \frac{\mathbf{I}_{\nu_{t_i}}}{n\tau^2}\right|^{-1/2}\left|\mathbf{R}_{t_i,11}\right|^{-1/2}} \left(\frac{\mathbf{U}_{\mathbf{R}_{t_i}}}{\mathbf{U}_{\mathbf{R}_{m_i}}}\right)^{\left(\frac{n+\nu}{2}\right)} \\
& \prod_{j=1}^{k_2}\left(\frac{2q}{\tau\sqrt{n}}\right)^{(\nu_{m_j}-\nu_{t_j})} \frac{\left|\frac{\mathbf{Z}_{m_j}'\mathbf{Z}_{m_j}}{n} + \frac{\mathbf{I}_{\nu_{m_j}}}{n\tau^2}\right|^{-1/2}}{\left|\frac{\mathbf{Z}_{t_i}'\mathbf{Z}_{t_i}}{n} + \frac{\mathbf{I}_{\nu_{t_i}}}{n\tau^2}\right|^{-1/2}} \left(\frac{R_{t_j}}{R_{m_j}}\right)^{\left(\frac{n}{2}+\alpha\right)} \\
= \quad & \prod_{i=1}^{k_1} B(\boldsymbol{m}_i,\boldsymbol{t}_i)\prod_{j=1}^{k_2} B^{\star}\left(\boldsymbol{m}_j,\boldsymbol{t}_j\right), \text{ say,}
\end{aligned}
\tag{S5.2}
$$

As the analysis of $B^{\star}\left(\boldsymbol{m}_j,\boldsymbol{t}_j\right)$ in (S5.2) corresponding to $k_2$ low frequency regressions is very similar to the analysis presented in Ghosh et al. (2021), we will only study the behaviour of $B\left(\boldsymbol{m}_i,\boldsymbol{t}_i\right)$ in detail that correspond to the high frequency terms. For $i = 1,\cdots,k_1$, let

$$
\tilde{\mathbf{R}}_i = \frac{1}{n}\tilde{\mathbf{Y}}_{i,\mathrm{H}}'\tilde{\mathbf{Y}}_{i,\mathrm{H}} - \left(\frac{1}{n}\tilde{\mathbf{Y}}_{i,\mathrm{H}}'\mathbf{Z}_i\right)\left(\frac{1}{n}\mathbf{Z}_i'\mathbf{Z}_i\right)^{-1}\left(\frac{1}{n}\mathbf{Z}_i'\tilde{\mathbf{Y}}_{i,\mathrm{H}}\right)
$$

Then the term $\mathbf{U}_{\tilde{\mathbf{R}}_i} = \tilde{\mathbf{R}}_{i,22} - \tilde{\mathbf{R}}_{i,12}{}'(\tilde{\mathbf{R}}_{i,11})^{-1}\tilde{\mathbf{R}}_{i,12}$, which is the schur complement of $\tilde{\mathbf{R}}_{i,11}$, represents residual sum of squares for regressing $\tilde{\boldsymbol{y}}_{i3}$ on $\tilde{\boldsymbol{\xi}}_{i1},\tilde{\boldsymbol{\xi}}_{i2}$ and $\mathbf{Z}_i$. The corresponding regression equation follows from (S2.11)

:

$$\tilde{\boldsymbol{y}}_{i3} = \|\boldsymbol{\delta}\|^2 \left(\mathbf{Z}_i, \tilde{\boldsymbol{\xi}}_{i1}, \tilde{\boldsymbol{\xi}}_{i2}\right) \begin{bmatrix} \mathbf{A}'_{i\cdot} \\ 0 \\ 0 \end{bmatrix} + \tilde{\boldsymbol{\xi}}_{i3} \text{ for } i = 1, \cdots k_1 \tag{S5.3}$$

Let us denote $\tilde{\mathbf{X}} = \left(\mathbf{Z}_{\mathcal{G},i}, \tilde{\boldsymbol{\xi}}_{i1}, \tilde{\boldsymbol{\xi}}_{i2}\right)$. We will keep on referring to (S5.3) for the analysis of $B\left(\boldsymbol{m}_i, \boldsymbol{t}_i\right)$ throughout the proof. It follows from Assumption A2 that

$$\begin{aligned} \tilde{\lambda}_1 &\leq \lambda_{\min}(\mathbf{C}_Z) \leq \lambda_{\max}(\mathbf{C}_Z) \leq \tilde{\lambda}_2 \text{ and} \\ \sigma_1 &\leq \lambda_{\min}\left(\mathbf{C}_{\begin{pmatrix} \tilde{\boldsymbol{\xi}}_{i1} \\ \tilde{\boldsymbol{\xi}}_{i2} \end{pmatrix}}\right) \leq \lambda_{\max}\left(\mathbf{C}_{\begin{pmatrix} \tilde{\boldsymbol{\xi}}_{i1} \\ \tilde{\boldsymbol{\xi}}_{i2} \end{pmatrix}}\right) \leq \sigma_2 \end{aligned} \tag{S5.4}$$

which leads to

$$\lambda'_1 \leq \lambda_{\min}(\mathbf{C}_{\tilde{X}}) \leq \lambda_{\max}(\mathbf{C}_{\tilde{X}}) \leq \lambda'_2 \tag{S5.5}$$

for appropriate constants $0 < \tilde{\lambda}_1 < \tilde{\lambda}_2 < \infty$, $0 < \lambda'_1 < \lambda'_2 < \infty$ and $0 < \sigma_1 < \sigma_2 < \infty$ not depending on $n$. $\mathbf{C}_{\begin{pmatrix} \tilde{\boldsymbol{\xi}}_{i1} \\ \tilde{\boldsymbol{\xi}}_{i2} \end{pmatrix}}$ denotes the population error covariance matrix for $\begin{pmatrix} \tilde{\boldsymbol{\xi}}_{i1} \\ \tilde{\boldsymbol{\xi}}_{i2} \end{pmatrix}$ and $\mathbf{C}_Z$ is such that the rows of $\mathbf{Z}$ is distributed as $\mathcal{N}(\mathbf{O}, \mathbf{C}_Z)$. Next, recall that by Assumption A2 there exist $0 < \sigma_{\min} < \sigma_{\max} < \infty$ such that

$$\sigma_{\min} < \lambda_{\min}(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}) \leq \lambda_{\max}(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}) < \sigma_{\max}$$

Then, define the following events:

$$
\begin{aligned}
G_{1,n} &:= \left\{ \left\| \frac{\mathbf{X}'\mathbf{X}}{n} - \mathbf{C}_X \right\|_{\max} \leq \mathcal{M}_n \sqrt{\frac{2\log p}{n}} \right\}, \\
G_{2,n} &:= \left\{ \left\| \frac{\mathbf{X}'\mathbf{E}}{n} \right\|_{\max} \leq \mathcal{M}_n \sqrt{\frac{\log p^2}{n}} \right\} \\
E_{1,n} &:= \bigcap_{i=1}^{3k_1+k_2} \left\{ \frac{7\sigma_{\min}}{8} \leq \frac{\boldsymbol{\xi}_i'\boldsymbol{\xi}_i}{n} \leq \frac{3\sigma_{\max}}{2} \right\} \qquad \text{and} \\
E_{2,n} &:= \bigcap_{m:1\leq|m|<\frac{n}{2}} \left\{ \left\| \frac{\mathbf{X}_m'\mathbf{X}_m}{n} - \mathbf{C}_{Xm} \right\| \leq 4\mathcal{M}_n \sqrt{\frac{\nu_m \log p}{cn}} \right\}
\end{aligned}
$$

where $c$ is a constant not depending on $n$. Let $E_n = E_{1,n} \cap E_{2,n}$. As followed from the arguments given in Ghosh et al. (2021), we know,

$$
\mathbb{P}(G_{1,n} \cap G_{1,2} \cap E_n) \to 1 \qquad \text{as } n \to \infty. \tag{S5.6}
$$

Now, define the next set of events:

$$
\begin{aligned}
\tilde{G}_{1,n} &:= \left\{ \left\| \frac{\mathbf{Z}'\mathbf{Z}}{n} - \mathbf{C}_Z \right\|_{\max} \leq \mathcal{M}_n \sqrt{\frac{2 \log p}{n}} \right\}, \\
\tilde{G}_{2,n} &:= \left\{ \left\| \frac{\mathbf{Z}'\mathbf{E}}{n} \right\|_{\max} \leq \mathcal{M}_n \sqrt{\frac{\log p^2}{n}} \right\} \\
\tilde{G}_{21,n} &:= \left\{ \left\| \frac{\mathbf{Z}'\tilde{\mathbf{E}}_1}{n} \right\|_{\max} \leq \mathcal{M}_n \sqrt{\frac{\log p^2}{n}} \right\} \\
\tilde{G}_{22,n} &:= \left\{ \left\| \frac{\mathbf{Z}'\tilde{\mathbf{E}}_2}{n} \right\|_{\max} \leq \mathcal{M}_n \sqrt{\frac{\log p^2}{n}} \right\} \\
\tilde{G}_{23,n} &:= \left\{ \left\| \frac{\mathbf{Z}'\tilde{\mathbf{E}}_3}{n} \right\|_{\max} \leq \mathcal{M}_n \sqrt{\frac{\log p^2}{n}} \right\} \\
\tilde{G}_{3,n} &:= \left\{ \left\| \frac{\tilde{\mathbf{X}}'\tilde{\mathbf{X}}}{n} - \mathbf{C}_{\tilde{X}} \right\|_{\max} \leq \mathcal{M}_n \sqrt{\frac{2 \log p}{n}} \right\} \\
\tilde{E}_{11,n} &:= \bigcap_{i=1}^{k_1} \left\{ \frac{7\sigma_{11}}{8} \leq \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i1}}{n} \leq \frac{3\sigma_{12}}{2} \right\} \\
\tilde{E}_{12,n} &:= \bigcap_{i=1}^{k_1} \left\{ \frac{7\sigma_{21}}{8} \leq \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{\xi}}_{i2}}{n} \leq \frac{3\sigma_{22}}{2} \right\} \\
\tilde{E}_{13,n} &:= \bigcap_{i=1}^{k_1} \left\{ \frac{7\sigma_{31}}{8} \leq \frac{\tilde{\boldsymbol{\xi}}_{i3}'\tilde{\boldsymbol{\xi}}_{i3}}{n} \leq \frac{3\sigma_{32}}{2} \right\} \\
\tilde{E}_{21,n} &:= \bigcap_{m:1 \leq |m| < \frac{n}{2}} \left\{ \left\| \frac{\mathbf{Z}_m'\mathbf{Z}_m}{n} - \mathbf{C}_{Zm} \right\| \leq 4c_1 \mathcal{M}_n \sqrt{\frac{\nu_m \log p}{cn}} \right\} \\
\tilde{E}_{22,n} &:= \bigcap_{m:1 \leq |m| < \frac{n}{2}} \left\{ \left\| \frac{\tilde{\mathbf{X}}_m'\tilde{\mathbf{X}}_m}{n} - \mathbf{C}_{\tilde{X}m} \right\| \leq 4c_1 \mathcal{M}_n \sqrt{\frac{\nu_m \log p}{cn}} \right\}
\end{aligned}
$$

where $\sigma_{11}, \sigma_{12}, \sigma_{21}, \sigma_{22}, \sigma_{31}, \sigma_{32}, c, c_1$ are constants not depending on $n$, defined later in this section. In the above events $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2$ and $\tilde{\mathbf{E}}_3$ are matrices whose i-th column is given by $\tilde{\boldsymbol{\xi}}_{i1}$, $\tilde{\boldsymbol{\xi}}_{i2}$ and $\tilde{\boldsymbol{\xi}}_{i3}$ respectively. Let $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\delta}$ be as defined in (S2.10) and let us write $\boldsymbol{a} = (a_1, a_2, a_3)$ and $\boldsymbol{b} = (b_1, b_2, b_3)$. Then $\tilde{\mathbf{E}}_1, \tilde{\mathbf{E}}_2$ and $\tilde{\mathbf{E}}_3$ can be written as $\mathbf{EP}_1, \mathbf{EP}_2$ and $\mathbf{EP}_3$ respectively where

$$\mathbf{P}_1 = \begin{bmatrix} a_1\mathbf{I}_{k_1\times k_1} \\ a_2\mathbf{I}_{k_1\times k_1} \\ a_3\mathbf{I}_{k_1\times k_1} \\ \mathbf{O}_{k_2\times k_1} \end{bmatrix}, \ \mathbf{P}_2 = \begin{bmatrix} b_1\mathbf{I}_{k_1\times k_1} \\ b_2\mathbf{I}_{k_1\times k_1} \\ b_3\mathbf{I}_{k_1\times k_1} \\ \mathbf{O}_{k_2\times k_1} \end{bmatrix}, \ \mathbf{P}_3 = \begin{bmatrix} \theta^2\mathbf{I}_{k_1\times k_1} \\ \mathbf{I}_{k_1\times k_1} \\ \theta\mathbf{I}_{k_1\times k_1} \\ \mathbf{O}_{k_2\times k_1} \end{bmatrix}$$

Note that $\mathbf{Z}$ can be written as,

$$\begin{aligned} \mathbf{Z} =& \mathbf{X} \left[ \begin{array}{c|c} \mathbf{I}_{k_1\times k_1} & \\ \theta^2\mathbf{I}_{k_1\times k_1} & \mathbf{O}_{3k_1\times k_2} \\ \theta\mathbf{I}_{k_1\times k_1} & \\ \hline \mathbf{O}_{k_2\times k_1} & \mathbf{I}_{k_2\times k_2} \end{array} \right] \\ =& \mathbf{XQ}, \text{ say} \end{aligned}$$

Then, by **Proposition B1, B3** and **Corollary B3, B4** of Ghosh et al. (2018), with $\boldsymbol{u} = \mathbf{Q}\mathbf{e}_i$ and $\boldsymbol{v} = \mathbf{Q}\mathbf{e}_j$ where $\mathbf{e}_i$ and $\mathbf{e}_j$ are $i-$th and $j-$th unit vector in $\mathcal{R}^{k_1+k_2}$ respectively, there exists $c$ (not depending on $n$) such that

$$\mathbb{P}\left(\sup_{1\le i,\ j\le(k_1+k_2)} \left|\mathbf{e}_i{}'\mathbf{Q}'\left(\frac{\mathbf{X}'\mathbf{X}}{n} - \mathbf{C}_X\right)\mathbf{Q}\mathbf{e}_j\right| > \frac{\lambda_{\max}(\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}})}{\mu_{\min}(\mathcal{A})}\eta\right) \le 2\exp\left(-cn\min\left\{\frac{\eta^2}{4}, \frac{\eta}{2}\right\} + 2\log p\right)$$

By taking $(\boldsymbol{u}, \boldsymbol{v}) = (\mathbf{Q}\mathbf{e}_i, \mathbf{e}_j)$ where $\mathbf{e}_i \in \mathcal{R}^{k_1+k_2}$ and $\mathbf{e}_j \in \mathcal{R}^{3k_1+k_2}$ gives the following inequality:

$$\mathbb{P}\left(\sup_{\substack{1\le i\le(k_1+k_2), \\ 1\le j\le(3k_1+k_2)}} \left|\mathbf{e}_i{}'\mathbf{Q}'\left(\frac{\mathbf{X}'\mathbf{E}}{n}\right)\mathbf{e}_j\right| > 2\pi\lambda_{\max}(\boldsymbol{\Sigma}_{\boldsymbol{\epsilon}})\left[1 + \frac{1+\mu_{\max}(\mathcal{A})}{\mu_{\min}(\mathcal{A})}\right]\eta\right) \le 6\exp\left(-cn\min\left\{\frac{\eta^2}{4}, \frac{\eta}{2}\right\} + \log p^2\right)$$

Again by taking $(\boldsymbol{u}, \boldsymbol{v}) = (\mathbf{Q}\mathbf{e}_i, \mathbf{P}_1\mathbf{e}_j), (\boldsymbol{u}, \boldsymbol{v}) = (\mathbf{Q}\mathbf{e}_i, \mathbf{P}_2\mathbf{e}_j)$ and $(\boldsymbol{u}, \boldsymbol{v}) = (\mathbf{Q}\mathbf{e}_i, \mathbf{P}_3\mathbf{e}_j)$ where $\mathbf{e}_i \in \mathcal{R}^{k_1+k_2}$ and $\mathbf{e}_j \in \mathcal{R}^{k_1}$ gives the following three inequalities:

$$\mathbb{P}\left(\sup_{\substack{1\le i\le(k_1+k_2),\\ 1\le j\le k_1}}\left|\mathbf{e}_i{}'\mathbf{Q}'\left(\frac{\mathbf{X}'\mathbf{E}}{n}\right)\mathbf{P}_1\mathbf{e}_j\right| > 2\pi\lambda_{\max}(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}})\left[1+\frac{1+\mu_{\max}(\mathcal{A})}{\mu_{\min}(\mathcal{A})}\right]\eta\right) \le 6\exp\left(-cn\min\left\{\frac{\eta^2}{4},\frac{\eta}{2}\right\}+\log p^2\right)$$

$$\mathbb{P}\left(\sup_{\substack{1\le i\le(k_1+k_2),\\ 1\le j\le k_1}}\left|\mathbf{e}_i{}'\mathbf{Q}'\left(\frac{\mathbf{X}'\mathbf{E}}{n}\right)\mathbf{P}_2\mathbf{e}_j\right| > 2\pi\lambda_{\max}(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}})\left[1+\frac{1+\mu_{\max}(\mathcal{A})}{\mu_{\min}(\mathcal{A})}\right]\eta\right) \le 6\exp\left(-cn\min\left\{\frac{\eta^2}{4},\frac{\eta}{2}\right\}+\log p^2\right)$$

$$\mathbb{P}\left(\sup_{\substack{1\le i\le(k_1+k_2),\\ 1\le j\le k_1}}\left|\mathbf{e}_i{}'\mathbf{Q}'\left(\frac{\mathbf{X}'\mathbf{E}}{n}\right)\mathbf{P}_3\mathbf{e}_j\right| > 2\pi\lambda_{\max}(\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}})\left[1+\frac{1+\mu_{\max}(\mathcal{A})}{\mu_{\min}(\mathcal{A})}\right]\eta\right) \le 6\exp\left(-cn\min\left\{\frac{\eta^2}{4},\frac{\eta}{2}\right\}+\log p^2\right)$$

Next, by setting an appropriate $\eta \sim \sqrt{\frac{\log p^2}{cn}}$ and under Assumption A1 we have $\mathbb{P}(\tilde{G}_{1,n}) \to 1$ , $\mathbb{P}(\tilde{G}_{2,n}) \to 1$, $\mathbb{P}(\tilde{G}_{21,n}) \to 1$, $\mathbb{P}(\tilde{G}_{22,n}) \to 1$ and $\mathbb{P}(\tilde{G}_{23,n}) \to 1$ as $n \to \infty$. It follows from Lemma 3.4 of Xiang et al. (2015) that,

$$\left\| \begin{bmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i1}}{n} & \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i2}}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{\xi}}_{i1}}{n} & \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{\xi}}_{i2}}{n} \end{bmatrix} - \mathbf{C}_{\begin{pmatrix} \tilde{\boldsymbol{\xi}}_{i1} \\ \tilde{\boldsymbol{\xi}}_{i2} \end{pmatrix}} \right\|_{\max} \le \mathcal{M}_n\sqrt{\frac{\log p}{n}} \tag{S5.7}$$

This along with $\mathbb{P}(\tilde{G}_{1,n}) \to 1$,$\mathbb{P}(\tilde{G}_{21,n}) \to 1$ and $\mathbb{P}(\tilde{G}_{22,n}) \to 1$ proves that $\mathbb{P}(\tilde{G}_{3,n}) \to 1$.

Note that the rows of $\mathbf{E}_{i,\mathrm{H}} = \left(\boldsymbol{\varepsilon}_{i,\mathrm{H}}^{t}, \boldsymbol{\varepsilon}_{i,\mathrm{H}}^{t-\frac{2}{3}}, \boldsymbol{\varepsilon}_{i,\mathrm{H}}^{t-\frac{1}{3}}\right), t = 1,\cdots,n$ are distributed as $\mathcal{N}\left(\mathbf{O}, \boldsymbol{\Sigma}_{i,\mathrm{H},0n}\right)$. Then for $i = 1,\cdots,k_1$,

$$\begin{aligned} &\mathbf{E}_{i,\mathrm{H}}\boldsymbol{\delta} = \tilde{\boldsymbol{\xi}}_{i3} \sim \mathcal{N}\left(\mathbf{O}, \tilde{\sigma}_{i3}^2\mathbf{I}_n\right) \text{ where } \tilde{\sigma}_{i3}^2 = \boldsymbol{\delta}'\boldsymbol{\Sigma}_{i,\mathrm{H},0n}\boldsymbol{\delta} \\ &\implies \frac{\tilde{\boldsymbol{\xi}}_{i3}}{\tilde{\sigma}_{i3}^2} \sim \mathcal{N}\left(\mathbf{O}, \mathbf{I}_n\right) \end{aligned}$$

An application of the Hanson-Wright inequality (Rudelson et al. (2013)) yields that there exists $K$ such that

$$\mathbb{P}\left(\left|\frac{\tilde{\boldsymbol{\xi}}_{i3}'\tilde{\boldsymbol{\xi}}_{i3}}{n\tilde{\sigma}_{i3}^2}-1\right|>\frac{1}{8}\right)\leq 2\mathrm{e}^{-cn/4K^4}.$$

That is $\frac{7\tilde{\sigma}_{i3}^2}{8}\leq\frac{\tilde{\boldsymbol{\xi}}_{i,3}'\tilde{\boldsymbol{\xi}}_{i,3}}{n}\leq\frac{3\tilde{\sigma}_{i3}^2}{2}$ happens with probability at least $1-2\mathrm{e}^{-cn/4K^4}$. Let $\boldsymbol{\Sigma}_{3,0n}$ be the $k_1\times k_1$ error covariance matrix of the errors $\tilde{\boldsymbol{\xi}}_{i,3}, i=1,\cdots,k_1$, having $\tilde{\sigma}_{i3}^2$ as its i-th diagonal element. It follows from Assumption A2 that the eigenvalues of the matrix $\boldsymbol{\Sigma}_{3,0n}$ will be bounded. Hence there exist $0<\sigma_{31}<\sigma_{32}<\infty$ not depending on $n$ such that $\sigma_{31}<\lambda_{\min}\left(\boldsymbol{\Sigma}_{3,0n}\right)<\lambda_{\max}\left(\boldsymbol{\Sigma}_{3,0n}\right)<\sigma_{32}$. Hence $\mathbb{P}(\tilde{E}_{13,n})\to 1$ as $n\to\infty$. Similarly it can be shown that

$$\mathbb{P}\left(\left|\frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i1}}{n\tilde{\sigma}_{i1}^2}-1\right|>\frac{1}{8}\right)\leq 2\mathrm{e}^{-cn/4K^4}\text{ and}$$
$$\mathbb{P}\left(\left|\frac{\tilde{\boldsymbol{\xi}}_{i,2}'\tilde{\boldsymbol{\xi}}_{i,2}}{n\tilde{\sigma}_{i2}^2}-1\right|>\frac{1}{8}\right)\leq 2\mathrm{e}^{-cn/4K^4}$$

where $\tilde{\sigma}_{i1}^2=\boldsymbol{a}'\boldsymbol{\Sigma}_{i,\mathrm{H},0n}\boldsymbol{a}$ and $\tilde{\sigma}_{i2}^2=\boldsymbol{b}'\boldsymbol{\Sigma}_{i,\mathrm{H},0n}\boldsymbol{b}$. Let $\boldsymbol{\Sigma}_{1,0n}$ and $\boldsymbol{\Sigma}_{2,0n}$ be the error covariance matrices of the errors $\tilde{\boldsymbol{\xi}}_{i1}$ and $\tilde{\boldsymbol{\xi}}_{i2}$ respectively, $i=1,\cdots,k_1$, whose eigen values are bounded. Hence there exist $0<\sigma_{11}<\sigma_{12}<\infty$ and $0<\sigma_{21}<\sigma_{22}<\infty$ not depending on $n$ such that

$$\sigma_{11}<\lambda_{\min}\left(\boldsymbol{\Sigma}_{1,0n}\right)<\lambda_{\max}\left(\boldsymbol{\Sigma}_{1,0n}\right)<\sigma_{12}$$

$$\sigma_{21}<\lambda_{\min}\left(\boldsymbol{\Sigma}_{2,0n}\right)<\lambda_{\max}\left(\boldsymbol{\Sigma}_{2,0n}\right)<\sigma_{22}$$

Hence $\mathbb{P}(\tilde{E}_{11,n}) \to 1, \mathbb{P}(\tilde{E}_{12,n}) \to 1$ as $n \to \infty$. Now on the set $E_{2n}$,

$$\begin{aligned} &\left\| \frac{\mathbf{X}_m'\mathbf{X}_m}{n} - \mathbf{C}_{Xm} \right\| \leq 4\mathcal{M}_n \sqrt{\frac{\nu_m \log p}{cn}} \\ \implies &\left\| \frac{\mathbf{Z}_m'\mathbf{Z}_m}{n} - \mathbf{C}_{Zm} \right\| \leq 4c_1\mathcal{M}_n \sqrt{\frac{\nu_m \log p}{cn}}, \end{aligned}$$

for appropriate constant $c_1$ not depending on $n$, as columns of $\mathbf{Z}$ are linear combinations of columns of $\mathbf{X}$. Also from (S5.7),

$$\left\| \begin{bmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i1}}{n} & \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i2}}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{\xi}}_{i1}}{n} & \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{\xi}}_{i2}}{n} \end{bmatrix} - \mathbf{C}_{\begin{pmatrix} \tilde{\boldsymbol{\xi}}_{i1} \\ \tilde{\boldsymbol{\xi}}_{i2} \end{pmatrix}} \right\| \leq 2 \left\| \begin{bmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i1}}{n} & \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i2}}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{\xi}}_{i1}}{n} & \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{\xi}}_{i2}}{n} \end{bmatrix} - \mathbf{C}_{\begin{pmatrix} \tilde{\boldsymbol{\xi}}_{i1} \\ \tilde{\boldsymbol{\xi}}_{i2} \end{pmatrix}} \right\|_{\max} \leq 2\mathcal{M}_n \sqrt{\frac{\log p}{n}} \tag{S5.8}$$

Hence it follows that,

$$\left\| \frac{\tilde{\mathbf{X}}_m'\tilde{\mathbf{X}}_m}{n} - \mathbf{C}_{\tilde{X}m} \right\| \leq 4c_1\mathcal{M}_n \sqrt{\frac{\nu_m \log p}{cn}}$$

Thus, $\mathbb{P}(\tilde{E}_{22,n}) \to 1$ as $n \to \infty$. We will restrict ourselves to the event $\tilde{G}_{1,n} \bigcap \tilde{G}_{21,n} \bigcap \tilde{G}_{22,n} \bigcap \tilde{G}_{23,n} \bigcap \tilde{G}_{3,n} \bigcap E_{21,n} \bigcap \tilde{E}_{12,n} \bigcap \tilde{E}_{13,n} \bigcap \tilde{E}_{21,n} \bigcap \tilde{E}_{22,n}$ for analysing the behaviour of $B(\boldsymbol{m}_i, \boldsymbol{t}_i)$ while dealing with the $k_1$ high frequency terms of (S5.2). Next we analyse the behaviour of $B(\boldsymbol{m}_i, \boldsymbol{t}_i)$ under different scenarios in the following sequence of lemmas (Lemma 2 - Lemma 4).

**Lemma 2.** *If $\boldsymbol{m}_i \subset \boldsymbol{t}_i$ then there exists $N_1$ (not depending on $i$ and $\mathcal{G}_m$) such that for all $n \geq N_1$ we have $B(\boldsymbol{m}_i, \boldsymbol{t}_i) \leq q_n^{\frac{\nu_{ti}-\nu_{mi}}{2}}$*

**Lemma 3.** *If $\boldsymbol{t}_i \subset \boldsymbol{m}_i$ then there exists $N_2$ (not depending on $i$ and $\mathcal{G}_m$) such that for all $n \geq N_2$ we have $B(\boldsymbol{m}_i, \boldsymbol{t}_i) \leq q_n^{\left(\nu_{m_i} - \nu_{t_i}\right)/2}$*

**Lemma 4.** *If*

$$\boldsymbol{t}_i \neq \boldsymbol{m}_i$$

$$\boldsymbol{t}_i \not\subseteq \boldsymbol{m}_i$$

$$\boldsymbol{t}_i \not\supseteq \boldsymbol{m}_i$$

*then there exists $N_3$ (not depending on $i$ and $\mathcal{G}_m$) such that for all $n \geq N_3, B(\boldsymbol{m}_i, \boldsymbol{t}_i) \leq q_n^{\left(\nu_{m_i} - \nu_{t_i}\right)/2}$ if $\nu_{m_i} > \nu_{t_i}$, and $B(\boldsymbol{m}_i, \boldsymbol{t}_i) \leq q_n^{1/2}$ if $\nu_{m_i} \leq \nu_{t_i}$.*

We now present the proofs of the above lemmas.

**Proof of Lemma 2:**

We start with the term

$$\frac{\left|\frac{\mathbf{Z}'_{m_i}\mathbf{Z}_{m_i}}{n} + \frac{\mathbf{I}_{\nu_{m_i}}}{n\tau^2}\right|}{\left|\frac{\mathbf{Z}'_{t_i}\mathbf{Z}_{t_i}}{n} + \frac{\mathbf{I}_{\nu_{t_i}}}{n\tau^2}\right|} \frac{\left|\mathbf{R}_{m_i,11}\right|}{\left|\mathbf{R}_{t_i,11}\right|}$$

Considering $\mathbf{C}^{-1}\mathbf{V}\mathbf{C}'^{-1} = \frac{\mathbf{I}}{\tilde{c}\tau^2}$ for some constant $\tilde{c}$ (i.e. taking $\mathbf{V}$ to be diagonal), it can be shown that

$$\left|\frac{\mathbf{Z}'_i\mathbf{Z}_i}{n} + \frac{\mathbf{I}_{\nu_i}}{n\tau^2}\right|\left|\mathbf{R}_{i,11}\right| = \left|\tilde{\mathbf{X}}'_i\tilde{\mathbf{X}}_i + \frac{\mathbf{I}}{c_2 n\tau^2}\right|$$

since $\left(\mathbf{Z}_i'\mathbf{Z}_i + \frac{\mathbf{I}}{n\tau^2}\right)$ is a submatrix of $\left(\tilde{\mathbf{X}}_i'\tilde{\mathbf{X}}_i + \frac{\mathbf{I}}{c_2 n\tau^2}\right)$, where,

$$\begin{aligned} c_2 &= 1 \text{ for first } \nu_i \text{ rows of } \left(\tilde{\mathbf{X}}_i'\tilde{\mathbf{X}}_i + \frac{\mathbf{I}}{c_2 n\tau^2}\right) \\ &= \tilde{c} \text{ for } (\nu_i+1) \text{ and } (\nu_i+2) \text{ th row of } \left(\tilde{\mathbf{X}}_i'\tilde{\mathbf{X}}_i + \frac{\mathbf{I}}{c_2 n\tau^2}\right) \end{aligned}$$

Hence from now onwards, we will deal with the ratio,

$$\frac{\left|\tilde{\mathbf{X}}_{m_i}'\tilde{\mathbf{X}}_{m_i} + \frac{\mathbf{I}_{\nu_{m_i}}}{c_2 n\tau^2}\right|}{\left|\tilde{\mathbf{X}}_{t_i}'\tilde{\mathbf{X}}_{t_i} + \frac{\mathbf{I}_{\nu_{t_i}}}{c_2 n\tau^2}\right|}$$

Since

$$\boldsymbol{m}_i = \{i_1, \cdots, i_{\nu_{m_i}}\} \subset \boldsymbol{t}_i = \{j_1, \cdots, j_{\nu_{t_i}}\} \quad \text{and} \quad \nu_{m_i} < \nu_{t_i} \tag{$\dagger$}$$

the submatrix $\tilde{\mathbf{X}}_{m_i}$ is also a submatrix of $\tilde{\mathbf{X}}_{t_i}$ which implies

$$\frac{\left|\frac{\tilde{\mathbf{X}}_{t_i}'\tilde{\mathbf{X}}_{t_i}}{n} + \frac{\mathbf{I}_{\nu_{t_i}}}{c_2 n\tau^2}\right|}{\left|\frac{\tilde{\mathbf{X}}_{m_i}'\tilde{\mathbf{X}}_{m_i}}{n} + \frac{\mathbf{I}_{\nu_{m_i}}}{c_2 n\tau^2}\right|} = |\mathbf{U}_{\tilde{\mathbf{X}}}(t_i, m_i)|$$

where $\mathbf{U}_{\tilde{\mathbf{X}}}(t_i, m_i)$ is the schur complement of $\left(\frac{\tilde{\mathbf{X}}_{m_i}'\tilde{\mathbf{X}}_{m_i}}{n} + \frac{\mathbf{I}_{\nu_{m_i}}}{c_2 n\tau^2}\right)$, defined by

$$\mathbf{U}_{\tilde{\mathbf{X}}}(t_i, m_i) = \mathbf{D} - \mathbf{B}'\left(\frac{\tilde{\mathbf{X}}_{m_i}'\tilde{\mathbf{X}}_{m_i}}{n} + \frac{\mathbf{I}_{\nu_{m_i}}}{c_2 n\tau^2}\right)^{-1}\mathbf{B}$$

for appropriate sub-matrices $\mathbf{B}$ and $\mathbf{D}$ of $\left(\frac{\tilde{\mathbf{X}}_{t_i}'\tilde{\mathbf{X}}_{t_i}}{n} + \frac{\mathbf{I}_{\nu_{t_i}}}{c_2 n\tau^2}\right)$.

Note that, on $\tilde{E}_{22,n}$,

$$\left\|\frac{\tilde{\mathbf{X}}_{t_i}'\tilde{\mathbf{X}}_{t_i}}{n} - \mathbf{C}_{\tilde{X}t_i}\right\| \le 4c_1\mathcal{M}_n\sqrt{\frac{\nu_{t_i}\log p}{cn}} \le 4c_1\mathcal{M}_n\sqrt{\frac{b_n\log p}{cn}}$$

Then, using ($S5.5$) and similar arguments as in Ghosh et al. (2021), it can be shown that,

$$\frac{\left|\tilde{\mathbf{X}}'_{m_i}\tilde{\mathbf{X}}_{m_i} + \frac{\mathbf{I}_{\nu_{m_i}}}{n\tau^2}\right|}{\left|\tilde{\mathbf{X}}'_{t_i}\tilde{\mathbf{X}}_{t_i} + \frac{\mathbf{I}_{\nu_{t_i}}}{n\tau^2}\right|} \leq (2\lambda'_2)^{\frac{\nu_{t_i}-\nu_{m_i}}{2}}$$

Next, we note that the term $\left(\frac{\mathbf{U}_{\mathbf{R}_{m_i}}}{\mathbf{U}_{\mathbf{R}_{t_i}}}\right)^{-(\frac{n+\nu}{2})}$ can be written as,

$$\left(\frac{\mathbf{U}_{\mathbf{R}_{m_i}}}{\mathbf{U}_{\mathbf{R}_{t_i}}}\right)^{-(\frac{n+\nu}{2})} = \left(1 - \frac{\mathbf{U}_{\mathbf{R}_{m_i}} - \mathbf{U}_{\mathbf{R}_{t_i}}}{\mathbf{U}_{\mathbf{R}_{m_i}}}\right)^{(\frac{n+\nu}{2})} = \left(1 - \frac{\mathbf{U}_{\tilde{\mathbf{R}}_{m_i}} - \mathbf{U}_{\tilde{\mathbf{R}}_{t_i}} + o(1)}{\mathbf{U}_{\mathbf{R}_{m_i}}}\right)^{(\frac{n+\nu}{2})} \tag{S5.9}$$

This is because,

$$\begin{aligned}
\mathbf{U}_{\mathbf{R}_i} - \mathbf{U}_{\tilde{\mathbf{R}}_i} &= \frac{1}{(\mathbf{R}_i^{-1})_{22}} - \frac{1}{(\tilde{\mathbf{R}}_i^{-1})_{22}} \\
&= \frac{|(\mathbf{R}_i^{-1})_{22} - (\tilde{\mathbf{R}}_i^{-1})_{22}|}{(\mathbf{R}_i^{-1})_{22}(\tilde{\mathbf{R}}_i^{-1})_{22}} \\
&\leq \frac{\left\|\mathbf{R}_i^{-1} - \tilde{\mathbf{R}}_i^{-1}\right\|}{(\mathbf{R}_i^{-1})_{22}(\tilde{\mathbf{R}}_i^{-1})_{22}} \\
&\leq c\left\|\mathbf{R}_i^{-1}\right\|\left\|\tilde{\mathbf{R}}_i^{-1}\right\|\left\|\mathbf{R}_i - \tilde{\mathbf{R}}_i\right\| \text{ for some appropriate constant c.}
\end{aligned}$$

Note that the quantities $\left\|\mathbf{R}_i^{-1}\right\|$ and $\left\|\tilde{\mathbf{R}}_i^{-1}\right\|$ are bounded above since

$$\left\|\tilde{\mathbf{R}}_i^{-1}\right\| \leq \left\|\frac{1}{n}\tilde{\mathbf{Y}}'_{i,\mathrm{H}}\tilde{\mathbf{Y}}_{i,\mathrm{H}}\right\|$$

which is bounded as from (S5.8), $\left\|\begin{bmatrix} \frac{\tilde{\boldsymbol{\xi}}'_{i1}\tilde{\boldsymbol{\xi}}_{i1}}{n} & \frac{\tilde{\boldsymbol{\xi}}'_{i1}\tilde{\boldsymbol{\xi}}_{i2}}{n} \\ \frac{\tilde{\boldsymbol{\xi}}'_{i2}\tilde{\boldsymbol{\xi}}_{i1}}{n} & \frac{\tilde{\boldsymbol{\xi}}'_{i2}\tilde{\boldsymbol{\xi}}_{i2}}{n} \end{bmatrix}\right\| \leq 2\sigma_2$ and also $\frac{\tilde{\boldsymbol{y}}'_{i3}\tilde{\boldsymbol{y}}_{i3}}{n} \leq$

$c_A$ (see (S5.12)). Also, $\left\|\mathbf{R}_i - \tilde{\mathbf{R}}_i\right\|$ is $o(1)$ since

$$\begin{aligned}
\left\|\mathbf{R}_i - \tilde{\mathbf{R}}_i\right\| &= \left\|\frac{\tilde{\mathbf{Y}}'_{i,\mathrm{H}}\mathbf{Z}_i}{n}\left[\left(\frac{1}{n}\mathbf{Z}'_i\mathbf{Z}_i + \frac{\mathbf{I}}{n\tau^2}\right)^{-1} - \left(\frac{1}{n}\mathbf{Z}'_i\mathbf{Z}_i\right)^{-1}\right]\frac{\mathbf{Z}'_i\tilde{\mathbf{Y}}_{i,\mathrm{H}}}{n} + \frac{\mathbf{C}^{-1}\mathbf{V}\mathbf{C}'^{-1}}{n}\right\| \\
&\leq \left\|\frac{\tilde{\mathbf{Y}}'_{i,\mathrm{H}}\mathbf{Z}_i}{n}\right\|\left\|\left(\frac{1}{n}\mathbf{Z}'_i\mathbf{Z}_i + \frac{\mathbf{I}}{n\tau^2}\right)^{-1} - \left(\frac{1}{n}\mathbf{Z}'_i\mathbf{Z}_i\right)^{-1}\right\|\left\|\frac{\mathbf{Z}'_i\tilde{\mathbf{Y}}_{i,\mathrm{H}}}{n}\right\| + \left\|\frac{\mathbf{I}}{\tilde{c}n\tau^2}\right\| \\
&\leq \left\|\frac{\tilde{\mathbf{Y}}'_{i,\mathrm{H}}\mathbf{Z}_i}{n}\right\|\left\|\frac{1}{n}\mathbf{Z}'_i\mathbf{Z}_i + \frac{\mathbf{I}}{n\tau^2}\right\|\left\|\frac{1}{n}\mathbf{Z}'_i\mathbf{Z}_i\right\|\left\|\frac{\mathbf{I}}{n\tau^2}\right\|\left\|\frac{\mathbf{Z}'_i\tilde{\mathbf{Y}}_{i,\mathrm{H}}}{n}\right\| + \left\|\frac{\mathbf{I}}{\tilde{c}n\tau^2}\right\| \\
&\leq \left\|\frac{\tilde{\mathbf{Y}}'_{i,\mathrm{H}}\mathbf{Z}_i}{n}\right\|\left(\left\|\frac{1}{n}\mathbf{Z}'_i\mathbf{Z}_i\right\| + \frac{1}{n\tau^2}\right)\left\|\frac{1}{n}\mathbf{Z}'_i\mathbf{Z}_i\right\|\left(\frac{1}{n\tau^2}\right)\left\|\frac{\mathbf{Z}'_i\tilde{\mathbf{Y}}_{i,\mathrm{H}}}{n}\right\| + \left(\frac{1}{\tilde{c}n\tau^2}\right) \\
&\leq \left\|\frac{\tilde{\mathbf{Y}}'_{i,\mathrm{H}}\mathbf{Z}_i}{n}\right\|\left(2\tilde{\lambda}_2 + \frac{1}{n\tau^2}\right)\left(\frac{2\tilde{\lambda}_2}{n\tau^2}\right)\left\|\frac{\mathbf{Z}'_i\tilde{\mathbf{Y}}_{i,\mathrm{H}}}{n}\right\| + \left(\frac{1}{\tilde{c}n\tau^2}\right) \\
&= o(1) \text{ for large } n
\end{aligned}$$

This is because:

$$\begin{aligned}
\left\|\frac{\tilde{\mathbf{Y}}'_{i,\mathrm{H}}\mathbf{Z}_i}{n}\right\| &\leq \left\|\frac{\tilde{\boldsymbol{\xi}}'_{i1}\mathbf{Z}_i}{n}\right\| + \left\|\frac{\tilde{\boldsymbol{\xi}}'_{i2}\mathbf{Z}_i}{n}\right\| + \left\|\frac{\tilde{\boldsymbol{y}}'_{i3}\mathbf{Z}_i}{n}\right\| \\
&\overset{(i)}{\leq} \mathcal{M}_n\sqrt{\frac{b_n\log p^2}{n}} + \mathcal{M}_n\sqrt{\frac{b_n\log p^2}{n}} + \left\|\frac{\tilde{\boldsymbol{y}}'_{i3}\mathbf{Z}_i}{n}\right\| \\
&\overset{(ii)}{\leq} 2\mathcal{M}_n\sqrt{\frac{b_n\log p^2}{n}} + \left(c_3 + \mathcal{M}_n\sqrt{\frac{b_n\log p^2}{n}}\right)
\end{aligned}$$

for an appropriate constant $c_3$ not depending on $n$. Here $(i)$ follows from the definition of $\tilde{\mathcal{G}}_{21,n}$ and $\tilde{\mathcal{G}}_{22,n}$ and $(ii)$ follows from the following inequality on $\tilde{\mathcal{G}}_{1,n} \cap \tilde{\mathcal{G}}_{23,n}$

$$\begin{aligned}
\left\|\frac{\tilde{\boldsymbol{y}}'_{i3}\mathbf{Z}_i}{n}\right\| &= \|\boldsymbol{\delta}\|\|\mathbf{A}_{0,i.}\|\left\|\frac{\mathbf{Z}'_i\mathbf{Z}_i}{n}\right\| + \left\|\frac{\tilde{\boldsymbol{\xi}}'_{i3}\mathbf{Z}_i}{n}\right\| \\
&\leq (\underbrace{2\tilde{\lambda}_2\|\boldsymbol{\delta}\|\|\mathbf{A}'_{0,i.}\|}_{=c_3,\text{ say}}) + \mathcal{M}_n\sqrt{\frac{b_n\log p^2}{n}} \qquad \text{(S5.10)}
\end{aligned}$$

Here we claim that the rows of the true parameter matrix $\mathbf{W}_0$ (and hence $\mathbf{A}_0$) is bounded above in $\ell_2$ norm by a constant not depending on $n$. This follows from the stationarity and stability of the underlying process $\boldsymbol{y}^t$. Specifically, from (2.6) of the main paper

$$
\begin{aligned}
\boldsymbol{y}^t &= \mathbf{W}_0 \boldsymbol{y}^{t-1} + \boldsymbol{\varepsilon}^t \\
\implies \boldsymbol{y}^t &= \underbrace{\left[\begin{array}{c|c} \boldsymbol{\delta} \otimes \mathbf{A}_{11_0} & \boldsymbol{\delta} \otimes \mathbf{A}_{12_0} \\ \hline \mathbf{A}_{21_0} & \mathbf{A}_{22_0} \end{array}\right]}_{=\mathbf{W}_{10},\text{ say}} \boldsymbol{z}^{t-1} + \boldsymbol{\varepsilon}^t \\
\implies \Gamma_X(0) &= \mathbf{W}_{10} \mathbf{C}_Z \mathbf{W}_{10}' + \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon},0n} \qquad \text{(S5.11)}
\end{aligned}
$$

By pre and post multiplication by unit vector $\mathbf{e}_i$ and using the Assumption A2 , we can justify that any column of true $\mathbf{W}_{10}'$ and hence $\mathbf{A}_{0_{i,\cdot}}'$ is bounded above in $\ell_2$ norm by a constant not depending on $n$.

Now, for all large n, $\mathbf{U}_{\mathbf{R}_{m_i}}$ is $O(1)$. This is because,

$$
\begin{aligned}
& \mathbf{R}_{mi} \preceq \frac{\tilde{\mathbf{Y}}_{i,\mathrm{H}}'\tilde{\mathbf{Y}}_{i,\mathrm{H}}}{n} + \frac{\mathbf{I}}{\tilde{c}n\tau^2} \\
\Rightarrow \quad & {\mathbf{R}_{mi}}^{-1} \succeq \left(\frac{\tilde{\mathbf{Y}}_{i,\mathrm{H}}'\tilde{\mathbf{Y}}_{i,\mathrm{H}}}{n} + \frac{\mathbf{I}}{\tilde{c}n\tau^2}\right)^{-1} \\
\Rightarrow \quad & \left({\mathbf{R}_{mi}}^{-1}\right)_{22} \geq \left(\left(\frac{\tilde{\mathbf{Y}}_{i,\mathrm{H}}'\tilde{\mathbf{Y}}_{i,\mathrm{H}}}{n} + \frac{\mathbf{I}}{\tilde{c}n\tau^2}\right)^{-1}\right)_{22} \\
\Rightarrow \quad & (\mathbf{U}_{\mathbf{R}_{mi}})^{-1} \geq \left(\mathbf{U}_{\begin{pmatrix} \boldsymbol{\xi}_{i1} \\ \boldsymbol{\xi}_{i2} \end{pmatrix}}\right)^{-1} \text{ where } \mathbf{U}_{\begin{pmatrix} \boldsymbol{\xi}_{i1} \\ \boldsymbol{\xi}_{i2} \end{pmatrix}} \text{ is the schur complement of} \\
& \qquad\qquad\qquad\qquad\qquad\qquad \begin{bmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i1}}{n} + \frac{1}{\tilde{c}n\tau^2} & \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i2}}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{\xi}}_{i1}}{n} & \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{\xi}}_{i2}}{n} + \frac{1}{\tilde{c}n\tau^2} \end{bmatrix} \\
\Rightarrow \quad & \mathbf{U}_{\mathbf{R}_{mi}} \leq \mathbf{U}_{\begin{pmatrix} \boldsymbol{\xi}_{i1} \\ \boldsymbol{\xi}_{i2} \end{pmatrix}} \leq \left(\frac{\tilde{\boldsymbol{y}}_{i3}'\tilde{\boldsymbol{y}}_{i3}}{n} + \frac{1}{\tilde{c}n\tau^2}\right)
\end{aligned}
$$

Now, $\frac{\tilde{\boldsymbol{y}}_{i3}'\tilde{\boldsymbol{y}}_{i3}}{n}$ can be further upper bounded by a constant $c_A$ (not depending on $n$) as follows:

$$
\begin{aligned}
\frac{\tilde{\boldsymbol{y}}_{i3}'\tilde{\boldsymbol{y}}_{i3}}{n} &= \frac{(\mathbf{A}_{0,i\cdot}', \mathbf{O})'\tilde{\mathbf{X}}_{t_i}'\tilde{\mathbf{X}}_{t_i}(\mathbf{A}_{0,i\cdot}', \mathbf{O})}{n} + \frac{2\mathbf{A}_{0,i\cdot}\mathbf{Z}_{t_i}'\tilde{\boldsymbol{\xi}}_{i3}}{n} + \frac{\tilde{\boldsymbol{\xi}}_{i3}'\tilde{\boldsymbol{\xi}}_{i3}}{n} \\
&\leq 2\lambda_2'\|\mathbf{A}_{0,i\cdot}'\|^2 + 2\left\|\mathbf{A}_{0,i\cdot}\right\|\left\|\frac{\tilde{\boldsymbol{\xi}}_{i3}'\mathbf{Z}_{t_i}}{n}\right\| + \frac{\tilde{\boldsymbol{\xi}}_{i3}'\tilde{\boldsymbol{\xi}}_{i3}}{n} \\
&\overset{(i)}{\leq} 2\lambda_2'\|\mathbf{A}_{0,i,\cdot}'\|^2 + 2\left\|\mathbf{A}_{0,i\cdot}'\right\|\mathcal{M}_n\sqrt{\frac{b_n\log p^2}{n}} + \frac{3\sigma_{32}}{2} \\
&\leq c_A
\end{aligned} \tag{S5.12}
$$

where $(i)$ follows from the definition of $\tilde{G}_{23,n}$ and $\tilde{E}_{13,n}$ along with the following inequality

$$\text{for each } i, \qquad \left\| \frac{\tilde{\boldsymbol{\xi}}_{i3}' \mathbf{Z}_{t_i}}{n} \right\| \leq \mathcal{M}_n \sqrt{\nu_{t_i} \frac{\log p^2}{n}} \leq \mathcal{M}_n \sqrt{\frac{b_n \log p^2}{n}}.$$

Now,

$$\begin{aligned}
&\mathbf{U}_{\tilde{\mathbf{R}}_{m_i}} - \mathbf{U}_{\tilde{\mathbf{R}}_{t_i}} \\
&= \frac{\tilde{\boldsymbol{y}}_{i3}'(\mathbf{P}_{t_i} - \mathbf{P}_{m_i})\tilde{\boldsymbol{y}}_{i3}}{n} \\
&= \frac{\tilde{\boldsymbol{\xi}}_{i3}'(\mathbf{P}_{t_i} - \mathbf{P}_{m_i})\tilde{\boldsymbol{\xi}}_{i3}}{n} + \frac{\mathbf{A}_{t_i,\cdot}\mathbf{Z}_{t_i}'(\mathbf{P}_{t_i} - \mathbf{P}_{m_i})\mathbf{Z}_{t_i}\mathbf{A}_{i\cdot}'}{n} + 2\frac{\mathbf{A}_{t_i,\cdot}\mathbf{Z}_{t_i}'(\mathbf{P}_{t_i} - \mathbf{P}_{m_i})\tilde{\boldsymbol{\xi}}_{i3}}{n}
\end{aligned} \tag{S5.13}$$

where, $\mathbf{P}_i : \frac{\tilde{\mathbf{x}}_i \left(\tilde{\mathbf{x}}_i'\tilde{\mathbf{x}}_{i}/n\right)^{-1} \tilde{\mathbf{x}}_i'}{n}$. Similar to Ghosh et al. (2021), our objective is to show that for all sufficiently large $n$, $\mathbf{U}_{\tilde{\mathbf{R}}_{m_i}} - \mathbf{U}_{\tilde{\mathbf{R}}_{t_i}} \geq \tilde{\lambda}_1(\nu_{t_i} - \nu_{m_i})s_n^2$. We justify our claim by analysing the terms in (S5.13) individually. Note that, $\mathbf{P}_{t_i}$ and $\mathbf{P}_{m_i}$ are projection on $\mathcal{C}\left(\mathcal{A} \bigcup \mathcal{B}\right)$ and $\mathcal{C}\left(\mathcal{B}\right)$ respectively where $\mathcal{A} = (\mathbf{Z}_{t_i - m_i})$ and $\mathcal{B} = \left(\mathbf{Z}_{m_i}, \tilde{\boldsymbol{\xi}}_{i1}, \tilde{\boldsymbol{\xi}}_{i2}\right)$. Hence $\mathbf{P}_{t_i} - \mathbf{P}_{m_i}$ is a projection on $\mathcal{C}\left(\mathcal{A} - \mathbf{P}_{\mathcal{B}}\mathcal{A}\right)$. On $\tilde{G}_{21,n} \cap \tilde{G}_{22,n} \cap \tilde{G}_{23,n}$ we have,

$$\left\| \frac{\tilde{\boldsymbol{\xi}}_{i3}' \mathbf{Z}_{t_i - m_i}}{n} \right\|^2 \leq \mathcal{M}_n^2 (\nu_{t_i} - \nu_{m_i}) \frac{\log p^2}{n} \tag{S5.14}$$

$$\left\| \frac{\tilde{\boldsymbol{\xi}}_{i1}' \mathbf{Z}_{t_i - m_i}}{n} \right\|^2 \leq \mathcal{M}_n^2 (\nu_{t_i} - \nu_{m_i}) \frac{\log p^2}{n} \tag{S5.15}$$

$$\left\| \frac{\tilde{\boldsymbol{\xi}}_{i2}' \mathbf{Z}_{t_i - m_i}}{n} \right\|^2 \leq \mathcal{M}_n^2 (\nu_{t_i} - \nu_{m_i}) \frac{\log p^2}{n} \tag{S5.16}$$

Also note that,

$$\mathbf{P}_{t_i} - \mathbf{P}_{m_i} = (\mathcal{A} - \mathbf{P}_{\mathcal{B}}\mathcal{A})\,\lambda_{\max}\left[(\mathcal{A} - \mathbf{P}_{\mathcal{B}}\mathcal{A})'\,(\mathcal{A} - \mathbf{P}_{\mathcal{B}}\mathcal{A})\right]^{-1}(\mathcal{A} - \mathbf{P}_{\mathcal{B}}\mathcal{A})$$
$$= \|\mathcal{A} - \mathbf{P}_{\mathcal{B}}\mathcal{A}\|_2^2\,\lambda_{\max}\left(\mathcal{A}'\,(\mathbf{I} - \mathbf{P}_{\mathcal{B}})\,\mathcal{A}\right)$$

Now using (S5.14), (S5.15) and (S5.16) and by some straight-forward calculation it can be shown that,

$$\frac{\tilde{\boldsymbol{\xi}}_{i3}'(\mathbf{P}_{t_i} - \mathbf{P}_{m_i})\tilde{\boldsymbol{\xi}}_{i3}}{n} \leq c_4(\nu_{t_i} - \nu_{m_i})\frac{\log p^2}{n} \text{ for some appropriate constant } c_4. \tag{S5.17}$$

Using the same line of arguments as in Ghosh et al. (2021), it can be shown that, $\frac{\mathbf{A}_{t_i,\cdot}\mathbf{Z}_{t_i}'(\mathbf{P}_{t_i}-\mathbf{P}_{m_i})\mathbf{Z}_{t_i}\mathbf{A}_{t_i,\cdot}'}{n} \geq \frac{\tilde{\lambda}_1}{2}(\nu_{t_i} - \nu_{m_i})s_n^2$ and $\left|\frac{\mathbf{A}_{t_i,\cdot}\mathbf{Z}_{t_i}'(\mathbf{P}_{t_i}-\mathbf{P}_{m_i})\tilde{\boldsymbol{\xi}}_{i3}}{n}\right|$ is bounded above, which implies that,

$$\mathbf{U}_{\tilde{\mathbf{R}}_{m_i}} - \mathbf{U}_{\tilde{\mathbf{R}}_{t_i}} \geq \frac{\tilde{\lambda}_1}{4}(\nu_{t_i} - \nu_{m_i})s_n^2.$$

Thus it can be shown using similar arguments as in Ghosh et al. (2021) that there exists $N_1$ such that for all $n \geq N_1$, we have

$$B(\boldsymbol{m}_i, \boldsymbol{t}_i) \leq q_n^{(\nu_{t_i} - \nu_{m_i})/2}$$

**Proof of Lemma 3:**

Since $\boldsymbol{t}_i$ (the set of neighbors of $i$ in the true model) is contained in $\boldsymbol{m}_i$ (the set of neighbors of $i$ in the candidate model), it can be shown that(using similar arguments as in Ghosh et al. (2021)),

$$\frac{\left|\frac{\tilde{\mathbf{X}}'_{m_i}\tilde{\mathbf{X}}_{m_i}}{n}+\frac{\mathbf{I}_{\nu_{m_i}}}{nc_2\tau^2}\right|^{-1/2}}{\left|\frac{\tilde{\mathbf{X}}'_{t_i}\tilde{\mathbf{X}}_{t_i}}{n}+\frac{\mathbf{I}_{\nu_{t_i}}}{nc_2\tau^2}\right|^{-1/2}} \leq \left(\tau\sqrt{c_2 n}\right)^{(\nu_{m_i}-\nu_{t_i})}.$$

Therefore,

$$\begin{aligned}
B(\boldsymbol{m}_i,\boldsymbol{t}_i) &= \left(\frac{2q_{\nu_{m_i}}}{\tau\sqrt{n}}\right)^{(\nu_{m_i}-\nu_{t_i})} \frac{\left|\frac{\tilde{\mathbf{X}}'_{m_i}\tilde{\mathbf{X}}_{m_i}}{n}+\frac{\mathbf{I}_{\nu_{m_i}}}{nc_2\tau^2}\right|^{-1/2}}{\left|\frac{\tilde{\mathbf{X}}'_{t_i}\tilde{\mathbf{X}}_{t_i}}{n}+\frac{\mathbf{I}_{\nu_{t_i}}}{nc_2\tau^2}\right|^{-1/2}} \left(\frac{\mathbf{U}_{\mathbf{R}_{m_i}}}{\mathbf{U}_{\mathbf{R}_{t_i}}}\right)^{-(\frac{n+\nu}{2})} \\
&\leq \left(2q_{\nu_{m_i}}\sqrt{c_2}\right)^{(\nu_{m_i}-\nu_{t_i})}\left(1+\frac{\mathbf{U}_{\mathbf{R}_{t_i}}-\mathbf{U}_{\mathbf{R}_{m_i}}}{\mathbf{U}_{\mathbf{R}_{m_i}}}\right)^{(\frac{n+\nu}{2})} \\
&= \left(2q_{\nu_{m_i}}\sqrt{c_2}\right)^{(\nu_{m_i}-\nu_{t_i})}\left(1+\frac{\mathbf{U}_{\tilde{\mathbf{R}}_{t_i}}-\mathbf{U}_{\tilde{\mathbf{R}}_{m_i}}+o(1)}{\mathbf{U}_{\tilde{\mathbf{R}}_{m_i}}+o(1)}\right)^{(\frac{n+\nu}{2})} \\
&= \left(2q_{\nu_{m_i}}\sqrt{c_2}\right)^{(\nu_{m_i}-\nu_{t_i})}\left(1+\frac{\tilde{\boldsymbol{y}}'_{i3}\left(\mathbf{P}_{m_i}-\mathbf{P}_{t_i}\right)\tilde{\boldsymbol{y}}_{i3}/n+o(1)}{\tilde{\boldsymbol{y}}'_{i3}(\mathbf{I}-\mathbf{P}_{m_i})\tilde{\boldsymbol{y}}_{i3}/n+o(1)}\right)^{(\frac{n+\nu}{2})} \\
&\stackrel{(i)}{=} \left(2q_{\nu_{m_i}}\sqrt{c_2}\right)^{(\nu_{m_i}-\nu_{t_i})}\left(1+\frac{\tilde{\boldsymbol{\xi}}'_{i3}(\mathbf{P}_{m_i}-\mathbf{P}_{t_i})\tilde{\boldsymbol{\xi}}_{i3}/n+o(1)}{\tilde{\boldsymbol{\xi}}'_{i3}(\mathbf{I}-\mathbf{P}_{m_i})\tilde{\boldsymbol{\xi}}_{i3}/n+o(1)}\right)^{(\frac{n+\nu}{2})}
\end{aligned}$$

Here (i) follows by substituting the true model (S5.3). We now analyse $B(\boldsymbol{m}_i,\boldsymbol{t}_i)$ in three different cases depending on the size of $\nu_{m_i}$.

**Case I:** $2\nu_{t_i} \leq \nu_{m_i}$.

Let us define $\mathbf{P}^{\star}_{m_i}$ to be the projection matrix on the column space of $\mathbf{Z}_{m_i}$ after removing the effect of $\tilde{\boldsymbol{\xi}}_{i1}$ and $\tilde{\boldsymbol{\xi}}_{i2}$. Let $\tilde{\boldsymbol{\xi}}^{\star}_{i3}$ be the LS residual after regressing $\tilde{\boldsymbol{\xi}}_{i3}$ on $\left[\tilde{\boldsymbol{\xi}}_{i1},\tilde{\boldsymbol{\xi}}_{i2}\right]$.

Then,

$$\begin{aligned}\frac{\tilde{\boldsymbol{\xi}}_{i3}'(\mathbf{P}_{m_i}-\mathbf{P}_{t_i})\tilde{\boldsymbol{\xi}}_{i3}}{n} &= \frac{\tilde{\boldsymbol{\xi}}_{i3}'\left(\mathbf{I}-\mathbf{P}_{t_i}\right)\tilde{\boldsymbol{\xi}}_{i3}-\tilde{\boldsymbol{\xi}}_{i3}'\left(\mathbf{I}-\mathbf{P}_{m_i}\right)\tilde{\boldsymbol{\xi}}_{i3}}{n}\\ &= \frac{\tilde{\boldsymbol{\xi}}_{i3}^{\star}{}'\left(\mathbf{I}-\mathbf{P}_{t_i}^{\star}\right)\tilde{\boldsymbol{\xi}}_{i3}^{\star}-\tilde{\boldsymbol{\xi}}_{i3}^{\star}{}'\left(\mathbf{I}-\mathbf{P}_{m_i}^{\star}\right)\tilde{\boldsymbol{\xi}}_{i3}}{n}\\ &= \frac{\tilde{\boldsymbol{\xi}}_{i3}^{\star}{}'(\mathbf{P}_{m_i}^{\star}-\mathbf{P}_{t_i}^{\star})\tilde{\boldsymbol{\xi}}_{i3}^{\star}}{n} \end{aligned} \tag{S5.18}$$

Note that, $(\mathbf{P}_{m_i}^{\star}-\mathbf{P}_{t_i}^{\star})$ turns out to be the projection matrix on $\mathcal{C}\left(\mathbf{A}-\mathbf{P_B}\mathbf{A}\right)$ where $\mathbf{A}=[\mathbf{Z}_{m_i-t_i}]$ and $\mathbf{B}=[\mathbf{Z}_{t_i}]$ after removing the effect of $\tilde{\boldsymbol{\xi}}_{i1},\tilde{\boldsymbol{\xi}}_{i2}$ from both $\mathbf{Z}_{m_i}$ and $\mathbf{Z}_{t_i}$.

Then,

$$\frac{\tilde{\boldsymbol{\xi}}_{i3}^{\star}{}'(\mathbf{P}_{m_i}^{\star}-\mathbf{P}_{t_i}^{\star})\tilde{\boldsymbol{\xi}}_{i3}^{\star}}{n} \le \left\|\frac{\tilde{\boldsymbol{\xi}}_{i3}^{\star}{}'\left(\mathbf{A}-\mathbf{P_B}\mathbf{A}\right)}{n}\right\|^2 \lambda_{\max}\left(\frac{\mathbf{A}'\left(\mathbf{I}-\mathbf{P_B}\right)\mathbf{A}}{n}\right)^{-1}$$

Now,

$$\begin{aligned}\left\|\frac{\tilde{\boldsymbol{\xi}}_{i3}^{\star}{}'\left(\mathbf{A}-\mathbf{P_B}\mathbf{A}\right)}{n}\right\|^2 &= \frac{\tilde{\boldsymbol{\xi}}_{i3}^{\star}{}'\mathbf{A}\left(\mathbf{I}-\mathbf{P_B}\right)\mathbf{A}'\tilde{\boldsymbol{\xi}}_{i3}^{\star}{}'}{n^2}\\ &\le \frac{\tilde{\boldsymbol{\xi}}_{i3}^{\star}{}'\mathbf{A}\mathbf{A}'\tilde{\boldsymbol{\xi}}_{i3}^{\star}}{n^2}\\ &= \left\|\frac{\tilde{\boldsymbol{\xi}}_{i3}^{\star}{}'\mathbf{A}}{n}\right\|^2\\ &= \left\|\frac{\tilde{\boldsymbol{\xi}}_{i3}^{\star}{}'\mathbf{Z}_{m_i-t_i}}{n}\right\|^2\\ &\le \mathcal{M}_n^2(\nu_{m_i}-\nu_{t_i})\frac{\log p^2}{n}\end{aligned}$$

Also,

$$\lambda_{\max}\left(\frac{\mathbf{A}'\left(\mathbf{I}-\mathbf{P_B}\right)\mathbf{A}}{n}\right)^{-1} = \frac{1}{\lambda_{\min}\left(\frac{\mathbf{A}'(\mathbf{I}-\mathbf{P_B})\mathbf{A}}{n}\right)} \leq \frac{2}{\tilde{\lambda}_1}$$

using the fact that $\mathbf{A}'\mathbf{P_B}\mathbf{A}$ is of smaller order compared to $\mathbf{A}'\mathbf{A}$. Now,

$$\frac{\tilde{\boldsymbol{\xi}}_{i3}'\mathbf{P}_{m_i}\tilde{\boldsymbol{\xi}}_{i3}}{n} = \left(\frac{\tilde{\boldsymbol{\xi}}_{i3}'\tilde{\boldsymbol{\xi}}_{i1}}{n}, \frac{\tilde{\boldsymbol{\xi}}_{i3}'\tilde{\boldsymbol{\xi}}_{i2}}{n}, \frac{\tilde{\boldsymbol{\xi}}_{i3}'\mathbf{Z}_{m_i}}{n}\right) \begin{bmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i1}}{n} & \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i2}}{n} & \frac{\tilde{\boldsymbol{\xi}}_{i1}'\mathbf{Z}_{m_i}}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{\xi}}_{i1}}{n} & \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{\xi}}_{i2}}{n} & \frac{\tilde{\boldsymbol{\xi}}_{i2}'\mathbf{Z}_{m_i}}{n} \\ \frac{\mathbf{Z}_{m_i}'\tilde{\boldsymbol{\xi}}_{i1}}{n} & \frac{\mathbf{Z}_{m_i}'\tilde{\boldsymbol{\xi}}_{i2}}{n} & \frac{\mathbf{Z}_{m_i}'\mathbf{Z}_{m_i}}{n} \end{bmatrix}^{-1} \begin{pmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i3}}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{\xi}}_{i3}}{n} \\ \frac{\mathbf{Z}_{m_i}{}'\tilde{\boldsymbol{\xi}}_{i3}}{n} \end{pmatrix}$$

Let us denote, $\mathbf{C} = \begin{bmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i1}}{n} & \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i2}}{n} & \frac{\tilde{\boldsymbol{\xi}}_{i1}'\mathbf{Z}_{m_i}}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{\xi}}_{i1}}{n} & \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{\xi}}_{i2}}{n} & \frac{\tilde{\boldsymbol{\xi}}_{i2}'\mathbf{Z}_{m_i}}{n} \\ \frac{\mathbf{Z}_{m_i}'\tilde{\boldsymbol{\xi}}_{i1}}{n} & \frac{\mathbf{Z}_{m_i}'\tilde{\boldsymbol{\xi}}_{i2}}{n} & \frac{\mathbf{Z}_{m_i}'\mathbf{Z}_{m_i}}{n} \end{bmatrix} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{12}' & \mathbf{C}_{22} \end{bmatrix}$ where,

$$\mathbf{C}_{11} = \begin{bmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i1}}{n} & \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i2}}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{\xi}}_{i1}}{n} & \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{\xi}}_{i2}}{n} \end{bmatrix}, \mathbf{C}_{12} = \begin{bmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\mathbf{Z}_{m_i}}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\mathbf{Z}_{m_i}}{n} \end{bmatrix}, \mathbf{C}_{22} = \frac{\mathbf{Z}_{m_i}'\mathbf{Z}_{m_i}}{n} \tag{S5.19}$$

Then by straightforward calculations,

$$\begin{aligned} &\frac{\tilde{\boldsymbol{\xi}}_{i3}'\mathbf{P}_{m_i}\tilde{\boldsymbol{\xi}}_{i3}}{n} \\ = &\left(\frac{\tilde{\boldsymbol{\xi}}_{i3}'\tilde{\boldsymbol{\xi}}_{i1}}{n}, \frac{\tilde{\boldsymbol{\xi}}_{i3}'\tilde{\boldsymbol{\xi}}_{i2}}{n}\right)\mathbf{C}_{11}^{-1}\begin{pmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i3}}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{\xi}}_{i3}}{n} \end{pmatrix} + \left(\frac{\tilde{\boldsymbol{\xi}}_{i3}'\tilde{\boldsymbol{\xi}}_{i1}}{n}, \frac{\tilde{\boldsymbol{\xi}}_{i3}'\tilde{\boldsymbol{\xi}}_{i2}}{n}\right)\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\mathbf{F}^{-1}\mathbf{C}_{12}'\mathbf{C}_{11}^{-1}\begin{pmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i3}}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{\xi}}_{i3}}{n} \end{pmatrix} - \\ &\frac{\tilde{\boldsymbol{\xi}}_{i3}'\mathbf{Z}_{m_i}}{n}\ \mathbf{F}^{-1}\mathbf{C}_{12}'\mathbf{C}_{11}^{-1}\begin{pmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i3}}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{\xi}}_{i3}}{n} \end{pmatrix} - \left(\frac{\tilde{\boldsymbol{\xi}}_{i3}'\boldsymbol{\xi}_{i1}}{n}, \frac{\tilde{\boldsymbol{\xi}}_{i3}'\tilde{\boldsymbol{\xi}}_{i2}}{n}\right)\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\mathbf{F}^{-1}\ \frac{\mathbf{Z}_{m_i}'\tilde{\boldsymbol{\xi}}_{i3}}{n} + \\ &\qquad\qquad \frac{\tilde{\boldsymbol{\xi}}_{i3}'\mathbf{Z}_{m_i}}{n}\ \mathbf{F}^{-1}\ \frac{\mathbf{Z}_{m_i}'\tilde{\boldsymbol{\xi}}_{i3}}{n} \end{aligned} \tag{S5.20}$$

where, $\mathbf{F} = \mathbf{C}_{22} - \mathbf{C}'_{12}\mathbf{C}_{11}^{-1}\mathbf{C}_{12}$

We claim that $\left(\frac{\tilde{\boldsymbol{\xi}}'_{i3}\tilde{\boldsymbol{\xi}}_{i1}}{n}, \frac{\tilde{\boldsymbol{\xi}}'_{i3}\tilde{\boldsymbol{\xi}}_{i2}}{n}\right)\mathbf{C}_{11}^{-1}\begin{pmatrix}\frac{\tilde{\boldsymbol{\xi}}'_{i1}\tilde{\boldsymbol{\xi}}_{i3}}{n}\\ \frac{\tilde{\boldsymbol{\xi}}'_{i2}\tilde{\boldsymbol{\xi}}_{i3}}{n}\end{pmatrix}$ is the dominating term which is bounded above and the other terms are of smaller order compared to it. We will justify our claim by analysing each term separately.

First consider the second term in (S5.20), $\left(\frac{\tilde{\boldsymbol{\xi}}'_{i3}\tilde{\boldsymbol{\xi}}_{i1}}{n}, \frac{\tilde{\boldsymbol{\xi}}'_{i3}\tilde{\boldsymbol{\xi}}_{i2}}{n}\right)\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\mathbf{F}^{-1}\mathbf{C}'_{12}\mathbf{C}_{11}^{-1}\begin{pmatrix}\frac{\tilde{\boldsymbol{\xi}}'_{i1}\tilde{\boldsymbol{\xi}}_{i3}}{n}\\ \frac{\tilde{\boldsymbol{\xi}}'_{i2}\tilde{\boldsymbol{\xi}}_{i3}}{n}\end{pmatrix}$.

Note that,

$$\begin{aligned}
&\left(\frac{\tilde{\boldsymbol{\xi}}'_{i3}\tilde{\boldsymbol{\xi}}_{i1}}{n}, \frac{\tilde{\boldsymbol{\xi}}'_{i3}\tilde{\boldsymbol{\xi}}_{i2}}{n}\right)\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\mathbf{F}^{-1}\mathbf{C}'_{12}\mathbf{C}_{11}^{-1}\begin{pmatrix}\frac{\tilde{\boldsymbol{\xi}}'_{i1}\tilde{\boldsymbol{\xi}}_{i3}}{n}\\ \frac{\tilde{\boldsymbol{\xi}}'_{i2}\tilde{\boldsymbol{\xi}}_{i3}}{n}\end{pmatrix}\\
&\leq \left\|\begin{pmatrix}\frac{\tilde{\boldsymbol{\xi}}'_{i1}\tilde{\boldsymbol{\xi}}_{i3}}{n}\\ \frac{\tilde{\boldsymbol{\xi}}'_{i2}\tilde{\boldsymbol{\xi}}_{i3}}{n}\end{pmatrix}\right\|^2 \left\|\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\mathbf{F}^{-1}\mathbf{C}'_{12}\mathbf{C}_{11}^{-1}\right\|\\
&\leq \left[\frac{\left(\tilde{\boldsymbol{\xi}}'_{i1}\tilde{\boldsymbol{\xi}}'_{i1}\right)}{n}\frac{\left(\tilde{\boldsymbol{\xi}}_{i3}{}'\tilde{\boldsymbol{\xi}}_{i3}\right)}{n} + \frac{\left(\tilde{\boldsymbol{\xi}}'_{i2}\tilde{\boldsymbol{\xi}}_{i2}\right)}{n}\frac{\left(\tilde{\boldsymbol{\xi}}_{i3}{}'\tilde{\boldsymbol{\xi}}_{i3}\right)}{n}\right]\left\|\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\mathbf{F}^{-1}\mathbf{C}'_{12}\mathbf{C}_{11}^{-1}\right\|\\
&\leq c_5\left\|\mathbf{C}_{11}^{-1}\right\|^2\left\|\mathbf{C}_{12}\right\|\left\|\mathbf{C}'_{12}\right\|\left\|\mathbf{F}^{-1}\right\| \qquad \text{(S5.21)}
\end{aligned}$$

for an appropriate constant $c_5$ not depending on $n$, on the event $\tilde{E}_{11,n} \cap \tilde{E}_{12,n} \cap \tilde{E}_{13,n}$. Now from (S5.4) and (S5.8), we know $\left\|\mathbf{C}_{11}^{-1}\right\| \leq \frac{2}{\sigma_1}$. Also,

$$\|\mathbf{C}_{12}\| = \|\mathbf{C}'_{12}\| \leq \sqrt{2\nu_{m_i}}\,\|\mathbf{C}_{12}\|_{\max} \leq \sqrt{2\nu_{m_i}}\mathcal{M}_n\sqrt{\frac{\log p^2}{n}} \qquad \text{(S5.22)}$$

Now consider $\left\|\mathbf{F}^{-1}\right\|$. Note that, $\mathbf{C}'_{12}\mathbf{C}_{11}^{-1}\mathbf{C}_{12}$ is of smaller order compared to $\mathbf{C}_{22}$ since,

$$\begin{aligned}\left\|\mathbf{C}_{12}'\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\right\|_{\max} &\leq \left\|\mathbf{C}_{12}'\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\right\| \\ &\leq \left\|\mathbf{C}_{12}\right\|^2 \left\|\mathbf{C}_{11}^{-1}\right\| \\ &\leq \frac{4}{\sigma_1}\nu_{mi}\mathcal{M}_n^2\frac{\log p^2}{n}\end{aligned}$$

Hence,

$$\mathbf{F} \approx \mathbf{C}_{22} = \frac{\mathbf{Z}_{m_i}'\mathbf{Z}_{m_i}}{n} \text{ and we can have, } \left\|\mathbf{F}^{-1}\right\| \leq \frac{2}{\tilde{\lambda}_1} \tag{S5.23}$$

where '≈' is used to denote that two terms are of same order. Thus, we obtain

$$\left(\frac{\tilde{\boldsymbol{\xi}}_{i3}'\tilde{\boldsymbol{\xi}}_{i1}}{n}, \frac{\tilde{\boldsymbol{\xi}}_{i3}'\tilde{\boldsymbol{\xi}}_{i2}}{n}\right)\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\mathbf{F}^{-1}\mathbf{C}_{12}'\mathbf{C}_{11}^{-1}\begin{pmatrix}\frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i3}}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{\xi}}_{i3}}{n}\end{pmatrix} \leq \frac{16c_5}{\sigma_1^2\tilde{\lambda}_1}\nu_{mi}\mathcal{M}_n^2\frac{\log p^2}{n}$$

which is o(1) for large $n$. Similarly,

$$
\begin{aligned}
& \left\| \left( \frac{\tilde{\boldsymbol{\xi}}_{i3}' \tilde{\boldsymbol{\xi}}_{i1}}{n}, \frac{\tilde{\boldsymbol{\xi}}_{i3}' \tilde{\boldsymbol{\xi}}_{i2}}{n} \right) \mathbf{C}_{11}^{-1} \mathbf{C}_{12} \mathbf{F}^{-1} \ \frac{\mathbf{Z}_{m_i}' \tilde{\boldsymbol{\xi}}_{i3}}{n} \right\|_{\max} \\
= \ & \left\| \frac{\tilde{\boldsymbol{\xi}}_{i3}' \mathbf{Z}_{m_i}}{n} \ \mathbf{F}^{-1} \mathbf{C}_{12}' \mathbf{C}_{11}^{-1} \begin{pmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}' \tilde{\boldsymbol{\xi}}_{i3}}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}' \tilde{\boldsymbol{\xi}}_{i3}}{n} \end{pmatrix} \right\|_{\max} \\
\leq \ & \left\| \frac{\tilde{\boldsymbol{\xi}}_{i3}' \mathbf{Z}_{m_i}}{n} \ \mathbf{F}^{-1} \mathbf{C}_{12}' \mathbf{C}_{11}^{-1} \begin{pmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}' \tilde{\boldsymbol{\xi}}_{i3}}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}' \tilde{\boldsymbol{\xi}}_{i3}}{n} \end{pmatrix} \right\| \\
\leq \ & \left\| \frac{\tilde{\boldsymbol{\xi}}_{i3}' \mathbf{Z}_{m_i}}{n} \right\| \left\| \mathbf{F}^{-1} \right\| \left\| \mathbf{C}_{12}' \right\| \left\| \mathbf{C}_{11}^{-1} \right\| \left\| \begin{pmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}' \tilde{\boldsymbol{\xi}}_{i3}}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}' \tilde{\boldsymbol{\xi}}_{i3}}{n} \end{pmatrix} \right\| \\
\leq \ & \frac{4\sqrt{2c_5}}{\sigma_1 \tilde{\lambda}_1} \left( \mathcal{M}_n^2 \nu_{m_i} \frac{\log p^2}{n} \right) \text{ using } (S5.21), (S5.22) \text{ and } (S5.23).
\end{aligned}
$$

which is o(1) for large $n$.

Also, $\frac{\tilde{\boldsymbol{\xi}}_{i3}' \mathbf{Z}_{m_i}}{n} \ \mathbf{F}^{-1} \ \frac{\mathbf{Z}_{m_i}' \tilde{\boldsymbol{\xi}}_{i3}}{n} \leq \left\| \frac{\tilde{\boldsymbol{\xi}}_{i3}' \mathbf{Z}_{m_i}}{n} \right\|^2 \left\| \mathbf{F}^{-1} \right\| \leq \frac{2}{\tilde{\lambda}_1} \left( \mathcal{M}_n^2 \nu_{m_i} \frac{\log p^2}{n} \right)$ which is o(1) for large $n$. Now, $\left( \frac{\tilde{\boldsymbol{\xi}}_{i3}' \tilde{\boldsymbol{\xi}}_{i1}}{n}, \frac{\tilde{\boldsymbol{\xi}}_{i3}' \tilde{\boldsymbol{\xi}}_{i2}}{n} \right) \mathbf{C}_{11}^{-1} \begin{pmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}' \tilde{\boldsymbol{\xi}}_{i3}}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}' \tilde{\boldsymbol{\xi}}_{i3}}{n} \end{pmatrix}$ is bounded above by a constant not depending on $n$ since

$$\begin{aligned}
&\left(\frac{\tilde{\boldsymbol{\xi}}_{i3}'\tilde{\boldsymbol{\xi}}_{i1}}{n}, \frac{\tilde{\boldsymbol{\xi}}_{i3}'\tilde{\boldsymbol{\xi}}_{i2}}{n}\right) \mathbf{C}_{11}^{-1} \begin{pmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i3}}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{\xi}}_{i3}}{n} \end{pmatrix} \\
&\leq \left\| \begin{pmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i3}}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{\xi}}_{i3}}{n} \end{pmatrix} \right\|^2 \left\|\mathbf{C}_{11}^{-1}\right\| \\
&\leq \frac{2c_5}{\sigma_1}, \text{ using previous arguments.}
\end{aligned}$$

Then, for sufficiently large $n$,

$$\begin{aligned}
&\frac{\tilde{\boldsymbol{\xi}}_{i3}'\tilde{\boldsymbol{\xi}}_{i3}}{n} - \frac{\tilde{\boldsymbol{\xi}}_{i3}'\mathbf{P}_{m_i}\tilde{\boldsymbol{\xi}}_{i3}}{n} + o(1) \\
\geq &\frac{\tilde{\boldsymbol{\xi}}_{i3}'\tilde{\boldsymbol{\xi}}_{i3}}{n} - \left(\tilde{\boldsymbol{\xi}}_{i3}'\tilde{\boldsymbol{\xi}}_{i1} \;\; \tilde{\boldsymbol{\xi}}_{i3}'\tilde{\boldsymbol{\xi}}_{i2}\right) \begin{bmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i1}}{n} & \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i2}}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{\xi}}_{i1}}{n} & \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{\xi}}_{i2}}{n} \end{bmatrix}^{-1} \begin{pmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i3}}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{\xi}}_{i3}}{n} \end{pmatrix} \\
= &\mathbf{U}_{\mathbf{C}_{11}} \text{ i.e. the schur complement of } \mathbf{C}_{11} \\
= &\frac{1}{(\tilde{\mathbf{E}}_i^{-1})_{33}} \text{ where } \tilde{\mathbf{E}}_i = \frac{1}{n}(\mathbf{E}_{i,\mathrm{H}}\mathbf{C})'\mathbf{E}_{i,\mathrm{H}}\mathbf{C} \text{ is the sample variance-covariaance matrix of } \begin{pmatrix} \tilde{\boldsymbol{\xi}}_{i1} \\ \tilde{\boldsymbol{\xi}}_{i2} \\ \tilde{\boldsymbol{\xi}}_{i3} \end{pmatrix} \\
\overset{(i)}{\geq} &\frac{1}{2(\tilde{\boldsymbol{\Sigma}}_i^{-1})_{33}} \text{ where } \tilde{\boldsymbol{\Sigma}}_i \text{ is the true variance-covariaance matrix of } \begin{pmatrix} \tilde{\boldsymbol{\xi}}_{i1} \\ \tilde{\boldsymbol{\xi}}_{i2} \\ \tilde{\boldsymbol{\xi}}_{i3} \end{pmatrix} \\
\geq &\frac{\sigma_{\min}}{2}, \text{ by Assumption A2}
\end{aligned}$$

where (i) holds for the following reason:

From Lemma 3.4 of Xiang et al. (2015) it follows that,

$$\mathbf{P}\left(\left|(\tilde{\mathbf{E}}_i^{-1})_{33} - (\tilde{\mathbf{\Sigma}}_i^{-1})_{33}\right| \geq \mathcal{M}_n\sqrt{\frac{\log p^2}{n}}\right) \leq m_1 p^{2-m_2 n^2}$$

where $m_1, m_2$ depends on $\lambda_{\min}(\tilde{\mathbf{\Sigma}}_i^{-1})$ . Hence,

$$\mathbf{P}\left(\left|\frac{1}{(\tilde{\mathbf{E}}_i^{-1})_{33}} - \frac{1}{(\tilde{\mathbf{\Sigma}}_i^{-1})_{33}}\right| \geq \mathcal{M}_n\sqrt{\frac{\log p^2}{n}}\right) \longrightarrow 0 \text{ for sufficiently large } n.$$

Thus,

$$\frac{1}{(\tilde{\mathbf{E}}_i^{-1})_{33}} \geq \frac{1}{(\tilde{\mathbf{\Sigma}}_i^{-1})_{33}} - \mathcal{M}_n\sqrt{\frac{\log p^2}{n}} \geq \frac{1}{2(\tilde{\mathbf{\Sigma}}_i^{-1})_{33}}$$

Hence we have,

$$\begin{aligned}
& B(\boldsymbol{m}_i, \boldsymbol{t}_i) \\
\leq \quad & (2q_n\sqrt{c_2})^{(\nu_{m_i}-\nu_{t_i})}\left(1 + \frac{4}{\tilde{\lambda}_1\sigma_{\min}}\mathcal{M}_n^2(\nu_{m_i}-\nu_{t_i})\frac{\log p^2}{n}\right)^{\left(\frac{n+\nu}{2}\right)} \\
\leq \quad & (q_n)^{\frac{(\nu_{m_i}-\nu_{t_i})}{2}} \exp\left\{-(\nu_{m_i}-\nu_{t_i})\left(4\mathcal{M}_n^2\frac{b_n\log p^2}{\tilde{\lambda}_1\sigma_{\min}} - \left(2+\frac{2\nu}{n}\right)\mathcal{M}_n^2\frac{\log p^2}{\tilde{\lambda}_1\sigma_{\min}}\right)\right\} \\
\leq \quad & (q_n)^{\frac{(\nu_{m_i}-\nu_{t_i})}{2}} \\
\leq \quad & (q_n)^{\frac{(\nu_{m_i}-\nu_{t_i})}{2}}
\end{aligned}$$

for large enough $n$.

**Case II:** $\nu_{m_i} < 2\nu_{t_i}$.

Using same argument as in Case I and $\nu_{m_i} - \nu_{t_i} \geq 1$, we have

$$\begin{aligned}
B(\boldsymbol{m}_i, \boldsymbol{t}_i) &\leq (2q_n\sqrt{c_2})^{(\nu_{m_i}-\nu_{t_i})}\left(1 + \frac{4}{\tilde{\lambda}_1\sigma_{\min}}\mathcal{M}_n^2(\nu_{m_i} - \nu_{t_i})\frac{\log p^2}{n}\right)^{(\frac{n+\nu}{2})} \\
&\leq (q_n)^{\frac{(\nu_{m_i}-\nu_{t_i})}{2}} \exp\left\{ - 4\mathcal{M}_n^2 \frac{b_n \log p^2}{\tilde{\lambda}_1\sigma_{\min}} + \left(2 + \frac{2\nu}{n}\right)\mathcal{M}_n^2 \frac{\log p^2}{\tilde{\lambda}_1\sigma_{\min}}\right\} \\
&\leq (q_n)^{\frac{(\nu_{m_i}-\nu_{t_i})}{2}} \\
&\leq (q_n)^{\frac{(\nu_{m_i}-\nu_{t_i})}{2}}
\end{aligned}$$

for large enough $n$.

**Proof of Lemma 4:**

Note that,

$$B(\boldsymbol{m}_i, \boldsymbol{t}_i) = \left(\frac{2q_{\nu_{m_i}}}{\tau\sqrt{n}}\right)^{(\nu_{m_i}-\nu_{t_i})} \frac{\left|\frac{\tilde{\mathbf{X}}'_{m_i}\tilde{\mathbf{X}}_{m_i}}{n} + \frac{\mathbf{I}_{\nu_{m_i}}}{nc_2\tau^2}\right|^{-1/2}}{\left|\frac{\tilde{\mathbf{X}}'_{t_i}\tilde{\mathbf{X}}_{t_i}}{n} + \frac{\mathbf{I}_{\nu_{t_i}}}{nc_2\tau^2}\right|^{-1/2}} \left(\frac{\mathbf{U}_{\mathbf{R}_{t_i}}}{\mathbf{U}_{\mathbf{R}_{m_i}}}\right)^{(\frac{n+\nu}{2})}$$

**Case I:** $\nu_{t_i} < \nu_{m_i}$.

Note that,

$$\frac{\left|\frac{\tilde{\mathbf{X}}'_{m_i}\tilde{\mathbf{X}}_{m_i}}{n} + \frac{\mathbf{I}_{\nu_{m_i}}}{nc_2\tau^2}\right|^{-1/2}}{\left|\frac{\tilde{\mathbf{X}}'_{t_i}\tilde{\mathbf{X}}_{t_i}}{n} + \frac{\mathbf{I}_{\nu_{t_i}}}{nc_2\tau^2}\right|^{-1/2}} \leq \frac{(2\lambda'_2)^{\frac{t_i}{2}}}{(\lambda'_1/2)^{\frac{m_i}{2}}} \leq C^{b_n(\nu_{m_i}-\nu_{t_i})}$$

for an appropriate constant $C$. Let, $\tilde{m}_i = m_i \cup t_i$. Since $\mathbf{U}_{\mathbf{R}_{\tilde{m}_i}} \leq \mathbf{U}_{\mathbf{R}_{m_i}}$, we have

$$
\begin{aligned}
B(\boldsymbol{m}_i, \boldsymbol{t}_i) &\leq \left(\frac{2q_{\nu_{m_i}}}{\tau\sqrt{n}}\right)^{(\nu_{m_i}-\nu_{t_i})} C^{b_n(\nu_{m_i}-\nu_{t_i})} \left(\frac{\mathbf{U}_{\mathbf{R}_{t_i}}}{\mathbf{U}_{\mathbf{R}_{\tilde{m}_i}}}\right)^{(\frac{n+\nu}{2})} \\
&\leq \left(\frac{2q_{\nu_{m_i}}}{\tau\sqrt{n}}\right)^{(\nu_{m_i}-\nu_{t_i})} C^{b_n(\nu_{m_i}-\nu_{t_i})} \left(1 + \frac{\tilde{\boldsymbol{\xi}}_{i3}'(\mathbf{P}_{\tilde{m}_i} - \mathbf{P}_{t_i})\tilde{\boldsymbol{\xi}}_{i3}/n + o(1)}{\tilde{\boldsymbol{\xi}}_{i3}'(\mathbf{I} - \mathbf{P}_{\tilde{m}_i})\tilde{\boldsymbol{\xi}}_{i3}/n + o(1)}\right)^{(\frac{n+\nu}{2})}
\end{aligned}
\tag{S5.24}
$$

Using similar arguments from Lemma 3,

$$
\tilde{\boldsymbol{\xi}}_{i3}'(\mathbf{P}_{\tilde{m}_i} - \mathbf{P}_{t_i})\tilde{\boldsymbol{\xi}}_{i3}/n = \frac{\tilde{\boldsymbol{\xi}}_{i3}^{\star\prime}(\mathbf{P}_{\tilde{m}_i}^{\star} - \mathbf{P}_{t_i}^{\star})\tilde{\boldsymbol{\xi}}_{i3}^{\star}}{n}
$$

where $(\mathbf{P}_{\tilde{m}_i}^{\star} - \mathbf{P}_{t_i}^{\star})$ turns out to be the projection matrix on $\mathcal{C}\left(\mathbf{A} - \mathbf{P}_{\mathbf{B}}\mathbf{A}\right)$ where, $\mathbf{A} = [\mathbf{Z}_{\tilde{m}_i - t_i}] = [\mathbf{Z}_{m_i \cap t_i{}^c}]$ and $\mathbf{B} = [\mathbf{Z}_{t_i}]$ after removing the effect of $\tilde{\boldsymbol{\xi}}_{i1}, \tilde{\boldsymbol{\xi}}_{i2}$ from both $\mathbf{Z}_{m_i}$ and $\mathbf{Z}_{t_i}$. Then,

$$
\begin{aligned}
\frac{\tilde{\boldsymbol{\xi}}_{i3}^{\star\prime}(\mathbf{P}_{\tilde{m}_i}^{\star} - \mathbf{P}_{t_i}^{\star})\tilde{\boldsymbol{\xi}}_{i3}^{\star}}{n} &\leq \left\|\frac{\tilde{\boldsymbol{\xi}}_{i3}^{\star\prime}\left(\mathbf{A} - \mathbf{P}_{\mathbf{B}}\mathbf{A}\right)}{n}\right\|^2 \lambda_{\max}\left(\frac{\mathbf{A}'\left(\mathbf{I} - \mathbf{P}_{\mathbf{B}}\right)\mathbf{A}}{n}\right)^{-1} \\
&\leq \frac{2}{\tilde{\lambda}_1}\mathcal{M}_n^2 \nu_{m_i \cap t_i{}^c} \frac{\log p^2}{n}
\end{aligned}
$$

Also,

$$
\frac{\tilde{\boldsymbol{\xi}}_{i3}'\tilde{\boldsymbol{\xi}}_{i3}}{n} - \frac{\tilde{\boldsymbol{\xi}}_{i3}'\mathbf{P}_{\tilde{m}_i}\tilde{\boldsymbol{\xi}}_{i3}}{n} + o(1) \geq \frac{1}{2(\tilde{\boldsymbol{\Sigma}}_i^{-1})_{33}}
$$

Using these bounds in (S5.24) we have,

$$
\begin{aligned}
& B(\boldsymbol{m}_i, \boldsymbol{t}_i) \\
\leq & \left(\frac{2q_n}{\tau\sqrt{n}}\right)^{(\nu_{m_i}-\nu_{t_i})} C^{b_n(\nu_{m_i}-\nu_{t_i})} \left(1 + \frac{4}{\tilde{\lambda}_1 \sigma_{\min}} \mathcal{M}_n^2 \nu_{m_i \cap t_i{}^{\mathrm{C}}} \frac{\log p^2}{n}\right)^{(\frac{n+\nu}{2})} \\
\leq & (q_n)^{\frac{(\nu_{m_i}-\nu_{t_i})}{2}} \exp\Bigg\{ -(\nu_{m_i} - \nu_{t_i}) \left(4\mathcal{M}_n^2 \frac{b_n \log p^2}{\tilde{\lambda}_1 \sigma_{\min}} - \left(2 + \frac{2\nu}{n}\right) \mathcal{M}_n^2 \frac{\log p^2}{\tilde{\lambda}_1 \sigma_{\min}} - b_n \log C\right) \Bigg\} \\
\leq & (q_n)^{\frac{(\nu_{m_i}-\nu_{t_i})}{2}}
\end{aligned}
$$

for large enough $n$.

**Case II:** $\nu_{m_i} \leq \nu_{t_i}$.

Let $a_i = t_i \cap m_i{}^c$. Then,

$$
\frac{\mathbf{A}_{t_i,\cdot} \mathbf{Z}'_{t_i} \left(\mathbf{P}_{t_i} - \mathbf{P}_{m_i}\right) \mathbf{Z}_{t_i} \mathbf{A}'_{t_i,\cdot}}{n} = \frac{\mathbf{A}_{a_i,\cdot} \mathbf{Z}'_{a_i} \left(\mathbf{I} - \mathbf{P}_{m_i}\right) \mathbf{Z}_{a_i} \mathbf{A}'_{a_i,\cdot}}{n}
$$

Note that, $\lambda_{\min}\left(\frac{\mathbf{Z}'_{a_i}\mathbf{Z}_{a_i}}{n}\right) \geq \frac{\tilde{\lambda}_1}{2}$

Using similar arguments as in Ghosh et al. (2021), it can be shown that,

$B(\boldsymbol{m}_i, \boldsymbol{t}_i) \leq (q_n)^{1/2}$.

We will now employ the above lemmas to prove Theorem 1.

**Proof of Theorem 1:** First we will show the following:

For the mixed frequency VAR model posited in (2.6) with lag $d = 1$ and the prior distributions on $\mathbf{A}, \mathcal{G}, \mathbf{\Sigma}_{\boldsymbol{\varepsilon}}$ specified in Section 3.1 of the main paper and fixed $\theta$, satisfying Assumptions A1-A4, for any "non-true" activity graph $\mathcal{G}_m$ with $n$ sufficiently large the following holds

$$\frac{\pi_{\text{pseudo}}(\mathcal{G}_m \mid \mathcal{Y})}{\pi_{\text{pseudo}}(\mathcal{G}_0 \mid \mathcal{Y})} \leq \ (p^2)^{-2D(m,t)},$$

Let $D(m,t)$ denote the total number of disagreements between the two activity graphs i.e. $D(m,t) = \left(\sum_{i=1}^{k_1} d(\boldsymbol{m}_i, \boldsymbol{t}_i) + \sum_{j=1}^{k_2} d(\boldsymbol{m}_j, \boldsymbol{t}_j)\right)$.

Note that $d(\boldsymbol{m}_i, \boldsymbol{t}_i) \leq 2b_n(\nu_{m_i} - \nu_{t_i})$ for $\nu_{m_i} > \nu_{t_i}$ and $d(\boldsymbol{m}_i, \boldsymbol{t}_i) \leq 2b_n$ for $\nu_{m_i} \leq \nu_{t_i}$. We will assume without loss of generality that $\sigma_{\min}$ and $\tilde{\lambda}_1$ are bounded above by 1. Then it follows by Lemma 2 - 4 and Assumption A3 that if $t_i \neq m_i$ and $t_j \neq m_j$ then for sufficiently large $n$,

$$B(\boldsymbol{m}_i, \boldsymbol{t}_i) \leq (p^2)^{-2d(\boldsymbol{m}_i, \boldsymbol{t}_i)} \text{ for } i = 1, \cdots, k_1$$

and,

$$B^*(\boldsymbol{m}_j, \boldsymbol{t}_j) \leq (p^2)^{-2d(\boldsymbol{m}_j, \boldsymbol{t}_j)} \text{ for } j = 1, \cdots, k_2$$

Also note that $B(\boldsymbol{m}_i, \boldsymbol{t}_i) = 1$ for $\boldsymbol{m}_i = \boldsymbol{t}_i$ and $B(\boldsymbol{m}_j, \boldsymbol{t}_j) = 1$ for $\boldsymbol{m}_j = \boldsymbol{t}_j$. Then for sufficiently large $n$,

$$\frac{1}{2}\frac{\pi(\mathcal{G}_m \mid \mathcal{Y})}{\pi(\mathcal{G}_0 \mid \mathcal{Y})} \leq \prod_{i=1}^{k_1} B\left(\boldsymbol{m}_i, \boldsymbol{t}_i\right) \prod_{j=1}^{k_2} B^{\star}\left(\boldsymbol{m}_j, \boldsymbol{t}_j\right) \leq (p^2)^{-2D(m,t)} \tag{S5.25}$$

The proof of Theorem 1 is straightforward using the above bound in (S5.25), following similar arguments in Ghosh et al. (2021).

**Remark S3.** *While the analysis of the $k_2$ terms $\{B^*(\boldsymbol{m}_j, \boldsymbol{t}_j)\}_{j=1}^{k_2}$ in (S5.2) corresponding to the low frequency regressions is very similar to the analysis in (Ghosh et al., 2021)), the main challenge and novelty in the above proof lies in the analysis of the the $k_1$ terms $\{B(\boldsymbol{m}_i, \boldsymbol{t}_i)\}_{i=1}^{k_1}$ corresponding to the high frequency regressions. Since $\boldsymbol{y}_{i,H}^t$ is a $3 \times 1$ vector whereas $y_{j,L}^t$ is a scalar, the ratio of residual sum of square terms $R_{t_j}/R_{m_j}$ in a low frequency term $B^*(\boldsymbol{m}_j, \boldsymbol{t}_j)$ is replaced by a more complex expression involving determinants and Schur complements of relevant matrices in the high frequency term $B(\boldsymbol{m}_i, \boldsymbol{t}_i)$. This complication, arising at heart due to the multivariate nature of the high frequency observations $\boldsymbol{y}_{i,H}^t$, leads to non-trivial mathematical challenges which are handled and resolved in a careful way in the proof above. The techniques used in this analysis can be leveraged to generalize existing consistency results in the literature for related contexts. For example, results in Narisetty and He (2014) on selection consistency for Bayesian univariate regression with spike-and-slab priors, can now be extended for Bayesian multivariate regression with spike-and-slab priors and a full error covariance matrix as long as the number of responses stays bounded with $n$. Alternatively, the number of responses can grow with $n$, but the error covariance matrix needs to have a block diagonal form with the dimension of each block staying bounded with $n$. These applications will* *be investigated as part of future research.*

**Remark S4.** *Note that in this high dimensional setting, we allow the number of high frequency and low frequency variables* $k_1 = k_{1,n}$ *and* $k_2 = k_{2,n}$ *respectively to vary with* $n$*. Next, we illustrate how the convergence rate* $\frac{1+\mu_{\max}(\mathcal{A})}{\mu_{\min}(\mathcal{A})}\sqrt{\frac{b_n \log(3k_1+k_2)}{n}}$ *appearing in Assumption A1 behaves as a function of* $k_1$ *and* $k_2$ *under different regimes. Note that* $(3k_{1,n} + k_{2,n}) = O(max\ (k_{1,n}, k_{2,n}))$*. In typical nowcasting applications, the majority of variables are high frequency and they are used to predict low frequency variables. In such settings, we would expect that* $k_{1,n}$ *converges to infinity faster than* $k_{2,n}$ *and hence* $(3k_{1,n} + k_{2,n}) = O(k_{1,n})$*. Note that* $b_n$ *denotes the maximum number of non-null entries within the rows of* $\boldsymbol{A}_{0n}$ *and is different from* $k_{1,n}$ *and* $k_{2,n}$*. Assume now that the number of non-zero entries in every row of* $\boldsymbol{A}_0$ *is a fraction* $r_n$ *of* $(k_{1,n} + k_{2,n})$*. Then* $b_n = O(r_n k_{1,n})$*. Hence, Assumption A1 will hold if* $\dfrac{1+\mu_{\max}(\mathcal{A})}{\mu_{\min}(\mathcal{A})}\sqrt{\dfrac{r_n k_1 \log k_1}{n}} \to 0$ *holds. In other situations wherein the number of high frequency variables grows at a slower rate compared to the number of low frequency variables, in the above analysis* $k_{1,n}$ *can be replaced by* $k_{2,n}$*.*

### S5.3 Proof of Theorem 2:

Throughout the proof, we will use $p$ to denote $(3k_1 + k_2)$. For ease of presentation, we denote $\Pi_{\text{pseudo}}(. \mid \mathcal{Y})$ by $\Pi(. \mid \mathcal{Y})$. Note that for any $\eta > 0$

$$\begin{aligned}
&\mathbb{E}_0\left(\Pi_n\left\{\|\mathbf{A} - \mathbf{A}_0\|_F \geq K\eta \mid \mathcal{Y}\right\}\right) \\
&= \sum_{\mathcal{G}} \mathbb{E}_0\left(\Pi_n\left\{\|\mathbf{A} - \mathbf{A}_0\|_F \geq K\eta \mid \mathcal{Y}, \mathcal{G}\right\} \pi_n(\mathcal{G} \mid \mathcal{Y})\right) \\
&\leq \mathbb{E}_0\left(\Pi_n\left\{\|\mathbf{A} - \mathbf{A}_0\|_F \geq K\eta \mid \mathcal{Y}, \mathcal{G}_0\right\}\right) + \mathbb{E}_0 \Pi_n(\mathcal{G} \neq \mathcal{G}_0 \mid \mathcal{Y})
\end{aligned}$$

By Theorem 1, it is enough to prove that $\mathbb{E}_0\left(\Pi_n\left\{\|\mathbf{A} - \mathbf{A}_0\|_F \geq K\eta \mid \mathcal{Y}, \mathcal{G}_0\right\}\right) \to 0$ as $n \to \infty$. Henceforth all the analysis is restricted to the true activity graph $\mathcal{G}_0$. Thus for ease of exposition we shall use $\mathbf{Z}_i, \nu_i$ to denote $\mathbf{Z}_{t_i}, \nu_{t_i}$ respectively. Using the same steps as in Ghosh et al. (2021), it can be shown that

$$\begin{aligned}
&\mathbb{E}_0\left(\Pi_n\left\{\|\mathbf{A} - \mathbf{A}_0\|_F \geq K\frac{1 + \mu_{\max}(\mathcal{A})}{\mu_{\min}(\mathcal{A})}\sqrt{\frac{\delta_n \log p}{n}}\middle| \mathcal{Y}, \mathcal{G}_0\right\}\right) \\
\leq\;& k_1 \max_{1 \leq i \leq k_1} \mathbb{E}_0\left(\Pi_n\left\{\|\mathbf{A}'_{i\cdot} - \mathbf{A}'_{0,i\cdot}\| \geq K\frac{1 + \mu_{\max}(\mathcal{A})}{\mu_{\min}(\mathcal{A})}\sqrt{\frac{\nu_i \log p}{n}}\middle| \mathcal{Y}, \mathcal{G}_0\right\}\right) + \\
& k_2 \max_{1 \leq j \leq k_2} \mathbb{E}_0\left(\Pi_n\left\{\|\mathbf{A}'_{j\cdot} - \mathbf{A}'_{0,j\cdot}\| \geq K\frac{1 + \mu_{\max}(\mathcal{A})}{\mu_{\min}(\mathcal{A})}\sqrt{\frac{\nu_{k_1+j} \log p}{n}}\middle| \mathcal{Y}, \mathcal{G}_0\right\}\right).
\end{aligned}$$

Following the same line of arguments as in Ghosh et al. (2021), it can be shown that

$$\mathbb{E}_0\left(\Pi_n\left\{\|\mathbf{A}'_{j\cdot} - \mathbf{A}'_{0,j\cdot}\| \geq K\frac{1 + \mu_{\max}(\mathcal{A})}{\mu_{\min}(\mathcal{A})}\sqrt{\frac{\nu_j \log p}{n}} \mid \mathcal{Y}, \mathcal{G}_0\right\}\right) \leq 5\mathrm{e}^{-2\log p} \text{ for } j = 1, \ldots, k_2$$

The proof will be complete by showing

$$\mathbb{E}_0\left(\Pi_n\left\{\|\mathbf{A}'_{i\cdot}-\mathbf{A}'_{0,i\cdot}\|\geq K\frac{1+\mu_{\max}(\mathcal{A})}{\mu_{\min}(\mathcal{A})}\sqrt{\frac{\nu_i\log p}{n}}\mid\mathcal{Y},\mathcal{G}_0\right\}\right)\leq 5\mathrm{e}^{-2\log p}$$

for $i=1,\ldots,k_1$. For the ease of exposition we denote $\frac{1+\mu_{\max}(\mathcal{A})}{\mu_{\min}(\mathcal{A})}\sqrt{\frac{\nu_i\log p}{n}}$ by $\eta_{n,i}$. Next, under the prior on $\mathbf{A}$ and $\mathcal{G}$ in (3.4), we have

$$\mathbf{A}'_{i\cdot}\big|\mathcal{G}_0,\boldsymbol{\Omega}_{i,\mathrm{H}},\mathcal{Y}\ \sim\ \mathcal{N}_{\nu_i}\left(\left(\mathbf{Z}'_i\mathbf{Z}_i+\frac{\mathbf{I}}{\tau^2}\right)^{-1}\mathbf{Z}'_i\tilde{\mathbf{Y}}_{i,\mathrm{H}}\tilde{\sigma}_i^2\mathbf{F}_{i,\cdot 3},\tilde{\sigma}_i^2\left(\mathbf{Z}'_i\mathbf{Z}_i+\frac{\mathbf{I}}{\tau^2}\right)^{-1}\right) \tag{S5.26}$$

where $\mathbf{F}_i$ is as defined in (S3.1). Let us denote $\left(\mathbf{Z}'_i\mathbf{Z}_i+\frac{\mathbf{I}}{\tau^2}\right)^{-1}\mathbf{Z}'_i\tilde{\mathbf{Y}}_{i,\mathrm{H}}\tilde{\sigma}_i^2\mathbf{F}_{i,\cdot 3}$ by $\mathbf{A}_{i,M}$. To derive the marginal posterior of $\boldsymbol{\Omega}_{i,\mathrm{H}}$, we consider the transformation:

$$\boldsymbol{\Omega}_{i,\mathrm{H}}\to\mathbf{F}_i=\mathbf{C}'\boldsymbol{\Omega}_{i,\mathrm{H}}\mathbf{C}$$

as described in (S3.1). Then following the same steps as in (S3.2), $\mathbf{F}_i$ can be written as

$$\mathbf{F}_i=\begin{bmatrix}\mathbf{B}_1+\mathbf{e}\mathbf{e}'f & \mathbf{e}f\\ \mathbf{e}'f & f\end{bmatrix}$$

where $\mathbf{e} = \frac{\mathbf{F}_{i,12}}{\mathbf{F}_{i,22}}, f = \mathbf{F}_{i,22}, \mathbf{B}_1 = \mathbf{F}_{i,11} - \mathbf{F}_{i,12}\mathbf{F}_{i,22}^{-1}\mathbf{F}_{i,21}$. Then, under the prior on $\mathbf{\Sigma}_{i,\mathrm{H}}$, we have

$$\begin{aligned} \mathbf{e}\Big| f, \mathcal{G}_0, \mathcal{Y} &\sim \mathcal{N}_2\left(-(\mathbf{R}_{i,11}^{\perp})^{-1}\mathbf{R}_{i,12}^{\perp}, \frac{1}{f}(\mathbf{R}_{i,11}^{\perp})^{-1}\right) \\ \tilde{\sigma}_i^2 = \frac{1}{f}\Big|\mathcal{G}_0, \mathcal{Y} &\sim \text{Inverse-Gamma}\left(\frac{n+\omega}{2}, \frac{\mathbf{R}_{i,22}^{\perp} - \mathbf{R}_{i,12}^{\perp}{}'(\mathbf{R}_{i,11}^{\perp})^{-1}\mathbf{R}_{i,12}^{\perp}}{2}\right) \end{aligned} \tag{S5.27}$$

where $\mathbf{R}_i^{\perp} = n\mathbf{R}_i$ and $\mathbf{R}_i$ is as defined in Section S2.5. First note that, for $i = 1, \ldots, k_1$

$$\begin{aligned} &\mathbb{E}_0\left(\Pi_n\left\{\|\mathbf{A}_{i\cdot}' - \mathbf{A}_{0,i\cdot}'\| \geq K\eta_{n,i} \mid \mathcal{Y}, \mathcal{G}_0\right\}\right) \\ &\leq \mathbb{E}_0\Pi_n\left(\|\mathbf{A}_{i\cdot}' - \mathbf{A}_{i,M}\| \geq \frac{K\eta_{n,i}}{2} \mid \mathcal{Y}, \mathcal{G}_0\right) + \mathbb{E}_0\Pi_n\left(\|\mathbf{A}_{i,M} - \mathbf{A}_{0,i\cdot}'\| \geq \frac{K\eta_{n,i}}{2} \mid \mathcal{Y}, \mathcal{G}_0\right) \end{aligned} \tag{S5.28}$$

Now,

$$\begin{aligned} \|\mathbf{A}_{i\cdot}' - \mathbf{A}_{i,M}\| &= \left\|\tilde{\sigma}_i\left(\mathbf{Z}_i'\mathbf{Z}_i + \frac{\mathbf{I}_{\nu_i}}{\tau^2}\right)^{-1/2} \frac{\left(\mathbf{Z}_i'\mathbf{Z}_i + \frac{\mathbf{I}_{\nu_i}}{\tau^2}\right)^{1/2}(\mathbf{A}_{i\cdot}' - \mathbf{A}_{i,M})}{\tilde{\sigma}_i}\right\| \\ &\leq \frac{\tilde{\sigma}_i}{\sqrt{\lambda_{\min}\left(\mathbf{Z}_i'\mathbf{Z}_i\right)}}\|\boldsymbol{z}\|, \end{aligned}$$

where $\boldsymbol{z}$ is $\nu_i \times 1$ standard normal random vector and the last step follows from (S5.26). Thus, for any $M^* > 0$

$$\begin{aligned}
&\mathbb{E}_0\Pi_n\left(\|\mathbf{A}_{i\cdot}' - \mathbf{A}_{i,M}\| \geq \frac{K\eta_{n,i}}{2} \mid \mathcal{Y}\right) \\
&\leq \mathbb{P}\left(\lambda_{\min}\left(\frac{\mathbf{Z}_i'\mathbf{Z}_i}{n}\right) < \tilde{\lambda}_1\right) + \mathbb{P}\left(\|\boldsymbol{z}\| > \sqrt{n}\eta_{n,i}K\sqrt{\frac{\tilde{\lambda}_1}{8(M^\star)^2}}\right) + \mathbb{E}_0\Pi_n\left(\tilde{\sigma}_i > M^* \mid \mathcal{Y}\right).
\end{aligned} \tag{S5.29}$$

On the good set $\tilde{E}_{21,n}$, there exists $n_1$ such that for all $n \geq n_1$ the first term above is upper bounded by $\mathrm{e}^{-2\log p}$. Using the same line of arguments as in Ghosh et al. (2021), it can be shown that the second term is also upper bounded by $\mathrm{e}^{-2\log p}$. For the third term, recall from the distribution of $\tilde{\sigma}_i^2 \big| \mathcal{G}_0, \mathcal{Y}$ in (S5.27) that $\frac{n+\omega}{2} \sim n$ and

$$\begin{aligned}
\frac{\mathbf{R}_{i,22}^{\perp} - {\mathbf{R}_{i,12}^{\perp}}'(\mathbf{R}_{i,11}^{\perp})^{-1}\mathbf{R}_{i,12}^{\perp}}{2n} &= \frac{1}{2}\mathbf{U}_{\mathbf{R}_i} \\
&\leq \frac{1}{2}\left(\frac{\tilde{\boldsymbol{y}}_{i3}'\tilde{\boldsymbol{y}}_{i3}}{n} + \frac{1}{\tilde{c}n\tau^2}\right) \\
&\overset{(i)}{\leq} c_A
\end{aligned}$$

where (i) follows from (S5.12). Hence the scale parameter $\frac{\mathbf{R}_{i,22}^{\perp} - {\mathbf{R}_{i,12}^{\perp}}'(\mathbf{R}_{i,11}^{\perp})^{-1}\mathbf{R}_{i,12}^{\perp}}{2}$ in the distribution of $\tilde{\sigma}_i^2 \big| \mathcal{G}_0, \mathcal{Y}$ is of order n. By choosing $M^*$ properly, we can make $\mathbb{E}_0\Pi_n\left(\tilde{\sigma}_i > M^* \mid \mathcal{Y}\right) \leq \mathrm{e}^{-2\log p}$ for all large $n$. Moving onto the second term in (S5.28), note that

$$
\begin{aligned}
\left\|\mathbf{A}_{i,M} - \mathbf{A}'_{0,i\cdot}\right\| &= \left\|\left(\mathbf{Z}_i'\mathbf{Z}_i + \frac{\mathbf{I}_{\nu_i}}{\tau^2}\right)^{-1} \mathbf{Z}_i'\tilde{\mathbf{Y}}_{i,\mathrm{H}}\tilde{\sigma}_i^2\mathbf{F}_{i,\cdot 3} - \mathbf{A}'_{0,i\cdot}\right\| \\
&\le \left\|\left(\mathbf{Z}_i'\mathbf{Z}_i + \frac{\mathbf{I}_{\nu_i}}{\tau^2}\right)^{-1}\right\| \left\|\mathbf{Z}_i'\tilde{\mathbf{Y}}_{i,\mathrm{H}}\tilde{\sigma_i}^2\mathbf{F}_{i,\cdot 3} - \left(\mathbf{Z}_i'\mathbf{Z}_i + \frac{\mathbf{I}_{\nu_i}}{\tau^2}\right)\mathbf{A}'_{0,i\cdot}\right\| \\
&\overset{(ii)}{\le} \left\|\left(\mathbf{Z}_i'\mathbf{Z}_i + \frac{\mathbf{I}_{\nu_i}}{\tau^2}\right)^{-1}\right\| \left\|\mathbf{Z}_i'\mathbf{Z}_i\mathbf{A}'_{0,i\cdot} + \mathbf{Z}_i'\mathbf{E}_{i,\mathrm{H}}\boldsymbol{\Omega}_{i,\mathrm{H}}\boldsymbol{\delta}\tilde{\sigma}_i^2 - \mathbf{Z}_i'\mathbf{Z}_i\mathbf{A}'_{0,i\cdot} - \frac{\mathbf{A}'_{0,i\cdot}}{\tau^2}\right\| \\
&\le \frac{1}{\lambda_{\min}(\mathbf{Z}_i'\mathbf{Z}_i/n)} \left(\frac{\left\|\mathbf{Z}_i'\mathbf{E}_{i,\mathrm{H}}\boldsymbol{\Omega}_{i,\mathrm{H}}\boldsymbol{\delta}\right\|\tilde{\sigma}_i^2}{n} + \frac{\left\|\mathbf{A}'_{0,i\cdot}\right\|}{n\tau^2}\right).
\end{aligned}
$$

where (ii) follows from the fact that

$$
\begin{aligned}
\mathbf{Z}_i'\tilde{\mathbf{Y}}_{i,\mathrm{H}}\mathbf{F}_{i,\cdot 3}\tilde{\sigma}_i^2 &= \mathbf{Z}_i'\left(\mathbf{Y}_{i,\mathrm{H}}\mathbf{C}\right)'\left(\mathbf{C}'\boldsymbol{\Omega}_{i,\mathrm{H}}\boldsymbol{\delta}\right)\tilde{\sigma}_i^2 \\
&= \mathbf{Z}_i\mathbf{Y}_{i,\mathrm{H}}\boldsymbol{\Omega}_{i,\mathrm{H}}\boldsymbol{\delta}\tilde{\sigma}_i^2 \\
&= \mathbf{Z}_i'\left(\mathbf{Z}_i\mathbf{A}'_{0,i\cdot}\boldsymbol{\delta}' + \mathbf{E}_{i,\mathrm{H}}\right)\boldsymbol{\Omega}_{i,\mathrm{H}}\boldsymbol{\delta}\tilde{\sigma}_i^2 \\
&= \mathbf{Z}_i'\mathbf{Z}_i\mathbf{A}_{0,i\cdot} + \mathbf{Z}_i'\mathbf{E}_{i,\mathrm{H}}\boldsymbol{\Omega}_{i,\mathrm{H}}\boldsymbol{\delta}\tilde{\sigma}_i^2
\end{aligned}
$$

Therefore,

$$\begin{aligned}
&\mathbb{E}_0\Pi_n\left(\|\mathbf{A}_{i,M}-\mathbf{A}'_{0,i\cdot}\| \geq \frac{K\eta_{n,i}}{2} \mid \mathcal{Y},\mathcal{G}_0\right)\\
\leq\;& \mathbb{E}_0\Pi_n\left[\frac{1}{\lambda_{\min}(\mathbf{Z}_i'\mathbf{Z}_i/n)}\left(\frac{\|\mathbf{Z}_i'\mathbf{E}_{i,\mathrm{H}}\boldsymbol{\Omega}_{i,\mathrm{H}}\boldsymbol{\delta}\|\,\tilde{\sigma}_i^2}{n}+\frac{\left\|\mathbf{A}'_{0,i\cdot}\right\|}{n\tau^2}\right)\geq\frac{K\eta_{n,i}}{2}\right]\\
\leq\;& \mathbb{P}_0\left[\lambda_{\min}\left(\frac{\mathbf{Z}_i'\mathbf{Z}_i}{n}\right)<\frac{\tilde{\lambda}_1}{2}\right]+\mathbb{E}_0\Pi_n\left(\frac{\|\mathbf{Z}_i'\mathbf{E}_{i,\mathrm{H}}\boldsymbol{\Omega}_{i,\mathrm{H}}\boldsymbol{\delta}\|\,\tilde{\sigma}_i^2}{n}\geq\frac{K\eta_{n,i}}{8}\mid\mathcal{Y}\right)+\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\mathbb{P}_0\left(\frac{\left\|\mathbf{A}'_{0,i\cdot}\right\|}{n\tau^2}>\frac{K\eta_{n,i}}{8}\right).\\
\leq\;& \mathbb{P}_0\left[\lambda_{\min}\left(\frac{\mathbf{Z}_i'\mathbf{Z}_i}{n}\right)<\frac{\tilde{\lambda}_1}{2}\right]+\mathbb{E}_0\Pi_n\left(\tilde{\sigma}_i^2>M_1^*\mid\mathcal{Y}\right)+\\
&\mathbb{E}_0\Pi_n\left(\frac{\|\mathbf{Z}_i'\mathbf{E}_{i,\mathrm{H}}\boldsymbol{\Omega}_{i,\mathrm{H}}\boldsymbol{\delta}\|}{n}\geq\frac{K\eta_{n,i}}{8M_1^\star}\mid\mathcal{Y}\right)+\mathbb{P}_0\left(\frac{\left\|\mathbf{A}'_{0,i\cdot}\right\|}{n\tau^2}>\frac{K\eta_{n,i}}{8}\right).
\end{aligned}\tag{S5.30}$$

for any $M_1^* > 0$. Now both the first and second term in (S5.30) can be upper bounded by $\mathrm{e}^{-2\log p}$ for all large $n$ using similar arguments discussed before. Note that

$$\begin{aligned}
&\mathbb{E}_0\Pi_n\left(\frac{\left\|\mathbf{Z}_i'\mathbf{E}_{i,\mathrm{H}}\mathbf{C}\mathbf{C}^{-1}\boldsymbol{\Omega}_{i,\mathrm{H}}\boldsymbol{\delta}\right\|}{n}\geq\frac{K\eta_{n,i}}{8M_1^\star}\mid\mathcal{Y}\right)\\
=\;& \mathbb{E}_0\Pi_n\left(\frac{\left\|\mathbf{Z}_i'\left(\tilde{\boldsymbol{\xi}}_{i1}\;\tilde{\boldsymbol{\xi}}_{i2}\;\tilde{\boldsymbol{\xi}}_{i3}\right)\mathbf{F}_{i,\cdot3}\right\|}{n}\geq\frac{K\eta_{n,i}}{8M_1^\star}\mid\mathcal{Y}\right)\\
\leq\;& \mathbb{P}_0\left(\frac{\left\|\mathbf{Z}_i'\tilde{\boldsymbol{\xi}}_{i1}\;\mathbf{Z}_i'\tilde{\boldsymbol{\xi}}_{i2}\;\mathbf{Z}_i'\tilde{\boldsymbol{\xi}}_{i3}\right\|}{n}\geq\frac{K\eta_{n,i}}{8M_1^\star M_2^\star}\right)+\\
&\qquad\qquad\qquad\mathbb{E}_0\Pi_n\left(\|\mathbf{F}_{i,\cdot3}\|\geq M_2^\star\mid\mathcal{Y}\right)\quad\text{for any } M_2^\star>0\\
\leq\;& \mathbb{P}_0\left(\frac{\left\|\mathbf{Z}_i'\tilde{\boldsymbol{\xi}}_{i1}\right\|}{n}+\frac{\left\|\mathbf{Z}_i'\tilde{\boldsymbol{\xi}}_{i2}\right\|}{n}+\frac{\left\|\mathbf{Z}_i'\tilde{\boldsymbol{\xi}}_{i3}\right\|}{n}\geq\frac{K\eta_{n,i}}{8M_1^\star M_2^\star}\right)+\\
&\qquad\qquad\qquad\mathbb{E}_0\Pi_n\left(\|\mathbf{F}_{i,\cdot3}\|\geq M_2^\star\mid\mathcal{Y}\right)
\end{aligned}\tag{S5.31}$$

On the good sets $\tilde{G}_{21,n}$,$\tilde{G}_{22,n}$ and $\tilde{G}_{23,n}$, the first term in (S5.31) can be upper bounded by $\mathrm{e}^{-2\log p}$ for all large $n$. Now,

$$\begin{aligned}
& \mathbb{E}_0\Pi_n\left(\|\mathbf{F}_{i,\cdot 3}\| \geq M_2^\star \mid \mathcal{Y}\right) \\
= \; & \mathbb{E}_0\Pi_n\left(\|\mathbf{e}f\, f\| \geq M_2^\star \mid \mathcal{Y}\right) \\
\leq \; & \mathbb{E}_0\Pi_n\left(\|\mathbf{e}\| \geq M_4^\star \mid \mathcal{Y}, f \leq M_3^\star\right) + \\
& \mathbb{E}_0\Pi_n\left(f \geq M_3^\star \mid \mathcal{Y}\right) \qquad \text{for any } M_3^\star > 0 \text{ and } M_4^\star = \frac{M_2^\star}{M_3^\star} - 1 \\
\leq \; & \mathbb{E}_0\Pi_n\left(\left\|-\left(\mathbf{R}_{i,11}^{\perp}\right)^{-1}\mathbf{R}_{i,12}^{\perp}\right\| + \frac{1}{M_3^\star}\left\|n\left(\mathbf{R}_{i,11}^{\perp}\right)^{-1}\right\|\frac{\|\boldsymbol{z}\|}{n} \geq M_4^\star \mid \mathcal{Y}\right) + \\
& \qquad\qquad\qquad \mathbb{E}_0\Pi_n\left(f \geq M_3^\star \mid \mathcal{Y}\right) \qquad\qquad (S5.32)
\end{aligned}$$

From the distribution of $\mathbf{e}\big|f, \mathcal{G}_0, \mathcal{Y}$, we will now show that $\left\|n(\mathbf{R}_{i,11}^{\perp})^{-1}\right\|$ and $\left\|-(\mathbf{R}_{i,11}^{\perp})^{-1}\mathbf{R}_{i,12}^{\perp}\right\|$ are upper bounded by constants not depending on $n$. Note that,

$$\left\|n\left(\mathbf{R}_{i,11}^{\perp}\right)^{-1}\right\| = \frac{1}{\lambda_{\min}\left(\mathbf{R}_{i,11}^{\perp}/n\right)}$$

where $\mathbf{R}_{i,11}^{\perp}/n$ can be expressed as

$$\mathbf{R}_{i,11}^{\perp}/n = \mathbf{C}_{11} - \begin{pmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\mathbf{Z}_i}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\mathbf{Z}_i}{n} \end{pmatrix} \left(\frac{\mathbf{Z}_i'\mathbf{Z}_i}{n} + \frac{\mathbf{I}}{n\tau^2}\right)^{-1} \begin{pmatrix} \frac{\mathbf{Z}_i'\tilde{\boldsymbol{\xi}}_{i1}}{n} & \frac{\mathbf{Z}_i'\tilde{\boldsymbol{\xi}}_{i2}}{n} \end{pmatrix} + \frac{\mathbf{I}_2}{n\tilde{c}\tau^2}$$

where $\mathbf{C}_{11}$ is as given in (S5.19).

Note that, $\frac{\mathbf{I}_2}{n\tilde{c}\tau^2}$ is $o(1)$ and $\begin{pmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\mathbf{Z}_i}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\mathbf{Z}_i}{n} \end{pmatrix} \left(\frac{\mathbf{Z}_i'\mathbf{Z}_i}{n} + \frac{\mathbf{I}}{n\tau^2}\right)^{-1} \begin{pmatrix} \frac{\mathbf{Z}_i'\tilde{\boldsymbol{\xi}}_{i1}}{n} & \frac{\mathbf{Z}_i'\tilde{\boldsymbol{\xi}}_{i2}}{n} \end{pmatrix}$ is of smaller order compared to $\mathbf{C}_{11}$ since

$$\begin{aligned}
&\left\| \begin{pmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\mathbf{Z}_i}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\mathbf{Z}_i}{n} \end{pmatrix} \left( \frac{\mathbf{Z}_i'\mathbf{Z}_i}{n} + \frac{\mathbf{I}}{n\tau^2} \right)^{-1} \begin{pmatrix} \frac{\mathbf{Z}_i'\tilde{\boldsymbol{\xi}}_{i1}}{n} & \frac{\mathbf{Z}_i'\tilde{\boldsymbol{\xi}}_{i2}}{n} \end{pmatrix} \right\|_{\max} \\
&\leq \left\| \begin{pmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\mathbf{Z}_i}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\mathbf{Z}_i}{n} \end{pmatrix} \left( \frac{\mathbf{Z}_i'\mathbf{Z}_i}{n} + \frac{\mathbf{I}}{n\tau^2} \right)^{-1} \begin{pmatrix} \frac{\mathbf{Z}_i'\tilde{\boldsymbol{\xi}}_{i1}}{n} & \frac{\mathbf{Z}_i'\tilde{\boldsymbol{\xi}}_{i2}}{n} \end{pmatrix} \right\| \\
&\leq \left\| \begin{pmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\mathbf{Z}_i}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\mathbf{Z}_i}{n} \end{pmatrix} \right\|^2 \lambda_{\max} \left( \frac{\mathbf{Z}_i'\mathbf{Z}_i}{n} + \frac{\mathbf{I}}{n\tau^2} \right)^{-1} \\
&\leq \left( \left\| \frac{\tilde{\boldsymbol{\xi}}_{i1}'\mathbf{Z}_i}{n} \right\| + \left\| \frac{\tilde{\boldsymbol{\xi}}_{i2}'\mathbf{Z}_i}{n} \right\| \right)^2 \frac{1}{\lambda_{\min}\left( \frac{\mathbf{Z}_i'\mathbf{Z}_i}{n} + \frac{\mathbf{I}}{n\tau^2} \right)} \\
&\leq \frac{2}{\tilde{\lambda}_1} \left( 4\mathcal{M}_n^2 b_n \frac{\log p^2}{n} \right)
\end{aligned}$$

Hence, $\mathbf{R}_{i,11}^{\perp}/n \approx \mathbf{C}_{11}$ and we can have, $\lambda_{\min}\left(\mathbf{R}_{i,11}^{\perp}/n\right) \geq \frac{\sigma_1}{2}$ where $\sigma_1$ is as given in (S5.4) and thus,

$$\left\| n \left(\mathbf{R}_{i,11}^{\perp}\right)^{-1} \right\| \leq \frac{2}{\sigma_1}.$$

Next we move on to the term $\left\| - \left(\mathbf{R}_{i,11}^{\perp}\right)^{-1} \mathbf{R}_{i,12}^{\perp} \right\|$. Note that,

$$\left\| - \left(\mathbf{R}_{i,11}^{\perp}\right)^{-1} \mathbf{R}_{i,12}^{\perp} \right\| \leq \left\| n \left(\mathbf{R}_{i,11}^{\perp}\right)^{-1} \right\| \left\| \frac{\mathbf{R}_{i,12}^{\perp}}{n} \right\| \tag{S5.33}$$

Now,

$$\frac{\mathbf{R}_{i,12}^{\perp}}{n} = \begin{pmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{y}}_{i3}}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{y}}_{i3}}{n} \end{pmatrix} - \begin{pmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\mathbf{Z}_i}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\mathbf{Z}_i}{n} \end{pmatrix} \left( \frac{\mathbf{Z}_i'\mathbf{Z}_i}{n} + \frac{\mathbf{I}}{n\tau^2} \right)^{-1} \frac{\mathbf{Z}_i'\tilde{\boldsymbol{y}}_{i3}}{n}$$

Note that, $\begin{pmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\mathbf{Z}_i}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\mathbf{Z}_i}{n} \end{pmatrix} \left(\frac{\mathbf{Z}_i'\mathbf{Z}_i}{n} + \frac{\mathbf{I}}{n\tau^2}\right)^{-1} \frac{\mathbf{Z}_i'\tilde{\boldsymbol{y}}_{i3}}{n}$ is of smaller order compared to $\begin{pmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{y}}_{i3}}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{y}}_{i3}}{n} \end{pmatrix}$, since

$$
\begin{aligned}
& \left\| \begin{pmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\mathbf{Z}_i}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\mathbf{Z}_i}{n} \end{pmatrix} \left(\frac{\mathbf{Z}_i'\mathbf{Z}_i}{n} + \frac{\mathbf{I}}{n\tau^2}\right)^{-1} \frac{\mathbf{Z}_i'\tilde{\boldsymbol{y}}_{i3}}{n} \right\|_{\max} \\
\leq \; & \left\| \begin{pmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\mathbf{Z}_i}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\mathbf{Z}_i}{n} \end{pmatrix} \left(\frac{\mathbf{Z}_i'\mathbf{Z}_i}{n} + \frac{\mathbf{I}}{n\tau^2}\right)^{-1} \frac{\mathbf{Z}_i'\tilde{\boldsymbol{y}}_{i3}}{n} \right\| \\
\leq \; & \frac{2}{\tilde{\lambda}_1} \left( 2\mathcal{M}_n \sqrt{b_n \frac{\log p^2}{n}} \right) \left\| \frac{\mathbf{Z}_i'\tilde{\boldsymbol{y}}_{i3}}{n} \right\| \\
\leq \; & \frac{2}{\tilde{\lambda}_1} \left( 2\mathcal{M}_n \sqrt{b_n \frac{\log p^2}{n}} \right) \left( c_3 + \mathcal{M}_n \sqrt{\frac{b_n \log p^2}{n}} \right) \text{ using (S5.10)}
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\left\|\mathbf{R}_{i,12}^{\perp}/n\right\| \approx & \left\| \begin{pmatrix} \frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{y}}_{i3}}{n} \\ \frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{y}}_{i3}}{n} \end{pmatrix} \right\| \\
\leq & \sqrt{ \left(\frac{\tilde{\boldsymbol{\xi}}_{i1}'\tilde{\boldsymbol{\xi}}_{i1}}{n}\right) \left(\frac{\tilde{\boldsymbol{y}}_{i3}'\tilde{\boldsymbol{y}}_{i3}}{n}\right) + \left(\frac{\tilde{\boldsymbol{\xi}}_{i2}'\tilde{\boldsymbol{\xi}}_{i2}}{n}\right) \left(\frac{\tilde{\boldsymbol{y}}_{i3}'\tilde{\boldsymbol{y}}_{i3}}{n}\right) } \\
\leq & \sqrt{ c_A \left( \frac{3\sigma_{12}}{2} + \frac{3\sigma_{22}}{2} \right) } \text{ from (S5.12) and on the event } \tilde{E}_{11,n} \cap \tilde{E}_{12,n}
\end{aligned}
$$

Then, from (S5.33) $\left\| -\mathbf{R}_{i,11}^{\perp}{}^{-1}\mathbf{R}_{i,12}^{\perp} \right\| \leq \frac{2}{\sigma_1}\sqrt{c_A\left(\frac{3\sigma_{12}}{2}+\frac{3\sigma_{22}}{2}\right)} \leq c_6,$ say, where $c_6$ does not depend on $n$. Then from (S5.32),

$$
\begin{aligned}
&\mathbb{E}_0\Pi_n\left(\|\mathbf{F}_{i,\cdot 3}\| \geq M_2^\star \mid \mathcal{Y}\right) \\
&\leq \mathbb{E}_0\Pi_n\left(\left\| -\left(\mathbf{R}_{i,11}^{\perp}\right)^{-1}\mathbf{R}_{i,12}^{\perp}\right\| + \frac{1}{M_3^\star}\left\| n\left(\mathbf{R}_{i,11}^{\perp}\right)^{-1}\right\| \frac{\|\boldsymbol{z}\|}{n} \geq M_4^\star \mid \mathcal{Y}\right) + \mathbb{E}_0\Pi_n\left(f \geq M_3^\star \mid \mathcal{Y}\right) \\
&\leq \mathbb{E}_0\Pi_n\left(\frac{\|\boldsymbol{z}\|}{n} \geq M_5^\star \mid \mathcal{Y}\right) + \mathbb{E}_0\Pi_n\left(f \geq M_3^\star \mid \mathcal{Y}\right) \quad \text{where } M_5^\star = \frac{\sigma_1 M_3^\star(M_4^\star - c_6)}{2}
\end{aligned}
$$

Using arguments from Vershynin (2010), we can select $n_2$ such that for all $n \geq n_2$ the first term can be upper bounded by $\mathrm{e}^{-2\log p}$ and the second term can also be upper bounded by $\mathrm{e}^{-2\log p}$ for large $n$ using similar arguments discussed before.

For the last term in (S5.30), note that any column of the true parameter matrix $\mathbf{A}_0$ is bounded in $l_2$ norm by a constant not depending on $n$. This follows from the same argument used in the discussion after (S5.11) which implies that $\frac{\|\mathbf{A}_{0,i\cdot}'\|}{n\tau^2} = o\left(\eta_{n,i}\right)$. By Assumption A1 the last term in (S5.30) can be upper bounded by $\leq \mathrm{e}^{-2\log p}$ for large $n$ and for a large enough choice of $K$. Hence, for an appropriate choice of $K$ we have

$$
\begin{aligned}
&\mathbb{E}_0\Pi_n\left(\|\mathbf{A}'_{i\cdot} - \mathbf{A}'_{0,i\cdot}\| \geq K\eta_{n,i} \mid \mathcal{Y}\right) \\
&\leq \mathbb{E}_0\Pi_n\left(\|\mathbf{A}'_{i\cdot} - \mathbf{A}_{i,M}\| \geq \frac{K\eta_{n,i}}{2} \mid \mathcal{Y}\right) + \mathbb{E}_0\Pi_n\left(\|\mathbf{A}_{i,M} - \mathbf{A}'_{0,i\cdot}\| \geq \frac{K\eta_{n,i}}{2} \mid \mathcal{Y}\right) \\
&\leq 3\mathrm{e}^{-2\log p} + 6\mathrm{e}^{-2\log p} \\
&= \frac{9}{p^2},
\end{aligned}
$$

which completes the proof.

# S6 Additional Simulation Results

## S6.1 Uncertainty quantification of the proposed Bayesian MF model through posterior credible intervals

In any Bayesian analysis, a standard way to capture uncertainty regarding key model parameters is to construct posterior credible intervals. To this end, we generate data from the Bayesian MF model as given in (2.6) with lag $d = 1$, $k_1 = 3, k_2 = 30$ and $\theta = 0.5$. The 'true' coefficient matrix $\mathbf{A}$ is generated with non-zero entries drawn from Unif $(0, 10) \cup$ Unif $(-10, 0)$, keeping edge density fixed at 4%. Then the entire $\mathbf{W}$ matrix as given in (2.7) is formed using $\mathbf{A}$ and $\theta = 0.5$. The spectral density of $\mathbf{W}$ is set to 0.72. The 'true' error covariance matrix $\mathbf{\Sigma}_{\boldsymbol{\varepsilon}}$ is generated (the generation process is discussed in the main paper) and rescaled to ensure that the process is

Table 1: True and estimated model parameters

| Block | Truth | Estimates obtained by Bayesian MF model |
|---|---|---|
| $\mathbf{A}_{11}$ | 0.167 | 0.160 |
| $\mathbf{A}_{12}$ | 0.777 | 0.713 |
| $\mathbf{A}_{21}$ | -0.828 | -0.679 |
| $\mathbf{A}_{22}$ | 0.858 | 0.825 |

stable with signal-to-noise ratio SNR $= 2$. Under this setting we generate $n = 100$ observations from the Bayesian MF model and estimate the activity graph associated with $\mathbf{A}$ using the Gibbs sampler algorithm. Next we pick 4 non-zero entries from $\mathbf{A}$ randomly, one from each block $\mathbf{A}_{11}, \mathbf{A}_{12}, \mathbf{A}_{21}$ and $\mathbf{A}_{22}$, for subsequent analysis; i.e. estimation and construction of posterior credible intervals. We repeat it 100 times. The estimates of the four selected entries obtained from the Bayesian MF model after averaging over the 100 replicates are given in Table 1 along with the true values of the parameters . The table indicates that the estimates are very close to the true values of the parameters for both the models. Next, we examine 95% posterior credible intervals for these 4 entries of $\mathbf{A}$. For each of these 4 parameters, the coverage of the credible intervals constructed using Bayesian MF model in 4 randomly chosen instances out of 100 replicates is depicted in Figure 2. The true values are marked by the blue circle. It can be clearly seen

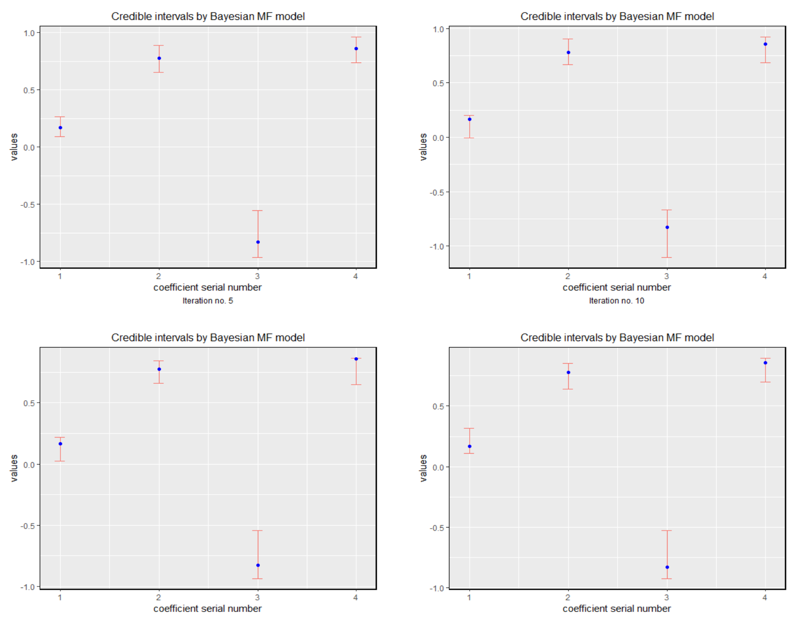


Figure 2: 95% Posterior credible intervals by Bayesian MF model for the 4 selected VAR coefficients.

from Figure 2 that the credible interval contains the true parameter and in most cases the true parameters are very close to the center of the credible interval.

### S6.2 MCMC Diagnostics

We monitor the convergence of the relevant Markov chains in our numerical illustrations by looking at trace plots and cumulative average plots. In Figure 3, we provide some selected such plots for setting 2 ($k_1 = 5, k_2 = 50$) in the simulation using a randomly chosen replicate .

### S6.3 Additional details on data generation from the Bayesian MF model as described in Section 6 of the main paper

Lemma 1 in Section S2.8 states that the stability of the VAR process depends on $\theta$. Hence, to ensure that the VAR process is stable with $\max_{1\le i\le(3k_1+k_2)} |\lambda_i(\mathbf{W})| < 1$, the upper bound of $\rho(\mathbf{A}_\theta)$ changes as $\theta$ increases if we keep $\max_{1\le i\le k_1} |\lambda_i(\mathbf{A}_{11})|$ to be fixed. For example, for settings 1 and 2 with dense $\mathbf{A}_{11}$ block, $\max_{1\le i\le k_1} |\lambda_i(\mathbf{A}_{11})|$ is fixed at 0.3 for $\theta = 0.2$ and $\theta = 0.5$ and other entries of $\mathbf{A}$ are adjusted in order to attain the desired signal. When $\theta$ increases to 0.8, $\max_{1\le i\le k_1} |\lambda_i(\mathbf{A}_{11})|$ is kept fixed at 0.1 in order to provide a reasonably good signal strength to $\mathbf{A}$.

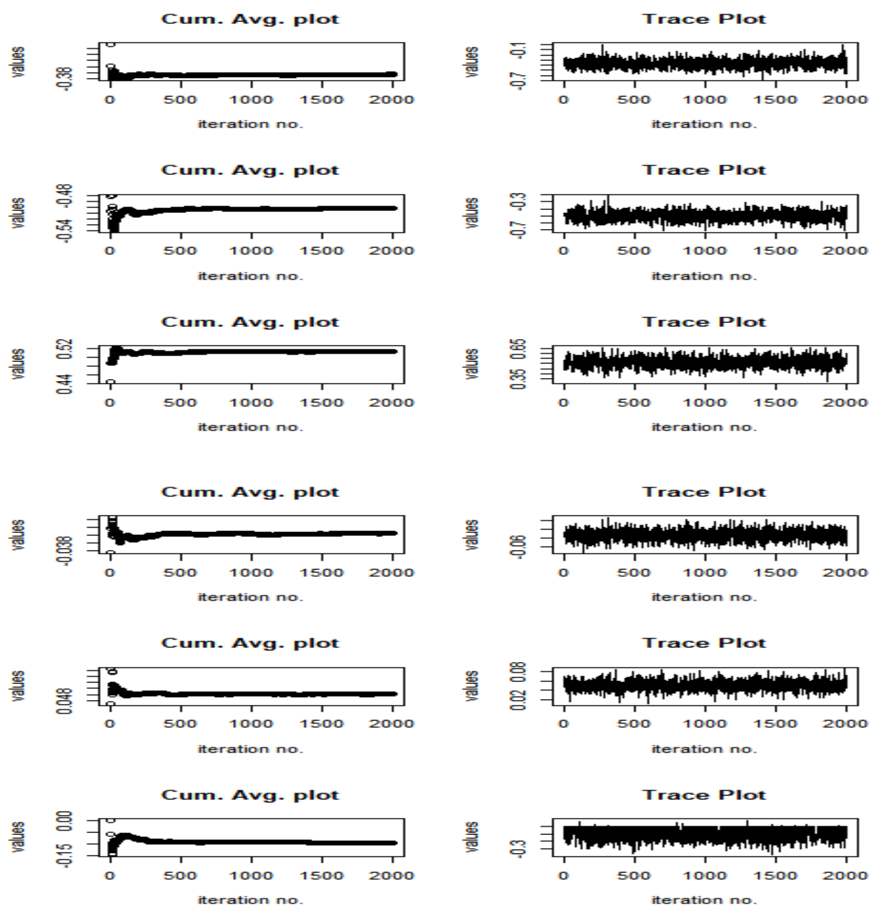


Figure 3: MCMC Diagnostics for some selected entries of $\mathbf{A}$ in the simulation setting $k_1 = 5, k_2 = 50$

## S6.4 Model selection and estimation consistency results for the Bayesian MF model

Table 2 and Table 3 provide the model selection and estimation performance results of the Bayesian MF model respectively for $\theta = 0.2, 0.5$ and 0.8, as mentioned in Section 6 of the main paper.

## S6.5 Additional nowcasting/forecasting results when the data are generated from the Bayesian MF model

Table 4 gives additional nowcasting/forecasting results for MFBVAR model corresponding to other priors (except the best performing one) when the data are generated from the Bayesian MF model and the true error covariance has a block diagonal structure. Table 5 provides the CRPS and LPS values for BMF and the best performing MFBVAR model that corresponds to the Minnesota prior, for all the settings.

## S6.6 Nowcasting/forecasting results when the data are generated from the Bayesian MF model and the true error covariance is not block-diagonal

Table 6 provides the relative RMSE values for the BMF, MFBVAR model with Minnesota (Minn), Steady-state (SS) and Hierarchical steady-state

Table 2: Sensitivity/Specificity for Bayesian MF for different values of $k_1, k_2, n$ & $\theta$

| $\theta = 0.2$ | $k_1 = 3, k_2 = 30$ | | $k_1 = 5, k_2 = 50$ | | $k_1 = 10, k_2 = 50$ | | $k_1 = 20, k_2 = 20$ | | $k_1 = 30, k_2 = 10$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| Sam. size | 100 | 150 | 200 | 400 | 200 | 400 | 100 | 150 | 100 | 150 |
| SN($\mathbf{A}_{11}$) | NA | NA | NA | NA | 0.66 | 0.87 | 0.87 | 0.90 | 0.79 | 0.85 |
| SP($\mathbf{A}_{11}$) | NA | NA | NA | NA | 0.99 | 0.99 | 0.92 | 0.93 | 0.92 | 0.93 |
| SN($\mathbf{A}_{12}$) | 1 | 1 | 1 | 1 | 1 | 1 | 0.95 | 0.96 | 0.99 | 1 |
| SP($\mathbf{A}_{12}$) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| SN($\mathbf{A}_{21}$) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.97 | 0.99 |
| SP($\mathbf{A}_{21}$) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| SN($\mathbf{A}_{22}$) | 0.96 | 0.99 | 0.98 | 1 | 0.96 | 0.98 | 0.96 | 0.97 | 1 | 1 |
| SP($\mathbf{A}_{22}$) | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.97 | 0.97 | 0.98 | 0.98 |
| SN($\mathbf{A}$) | 0.97 | 0.99 | 0.98 | 1 | 0.95 | 0.98 | 0.93 | 0.95 | 0.86 | 0.90 |
| SP($\mathbf{A}$) | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 1 | 0.97 | 0.98 | 0.95 | 0.96 |
| $\theta = 0.5$ | | | | | | | | | | |
| SN($\mathbf{A}_{11}$) | NA | NA | NA | NA | 0.90 | 0.99 | 0.92 | 0.94 | 0.87 | 0.90 |
| SP($\mathbf{A}_{11}$) | NA | NA | NA | NA | 1 | 1 | 0.93 | 0.94 | 0.92 | 0.93 |
| SN($\mathbf{A}_{12}$) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| SP($\mathbf{A}_{12}$) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| SN($\mathbf{A}_{21}$) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| SP($\mathbf{A}_{21}$) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| SN($\mathbf{A}_{22}$) | 0.90 | 0.95 | 0.97 | 0.99 | 0.94 | 0.97 | 0.84 | 0.92 | 0.96 | 0.98 |
| SP($\mathbf{A}_{22}$) | 0.97 | 0.97 | 0.98 | 0.98 | 0.98 | 0.99 | 0.97 | 0.97 | 0.96 | 0.97 |
| SN($\mathbf{A}$) | 0.92 | 0.97 | 0.97 | 0.99 | 0.95 | 0.98 | 0.91 | 0.95 | 0.89 | 0.92 |
| SP($\mathbf{A}$) | 0.98 | 0.98 | 0.98 | 0.99 | 0.99 | 0.99 | 0.97 | 0.98 | 0.95 | 0.96 |
| $\theta = 0.8$ | | | | | | | | | | |
| SN($\mathbf{A}_{11}$) | NA | NA | NA | NA | 0.98 | 1 | 0.89 | 0.93 | 0.79 | 0.85 |
| SP($\mathbf{A}_{11}$) | NA | NA | NA | NA | 1 | 1 | 0.97 | 0.97 | 0.97 | 0.97 |
| SN($\mathbf{A}_{12}$) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.98 | 1 |
| SP($\mathbf{A}_{12}$) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| SN($\mathbf{A}_{21}$) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0.99 | 1 |
| SP($\mathbf{A}_{21}$) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| SN($\mathbf{A}_{22}$) | 0.95 | 0.98 | 0.96 | 0.99 | 0.92 | 0.97 | 0.74 | 0.83 | 1 | 1 |
| SP($\mathbf{A}_{22}$) | 0.95 | 0.95 | 0.98 | 0.98 | 0.96 | 0.97 | 0.97 | 0.97 | 0.94 | 0.95 |
| SN($\mathbf{A}$) | 0.97 | 0.99 | 0.97 | 0.99 | 0.93 | 0.97 | 0.84 | 0.89 | 0.86 | 0.90 |
| SP($\mathbf{A}$) | 0.96 | 0.96 | 0.98 | 0.99 | 0.97 | 0.98 | 0.98 | 0.99 | 0.98 | 0.98 |

Table 3: Relative Error in Bayesian MF for different combinations of $k_1, k_2, n$ and $\theta$

| $\theta = 0.2$ | $k_1 = 3, k_2 = 30$ | | $k_1 = 5, k_2 = 50$ | | $k_1 = 10, k_2 = 50$ | | $k_1 = 20, k_2 = 20$ | | $k_1 = 30, k_2 = 10$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| Sam. size | 100 | 150 | 200 | 400 | 200 | 400 | 100 | 150 | 100 | 150 |
| Rel. Err($\mathbf{A}_{11}$) | 0.66 | 0.55 | 0.44 | 0.33 | 0.66 | 0.44 | 0.65 | 0.54 | 0.79 | 0.65 |
| Rel. Err($\mathbf{A}_{12}$) | 0.13 | 0.09 | 0.07 | 0.04 | 0.11 | 0.08 | 0.22 | 0.19 | 0.17 | 0.12 |
| Rel. Err($\mathbf{A}_{21}$) | 0.21 | 0.17 | 0.12 | 0.08 | 0.13 | 0.09 | 0.20 | 0.14 | 0.22 | 0.16 |
| Rel. Err($\mathbf{A}_{22}$) | 0.26 | 0.21 | 0.20 | 0.13 | 0.18 | 0.12 | 0.37 | 0.30 | 0.41 | 0.33 |
| Rel. Err($\mathbf{A}$) | 0.26 | 0.20 | 0.20 | 0.13 | 0.17 | 0.12 | 0.39 | 0.33 | 0.47 | 0.38 |
| Rel. Err($\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$) | 0.18 | 0.14 | 0.11 | 0.08 | 0.12 | 0.08 | 0.20 | 0.16 | 0.22 | 0.18 |
| Est. of $\theta$ | 0.18 | 0.18 | 0.17 | 0.16 | 0.20 | 0.20 | 0.11 | 0.11 | 0.11 | 0.11 |
| $\theta = 0.5$ | | | | | | | | | | |
| Rel. Err($\mathbf{A}_{11}$) | 0.52 | 0.41 | 0.33 | 0.23 | 0.34 | 0.17 | 0.48 | 0.40 | 0.60 | 0.48 |
| Rel. Err($\mathbf{A}_{12}$) | 0.16 | 0.12 | 0.09 | 0.06 | 0.10 | 0.06 | 0.21 | 0.18 | 0.21 | 0.16 |
| Rel. Err($\mathbf{A}_{21}$) | 0.24 | 0.14 | 0.09 | 0.06 | 0.11 | 0.07 | 0.18 | 0.18 | 0.20 | 0.18 |
| Rel. Err($\mathbf{A}_{22}$) | 0.40 | 0.32 | 0.30 | 0.20 | 0.27 | 0.18 | 0.55 | 0.45 | 0.83 | 0.68 |
| Rel. Err($\mathbf{A}$) | 0.39 | 0.31 | 0.29 | 0.19 | 0.24 | 0.16 | 0.43 | 0.36 | 0.54 | 0.43 |
| Rel. Err($\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$) | 0.17 | 0.14 | 0.11 | 0.08 | 0.12 | 0.08 | 0.20 | 0.16 | 0.21 | 0.17 |
| Est. of $\theta$ | 0.50 | 0.50 | 0.51 | 0.50 | 0.51 | 0.51 | 0.50 | 0.51 | 0.48 | 0.49 |
| $\theta = 0.8$ | | | | | | | | | | |
| Rel. Err($\mathbf{A}_{11}$) | 0.71 | 0.54 | 0.62 | 0.40 | 0.19 | 0.13 | 0.40 | 0.32 | 0.60 | 0.48 |
| Rel. Err($\mathbf{A}_{12}$) | 0.13 | 0.10 | 0.08 | 0.06 | 0.10 | 0.06 | 0.21 | 0.15 | 0.17 | 0.14 |
| Rel. Err($\mathbf{A}_{21}$) | 0.08 | 0.05 | 0.07 | 0.04 | 0.10 | 0.06 | 0.18 | 0.15 | 0.17 | 0.12 |
| Rel. Err($\mathbf{A}_{22}$) | 0.43 | 0.34 | 0.30 | 0.20 | 0.41 | 0.27 | 0.69 | 0.59 | 0.79 | 0.64 |
| Rel. Err($\mathbf{A}$) | 0.40 | 0.32 | 0.28 | 0.19 | 0.37 | 0.24 | 0.51 | 0.43 | 0.37 | 0.29 |
| Rel. Err($\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$) | 0.17 | 0.14 | 0.11 | 0.08 | 0.12 | 0.08 | 0.20 | 0.16 | 0.21 | 0.17 |
| Est. of $\theta$ | 0.81 | 0.81 | 0.80 | 0.80 | 0.82 | 0.81 | 0.79 | 0.81 | 0.78 | 0.79 |

Table 4: Relative RMSE values for the MFBVAR model with steady-state (SS) and hierarchical steady-state (Hier. SS) prior using data generated from the Bayesian MF model when true error covariance is block-diagonal

| $\theta = 0.5$ | $k_1 = 3, k_2 = 30$ | | $k_1 = 5, k_2 = 50$ | | $k_1 = 10, k_2 = 50$ | | $k_1 = 20, k_2 = 20$ | | $k_1 = 30, k_2 = 10$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| | SS | Hier. SS | SS | Hier. SS | SS | Hier. SS | SS | Hier. SS | SS | Hier. SS |
| $h = 1/3$ | 2.93 | 1.74 | 2.47 | 2.66 | 2.12 | 1.76 | 0.99 | 1.10 | 1.03 | 0.94 |
| $h = 2/3$ | 2.94 | 1.69 | 2.41 | 2.64 | 2.06 | 1.77 | 0.89 | 1.03 | 0.94 | 0.86 |
| $h = 1$ | 2.87 | 1.56 | 2.46 | 2.54 | 2.03 | 1.70 | 0.87 | 0.99 | 0.91 | 0.85 |
| $h = 4/3$ | 2.86 | 1.63 | 2.36 | 2.51 | 1.87 | 1.57 | 0.89 | 1.05 | 0.82 | 0.84 |
| $h = 5/3$ | 2.85 | 1.58 | 2.30 | 2.50 | 1.80 | 1.56 | 0.86 | 1.02 | 0.79 | 0.81 |
| $h = 2$ | 2.90 | 1.50 | 2.30 | 2.43 | 1.82 | 1.51 | 0.83 | 0.97 | 0.78 | 0.81 |

Table 5: CRPS and LPS values of BMF and MFBVAR using data generated from the proposed BMF model when true error covariance is block-diagonal

| CRPS | $k_1 = 3, k_2 = 30$ | | $k_1 = 5, k_2 = 50$ | | $k_1 = 10, k_2 = 50$ | | $k_1 = 20, k_2 = 20$ | | $k_1 = 30, k_2 = 10$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| | BMF | MFBVAR | BMF | MFBVAR | BMF | MFBVAR | BMF | MFBVAR | BMF | MFBVAR |
| $h = 1/3$ | 0.29 | 0.86 | 0.30 | 1.12 | 0.34 | 0.91 | 0.25 | 0.39 | 0.22 | 0.30 |
| $h = 2/3$ | 0.29 | 0.79 | 0.30 | 1.10 | 0.32 | 0.86 | 0.25 | 0.36 | 0.22 | 0.27 |
| $h = 1$ | 0.29 | 0.72 | 0.29 | 1.08 | 0.31 | 0.83 | 0.24 | 0.34 | 0.22 | 0.26 |
| $h = 4/3$ | 0.34 | 0.75 | 0.36 | 1.04 | 0.40 | 0.78 | 0.26 | 0.32 | 0.23 | 0.25 |
| $h = 5/3$ | 0.34 | 0.72 | 0.36 | 0.99 | 0.40 | 0.76 | 0.26 | 0.31 | 0.23 | 0.25 |
| $h = 2$ | 0.34 | 0.67 | 0.35 | 0.98 | 0.38 | 0.72 | 0.26 | 0.31 | 0.23 | 0.25 |
| Log score | | | | | | | | | | |
| $h = 1/3$ | 0.74 | 1.81 | 0.75 | 2.08 | 0.85 | 1.90 | 0.58 | 1.04 | 0.49 | 0.75 |
| $h = 2/3$ | 0.74 | 1.73 | 0.74 | 2.03 | 0.81 | 1.80 | 0.57 | 0.95 | 0.48 | 0.65 |
| $h = 1$ | 0.73 | 1.66 | 0.73 | 2.00 | 0.77 | 1.77 | 0.54 | 0.92 | 0.48 | 0.62 |
| $h = 4/3$ | 0.90 | 1.68 | 0.93 | 1.98 | 1.04 | 1.72 | 0.62 | 0.91 | 0.50 | 0.60 |
| $h = 5/3$ | 0.90 | 1.68 | 0.92 | 1.93 | 1.03 | 1.71 | 0.62 | 0.88 | 0.49 | 0.59 |
| $h = 2$ | 0.90 | 1.65 | 0.92 | 1.92 | 0.99 | 1.68 | 0.61 | 0.87 | 0.49 | 0.61 |

prior (Hier. SS) and MIDAS regression models for all the settings when the data are generated from the BMF model and the true error covariance does not have any specific structure. It can be clearly seen that BMF outperforms all other models across all the settings even when the true error covariance does not have a block diagonal structure. Table 7 provides the CRPS and LPS values for BMF and the best performing MFBVAR model that corresponds to the Minnesota prior, across all the settings.

## S6.7 Additional nowcasting/forecasting results for data generated by Ghysels (2016)'s model

Table 8 gives additional nowcasting/forecasting results for MFBVAR model (except the best performing one) when data are generated from Ghysels (2016)'s model. Table 8 shows that some of the relative RMSE values corresponding to Steady-state and Hierarchical steady state prior are quite large. We examined this carefully and found that, for few replicates these priors produce extremely high RMSE values whereas for other replicates the RMSE values are quite low, which in turn produces these high relative RMSE values, while we take average over the replicates. We also found that for $10 - 15\%$ of cases the MFBVAR model with Steady-state and Hierarchical steady state prior did not produce any result and stopped

Table 6: Relative RMSE values (benchmarked to a random walk model with drift) using data generated from the proposed BMF model when the true error covariance is not block-diagonal

| $k_1 = 3, k_2 = 30$ | $h = 1/3$ | $h = 2/3$ | $h = 1$ | $h = 4/3$ | $h = 5/3$ | $h = 2$ |
|---|---|---|---|---|---|---|
| BMF | 0.50 | 0.51 | 0.50 | 0.59 | 0.59 | 0.59 |
| MFBVAR (Minn) | 1.98 | 1.86 | 1.76 | 1.91 | 1.77 | 1.53 |
| MFBVAR (SS) | 2.04 | 1.90 | 1.76 | 2.01 | 1.83 | 1.74 |
| MFBVAR (Hier. SS) | 2.25 | 2.22 | 2.14 | 2.23 | 2.22 | 2.08 |
| U-MIDAS | 0.91 | 0.92 | 0.94 | 1.01 | 1.01 | 1.01 |
| MIDAS(Res.) | 0.90 | 0.92 | 0.93 | 0.98 | 0.99 | 0.98 |
| $k_1 = 5, k_2 = 50$ | | | | | | |
| BMF | 0.53 | 0.53 | 0.52 | 0.65 | 0.65 | 0.64 |
| MFBVAR (Minn) | 1.93 | 1.79 | 1.76 | 1.81 | 1.68 | 1.71 |
| MFBVAR (SS) | 1.83 | 1.78 | 1.75 | 1.68 | 1.70 | 1.65 |
| MFBVAR (Hier. SS) | 2.16 | 2.04 | 1.96 | 2.01 | 1.93 | 1.89 |
| U-MIDAS | 0.91 | 0.90 | 0.89 | 0.93 | 0.93 | 0.93 |
| MIDAS(Res.) | 0.63 | 0.63 | 0.62 | 0.65 | 0.65 | 0.65 |
| $k_1 = 10, k_2 = 50$ | | | | | | |
| BMF | 0.52 | 0.49 | 0.48 | 0.57 | 0.56 | 0.55 |
| MFBVAR (Minn) | 1.35 | 1.24 | 1.19 | 1.06 | 1.02 | 0.98 |
| MFBVAR (SS) | 2.88 | 2.76 | 2.74 | 2.47 | 2.45 | 2.43 |
| MFBVAR (Hier. SS) | 2.06 | 1.98 | 1.86 | 1.72 | 1.70 | 1.66 |
| U-MIDAS | 0.61 | 0.60 | 0.60 | 0.56 | 0.56 | 0.56 |
| MIDAS(Res.) | 0.59 | 0.58 | 0.56 | 0.54 | 0.54 | 0.53 |
| $k_1 = 20, k_2 = 20$ | | | | | | |
| BMF | 0.93 | 0.89 | 0.69 | 0.92 | 0.91 | 0.73 |
| MFBVAR (Minn) | 1.12 | 1.01 | 0.94 | 0.91 | 0.84 | 0.81 |
| MFBVAR (SS) | 1.09 | 1.00 | 0.92 | 0.90 | 0.83 | 0.83 |
| MFBVAR (Hier. SS) | 1.19 | 1.07 | 1.03 | 0.98 | 0.94 | 0.91 |
| U-MIDAS | 1.44 | 1.46 | 1.41 | 1.48 | 1.49 | 1.45 |
| MIDAS(Res.) | 0.91 | 0.95 | 0.93 | 1.05 | 1.04 | 0.99 |
| $k_1 = 30, k_2 = 10$ | | | | | | |
| BMF | 0.98 | 0.95 | 0.74 | 0.96 | 0.96 | 0.93 |
| MFBVAR (Minn) | 1.11 | 0.94 | 0.96 | 1.02 | 1.02 | 0.99 |
| MFBVAR (SS) | 1.09 | 0.97 | 1.00 | 1.07 | 1.06 | 1.02 |
| MFBVAR (Hier. SS) | 1.13 | 1.03 | 1.06 | 1.05 | 1.07 | 1.04 |
| U-MIDAS | 1.02 | 1.05 | 1.13 | 1.05 | 1.15 | 1.02 |
| MIDAS(Res.) | 0.81 | 0.78 | 0.82 | 0.77 | 0.87 | 0.82 |

Table 7: CRPS and LPS values of BMF and MFBVAR using data generated from the proposed BMF model when the true error covariance is not block-diagonal

| CRPS | $k_1 = 3, k_2 = 30$ | | $k_1 = 5, k_2 = 50$ | | $k_1 = 10, k_2 = 50$ | | $k_1 = 20, k_2 = 20$ | | $k_1 = 30, k_2 = 10$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| | BMF | MFBVAR | BMF | MFBVAR | BMF | MFBVAR | BMF | MFBVAR | BMF | MFBVAR |
| $h = 1/3$ | 0.28 | 1.05 | 0.32 | 1.14 | 0.35 | 0.92 | 0.36 | 0.43 | 0.34 | 0.41 |
| $h = 2/3$ | 0.28 | 1.01 | 0.32 | 1.06 | 0.34 | 0.83 | 0.35 | 0.39 | 0.33 | 0.35 |
| $h = 1$ | 0.28 | 0.94 | 0.32 | 1.03 | 0.33 | 0.81 | 0.27 | 0.37 | 0.27 | 0.35 |
| $h = 4/3$ | 0.31 | 0.95 | 0.38 | 1.03 | 0.42 | 0.79 | 0.36 | 0.36 | 0.37 | 0.40 |
| $h = 5/3$ | 0.31 | 0.90 | 0.38 | 0.97 | 0.41 | 0.76 | 0.35 | 0.34 | 0.37 | 0.40 |
| $h = 2$ | 0.31 | 0.80 | 0.38 | 0.98 | 0.40 | 0.74 | 0.28 | 0.33 | 0.36 | 0.39 |
| Log score | | | | | | | | | | |
| $h = 1/3$ | 0.73 | 2.00 | 0.86 | 2.12 | 0.95 | 1.91 | 0.96 | 1.14 | 0.93 | 1.10 |
| $h = 2/3$ | 0.73 | 1.96 | 0.86 | 2.02 | 0.91 | 1.79 | 0.93 | 1.04 | 0.91 | 0.97 |
| $h = 1$ | 0.72 | 1.91 | 0.85 | 1.99 | 0.89 | 1.78 | 0.68 | 1.01 | 0.70 | 0.99 |
| $h = 4/3$ | 0.82 | 1.91 | 1.02 | 1.98 | 1.10 | 1.76 | 0.95 | 1.00 | 0.98 | 1.06 |
| $h = 5/3$ | 0.82 | 1.87 | 1.02 | 1.94 | 1.09 | 1.75 | 0.95 | 0.98 | 0.97 | 1.07 |
| $h = 2$ | 0.81 | 1.80 | 1.01 | 1.96 | 1.07 | 1.73 | 0.73 | 0.97 | 0.95 | 1.05 |

execution in between. This occurred irrespective of whether data were generated from Bayesian MF or Ghysels 2016's model. Hence in those cases the results were based on averages of rest $\sim 85\%$ cases. Also, Table 9 provides the CRPS and LPS values for BMF and the best performing MFBVAR model that corresponds to the Minnesota prior, across all the settings.

Table 8: Relative RMSE values for the MFBVAR model with different prior distributions (Minnesota (Minn), steady-state (SS) or hierarchical steady-state (Hier. SS)) under a neutral data generating setting

| | $k_1 = 3, k_2 = 30$ | | $k_1 = 5, k_2 = 50$ | | $k_1 = 10, k_2 = 50$ | | $k_1 = 20, k_2 = 20$ | | $k_1 = 30, k_2 = 10$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Minn | SS | SS | Hier. SS | SS | Hier. SS | SS | Hier. SS | SS | Hier. SS |
| $h = 1/3$ | 2.89 | 904.03 | 4.49 | 171.70 | 7.47 | 10.20 | 1.41 | 1.46 | 0.86 | 0.73 |
| $h = 2/3$ | 2.77 | 664.57 | 5.86 | 178.67 | 8.27 | 11.23 | 1.23 | 1.24 | 0.81 | 0.67 |
| $h = 1$ | 2.67 | 888.04 | 4.62 | 175.44 | 7.03 | 11.49 | 1.16 | 1.14 | 0.77 | 0.65 |
| $h = 4/3$ | 2.58 | 949.02 | 4.25 | 167.32 | 7.72 | 10.64 | 1.11 | 1.18 | 0.93 | 0.81 |
| $h = 5/3$ | 2.65 | 697.72 | 5.57 | 174.29 | 8.56 | 11.72 | 1.08 | 1.12 | 0.88 | 0.74 |
| $h = 2$ | 2.53 | 932.21 | 4.36 | 170.95 | 7.27 | 12.01 | 1.08 | 1.11 | 0.89 | 0.75 |

Table 9: CRPS and LPS values of BMF and MFBVAR using data generated under a neutral data generating setting

| CRPS | $k_1 = 3, k_2 = 30$ | | $k_1 = 5, k_2 = 50$ | | $k_1 = 10, k_2 = 50$ | | $k_1 = 20, k_2 = 20$ | | $k_1 = 30, k_2 = 10$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| | BMF | MFBVAR | BMF | MFBVAR | BMF | MFBVAR | BMF | MFBVAR | BMF | MFBVAR |
| $h = 1/3$ | 0.18 | 0.70 | 0.17 | 0.62 | 0.16 | 0.54 | 0.14 | 0.28 | 0.14 | 0.17 |
| $h = 2/3$ | 0.19 | 0.65 | 0.17 | 0.57 | 0.16 | 0.49 | 0.14 | 0.23 | 0.14 | 0.16 |
| $h = 1$ | 0.19 | 0.64 | 0.17 | 0.56 | 0.16 | 0.46 | 0.14 | 0.21 | 0.14 | 0.15 |
| $h = 4/3$ | 0.19 | 0.57 | 0.16 | 0.51 | 0.16 | 0.45 | 0.15 | 0.20 | 0.16 | 0.16 |
| $h = 5/3$ | 0.19 | 0.59 | 0.16 | 0.52 | 0.16 | 0.43 | 0.15 | 0.19 | 0.15 | 0.15 |
| $h = 2$ | 0.19 | 0.57 | 0.16 | 0.50 | 0.15 | 0.40 | 0.15 | 0.18 | 0.15 | 0.16 |
| Log score | | | | | | | | | | |
| $h = 1/3$ | 0.28 | 1.60 | 0.19 | 1.53 | 0.13 | 1.43 | 0.03 | 0.69 | 0.02 | 0.17 |
| $h = 2/3$ | 0.30 | 1.50 | 0.19 | 1.42 | 0.13 | 1.27 | 0.03 | 0.51 | 0.02 | 0.14 |
| $h = 1$ | 0.31 | 1.48 | 0.18 | 1.38 | 0.13 | 1.19 | 0.03 | 0.46 | 0.03 | 0.11 |
| $h = 4/3$ | 0.31 | 1.39 | 0.16 | 1.30 | 0.11 | 1.17 | 0.05 | 0.42 | 0.10 | 0.18 |
| $h = 5/3$ | 0.32 | 1.40 | 0.16 | 1.31 | 0.11 | 1.12 | 0.04 | 0.39 | 0.10 | 0.15 |
| $h = 2$ | 0.32 | 1.38 | 0.16 | 1.29 | 0.11 | 1.07 | 0.04 | 0.39 | 0.10 | 0.17 |

## S6.8 Nowcasting/forecasting results for data generated from Schorfheide and Song (2015)'s model

Next, we generate data at a monthly level (as assumed in Schorfheide and Song (2015)'s model). In particular, we generate a $3(k_1 + k_2)$ dimensional true VAR transition matrix with non-zero entries drawn from Unif $(0, 10) \cup$ Unif $(-10, 0)$. After generating data using this transition matrix, we then discard the observations corresponding to the 1st and 2nd months of each quarter for the quarterly variables, pretending that they were unobserved, which leads to a $(3k_1 + k_2)$ dimensional data set. For the generated data, BMF and other competing model provide nowcasts/forecasts across the forecasting horizon $h = 1/3, 2/3, 1, 4/3, 5/3, 2$. Table 10 provides the relative RMSE values for the BMF, MFBVAR model with Minnesota (Minn), Steady-state (SS) and Hierarchical steady-state prior (Hier. SS) and MIDAS regression models for all the settings. It can be clearly seen that even in this setup our model performs significantly better in terms of out-of-sample forecasts than other competing models. Table 11 provides the CRPS and LPS values for BMF and the best performing MFBVAR model that corresponds to the Minnesota prior, across all the settings.

Table 10: Relative RMSE values (benchmarked to a random walk model with drift) using data generated from Schorfheide and Song (2015)'s model

| $k_1 = 3, k_2 = 30$ | $h = 1/3$ | $h = 2/3$ | $h = 1$ | $h = 4/3$ | $h = 5/3$ | $h = 2$ |
|---|---|---|---|---|---|---|
| BMF | 0.67 | 0.67 | 0.68 | 0.72 | 0.72 | 0.72 |
| MFBVAR (Minn) | 2.60 | 2.43 | 2.35 | 2.17 | 2.07 | 2.07 |
| MFBVAR (SS) | 2.79 | 2.53 | 2.48 | 2.49 | 2.46 | 2.36 |
| MFBVAR (Hier. SS) | 1106.48 | 1070.01 | 1347.61 | 1124.84 | 1088.12 | 1369.98 |
| U-MIDAS | 0.88 | 0.90 | 0.88 | 0.92 | 0.93 | 0.94 |
| MIDAS(Res.) | 0.90 | 0.91 | 0.89 | 0.87 | 0.88 | 0.88 |
| $k_1 = 5, k_2 = 50$ | | | | | | |
| BMF | 0.73 | 0.73 | 0.73 | 0.71 | 0.71 | 0.71 |
| MFBVAR (Minn) | 2.85 | 2.72 | 2.62 | 2.56 | 2.44 | 2.36 |
| MFBVAR (SS) | 83.84 | 83.44 | 74.11 | 83.94 | 83.51 | 74.19 |
| MFBVAR (Hier. SS) | 356.68 | 362.33 | 309.61 | 357.52 | 363.16 | 310.41 |
| U-MIDAS | 0.89 | 0.89 | 0.89 | 0.88 | 0.88 | 0.87 |
| MIDAS(Res.) | 0.85 | 0.85 | 0.85 | 0.85 | 0.86 | 0.85 |
| $k_1 = 10, k_2 = 50$ | | | | | | |
| BMF | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 | 0.63 |
| MFBVAR (Minn) | 2.01 | 1.92 | 1.84 | 1.71 | 1.55 | 1.53 |
| MFBVAR (SS) | 233.81 | 201.25 | 162.84 | 228.64 | 196.82 | 159.22 |
| MFBVAR (Hier. SS) | 105.00 | 91.45 | 117.35 | 102.70 | 89.52 | 114.90 |
| U-MIDAS | 1.21 | 1.23 | 1.21 | 1.15 | 1.15 | 1.21 |
| MIDAS(Res.) | 0.92 | 0.97 | 0.89 | 0.84 | 0.85 | 0.88 |
| $k_1 = 20, k_2 = 20$ | | | | | | |
| BMF | 0.69 | 0.69 | 0.69 | 0.68 | 0.68 | 0.68 |
| MFBVAR (Minn) | 1.34 | 1.10 | 1.06 | 0.90 | 0.87 | 0.81 |
| MFBVAR (SS) | 1.37 | 1.10 | 1.08 | 0.98 | 0.94 | 0.90 |
| MFBVAR (Hier. SS) | 1.56 | 1.32 | 1.39 | 1.27 | 1.25 | 1.23 |
| U-MIDAS | 1.70 | 1.61 | 1.59 | 1.56 | 1.61 | 1.65 |
| MIDAS(Res.) | 0.99 | 1.01 | 1.00 | 1.05 | 1.03 | 1.11 |
| $k_1 = 30, k_2 = 10$ | | | | | | |
| BMF | 0.72 | 0.72 | 0.71 | 0.67 | 0.67 | 0.67 |
| MFBVAR (Minn) | 0.94 | 0.84 | 0.75 | 0.76 | 0.74 | 0.69 |
| MFBVAR (SS) | 0.96 | 0.87 | 0.78 | 0.81 | 0.79 | 0.76 |
| MFBVAR (Hier. SS) | 1.14 | 1.07 | 1.00 | 0.94 | 0.94 | 0.89 |
| U-MIDAS | 1.21 | 1.23 | 1.21 | 1.15 | 1.15 | 1.21 |
| MIDAS(Res.) | 0.92 | 0.97 | 0.89 | 0.84 | 0.85 | 0.88 |

Table 11: CRPS and LPS values of BMF and MFBVAR using data generated from Schorfheide and Song (2015)'s model

| CRPS | $k_1 = 3, k_2 = 30$ | | $k_1 = 5, k_2 = 50$ | | $k_1 = 10, k_2 = 50$ | | $k_1 = 20, k_2 = 20$ | | $k_1 = 30, k_2 = 10$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| | BMF | MFBVAR | BMF | MFBVAR | BMF | MFBVAR | BMF | MFBVAR | BMF | MFBVAR |
| $h = 1/3$ | 0.17 | 0.62 | 0.17 | 0.66 | 0.16 | 0.92 | 0.16 | 0.31 | 0.18 | 0.24 |
| $h = 2/3$ | 0.17 | 0.59 | 0.17 | 0.63 | 0.16 | 0.83 | 0.16 | 0.26 | 0.18 | 0.21 |
| $h = 1$ | 0.17 | 0.57 | 0.17 | 0.60 | 0.16 | 0.81 | 0.16 | 0.25 | 0.17 | 0.19 |
| $h = 4/3$ | 0.18 | 0.52 | 0.17 | 0.59 | 0.16 | 0.79 | 0.16 | 0.22 | 0.17 | 0.20 |
| $h = 5/3$ | 0.18 | 0.50 | 0.17 | 0.56 | 0.16 | 0.76 | 0.16 | 0.21 | 0.17 | 0.20 |
| $h = 2$ | 0.18 | 0.49 | 0.17 | 0.54 | 0.16 | 0.74 | 0.16 | 0.20 | 0.17 | 0.19 |
| Log score | | | | | | | | | | |
| $h = 1/3$ | 0.20 | 1.54 | 0.22 | 1.64 | 0.17 | 1.91 | 0.13 | 0.77 | 0.20 | 0.52 |
| $h = 2/3$ | 0.20 | 1.44 | 0.22 | 1.53 | 0.17 | 1.79 | 0.12 | 0.61 | 0.20 | 0.43 |
| $h = 1$ | 0.20 | 1.39 | 0.22 | 1.45 | 0.17 | 1.78 | 0.12 | 0.58 | 0.20 | 0.36 |
| $h = 4/3$ | 0.23 | 1.31 | 0.19 | 1.42 | 0.17 | 1.76 | 0.14 | 0.51 | 0.21 | 0.40 |
| $h = 5/3$ | 0.23 | 1.26 | 0.19 | 1.37 | 0.17 | 1.75 | 0.14 | 0.50 | 0.21 | 0.39 |
| $h = 2$ | 0.23 | 1.26 | 0.19 | 1.35 | 0.17 | 1.73 | 0.14 | 0.48 | 0.21 | 0.37 |

## S6.9 Comparison of nowcasting/forecasting performance of BMF and MFBVAR in a very low dimensional setting

In our high/moderate dimensional simulations, the forecasting performance of the proposed BMF approach is better than the state-space based MF-BVAR approach uniformly across different data generating mechanisms, including the BMF, the Ghysels (2016) and a monthly VAR fitted into a quarterly level mechanisms. Following a reviewer's recommendation, we performed simulations for a very low-dimensional setting where $k_1 = k_2 = 2$ and $n = 20$. We first consider a scenario when the data are generated from the model in Ghysels (2016) and second, when data are generated from the

BMF model, but the true coefficient matrix $\mathbf{W}$ deviates from the model specified structure in the sense that the dampening effect of different high frequency variables is different. Then, the performance of the MFBVAR model becomes slightly better than the proposed BMF model in more important short-term forecasting horizon in this very low dimensional setting ($k_1 = k_2 = 2$), which is expected. The full forecasting results in terms of relative RMSE values, CRPS and LPS values are provided in Table 12. As before, we consider a random walk model with drift as a benchmark for obtaining the relative RMSE values. For the MFBVAR model, we report all the metrics for the best performing MFBVAR model w.r.t the choice of priors which corresponds to the hierarchical steady-state prior in this setting. In Table 12, case I corresponds to the scenario when the data are generated from the model in Ghysels (2016) and case II corresponds to the scenario when data are generated from BMF and the true theta =0.5, except for one high-frequency variable for which the dampening rate is 0.01. This validates the intuition that while the structure underlying the proposed BMF model controls parameter proliferation and provides better results for sample starved high-dimensional settings, these advantages may not hold in low-dimensional settings.

Table 12: Relative RMSE, CRPS and LPS values of BMF and MFBVAR when $k_1 = k_2 = 2$ when (I) data is generated from Ghysels (2016)'s model or (II) from a variant of BMF that uses different $\theta$ for different high-frequency variables for constructing $\mathbf{W}$

| Case I: | BMF | | | MFBVAR | | |
|---|---|---|---|---|---|---|
| | Rel. RMSE | CRPS | LPS | Rel. RMSE | CRPS | LPS |
| $h = 1/3$ | 0.82 | 0.29 | 0.78 | 0.79 | 0.28 | 0.75 |
| $h = 2/3$ | 1.05 | 0.38 | 1.15 | 0.79 | 0.28 | 0.76 |
| $h = 1$ | 0.85 | 0.28 | 0.75 | 0.83 | 0.29 | 0.77 |
| $h = 4/3$ | 0.60 | 0.26 | 0.76 | 0.84 | 0.36 | 0.94 |
| $h = 5/3$ | 0.57 | 0.25 | 0.61 | 0.82 | 0.36 | 0.93 |
| $h = 2$ | 0.80 | 0.35 | 0.94 | 0.81 | 0.36 | 0.93 |
| Case II: | BMF | | | MFBVAR | | |
| | Rel. RMSE | CRPS | LPS | Rel. RMSE | CRPS | LPS |
| $h = 1/3$ | 0.81 | 0.14 | 0.06 | 0.34 | 0.10 | 0.06 |
| $h = 2/3$ | 0.73 | 0.12 | 0.05 | 0.26 | 0.10 | 0.04 |
| $h = 1$ | 0.36 | 0.09 | 0.23 | 0.40 | 0.12 | 0.09 |
| $h = 4/3$ | 1.45 | 0.26 | 0.58 | 1.18 | 0.21 | 0.41 |
| $h = 5/3$ | 1.49 | 0.27 | 0.62 | 1.24 | 0.20 | 0.39 |
| $h = 2$ | 1.15 | 0.22 | 0.44 | 1.17 | 0.19 | 0.34 |

# S7 Additional Empirical Data Analysis Results

The estimated Granger causal relationships among 77 monthly and 9 quarterly variables based on the sparsity pattern in $\mathbf{W}$ provided by the BMF approach are depicted in Figure 4. Table 16 shows the number of edges in different sub-blocks of the estimated $\mathbf{A}$ matrix. The monthly and quarterly variables are displayed by dark grey and light grey vertices respectively.

The names of the variables corresponding to the vertices of the network in Figure 4 are provided in Table 13 and Table 14.

The quarterly variables used in all the empirical data analysis are given in Table 13. The full set of monthly variables used in Data 3 are given in Table 14 and those which were part of Data 1 and Data 2 are marked using indicators. Table 15 shows the transformation needed for each economic series. We code the transformations as follows: (1) no transformation (2) $\triangle x_t$ (3) $\triangle \log(x_t)$ (4) $\triangle^2 \log(x_t)$.

Table 13: Quarterly variables used in the empirical data analysis and their abbreviations

| ID | Series abbreviation | Description |
|---|---|---|
| Q1 | GDPC1 | Real Gross Domestic Product |
| Q2 | PCECC96 | Real Personal Consumption Expenditures |
| Q3 | GPDIC1 | Real Gross Private Domestic Investment |
| Q4 | GCEC1 | Real Government Consumption Expenditures and Gross Investment |
| Q5 | IPDBS | Implicit Price Deflator |
| Q6 | RCPHBS | Real Compensation Per Hour |
| Q7 | OPHPBS | Real Output Per Hour of All Persons |
| Q8 | TABSHNO | Real Total Assets of Households and Nonprofit Organizations |
| Q9 | TLBSNNCB | Real Nonfinancial Corporate Business Sector Liabilities |

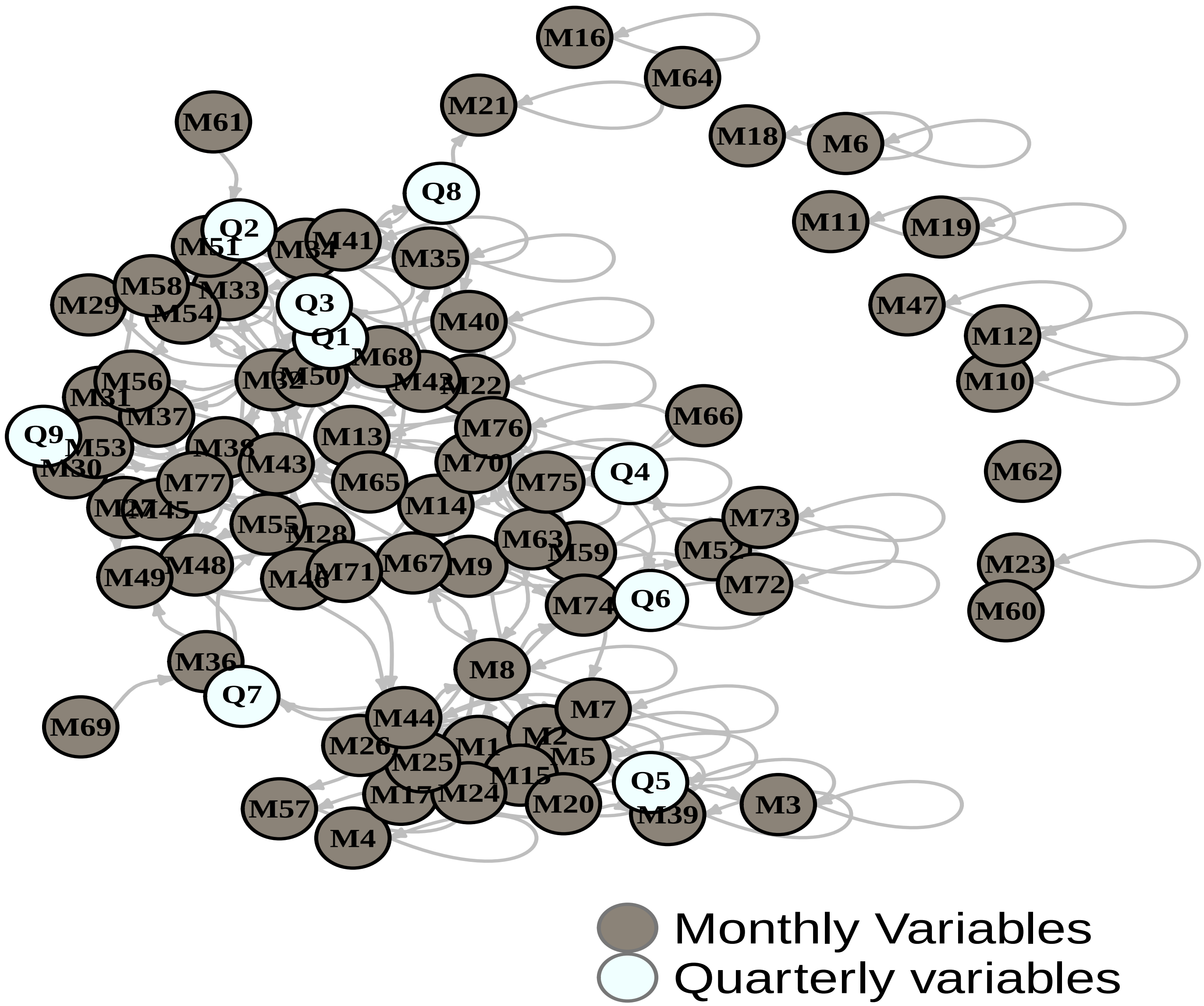


Figure 4: Network representation of the estimated transition matrix

Table 14: Monthly variables used in the empirical data analysis and their abbreviations

| ID | Series abbreviation | Description | Data 1 use | Data 2 use |
|---|---|---|---|---|
| M1 | CPIAUCSL | Consumer Price Index for All Urban Consumers: All Items in U.S. City Average | 1 | 1 |
| M4 | PCEPI | Personal Consumption Expenditures | 1 | 1 |
| M5 | CPITRNSL | Consumer Price Index for All Urban Consumers | 1 | 1 |
| M6 | DSERRG3M086SBEA | Personal consumption expenditures | 1 | 1 |
| M7 | MZMSL | MZM Money Stock | 1 | 1 |
| M8 | TOTRESNS | Total Reserves of Depository Institutions | 1 | 1 |
| M9 | BUSLOANS | Commercial and Industrial Loans, All Commercial Banks | 1 | 1 |
| M10 | NONREVSL | Nonrevolving Consumer Credit Owned and Securitized | 1 | 1 |
| M27 | INDPRO | Industrial Production: Total Index | 1 | 1 |
| M28 | IPDMAT | Industrial Production: Durable Goods Materials | 1 | 1 |

*Continued on next page*

Table 14 – *Continued from previous page*

| ID | Series abbreviation | Description | Data 1 use | Data 2 use |
|---|---|---|---|---|
| M29 | RPI | Real Personal Income | 1 | 1 |
| M30 | IPBUSEQ | Industrial Production: Equipment: Business Equipment | 1 | 1 |
| M31 | W875RX1 | Real Personal Income Excluding Current Transfer Receipts | 1 | 1 |
| M32 | DPCERA3M086SBEA | Real Personal Consumption Expenditures (chain-type quantity index) | 1 | 1 |
| M33 | USTPU | All Employees, Trade, Transportation, and Utilities | 1 | 1 |
| M34 | CE16OV | Employment Level | 1 | 1 |
| M35 | CLF16OV | Civilian Labor Force Level | 1 | 1 |
| M59 | FEDFUNDS | Effective Federal Funds Rate | 1 | 1 |
| M60 | TB3MS | 3-Month Treasury Bill | 1 | 1 |
| M62 | GS5 | 5-Year Treasury Constant Maturity Rate | 1 | 1 |
| M63 | GS10 | 10-Year Treasury Constant Maturity Rate | 1 | 1 |
| M64 | TB6MS | 6-Month Treasury Bill | 1 | 1 |
| M65 | UNRATE | Civilian Unemployment Rate | 1 | 1 |
| M77 | SP500 | S&P 500 | 1 | 1 |
| M18 | CPIMEDSL | Consumer Price Index for All Urban Consumers:Medical Care in U.S. City Average | | 1 |
| M11 | PPICMM | Producer Price Index by Commodity | | 1 |
| M15 | CUSR0000SA0L2 | Consumer Price Index for All Urban Consumers: All Items Less Shelter in U.S. City Average | | 1 |
| M17 | DNDGRG3M086SBEA | Personal consumption expenditures: Nondurable goods (chain-type price index) | | 1 |
| M12 | DTCTHFNM | Total Consumer Loans and Leases Owned and Securitized by Finance Companies, Level | | 1 |
| M13 | CES3000000008 | Average Hourly Earnings of Production and Nonsupervisory Employees, Manufacturing | | 1 |
| M14 | CES0600000008 | Average Hourly Earnings of Production and Nonsupervisory Employees, Goods-Producing | | 1 |
| M3 | CPIAPPSL | Consumer Price Index for All Urban Consumers: Apparel in U.S. City Average | | 1 |
| M2 | CPIULFSL | Consumer Price Index for All Urban Consumers: All Items Less Food in U.S. City Average | | 1 |
| M16 | DTCOLNVHFNM | Consumer Motor Vehicle Loans Owned by Finance Companies, Level | | 1 |
| M37 | IPFINAL | Industrial Production: Final Products | | 1 |
| M38 | IPFPNSS | Industrial Production: Final Products and Nonindustrial Supplies | | 1 |
| M40 | IPNCONGD | Industrial Production: Non-Durable Consumer Goods | | 1 |
| M44 | M2REAL | Real M2 Money Stock | | 1 |
| M41 | USFIRE | All Employees, Financial Activities | | 1 |
| M42 | USCONS | All Employees, Construction | | 1 |
| M39 | IPFUELS | Industrial Production: Non-Durable Consumer Energy Products: Fuels | | 1 |
| M36 | IPNMAT | Industrial Production: Non-Durable Goods Materials | | 1 |
| M45 | DMANEMP | All Employees, Durable Goods | | 1 |
| M66 | AAA | Moody's Seasoned Aaa Corporate Bond Yield | | 1 |
| M61 | GS1 | Market Yield on U.S. Treasury Securities at 1-Year Constant Maturity | | 1 |
| M69 | AWOTMAN | Average Weekly Overtime Hours of Production and Nonsupervisory Employees, Manufacturing | | 1 |
| M67 | BAA | Moody's Seasoned Baa Corporate Bond Yield | | 1 |
| M68 | UEMPMEAN | Average Weeks Unemployed | | 1 |
| M43 | USGOOD | All Employees, Goods-Producing | | 1 |
| M19 | DDURRG3M086SBEA | Personal consumption expenditures: Durable goods (chain-type price index) | | |
| M20 | M2SL | M2 money stock | | |
| M21 | INVEST | Securities in Bank Credit | | |
| M22 | CES2000000008 | Average Hourly Earnings of Production and Nonsupervisory Employees, Construction | | |
| M23 | CUSR0000SAS | Consumer Price Index for All Urban Consumers: Services in U.S. City Average | | |

*Continued on next page*

Table 14 – *Continued from previous page*

| ID | Series abbreviation | Description | Data 1 use | Data 2 use |
|---|---|---|---|---|
| M24 | CUSR0000SA0L5 | Consumer Price Index for All Urban Consumers: All Items Less Medical Care in U.S. City Average | | |
| M25 | CUSR0000SAC | Consumer Price Index for All Urban Consumers: Commodities in U.S. City Average | | |
| M26 | M1SL | M1 money stock | | |
| M46 | IPMAT | Industrial Production: Materials | | |
| M47 | IPB51222S | Industrial Production: Non-Durable Consumer Energy Products: Residential Utilities | | |
| M48 | MANEMP | All Employees, Manufacturing | | |
| M49 | NDMANEMP | All Employees, Nondurable Goods | | |
| M50 | PAYEMS | All Employees, Total Nonfarm | | |
| M51 | SRVPRD | All Employees, Service-Providing | | |
| M52 | USGOVT | All Employees, Government | | |
| M53 | IPMANSICS | Industrial Production: Manufacturing (SIC) | | |
| M55 | IPDCONGD | Industrial Production: Durable Consumer Goods | | |
| M56 | IPCONGD | Industrial Production: Consumer Goods | | |
| M57 | CES1021000001 | All Employees, Mining | | |
| M58 | USWTRADE | All Employees, Wholesale Trade | | |
| M54 | USTRADE | All Employees, Retail Trade | | |
| M70 | T10YFFM | 10-Year Treasury Constant Maturity Minus Federal Funds Rate | | |
| M71 | TB3SMFFM | 3-Month Treasury Bill Minus Federal Funds Rate | | |
| M72 | TB6SMFFM | 6-Month Treasury Bill Minus Federal Funds Rate | | |
| M73 | T1YFFM | 1-Year Treasury Constant Maturity Minus Federal Funds Rate | | |
| M74 | BAAFFM | Moody's Seasoned Baa Corporate Bond Minus Federal Funds Rate | | |
| M75 | AAAFFM | Moody's Seasoned Aaa Corporate Bond Minus Federal Funds Rate | | |
| M76 | T5YFFM | 5-Year Treasury Constant Maturity Minus Federal Funds Rate | | |

The relative RMSE values corresponding to the forecasts obtained by different competing models using Data 3 (refer to Table 4 of the main paper for results using Data 1 and 2) are provided in Table 17.

## S7.1 Diebold-Mariano-West (DMW) test in forecasting GDP

We have performed the Diebold-Mariano-West (DMW) test (Diebold and Mariano (1995)) for differences in predictive accuracy of the proposed BMF model with respect to MFBVAR model with different choices of priors for

Table 15: Series transformation

| Transformation Code | Series Name |
|---|---|
| 1 | T10YFFM,TB3SMFFM,TB6SMFFM,T1YFFM,BAAFFM,AAAFFM,T5YFFM |
| 2 | FEDFUNDS,TB3MS,GS1,GS5,GS10,TB6MS,UNRATE,AAA,BAA,UEMPMEAN,AWOTMAN |
| 3 | GDPC1, PCECC96, GPDIC1, GCEC1, RCPHBS, OPHPBS, TABSHNO, TLBSNNCB, INDPRO, IPDMAT, RPI, IPBUSEQ, W875RX1, DPCERA3M086SBEA, USTPU, CE16OV, CLF16OV, USCONS, USGOOD, M2REAL, DMANEMP, IPMAT, IPB51222S, MANEMP, NDMANEMP, PAYEMS, SRVPRD, USGOVT, IPMANSICS, USTRADE, IPDCONGD, IPCONGD, CES1021000001, USWTRADE |
| 4 | CPIAUCSL, CPIULFSL, CPIAPPSL, PCEPI, CPITRNSL, DSERRG3M086SBEA, MZMSL, TOTRESNS, BUSLOANS, NONREVSL, PPICMM, DTCTHFNM, CES3000000008, CES0600000008, CUSR0000SA0L2, DTCOLNVHFNM, DNDGRG3M086SBEA, CPIMEDSL, DDURRG3M086SBEA, M2SL, INVEST, CES2000000008, CUSR0000SAS, CUSR0000SA0L5, CUSR0000SAC, M1SL, IPDBS |

Table 16: Edge density in different blocks of $\mathbf{A}$

| Block | No. of edges |
|---|---|
| $\mathbf{A}_{11}$ | 186 |
| $\mathbf{A}_{12}$ | 9 |
| $\mathbf{A}_{21}$ | 13 |
| $\mathbf{A}_{22}$ | 4 |

Table 17: Relative RMSE values for overall comparison of nowcasting/forecasting performance of competing models

| Data 3 | BMF | MFBVAR(Minn) | MFBVAR(SS) | MFBVAR(Hier. SS) | U-MIDAS | MIDAS(Res.) | QVAR(Equal wt.) | QVAR(Skewed) |
|---|---|---|---|---|---|---|---|---|
| $h = 1/3$ | 0.63 | NA | NA | NA | NA | NA | 1.04 | 1.00 |
| $h = 2/3$ | 0.65 | NA | NA | NA | NA | NA | 1.05 | 1.01 |
| $h = 1$ | 0.69 | NA | NA | NA | NA | NA | 1.11 | 1.10 |
| $h = 4/3$ | 0.73 | NA | NA | NA | NA | NA | 1.08 | 1.03 |
| $h = 5/3$ | 0.76 | NA | NA | NA | NA | NA | 0.91 | 0.88 |
| $h = 2$ | 0.81 | NA | NA | NA | NA | NA | 1.12 | 1.30 |

forecasting GDP using Data 1 and Data 2. The results are given in Table 18.

## S7.2 Forecasting using the initial vintage of variables

Note that in the forecasting analysis described in Section 7 of the main paper, we have used the most recently updated values (available as of September 2021 in FRED database) of all the variables for all the time points. Instead of using the current vintage values of the variables, we perform an additional forecasting exercise using the first published estimates of the variables (to respect the release calendar as much as possible). To perform this analysis here we construct the data sets using initial estimates available for all the variables corresponding to each time point. The forecasting performance of BMF using the first vintage values of all variables, is depicted in Table 19 in terms of RMSE, CRPS and LPS values. Note that the forecasts are still compared to the corresponding final vintage values

Table 18: p-values corresponding to the Diebold-Mariano-West test for equal predictive accuracy with respect to the proposed BMF model

| Data 1 | $h = 1/3$ | $h = 2/3$ | $h = 1$ | $h = 4/3$ | $h = 5/3$ | $h = 2$ |
|---|---|---|---|---|---|---|
| MFBVAR(Minn.) | 0.01 | 0.01 | 0.03 | 0.11 | 0.00 | 0.38 |
| MFBVAR(SS) | 0.00 | 0.00 | 0.01 | 0.00 | 0.00 | 0.03 |
| MFBVAR(Hier. SS) | 0.00 | 0.00 | 0.01 | 0.00 | 0.00 | 0.01 |
| Data 2 | $h = 1/3$ | $h = 2/3$ | $h = 1$ | $h = 4/3$ | $h = 5/3$ | $h = 2$ |
| MFBVAR(Minn.) | 0.00 | 0.00 | 0.02 | 0.52 | 0.52 | 0.59 |
| MFBVAR(SS) | 0.00 | 0.00 | 0.01 | 0.46 | 0.42 | 0.52 |
| MFBVAR(Hier. SS) | 0.00 | 0.00 | 0.00 | 0.28 | 0.26 | 0.38 |

of the quarterly variables. Comparing the performance of BMF with the results in Table 3 of the main paper, it can be clearly seen that the forecasting performance of BMF deteriorates when we construct the training data set using the initial releases of the variables for each time point and use it for forecasting. This is expected, as the initial estimates are generally not as accurate, and we found a similar pattern for other methods such as MFBVAR. We also compute the RMSE values for the vector of quarterly forecasts relative to the benchmark model for the proposed BMF as well as MFBVAR, MIDAS and Quarterly VAR model across the forecasting horizon and report them in Table 20. As MFBVAR and MIDAS fails to run for Data 3, we skip the results for Data 3 in Table 20. The results show that for the data set constructed with first releases of variables, our model still performs significantly better than all its competitors. We also provide the CRPS and LPS values of BMF and MFBVAR in Table 21 which shows that BMF still performs significantly better compared to MFBVAR in terms of evaluation of the proposed framework for its posterior predictive distribution.

Table 19: RMSE, CRPS and Log score values for comparing the nowcasting performance of BMF with different no. of monthly variables, using initial releases of all the variables

| | Data 1 | | | Data 2 | | | Data 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| | RMSE | CRPS | log score | RMSE | CRPS | log score | RMSE | CRPS | log score |
| $h = 1/3$ | 1.79 | 1.10 | 2.20 | 1.66 | 1.10 | 2.18 | 1.67 | 1.09 | 2.18 |
| $h = 2/3$ | 1.90 | 1.14 | 2.22 | 1.71 | 1.12 | 2.24 | 1.65 | 1.10 | 2.21 |
| $h = 1$ | 1.98 | 1.16 | 2.23 | 2.83 | 1.27 | 2.27 | 2.97 | 1.26 | 2.26 |
| $h = 4/3$ | 1.68 | 1.11 | 2.16 | 1.65 | 1.11 | 2.17 | 1.66 | 1.11 | 2.18 |
| $h = 5/3$ | 1.74 | 1.12 | 2.17 | 1.69 | 1.12 | 2.19 | 1.66 | 1.11 | 2.17 |
| $h = 2$ | 1.71 | 1.12 | 2.17 | 1.80 | 1.15 | 2.17 | 1.90 | 1.15 | 2.17 |

Table 20: Relative RMSE values for overall comparison of nowcasting/forecasting performance of competing models, using initial releases of all the variables

| Data 1 | BMF | MFBVAR(Minn) | MFBVAR(SS) | MFBVAR(Hier. SS) | U-MIDAS | MIDAS(Res.) | QVAR(Equal wt.) | QVAR(Skewed) |
|---|---|---|---|---|---|---|---|---|
| $h = 1/3$ | 0.50 | 1.41 | 1.42 | 1.46 | 1.85 | 1.54 | 1.19 | 1.33 |
| $h = 2/3$ | 0.53 | 1.62 | 1.52 | 1.47 | 2.21 | 1.51 | 1.10 | 1.28 |
| $h = 1$ | 0.55 | 1.37 | 1.41 | 1.49 | 2.74 | 2.05 | 1.91 | 1.60 |
| $h = 4/3$ | 0.45 | 0.85 | 0.88 | 0.86 | 1.86 | 1.41 | 0.64 | 0.59 |
| $h = 5/3$ | 0.47 | 1.14 | 1.12 | 1.26 | 2.11 | 1.43 | 0.55 | 0.63 |
| $h = 2$ | 0.46 | 0.84 | 0.99 | 1.14 | 2.68 | 1.80 | 0.91 | 0.98 |
| Data 2 | | | | | | | | |
| $h = 1/3$ | 0.47 | 2.55 | 1.99 | 2.24 | 3.84 | 1.91 | 1.80 | 1.41 |
| $h = 2/3$ | 0.48 | 1.77 | 1.51 | 1.38 | 4.81 | 2.07 | 1.46 | 1.63 |
| $h = 1$ | 0.79 | 1.84 | 1.66 | 1.61 | 6.95 | 2.20 | 2.09 | 1.60 |
| $h = 4/3$ | 0.44 | 1.28 | 1.03 | 1.31 | 3.69 | 1.90 | 1.09 | 1.10 |
| $h = 5/3$ | 0.45 | 1.33 | 1.16 | 1.16 | 4.90 | 1.90 | 1.15 | 1.03 |
| $h = 2$ | 0.48 | 0.91 | 0.84 | 0.91 | 7.04 | 2.36 | 1.96 | 1.23 |

Table 21: Nowcasting/forecasting performance of BMF and MFBVAR using CRPS and LPS values, using initial releases of all the variables

| CRPS | Data 1 | | | | Data 2 | | | |
|---|---|---|---|---|---|---|---|---|
| | BMF | MFBVAR(Minn) | MFBVAR(SS) | MFBVAR(Hier. SS) | BMF | MFBVAR(Minn) | MFBVAR(SS) | MFBVAR(Hier. SS) |
| $h = 1/3$ | 1.10 | 2.34 | 2.38 | 2.46 | 1.10 | 3.04 | 2.72 | 2.94 |
| $h = 2/3$ | 1.14 | 2.45 | 2.39 | 2.42 | 1.12 | 2.54 | 2.40 | 2.27 |
| $h = 1$ | 1.16 | 2.31 | 2.33 | 2.47 | 1.27 | 2.40 | 2.36 | 2.28 |
| $h = 4/3$ | 1.11 | 2.11 | 2.09 | 2.18 | 1.11 | 2.10 | 2.08 | 2.18 |
| $h = 5/3$ | 1.12 | 2.20 | 2.19 | 2.36 | 1.12 | 2.13 | 2.12 | 2.10 |
| $h = 2$ | 1.12 | 2.06 | 2.16 | 2.28 | 1.15 | 1.91 | 1.93 | 1.94 |
| Log score | | | | | | | | |
| $h = 1/3$ | 2.20 | 2.66 | 2.67 | 2.69 | 2.18 | 2.70 | 2.69 | 2.70 |
| $h = 2/3$ | 2.22 | 2.68 | 2.69 | 2.70 | 2.24 | 2.67 | 2.68 | 2.65 |
| $h = 1$ | 2.23 | 2.69 | 2.70 | 2.74 | 2.27 | 2.62 | 2.65 | 2.62 |
| $h = 4/3$ | 2.16 | 2.70 | 2.71 | 2.73 | 2.17 | 2.60 | 2.64 | 2.63 |
| $h = 5/3$ | 2.17 | 2.69 | 2.70 | 2.73 | 2.19 | 2.60 | 2.64 | 2.62 |
| $h = 2$ | 2.17 | 2.69 | 2.71 | 2.72 | 2.17 | 2.59 | 2.63 | 2.61 |

# References

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

Diebold, F. X. and R. S. Mariano (1995). Com paring predictive accu racy. *Journal of Business and Economic Statistics 13*(3), 253–263.

Ghosh, S., K. Khare, and G. Michailidis (2018). High-dimensional posterior consistency in bayesian vector autoregressive models. *Journal of the American Statistical Association*.

Ghosh, S., K. Khare, and G. Michailidis (2021). Strong selection consistency of bayesian vector autoregressive models based on a pseudo-likelihood approach. *Annals of Statistics*.

Ghysels, E. (2016). Macroeconomics and the reality of mixed frequency data. *Journal of Econometrics 193*(2), 294–314.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis.* Springer.

McCracken, M. W., M. Owyang, and T. Sekhposyan (2015). Real-time forecasting with a large, mixed frequency, bayesian var. *FRB St. Louis Working Paper* (2015-30).

Narisetty, N. and X. He (2014). Bayesian variable selection with shrinking and diffusing priors. *Annals of Statistics 42*, 789–817.

Rudelson, M., R. Vershynin, et al. (2013). Hanson-wright inequality and sub-gaussian concentration. *Electronic Communications in Probability 18*.

Schorfheide, F. and D. Song (2015). Real-time forecasting with a mixed-frequency var. *Journal of Business & Economic Statistics 33*(3), 366–380.

Vershynin, R. (2010). Introduction to the non-asymptotic analysis of random matrices. *arXiv preprint arXiv:1011.3027*.

Xiang, R., K. Khare, M. Ghosh, et al. (2015). High dimensional posterior convergence rates for decomposable graphical models. *Electronic Journal of Statistics 9*(2), 2828–2854.