

Alternative Bayesian Regularizations for Factor Correlations and Loadings: Lasso, Spike-and-Slab, and Horseshoe

Yifan Zhang and Jinsong Chen

The University of Hong Kong

ABSTRACT

We embed Bayesian graphical modeling within partially confirmatory factor analysis (PCFA) to estimate sparse latent factor correlations, thereby overcoming the limitations of the inverse-Wishart prior. We regularize the factor correlation matrix using graphical lasso, horseshoe, and spike-and-slab priors, yielding a unified framework that applies sparsity to loadings and factor correlations. For complex structures, we further introduce partial regularization for multitrait-multimethod and extended bifactor models. Across simulation studies, the spike-and-slab prior achieves superior false-positive control and high accuracy, mainly in large-sample, complex, or sparse settings. The horseshoe prior demonstrates stable performance across sample sizes but yields more false positives, whereas the lasso prior shows moderate and reliable performance. For model identification, significance-based criteria outperform threshold-based rules for regularized estimators, especially in complex designs and empirical analyses. A personality-inventory application shows that partial regularization improves interpretability and model fit in estimating factor correlations and loadings.

KEYWORDS

Bayesian graphical model; factor correlation; horseshoe; lasso; spike-and-slab

1. Introduction

Latent factor models underpin measurement and structural modeling in the social, behavioral, and educational sciences. Among their key components, factor correlations play a critical role as they capture the relationships among latent constructs and guide theoretical interpretations of higher-order dimensions, redundancy, and discriminant validity. In many empirical applications, the factor correlation matrix is sparse: although a subset of factor pairs shows substantial associations, most others are only weakly related (Mahnke et al., 2014; Matzler et al., 2013). More broadly, these patterns may arise from complex latent structures or theory-driven specifications, in which associations are expected among selected factor pairs while others remain weakly related or approximately independent. When such sparsity is not explicitly modeled, conventional estimation approaches may fail to recover the underlying structure accurately, leading to inflated noise correlations, reduced interpretability, and overfitting.

Partially Confirmatory Factor Analysis (PCFA) offers a flexible Bayesian framework by integrating lasso regularization into confirmatory modeling (Chen, 2020). In PCFA, factor loadings can be estimated using either conventional Gaussian priors or regularization that shrinks small or spurious values toward zero. PCFA also incorporates graphical lasso for the residual covariance matrix, enabling

regularization of local dependencies. However, existing PCFA implementations estimate the factor correlation matrix using an inverse-Wishart prior, which, consistent with prior work on Bayesian factor analysis (Lee, 1981), neither encourages sparsity nor accommodates structural hypotheses about factor relationships.

In this research, we extend PCFA by introducing PCFA with Regularized Factor Correlation (PCFAwRFC), which applies Bayesian graphical modeling to the latent factor correlation matrix. In recent years, a number of methods have been proposed for estimating sparse graphical models in genetic data analysis, genomics, and other high-dimensional biological applications (Friedman et al., 2008; Shaddox et al., 2018; Wang, 2012). Bayesian graphical regularization is built on the covariance (or precision) matrix and offers a powerful framework for representing conditional dependencies among variables (Khondker et al., 2013). In Gaussian graphical models, a zero entry in the precision matrix indicates conditional independence between two variables, which is a stronger and more informative statement than a zero entry in the covariance matrix reflecting only marginal independence. Building on this idea, Pan et al. (2017) applied graphical lasso to the residual covariance structure in CFA to identify sparse dependencies among residuals.

To adapt these ideas to factor structure, we treat the factor covariance matrix as a graphical structure in which



latent factors serve as nodes and factor correlations represent edges. Regularizing this matrix therefore allows us to identify interpretable latent-factor relationships while maintaining the substantive meaning of factor correlations in the factor model. Our work differs from most existing approaches in that it targets the factor covariance structure rather than the residual covariance among observed items. By modeling at the latent level, PCFAwRFC strikes an optimal balance between sparsity and fidelity: near-zero correlations are shrunk toward zero, whereas substantial associations are retained. This also aligns the treatment of factor correlations with that of loadings and residual covariances in PCFA.

Regularizing factor correlations is particularly valuable in complex latent factor structures. Schmid and Leiman (1957) showed that sparse factor structures naturally extend to hierarchical or multifaceted models that better reflect the nature of the data. For instance, multitrait-multimethod (MTMM) designs introduce both trait and method factors, leading to intricate cross-correlation patterns that are challenging to identify and interpret (Campbell & Fiske, 1959). Widaman (1985) categorized MTMM structures into 16 forms based on how trait and method factors are allowed to correlate, underscoring the structural diversity in such models. Similarly, bifactor models represent variance attributable to a general factor alongside group (specific or special) factors. Classic bifactor models assume orthogonality among latent factors for parsimony (Holzinger & Swineford, 1937). In contrast, oblique bifactor models relax this assumption by allowing correlated group factors, enabling a more flexible representation of residual associations among group constructs (Reise, 2012). Recent extensions of bifactor with multiple general factors further incorporate multiple correlated general factors with sparse correlated group factors to capture the intricate dependencies in empirical data (Jiménez et al., 2023). Collectively, these developments motivate modeling complex factor structures under (partially) sparse assumptions, in which most factors remain nearly independent while preserving key dependencies that reflect the underlying data structure.

Zhang and Chen (2024) extended PCFA to accommodate multiple general factors and orthogonal special factors. However, strict orthogonality among special factors may obscure meaningful relationships. Moreover, Zhang and Chen (2024) reported that reliance on the inverse-Wishart prior for factor correlation estimation can hinder recovery, particularly in complex latent structures. To address these limitations, we propose PCFA with Partial Regularized Factor Correlation (PCFAwPRFC): correlations among general factors are estimated using conventional inverse-Wishart prior, while correlations among special factors are regularized to induce partial sparsity. This design preserves important relationships among general constructs while improving interpretability and parsimony for special factors, thereby aligning model flexibility with empirical complexity.

In Bayesian regularization, penalty priors can be specified in different forms. Growing evidence suggests that the Bayesian lasso (Park & Casella, 2008) used in factor analysis

may underperform relative to alternative shrinkage priors in some settings (Harra & Kaplan, 2024; Yamaguchi & Zhang, 2023). Beyond lasso, we consider two additional shrinkage priors: the horseshoe prior (Carvalho et al., 2010) and the spike-and-slab prior (SSP, Mitchell & Beauchamp, 1988). The horseshoe prior adopts global-local shrinkages to aggressively shrink near-zero parameters while preserving large effects, whereas the spike-and-slab prior employs a two-component mixture that explicitly separates negligible effects (the spike) from substantial coefficients (the slab). Although these priors are well established in Bayesian variable selection (van Erp, 2023), their simultaneous application to three parameter matrices (factor loading, factor correlation, and residual covariance) is novel and technically nontrivial. In this study, we introduce graphical regularization to factor correlations using three priors (lasso, horseshoe, and SSP) and extend the original PCFA framework by incorporating horseshoe and spike-and-slab priors into the loading-regularization component. These priors share parallel hyperparameter structures across loadings and factor correlations, enabling unified comparison within the PCFA framework. Because a full investigation across all parameter spaces is beyond the scope of this article, we focus on regularizing factor correlations and factor loadings, leaving residual covariance regularization for future work.

The remainder of the article is organized as follows. We first introduce PCFA with Regularized Factor Correlation (PCFAwRFC) and its estimation procedures with Bayesian lasso, horseshoe and spike-and-slab priors. We then propose its extension, PCFA with Partial Regularized Factor Correlation (PCFAwPRFC), for complex factor structures. This is followed by simulation studies to evaluate model performance under varied structural conditions, as well as an empirical analysis using real data. Finally, we conclude with a discussion of findings and implications for future research.

2. Theoretical Framework

2.1. Partially Confirmatory Factor Analysis with Regularized Factor Correlation (PCFAwRFC)

The PCFA framework strikes a balance between purely exploratory and confirmatory factor analysis using Bayesian regularization (Chen, 2020). In PCFA, $y_1, y_2, \dots, y_N$ are independent observations that satisfy

$$y_i = \mu + \Lambda f_i + \varepsilon_i, i = 1, 2, \dots, N \quad (1)$$

where μ denotes the intercept and Λ denotes the loading matrix. Researchers can fix or free selected elements of the factor loading matrix Λ based on substantive theory while allowing other elements to be estimated with Bayesian lasso, enabling flexible model specification without the rigidity of fully confirmatory approaches. The latent factors follow $f_i \sim \mathcal{N}_K(0, \Phi)$ and residuals follow $\varepsilon_i \sim \mathcal{N}_J(0, \Psi)$. Here $\Phi = (\phi_{kk'})$ is the $K \times K$ factorial covariance matrix and Ψ is the $J \times J$ residual covariance matrix. PCFA generally assumes a multivariate normal distribution for the factor scores and

places an inverse-Wishart prior on the factor covariance matrix Φ .

In PCFAwRFC, we adopt the Bayesian lasso for factor loadings in the subsequent simulations. The full conditional distribution for the unspecified loadings in PCFA is provided in [Appendix A](#). For more details of PCFA, please refer to [Chen \(2020\)](#).

Bayesian graphical models consider a p-dimensional random vector $\mathbf{y}$ following a normal distribution with zero mean and covariance Σ as $\mathbf{y} \sim \mathcal{N}(0, \Sigma)$. The goal of graphical models is to maximize the penalized likelihood

$$\log(\det(\Omega)) - \text{tr}\left(\frac{\mathbf{S}}{n}\Omega\right) - \text{pen}(\Omega) \quad (2)$$

where $\mathbf{S} = \mathbf{Y}'\mathbf{Y}$, and $\mathbf{Y}$ is the data matrix with n independent samples $\mathbf{y}$. The Ω in penalized function is the inverse matrix of covariance matrix Σ . The penalty term $\text{pen}(\Omega)$ includes different shrinkage parameters and norm functions. For instance, Bayesian lasso uses $\rho\|\Omega\|_1$ as the L_1 -norm of Ω .

PCFAwRFC extends the PCFA framework by replacing the inverse-Wishart prior on Φ with graphical model-based priors (e.g., lasso, horseshoe and SSP), thereby integrating sparse structure estimation for factor correlations directly into the model. The graphical model framework assigns priors to the inverse factor covariance matrix $\Omega = (\omega_{kk'}) = \Phi^{-1}$, rather than on Φ itself. In Equation (2), we define the sum-of-products matrix $\mathbf{S} = \mathbf{F}'\mathbf{F}$, where $\mathbf{F} \in \mathbb{R}^{I \times K}$ denotes the matrix of latent factor scores. Unlike classical graphical models that operate on observed variables, this formulation applies to unobserved latent factors, making $\mathbf{S}$ a latent-level covariance summary.

2.1.1. Graphical Lasso

We adopt the lasso prior within the graphical model for factor covariance matrix. Bayesian graphical lasso integrates double exponential density for factor covariance (off-diagonal elements of Φ) and exponential density for factor variances (diagonal elements of Φ). As a result, graphical lasso estimators are equivalent to maximizing a posterior estimation:

$$p(\Omega|\lambda) = \prod_{k < k'} \{DE(\omega_{kk'}|\lambda)\} \prod_{k=1}^K \left\{ \text{Exp}\left(\omega_{kk} \middle| \frac{\lambda}{2}\right) \right\} 1_{\Omega \in M^+} \quad (3)$$

where M^+ denotes the space of positive-definite matrices and λ is the shrinkage parameter. To enable efficient Gibbs sampling, we follow the hierarchical representation of [Wang \(2012\)](#), introducing latent scale parameters τ :

$$\begin{aligned} \omega_{kk} &\sim \text{Exp}\left(\omega_{kk} \middle| \frac{\lambda}{2}\right), \\ \omega_{kk'}, k < k' &\sim N(\omega_{kk'} | 0, \tau_{kk'}), \\ \tau_{kk'}, k < k' &\sim \text{Exp}\left(\tau_{kk'} \middle| \frac{\lambda^2}{2}\right), \\ \lambda &\sim \text{gama}(a, b). \end{aligned} \quad (4)$$

Let $\omega = \{\omega_{kk'}\}_{k \leq k'}$ denote the vector elements of Ω , and the joint prior for ω conditional on τ and λ is:

$$\begin{aligned} p(\omega|\tau, \lambda) &= \prod_{k < k'} \frac{1}{\sqrt{2\pi\tau_{kk'}}} \exp\left(-\frac{\omega_{kk'}^2}{2\tau_{kk'}}\right) \prod_{k=1}^K \frac{\lambda}{2} \\ &\exp\left(-\frac{\lambda}{2}\omega_{kk}\right) \cdot 1_{\Omega \in M^+} \end{aligned} \quad (5)$$

Posterior samples are drawn by a block Gibbs sampling by partition Ω and $\mathbf{S}$ as follows:

$$\Omega = \begin{pmatrix} \Omega_{-kk} & \omega_k \\ \omega_k^T & \omega_{kk} \end{pmatrix}, \mathbf{S} = \begin{pmatrix} \mathbf{S}_{-kk} & \mathbf{s}_k \\ \mathbf{s}_k^T & s_{kk} \end{pmatrix}, \quad (6)$$

where $-k$ denotes the set of all indices except for k and $\Omega_{-kk}/\mathbf{S}_{-kk}$ is the submatrix of $\Omega/\mathbf{S}$ without the k -th row and column. Here $\omega_k/\mathbf{s}_k$ is the k -th row of $\Omega/\mathbf{S}$ excluding the k -th element itself. The conditional distribution of the k -th column of Ω is:

$$\begin{aligned} &p(\omega_k, \omega_{kk} | \Omega_{-kk}, \text{rest}) \\ &\propto (\omega_{kk} - \omega_k^T \Omega_{-kk}^{-1} \omega_k)^{\frac{N}{2}} \times \exp\left\{-\frac{1}{2} [\omega_k^T \mathbf{D}_\tau \omega_k + 2\mathbf{s}_k^T \omega_k + (s_{kk} + \lambda)\omega_{kk}]\right\}, \end{aligned} \quad (7)$$

where $\mathbf{D}_\tau = \text{diag}(\tau_k)$ is the diagonal matrix with diagonal elements $\tau_{(k1)}, \dots, \tau_{(k, k-1)}, \tau_{(k, k+1)}, \dots, \tau_{(kK)}$. We make a change of variables

$$(\omega_k, \omega_{kk}) \rightarrow (\beta = \omega_k, \mathbf{k} = \omega_{kk} - \omega_k^T \Omega_{-kk}^{-1} \omega_k), \quad (8)$$

then we have:

$$\begin{aligned} &p(\beta, \gamma | \Omega_{-kk}, \text{rest}) \\ &\propto \gamma^{\frac{N}{2}} \exp\left(-\frac{s_{kk} + \lambda}{2}\gamma\right) \times \exp\left(-\frac{1}{2} [\beta^T \{\mathbf{D}_\tau^{-1} + (s_{kk} + \lambda)\Omega_{-kk}^{-1}\} \beta + 2\mathbf{s}_k^T \beta]\right). \end{aligned} \quad (9)$$

This implies that:

$$p(\beta, \gamma | \Omega_{-kk}, \text{rest}) \propto N(-\mathbf{A}\mathbf{s}_k, \mathbf{A}) \text{Gamma}\left(\frac{N}{2} + 1, \frac{s_{kk} + \lambda}{2}\right), \quad (10)$$

where $\mathbf{A} = [(s_{kk} + \lambda)\Omega_{-kk}^{-1} + \mathbf{D}_\tau^{-1}]^{-1}$.

After sampling from the above conditional distributions, we can obtain $\omega_k = \beta^T$ and $\omega_{kk} = \gamma + \omega_k^T \Omega_{-kk}^{-1} \omega_k$. The latent scale parameters τ and the shrinkage parameter λ are sampled:

$$p\left(\frac{1}{\tau_{kk'}} | \text{rest}\right) \propto \text{Inv-Gaussian}\left(\sqrt{\frac{\lambda^2}{\omega_{kk'}^2}}, \lambda^2\right), \quad (11)$$

$$p(\lambda | \text{rest}) \propto \text{Gamma}\left(a + \frac{K(K+1)}{2}, b + \frac{1}{2} \sum_{k=1}^K \sum_{k'=1}^K |\omega_{kk'}|\right). \quad (12)$$

To ensure consistency with the lasso prior used in the loading part of the PCFA model, we adopted the same default hyperparameter settings in the PCFA context, with $a = 1$ and $b = 0.1$ in the Gamma distribution.



2.1.2. Graphical Horseshoe

The horseshoe prior combines local shrinkage parameters $\tau_{kk'}$ for each dimension with a global shrinkage parameter λ shared across all dimensions. Unlike the lasso, which applies uniform shrinkage, the horseshoe prior can strongly shrink near-zero coefficients while preserving large signals. Li et al. (2019) proposed the Bayesian graphical horseshoe, which assigns horseshoe priors to off-diagonal elements of the precision matrix while placing an uninformative prior on the diagonal elements. To ensure comparability with the other priors examined in this study, we instead assign an exponential prior to the diagonal elements, consistent with the specification used under the lasso framework. Considering alternative specifications by applying shrinkage only to the off-diagonal elements while leaving factor variances unregularized is a worthwhile direction for future research and could also be explored for the lasso and spike-and-slab priors, although it is not included in the present study. The prior for the inverse factor covariance matrix Ω is defined as:

$$p(\Omega|\lambda) \propto \prod_{k < k'} \left\{ \mathcal{N}(\omega_{kk'} | 0, \tau_{kk'}^2 \lambda^2) \right\} \prod_{k=1}^K \left\{ \text{Exp} \left(\omega_{kk} | \frac{\lambda^2}{2} \right) \right\} 1_{\Omega \in M^+} \quad (13)$$

The element-wise priors are specified as follows:

$$\begin{aligned} \omega_{kk} &\sim \text{Exp} \left(\omega_{kk} | \frac{\lambda^2}{2} \right), \\ \omega_{kk'}, k < k' &\sim N(0, \tau_{kk'}^2 \lambda^2), \\ \tau_{kk'}, k < k' &\sim \mathcal{C}^+(0, 1), \\ \lambda &\sim \mathcal{C}^+(0, 1) \end{aligned} \quad (14)$$

where $\mathcal{C}^+(0, 1)$ denotes half-Cauchy random variable.

The block Gibbs sampling scheme is similar to the procedure used for the graphical lasso (detailed in Appendix). To ensure coherence, we apply the same horseshoe specification to the regularized loading component in PCFA. This integrated framework enables us to investigate whether using consistent versus different regularization across loadings and factor correlations improves estimation performance. The corresponding full conditional distributions for the loading parameters are provided in Appendix B, and the comparative effectiveness of these strategies is evaluated in the simulation studies presented later.

2.1.3. Graphical Spike-and-Slab

The spike-and-slab prior is another popular Bayesian shrinkage method that allows selective shrinkage by mixing a sharply peaked “spike” component (encouraging near-zero estimates) with a broader “slab” component (permitting substantial non-zero values). Here, we introduce the graphical spike-and-slab prior (Wang, 2015) for the inverse factor covariance matrix.

The graphical SSP integrates two-component normal mixture densities for factor covariance, and exponential density for factor variance:

$$p(\Omega|\lambda) = \prod_{k < k'} \left\{ (1 - \pi) \mathcal{N}(\omega_{kk'} | 0, v_0^2) + \pi \mathcal{N}(\omega_{kk'} | 0, v_1^2) \right\} \cdot \prod_{k=1}^K \left\{ \text{Exp} \left(\omega_{kk} | \frac{\lambda^2}{2} \right) \right\} 1_{\Omega \in M^+} \quad (15)$$

The binary indicators π determine whether a correlation is drawn from the spike or the slab distribution and a common choice in the literature is $\pi = 2/(K - 1)$. However, unlike conventional graphical models that often involve a large number of edges, factor analytic models typically include a relatively small number of factors. To prevent the prior inclusion probability π from exceeding 1, we fixed $\pi = 0.3$ for models with fewer than four factors. Other hyperparameters are fixed at $v_0 = 0.02$, $v_1 = 1$ and $\lambda = 1$, following the guidelines of Wang (2015). Posterior sampling and corresponding prior for loading estimation are provided in Appendix C.

2.2. Standardization and Sampling of the Factor Covariance Matrix

Traditionally, two common identification strategies are used in factor analysis: (1) fixing factor variances to 1 (*standardized* format with factor correlation matrix), or (2) fixing one loading per factor to 1 (*non-standardized* format with factor covariance matrix). Standardization is often beneficial for improving interpretability, facilitating comparisons across parameters, and enhancing the clarity of the latent structure.

In the present framework, both the inverse-Wishart prior and the proposed regularized prior in PCFA are specified on the factor covariance matrix rather than directly on the correlation matrix. Consequently, we adopt the Metropolis–Hastings (MH)–based standardization strategy, which is commonly used in traditional Bayesian factor models, to obtain interpretable factor correlations (Chen, 2022). At each iteration, samples are drawn from the unconstrained posterior of the factor covariance matrix, and the proposed draw is accepted or rejected based on the determinant of the implied correlation matrix relative to the previous draw. Our preliminary analyses indicate that applying MH-based standardization within PCFAwRFC yields stable estimation. In addition, the MH strategy is consistent with the inverse-Wishart prior specification for factor correlations in traditional PCFA, which facilitates comparability across prior families. For these reasons, and to ensure both interpretability and comparability of the estimated latent factors, all simulation studies in this work adopt MH-based standardization.

2.3. Partially Confirmatory Factor Analysis with Partial Regularized Factor Correlation (PCFAwPRFC)

Zhang and Chen (2024) have extended PCFA to accommodate complex latent structures for special effects involving multiple general (construct/trait) factors and special (group/method) factors. Factors in Equation (1) are separated into two parts, K_g general factors and K_s special factors:

$$y_i = \mu + \Lambda_g f_{gi} + \Lambda_s f_{si} + \epsilon_i, i = 1, 2, \dots, N \quad (16)$$

Factor covariance matrix Φ is partitioned into Φ_g , a $K_g \times K_g$ covariance matrix for the general factors, and Φ_s , a

$K_s \times K_s$ covariance matrix for the special factors accordingly:

$$\Phi = \begin{bmatrix} \Phi_g & 0 \\ 0 & \Phi_s \end{bmatrix} \quad (17)$$

In the literature, Φ_s is typically restricted to a diagonal matrix for simplicity and computational convenience, thereby assuming that special effects are independent of each other. While this orthogonality constraint facilitates estimation, it can be overly restrictive and may obscure meaningful residual relationships among special factors.

Building on the PCFAwRFC framework, we propose a partially regularized factor correlation approach, PCFAwPRFC, to better accommodate complex factor structures. PCFAwPRFC relaxes the strict orthogonality of special factors, allowing for small but potentially important correlations to be estimated, while still encouraging them to remain nearly independent through sparsity-inducing priors. Under this approach, Φ_g is estimated without sparsity

constraints using inv-Wishart distribution to capture potentially rich associations among general factors. Φ_s is modeled using a sparse covariance structure with regularization, striking a balance between the simplicity of orthogonal assumptions and the flexibility to recover meaningful inter-factor dependencies.

Figure 1 illustrates an example of the PCFAwPRFC framework. Solid lines represent substantive knowledge: for factor loadings, they indicate zero-fixed or freely estimated loadings, and for factor correlations (general factors), they indicate parameters sampled under the inverse-Wishart prior. Dashed lines represent regularized components: for factor loadings, they indicate unspecified loadings estimated through regularization, and for factor correlations (special factors), they indicate regularized correlations. Both regularized components can adopt different shrinkage priors, including the lasso, horseshoe, or spike-and-slab priors.

Figure 2 presents the full framework of factor analysis under PCFA paradigm, highlighting the extensions introduced in this study with blue-highlighted boxes. All the three main components—factor loadings (Λ), factor correlations (Φ), and residual covariances (Ψ)—can be estimated with or without regularization depending on user-specified model structures. Dashed arrows indicate pathways where regularization can be applied. This integrated framework allows flexible prior selection and parameter regularization across all three components, enabling improved estimation for both simple and complex latent structures.

3. Simulation and Empirical Studies

This article presents a series of studies, comprising four simulation scenarios and one real-data application to evaluate the performance and practical utility of the proposed methods. These studies are designed to progressively assess the methods across varying levels of complexity in latent structures—ranging from low-dimensional models to complex, high-dimensional settings with intricate sparsity patterns in the factor correlation matrix. Estimation approaches

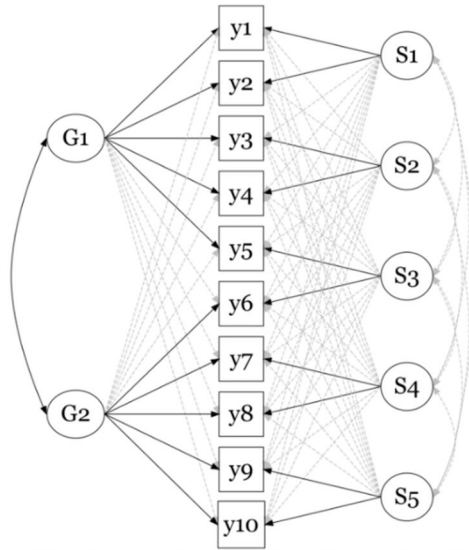


Figure 1. Illustration of the PCFAwPRFC framework, with dashed lines indicating regularized estimates.

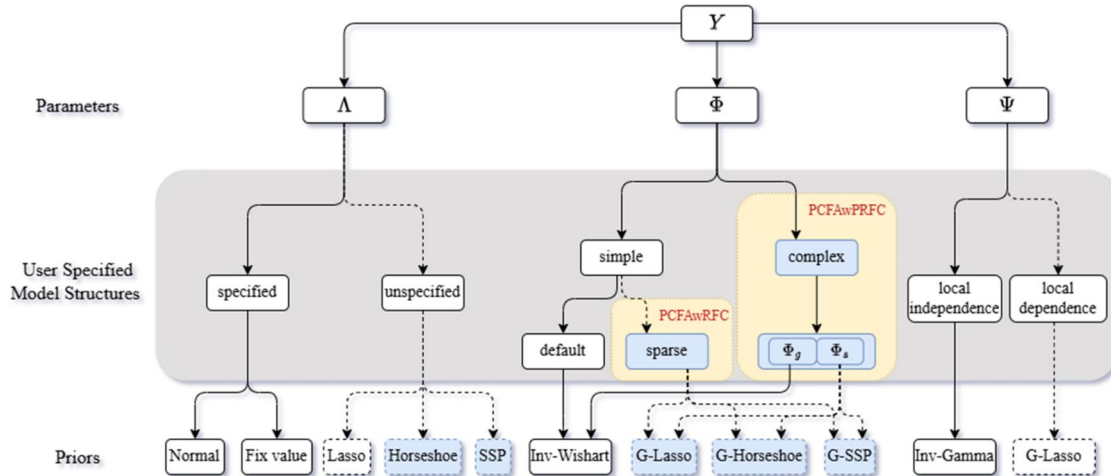


Figure 2. PCFA-based framework with extensions for regularized and partial regularized factor correlation estimation. The dashed lines depict regularized components.



for comparison include the traditional Inverse-Wishart prior (IW); PCFAwRFC framework with three priors (lasso, horseshoe, SSP) and the PCFAwPRFC framework for complex models with different priors in special factor correlation (Par-Lasso, Par-Horseshoe, Par-SSP). Here, “Par-” refers to partial regularization of special factor correlations, under which different priors are applied within the PCFAwPRFC framework.

The first two simulation studies evaluated simple factor models under fully correlated and sparse factor-correlation structures, comparing the traditional IW estimator with the proposed PCFAwRFC framework for estimating the factor-correlation matrix. In both studies, the loading structure was fully specified and estimated without regularization. The third study investigated the interaction between loading regularization and factor-correlation regularization. It used the same factor-correlation estimators as in Study 1 and 2 while applying regularization to the unspecified loadings under partially specified loading structures. The fourth study extended the evaluation to complex models with multiple general and special factors, estimating factor correlations using the PCFAwPRFC framework and regularizing loadings according to the designed Q-specifications.

Each simulation condition was replicated 100 times using a Markov Chain Monte Carlo (MCMC) algorithm. For Bayesian estimation, we performed 10,000 iterations following a burn-in period of 10,000 iterations. Convergence of MCMC chains was assessed using the estimated potential scale reduction (EPSR), with values below 1.1 considered indicative of satisfactory convergence. All simulation studies were conducted in the R computing environment (R Core Team, 2021) and extend the original PCFA implementation from the *LAWBL* package (Chen, 2022).

Model performance was evaluated using several widely accepted metrics, including BIAS, root mean squared error (RMSE), standard error (SE), and the percentage of significant estimates at the .05 level (SIG%) based on the concept of highest posterior density (HPD) intervals. The SIG% criterion reflects a key strength of Bayesian regularization because it is derived from credible intervals obtained from the posterior distributions (Casella et al., 2010). To evaluate performance in identifying sparse factor correlations, we further examined false positive detection for true zero correlations under two identification criteria: the significance rule (SIG%) and a threshold-based rule, which is commonly used in point-estimate-based analyses. Other criteria have been proposed in the literature, such as using marginal inclusion probabilities under spike-and-slab priors for variable selection (e.g., Bainter et al., 2023; Barbieri & Berger, 2004). These approaches may be valuable for further investigation, but they are not considered in the present study in order to maintain comparability across priors. Details of each simulation design are provided in the respective subsections. The final empirical example demonstrates the real-world applicability of partial regularized factor correlation in modeling complex psychological data.

3.1. Study 1: Simple CFA with Fully Correlated Factor Structure

The first study evaluates the performance of PCFAwRFC under a conventional confirmatory factor analysis (CFA) setting with a low-dimensional, fully correlated factor structure. In this CFA configuration, only the factor correlation matrix is subjected to regularization, while all other parameters remain freely estimated and all zero loadings are fixed to zero. This baseline scenario is designed to assess the reliability and precision of parameter estimation under typical modeling conditions.

We varied the number of latent factors ($K = 3, 5$), sample size ($N = 250, 500, 1000$), and factor correlation ($\phi_{kk'} = .3, .5$). Each factor measured by 6 items and accompanied by 2 cross-loadings. True primary factor loadings (λ_m) were fixed at 0.7, while cross-loadings (λ_c) were set at 0.3.

Figure 3 summarizes parameter recovery for Study 1, with complete results reported in Appendix D. Across all design conditions, estimation quality was broadly stable with respect to the number of factors and the level of factor correlation, and larger samples yielded modest improvements over smaller samples for all priors. The spike-and-slab prior showed slightly degraded performance in terms of BIAS and RMSE at the smaller sample size, although its accuracy remained acceptable overall. By contrast, the remaining priors exhibited comparable and satisfactory recovery across metrics. Taken together, these findings indicate that, under conventional simple factor-model settings without sparsity in the factor-correlation structure, both traditional and regularized Bayesian approaches recover the underlying parameters accurately, although the spike-and-slab prior is not recommended when sample sizes are small. From a computational perspective, the traditional inverse-Wishart prior remains the more efficient and practical choice in such settings.

3.2. Study 2: High-Dimensional Sparse Factor Correlation Structure

The second study evaluates the performance of PCFAwRFC in high-dimensional settings where the true factor correlation matrix is sparse. This scenario is designed to reflect psychological measurement settings in which most latent constructs are independent, with only a few exhibiting meaningful inter-factor associations.

We considered two levels of sample size ($N = 250, 1000$); factor dimensionality ($K = 5, 10$), and three levels of sparsity (Sparsity = 1, 3, 5). The sparsity of the factor correlation matrix was manipulated by varying the number of nonzero off-diagonal correlations. For example, a sparsity level of 1 in a 5-factor ($K = 5$) population indicates that only one of $5 \times 4/2 = 10$ unique factor correlations is nonzero, while the remaining 9 are set to zero. We fix the nonzero correlation at .5. All other parameters and estimation structure were the same as in Study 1.

Figure 4 presents the parameter recovery results with detailed numerical outcomes provided in the Appendix E. Except for factor correlations, the estimation accuracy for factor loadings and residual variances was generally satisfactory and comparable across all four correlation estimators.

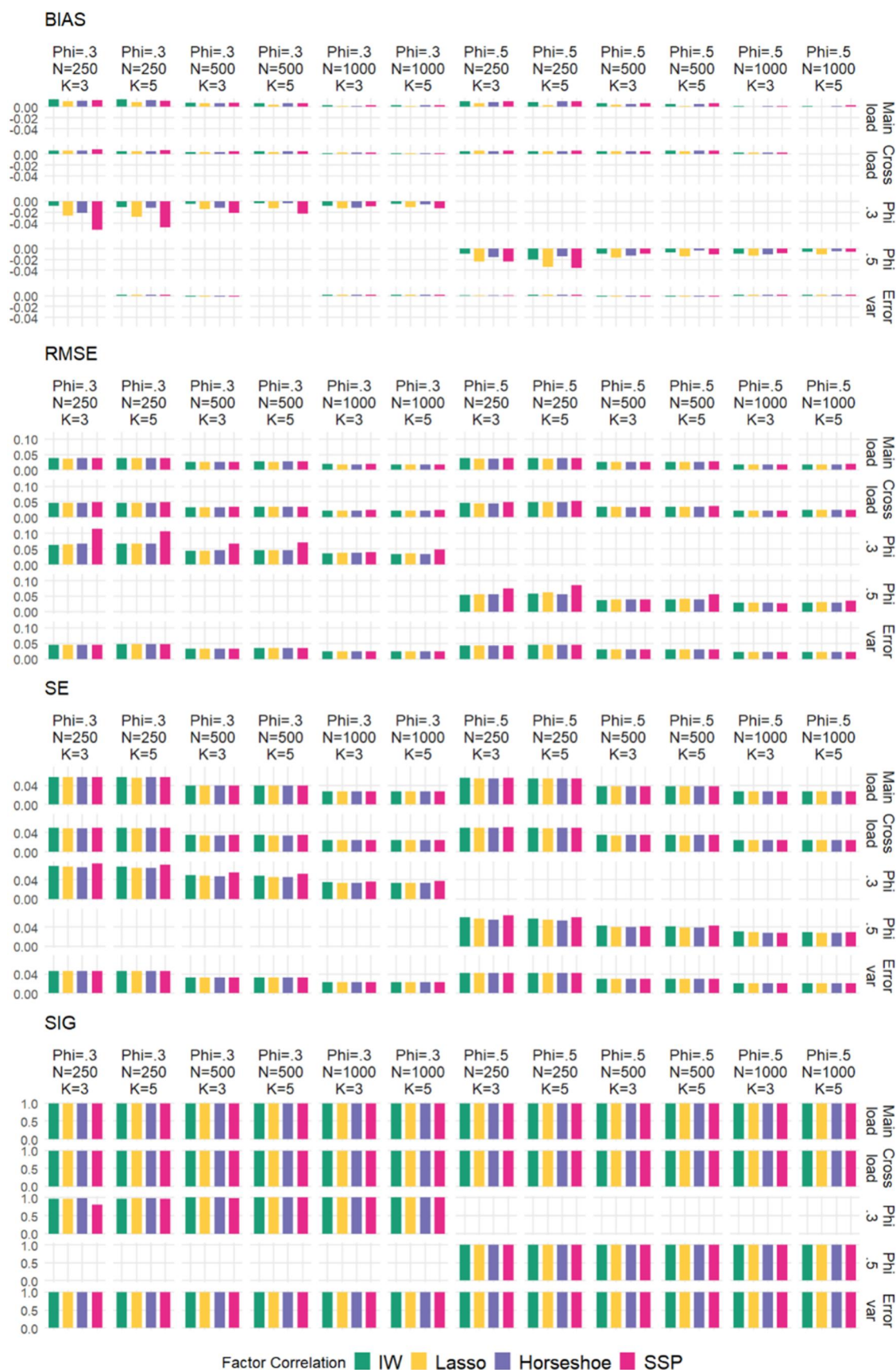


Figure 3. Parameter recovery results for Study 1.

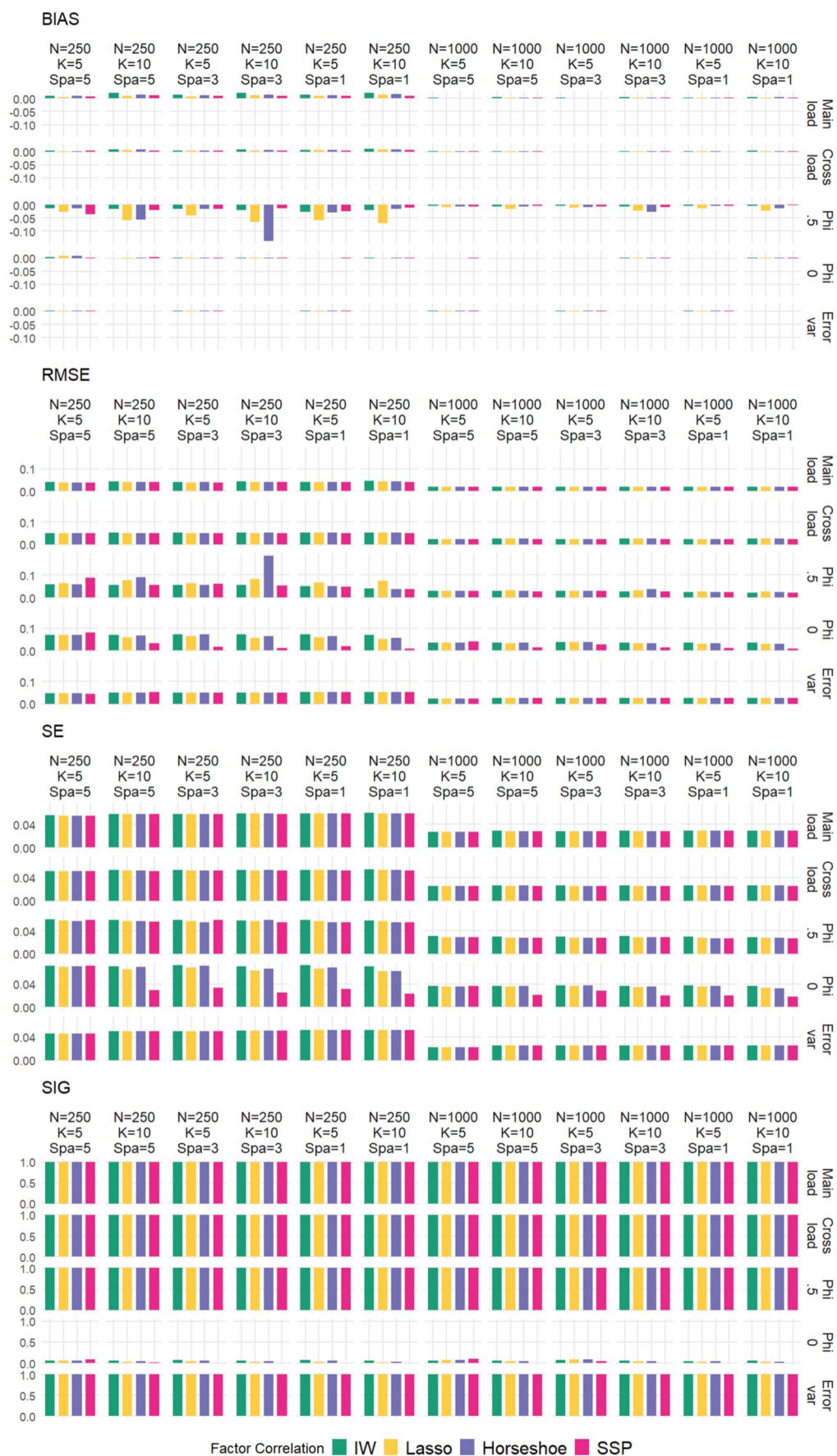


Figure 4. Parameter recovery results for study 2.

Similar to Study 1, larger sample sizes yielded better recovery performance than smaller ones. For factor correlations, only one condition (small sample size with $K=10$ and sparse = 3) showed moderate RMSE (around 0.1) for non-zero correlations when using the horseshoe prior.

Although the average recovery of zero correlations appeared broadly acceptable across methods, this aggregate view conceals notable differences as the number of true zero correlations increases. To further evaluate the performance of each method under varying levels of sparsity, we examined false positive indicators. Figure 5 presents false positive metrics under two identification criteria: (1) significance at the 0.05 level (95% SIG; top panel), and (2) an absolute estimated correlation exceeding 0.1 (Threshold = 0.1; bottom panel). Overall, all regularized estimators outperformed the traditional inverse-Wishart (IW) prior in high-dimensional and sparse settings, whereas the IW prior performed slightly better in low-dimensional and less sparse conditions. Among the regularized methods, the spike-and-slab prior exhibited this trend most prominently showing a relatively high false positive rate under the moderate condition ($K=5$, sparse = 5) and achieving the lowest rates as dimension and sparsity increased. The lasso and horseshoe methods produced comparable results, with lasso showing slightly greater stability. The significance-based criterion (95% SIG) appeared more robust across different sample sizes than the threshold-based criterion.

3.3. Study 3: Impact of Regularized Loadings on Factor Correlation Estimation

This study evaluates the interaction between loading and factor correlation regularization to examine how different prior specifications influence estimation performance when

both components are regularized simultaneously. By extending the unspecified loading prior from the lasso to the horseshoe and spike-and-slab priors, both the loading and factor correlation structures can be modeled under three alternative priors for comprehensive comparison.

Preliminary analyses revealed convergence issues when inconsistent priors were applied to different model components. For example, using a lasso prior for the loading matrix and a spike-and-slab prior for the factor-correlation matrix resulted in non-convergence in more than half of the replications. One plausible explanation is that different priors impose shrinkage through different mechanisms and at different rates across parameter spaces, which can create mismatched regularization pressures during joint estimation. When the loading matrix and the factor-correlation matrix are regularized in incompatible ways, the sampler may have difficulty balancing cross-loading shrinkage against latent-correlation shrinkage, leading to unstable posterior exploration.

Therefore, in this study, we adopted consistent priors for both the loading and factor-correlation matrices when applying the PCFAwRFC framework. For comparison, the traditional IW estimator for factor correlations was also combined with three different loading priors (lasso, horseshoe, and SSP), resulting in six total prior combinations: Lasso-IW, Horseshoe-IW, SSP-IW, Lasso-Lasso, Horseshoe-Horseshoe, and SSP-SSP.

We assessed each method under two sample sizes ($N = 250, 1000$) and two factor-correlation structures: sparse (sparsity = 1) and non-sparse. Simulated data were generated from a five-factor model with six indicators per factor and two cross-loadings per factor. Primary factor loadings were fixed at 0.7, cross-loadings at 0.3, and nonzero correlations under the sparse condition were set to 0.3. To guide



Figure 5. False positive detection for zero correlations in study 2.

Note. Top: significance at the 0.05 level. Bottom: absolute estimated correlation exceeding 0.1. Three metrics are reported: false positive rate (FPR), the average estimated value of falsely detected nonzero correlations (FPest), and the average standard error of these false positives (FPse)



regularization of the loading matrix, we employed two configurations of the special matrix Q : “main” specified design, in which all main loadings are freely estimated while all others are regularized; and “halfmain” specified, in which half of the main loadings are specified. The true loading matrix and the two Q specification matrices are provided in the Appendix.

Figures 6 and 7 present the recovery results for the non-sparse and sparse factor-correlation structures, with detailed numerical results provided in the Appendix F. Figure 8 summarizes the false positive detection of zero factor correlations. Under the non-sparse factor-correlation structure, all methods achieved acceptable parameter recovery. In this setting, regularizing the factor-correlation matrix does not provide a clear advantage over the traditional IW prior, although differences emerged in the choice of loading priors: both the horseshoe and spike-and-slab (SSP) priors yielded notably better recovery than the lasso prior. Thus, while factor-correlation regularization is not essential in non-sparse models, the selection of loading priors still meaningfully influences estimation quality.

Under the sparse factor-correlation structure, most conditions showed acceptable parameter recovery, and the traditional IW estimator performed comparably to the PCFAwRFC framework in many cases. A notable exception was the spike-and-slab prior in the small-sample condition, which produced moderate RMSEs for nonzero correlations and yielded lower significance rates (SIG). In contrast, the horseshoe prior showed better recovery in small samples. For zero-correlation detection, the PCFAwRFC framework consistently achieved lower false positive rates than the traditional IW prior, with the spike-and-slab prior providing the strongest specificity, and the lasso also showing good zero-correlation detection under the significance rule. Larger sample sizes naturally improved zero-correlation detection, particularly under the threshold-based criterion.

An interesting pattern emerged for the traditional IW estimator: under the significance criterion, the half-main loading specification produced fewer false detections of zero correlations than the fully specified main structure, a difference not observed under the threshold criterion. This occurs because the main specification leaves all primary loadings unregularized, increasing sampling noise that inflates small correlations, whereas the IW prior provides no shrinkage to counteract this noise. The half-main design, by shrinking more loadings, reduces noise at the source and therefore improves zero-correlation detection. Overall, the PCFAwRFC framework is recommended for more reliable identification of sparse factor correlations.

3.4. Study 4: Evaluating PCFAwPRFC in Complex Factor Structures

The fourth simulation study evaluates the performance of the proposed PCFAwPRFC framework under a complex factor structure involving multiple general and special factors. In this framework, factor correlations are partially regularized: the correlations among general factors are estimated

using the traditional IW prior, whereas the correlations among special factors are regularized using one of three priors. For convenience, we refer to these partially regularized specifications as Par-Lasso, Par-Horseshoe, and Par-SSP. Specifically, the simulated structure includes two general factors and ten special factors, collectively measured by a total of 40 items, with two sample sizes ($N=250, 1000$). The general factor loadings are drawn uniformly from the range 0.6 to 0.8, and special factor loadings from 0.3 to 0.5. The two general factors are set to be correlated at 0.5. In contrast, the correlation matrix among the special factors is assumed to be sparse with two sparsity levels. For loading estimation, we evaluated two Q -matrix specifications, denoted as “GparSpar” and “GparScfa.” In “GparSpar,” both general and special factors are partially specified with all true nonzero loadings freely estimated and all other loadings unspecified (i.e., regularized). In “GparScfa,” the general factor remains partially specified in the same way, while the special factors follow a strict CFA specification in which all true zero loadings are fixed to zero. In practice, general factors typically represent broad latent traits, whereas special factors reflect method or format effects that are well-defined by the study design. All unspecified loadings are estimated under the same prior used for the partially regularized factor-correlation structure to maintain consistency. The true loading matrix and corresponding design matrix Q are provided in the Appendix G.

Figures 9 and 10 present the results for parameter recovery and false positive detection of true zero special-factor correlations (detailed numerical results are provided in the Appendix G). Overall, all methods show satisfactory estimation performance for factor loadings, general factor correlations and residuals, with little sensitivity to the sparsity level in this study. For zero-correlation detection, the significance-based criterion (SIG) yields results that align more closely with the true model structure compared to the threshold-based rule.

For Lasso and SSP priors, the number of unspecified loadings influences how strongly the model shrinks the factor-correlation matrix. Under GparSpar, the Lasso and SSP priors shrink more special-factor correlations toward zero. This results in better detection of true zero correlations, but at the cost of underestimating substantial correlations, especially in small samples.

In contrast, the horseshoe prior is less sensitive to sample size and Q -matrix specification, but exhibits consistently higher false positive rates when regularizing special-factor correlations. This pattern may arise from its global-local shrinkage mechanism, which automatically adapts to the amount of uncertainty introduced by unspecified loadings. This adaptive behavior stabilizes estimation whether many or few loadings are regularized. However, its heavy-tailed global-local structure might also allow some spurious correlations to escape shrinkage when they are interpreted as potential signals, contributing to higher false positive rates.

Taken together, these findings suggest that PCFAwPRFC is capable of recovering complex structures, although differences in estimation accuracy across priors remain modest.

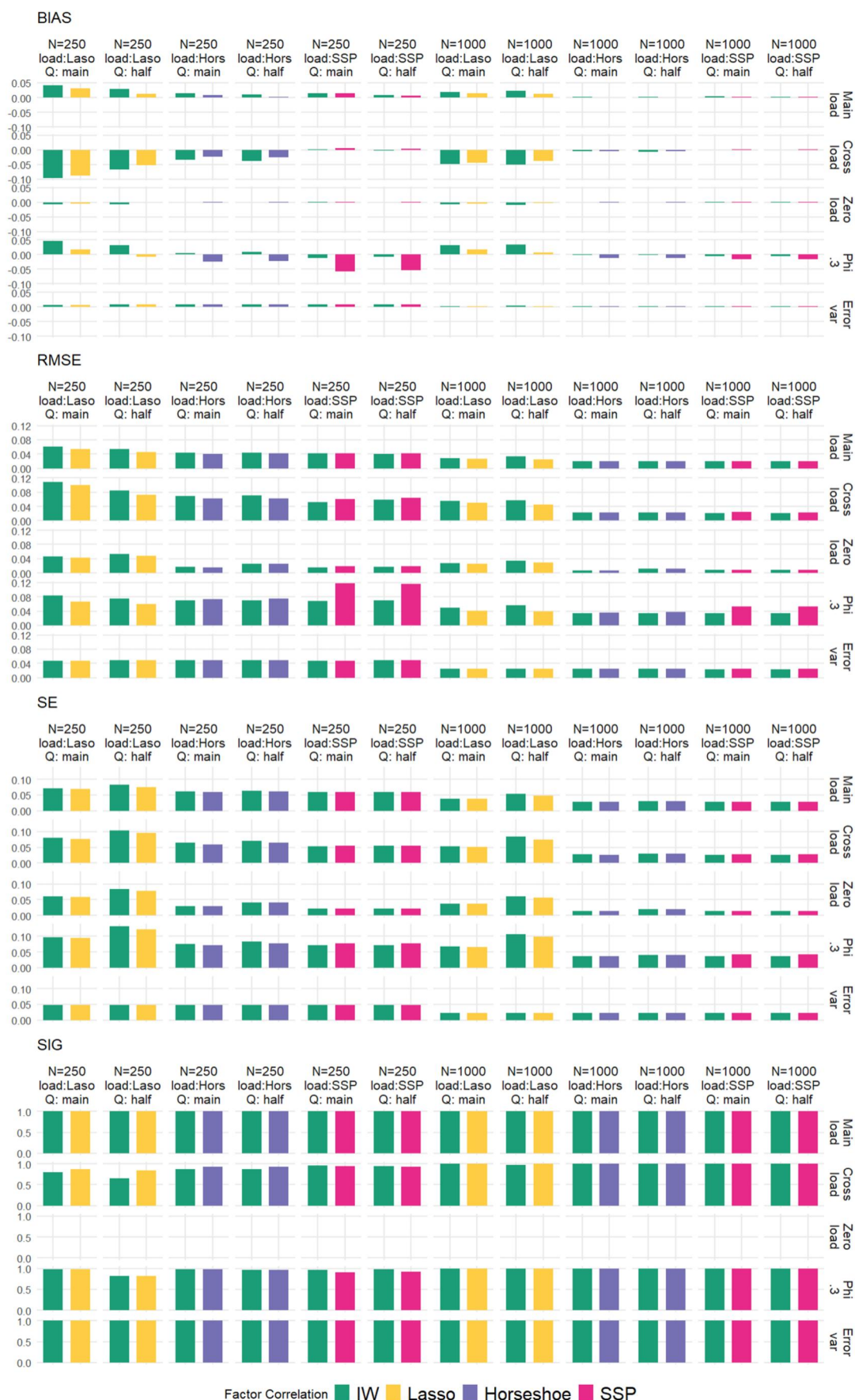


Figure 6. Parameter recovery results for study 3 under non-sparse factor structures.

Note. "load: Lasso/Hors/SSP" indicates that unspecified loadings were estimated using the lasso, horseshoe, or spike-and-slab priors, respectively. "Q: main/half" refers to loading specifications in which all primary loadings (main) or half of the primary loadings (half) were fixed according to the theoretical structure

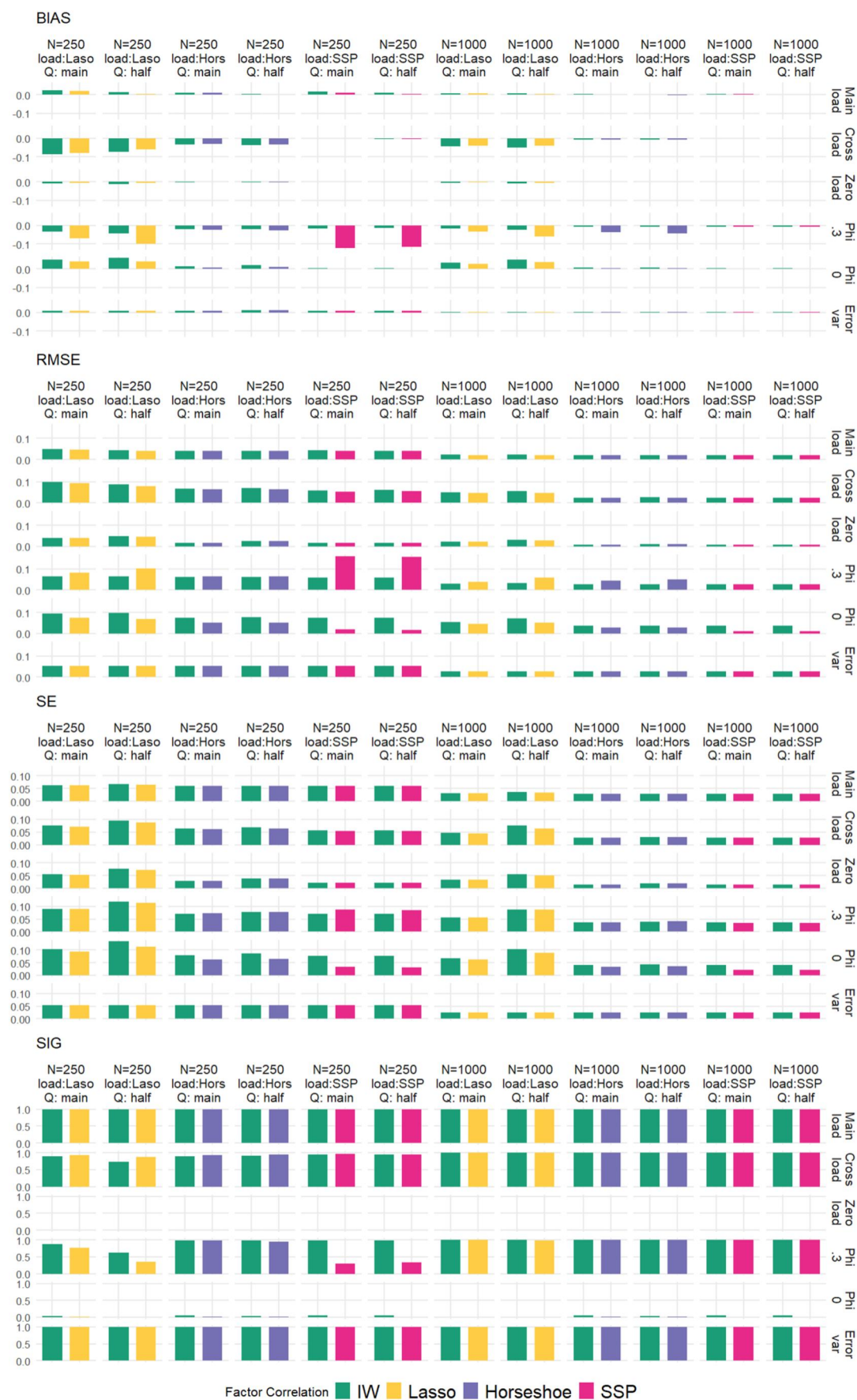


Figure 7. Parameter recovery results for study 3 under sparse factor structures.

Note. "load: Lasso/Hors/SSP" indicates that unspecified loadings were estimated using the lasso, horseshoe, or spike-and-slab priors, respectively. "Q: main/half" refers to loading specifications in which all primary loadings (main) or half of the primary loadings (half) were fixed according to the theoretical structure

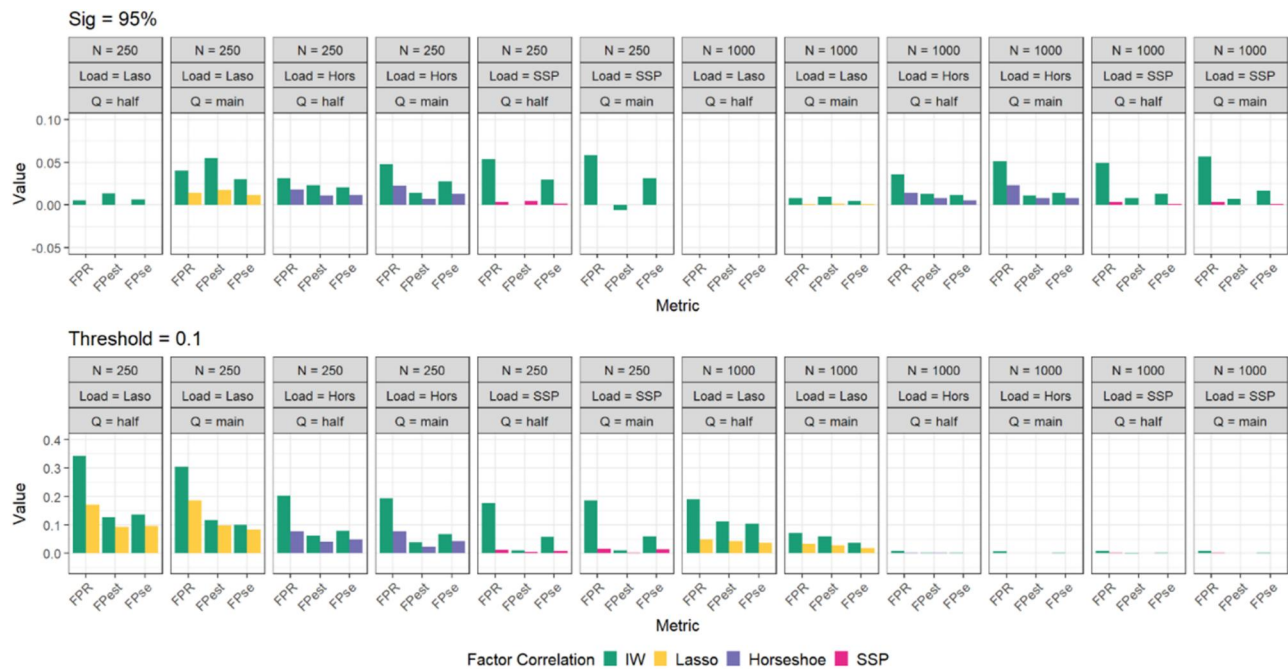


Figure 8. False positive detection for zero correlations in study 3.

Note. Top: significance at the 0.05 level. Bottom: absolute estimated correlation exceeding 0.1. Three metrics are reported: false positive rate (FPR), the average estimated value of falsely detected nonzero correlations (FPest), and the average standard error of these false positives (FPse). “load: Lasso/Hors/SSP” indicates that unspecified loadings were estimated using the lasso, horseshoe, or spike-and-slab priors, respectively. “Q: main/half” refers to loading specifications in which all primary loadings (main) or half of the primary loadings (half) were fixed according to the theoretical structure

The regularization of complex structure reflects a tradeoff between accurate estimation of non-zero correlations and reliable detection of zero correlations, without indicating a uniformly superior choice. A larger sample size is generally beneficial for stable estimation. When sample sizes are limited, the horseshoe prior may offer advantages in terms of stability and recovery of substantial correlations. Under GparScfa, substantial correlations are recovered more accurately, but zero-correlation detection becomes slightly worse. Finally, the significance-based criterion provides more reliable zero-correlation detection than the threshold-based rule.

3.5. Study 5: Real-Data Applications (PID-5-SF)

To evaluate the practical utility of the proposed partially regularized factor correlation framework, we applied it to an empirical dataset from the personality inventory DSM-5 Short Form (PID-5-SF, Jiménez et al., 2023; Roskam et al., 2015). The dataset consists of responses from 2,532 participants on a French personality inventory comprising 5 general factors and 15 special factors¹, measured by a total of 60 items.

In previous literature, researchers either restricted the special factors to be orthogonal, or targeted the entire correlation matrix toward zero except the general factor correlations. These assumptions may oversimplify the true latent structure and

potentially ignore subtle interdependencies among special factors. PCFAwPRFC addresses this limitation by flexibly modeling the general and special factor blocks separately. Specifically, we estimate the general factor correlations freely, while imposing a sparse prior (via lasso/horseshoe/SSP) on the special factor correlation block, thereby aligning with the assumption of near orthogonality without enforcing it strictly, and restricting independence between general and special factors. This formulation allows more nuanced recovery of factor relationships while preserving theoretical expectations.

To maintain consistency with simulation studies, we adopted the same prior settings for both the loading and factor-correlation components and applied the two Q-matrix specifications used in Study 4, with all theory-based loadings specified. To evaluate the practical utility of the estimated models, we re-estimated standard CFA models in lavaan package (Rosseel, 2012), using model structures determined by two selection criteria examined in the simulation studies (SIG and threshold = 0.2).

Model fits are summarized in Table 1. All regularized methods achieve acceptable model fit and consistently outperform the theory-based model, indicating that the proposed PCFAwPRFC substantially improves model fit and yields more informative representations of the latent structure. The empirical findings are consistent with the simulation results in several aspects, providing additional insights into practical model selection and interpretation.

First, the significance-based criterion (SIG) consistently yields more reliable zero-correlation detection than the threshold-based rule, aligning with the simulation findings. Second, the sparsity patterns across priors closely mirror those observed in the simulation. The general loading

¹5 general factors and 15 special factors include: Negative Affect (G1), Emotional Lability (S1), Anxiousness (S2) and Separation Insecurity (S3); Detachment (G2), Withdrawal (S4), Anhedonia (S5) and Intimacy Avoidance (S6); Antagonism (G3), Manipulativeness (S7), Deceitfulness (S8) and Grandiosity (S9); Disinhibition (G4), Irresponsibility (S10), Impulsivity (S11) and Distractibility (S12); Psychoticism (G5), Unusual Beliefs (S13), Eccentricity (S14) and Perceptual Dysregulation (S15).

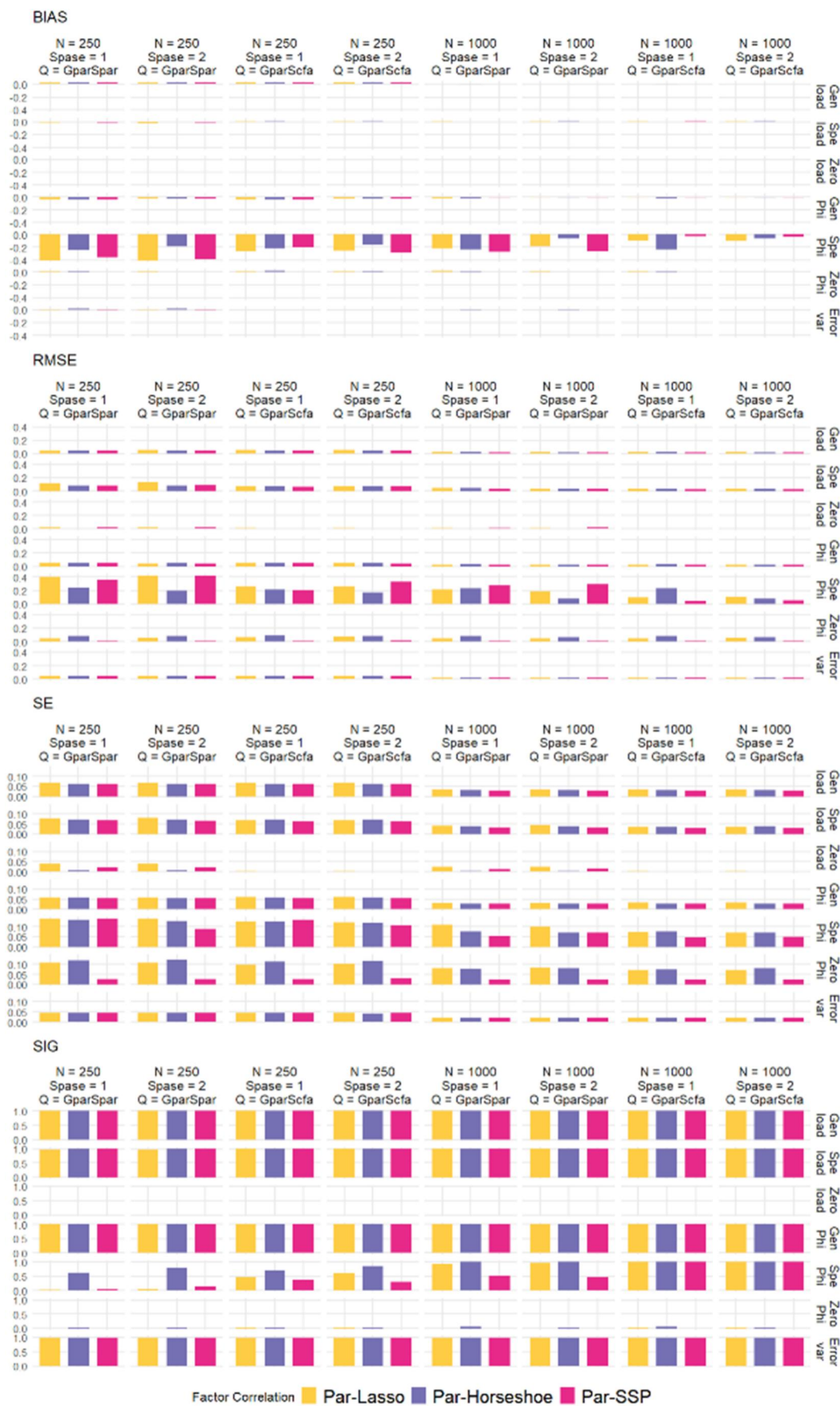


Figure 9. Parameter recovery results for study 4.

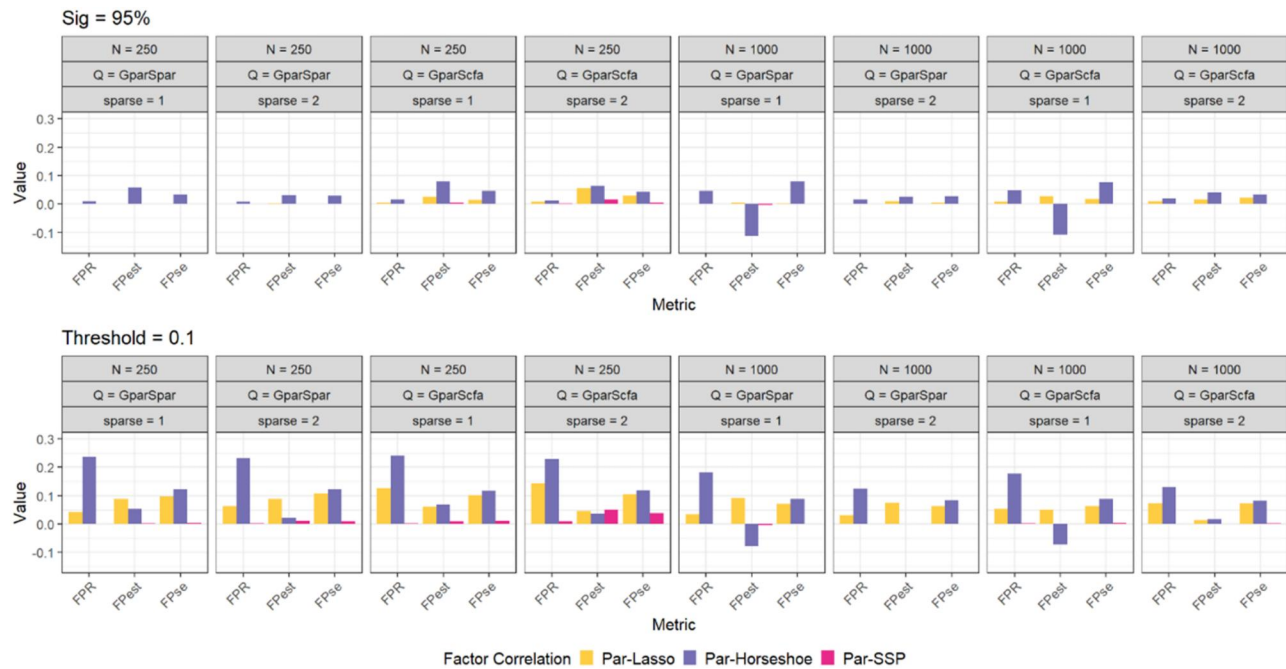


Figure 10. False positive detection for zero factor correlations in Study 4.

Note. Top: significance at the 0.05 level. Bottom: absolute estimated correlation exceeding 0.1. Three metrics are reported: false positive rate (FPR), the average estimated value of falsely detected nonzero correlations (FPest), and the average standard error of these false positives (FPse)

Table 1. Model fit summary of PID-5-SF data.

Model	Q	Criteria	RMSE	CI	CFI	TLI	SRMR	BIC	df	Chi2
Theory			.034	[.033,.035]	.920	.913	.053	317807.1	1640	6537
Lasso	GparSpar	0.2	.032	[.031,.033]	.932	.926	.054	317057.7	1642	5803
Horseshoe	GparSpar	0.2	.030	[.029,.031]	.937	.932	.049	316753.3	1638	5467
SSP	GparSpar	0.2	.030	[.029,.031]	.937	.932	.049	316753.3	1638	5467
Lasso	GparScfa	0.2	.033	[.032,.034]	.927	.921	.057	317395.5	1637	6102
Horseshoe	GparScfa	0.2	.030	[.029,.031]	.940	.935	.059	316665.1	1623	5262
SSP	GparScfa	0.2	.032	[.031,.033]	.931	.925	.060	317211.1	1625	5823
Lasso	GparSpar	SIG	.025	[.025,.026]	.957	.952	.036	315885.6	1586	4192
Horseshoe	GparSpar	SIG	.025	[.024,.026]	.959	.955	.037	315756.1	1586	4063
SSP	GparSpar	SIG	.022	[.021,.023]	.968	.963	.030	315503.0	1548	3512
Lasso	GparScfa	SIG	.024	[.023,.025]	.964	.958	.028	315759.1	1547	3760
Horseshoe	GparScfa	SIG	.024	[.023,.025]	.962	.956	.028	315832.5	1554	3888
SSP	GparScfa	SIG	.025	[.024,.026]	.959	.954	.037	315902.4	1563	4029

structures are similar across priors, while the estimated factor correlations differ substantially in sparsity (see Appendix H for details). The SSP produces markedly sparser factor structures, whereas the lasso and horseshoe priors retain more substantial correlations (Figure 11).

Third, the influence of Q-matrix specification on model fit is also consistent with the simulation results. The horseshoe prior shows relatively stable performance across Q specifications, whereas the other two priors exhibit greater variation, although these differences are generally modest. Meanwhile, the SSP tends to perform better with less restrictive loading specifications (e.g., GparSpar), whereas the lasso and horseshoe priors achieve better fit under more certain structures (e.g., GparScfa).

Overall, these results suggest that the proposed PCFAwPRFC framework provides a flexible approach that allows researchers to tailor model specification to their substantive goals, balancing interpretability, parsimony, and structural richness.

4. Discussion

This study introduces regularization of the factor correlation matrix, using graphical lasso, horseshoe and spike-and-slab priors, to overcome the limitations of the traditional inverse-Wishart prior in recovering sparse or partially sparse latent structures. Within a Partially Confirmatory Factor Analysis framework, our estimator (PCFAwRFC) extends PCFA in two key directions. First, it expands the range of parameter matrices subject to regularization. Whereas PCFA traditionally focuses on regularizing factor loadings via lasso and, optionally, residual covariances via graphical lasso, our approach also allows regularization of the factor correlation matrix. This creates a unified framework in which loadings, residual covariances, and factor correlations can all be flexibly estimated under sparsity-inducing priors.

Second, we expand the priors available for factor loadings from the Bayesian lasso to horseshoe and spike-and-slab. This extension allows us to compare the three different Bayesian shrinkage families both within a parameter type and across

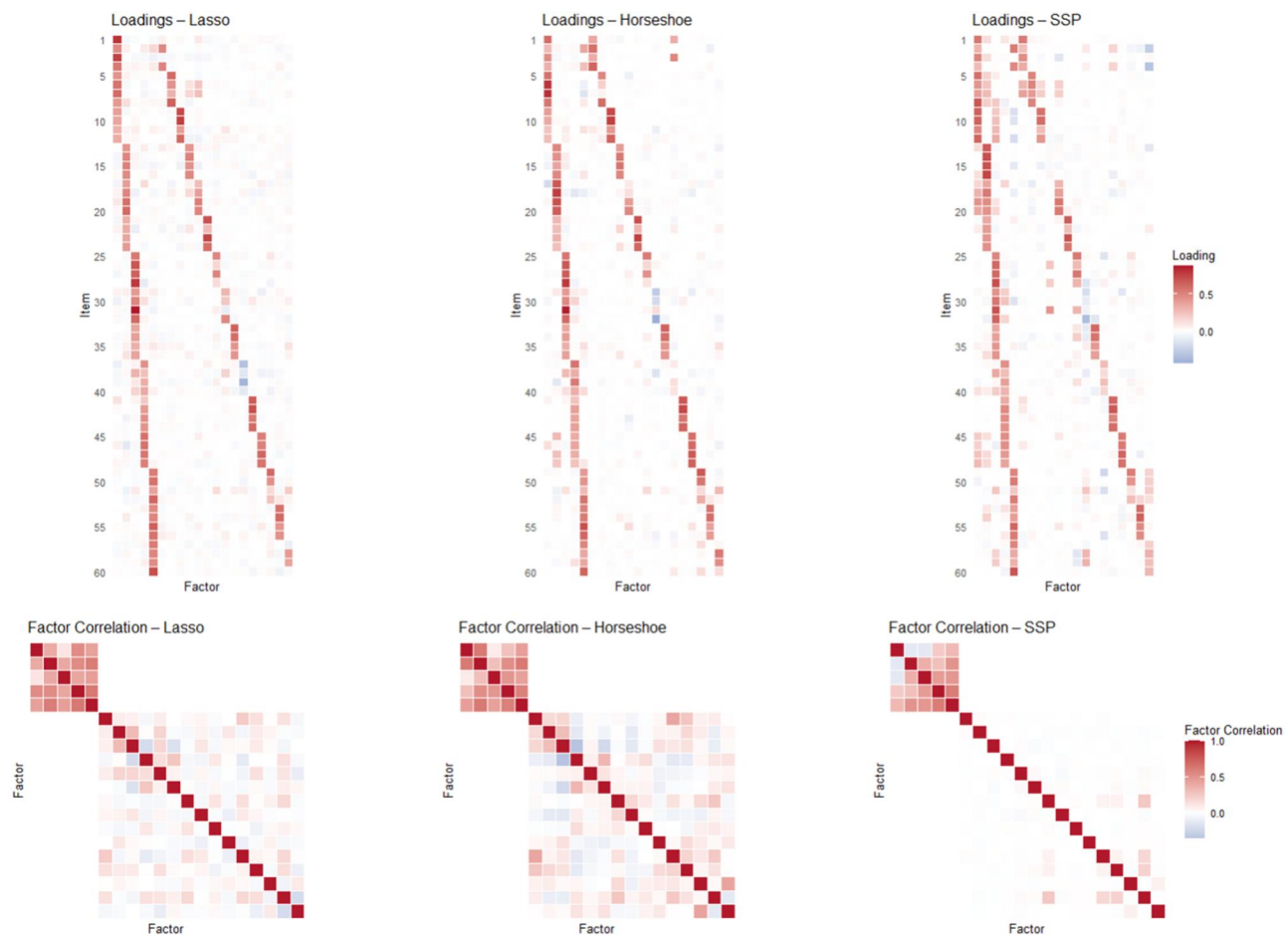


Figure 11. Heatmap of loading and factor correlation estimation for PID-5-SF data.

parameter spaces (e.g., loadings and factor correlations) under different prior combinations. While our current implementation focuses on loadings and factor correlations, the same framework can be further extended to residual covariance priors, thereby enabling spike-and-slab or other flexible shrinkage priors to be applied to all major parameter spaces in the model. This unified approach provides a coherent basis for systematically studying how prior selection and combination influence recovery, parsimony, and interpretability under diverse structural conditions in Bayesian regularized factor analysis.

A particularly relevant application of regularized factor correlation is in complex latent structures such as multi-trait-multimethod and extended bifactor models, where theoretical constraints often allow general factors to correlate while restricting method or special factors toward orthogonality. By combining inverse-Wishart prior for general-factor correlations with sparse regularization for special-factor correlations, our extension, PCFAwPRFC, accommodates such partial correlation patterns while retaining interpretability.

Across the simulation studies, Bayesian graphical modeling of factor correlations showed clear improvements over the inverse-Wishart prior in sparse structures. Under non-sparse and simple factor models, the traditional IW prior remains a practical choice due to its computational efficiency and comparable accuracy. We also found that when

IW is used for the factor-correlation structure, employing horseshoe or spike-and-slab priors for the unspecified loadings further improves estimation relative to the original PCFA framework, which relied solely on lasso loading prior.

Among the priors in PCFAwRFC and PCFAwPRFC framework, lasso prior demonstrated moderate performance, offering stable estimation across a variety of conditions but without clear superiority in any particular scenario. Its consistency and lack of extreme behavior make it a reasonable and conservative choice for preliminary analyses or for situations where a more cautious regularization approach is desired. The spike-and-slab prior, in contrast, showed greater sensitivity to small sample sizes and produced fewer stable estimates in those settings. The horseshoe achieved stable recovery across different sample sizes but tended to produce relatively higher false positive rates, especially in large-sample, complex models. In evaluating model identification rules, significance testing proved more effective than threshold-based criteria. Empirically, permitting partial correlations among special factors improved overall fit and uncovered nuanced dependencies beyond a priori expectations, offering richer and more interpretable insights into psychological constructs.

The present study focuses on graphical-model regularization via the inverse covariance (precision) matrix, whereas another promising direction is to impose sparsity directly

on the factor correlation matrix itself. For example, Bien and Tibshirani (2011) proposed penalizing individual covariance entries, and Barnard et al. (2000) introduced a decomposition of covariance matrices that allows placing shrinkage priors directly on correlations, providing a flexible alternative to inverse-Wishart prior. Extending PCFA to incorporate direct covariance- or correlation-level sparsity may provide an alternative framework for identifying interpretable factor structures and represents a valuable avenue for future research.

While our evaluation covered a broad set of structural conditions, many plausible combinations of priors, parameterizations, and sparsity patterns remain for future work. Future work could also explore sensitivity analyses for hyperparameters to better understand the interplay between regularization strength, model complexity, and convergence behavior. Moreover, adaptive shrinkage strategies and faster posterior approximations may further broaden applicability of this extension in large-scale scenarios. Although we focus on continuous data here, generalized PCFA (GPCFA, Chen, 2022) extends to categorical and mixed outcomes, providing a natural path for further development.

References

- Bainter, S. A., McCauley, T. G., Fahmy, M. M., Goodman, Z. T., Kupis, L. B., & Rao, J. S. (2023). Comparing Bayesian variable selection to lasso approaches for applications in psychology. *Psychometrika*, 88, 1032–1055. <https://doi.org/10.1007/s11336-023-09914-9>
- Barnard, J., McCulloch, R., & Meng, X.-L. (2000). Modeling covariance matrices in terms of standard deviations and correlations, with application to shrinkage. *Statistica Sinica*, 10, 1281–1311. <http://www.jstor.org/stable/24306780>
- Barbieri, M. M., & Berger, J. O. (2004). Optimal predictive model selection. *Annals of Statistics*, 32, 870–897.
- Bien, J., & Tibshirani, R. J. (2011). Sparse estimation of a covariance matrix. *Biometrika*, 98, 807–820. <https://doi.org/10.1093/biomet/asr054>
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81–105. <https://doi.org/10.1037/h0046016>
- Carvalho, C. M., Polson, N. G., & Scott, J. G. (2010). The horseshoe estimator for sparse signals. *Biometrika*, 97, 465–480. <https://doi.org/10.1093/biomet/asq017>
- Casella, G., Ghosh, M., Gill, J., & Kyung, M. (2010). Penalized regression, standard errors, and Bayesian lassos. *Bayesian Analysis*, 5, 369–412. <https://doi.org/10.1214/10-BA607>
- Chen, J. (2020). A partially confirmatory approach to the multidimensional item response theory with the Bayesian Lasso. *Psychometrika*, 85, 738–774. <https://doi.org/10.1007/s11336-020-09724-3>
- Chen, J. (2022). A generalized partially confirmatory factor analysis framework with mixed Bayesian Lasso methods. *Multivariate Behavioral Research*, 57, 879–894. <https://doi.org/10.1080/00273171.2021.1925520>
- Chen, J. (2022). Partially confirmatory approach to factor analysis with Bayesian learning: A LAWBL tutorial. *Structural Equation Modeling*.
- Wang, H. (2015). Scaling it up: Stochastic search structure learning in graphical models.
- Widaman, K. F. (1985). Hierarchically nested covariance structure models for multitrait-multimethod data. *Applied Psychological Measurement*, 9, 1–26. <https://doi.org/10.1177/014662168500900101>
- A *Multidisciplinary Journal*, 29, 800–816. <https://doi.org/10.1080/10705511.2022.2039660>
- Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics (Oxford, England)*, 9, 432–441. <https://doi.org/10.1093/biostatistics/kxm045>
- Harra, K., & Kaplan, D. (2024). On the performance of horseshoe priors for inducing sparsity in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 31, 667–684. <https://doi.org/10.1080/10705511.2023.2280895>
- Holzing, K. J., & Swineford, F. (1937). The bi-factor method. *Psychometrika*, 2, 41–54. <https://doi.org/10.1007/BF02287965>
- Jiménez, M., Abad, F. J., García-Garzon, E., & Garrido, L. E. (2023). Exploratory bi-factor analysis with multiple general factors. *Multivariate Behavioral Research*, 58, 1072–1089. <https://doi.org/10.1080/00273171.2023.2189571>
- Khondker, Z. S., Zhu, H., Chu, H., Lin, W., & Ibrahim, J. G. (2013). The Bayesian covariance lasso. *Statistics and Its Interface*, 6, 243–259. <https://doi.org/10.4310/sii.2013.v6.n2.a8>
- Lee, S.-Y. (1981). A Bayesian approach to confirmatory factor analysis. *Psychometrika*, 46, 153–160. <https://doi.org/10.1007/BF02293896>
- Li, Y., Craig, B. A., & Bhadra, A. (2019). The graphical horseshoe estimator for inverse covariance matrices. *Journal of Computational and Graphical Statistics*, 28, 747–757. <https://doi.org/10.1080/10618600.2019.1575744>
- Mahnke, R., Benlian, A., & Hess, T. (2014). Flow experience in information systems research: Revisiting its conceptualization, conditions, and effects. <https://doi.org/10.13140/2.1.4852.0486>
- Matzler, K., Faullant, R., & Mooradian, T. (2013). The thrill of victory, the agony of defeat: Personality, perceptions of price (Un)fairness, emotions, and price satisfaction. In *Psychology of emotions: Recent theoretical perspectives and novel empirical findings* (Vol. 2, pp. 13). Nova Science Publishers, Inc.
- Mitchell, T. J., & Beauchamp, J. J. (1988). Bayesian variable selection in linear regression. *Journal of the American Statistical Association*, 83, 1023–1032. <https://doi.org/10.1080/01621459.1988.10478694>
- Pan, J., Ip, E. H., & Dubé, L. (2017). An alternative to post hoc model modification in confirmatory factor analysis: The Bayesian lasso. *Psychological Methods*, 22, 687–704. <https://doi.org/10.1037/met0000112>
- Park, T., & Casella, G. (2008). The bayesian lasso. *Journal of the American Statistical Association*, 103, 681–686. <https://doi.org/10.1198/016214508000000337>
- Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, 47, 667–696. <https://doi.org/10.1080/00273171.2012.715555>
- Roskam, I., Galdiolo, S., Hansenne, M., Massoudi, K., Rossier, J., Gicquel, L., & Rolland, J.-P. (2015). The psychometric properties of the French version of the Personality Inventory for DSM-5. *PLoS One*, 10, e0133413. <https://doi.org/10.1371/journal.pone.0133413>
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48, 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika*, 22, 53–61. <https://doi.org/10.1007/BF02289209>
- Shaddox, E., Stingo, F. C., Peterson, C. B., Jacobson, S., Cruickshank-Quinn, C., Kechris, K., Bowler, R., & Vannucci, M. (2018). A Bayesian approach for learning gene networks underlying disease severity in COPD. *Statistics in Biosciences*, 10, 59–85. <https://doi.org/10.1007/s12561-016-9176-6>
- R Core Team. (2021). *R: A language and environment for statistical computing*. Foundation for Statistical Computing. <https://www.R-project.org/>
- van Erp, S. (2023). Bayesian regularized SEM: Current capabilities and constraints. *Psych*, 5, 814–835. <https://doi.org/10.3390/psych5030054>
- Wang, H. (2012). Bayesian graphical lasso models and efficient posterior computation. *Bayesian Analysis*, 7, 867–886. <https://doi.org/10.1214/12-BA729>



Yamaguchi, K., & Zhang, J. (2023). Fully Gibbs sampling algorithms for Bayesian variable selection in latent regression models. *Journal of Educational Measurement*, 60, 202–234. <https://doi.org/10.1111/jedm.12348>

Zhang, Y., & Chen, J. (2024). Accommodating and extending various models for special effects within the generalized partially confirmatory factor analysis framework. *Applied Psychological Measurement*, 48, 208–229. <https://doi.org/10.1177/01466216241261704>

Appendix A.

Lasso Conditional Distribution for Factor Loading

The conditional distribution for unspecified loadings Λ_j'' follows the formulation of Chen (2020) and can be expressed as:

$$\begin{aligned} p(\Lambda_j'' | \mathbf{Y}, \mathbf{F}, \Phi, \Psi) &\propto p(y_j | \Lambda_j'', \mathbf{F}, \Psi) p(\Lambda_j'' | \psi_{jj}) \\ &\propto \prod_{i=1}^N \exp \left\{ -\frac{1}{2} (y_{ij} - \Lambda_j''^T \omega_i'')^T \psi_{jj}^{-1} (y_{ij} - \Lambda_j''^T \omega_i'') \right\} \times p(\Lambda_j'' | \psi_{jj}) \\ &\Rightarrow p(\Lambda_j'' | \text{rest}) \propto N(m, \text{var}) \end{aligned}$$

where $m = (\mathbf{F}''(\mathbf{F}'')^T + \mathbf{D}_{\tau_j}^{-1})^{-1} (\mathbf{F}'' \mathbf{Y}_j'')$ and $\text{var} = \psi_{jj} (\mathbf{F}''(\mathbf{F}'')^T + \mathbf{D}_{\tau_j}^{-1})^{-1}$.

The full conditional distribution for shrinkage parameters is sampled:

$$\begin{aligned} p\left(\frac{1}{\tau_{jk}^2} | \text{rest}\right) &\propto \text{Inv} - \text{Gamma} \left(\sqrt{\frac{\delta_j^2}{\Lambda_{jk}^2} \psi_{jj}}, \delta_j^2 \right) \\ p(\delta_j^2 | \text{rest}) &\propto \text{Gamma} \left(a_l + 1, b_l + \frac{\tau_{jk}^2}{2} \right) \end{aligned}$$

where $a_l = 1, b_l = 0.1$.

The full conditional distribution for ψ_{jj} can be expressed as:

$$p(\psi_{jj}^{-1} | \text{rest}) \propto \text{Gamma} \left(a_r + \frac{N}{2}, b_r + \frac{1}{2} \left[\sum_{i=1}^N (y_{ij} - \Lambda_j''^T \omega_i'')^2 + \Lambda_j''^T \mathbf{D}_{\tau_j}^{-1} \Lambda_j'' \right] \right)$$

where $a_r = 1, b_r = 0.1$.

Appendix B.

Horseshoe Block Gibbs Scheme for Factor Correlation and Conditional Distribution for Factor Loading

Let Ω and $\mathbf{S}$ be partitioned as in Equation (6). The conditional distribution of the (ω_k, ω_{kk}) for horseshoe is:

$$\begin{aligned} &p(\omega_k, \omega_{kk} | \Omega_{-kk}, \text{rest}) \\ &\propto (\omega_{kk} - \omega_k^T \Omega_{-kk}^{-1} \omega_k)^{\frac{N}{2}} \times \exp \left\{ -\frac{1}{2} \left[\omega_k^T (\mathbf{D}_{\tau}^{-1} \lambda^2) \omega_k + 2\mathbf{s}_k^T \omega_k + (\mathbf{s}_{kk} + \lambda^2) \omega_{kk} \right] \right\} \end{aligned}$$

The changed variables (β, γ) follow:

$$\begin{aligned} &p(\beta, \gamma | \Omega_{-kk}, \text{rest}) \\ &\propto \gamma^{\frac{N}{2}} \exp \left(-\frac{\mathbf{s}_{kk} + \lambda^2}{2} \gamma \right) \times \exp \left(-\frac{1}{2} \left[\beta^T \left\{ (\mathbf{D}_{\tau}^{-1} \lambda^2) + (\mathbf{s}_{kk} + \lambda^2) \Omega_{-kk}^{-1} \right\} \beta + 2\mathbf{s}_k^T \beta \right] \right) \end{aligned}$$

The local and global shrinkage parameters are sampled:

$$\begin{aligned} p(\tau_k^2 | \text{rest}) &\propto \text{Inv} - \text{Gamma} \left(1, \frac{1}{\mathbf{v}_k} + \frac{\omega_k^T \omega_k}{2\lambda^2} \right) \\ p(\mathbf{v}_k | \text{rest}) &\propto \text{Inv} - \text{Gamma} \left(1, 1 + \frac{1}{\tau_k^2} \right) \\ p(\lambda^2 | \text{rest}) &\propto \text{Inv} - \text{Gamma} \left(\frac{K(K+1)}{2}, \frac{1}{\xi} + \sum_{k, k': k < k'} \frac{\omega_{kk'}^2}{2\tau_{kk'}^2} \right) \\ p(\xi | \text{rest}) &\propto \text{Inv} - \text{Gamma} \left(1, 1 + \frac{1}{\lambda^2} \right) \end{aligned}$$

The conditional distribution for unspecified loading Λ_j'' can be expressed as:

$$p(\Lambda_j'' | \text{rest}) \propto N(m, \text{var})$$

where $m = (\mathbf{F}''(\mathbf{F}'')^T + \mathbf{D}_{\tau_j}^{-1} \frac{1}{\lambda^2})^{-1} (\mathbf{F}'' \mathbf{Y}_j'')$ and $\text{var} = \psi_{jj} (\mathbf{F}''(\mathbf{F}'')^T + \mathbf{D}_{\tau_j}^{-1} \frac{1}{\lambda^2})^{-1}$.

The local and global shrinkage parameters are sampled:

$$\begin{aligned}
 p(\tau_{jk}^2 | rest) &\propto \text{Inv-Gamma} \left(1, \frac{1}{v_{jk}} + \frac{\Lambda_{jk}^T \Lambda_{jk}}{2\lambda^2 \psi_{jj}} \right) \\
 p(v_{jk} | rest) &\propto \text{Inv-Gamma} \left(1, 1 + \frac{1}{\tau_{jk}^2} \right) \\
 p(\lambda^2 | rest) &\propto \text{Inv-Gamma} \left(\frac{J \times K + 1}{2}, \frac{1}{\xi} + \frac{1}{2} \sum_{j=1}^J \frac{1}{\psi_{jj}} \sum_{k=1}^K \frac{\Lambda_{jk}^2}{\tau_{jk}^2} \right) \\
 p(\xi | rest) &\propto \text{Inv-Gamma} \left(1, 1 + \frac{1}{\lambda^2} \right)
 \end{aligned}$$

The full conditional distribution for ψ_{jj} can be expressed as:

$$p(\psi_{jj}^{-1} | rest) \propto \text{Gamma} \left(a_r + \frac{N}{2}, b_r + \frac{1}{2} \left[\sum_{i=1}^N (y_{ij} - \Lambda_j^T \omega_i'')^2 + (\Lambda_j^T \mathbf{D}_{\tau_j}^{-1} \Lambda_j'') \frac{1}{\lambda^2} \right] \right)$$

where $a_r = 1, b_r = 0.1$.

Appendix C.

Spike-and-Slab Block Gibbs Scheme for Factor Correlation and Conditional Distribution for Factor Loading

The conditional distribution of the (ω_k, ω_{kk}) is:

$$\begin{aligned}
 &p(\omega_k, \omega_{kk} | \Omega_{-kk}, rest) \\
 &\propto (\omega_{kk} - \omega_k^T \Omega_{-kk}^{-1} \omega_k)^{\frac{N}{2}} \times \exp \left\{ -\frac{1}{2} [\omega_k^T \mathbf{D}_{\tau}^{-1} \omega_k + 2\mathbf{s}_k^T \omega_k + (s_{kk} + \lambda) \omega_{kk}] \right\}
 \end{aligned}$$

The changed variables (β, γ) follows:

$$\begin{aligned}
 &p(\beta, \gamma | \Omega_{-kk}, rest) \\
 &\propto \gamma^{\frac{N}{2}} \exp \left(-\frac{s_{kk} + \lambda}{2} \gamma \right) \times \exp \left(-\frac{1}{2} [\beta^T \{ \mathbf{D}_{\tau}^{-1} + (s_{kk} + \lambda) \Omega_{-kk}^{-1} \} \beta + 2\mathbf{s}_k^T \beta] \right)
 \end{aligned}$$

The latent scale parameters τ are based on independent Bernoulli $z_{kk'}$:

$$p(z_{kk'} = 1 | rest) = \frac{N(\omega_{kk'} | 0, v_1^2) \pi}{N(\omega_{kk'} | 0, v_1^2) \pi + N(\omega_{kk'} | 0, v_0^2) (1 - \pi)}$$

The conditional distribution for unspecified loading Λ_j'' can be expressed as:

$$p(\Lambda_j'' | rest) \propto N(m, var)$$

where $m = (\mathbf{F}''(\mathbf{F}'')^T + \mathbf{D}_{\tau_j}^{-1})^{-1} (\mathbf{F}'' \mathbf{Y}_j'')$ and $var = \psi_{jj} (\mathbf{F}''(\mathbf{F}'')^T + \mathbf{D}_{\tau_j}^{-1})^{-1}$.

The latent scale parameters τ follow the Bernoulli z_{jk} like above graphical SSP:

$$p(z_{jk} = 1 | rest) = \frac{N(\Lambda_{jk} | 0, v_1^2) \pi}{N(\Lambda_{jk} | 0, v_1^2) \pi + N(\Lambda_{jk} | 0, v_0^2) (1 - \pi)}$$

Here we fix $\pi = 0.5, v_0 = 0.02, v_1 = 1$.

The full conditional distribution for ψ_{jj} can be expressed as:

$$p(\psi_{jj}^{-1} | rest) \propto \text{Gamma} \left(a_r + \frac{N}{2}, b_r + \frac{1}{2} \left[\sum_{i=1}^N (y_{ij} - \Lambda_j^T \omega_i'')^2 + \Lambda_j^T \mathbf{D}_{\tau_j}^{-1} \Lambda_j'' \right] \right)$$

where $a_r = 1, b_r = 0.1$.