Measuring Time-Varying Connectedness: A Bayesian Approach

Author: Sebestyén Kártyás
Consultant: Milán Csaba Badics

May 7, 2024

Motivations

Paradigm shift in financial modeling

- Lessons from the GFC: **interconnectedness** of assets and emergence of abrupt **structural breaks**
- Growing **importance of networks during turbulent periods** as risks and shocks spread across financial infrastructure
- New evidences: Covid-19 Pandemic *(Lenza and Primiceri, 2020, Carriero et al., 2021)*, Russo-Ukrainian war (*Umar et al., 2022, Karkowska and Urjasz, 2023*)

Need for new approaches

- No existing framework to capture structural breaks when measuring connectedness among participants of the network
- **I propose a Bayesian approach to estimate connectedness in the presence of structural breaks**

Preceding Studies

Capturing systemic risks

- Conditional Value at Risk *(Adrian and Brunnermeier, 2011)*
- Systemic and Marginal Expected Shortfall *(Acharya et al., 2012)*
- Granger-causality-based networks *(Billio et al., 2011)*
- **Forecast Error Variance Decomposition based measures** *(Diebold and Yilmaz, 2009, 2012, 2014)*

Dynamic analysis of financial networks

(i) Rolling-window VAR models *(Diebold and Yilmaz, 2012)*

(ii) TVP-VAR with approximate Kalman-filtering method and forgetting factors *(Koop and Korobilis, 2013, Antonakakis and Gabauer, 2017)*

(iii) Large Bayesian TVP-VAR model *(Korobilis and Yilmaz, 2012)*

Critics of existing methods

(i) slow adjustment to changes, arbitrary window size, loss of observation

(ii) overestimates real time-varying tendencies

(iii) can not account for larger, abrupt breaks in the system

This Study

Value added

- TVP-VAR-SV model with hierarchical prior specifications *(Prüser, 2021)* mixed with the DY-framework
- Assessing estimation efficiency of the two priors through a three-step Monte Carlo simulation

Monte Carlo simulations

1. **Univariate unobserved component model** to analyze the properties of the priors on a broader scale than *Prüser (2021)* has done it
2. **Bivariate TVP-VAR-SV(1)** with three different structural break regimes
3. Applying the **DY-framework** on the time series generated in Step 2

Results

- The accuracy of the two prior specifications heavily dependent on variance levels and break-profile of the system
- The accuracy patterns descend when we estimate connectedness
- The horseshoe prior is more accurate around the structural break

Methodology

Time-Varying Parameter Vector Autoregression with Stochastic Volatility (TVP-VAR-SV) following Eisenstat et al. (2015)

Structural form:

$$\mathbf{B}_{0t}\mathbf{y}_t = \boldsymbol{\mu}_t + \mathbf{B}_{1t}\mathbf{y}_{t-1} + \cdots + \mathbf{B}_{pt}\mathbf{y}_{t-p} + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_t)$$

Reduced form:

$$\mathbf{y}_t = \tilde{\mathbf{X}}_t\boldsymbol{\beta}_t + \mathbf{W}_t\gamma_t + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_t)$$

State-space representation:

$$\mathbf{y}_t = \mathbf{X}_t\boldsymbol{\theta}_t + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_t)$$

Time varying parameters:

$$\theta_{j,t} = \theta_{j,t-1} + \eta_{j,t}, \quad \eta_{j,t} \sim \mathcal{N}(0, V_{\theta_{j,t}})$$

Stochastic volatility:

$$\boldsymbol{\Sigma}_t = diag(exp(h_{1,t}), \ldots, exp(h_{n,t})) \text{ and}$$
$$h_{i,t} = h_{i,t-1} + \xi_{i,t}, \quad \xi_{i,t} \sim \mathcal{N}(0, V_{h_{i,t}})$$

Bayesian Framework II - Prior Specifications

Chan and Eisenstat (2018), Prüser (2021)

Inverse Gamma Prior

- Assume the prior variances of $\boldsymbol{\theta}_t$ and to be constant
- $V_{\theta_j} \sim IG(\nu_{\theta_j}, S_{\theta_j}), \quad V_{h_j} \sim IG(\nu_{h_j}, S_{h_j})$
- Shape parameters: $\nu_{\theta_j} = \nu_{h_j} = 5$
- Scale parameters S_{θ_j} and S_{h_j} are set to produce means of 0.01^2 and 0.1^2
- Favors many smaller, gradual changes

Horseshoe Prior

- Global-local shrinkage prior, often used with sparse settings (NOT HERE!)
- $V_{x_{j,t}} = \tau_{x_j} \lambda_{x_{j,t}}, \quad \sqrt{\tau_{x_j}} \sim \mathcal{C}^+(0,1), \quad \sqrt{\lambda_{x_{j,t}}} \sim \mathcal{C}^+(0,1), \quad x \in \{\theta, h\}$
- τ global shrinkage parameters, λ local shrinkage parameters
- Allows for many smaller gradual changes, few larger abrupt breaks or the mix of these

Diebold-Yilmaz Framework

Advantages

- **No need for special conditions** apart from the ones necessary for the identification of the VAR model
- **Easy interpretation** of the results: weighted and directed graphs
- Suitable for both **macro- and micro-level analysis**
- Generalization of Granger-causality-based networks

Concept

- Base: **VAR model** in its Wold-representation form
- **Forecast Error Variance Decomposition** (FEVD) from non-linear transformation of the **coefficient** and **covariance matrices**
- The transmission of shocks among time series can be analyzed by this decomposition
- Summarize the FEVDs for all time series in the **DY-connectedness table**
- Aggregation by row: shocks received from the system, aggregation by column: shocks emitted to the system

Simulation Study

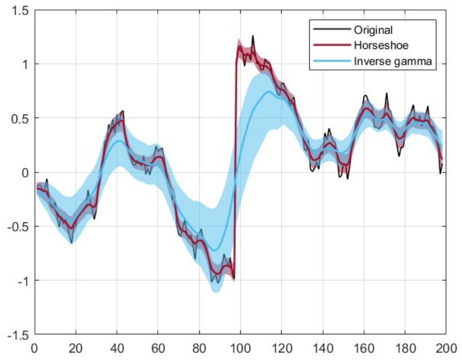
Univariate Example I - Break in Unobserved Component

Univariate Unobserved Component Model with Stoch Vol (Prüser, 2021)

$$y_t = \tau_t + \epsilon_t^y, \quad \epsilon_t^y \sim \mathcal{N}(0, exp(h_t)) \tag{1}$$

$$\tau_t = \tau_{t-1} + c\mathbf{1}(t = t_b) + \epsilon_t^\tau, \quad \epsilon_t^\tau \sim \mathcal{N}(0, \sigma_\tau^2), \quad c = 2, \quad \sigma_\tau^2 = 0.1^2 \tag{2}$$

$$h_t = h_{t-1} + \epsilon_t^h, \quad \epsilon_t^h \sim \mathcal{N}(0, \sigma_h^2), \quad \sigma_h^2 = 0.1^2 \tag{3}$$



$$MAE_\tau = \frac{1}{T \times S} \sum_{t=1}^{T} \sum_{s=1}^{S} |\tau_t - \hat{\tau}_{s,t}|$$

Table 1: Simulation results: Univariate example, abrupt break in a coefficient

	horseshoe	inverse gamma
c = 0	0.07	0.07
c = 2	**0.05**	**0.11**

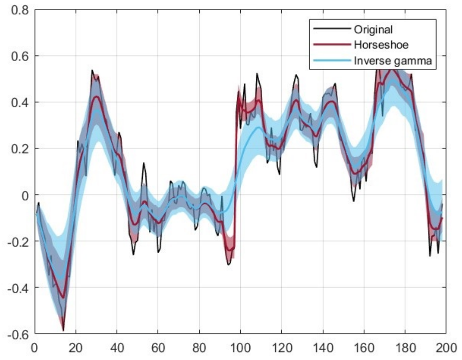
Univariate Example II - Break in Stochastic Volatility

Univariate Unobserved Component Model with Stoch Vol

$$y_t = \tau_t + \epsilon_t^y, \quad \epsilon_t^y \sim \mathcal{N}(0, exp(h_t)) \tag{4}$$

$$\tau_t = \tau_{t-1} + \epsilon_t^\tau, \quad \epsilon_t^\tau \sim \mathcal{N}(0, \sigma_\tau^2), \quad \sigma_\tau^2 = 0.1^2 \tag{5}$$

$$h_t = h_{t-1} + c_h \mathbf{1}(t = t_b) + \epsilon_t^h, \quad \epsilon_t^h \sim \mathcal{N}(0, \sigma_h^2), \quad c_h = 0.5, \quad \sigma_h^2 = 0.1^2 \tag{6}$$



$$MAE_h = \frac{1}{T \times S} \sum_{t=1}^{T} \sum_{s=1}^{S} |h_t - \hat{h}_{s,t}|$$

Table 2: Simulation results: Univariate example, abrupt break in a coefficient

	horseshoe	inverse gamma
c = 0	0.07	0.08
c = 2	**0.06**	**0.09**

Bivariate TVP-VAR-SV I - Break in One Coefficient

Model Equations (Antonakakis et al., 2020, Prüser, 2021)

$$\mathbf{y}_t = \boldsymbol{B}_t \mathbf{y}_{t-1} + \boldsymbol{\epsilon}_t^y, \quad \boldsymbol{\epsilon}_t^y \sim \mathcal{N}(\mathbf{0}, diag(\boldsymbol{h}_t)) \tag{7}$$

$$\boldsymbol{B}_t = \boldsymbol{B}_{t-1} + \boldsymbol{C}_{b,t}\mathbf{1}(t = t_b) + \boldsymbol{\epsilon}_t^b, \quad \epsilon_{jk,t}^b \sim \mathcal{N}(0, \sigma_b^2) \tag{8}$$

$$\boldsymbol{h}_t = \boldsymbol{h}_{t-1} + \boldsymbol{\epsilon}_t^h, \quad \boldsymbol{\epsilon}_t^h \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}^h) \tag{9}$$

$$\boldsymbol{B}_0 = \begin{bmatrix} 0.6 & 0.3 \\ 0.3 & 0.6 \end{bmatrix}, \quad \boldsymbol{C}_{b,t} = \begin{bmatrix} 0.2 & 0 \\ 0 & 0 \end{bmatrix}, \quad \boldsymbol{\Sigma}^h = \begin{bmatrix} 0.1^2 & 0 \\ 0 & 0.1^2 \end{bmatrix}$$

Mean Absolute Error

$$MAE_b = \frac{1}{T \times S \times 4} \sum_{s=1}^{S} \sum_{t=1}^{T} \sum_{k=1}^{K} \sum_{j=1}^{J} |b_{jk,t,s} - \hat{b}_{jk,t,s}|$$

Table 3: Simulation results: VAR, abrupt break in one coefficient

	c = 0.2				c = 0			
σ_b^2	0.02	0.01	0.001	0.0001	0.02	0.01	0.001	0.0001
horseshoe	**0.0100**	**0.0061**	0.0019	0.0008	**0.0098**	**0.0061**	0.0020	0.0009
inverse gamma	0.0123	0.0065	**0.0010**	**0.0003**	0.0118	0.0063	**0.0009**	**0.0002**

- Results differ from those of *Prüser (2021)*: the superiority of the horseshoe prior is not that straightforward
- With relatively higher parameter variance, the horseshoe prior is more accurate, the inverse gamma prior is better with lower variance levels
- With a given level of variance, the horseshoe prior is better in the presence of breaks and the inverse gamma prior is better in their absence

Model Equations (Antonakakis et al., 2020, Prüser, 2021)

$$\mathbf{y}_t = \boldsymbol{B}_t \mathbf{y}_{t-1} + \boldsymbol{\epsilon}_t^y, \quad \boldsymbol{\epsilon}_t^y \sim \mathcal{N}(\mathbf{0}, diag(exp(\boldsymbol{h}_t))) \tag{10}$$

$$\boldsymbol{B}_t = \boldsymbol{B}_{t-1} + \boldsymbol{C}_{b,t}\mathbf{1}(t = t_b) + \boldsymbol{\epsilon}_t^b, \quad \epsilon_{jk,t}^b \sim \mathcal{N}(0, \sigma_b^2) \tag{11}$$

$$\boldsymbol{h}_t = \boldsymbol{h}_{t-1} + \boldsymbol{\epsilon}_t^h, \quad \boldsymbol{\epsilon}_t^h \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}^h) \tag{12}$$

$$\boldsymbol{B}_0 = \begin{bmatrix} 0.6 & 0.3 \\ 0.3 & 0.6 \end{bmatrix}, \quad \boldsymbol{C}_{b,t} = \begin{bmatrix} 0.2 & -0.2 \\ -0.2 & 0.2 \end{bmatrix}, \quad \boldsymbol{\Sigma}^h = \begin{bmatrix} 0.1^2 & 0 \\ 0 & 0.1^2 \end{bmatrix}$$

Mean Absolute Error

$$MAE_b = \frac{1}{T \times S \times 4} \sum_{s=1}^{S} \sum_{t=1}^{T} \sum_{k=1}^{K} \sum_{j=1}^{J} |b_{jk,t,s} - \hat{b}_{jk,t,s}|$$

Bivariate TVP-VAR-SV II - Results

Table 4: Simulation results: VAR, abrupt break in all coefficients

	c = 0.2			
σ_b^2	0.02	0.01	0.001	0.0001
horseshoe	**0.0091**	**0.0058**	0.0017	0.0007
inverse gamma	0.0120	0.0071	**0.0012**	**0.0005**

- Similar patterns to the previous results
- The horse prior is superior when the parameter variance is higher, the inverse gamma prior is superior with lower variance levels
- With higher variance settings, the advantage of the horseshoe prior is larger
- With lower variance settings, the advantage of the inverse gamma prior is smaller
- The accuracy of both priors are inversely proportional to the variance levels

DY-Spillover Table Estimation

Requisite equations

- Bivariate TVP-VAR-SV II from Frame 11: $\mathbf{y}_t = \mathbf{B}_{1,t}\mathbf{y}_{t-1} + \boldsymbol{\epsilon}_t^y$, $\mathcal{N} \sim (\mathbf{0}, \mathbf{\Sigma}_t)$, $\mathbf{\Sigma}_t = diag(exp(h_{1,t}), exp(h_{2,t}))$
- TVP-VMA representation: $\mathbf{y}_t = \sum_{j=0}^{\infty} \mathbf{A}_{j,t}\boldsymbol{u}_{t-j}$
- Spillover from j to i:

$$\phi_{ij,t}(H) = \frac{\Sigma_{jj,t}^{-1} \sum_{h=0}^{H-1} (\mathbf{e}_i'\mathbf{A}_{h,t}\mathbf{\Sigma}_t\mathbf{e}_j)^2}{\sum_{h=0}^{H-1} \mathbf{e}_i'\mathbf{A}_{h,t}\mathbf{\Sigma}_t\mathbf{A}_{h,t}'\mathbf{e}_j}$$

Mean Absolute Error

$$MAE_{\tilde{\phi}} = \frac{1}{S \times T \times 4} \sum_{s=1}^{S} \sum_{t=1}^{T} \sum_{k=1}^{K} \sum_{j=1}^{J} |\tilde{\phi}_{jk,t,s} - \hat{\tilde{\phi}}_{jk,t,s}|$$

DY-Spillover Table Estimation - Results

Complete time frame

Table 5: Simulation results: VAR, abrupt break in all coefficients, Diebold-Yilmaz spillover table

	c = 0.2			
σ_b^2	0.02	0.01	0.001	0.0001
horseshoe	**2.13**	**1.22**	0.99	0.84
inverse gamma	2.73	1.42	**0.86**	**0.74**

- The horseshoe prior estimates the network better when the variance is higher
- The inverse gamma is the better when the parameters are less volatile
- The accuracy of both priors is inversely proportional to the variance level of the coefficients

Around the structural break

Table 6: Simulation results: VAR, abrupt break in all coefficients, Diebold-Yilmaz spillover table between t=95 and t=105

	c = 0.2			
σ_b^2	0.02	0.01	0.001	0.0001
horseshoe	**1.75**	**1.77**	**1.39**	**0.53**
inverse gamma	1.87	2.23	3.16	2.23

- The horseshoe prior is better with every setting
- The clear inverse relationship between accuracy and variance levels weakens, and it completely disappears with the inverse gamma prior

Wrap Up

Findings

Behaviour of the priors

- With a given state-switching regime, the **horseshoe prior** performs better when **variance of the coefficients is higher**, and also more effective **in the presence of structural breaks**
- The accuracy of both priors is inversely proportional to the variance level

Estimating the DY-spillover table

With wider time window

- The **horseshoe prior** is superior with **larger variance** levels
- The **inverse gamma prior** is superior with **smaller variance** levels
- The accuracy of both priors is inversely proportional to the variance levels

With shorter time window around the structural break

- The **horseshoe prior** is superior **regardless variance levels**
- The clear **inverse relationship** between accuracy and variance levels **weakens**, and it completely disappears with the inverse gamma prior

Questions

Diebold-Yilmaz Framework II - Equations

- Base model (TVP-VAR-SV):

$$\mathbf{B}_{0t}\mathbf{y}_t = \boldsymbol{\mu}_t + \mathbf{B}_{1t}\mathbf{y}_{t-1} + \cdots + \mathbf{B}_{pt}\mathbf{y}_{t-p} + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_t)$$

- Compact form:

$$\mathbf{y}_t = \tilde{\boldsymbol{\mu}}_t + \widetilde{\mathbf{B}}_t\mathbf{z}_{t-1} + \boldsymbol{u}_t, \quad \boldsymbol{u}_t \sim \mathcal{N}(\mathbf{0}, \widetilde{\boldsymbol{\Sigma}}_t)$$

- Transforming into its TVP-VMA representation with recursive substitution:

$$\mathbf{y}_t = \tilde{\boldsymbol{\mu}}_t + \sum_{j=0}^{\infty} \mathbf{A}_{j,t}\boldsymbol{u}_{t-j}$$

- Generalized impulse response function of a shock in the jth variable:

$$\boldsymbol{\Psi}_{j,t}(H) = \widetilde{\Sigma}_{jj,t}^{-\frac{1}{2}}\mathbf{A}_{H,t}\widetilde{\boldsymbol{\Sigma}}_t\mathbf{e}_j$$

- H-step-ahead forecast error variance:

$$\mathbf{FEV}_t(H) = \sum_{h=0}^{H-1}(\mathbf{A}_{h,t}\widetilde{\boldsymbol{\Sigma}}_t\mathbf{A}'_{h,t})$$

- The proportion of the H-step-ahead forecast error variance of i which is due to innovations in j:

$$\phi_{ij,t}(H) = \frac{\sum_{h=0}^{H-1}\Psi_{ij,t}^2(h)}{\sum_{h=0}^{H-1}FEV_{ij,t}(h)}$$

Bivariate TVP-VAR-SV - Equations in coordinate form

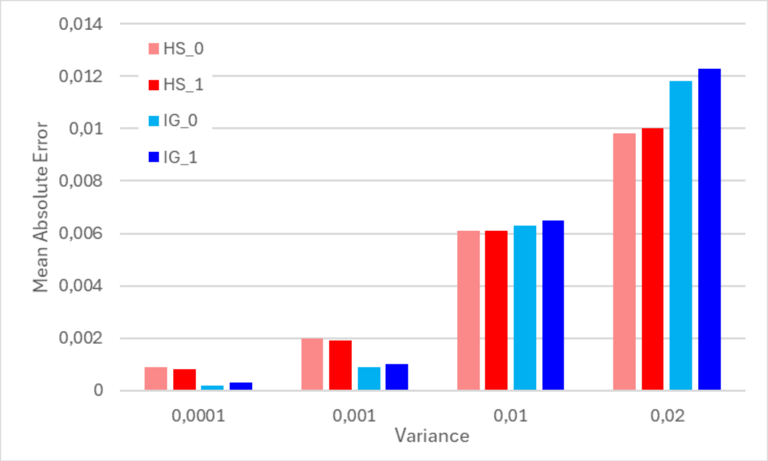
$$y_{1,t} = y_{1,t-1}b_{11,t} + y_{2,t-1}b_{12,t} + \epsilon^y_{1,t}, \quad \epsilon^y_{1,t} \sim \mathcal{N}(0, exp(h_{1,t})) \tag{13}$$

$$y_{2,t} = y_{1,t-1}b_{21,t} + y_{2,t-1}b_{22,t} + \epsilon^y_{2,t}, \quad \epsilon^y_{2,t} \sim \mathcal{N}(0, exp(h_{2,t})) \tag{14}$$

$$b_{jk,t} = b_{jk,t-1} + c_{b,jk}\mathbf{1}(t = t_b) + \epsilon^b_{jk,t}, \quad \epsilon^b_{jk,t} \sim \mathcal{N}(0, \sigma^2_b) \tag{15}$$

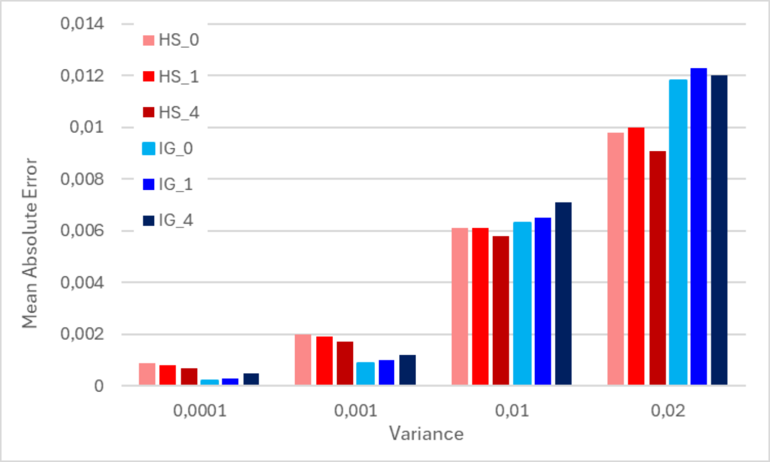
$$h_{i,t} = h_{i,t-1} + \epsilon^h_{i,t}, \quad \epsilon^h_{i,t} \sim \mathcal{N}(0, \sigma^2_h), \quad \sigma^2_h = 0.1^2 \tag{16}$$

Figure 1: Mean Absolute Error of the two prior specifications without break and with break in one coefficient



Bivariate TVP-VAR-SV II - Visualized Results

Figure 2: Mean Absolute Error of the two prior specifications without break, with break in one coefficient, and with break in all coefficients



DY-Spillover Table Estimation - Visualized Results

Figure 3: Mean Absolute Error of the two prior specifications in the estimation of the DY-spillover table in complete time frame

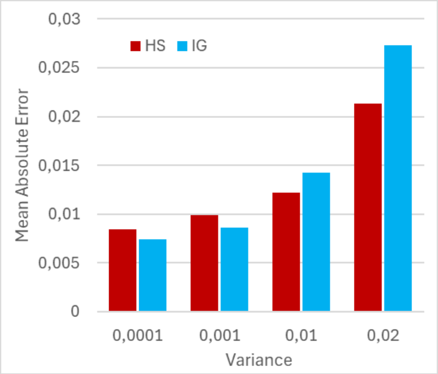
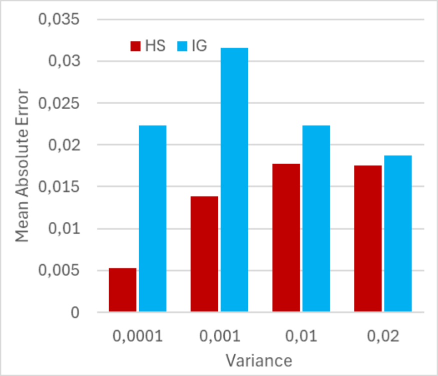


Figure 4: Mean Absolute Error of the two prior specifications in the estimation of the DY-spillover table between t=95 and t=105



Gibbs-sampler I

Following *Chan and Eisenstat (2018)* and *Prüser (2021)*

Let $\boldsymbol{y} = (\boldsymbol{y}_1', \ldots, \boldsymbol{y}_T')'$, $\boldsymbol{\theta} = (\boldsymbol{\theta}_1', \ldots, \boldsymbol{\theta}_T')'$,
$\boldsymbol{\Sigma}_\theta = diag(V_{\theta_{11}}, \ldots, V_{\theta_{1k_\theta}}, \ldots, V_{\theta_{T1}}, \ldots, V_{\theta_{Tk_\theta}})$, and
$\boldsymbol{\Sigma}_h = diag(V_{h_{11}}, \ldots, V_{h_{1n}}, \ldots, V_{h_{T1}}, \ldots, V_{h_{Tn}})$. A sample of the posterior can be obtained by sequentially drawing from the following conditional posterior distributions:

1. $p(\boldsymbol{\theta}|\boldsymbol{y}, \boldsymbol{h}, \boldsymbol{\Sigma}_\theta, \boldsymbol{\Sigma}_h, \boldsymbol{\theta}_0, \boldsymbol{h}_0)$
2. $p(\boldsymbol{h}|\boldsymbol{y}, \boldsymbol{\theta}, \boldsymbol{\Sigma}_\theta, \boldsymbol{\Sigma}_h, \boldsymbol{\theta}_0, \boldsymbol{h}_0)$
3. $p(\boldsymbol{\theta}_0, \boldsymbol{h}_0|\boldsymbol{y}, \boldsymbol{\theta}, \boldsymbol{h}, \boldsymbol{\Sigma}_\theta, \boldsymbol{\Sigma}_h)$
4. $p(\boldsymbol{\Sigma}_\theta, \boldsymbol{\Sigma}_h|\boldsymbol{y}, \boldsymbol{\theta}, \boldsymbol{h}, \boldsymbol{\theta}_0, \boldsymbol{h}_0)$

Step 1

Rewrite the TVP-VAR model as $\boldsymbol{y} = \boldsymbol{X\theta} + \boldsymbol{\epsilon}$, $\boldsymbol{\epsilon} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$, where $\boldsymbol{\epsilon} = (\boldsymbol{\epsilon}_1', \ldots, \boldsymbol{\epsilon}_T')'$, $\boldsymbol{\Sigma} = diag(\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_T)$, and $\boldsymbol{X} = diag(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_T)$. We can also rewrite the hierarchical prior as $\boldsymbol{H}_\theta\boldsymbol{\theta} = \tilde{\boldsymbol{\alpha}}_\theta + \boldsymbol{\eta}$, $\boldsymbol{\eta} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_\theta)$, where

$$\tilde{\boldsymbol{\alpha}}_\theta = (\boldsymbol{\theta}_0', \mathbf{0}, \dots, \mathbf{0})' \text{ and } \boldsymbol{H}_\theta = \begin{pmatrix} \boldsymbol{I}_{k_\theta} & \mathbf{0} & \dots & \mathbf{0} \\ -\boldsymbol{I}_{k_\theta} & \boldsymbol{I}_{k_\theta} & \ddots & \mathbf{0} \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \dots & -\boldsymbol{I}_{k_\theta} & \boldsymbol{I}_{k_\theta} \end{pmatrix}.$$

It follows that $(\boldsymbol{\theta}|\boldsymbol{\Sigma}_\theta, \boldsymbol{\theta}_0) \sim \mathcal{N}(\boldsymbol{\alpha}_\theta, (\boldsymbol{H}_\theta' \boldsymbol{\Sigma}_\theta^{-1} \boldsymbol{H}_\theta)^{-1})$, where $\boldsymbol{\alpha}_\theta = \boldsymbol{H}_\theta^{-1} \tilde{\boldsymbol{\alpha}}_\theta$. It can be shown that

$$(\boldsymbol{\theta}|\boldsymbol{y}, \boldsymbol{h}, \boldsymbol{\Sigma}_\theta, \boldsymbol{\Sigma}_h, \boldsymbol{\theta}_0, \boldsymbol{h}_0) \sim \mathcal{N}(\hat{\boldsymbol{\theta}}, \boldsymbol{K}_\theta^{-1}),$$

where $\hat{\boldsymbol{\theta}} = \boldsymbol{K}_\theta^{-1} \boldsymbol{d}_\theta$ with $\boldsymbol{K}_\theta = \boldsymbol{H}_\theta' \boldsymbol{\Sigma}_\theta^{-1} \boldsymbol{H}_\theta + \boldsymbol{X}' \boldsymbol{\Sigma}^{-1} \boldsymbol{X}$ and $\boldsymbol{d}_\theta = \boldsymbol{H}_\theta' \boldsymbol{\Sigma}_\theta^{-1} \boldsymbol{H}_\theta \boldsymbol{\alpha}_\theta + \boldsymbol{X}' \boldsymbol{\Sigma}^{-1} \boldsymbol{y}$.

Step 2

The auxiliary mixture sampler of *Kim et al. (1998)* in combination with the precision sampler of *Chan and Jeliazkov (2009)* to sequantially draw each slice of $\boldsymbol{h}_i = (h_{i1}, \dots, h_{iT})'$, $i = 1, \dots, n$.

Step 3

$$(\boldsymbol{\theta}_0|\boldsymbol{y}, \boldsymbol{\theta}, \boldsymbol{h}, \boldsymbol{\Sigma}_\theta, \boldsymbol{\Sigma}_h) \sim \mathcal{N}(\hat{\boldsymbol{\theta}}_0, \boldsymbol{K}_{\theta_0}^{-1}), \quad (\boldsymbol{h}_0|\boldsymbol{y}, \boldsymbol{\theta}, \boldsymbol{h}, \boldsymbol{\Sigma}_\theta, \boldsymbol{\Sigma}_h) \sim \mathcal{N}(\hat{\boldsymbol{h}}_0, \boldsymbol{K}_{h_0}^{-1}),$$

where $\boldsymbol{K}_{\theta_0} = diag(V_{\theta_{11}}, \ldots, V_{\theta_{1k_\theta}})^{-1} + \frac{1}{10}\boldsymbol{I}_{k_\theta}$,
$\hat{\boldsymbol{\theta}}_0 = \boldsymbol{K}_{\theta_0}^{-1} diag(V_{\theta_{11}}, \ldots, V_{\theta_{1k_\theta}})^{-1}\boldsymbol{\theta}_1$, $\boldsymbol{K}_{h_0} = diag(V_{h_{11}}, \ldots, V_{h_{1n}})^{-1} + \frac{1}{10}\boldsymbol{I}_n$,
$\hat{\boldsymbol{h}}_0 = \boldsymbol{K}_{h_0}^{-1} diag(V_{h_{11}}, \ldots, V_{h_{1n}})^{-1}\boldsymbol{h}_1$.

Step 4 - Inverse Gamma Prior

$$(V_{\theta_j}|\boldsymbol{y}, \boldsymbol{\theta}, \boldsymbol{h}, \boldsymbol{\theta}_0, \boldsymbol{h}_0) \sim IG\left(\nu_{\theta_j} + \frac{T}{2}, S_{\theta_j} + \frac{1}{2}\sum_{t=1}^{T}(\theta_{j,t} - \theta_{j,t-1})^2\right),\ j = 1, \ldots, k_\theta$$

$$(V_{h_i}|\boldsymbol{y}, \boldsymbol{\theta}, \boldsymbol{h}, \boldsymbol{\theta}_0, \boldsymbol{h}_0) \sim IG\left(\nu_{h_i} + \frac{T}{2}, S_{h_i} + \frac{1}{2}\sum_{t=1}^{T}(h_{i,t} - h_{i,t-1})^2\right),\ i = 1, \ldots, n$$

Step 4 - Horseshoe Prior

Following *Makalic and Schmidt (2016)*, one can use the scalar mixture representation of the half-Cauchy distribution: if X and w are random variables such that $X^2|w \sim IG(\frac{1}{2}, \frac{1}{w})$ and $w \sim IG(\frac{1}{2}, 1)$, then $X \sim C^+(0, 1)$.

Gibbs-sampler IV

$$(\tau_{\theta_j}|\boldsymbol{\theta}, \boldsymbol{\theta}_0, \nu_{\tau_{\theta_j}}, \lambda_{\theta_{t,j}}) \sim IG\left(\frac{T+1}{2}, \frac{1}{\nu_{\tau_{\theta_j}}} + \frac{1}{2}\sum_{t=1}^{T}\frac{(\theta_{j,t} - \theta_{j,t-1})^2}{\lambda_{\theta_{t,j}}}\right),\ j = 1, \ldots, k_\theta$$

$$(\lambda_{\theta_{t,j}}|\boldsymbol{\theta}, \boldsymbol{\theta}_0, \nu_{\tau_{\theta_j}}, \tau_{\theta_j}) \sim IG\left(1, \frac{1}{\nu_{\lambda_{\theta_{t,j}}}} + \frac{1}{2}\sum_{t=1}^{T}\frac{(\theta_{j,t} - \theta_{j,t-1})^2}{\tau_{\theta_j}}\right),\ j = 1, \ldots, k_\theta$$

$$(\nu_{\tau_{\theta_j}}|\tau_{\theta_j}) \sim IG\left(1, 1 + \frac{1}{\tau_{\theta_j}}\right)\ j = 1, \ldots, k_\theta$$

$$(\nu_{\lambda_{\theta_{t,j}}}|\lambda_{\theta_{t,j}}) \sim IG\left(1, 1 + \frac{1}{\lambda_{\theta_{t,j}}}\right)\ j = 1, \ldots, k_\theta$$

Gibbs-sampler V

$$(\tau_{h_j}|\boldsymbol{h},\boldsymbol{h}_0,\nu_{\tau_{h_i}},\lambda_{h_{t,i}}) \sim IG\left(\frac{T+1}{2},\frac{1}{\nu_{\tau_{h_i}}}+\frac{1}{2}\sum_{t=1}^{T}\frac{(h_{i,t}-h_{i,t-1})^2}{\lambda_{h_{t,i}}}\right),\ i=1,\dots,n$$

$$(\lambda_{h_{t,i}}|\boldsymbol{h},\boldsymbol{h}_0,\nu_{\tau_{h_i}},\tau_{h_i}) \sim IG\left(1,\frac{1}{\nu_{\lambda_{h_{t,i}}}}+\frac{1}{2}\sum_{t=1}^{T}\frac{(h_{i,t}-h_{i,t-1})^2}{\tau_{h_i}}\right),\ i=1,\dots,n$$

$$(\nu_{\tau_{h_i}}|\tau_{h_i}) \sim IG\left(1,1+\frac{1}{\tau_{h_i}}\right)\ i=1,\dots,n$$

$$(\nu_{\lambda_{h_{t,i}}}|\lambda_{h_{t,i}}) \sim IG\left(1,1+\frac{1}{\lambda_{h_{t,i}}}\right)\ i=1,\dots,n$$