Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Lag-Adaptive Hierarchical Sparse Bayesian Panel VAR Modeling for High-Dimensional Time Series Systems

Miftahul Irfan

miftahul.irfan@at.itera.ac.id

Sumatera Institute of Technology

Nanda Azzanina

Sumatera Institute of Technology

Erica Grace Simanjuntak

Sumatera Institute of Technology

Wardhani Utami Dewi

Universitas Muhammadiyah Metro

Research Article

Keywords: Bayesian Panel VAR, High-dimensional time series, Sparse Bayesian learning, Horseshoe prior, Lag-adaptive shrinkage, Forecasting

Posted Date: November 2nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10040186/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Lag-Adaptive Hierarchical Sparse Bayesian Panel VAR Modeling for High-Dimensional Time Series Systems

Miftahul Irfan[a,*], Nanda Azzanina[b], Erica Grace Simanjuntak[c], Wardhani Utami Dewi[d]

[a, b, c]Actuarial Science, Institut Teknologi Sumatera, Lampung, Indonesia
[d]Universitas Muhammadiyah Metro, Lampung, Indonesia

Email: miftahul.irfan@at.itera.ac.id

Abstract: High-dimensional panel vector autoregressive (PVAR) models have become increasingly important for analyzing complex dynamic interactions in multivariate time series systems. However, conventional estimation methods often suffer from overparameterization, multicollinearity, and poor forecasting performance, particularly when the number of parameters grows rapidly relative to the sample size. To address these challenges, this study proposes a Lag-Adaptive Hierarchical Sparse Bayesian Panel Vector Autoregressive (LAHS-BPVAR) model that integrates adaptive lag regularization, hierarchical global-local shrinkage, and sparse Bayesian learning within a unified framework. The proposed model introduces a novel prior covariance structure in which coefficient variances decay according to a hybrid lag-adaptive mechanism combining polynomial and exponential shrinkage, allowing automatic penalization of distant lags while preserving relevant short-run dynamics. A hierarchical sparse prior based on global-local shrinkage further enables adaptive variable selection in high-dimensional settings. Theoretical analysis establishes posterior contraction and sparse support recovery consistency under standard regularity conditions, providing formal asymptotic justification for the proposed estimator. Estimation is performed using a Gibbs sampler with Metropolis-Hastings updates for non-conjugate hyperparameters. Monte Carlo simulation results demonstrate competitive estimation and forecasting performance relative to conventional VAR, Minnesota Bayesian VAR, LASSO VAR, and standard Horseshoe Bayesian VAR. The proposed framework consistently achieves improved sparse support recovery through more effective control of false positive coefficient selection while maintaining estimation accuracy. An empirical application to Indonesian macroeconomic time series further demonstrates the practical effectiveness of the proposed framework in capturing dynamic economic relationships while maintaining model parsimony. The findings suggest that the proposed methodology provides a robust and flexible solution for high-dimensional panel time series modeling.

Keywords: Bayesian Panel VAR; High-dimensional time series; Sparse Bayesian learning; Horseshoe prior; Lag-adaptive shrinkage; Forecasting.

1. Introduction

The rapid expansion of high-dimensional multivariate time series data in economics, finance, and policy analysis has significantly increased the demand for statistical models capable of capturing complex dynamic interdependencies while maintaining estimation stability and forecasting reliability. Vector autoregressive (VAR) models remain among the most widely adopted tools for multivariate dynamic analysis due to their flexible structure and interpretability. However, as the dimensionality of the system increases, conventional VAR estimation rapidly becomes infeasible because the number of parameters grows quadratically with the number of endogenous variables and linearly with the lag order. This well-known curse of dimensionality often leads to severe overparameterization, unstable coefficient estimates, multicollinearity, and degraded out-of-sample forecasting performance, especially when the available sample size is limited relative to model complexity. Recent evidence continues to confirm that overfitting remains a central obstacle in large-scale VAR modeling, motivating the continued development of structured regularization methods (Gruber & Kastner, 2025).

Bayesian vector autoregression (BVAR) has emerged as one of the most successful approaches for addressing these challenges by incorporating prior information to regularize parameter estimation. The success of Bayesian shrinkage stems from its ability to stabilize inference while preserving predictive information contained in large dynamic systems. Traditional shrinkage specifications, particularly Minnesota-type priors, have demonstrated strong empirical performance in macroeconomic forecasting by

shrinking autoregressive coefficients toward parsimonious structures. Nevertheless, fixed prior shrinkage schemes often impose rigid assumptions regarding coefficient behavior across variables and lag structures. In modern high-dimensional settings, where dynamic dependence structures may vary substantially across variables and time horizons, such fixed shrinkage mechanisms may become overly restrictive (Chan, 2023; Chan & Eisenstat, 2018). Recent developments in large Bayesian VAR modeling increasingly emphasize adaptive shrinkage frameworks that improve forecasting performance while allowing greater flexibility in parameter regularization (Gefang et al., 2023). Recent developments in high-dimensional econometrics have emphasized sparse regularization and structured estimation strategies to improve statistical efficiency and interpretability in complex systems (Kutateladze & Seregina, 2024).

An important development in modern Bayesian regularization is the emergence of hierarchical shrinkage priors, particularly global-local shrinkage mechanisms. These priors introduce adaptive regularization by combining a global shrinkage parameter controlling overall model complexity with coefficient-specific local shrinkage parameters that selectively preserve strong signals while aggressively suppressing noise (Hamura et al., 2022; Qin & Ghosh, 2024). Among these approaches, the horseshoe prior has attracted considerable attention due to its desirable combination of strong concentration near zero and heavy-tailed behavior, enabling simultaneous sparsity recovery and reduced bias for important coefficients (Carvalho et al., 2010; van der Pas et al., 2014). In the context of Bayesian VARs, horseshoe-based shrinkage has shown competitive forecasting performance relative to conventional shrinkage methods while providing more interpretable sparse dynamic structures. More recent work further suggests that hierarchical shrinkage methods remain effective even in increasingly large and computationally demanding VAR environments (Follett & Yu, 2019; Gruber & Kastner, 2025).

Despite these advances, several unresolved methodological challenges remain. First, most existing Bayesian VAR shrinkage frameworks primarily regulate lag behavior through relatively simple deterministic decay structures, often assuming a fixed polynomial decline in prior variance as lag order increases. While this assumption is computationally convenient, it may inadequately represent heterogeneous temporal dynamics observed in real systems. Economic and financial time series frequently exhibit mixed dependence structures, where some dynamic effects decay gradually while others vanish rapidly after short horizons. A single deterministic lag decay mechanism may therefore be insufficient to capture such complexity. Recent advances in Bayesian time-series modeling have explored graphical structures and dynamic sparsity mechanisms to address parameter proliferation in large VAR systems (Vahap, 2025). However, relatively limited attention has been devoted to adaptive lag-dependent shrinkage within Bayesian panel VAR frameworks.

Second, although sparse hierarchical priors have improved coefficient regularization, most existing frameworks focus primarily on variable-level shrinkage rather than explicitly integrating adaptive lag-dependent shrinkage into the hierarchical Bayesian structure. This distinction is particularly relevant in high-dimensional panel or multivariate dynamic systems, where lag proliferation can dominate total parameter dimensionality. Without an explicit lag-adaptive mechanism, sparse priors may still retain unnecessary long-horizon coefficients, reducing interpretability and computational efficiency (Nagel et al., 2024). Third, while computational advances have substantially improved the feasibility of estimating high-dimensional Bayesian VAR models, theoretical guarantees for structured adaptive shrinkage remain relatively limited. In particular, establishing asymptotic justification for sparse hierarchical Bayesian estimators under dynamic lag-dependent shrinkage remains an open challenge, especially in settings involving high-dimensional autoregressive dependence (Makalic & Schmidt, 2016).

A growing body of recent research has sought to improve the scalability and forecasting performance of Bayesian VAR models through increasingly sophisticated shrinkage structures. One important line of development focuses on hierarchical adaptive priors that combine the predictive robustness of classical Minnesota shrinkage with the flexibility of modern sparse Bayesian regularization. For example, Chan

(2021) develops Minnesota-type adaptive hierarchical priors that integrate lag-specific shrinkage with heavy-tailed adaptive behavior, demonstrating improved forecasting performance relative to both conventional Minnesota priors and several alternative hierarchical specifications. This work is particularly relevant because it highlights a central limitation of many global-local shrinkage priors: while theoretically attractive, they often fail to explicitly encode economically plausible lag structures (Chan, 2021; Khare & Ghosh, 2022; Lee et al., 2016).

Parallel work has emphasized the computational scalability of Bayesian shrinkage in increasingly large VAR systems. Gefang et al. (2023) show that variational Bayesian methods substantially reduce the computational burden associated with large BVAR estimation while preserving strong forecasting performance under hierarchical shrinkage priors. Similarly, recent methodological developments demonstrate that efficient posterior inference in large dynamic systems increasingly depends on exploiting structured shrinkage assumptions rather than relying solely on brute-force computation. These advances confirm that shrinkage design remains central not only to statistical regularization but also to computational feasibility in high-dimensional time series environments (Chan, 2021; Gefang et al., 2023; Prüser, 2023).

Another influential research direction concerns explicitly sparse Bayesian VAR modeling. Follett and Yu (2019) introduce the horseshoe prior within Bayesian VAR frameworks and show that global-local shrinkage can effectively suppress irrelevant coefficients while preserving substantively important signals (Follett & Yu, 2019). Their empirical results indicate that sparse hierarchical priors can compete favorably with traditional shrinkage schemes, particularly when the underlying data-generating process exhibits meaningful sparsity. More broadly, Griffin and Brown (2021) review global-local shrinkage methods and emphasize their theoretical attractiveness for high-dimensional Bayesian inference, especially due to their strong concentration around zero combined with heavy-tailed robustness for significant coefficients (Griffin & Brown, 2021). Recent Bayesian approaches have increasingly focused on identifying dynamic dependence structures and causal interactions within multivariate time-series systems, including flexible Bayesian procedures for Granger causality assessment (Gutiérrez et al., 2024).

Despite these methodological gains, evidence regarding the empirical superiority of sparse hierarchical shrinkage in macroeconomic VAR forecasting remains mixed. Cross et al. (2020), for instance, report that global-local shrinkage priors such as horseshoe and normal-gamma do not consistently outperform hierarchical Minnesota priors in macroeconomic forecasting exercises. This finding suggests an important unresolved issue: sparsity alone may not be sufficient when dynamic dependence structures remain inadequately modeled. In other words, aggressive coefficient shrinkage does not automatically imply optimal temporal regularization (Kohns & Szendrei, 2024). This observation is highly relevant to the present study because it motivates the integration of structured lag-dependent shrinkage rather than relying exclusively on coefficient-level sparsity.

A core weakness of many existing shrinkage frameworks lies in their treatment of lag decay. Classical Minnesota-type priors typically impose deterministic polynomial decay, assigning progressively smaller prior variances to coefficients associated with higher lag orders. While economically intuitive, such formulations implicitly assume a single homogeneous temporal decay mechanism across all variables and dynamic relationships. In practical macroeconomic and financial systems, however, dependence structures often display heterogeneous persistence profiles. Some shocks dissipate almost immediately, whereas others exhibit medium-term propagation effects. Fixed polynomial lag shrinkage may therefore impose inappropriate regularization patterns, particularly in heterogeneous high-dimensional systems.

Recent literature increasingly recognizes the need for more flexible prior structures. Data-driven priors, adaptive hierarchical shrinkage, and structured Bayesian regularization have all been proposed to improve model adaptability in large-scale dynamic systems. However, most existing approaches continue to treat lag penalization as either fixed or only indirectly adaptive through hyperparameter tuning, rather than

explicitly modeling lag-specific decay as an estimable hierarchical structure. This distinction is critical because in large VAR and panel VAR systems, the proliferation of lagged coefficients often represents the dominant source of dimensionality. Recent empirical evidence has demonstrated the importance of modeling complex multivariate dynamic dependence structures in large interconnected systems, particularly in applications involving macroeconomic and energy interactions (Irfan et al., 2026). However, empirical dynamic modeling alone does not fully resolve the methodological challenges posed by high-dimensional parameter proliferation, sparse dependence structures, and adaptive lag regularization, thereby motivating the development of more flexible Bayesian shrinkage frameworks (Irfan et al., 2024).

The panel VAR context introduces additional complexity. Compared with standard multivariate VAR systems, panel VAR models must simultaneously account for dynamic dependence across time and heterogeneity across observational units (Chiu et al., 2017). This substantially enlarges the parameter space and amplifies the risks of overparameterization, weak identification, and unstable inference. While Bayesian shrinkage has been successfully applied to large VAR systems, comparatively fewer studies explicitly address structured sparse shrinkage in high-dimensional panel dynamic environments, particularly under theoretically justified adaptive lag penalization (Korobilis, 2016).

From a theoretical standpoint, an additional gap concerns asymptotic guarantees. Bayesian high-dimensional modeling has developed strong general results on posterior concentration and sparse recovery, yet these results are often established in generic regression frameworks rather than structured dynamic multivariate systems. Foundational asymptotic Bayesian theory provides general tools for posterior contraction analysis under suitable prior concentration and testing conditions, but translating these guarantees into dynamic Bayesian autoregressive systems with hierarchical lag-dependent shrinkage remains nontrivial. In particular, establishing oracle-type behavior requires simultaneously controlling model dimensionality, shrinkage bias, and sparsity recovery under dynamic dependence.

This theoretical gap is important because methodological novelty without formal statistical justification often faces skepticism in econometric and statistical review processes. A shrinkage prior may perform well empirically while lacking clear asymptotic interpretation, limiting confidence in its general applicability. Consequently, there remains a strong need for Bayesian dynamic modeling frameworks that combine practical forecasting effectiveness with formal theoretical support. Taken together, the literature reveals three major unresolved limitations. First, existing Bayesian VAR shrinkage methods inadequately address flexible lag adaptivity. Most models either impose fixed polynomial lag decay or rely on indirect hyperparameter regularization, limiting adaptability to heterogeneous temporal dependence structures. Second, sparse Bayesian frameworks rarely integrate lag-adaptive shrinkage and hierarchical global-local sparsity within a unified high-dimensional panel dynamic architecture (Hamura et al., 2022; Khare & Ghosh, 2022). Existing methods generally emphasize either structured lag shrinkage or coefficient sparsity, but seldom both simultaneously. Third, formal asymptotic guarantees for such adaptive sparse dynamic Bayesian systems remain limited, particularly regarding posterior contraction and oracle-type sparse recovery. These gaps motivate the methodological contribution of the present study.

To address these limitations, this study proposes a Lag-Adaptive Hierarchical Sparse Bayesian Panel Vector Autoregressive (LAHS-BPVAR) framework for high-dimensional multivariate dynamic systems. The proposed methodology integrates three complementary mechanisms within a unified Bayesian architecture. First, we introduce a lag-adaptive prior covariance structure that explicitly models temporal shrinkage using a hybrid decay mechanism:

$$w_l = l^{-\alpha}\exp(-\delta l)$$

where $l$denotes lag order, αcontrols long-memory polynomial decay, and δgoverns short-horizon exponential penalization. This structure generalizes conventional deterministic lag shrinkage by allowing more flexible temporal regularization. The hybrid specification is particularly attractive because it

accommodates heterogeneous persistence structures: slowly decaying dynamics remain supported through the polynomial component, while rapidly vanishing effects are efficiently suppressed through the exponential component. This formulation directly addresses a central limitation of conventional Bayesian VAR shrinkage priors, which typically impose overly rigid lag structures.

Second, the proposed framework embeds this lag-adaptive mechanism within a hierarchical global-local sparse Bayesian structure. Specifically, each autoregressive coefficient receives both a global shrinkage parameter and a coefficient-specific local shrinkage parameter, enabling adaptive regularization across the parameter space. This design combines structured temporal shrinkage with sparse Bayesian learning, allowing the model to aggressively suppress irrelevant coefficients while preserving substantively important dynamic relationships. In contrast to conventional shrinkage methods that primarily regularize coefficients uniformly, the proposed framework simultaneously adapts across both lag order and coefficient relevance.

Third, this study provides a formal theoretical justification for the proposed methodology. Under suitable regularity conditions, the lag-adaptive shrinkage structure induces summability of prior variances and effectively reduces the operational dimensionality of the model. Formal asymptotic justification through posterior contraction and oracle-type sparse recovery behavior under high-dimensional conditions. While prior literature has established general asymptotic results for sparse Bayesian inference, explicit theoretical treatment of adaptive lag-dependent sparse Bayesian Panel VAR systems remains limited. The present work therefore contributes not only computational methodology but also asymptotic statistical justification. Recent advances in high-dimensional econometric modeling have focused on simultaneous estimation, variable selection, and statistical inference under complex model structures (Chatla & Bhattacharya, 2025).

Fourth, we develop a practical Bayesian posterior computation algorithm based on Gibbs sampling combined with Metropolis-Hastings updates for non-conjugate hyperparameters. The estimation framework remains computationally tractable in moderately large dynamic systems while preserving the flexibility of the hierarchical prior structure. This hybrid estimation design enables practical implementation without sacrificing methodological sophistication. The contribution of this study can therefore be summarized in four dimensions:

1. Methodological contribution: A novel lag-adaptive hierarchical sparse Bayesian Panel VAR framework integrating hybrid temporal shrinkage and global-local sparse regularization.
2. Theoretical contribution: Formal asymptotic justification through posterior contraction and sparse support recovery consistency under high-dimensional conditions.
3. Computational contribution: A practical Gibbs-Metropolis posterior inference algorithm suitable for structured hierarchical Bayesian estimation.
4. Empirical contribution: Demonstration of improved estimation, sparsity recovery, and forecasting performance through simulation experiments and macroeconomic application.

The empirical relevance of the proposed methodology is particularly strong in macroeconomic and financial environments characterized by large dynamic systems, heterogeneous persistence structures, and limited effective sample sizes. In such contexts, traditional unrestricted VAR estimation is often infeasible, while overly rigid shrinkage priors may produce inefficient regularization. By combining adaptive lag shrinkage and sparse Bayesian regularization, the proposed approach offers a more flexible and statistically grounded alternative. The remainder of this paper is organized as follows. Section 2 introduces the proposed LAHS-BPVAR model specification and prior structure. Section 3 establishes the theoretical properties of the proposed estimator, including posterior contraction, sparse support recovery consistency, and asymptotic implications of structured shrinkage. Section 4 presents the Bayesian posterior computation algorithm. Section 5 reports Monte Carlo simulation results evaluating estimation accuracy, sparse recovery, and forecasting performance. Section 6 presents an empirical application to macroeconomic time series data. Section 7 concludes.

2. Model Specification

2.1 Panel VAR Framework

Suppose there are N cross-section units and T time periods. For each unit, the VAR(p) Panel model is $i = 1, \dots, N$ defined as:

$$y_{i,t} = \mu_i + \sum_{l=1}^{p} A_l \, y_{i,t-l} + u_{i,t}, \qquad u_{i,t} \sim N(0, \Sigma)$$

where $y_{i,t} \in \mathbb{R}^K$ is an endogenous variable vector, $A_l \in \mathbb{R}^{K \times K}$ matriks koefisien lag ke-l, $\Sigma \in \mathbb{R}^{K \times K}$ denotes the common innovation covariance matrix shared across all panel units., and μ_i unit-specific fixed effects.

To ensure stationarity, the characteristic polynomial satisfies:

$$\det\left(I_K - \sum_{l=1}^{p} A_l \, z^l\right) \neq 0, \qquad \forall |z| \leq 1$$

which guarantees a stable dynamic system.

2.2 Stacked Regression Representation

For notational convenience, define:

$$x_{i,t} = \left(y_{i,t-1}^{\top}, y_{i,t-2}^{\top}, \dots, y_{i,t-p}^{\top}\right)^{\top} \in \mathbb{R}^{Kp}$$

and coefficient matrix:

$$B = \left(A_1, A_2, \dots, A_p\right) \in \mathbb{R}^{K \times Kp}$$

Then model (1) becomes:

$$y_{i,t} = \mu_i + B x_{i,t} + u_{i,t} \tag{1}$$

Stacking over all units and time periods:

$$Y = M + XB + U$$

Where $Y \in \mathbb{R}^{NT \times K}$, $X \in \mathbb{R}^{NT \times Kp}$, M collects fixed effects, and Ucontains innovations. Using vectorization:

$$\mathrm{vec}(Y) = \mathrm{vec}(M) + Z\beta + \varepsilon$$

where:

$$\beta = \mathrm{vec}(B) \in \mathbb{R}^{K^2 p}$$
$$Z = I_K \otimes X$$

and $\varepsilon \sim N(0, \Sigma \otimes I_{NT})$. This representation transforms the PVAR into a high-dimensional Bayesian regression framework.

2.3 Likelihood Function

Assume error Gaussian:

$$U \sim MN(0, I_{NT}, \Sigma)$$

Then likelihood:

$$Y \mid \beta, \Sigma \sim MN(X\beta, I_{NT}, \Sigma)$$

or equivalent:

$$vec(Y) \mid \beta \sim N\big((I_K \otimes X)\beta, \Sigma \otimes I_{NT}\big)$$

2.4 Lag-Adaptive Hierarchical Sparse Prior

Sparse modeling approaches have been developed for a variety of high-dimensional problems, including covariance estimation through low-rank and sparse decompositions (Kutateladze and Seregina, 2024). To control parameter proliferation in high-dimensional settings, we introduce a hierarchical sparse Bayesian prior for each coefficient:

$$\beta_j \mid \tau, \lambda_j, \alpha, \delta \sim N\big(0, \tau^2 \lambda_j^2 w_j\big)$$

Where $\tau > 0$ is a global shrinkage parameter, $\lambda_j > 0$ is a coefficient-specific local shrinkage parameter, and w_j is a lag-adaptive prior weight (Nagel et al., 2024).

(a) Lag-Adaptive Structure

To explicitly regulate temporal dependence, prior variances are modulated through:

$$w_j = l_j^{-2\alpha} \cdot \exp\big(-\delta l_j\big)$$

where $l_j \in \{1, \dots, p\}$ denotes the lag index corresponding to coefficient j, $\alpha > 0$ controls polynomial decay, and $\delta > 0$ controls exponential decay.

This hybrid lag-adaptive structure combines:

Polynomial shrinkage

$$l_j^{-2\alpha}$$

which preserves slowly decaying long-memory effects, and

Exponential shrinkage

$$\exp\big(-\delta l_j\big)$$

which aggressively suppresses distant lags.

Hence:

$$\mathrm{Var}\big(\beta_j\big) = \tau^2 \lambda_j^2 l_j^{-2\alpha} \exp\big(-\delta l_j\big)$$

This structure generalizes conventional Minnesota-type lag decay by allowing adaptive temporal regularization.

(b) Hierarchical Global-Local Shrinkage

To induce sparsity, we employ horseshoe-type hierarchical shrinkage:

Local shrinkage

$$\lambda_j \sim C^+(0,1)$$

Global shrinkage

$$\tau \sim C^+(0,1)$$

The half-Cauchy specification provides strong shrinkage near zero, heavy tails for important signals, and adaptive sparse regularization. Thus the prior simultaneously allows coefficient sparsity, signal preservation, and adaptive model complexity control (Qin & Ghosh, 2024).

2.5 Fixed Effect Prior

To account for cross-sectional heterogeneity, unit effects receive Gaussian priors:

$$\mu_i \sim N\left(0, \sigma_\mu^2 I_K\right)$$

Where σ_μ^2 is chosen sufficiently large to represent weak prior information. This specification preserves the panel structure while controlling incidental parameter inflation.

2.6 Prior Covariance Structure

For the innovation covariance matrix Σ, we adopt an inverse-Wishart prior:

$$\Sigma \sim IW(\nu_0, S_0)$$

where $\nu_0 > K - 1$ and $S_0 > 0$. This prior ensures positive definiteness and conjugate posterior updating.

2.7 Thresholded Posterior Support Recovery

Because continuous shrinkage priors do not produce exact zeros, variable selection is defined via thresholding:

$$\hat{S} = \left\{j : \left|\hat{\beta}_j\right| > \gamma_{NT}\right\}$$

Where $\hat{\beta}_j = E\left(\beta_j \middle| Y\right)$ and threshold $\gamma_{NT} \to 0$, with $\sqrt{NT}\gamma_{NT} \to \infty$. This construction enables formal asymptotic variable selection consistency required for asymptotic support recovery analysis.

3. Theoretical Properties

3.1 Asymptotic Framework and Regularity Conditions

This section establishes the asymptotic properties of the proposed Lag-Adaptive Hierarchical Sparse Bayesian Panel VAR estimator. Let $\beta_0 \in \mathbb{R}^p$ denote the true parameter vector, where $p = K^2 L$ is the total number of autoregressive coefficients. Define the true active set:

$$S_0 = \left\{j : \beta_{0j} \neq 0\right\}$$

with sparsity level $s = |S_0|$. We consider the asymptotic regime N fixed, $T \to \infty$ so that the effective sample size is $n = NT$ and allow dimensional growth $\log p = o(n)$ This accommodates high-dimensional panel dynamic settings.

Assumption A1 (Stability)

The true Panel VAR process is stable:

$$\det\left(I_K - \sum_{l=1}^{L} A_l z^l\right) \neq 0, \qquad \forall\, |z| \leq 1$$

This ensures stationarity and ergodicity.

Assumption A2 (Restricted Eigenvalue Condition)

For some constant $\kappa > 0$, the design matrix Zsatisfies the restricted eigenvalue condition:

$$\frac{1}{n}\left|\left|Z\Delta\right|\right|_2^2 \geq \kappa\left|\left|\Delta\right|\right|_2^2$$

for all vectors $\Delta \in \mathbb{R}^{p_\beta}$ satisfying:

$$\left|\left|\Delta_{S_0^c}\right|\right|_1 \leq 3\left|\left|\Delta_{S_0}\right|\right|_1$$

where S_0 denotes the true active coefficient set and S_0^c its complement. This condition ensures sufficient identifiability of sparse coefficient structures and prevents excessive collinearity among relevant predictors in high-dimensional estimation.

Assumption A3 (Innovation Covariance)
The innovation covariance matrix satisfies:

$$0 < c_{\min} \leq \lambda_{\min}(\Sigma) \leq \lambda_{\max}(\Sigma) \leq c_{\max} < \infty$$

Assumption A4 (Mixing Dependence)
The Panel VAR process is geometrically α-mixing:

$$\alpha(m) \leq C\rho^m$$

for constants $0 < \rho < 1$. This ensures weak temporal dependence.

Assumption A5 (Beta-Min Condition)
For all active coefficients:

$$\min_{j \in S_0} |\beta_{0j}| > C\epsilon_n$$

where:

$$\epsilon_n = \sqrt{\frac{s \log p}{n}}$$

Assumption A6 (Sparse Truth)
The true model is sparse:

$$s = o\left(\frac{n}{\log p}\right)$$

Assumption A7 (Shrinkage Hyperparameter Behavior)
The global shrinkage satisfies:

$$\tau_n \to 0$$

while:

$$\sqrt{n}\,\tau_n \to \infty$$

ensuring sufficient shrinkage without asymptotic over-penalization.

3.2 Lemma 1 (Lag-Adaptive Summability)

For $w_l = l^{-2\alpha}\exp(-\delta l)$ with $\alpha > 0,\ \delta > 0$ the lag-adaptive prior weights are summable:

$$\sum_{l=1}^{\infty} w_l < \infty$$

Proof: See Appendix A.1

3.3 Lemma 2 (Prior Concentration Around Sparse Truth)

Let the true sparse parameter vector be $\beta_0 \in \mathbb{R}^{p_\beta}$ with active support $S_0 = \{j : \beta_{0j} \neq 0\}$ and sparsity level $s = |S_0|$ (van der Pas et al., 2014). Under the proposed hierarchical lag-adaptive horseshoe prior:

$$\beta_j | \tau, \lambda_j, \alpha, \delta \sim N\left(0, \tau^2 \lambda_j^2 w_j\right)$$

Where $w_j = l_j^{-2\alpha}\exp(-\delta l_j)$ with $\alpha > 0, \delta > 0$ and horseshoe shrinkage priors:

$$\lambda_j \sim C^+(0,1), \tau \sim C^+(0,1)$$

suppose the lag-adaptive weights satisfy $0 < w_j \leq C_w < \infty$ for all $j = 1, \ldots, p_\beta$. Then, for sufficiently small:

$$\epsilon_n = \sqrt{\frac{s\log p_\beta}{n}}$$

the prior assigns sufficient probability mass around the true sparse parameter vector:

$$\Pi\left(\left|\left|\beta - \beta_0\right|\right|_2 < \epsilon_n\right) \geq \exp(-Cn\epsilon_n^2)$$

for some constant $C > 0$.
Proof: See Appendix A.2

3.4 Theorem 1 (Posterior Contraction)

Under Assumptions A1-A7, suppose the true sparse parameter vector $\beta_0 \in \mathbb{R}^{p_\beta}$ satisfies $|S_0| = s$ with:

$$s = o\left(\frac{n}{\log p_\beta}\right)$$

and define the contraction rate:

$$\epsilon_n = \sqrt{\frac{s\log p_\beta}{n}}$$

Then there exists a sufficiently large constant $M > 0$ such that the posterior distribution satisfies:

$$\Pi\left(\left|\left|\beta - \beta_0\right|\right|_2 > M\epsilon_n \middle| Y\right) \to 0$$

in probability as $n \to \infty$
Proof: See Appendix A.3

3.5 Theorem 2 (Selection Consistency)

Suppose Assumptions A1-A7 hold and Theorem 1 establishes posterior contraction at rate:

$$\epsilon_n = \sqrt{\frac{s\log p_\beta}{n}}$$

Define the thresholded posterior support estimator $\hat{S} = \{j: |\hat{\beta}_j| > \gamma_n\}$ where $\hat{\beta}_j = E(\beta_j \mid Y)$ and the threshold sequence satisfies $\gamma_n \to 0$ with $\epsilon_n = o(\gamma_n)$. Assume the beta-min condition:

$$\min_{j \in S_0} |\beta_{0j}| > 2\gamma_n$$

Then $P(\hat{S} = S_0) \to 1$ as $n \to \infty$

3.6 Oracle-Type Asymptotic Implications

The asymptotic results established in Theorem 1 and Theorem 2 provide a foundation for understanding the potential oracle-type behavior of the proposed estimator. Specifically, Theorem 1 establishes posterior contraction around the true sparse parameter vector at rate:

$$\epsilon_n = \sqrt{\frac{s\log p_\beta}{n}}$$

while Theorem 2 guarantees consistent recovery of the true sparse support:

$$P(\hat{S} = S_0) \to 1$$

Under these conditions, the proposed estimator exhibits oracle-type implications in the sense that estimation increasingly concentrates on the true active parameter subspace after correct support recovery. However, a formal oracle efficiency result would require additional assumptions and remains beyond the scope of the present study. In particular, once irrelevant coefficients are consistently excluded and relevant coefficients are correctly retained, estimation effectively concentrates on the true active subspace. A full oracle property result, including asymptotic normality and formal efficiency equivalence, would require stronger regularity conditions and a more detailed treatment of shrinkage bias correction under dependent high-dimensional dynamic settings. Establishing such results remains an important direction for future theoretical research.

3.7 Discussion

The theoretical results provide formal asymptotic justification for the proposed LAHS-BPVAR framework and clarify the distinct roles of its structural components. First, Lemma 1 establishes the summability of the lag-adaptive prior weights:

$$w_l = l^{-2\alpha}\exp(-\delta l).$$

This property is important because it ensures that distant lag coefficients receive increasingly strong regularization, preventing uncontrolled growth of effective model complexity as lag order increases. Unlike conventional polynomial lag decay used in Minnesota-type priors, the hybrid polynomial-exponential structure provides stronger tail control while preserving flexibility for moderately persistent dynamics.

Second, Lemma 2 shows that the hierarchical sparse prior assigns sufficient probability mass around sparse true parameter vectors. This result is fundamental for Bayesian consistency, as posterior contraction requires the prior to adequately represent the true sparse data-generating structure. The combination of heavy-tailed local shrinkage and adaptive global regularization allows the prior to simultaneously accommodate strong signals and aggressive shrinkage of irrelevant coefficients. Third, Theorem 1 establishes posterior contraction at rate

$$\epsilon_n = \sqrt{\frac{s\log p}{n}},$$

showing that posterior uncertainty shrinks toward the true parameter vector as sample size increases. This provides statistical consistency for the proposed estimator in high-dimensional sparse dynamic settings.

Fourth, Theorem 2 establishes support recovery consistency for the thresholded posterior estimator. Because continuous shrinkage priors do not generate exact zeros, thresholding provides a practical mechanism for sparse variable selection. Under suitable threshold calibration and beta-min separation, the estimated support converges to the true sparse structure with probability approaching one. Finally, the asymptotic results suggest potential oracle-type implications under correct support recovery. Once irrelevant coefficients are consistently excluded and relevant coefficients are correctly retained, estimation effectively concentrates on the true active parameter subspace. However, a full oracle efficiency result,

including asymptotic normality and formal efficiency equivalence, lies beyond the scope of the present study and remains an important direction for future research.

Taken together, these results suggest that the proposed framework is not merely computationally convenient, but also statistically principled. The theory supports the integration of lag-adaptive regularization and hierarchical sparse shrinkage within structured Bayesian dynamic modeling. However, these asymptotic guarantees rely on standard regularity assumptions, including stability, geometric mixing dependence, sparse truth, and sufficient signal separation. Extensions to more general dependence structures, time-varying parameter systems, or ultra-high-dimensional regimes remain important directions for future work.

4. Estimation: Gibbs Sampling with Metropolis-Hastings Updates

4.1 Posterior Structure

Let:

$$\Theta = \{\beta, \mu, \Sigma, \lambda, \nu, \tau, \xi, \alpha, \delta\}$$

denote the full parameter set, where:

- B : autoregressive coefficients,
- $\mu = \{\mu_i\}_{i=1}^{N}$: panel fixed effects,
- Σ : innovation covariance matrix,
- $\lambda = \{\lambda_j\}_{j=1}^{p}$: local shrinkage parameters,
- $\nu = \{v_j\}_{j=1}^{p\beta}$: auxiliary local-scale variables,
- τ : global shrinkage parameter,
- ξ : auxiliary global-scale variable,
- α, δ : lag-adaptive hyperparameters.

The joint posterior distribution is:

$$p(\Theta \mid Y) \propto p(Y \mid \beta, \mu, \Sigma)\, p(\beta \mid \lambda, \tau, \alpha, \delta)\, p(\lambda, \nu)\, p(\tau, \xi)\, p(\mu)\, p(\Sigma)\, p(\alpha)\, p(\delta).$$

Because the posterior does not admit a fully closed-form joint solution, inference is performed using a hybrid Markov Chain Monte Carlo (MCMC) scheme combining Gibbs sampling for conjugate components and Metropolis-Hastings updates for non-conjugate hyperparameters.

4.2 Likelihood Function

From Section 2:

$$vec(Y) = vec(M) + Z\beta + \varepsilon$$

with:

$$\varepsilon \sim N(0, \Sigma \otimes I_n)$$

Where $n = NT$ Thus:

$$Y \mid \beta, \mu, \Sigma \sim N(vec(M) + Z\beta, \Sigma \otimes I_n)$$

with likelihood:

$$L(\beta, \mu, \Sigma) \propto |\Sigma|^{-\frac{n}{2}} \exp\left(-\frac{1}{2} r^{\top} (\Sigma^{-1} \otimes I_n) r\right)$$

where:

$$r = vec(Y) - vec(M) - Z\beta$$

4.3 Conditional Posterior for Regression Coefficients

Under the proposed hierarchical lag-adaptive shrinkage prior, each autoregressive coefficient follows

$$\beta_j | \tau, \lambda_j, \alpha, \delta \sim N\left(0, \tau^2 \lambda_j^2 w_j\right),$$

where the lag-adaptive variance weight is defined as $w_j = l_j^{-2\alpha} \exp(-\delta l_j)$, with l_j denoting the lag order associated with coefficient j. Let the prior covariance matrix be

$$D = \text{diag}\left(\tau^2 \lambda_1^2 w_1, \tau^2 \lambda_2^2 w_2, \ldots, \tau^2 \lambda_{p_\beta}^2 w_{p_\beta}\right),$$

where p_β denotes the total number of autoregressive coefficients. From Section 4.2, the Gaussian likelihood is

$$\text{vec}(Y) \mid \beta, \mu, \Sigma \sim N(\text{vec}(M) + Z\beta, \Sigma \otimes I_n).$$

Combining the likelihood with the Gaussian prior yields the conditional posterior distribution

$$\beta | - \sim N(m_\beta, V_\beta),$$

with posterior covariance matrix

$$V_\beta = [Z'(\Sigma^{-1} \otimes I_n)Z + D^{-1}]^{-1},$$

and posterior mean

$$m_\beta = V_\beta Z'(\Sigma^{-1} \otimes I_n)(\text{vec}(Y) - \text{vec}(M)).$$

This update jointly incorporates dependence across endogenous variables through the covariance structure Σ, coefficient-specific adaptive shrinkage via local horseshoe scales λ_j, global sparsity regularization through τ, and lag-dependent shrinkage through the adaptive weights w_j. As a result, posterior sampling of β balances dynamic dependence modeling and structured sparsity regularization within a unified Bayesian framework.

4.4 Conditional Posterior for Fixed Effects

To account for unit-specific heterogeneity, each panel unit is assigned an individual fixed-effect vector

$$\mu_i \sim N\left(0, \sigma_\mu^2 I_K\right),$$

for $i = 1, \ldots, N$. This prior provides regularization for panel-specific intercept terms while preserving cross-sectional heterogeneity in the dynamic system. Conditioning on the autoregressive structure, define the partial residual for unit i at time t as

$$e_{i,t} = y_{i,t} - \sum_{l=1}^{p} A_l \, y_{i,t-l}.$$

Then the observation equation becomes $e_{i,t} = \mu_i + u_{i,t}$, with $u_{i,t} \sim N(0, \Sigma)$. Because the Gaussian prior is conjugate to the Gaussian likelihood, the conditional posterior for each fixed-effect vector is

$$\mu_i | - \sim N(m_{\mu_i}, V_{\mu_i}),$$

where the posterior covariance matrix is

$$V_{\mu_i} = \left(T\Sigma^{-1} + \sigma_\mu^{-2} I_K\right)^{-1},$$

and the posterior mean is

$$m_{\mu_i} = V_{\mu_i}\Sigma^{-1}\sum_{t=1}^{T} e_{i,t}\,.$$

Equivalently,

$$m_{\mu_i} = V_{\mu_i}\Sigma^{-1}\sum_{t=1}^{T}\left(y_{i,t} - \sum_{l=1}^{p} A_l\, y_{i,t-l}\right).$$

This update allows each panel unit to retain its own intercept structure while maintaining a common dynamic coefficient system across the pooled panel model. The Bayesian shrinkage prior on μ_i also stabilizes estimation in dynamic panel settings by preventing excessive variability in unit-specific intercept estimates.

4.5 Conditional Posterior for Covariance Matrix

The innovation covariance matrix captures contemporaneous dependence among endogenous variables in the panel VAR system. Under the pooled specification adopted in this study, a common covariance structure is assumed across all panel units:

$$u_{i,t} \sim N(0, \Sigma).$$

To complete the Bayesian specification, the covariance matrix is assigned an inverse-Wishart prior:

$$\Sigma \sim IW(\nu_0, S_0),$$

Where ν_0 is the prior degrees of freedom, and $S_0$is the prior scale matrix. Define the model residuals as

$$e_{i,t} = y_{i,t} - \mu_i - \sum_{l=1}^{p} A_l\, y_{i,t-l},$$

for $i = 1, \ldots, N, t = 1, \ldots, T$. Given Gaussian innovations, the likelihood contribution of the residual matrix follows a multivariate normal structure, yielding a conjugate inverse-Wishart posterior. Accordingly, the conditional posterior distribution is

$$\Sigma|-\sim IW(\nu_n, S_n),$$

with updated degrees of freedom:

$$\nu_n = \nu_0 + NT,$$

and updated scale matrix:

$$S_n = S_0 + \sum_{i=1}^{N}\sum_{t=1}^{T} e_{i,t}\, e_{i,t}'\,.$$

Substituting the residual definition:

$$S_n = S_0 + \sum_{i=1}^{N}\sum_{t=1}^{T}\left(y_{i,t} - \mu_i - \sum_{l=1}^{p} A_l\, y_{i,t-l}\right)\left(y_{i,t}-\mu_i-\textstyle\sum_{l=1}^{p} A_l\, y_{i,t-l}\right)'\,.$$

This update allows the model to learn the contemporaneous covariance structure directly from the data while preserving conjugate Gibbs sampling efficiency. Because the covariance matrix is shared across panel units, this formulation remains computationally tractable while retaining multivariate dependence across endogenous variables.

4.6 Horseshoe Shrinkage Augmentation

The proposed lag-adaptive shrinkage prior employs a horseshoe global-local hierarchical structure to enable aggressive shrinkage of irrelevant coefficients while preserving large signals (Kohns & Szendrei, 2024). For each autoregressive coefficient:

$$\beta_j | \tau, \lambda_j, \alpha, \delta \sim N\left(0, \tau^2 \lambda_j^2 w_j\right),$$

where $w_j = l_j^{-2\alpha} \exp\left(-\delta l_j\right)$, and λ_j is the local shrinkage parameter, and τ is the global shrinkage parameter (Qin & Ghosh, 2024). The local and global shrinkage parameters follow half-Cauchy priors:

$$\lambda_j \sim C^+(0,1),$$
$$\tau \sim C^+(0,1).$$

Direct sampling from half-Cauchy priors is non-conjugate and computationally inconvenient within Gibbs sampling. To obtain tractable conditional updates, the horseshoe prior is expressed using the auxiliary-variable scale-mixture representation of Makalic and Schmidt (Makalic & Schmidt, 2016).

Under this representation, the local shrinkage hierarchy becomes:

$$\lambda_j^2 | \nu_j \sim IG\left(\frac{1}{2}, \frac{1}{\nu_j}\right),$$

with:

$$\nu_j \sim IG\left(\frac{1}{2}, 1\right).$$

Similarly, the global shrinkage hierarchy becomes:

$$\tau^2 | \xi \sim IG\left(\frac{1}{2}, \frac{1}{\xi}\right),$$

with:

$$\xi \sim IG\left(\frac{1}{2}, 1\right).$$

This hierarchical reparameterization transforms the non-conjugate horseshoe prior into a fully conditionally conjugate structure, enabling efficient Gibbs updates for both local and global shrinkage components. The resulting formulation preserves the desirable statistical properties of the horseshoe prior-namely strong concentration near zero and heavy tails-while substantially improving posterior computation.

4.7 Conditional Posterior for Local Shrinkage

Under the augmented horseshoe representation introduced in Section 4.6, the local shrinkage parameters admit conjugate inverse-gamma conditional posterior distributions. For each autoregressive coefficient, the hierarchical structure is:

$$\beta_j | \tau, \lambda_j, w_j \sim N\left(0, \tau^2 \lambda_j^2 w_j\right),$$

with:

$$\lambda_j^2 | \nu_j \sim IG\left(\frac{1}{2}, \frac{1}{\nu_j}\right),$$

and:

$$\nu_j \sim IG\left(\frac{1}{2}, 1\right).$$

Combining the Gaussian prior contribution from β_j with the inverse-gamma augmentation yields the conditional posterior:

$$\lambda_j^2|-\sim IG\left(1,\frac{\beta_j^2}{2\tau^2 w_j}+\frac{1}{\nu_j}\right).$$

Similarly, the auxiliary latent variable updates as:

$$\nu_j|-\sim IG\left(1,1+\frac{1}{\lambda_j^2}\right).$$

These updates are performed independently for $j=1,\ldots,p_\beta$.

The local shrinkage parameters provide coefficient-specific adaptivity, allowing the model to strongly regularize negligible coefficients while permitting genuinely important signals to escape excessive shrinkage. This local adaptivity is particularly important in high-dimensional dynamic systems, where sparse structure may vary substantially across lag positions and variable interactions.

4.8 Conditional Posterior for Global Shrinkage

The global shrinkage parameter governs the overall degree of sparsity across the autoregressive coefficient system by controlling the collective shrinkage intensity applied to all coefficients. Under the augmented horseshoe hierarchy:

$$\beta_j|\,\tau,\lambda_j,w_j\sim N\left(0,\tau^2\lambda_j^2 w_j\right),$$

for $j=1,\ldots,p_\beta$, with the global augmentation:

$$\tau^2|\,\xi\sim IG\left(\frac{1}{2},\frac{1}{\xi}\right),$$

and:

$$\xi\sim IG\left(\frac{1}{2},1\right).$$

Combining the Gaussian prior contribution across all coefficients with the inverse-gamma augmentation yields the conditional posterior:

$$\tau^2|\,-\sim IG\left(\frac{p_\beta+1}{2},\frac{1}{2}\sum_{j=1}^{p_\beta}\frac{\beta_j^2}{\lambda_j^2 w_j}+\frac{1}{\xi}\right).$$

The auxiliary global latent variable updates as:

$$\xi|\,-\sim IG\left(1,1+\frac{1}{\tau^2}\right).$$

The global shrinkage component plays a complementary role to the local shrinkage parameters (Hamura et al., 2022). While the local scales λ_j allow coefficient-specific adaptivity, the global parameter τ regulates the overall sparsity level of the model. This hierarchical combination enables aggressive regularization in sparse settings while preserving robustness against over-shrinkage of genuinely important dynamic coefficients.

4.9 Metropolis-Hastings Updates for Lag-Adaptive Hyperparameters

The lag-adaptive shrinkage mechanism is governed by the hyperparameters α and δ, which determine the prior variance decay structure:

$$w_j = l_j^{-2\alpha} \exp(-\delta l_j),$$

where l_j denotes the lag order associated with coefficient j. These parameters control the strength and shape of lag-dependent shrinkage α regulates polynomial decay, and δ controls exponential attenuation. Because the conditional posterior distributions of α and δ do not admit closed-form conjugate updates, they are sampled using Metropolis-Hastings steps within the Gibbs sampler.

To preserve positivity constraints $\alpha > 0, \delta > 0$, random-walk proposals are generated on the logarithmic scale:

$$\log \alpha^* = \log \alpha + \epsilon_\alpha,$$
$$\log \delta^* = \log \delta + \epsilon_\delta,$$

where:

$$\epsilon_\alpha \sim N(0, s_\alpha^2),$$
$$\epsilon_\delta \sim N(0, s_\delta^2),$$

with tuning variances s_α^2 and s_δ^2.

The proposed values are then obtained via:

$$\alpha^* = \exp(\log \alpha^*),$$
$$\delta^* = \exp(\log \delta^*).$$

Independent Gamma priors are assigned:

$$\alpha \sim Gamma(a_\alpha, b_\alpha),$$
$$\delta \sim Gamma(a_\delta, b_\delta).$$

The Metropolis-Hastings acceptance probability is:

$$A = \min\left(1, \frac{p(\beta \mid \alpha^*, \delta^*)\, p(\alpha^*)\, p(\delta^*)}{p(\beta \mid \alpha, \delta)\, p(\alpha)\, p(\delta)} \times \frac{\alpha^* \delta^*}{\alpha \delta}\right).$$

The Jacobian adjustment:

$$\frac{\alpha^* \delta^*}{\alpha \delta}$$

arises from the logarithmic parameter transformation.

These updates allow the lag-decay structure to be learned adaptively from the data rather than fixed a priori, improving flexibility relative to conventional static lag-shrinkage specifications. Practical tuning of proposal variances is performed to maintain stable acceptance rates, typically within the range of 20%-40%, consistent with standard MCMC practice.

4.10 Posterior Support Recovery

Because continuous shrinkage priors such as the horseshoe do not produce exact zero estimates, explicit support identification requires an additional thresholding rule (Kohns & Szendrei, 2024). Accordingly, theoretical support recovery is defined through a thresholded posterior estimator:

$$\hat{S} = \{j: |\hat{\beta}_j| > \gamma_n\},$$

where:

$$\hat{\beta}_j = E(\beta_j|Y)$$

denotes the posterior mean of coefficient j, and $\gamma_n$is a vanishing threshold sequence. Consistent with the asymptotic theory developed in Section 3, the threshold is chosen to satisfy $\gamma_n \to 0$, and $\epsilon_n = o(\gamma_n)$, where the posterior contraction rate is:

$$\epsilon_n = \sqrt{\frac{s \log p_\beta}{n}}.$$

Under these conditions, thresholding asymptotically separates negligible coefficients from genuinely active signals.

For finite-sample implementation, the threshold should be selected sufficiently above the posterior contraction rate to maintain stable separation between inactive and active coefficients. In practice, the threshold may be calibrated empirically using posterior uncertainty summaries, sensitivity analysis, or model-selection diagnostics. This construction provides an interpretable bridge between posterior shrinkage estimation and sparse support recovery. It is important to emphasize that this thresholding rule serves primarily as a theoretical device for establishing selection consistency and oracle-type asymptotic behavior. In practical empirical applications, posterior credible intervals or alternative Bayesian variable selection diagnostics may also be employed.

5. Simulation Study

A comprehensive simulation study was conducted to evaluate the performance of the proposed Lag-Adaptive Horseshoe (LAHS) prior under various high-dimensional panel VAR settings. The competing methods include ordinary least squares (OLS), the Minnesota prior, the Horseshoe prior, and Bayesian LASSO. The simulation experiments were designed to assess estimation accuracy, coefficient recovery, and lag-selection performance under different structural characteristics commonly encountered in panel VAR applications. The panel VAR model is defined as

$$Y_{it} = \sum_{l=1}^{p} A_l \, Y_{i,t-l} + \mu_i + \varepsilon_{it},$$

where $A_l$denotes the coefficient matrix associated with lag order l, $\mu_i$represents individual-specific effects, and ε_{it}denotes the innovation term. Model performance was evaluated using Root Mean Squared Error (RMSE), True Positive Rate (TPR), False Positive Rate (FPR), and F1-score. All reported results are based on 100 Monte Carlo replications.

Four simulation scenarios were considered. Scenario A represents a sparse coefficient structure, Scenario B introduces weak coefficient signals, Scenario C considers strong cross-variable correlations, and Scenario D examines a denser coefficient structure. The corresponding results are reported in Tables 1-4.

Table 1. Simulation Results for Scenario A (Sparse Structure)

Method	RMSE	TPR	FPR	F1
OLS	0.0567	0.6727	0.1822	0.5144
Minnesota	0.0556	0.6710	0.1748	0.5205
Horseshoe	0.0559	0.6715	0.1735	0.5227
Bayesian LASSO	0.0567	0.6727	0.1813	0.5152
LAHS	**0.0554**	0.6715	**0.1586**	**0.5390**

The results in Table 1 indicate that LAHS achieves the lowest RMSE and the highest F1-score among all competing methods under sparse coefficient structures. Although the true positive rates are relatively similar across methods, LAHS produces the lowest false positive rate, suggesting more effective elimination of irrelevant coefficients. This reduction in false discoveries directly contributes to improved overall recovery performance. The results demonstrate that adaptive lag-dependent shrinkage is particularly beneficial when the underlying coefficient matrix is highly sparse. By selectively shrinking unimportant lag coefficients toward zero, LAHS is able to maintain estimation accuracy while improving model parsimony. These findings support the suitability of the proposed prior for high-dimensional panel VAR applications characterized by sparse dynamic relationships..

Table 2. Simulation Results for Scenario B (Weak Signals)

Method	RMSE	TPR	FPR	F1
OLS	0.0409	0.3959	0.0562	0.3742
Minnesota	0.0395	0.3922	**0.0490**	0.3883
Horseshoe	0.0403	0.3939	0.0538	0.3782
Bayesian LASSO	0.0408	0.3957	0.0563	0.3738
LAHS	**0.0395**	0.3946	0.0507	**0.3892**

Table 2 presents the results under weak-signal conditions, where relevant coefficients are more difficult to distinguish from noise. As expected, all methods exhibit lower F1-scores than in Scenario A, reflecting the increased estimation difficulty. Nevertheless, LAHS attains the lowest RMSE and the highest F1-score among all competing approaches. The improvement suggests that the proposed prior is capable of preserving weak but informative signals while simultaneously controlling false positive selections. This property is particularly important in empirical applications where many dynamic effects are relatively small in magnitude. Overall, the results indicate that adaptive lag-dependent shrinkage provides additional protection against excessive shrinkage of relevant coefficients.

Table 3. Simulation Results for Scenario C (Correlated Structure)

Method	RMSE	TPR	FPR	F1
OLS	0.0400	0.3656	0.0497	0.3642
Minnesota	**0.0390**	0.3644	**0.0449**	**0.3748**
Horseshoe	0.0396	0.3642	0.0477	0.3678

Method	RMSE	TPR	FPR	F1
Bayesian LASSO	0.0400	0.3656	0.0497	0.3643
LAHS	0.0391	**0.3659**	0.0469	0.3712

The results in Table 3 evaluate performance under strong cross-variable correlations. Such settings are known to complicate estimation because correlated predictors may increase coefficient uncertainty and reduce variable-selection accuracy. Although Minnesota achieves a marginally higher F1-score, the difference relative to LAHS remains small. Moreover, LAHS attains the highest true positive rate while maintaining competitive RMSE and false positive rates. These findings suggest that the proposed prior remains robust in the presence of substantial dependence among variables. Therefore, the adaptive lag-specific shrinkage mechanism continues to provide reliable coefficient recovery even when the data exhibit strong correlation structures.

Table 4. Simulation Results for Scenario D (Dense Structure)

Method	RMSE	TPR	FPR	F1
OLS	0.0418	0.4055	0.0657	0.3602
Minnesota	0.0409	0.4029	**0.0586**	0.3734
Horseshoe	0.0414	0.4044	0.0625	0.3661
Bayesian LASSO	0.0418	0.4054	0.0656	0.3603
LAHS	**0.0408**	**0.4041**	0.0590	**0.3736**

Table 4 reports the results for the dense coefficient structure, where a larger proportion of coefficients are active. Compared with sparse settings, this scenario presents additional challenges because shrinkage methods must avoid eliminating genuinely important coefficients. The results show that LAHS achieves the lowest RMSE and the highest F1-score among all competing methods. At the same time, its false positive rate remains comparable to that of Minnesota. These findings indicate that LAHS is capable of balancing coefficient recovery and model sparsity even when the assumption of extreme sparsity is relaxed. Consequently, the proposed prior appears suitable for a broader range of panel VAR structures beyond highly sparse environments. The corresponding lag recovery rates are reported in Table 5.

Table 5. Lag Recovery under Alternative Memory Structures

	Lag	Minnesota	LAHS
Short-Memory Structure	1	1.000	1.000
	2	0.000	0.000
	3	0.000	0.000
	4	0.000	0.000
Long-Memory Structure	1	1.000	1.000
	2	1.000	1.000
	3	0.750	**0.975**
	4	0.175	**0.250**

The lag recovery experiment provides the most direct evaluation of the adaptive lag-dependent shrinkage mechanism that constitutes the central methodological contribution of the proposed LAHS-BPVAR

framework. Unlike the previous simulation scenarios, which primarily assess overall estimation accuracy and sparse coefficient recovery, Table 5 specifically examines the ability of competing priors to distinguish between short-memory and long-memory dynamic structures. Under the short-memory setting, both Minnesota and LAHS achieve perfect recovery of the true lag pattern. Because only the first lag remains active, higher-order lag coefficients should be aggressively shrunk toward zero. The identical recovery performance indicates that the proposed adaptive lag structure does not introduce unnecessary complexity when the underlying dynamic process is dominated by short-run dependence. In this setting, the exponential shrinkage component effectively suppresses irrelevant distant lags while preserving the true first-order dynamic relationship.

A substantially different pattern emerges under the long-memory structure. While both approaches perfectly recover the first and second lag orders, the recovery rates diverge noticeably for higher-order lags. In particular, LAHS improves lag-3 recovery from 0.750 to 0.975 and lag-4 recovery from 0.175 to 0.250 relative to the Minnesota prior. These gains indicate that the proposed adaptive shrinkage mechanism is more effective at preserving weak but genuinely relevant higher-order dynamic effects that may be excessively penalized under conventional deterministic lag-decay specifications. The observed behavior is consistent with the theoretical design of the proposed prior covariance structure. The polynomial component $l^{-2\alpha}$allows slowly decaying dynamic effects to remain sufficiently supported, whereas the exponential component $\exp(-\delta l)$continues to penalize irrelevant distant lags. Consequently, the proposed prior can adapt to heterogeneous persistence patterns rather than imposing a single predetermined decay profile across all lag orders. This flexibility becomes particularly valuable in economic and financial systems where some shocks dissipate rapidly while others propagate over longer horizons.

Taken together, the results in Table 5 provide direct empirical evidence that the performance improvements observed in the broader simulation study are not solely attributable to sparse horseshoe regularization. Instead, they suggest that the adaptive lag-dependent shrinkage mechanism contributes independently by improving recovery of persistent dynamic structures while retaining effective suppression of irrelevant higher-order lags. This finding provides strong empirical support for the primary methodological innovation introduced in this study.

6. Empirical Application

6.1 Data Description and Empirical Setup

To illustrate the practical applicability of the proposed LAHS-BPVAR framework, an empirical analysis was conducted using annual macroeconomic data from five ASEAN economies, namely Indonesia, Malaysia, the Philippines, Singapore, and Thailand. The sample period spans 2000–2020, yielding a balanced panel consisting of 105 observations. Five key macroeconomic and financial indicators were selected to characterize regional economic dynamics: GDP growth, inflation, unemployment, foreign direct investment (FDI), and domestic credit to the private sector. GDP growth captures overall economic performance, inflation reflects price stability, unemployment represents labor market conditions, FDI measures external investment activity, and domestic credit serves as an indicator of financial sector development. These variables are widely used in macroeconomic panel VAR studies and collectively capture important channels of economic transmission within ASEAN economies.

Prior to estimation, all variables were standardized to eliminate scale differences and facilitate comparison across variables. Missing observations were interpolated using linear interpolation to obtain a balanced panel structure. The empirical model was estimated using a first-order panel VAR specification. Three competing approaches were considered: ordinary least squares (OLS), Bayesian panel VAR with Minnesota prior, and the proposed LAHS-BPVAR estimator. To evaluate the sparsity properties of competing estimators, active coefficients were identified using a threshold value of 0.20. Coefficients with absolute posterior estimates exceeding the threshold were classified as active dynamic relationships.

6.2 Empirical Results and Discussion

Table 5 reports the number of active coefficients identified by each estimation method.

Table 6. Active Coefficients Identified by Alternative Estimators

Method	Active Coefficients
OLS	5
Minnesota BPVAR	5
LAHS-BPVAR	9

The results indicate substantial differences in the dynamic structures retained by the competing estimators. Both OLS and Minnesota BPVAR identified only five active autoregressive relationships, whereas the proposed LAHS-BPVAR retained nine active coefficients. This finding suggests that the adaptive lag-dependent shrinkage mechanism preserves additional dynamic relationships that may be excessively shrunk toward zero under conventional global shrinkage priors. At the same time, the resulting model remains highly parsimonious, retaining only 9 of the 25 possible dynamic interactions. To further investigate the economic implications of the retained coefficients, Table 7 summarizes the dynamic relationships selected by LAHS-BPVAR.

Table 7. Dynamic Relationships Selected by LAHS-BPVAR

Lagged Variable ($t-1$)	Current Variable (t)
Inflation	GDP Growth
Unemployment	GDP Growth
Inflation	Inflation
Unemployment	Inflation
Credit	Inflation
Unemployment	Unemployment
Unemployment	FDI
FDI	FDI
Credit	Credit

Several economically meaningful patterns emerge from the estimated dynamic structure. First, both inflation and unemployment exhibit predictive effects on subsequent GDP growth, indicating that macroeconomic stability and labor market conditions play important roles in determining economic performance across ASEAN economies. The presence of these links suggests that growth dynamics are influenced not only by past output conditions but also by broader macroeconomic fundamentals. Second, inflation displays substantial persistence through its own autoregressive effect, a result consistent with the well-documented tendency of inflationary pressures to evolve gradually over time. In addition, both unemployment and domestic credit influence future inflation, suggesting the existence of labor market and

financial transmission channels within the regional economy. These findings are broadly consistent with macroeconomic theories linking labor market conditions, credit expansion, and price dynamics. Third, strong persistence is observed in FDI and domestic credit. The autoregressive effects of these variables indicate that investment flows and financial development evolve gradually and are characterized by medium-term dynamics rather than short-run fluctuations. Such persistence is commonly observed in emerging and developing economies where structural adjustments occur over extended periods.

An additional finding is the central role of unemployment within the estimated system. Unemployment not only exhibits strong persistence but also influences subsequent GDP growth, inflation, and FDI. This pattern suggests that labor market conditions constitute an important transmission mechanism connecting real economic activity, price dynamics, and investment decisions across ASEAN economies. Overall, the empirical results demonstrate that the proposed LAHS-BPVAR framework produces a parsimonious yet economically interpretable dynamic structure. Compared with conventional shrinkage approaches, the adaptive lag-dependent prior retains a richer set of meaningful dynamic relationships while preserving model sparsity. These empirical findings complement the simulation evidence and illustrate the practical usefulness of LAHS-BPVAR for analyzing macroeconomic interactions in panel data environments.

The empirical application should be viewed primarily as an illustration of the proposed methodology. Given the limited number of countries and annual observations, the empirical findings should not be interpreted as definitive evidence regarding regional macroeconomic dynamics. Future work may consider larger macroeconomic or financial datasets with higher observation frequencies.

7. Conclusion

This paper proposed a Lag-Adaptive Hierarchical Sparse Bayesian Panel Vector Autoregressive (LAHS-BPVAR) framework for estimating dynamic relationships in panel VAR systems. The proposed approach combines adaptive lag-dependent shrinkage with hierarchical global-local regularization, allowing shrinkage intensity to vary across lag orders while simultaneously encouraging sparse coefficient structures. Theoretical analysis established posterior contraction and sparse support recovery consistency under standard regularity conditions, providing formal asymptotic justification for the proposed estimator. Simulation studies demonstrated competitive estimation and forecasting performance across a variety of dynamic environments. In particular, the lag recovery experiments showed that the adaptive lag-dependent prior more effectively preserves higher-order dynamic effects in long-memory systems while maintaining appropriate shrinkage under short-memory structures. These findings provide direct evidence that the proposed lag-adaptive mechanism contributes beyond conventional sparse Bayesian regularization.

An empirical application using macroeconomic data from five ASEAN economies further illustrated the practical usefulness of the proposed framework. The estimated model identified several economically interpretable dynamic relationships among economic growth, inflation, unemployment, foreign direct investment, and domestic credit. Compared with conventional shrinkage approaches, LAHS-BPVAR retained a richer yet still parsimonious dynamic structure, highlighting its potential value for empirical macroeconomic analysis. Overall, the findings indicate that combining adaptive lag-sensitive regularization with hierarchical sparsity provides a flexible and effective approach for Bayesian panel VAR estimation. The proposed methodology offers a useful alternative to existing shrinkage priors by balancing model complexity, estimation accuracy, and interpretability.

Several directions remain for future research. First, the framework may be extended to accommodate heterogeneous autoregressive coefficients across panel units. Second, alternative global-local shrinkage structures and dynamic covariance specifications may be explored. Third, future work may consider applications involving larger macroeconomic and financial datasets with mixed-frequency observations and

more complex dependence structures. These extensions may further enhance the applicability of adaptive shrinkage methods in high-dimensional panel time-series analysis.

Reference

Carvalho, C. M., Polson, N. G., & Scott, J. G. (2010). The horseshoe estimator for sparse signals. *Biometrika*, *97*(2), 465–480. https://doi.org/10.1093/biomet/asq017

Chan, J. C. C. (2021). Minnesota-type adaptive hierarchical priors for large Bayesian VARs. *International Journal of Forecasting*, *37*(3), 1212–1226. https://doi.org/10.1016/j.ijforecast.2021.01.002

Chan, J. C. C. (2023). Comparing stochastic volatility specifications for large Bayesian VARs. *Journal of Econometrics*, *235*(2), 1419–1446. https://doi.org/10.1016/j.jeconom.2022.11.003

Chan, J. C. C., & Eisenstat, E. (2018). Comparing hybrid time-varying parameter VARs. *Economics Letters*, *171*, 1–5. https://doi.org/10.1016/j.econlet.2018.06.031

Chatla, S. B., & Bhattacharya, I. (2025). Estimation and inference of quantile treatment effects in high-dimensional single-index model. *Econometrics and Statistics*. https://doi.org/10.1016/j.ecosta.2025.02.003

Chiu, C.-W. (Jeremy), Mumtaz, H., & Pintér, G. (2017). Forecasting with VAR models: Fat tails and stochastic volatility. *International Journal of Forecasting*, *33*(4), 1124–1143. https://doi.org/10.1016/j.ijforecast.2017.03.001

Follett, L., & Yu, C. (2019). Achieving parsimony in Bayesian vector autoregressions with the horseshoe prior. *Econometrics and Statistics*, *11*, 130–144. https://doi.org/10.1016/j.ecosta.2018.12.004

Gefang, D., Koop, G., & Poon, A. (2023). Forecasting using variational Bayesian inference in large vector autoregressions with hierarchical shrinkage. *International Journal of Forecasting*, *39*(1), 346–363. https://doi.org/10.1016/j.ijforecast.2021.11.012

Ghosal, S., Ghosh, J. K., & van der Vaart, A. W. (2000). Convergence rates of posterior distributions. *The Annals of Statistics*, *28*(2). https://doi.org/10.1214/aos/1016218228

Griffin, J. E., & Brown, P. J. (2021). Bayesian global-local shrinkage methods for regularisation in the high dimension linear model. *Chemometrics and Intelligent Laboratory Systems*, *210*, 104255. https://doi.org/10.1016/j.chemolab.2021.104255

Gruber, L., & Kastner, G. (2025). Forecasting macroeconomic data with Bayesian VARs: Sparse or dense? It depends! *International Journal of Forecasting*, *41*(4), 1589–1619. https://doi.org/10.1016/j.ijforecast.2025.02.001

Gutiérrez, I., Alvares, D., & Gutiérrez, L. (2024). A Bayesian flexible model for testing Granger causality. *Econometrics and Statistics*. https://doi.org/10.1016/j.ecosta.2024.08.001

Hamura, Y., Irie, K., & Sugasawa, S. (2022). On Global-Local Shrinkage Priors for Count Data. *Bayesian Analysis*, *17*(2). https://doi.org/10.1214/21-BA1263

Irfan, M., Usman, M., Warsono, & Dewi, W. U. (2026). Modeling the dynamic nexus between fossil fuels, renewables, and energy demand in G20 countries using extended panel VAR-GMM approach. *Next Research*, *3*, 101192. https://doi.org/10.1016/j.nexres.2025.101192

Irfan, M., Usman, M., Warsono, & Wamiliana. (2024). Modelling and forecasting data with panel vector autoregressive GMM estimation on economic growth and inflation data in 10 ASEAN countries. *AIP Conference Proceedings*, *2970*(1), 030005. https://doi.org/10.1063/5.0208249

Khare, K., & Ghosh, M. (2022). MCMC Convergence for Global-Local Shrinkage Priors. *Journal of Quantitative Economics*, *20*(S1), 211–234. https://doi.org/10.1007/s40953-022-00311-0

Kohns, D., & Szendrei, T. (2024). Horseshoe prior Bayesian quantile regression. *Journal of the Royal Statistical Society Series C: Applied Statistics*, *73*(1), 193–220. https://doi.org/10.1093/jrsssc/qlad091

Korobilis, D. (2016). Prior selection for panel vector autoregressions. *Computational Statistics & Data Analysis*, *101*, 110–120. https://doi.org/10.1016/j.csda.2016.02.011

Kutateladze, V., & Seregina, E. (2024). Doubly Sparse estimator for high-dimensional covariance matrices. *Econometrics and Statistics*. https://doi.org/10.1016/j.ecosta.2024.06.001

Lee, N., Choi, H., & Kim, S.-H. (2016). Bayes shrinkage estimation for high-dimensional VAR models with scale mixture of normal distributions for noise. *Computational Statistics & Data Analysis*, *101*, 250–276. https://doi.org/10.1016/j.csda.2016.03.007

Makalic, E., & Schmidt, D. F. (2016). *High-Dimensional Bayesian Regularised Regression with the BayesReg Package*.

Nagel, M., Fischer, L., Pawlowski, T., & Kelava, A. (2024). An Alternative Prior for Estimation in High-Dimensional Settings. *Structural Equation Modeling: A Multidisciplinary Journal*, *31*(6), 939–951. https://doi.org/10.1080/10705511.2023.2281279

Prüser, J. (2023). Data-based priors for vector error correction models. *International Journal of Forecasting*, *39*(1), 209–227. https://doi.org/10.1016/j.ijforecast.2021.10.007

Qin, Z., & Ghosh, M. (2024). Global-Local Shrinkage Priors for Asymptotic Point and Interval Estimation of Normal Means under Sparsity. *Sankhya A*, *86*(1), 93–137. https://doi.org/10.1007/s13171-023-00315-9

Vahap, S. (2025). Bayesian dynamic graphical models for high-dimensional vector autoregressions with time-varying parameters and volatility discounting. *Econometrics and Statistics*. https://doi.org/10.1016/j.ecosta.2025.07.001

van der Pas, S. L., Kleijn, B. J. K., & van der Vaart, A. W. (2014). The horseshoe estimator: Posterior concentration around nearly black vectors. *Electronic Journal of Statistics*, *8*(2). https://doi.org/10.1214/14-EJS962

Appendix

A.1 Lemma 1

Proof

Consider $w_l = l^{-2\alpha}\exp(-\delta l)$, with $\alpha > 0, \delta > 0$. Since $l \geq 1$, and $\alpha > 0$, it follows that $l^{-2\alpha} \leq 1$. Therefore $w_l = l^{-2\alpha}\exp(-\delta l) \leq \exp(-\delta l)$.Hence:

$$\sum_{l=1}^{\infty} w_l = \sum_{l=1}^{\infty} l^{-2\alpha}\exp(-\delta l) \leq \sum_{l=1}^{\infty} \exp(-\delta l).$$

The right-hand side is a geometric series with ratio $e^{-\delta}$, where $0 < e^{-\delta} < 1$ because $\delta > 0$. Thus:

$$\sum_{l=1}^{\infty} \exp(-\delta l) = \frac{e^{-\delta}}{1-e^{-\delta}}.$$

Since this quantity is finite,

$$\frac{e^{-\delta}}{1-e^{-\delta}} < \infty,$$

we conclude:

$$\sum_{l=1}^{\infty} w_l < \infty.$$

Therefore, the lag-adaptive prior weights are absolutel summable. ■

A.2 Lemma 2

Partition the parameter space into active and inactive coordinates $S_0 = \{j: \beta_{0j} \neq 0\}$ and $S_0^c = \{j: \beta_{0j} = 0\}$. Then:

$$\| \beta - \beta_0 \|_2^2 = \sum_{j \in S_0} (\beta_j - \beta_{0j})^2 + \sum_{j \in S_0^c} \beta_j^2$$

For the active coefficients $j \in S_0$ the horseshoe prior possesses heavy tails, ensuring positive prior density in neighborhoods around nonzero true coefficients (van der Pas et al., 2014). Because $0 < w_j \leq C_w$ the lag-adaptive scaling only modifies the variance magnitude by a bounded factor and therefore does not eliminate prior support near β_{0j}. Hence, there exists a constant $c_1 > 0$ such that:

$$\Pi(|\beta_j - \beta_{0j}| < \epsilon_n) \geq c_1\epsilon_n$$

for each $j \in S_0$. Thus:

$$\Pi\left(\sum_{j \in S_0} (\beta_j - \beta_{0j})^2 < \frac{\epsilon_n^2}{2}\right) \geq (c_1\epsilon_n)^s$$

For inactive coefficients $j \in S_0^c$ the horseshoe prior strongly concentrates near zero due to its sharp spike at the origin. Therefore, there exists $c_2 > 0$ such that:

$$\Pi\left(\sum\nolimits_{j \in S_0^c} \beta_j^2 < \frac{\epsilon_n^2}{2}\right) \geq \exp(-c_2 s \log p_\beta)$$

Combining both parts:

$$\Pi(\| \beta - \beta_0 \|_2 < \epsilon_n) \geq \exp(-Cs\log p_\beta)$$

for some $C > 0$. Since $n\epsilon_n^2 = s\log p_\beta$ it follows that:

$$\Pi(\| \beta - \beta_0 \|_2 < \epsilon_n) \geq \exp(-Cn\epsilon_n^2)$$

which establishes the required prior concentration property.

A.3 Thoerem 1 (Posterior Construction)

The proof follows standard Bayesian posterior contraction arguments adapted to sparse dynamic regression settings. From Lemma 2, the proposed lag-adaptive hierarchical horseshoe prior assigns sufficient prior mass around the true sparse parameter vector:

$$\Pi(\| \beta - \beta_0 \|_2 < \epsilon_n) \geq \exp(-Cn\epsilon_n^2)$$

for some constant $C > 0$. Thus, the prior concentration condition is satisfied. Next, under Assumption A2 (restricted eigenvalue condition), the sparse parameter space remains identifiable, preventing severe collinearity among relevant predictors. Under Assumptions A1 and A4, the panel VAR process is stable, stationary, and geometrically mixing, ensuring weak temporal dependence. Consequently, exponentially consistent tests exist for distinguishing:

$$H_0 : \beta = \beta_0$$

against alternatives:

$$H_1 : \| \beta - \beta_0 \|_2 > M\epsilon_n$$

as required by Bayesian contraction theory. Furthermore, Assumption A3 ensures bounded innovation covariance, which stabilizes likelihood behavior. Since:

$$s = o\left(\frac{n}{\log p_\beta}\right)$$

the effective model complexity remains sufficiently controlled, implying that the entropy of the sparse parameter class satisfies:

$$\log N(\epsilon_n, \Theta_n, \|\cdot\|_2) \leq Cn\epsilon_n^2$$

for an appropriate sieve Θ_n containing sparse coefficient vectors. Therefore, the three standard posterior contraction conditions are satisfied:

1. sufficient prior mass near truth,
2. existence of exponentially consistent tests,
3. manageable model entropy.

Applying the general Bayesian posterior contraction theorem (Ghosal et al., 2000), it follows that:

$$\Pi(\| \beta - \beta_0 \|_2 > M\epsilon_n \mid Y) \to 0$$

in probability.

■

A.4 Proof of Theorem 2

From Theorem 1, posterior contraction implies:

$$\Pi(\| \beta - \beta_0 \|_2 > M\epsilon_n \mid Y) \to 0$$

in probability. This implies that the posterior mean estimator satisfies:

$$\| \hat{\beta} - \beta_0 \|_2 = o_p(\gamma_n)$$

Since $\epsilon_n = o(\gamma_n)$ We now consider inactive and active coordinates separately.

Inactive coefficients

For $j \in S_0^c$, we have $\beta_{0j} = 0$. Thus:

$$| \hat{\beta}_j | = | \hat{\beta}_j - \beta_{0j} | \leq \| \hat{\beta} - \beta_0 \|_2$$

Since $\| \hat{\beta} - \beta_0 \|_2 = o_p(\gamma_n)$ it follows that:

$$P(| \hat{\beta}_j | > \gamma_n) \to 0$$

Therefore, inactive coefficients are asymptotically excluded:

$$P(S_0^c \cap \hat{S} = \emptyset) \to 1$$

Active coefficients

For $j \in S_0$ the beta-min condition gives $| \beta_{0j} | > 2\gamma_n$. Using the triangle inequality:

$$| \hat{\beta}_j | \geq | \beta_{0j} | - | \hat{\beta}_j - \beta_{0j} |$$

Thus:

$$| \hat{\beta}_j | > 2\gamma_n - o_p(\gamma_n)$$

which implies $| \hat{\beta}_j | > \gamma_n$ with probability approaching one. Hence:

$$P(S_0 \subseteq \hat{S}) \to 1$$

Combining both results:

$$P(\hat{S} = S_0) \to 1$$

Therefore, the thresholded posterior estimator consistently recovers the true sparse support.

B.1 Sensitivity to Threshold Selection

Because continuous shrinkage priors such as the horseshoe do not produce exact zero coefficients, empirical support recovery depends on the selected threshold level used for coefficient screening. To assess robustness, alternative threshold values were considered:

$$\gamma \in \{0.10,\ 0.20\}$$

Table B1 reports the resulting number of retained dynamic interactions.

Model	**Threshold = 0.10**	**Threshold = 0.20**
OLS	5	5
Minnesota BPVAR	5	5
LAHS-BPVAR	13	9

The results indicate that the proposed LAHS-BPVAR consistently preserves a richer dynamic dependence structure than conventional benchmark estimators across alternative threshold specifications. While the absolute number of retained interactions decreases under stricter thresholding, the qualitative interpretation remains unchanged. This suggests that the empirical findings are not driven by an arbitrary threshold choice.

B.2 MCMC Convergence Diagnostics

Posterior estimation of the proposed LAHS-BPVAR model relies on Gibbs sampling combined with Metropolis–Hastings updates for the lag-adaptive hyperparameters. The observed acceptance rates were:

$$Acceptance(\alpha) = 0.449$$

and

$$Acceptance(\delta) = 0.063$$

The acceptance rate for the primary lag-adaptive shrinkage parameter αfalls within an acceptable range for Metropolis-Hastings sampling, indicating stable posterior exploration. The lower acceptance rate for δsuggests stronger posterior concentration around the exponential lag-decay component, which is not unexpected given the short annual time dimension of the empirical sample. Overall, posterior computation remained numerically stable throughout estimation.

B.3 Robustness Interpretation

Taken together, the robustness checks indicate that the empirical conclusions remain qualitatively stable under alternative threshold specifications and posterior sampling diagnostics. These results reinforce the interpretation that the proposed LAHS-BPVAR framework provides a flexible and computationally feasible mechanism for uncovering heterogeneous dynamic dependence structures in macroeconomic panel systems.