Biometrics, 2025, 81(1), ujae160
https://doi.org/10.1093/biomtc/ujae160
Biometric Methodology

Robust Bayesian graphical regression models for assessing tumor heterogeneity in proteomic networks

Tsung-Hung Yao [1,*], Yang Ni[2], Anindya Bhadra [3], Jian Kang [1], Veerabhadran Baladandayuthapani [1,*]

[1]Department of Biostatistics, University of Michigan at Ann Arbor, Ann Arbor, MI 48109, United States, [2]Department of Statistics, Texas A&M University, College Station, TX 77843, United States, [3]Department of Statistics, Purdue University, West Lafayette, IN 47907, United States

*Corresponding authors: Tsung-Hung Yao, Department of Biostatistics, University of Michigan at Ann Arbor, Ann Arbor, MI 48109, USA (yaots@umich.edu); Veerabhadran Baladandayuthapani, Department of Biostatistics, University of Michigan at Ann Arbor, Ann Arbor, MI 48109, USA (veerab@umich.edu).

ABSTRACT

Graphical models are powerful tools to investigate complex dependency structures in high-throughput datasets. However, most existing graphical models make one of two canonical assumptions: (i) a homogeneous graph with a common network for all subjects or (ii) an assumption of normality, especially in the context of Gaussian graphical models. Both assumptions are restrictive and can fail to hold in certain applications such as proteomic networks in cancer. To this end, we propose an approach termed robust Bayesian graphical regression (rBGR) to estimate heterogeneous graphs for non-normally distributed data. rBGR is a flexible framework that accommodates non-normality through random marginal transformations and constructs covariate-dependent graphs to accommodate heterogeneity through graphical regression techniques. We formulate a new characterization of edge dependencies in such models called conditional sign independence with covariates, along with an efficient posterior sampling algorithm. In simulation studies, we demonstrate that rBGR outperforms existing graphical regression models for data generated under various levels of non-normality in both edge and covariate selection. We use rBGR to assess proteomic networks in lung and ovarian cancers to systematically investigate the effects of immunogenic heterogeneity within tumors. Our analyses reveal several important protein–protein interactions that are differentially associated with the immune cell abundance; some corroborate existing biological knowledge, whereas others are novel findings.

KEYWORDS: Bayesian graphical models; cancer; conditional sign independence; covariate-dependent graphs; protein—protein interactions.

1 INTRODUCTION

Graphical models are ubiquitous and powerful tools to investigate complex dependency structures in high-throughput biomedical datasets such as genomics and proteomics (Airoldi, 2007). They allow for holistic exploration of biologically relevant patterns that can be used for deciphering biological processes and formulating new testable hypotheses. However, most existing graphical models make one of two canonical assumptions: (i) a homogeneous graph that is common to all subjects or (ii) an assumption of normality, as in the context of Gaussian graphical models (Ni et al., 2022). However, in some biomedical applications, such as the inference of cancer proteomic networks, both assumptions fail, as we show next.

Proteomic networks and tumor heterogeneity. Proteins control many fundamental cellular processes through a complex but organized system of interactions, termed protein–protein interactions (PPIs; Cheng et al., 2020). Moreover, aberrant PPIs are associated with various diseases, including cancer, and investigating PPIs can lead to effective strategies and treatments, including immunotherapies, tailored to different individuals (Cheng et al., 2020). Consequently, it is highly desirable to elucidate PPIs in cancer and construct flexible graphical models that can identify multiple types and ranges of dependencies. Modern data collection methods have allowed systematic assessment of multiple proteins simultaneously in the same tumor samples, often referred to as high-throughput proteomics (Baladandayuthapani et al., 2014). However, the resulting data are typically not normally distributed, even after extensive preprocessing and data transformations (eg, logarithmic). As an illustration, Panel (B) of Figure 1 shows the level of non-normality in protein expression data after logarithmic transformation for 2 cancers: lung adenocarcinoma (LUAD) and ovarian cancer (OV) samples from The Cancer Genome Atlas (TCGA; Weinstein et al., 2013), which are used as case studies in this paper. Specifically, in 4 exemplar proteins, namely Akt and PTEN for LUAD and E-cadherin and Rb for OV, both the empirical density distributions and quantile–quantile (q-q) plots demonstrate deviations from the normal distribution with heavier tails. The level of non-normality is further quantified using the *H*-score, defined as $H(\boldsymbol{Y}) = 2\Phi[\log\{1 - \text{pval}(\boldsymbol{Y})\}]$, where Φ is the cumulative distribution function of the standard normal distribution, and $\text{pval}(\boldsymbol{Y})$ is the *P*-value of the

Received: October 24, 2023; Revised: November 25, 2024; Accepted: December 19, 2024

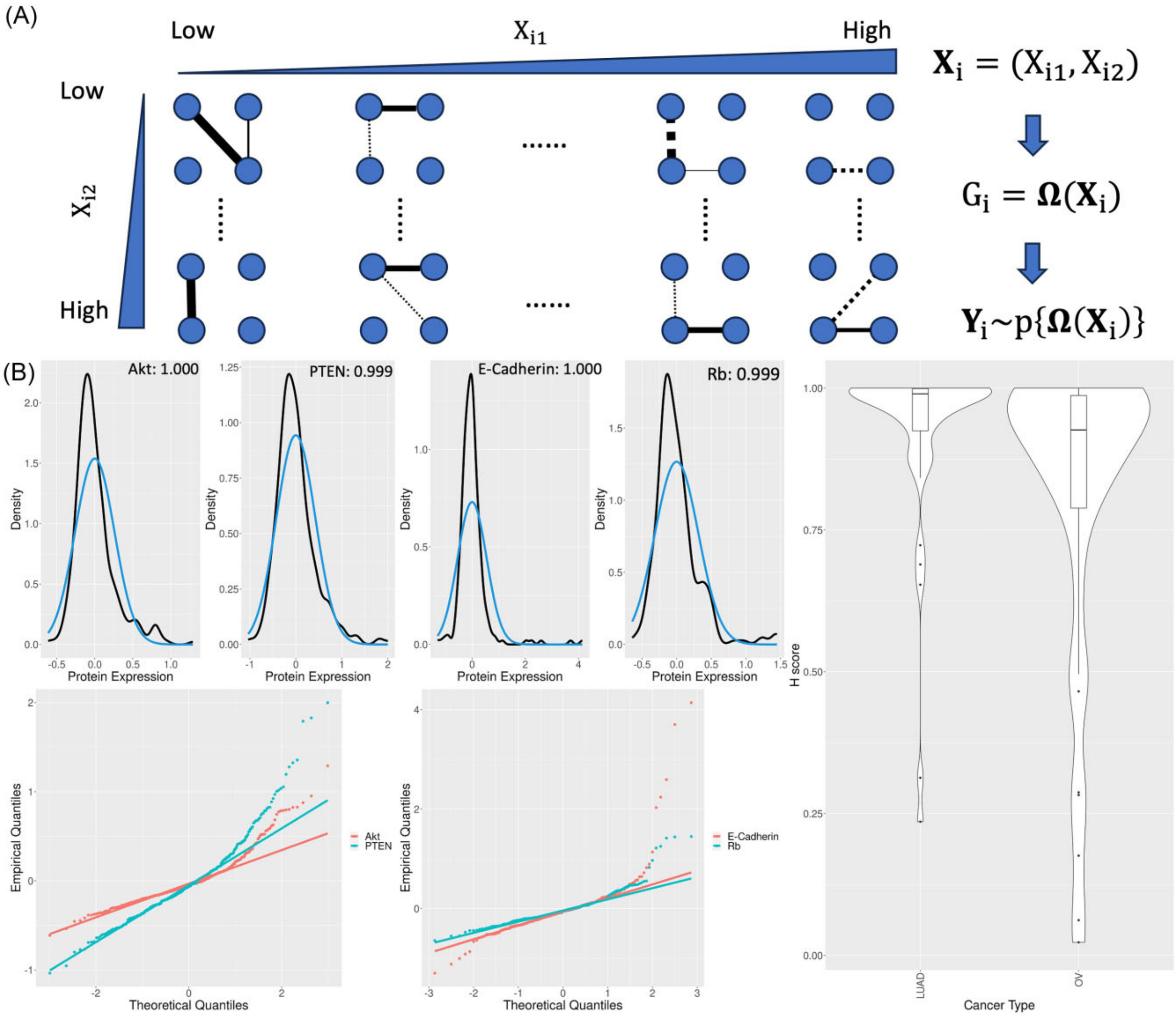

FIGURE 1 (A) A schematic diagram showing the heterogeneity in networks with observed responses based on heterogeneous networks. Heterogeneity in networks is characterized by covariates, $\boldsymbol{X}_i = (X_{i1}, X_{i2})$ as $G_i = \boldsymbol{\Omega}(\boldsymbol{X}_i)$. The width of edges in the network implies the magnitude of association between 2 nodes, and the 2 types of edges represent negative (dotted) and positive (solid) associations. Responses $\boldsymbol{Y}_i$ are obtained on heterogeneous networks $\boldsymbol{\Omega}(\boldsymbol{X}_i)$ that vary on covariate with $\boldsymbol{Y}_i \sim p\{\boldsymbol{\Omega}(\boldsymbol{X}_i)\}$, where p is a normal scale mixture distribution. (B) Examples of the non-normality levels of protein expression data from lung adenocarcinoma (LUAD) and ovarian cancer (OV). Empirical density from actual data (black) and the normal distribution (blue) for the expression of 4 proteins are shown on the top left with *H*-scores shown in the upper right of each protein. Quantile–quantile plots of the expression of 4 proteins in LUAD (Akt and PTEN) and OV (E-Cadherin and Rb) are illustrated in the lower left. The violin plot of *H*-scores across all proteins in LUAD and OV cancers is shown on the right. *H*-scores are bounded between 0 and 1, and a higher *H*-score implies a higher level of non-normality (see manuscript for more details). Two core assumptions of existing graphical models are violated in proteomic networks with heterogeneous networks (A) and non-normally distributed data (B).

Kolmogorov–Smirnov test for the normality of $\boldsymbol{Y}$ (Chakraborty et al., 2024). The *H*-score is bounded between 0 and 1, with a higher *H*-score implying increased departure from normality. The *H*-scores for all 4 proteins are > 0.999, consistent with the conclusions from the empirical and q-q plots. Panel (B) also shows the *H*-score across all the proteins in our datasets, indicating a high degree of non-normality across both cancers.

Another axis of complexity in cancer research is *tumor heterogeneity*. It is now well known that tumors are heterogeneous, with distinct proteomic aberrations even for the same type of cancer across different patients (Janku, 2014). Accumulating evidence suggests that considering tumor heterogeneity, both in general and specifically at the level of PPIs, can enhance our understanding of the progression of tumor growth and aid the development of anti-cancer treatments (Cheng et al., 2020). Specifically, tumor heterogeneity is characterized by different PPIs across patients and could potentially result in varied treatment responses (Cheng et al., 2020). Hence, incorporating patient-specific information, that is, accounting for tumor heterogeneity, could

TABLE 1 Comparison of existing methods and our proposed rBGR method across 4 different properties.

Method	Uncertainty quantification	Undirected	Sample-specific	Non-normality
GGMx (Ni et al., 2022)	✓	✓	✓	✗
RegGMM (Zhang and Li, 2022)	✗	✓	✓	✗
GSM (Bhadra et al., 2018)	✓	✓	✗	✓
BGR (Ni et al., 2019)	✓	✗	✓	✗
RCGM (Chakraborty et al., 2024)	✓	✓	✗	✓
rBGR (the proposed)	✓	✓	✓	✓

Abbreviations: BGR, Bayesian graphical regression; GGMx, Bayesian Gaussian graphical models with covariates; GSM, Gaussian scale mixture; rBGR, robust Bayesian graphical regression; RCGM, Bayesian robust chain graph model; RegGMM, Gaussian graphical regression models with covariates.

provide valuable clues to identify PPIs disrupted during carcinogenesis. In summary, constructing PPI networks poses 2 main statistical challenges simultaneously: (i) coherently accounting for non-normality in proteomic networks and (ii) incorporating heterogeneous patient-specific information in graphical modeling.

To encapsulate the tumor heterogeneity manifested in PPIs and the corresponding proteomic networks, here we model the edges of the proteomic network as a function of patient-specific covariates, resulting in networks that continuously vary based on these covariates. Consequently, each patient is associated with a unique proteomic network, and the observed proteomic expression depends on the underlying proteomic network. To illustrate, Panel (A) of Figure 1 shows heterogeneous networks that are characterized by 2 covariates with the corresponding observed proteomic expression.

Existing methods and modeling background. Most existing methods address the aforementioned challenges separately, that is, either accommodating non-normality without accounting for the sample-specific information (eg, Pitt et al., 2006; Dobra and Lenkoski, 2011) or requiring normality when incorporating patient-specific information (Ni et al., 2022). To accommodate non-normality, existing approaches transform the original variables into normal variables either via deterministic functions (eg, Dobra and Lenkoski, 2011; Liu et al., 2012; Chung et al., 2022) or via random transformations (eg, Finegold and Drton, 2011, 2014). For instance, Bhadra et al. (2018) use a Gaussian scale mixture technique that generalizes the t-distribution and introduce a new graph characterization for undirected graphs. Chakraborty et al. (2024) further generalize this concept to characterize chain graphs with both directed and undirected edges. However, all existing models that accommodate non-normality assume a common graph across all patients and fail to incorporate the subject-specific information.

More recently, several studies incorporated the subject-specific information under explicit Gaussian assumptions. Multiple Gaussian graphical models were first proposed to estimate graphs that vary across heterogeneous subpopulations (eg, Peng et al., 2009; Danaher et al., 2014; Peterson et al., 2015). Ni et al. (2019) introduce a more general framework called "graphical regression" that constructs covariate-dependent graphs through a regression model and incorporates both continuous and discrete covariates, in directed as well as undirected settings (Ni et al., 2022). Similarly, Zhang and Li (2022) provide a penalized procedure to estimate undirected graphs through Gaussian graphical regression and introduce continuous covariates in both the mean and the covariance structures. However, all these models are developed under the normality assumption for inferential and computational reasons (Table 1). To the best of our knowledge, no existing method incorporates subject-specific information under non-Gaussian settings, which motivates the development of a new methodology.

To address these challenges simultaneously, we developed a unified and flexible modeling strategy called robust Bayesian graphical regression (rBGR), which allows the construction of subject-specific graphical models for non-normally distributed continuous data. rBGR makes 3 main contributions:

1. *Robust framework to build subject-specific graphs for non-normal data.* rBGR robustifies the normal assumption via random transformation and incorporates covariates employing graphical regression strategies. By accommodating non-normality via random scale transformations, we obtain a Gaussian scale mixture, which presumes an underlying latent Gaussian variable, allows explicit incorporation of covariates in the precision matrix (Section 2.2), and admits efficient posterior sampling procedures (Supplementary Material Section S2).
2. *New characterization of dependency structures for non-normal graphical models.* The introduction of the random marginal transformations engenders a new type of edge characterization of the conditional dependence for non-normal data, called conditional sign independence with covariates (CSIx, Proposition 2). CSIx is a generalization of the notion of conditional sign independence (CSI), introduced by Bhadra et al. (2018), which explicitly characterizes the sign dependence between 2 variables that holds for a much broader class of models than Gaussian graphical models. We demonstrate via multiple simulations that rBGR can accurately recover dependency structures under different levels of non-normality and against competing graphical regression approaches that assume normality (Section 4).
3. *Deciphering the association of immunogenic heterogeneity with proteomic networks.* We use rBGR to assess proteomic networks across 2 cancers, lung and ovarian, to systematically investigate the effects of the

inherent immunogenic heterogeneity within tumors. Specifically, we quantify immune cell abundance across tumors and build PPI networks that vary across different immune cell abundances. Our analyses reveal several important PPIs that are associated with immune cell abundance; some corroborate existing biological knowledge, whereas others are novel associations (Section 5).

The rest of the paper is organized as follows: We introduce rBGR models and characterization in Section 2. Section 3 focuses on priors and estimation. In Section 4, we conduct a series of simulations to evaluate the operating characteristics of rBGR against competing approaches. Section 5 provides a detailed analysis of the TCGA dataset, results, biological interpretations, and implications. This paper concludes by discussing implications of the findings, limitations, and future directions in Section 6. A general-purpose R package and datasets used in this paper for constructing PPI networks are provided in the Supplementary Material.

2 ROBUST BAYESIAN GRAPHICAL REGRESSION

We start with the Gaussian graphical regression (Section 2.1), which is a special case of rBGR under the normality assumption, and then generalize it to the robust case by random transformations (Section 2.2). Subsequently, the introduction of the random scale changes the interpretation of the graph and motivates a new edge characterization (Section 2.3).

2.1 Gaussian graphical regression

Consider p-dimensional random vectors $\boldsymbol{Y}_i = (Y_{i1}, \ldots, Y_{ip})^{\mathrm{T}} \in \mathbb{R}^p$ as (continuous) responses with q-dimensional random vectors of $\boldsymbol{X}_i = (X_{i1}, \ldots, X_{iq})^{\mathrm{T}} \in \mathbb{R}^q$ as covariates for subject $i = 1, \ldots, n$. A subject-specific PPI network from proteomics data $\boldsymbol{Y}_i$ is constructed to vary based on the immune cell abundance $\boldsymbol{X}_i$ (Section 5). Let $G_i = (V, E_i)$ be an undirected graph over p nodes, where $V = \{1, \ldots, p\}$ is the set of nodes representing, $\boldsymbol{Y}_i$ and $E_i \subset V \times V$ is the set of undirected edges in the graph for subject i. An undirected edge exists between nodes j and k if $\{j, k\} \in E_i$. Under the Gaussian assumption, given the covariates $\boldsymbol{X}_i$, suppose $\boldsymbol{Y}_i$ follows a multivariate normal distribution,

$$\boldsymbol{Y}_i \mid \boldsymbol{X}_i \sim \boldsymbol{N}_p(\mathbf{0}, \widetilde{\boldsymbol{\Omega}}^{-1}(\boldsymbol{X}_i)), \text{ for } i = 1, \ldots, n, \qquad (1)$$

where $\widetilde{\boldsymbol{\Omega}}(\boldsymbol{X}_i) = \{\widetilde{\omega}^{j,k}(\boldsymbol{X}_i)\}_{p\times p}$, $j, k \in V$ is a functional precision matrix (of covariates) with each element $\widetilde{\omega}^{j,k}(\boldsymbol{X}_i)$ as a function that depends on $\boldsymbol{X}_i$. The functional precision matrix characterizes the graph G_i through zero precision elements. Specifically, a zero element of the precision matrix represents a missing edge in the graph; for example, for the case of a scalar precision matrix, $\widetilde{\omega}^{j,k}(\boldsymbol{X}_i) = \widetilde{\omega}^{j,k}$, zero precision implies conditional independence (CI) and a missing edge in the graph of CI under Gaussianity (Lauritzen, 1996). For the functional precision matrix, Ni et al. (2022) introduced covariate-dependent graphs in G and generalized the concept of CI to CI with covariates (CIx, henceforth). In essence, given a covariate $\boldsymbol{X}_i$, the zero precision of $\widetilde{\omega}^{j,k}(\boldsymbol{X}_i) = 0$ implies a missing edge of CIx between nodes j and k. Contrarily, when the functional precision is non-zero $\widetilde{\omega}^{j,k}(\boldsymbol{X}_i) \neq 0$, Y_j and Y_k are conditional dependent with covariates (CDx, henceforth), and an edge exists between nodes j and k given the covariate $\boldsymbol{X}_i$. By modeling the functional precision matrix, CIx defines covariate-specific graphs that vary based on different covariates.

2.2 Robust graphical regression via random transformation

In practice, the normal assumption does not always hold (Figure 1B). Violation of the normal assumption results in the failure of modeling graphs through normal precision matrices and motivates new modeling strategies (Finegold and Drton, 2011; Bhadra et al., 2018). Here, we adapt the random transformation approach (Bhadra et al., 2018) to allow for various non-normal distributions with different tail behaviors. We focus on continuous distributions with heavy tails as observed in our motivating data. To this end, let $0 < d_j < \infty$ for $j = 1, \ldots, p$ be independent positive random scales and have distribution as $d_j \sim p_j$ with $\int dp(d_j) < \infty$ almost surely. Let $\mathbf{D}_i = \mathrm{diag}(1/d_{i1}, \ldots, 1/d_{ip})$ be a diagonal matrix for subject i. Given random scales d_{ij}, $j = 1, \ldots, p$ and the covariates $\boldsymbol{X}_i$, we assume the distribution of $\mathbf{D}_i\boldsymbol{Y}_i$ conditional on $\mathbf{D}_i$ and $\boldsymbol{X}_i$ follows a multivariate distribution,

$$\mathbf{D}_i\boldsymbol{Y}_i = \left[\frac{Y_{i1}}{d_{i1}}, \ldots, \frac{Y_{ip}}{d_{ip}}\right]^{\mathrm{T}} \sim \boldsymbol{N}_p(\mathbf{0}, \boldsymbol{\Omega}^{-1}(\boldsymbol{X}_i)), \quad \text{for } i = 1, \ldots, n, \qquad (2)$$

where $\boldsymbol{\Omega}(\boldsymbol{X}_i) = \{\omega^{j,k}(\boldsymbol{X}_i)\}_{p\times p}$, $j, k \in V$ is the functional precision matrix that characterizes the graph with the covariates $\boldsymbol{X}_i$.

The model in (2) generalizes several existing approaches: (i) Equation 1 is a special case of Equation 2 with d_{ij} as a degenerated distribution of a point mass at 1; (ii) when $d_1 = \ldots = d_p = \tau$ with τ^2 following an inverse gamma distribution, Equation 2 is reduced to a multivariate t-distribution on $\boldsymbol{Y}$ as used by Finegold and Drton (2014); and (iii) for general d_{ij}, (2) establishes a rich family of Gaussian scale mixtures for the marginal distribution of Y_j with the density $p(Y_j) = \int (2\pi d_j)^{-1/2} \exp\{-y_j^2/(2d_j)\} dp(d_j)$.

The introduction of random scales in Equation 2 allows us to construct various marginal distributions of Y_j with different tail behaviors. Specifically, by matching tail behaviors of random scales and the target distribution, random scales allow us to model different marginal distributions that the data might exhibit. For example, letting Y_j decay polynomially, Y_j/d_j follows a normal distribution if the random scale d_j also has a polynomial tail. In Figure 2, Panel (A) shows that the target distribution of Y with a polynomial tail deviates from the normal distribution, but with the introduction of random scales, the distribution of Y/d is normally distributed. A similar idea can be used for target distribution with other tail behaviors, for example, exponential tails. Although the random scales robustify the model to accommodate non-normality, the resulting functional

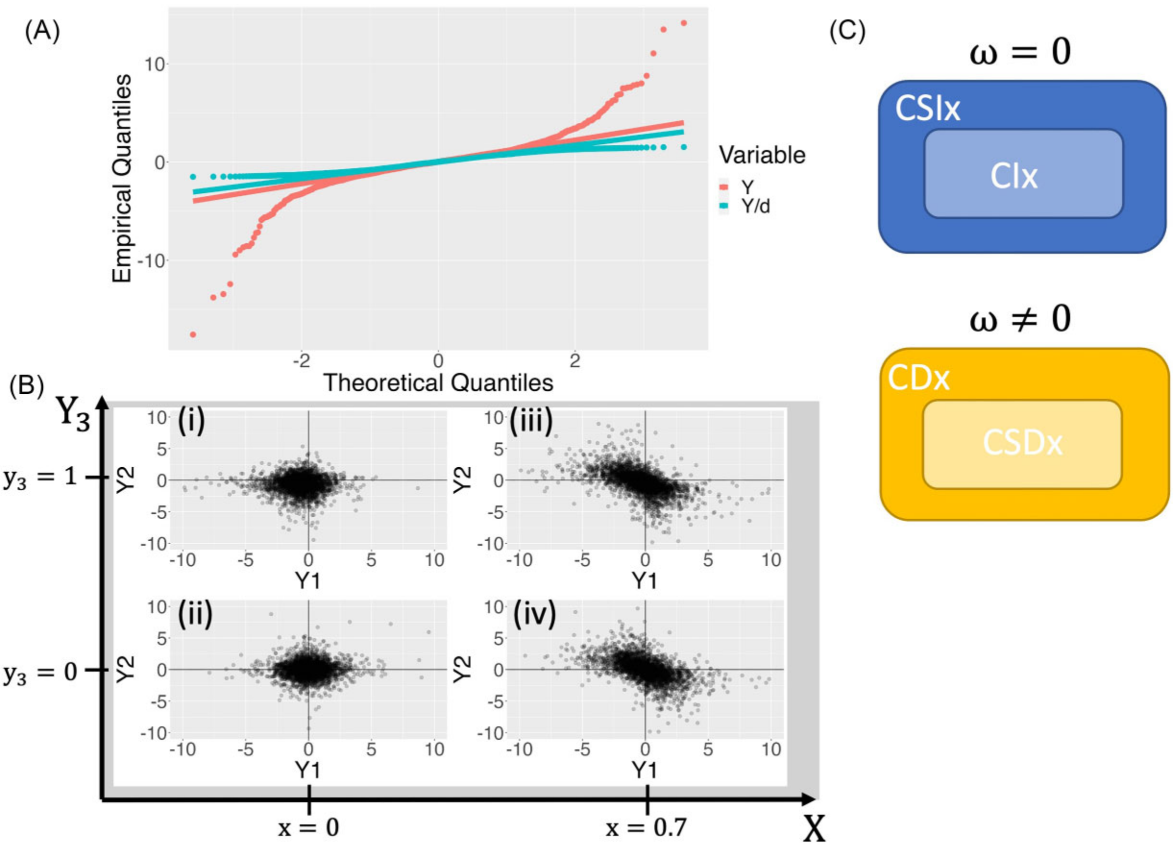

FIGURE 2 Robustification of non-normal distribution with random scales and the visualization of conditional sign independence with covariates (CSIx) and CSDx. Panel (A) shows the qq plot to illustrate that random scale d accommodates the non-normal distribution Y with Y/d following the normal distribution. Panel (B) demonstrates CSIx (Cases (i) and (ii)) and CSDx (Cases (iii) and (iv)) of Y_1 and Y_2 with the partial correlation $\omega^{1,2}(X_i) = X_i$ conditioning on Y_3. Cases (i) and (ii) represent 2 examples of CSIx with zero precision of $X_i = 0$ given $Y_3 = 1$ and 0. Cases (iii) and (iv) demonstrate the cases of CSDx with non-zero precision of $X_i = 0.7$ given $Y_3 = 1$ and 0. Panel (B) is centered on the values between $[-10, 10]$. Panel (C) shows the nested relationship between CSIx and CIx (top) and CSDx and CDx (bottom). See more details in Section 2.3.

precision matrix $\boldsymbol{\Omega}(\boldsymbol{X}_i)$ requires careful characterization and interpretation.

2.3 Characterization of functional precision matrix

The functional precision matrix in (2) determines the graphical dependence as a function of covariates, but the random (marginal) scale changes the standard CI interpretations in the resulting precision matrix, which requires a new characterization. Bhadra et al. (2018) introduced the concept of CSI in non-normal graphs, defined as follows. Consider the expression of 2 proteins of interest as random variables Y_1 and Y_2, with the expression data from the rest of the proteins denoted by a random vector $\boldsymbol{Y}_3$. Given $\boldsymbol{Y}_3$, Y_1 and Y_2 are conditional sign independent if $\mathcal{P}(Y_2 > 0 \mid Y_1, \boldsymbol{Y}_3) = \mathcal{P}(Y_2 > 0 \mid \boldsymbol{Y}_3)$ and $\mathcal{P}(Y_1 > 0 \mid Y_2, \boldsymbol{Y}_3) = \mathcal{P}(Y_1 > 0 \mid \boldsymbol{Y}_3)$. Otherwise, Y_1 and Y_2 are conditional sign dependent (CSD) given $\boldsymbol{Y}_3$. The CSI of Y_1 and Y_2 implies that the information of Y_1 does not affect the sign of Y_2 given $\boldsymbol{Y}_3$. That is, conditioning on the rest of the protein expression data $\boldsymbol{Y}_3$, the probability of over- or under-expression for Y_2 is independent of the expression level of Y_1. Under the multivariate distribution of (2) with a constant precision matrix $\boldsymbol{\Omega}(\boldsymbol{X}_i) = \boldsymbol{\Omega}$, zero precision of $\omega^{j,k} = 0$ and the CSI of Y_j and Y_k given the rest are equivalent, which can be represented by a missing edge between nodes j and k in an undirected graph (Bhadra et al., 2018; Chakraborty et al., 2024).

In this paper, we generalize the concept of CSI to incorporate covariates and consider the subject-specific CSI of 2 random variables given all the other random variables and a realization of covariates $\boldsymbol{X}_i$, as formalized in the following proposition:

Proposition 1 (Conditional Sign Independence with Covariate (CSIx)) *Given random scales* $\mathbf{D}_i = diag(1/d_{i1}, \ldots, 1/d_{ip})$ *and the covariates* $\boldsymbol{X}_i$*, consider the conditional distribution of* $\mathbf{D}_i\boldsymbol{Y}_i$ *as Equation 2 with functional precision matrix* $\boldsymbol{\Omega}(\boldsymbol{X}_i)$*. If* $\omega^{j,k}(\boldsymbol{X}_i) = 0$*, then* Y_{ij} *and* Y_{ik} *are CSI. Otherwise, when* $\omega^{j,k}(\boldsymbol{X}_i) \neq 0$*, then* Y_{ij} *and* Y_{ik} *are CSD.*

The proof of Proposition 1 follows the fact that $\omega^{j,k}(\boldsymbol{X}_i) = 0$ implies the CSI of Y_{ij} and Y_{ik} given $\boldsymbol{X}_i$, and we call Y_{ij} and Y_{ik} conditional sign independent with covariates $\boldsymbol{X}_i$ to highlight the role of the covariates in the graph. Otherwise, Y_{ij} and Y_{ik} are called conditional sign dependent with covariates (CSDx). See Supplementary Material Section S1 for details.

An example CSIx. We use a simple low-dimensional example to visually demonstrate and interpret CSIx and CSDx. Following Proposition 1, we show 2 examples with a general functional precision matrix $\boldsymbol{\Omega}(\boldsymbol{X}_i)$. Consider a trivariate distribution of (2) with unit diagonal elements and $\omega^{1,2}(X_i) = X_i$. We illustrate 2 scenarios in Panel (B) of Figure 2:

- When $X_i = 0$, we obtain the CSIx of Y_1 and Y_2 given 2 different values of $Y_3 = 0$ (Case (i)) and 1 (Case (ii)).
- When $X_i = 0.7$, Y_1 and Y_2 are CSDx, and we observe that the distribution of the sign of Y_2 varies based on the value of Y_1 (Case (iii) and Case (iv)). Specifically, as Y_1 increases, Y_2 tends to be negative.

By modeling the functional precision matrix, we can build a covariate-specific precision matrix that depends on the different realizations of the covariates $\boldsymbol{X}_i$. Consequently, we can construct a graph of CSI corresponding to the precision matrix and the covariates.

We can now conceptually compare models (1) and (2). Both models incorporate the covariates in the functional precision matrix, which characterize the covariate-specific graph. However, the interpretation of the graph differs. The graph from model (2) encodes CSIx, whereas the graph from model (1) encodes CIx. We visualize the relationship between CSIx and CIx in Panel (C) of Figure 2 and summarize it as follows:

- For $\omega = 0$, CSIx is a weaker condition than CIx since CSIx considers only the sign, while CIx depends on both the sign and the magnitude.
- When $\omega \neq 0$, CSDx is a stronger condition than CDx since CDx allows either magnitude or the sign to be dependent, while CSDx focuses on only the sign.

3 PRIORS AND ESTIMATION

The functional precision matrix $\boldsymbol{\Omega}(\boldsymbol{X}_i)$ lives in a high-dimensional space. For example, the PPI network for OV from our application considers $\mathcal{O}(np(p-1)/2) = 197,620$ possible edges, which makes joint estimation difficult if not untenable, especially since variability across each subject i is allowed. Hence, we employ a neighborhood selection procedure (Meinshausen and Bühlmann, 2006) to estimate the functional precision matrix that has been used in several graphical models (eg, Ni et al., 2019; Zhang and Li, 2022). This procedure offers 3 main benefits: (i) tractable estimation, (ii) reduced computation burden, and (iii) flexible prior elicitation. Specifically, we regress 1 node Y_j on the rest nodes Y_k, $k \neq j$ and build the graph based on zero coefficients (Sections 3.1 and 3.2). This use of neighborhood selection, which employs conditional estimation as opposed to joint estimation in (2), effectively reduces the number of edges to $\mathcal{O}(qp(p-1)/2) = 3280$—a 60-folds decrease. Moreover, the effective number of edges can be reduced by different model specifications, such as a thresholding mechanism (Section 3.3), and different priors, such as spike-and-slab (Section 3.4).

3.1 Regression-based approach for functional precision matrix estimation

The rBGR model leverages a regression-based framework on model (2) to relate the regression coefficients and precision matrix. Given random scales $\boldsymbol{D}_i$, we regress 1 variable on all other variables and relate the partial correlation with regression coefficients. Zero coefficients are then equivalent to zero partial correlations (Meinshausen and Bühlmann, 2006). Specifically, we construct node-specific regressions as

$$\frac{Y_{ij}}{d_{ij}} = \sum_{k\neq j}^{p} \beta_{j,k}(\mathbf{X}_\mathbf{i})\frac{Y_{ik}}{d_{ik}} + \epsilon_{ij}, \tag{3}$$

where $\epsilon_{ij} \sim N(0, 1/\omega^{j,j}(\mathbf{X}_i))$ and the functional coefficient $\beta_{j,k}(\boldsymbol{X}_i) = -\frac{\omega^{j,k}(\mathbf{X}_i)}{\omega^{j,j}(\mathbf{X}_i)}$. Under this specification, $\beta_{j,k}(\boldsymbol{X}_i) = 0$ if and only if $\omega^{j,k}(\boldsymbol{X}_i) = 0$, which enables the functional coefficients to characterize the covariate-specific graphs. However, the interpretation of the coefficients changes from the standard Gaussian graphical models due to the introduction of the random scales, which is detailed in the next section.

3.2 Graph construction through regression coefficients

We build graphs with a missing edge between nodes j and k when Y_j and Y_k are CSIx given the remaining variables and the covariates $\boldsymbol{X}_i$. Consider $\boldsymbol{Y}_i$ and $\boldsymbol{X}_i$ with the regression (3). We call $\beta_{j,k}(\boldsymbol{X}_i)$ the conditional sign independence function (CSIF) because zero CSIF, $\beta_{j,k}(\mathbf{X}_i) = 0$, implies that Y_j and Y_k are CSIx given all the other nodes $\boldsymbol{Y}_{-\{j,k\}}$ and covariates $\boldsymbol{X}_i$, as formally characterized in the following proposition.

Proposition 2 *Consider model (3). If* $\beta_{j,k}(\mathbf{X}_i) = 0$, *then* $\mathcal{P}(Y_j > 0 \mid Y_k, \boldsymbol{Y}_{-\{j,k\}}, \mathbf{X}_i) = \mathcal{P}(Y_j > 0 \mid \boldsymbol{Y}_{-\{j,k\}}, \mathbf{X}_i)$ *and* $\mathcal{P}(Y_k > 0 \mid Y_j, \boldsymbol{Y}_{-\{j,k\}}, \mathbf{X}_i) = \mathcal{P}(Y_k > 0 \mid \boldsymbol{Y}_{-\{j,k\}}, \mathbf{X}_i)$.

We sketch the proof here and leave the details in Supplementary Section S1. The proof follows from the fact that the CSIF $\beta_{j,k}(\boldsymbol{X}_i) = -\frac{\omega^{j,k}(\boldsymbol{X}_i)}{\omega^{j,j}(\boldsymbol{X}_i)}$ is related to the partial correlation, and a zero partial correlation is equivalent to a zero precision of $\omega^{j,k}(\boldsymbol{X}) = 0$, which ensures the CSIx between Y_j and Y_k (see the example in Section 2.3). Therefore, zero CSIF indicates the CSIx between Y_j and Y_k given the remaining response variables $\boldsymbol{Y}_{-\{j,k\}}$ and covariates $\boldsymbol{X}_i$. In this paper, we further assume $\omega^{j,j}$ in CSIF to be scalar, $\beta_{j,k}(\boldsymbol{X}_i) = -\frac{\omega^{j,k}(\boldsymbol{X}_i)}{\omega^{j,j}}$, for ease of computation. Note that our main interest is edge selection, and the CSIF is 0 if and only if $\omega^{j,k}(\boldsymbol{X}_i) = 0$, which is unrelated to $\omega^{j,j}$.

3.3 Modeling the CSIF

Proposition 2 transforms the problem of robust graph construction to a more tractable regression coefficient selection (ie, selecting which part of CSIF is exactly 0). Therefore, modeling the CSIF is crucial to the graph estimation. To this end, we parameterize the CSIF as a product of 2 components:

$$\beta_{j,k}(\boldsymbol{X}_i) = \underbrace{\theta_{j,k}(\boldsymbol{X}_i)}_{\text{Covariate function}} \; \underbrace{\mathbb{I}(|\theta_{j,k}(\boldsymbol{X}_i)| > t_j)}_{\text{Thresholding function}}. \tag{4}$$

We elaborate the role and justification of each component below.

Covariate functions $[\theta_\bullet(\bullet)]$. For simplicity, we consider only the linear effects of covariates $\boldsymbol{X}_i$, $\theta_{j,k}(\boldsymbol{X}_i) = \sum_{h=1}^{q} \alpha_{j,k,h} X_{ih}$, where $\alpha_{j,k,h}$ represents the coefficients for the hth covariate. The covariate function allows similar edge sets for individuals with a similar level of $\boldsymbol{X}_i$ and varies the graph, thus borrowing strength. If desired, it is relatively straightforward to extend $\alpha_{j,k,h}$ to nonlinear effects by, for example, using basis expansion techniques such as splines.

Thresholding functions $[\mathbb{I}(|\,\theta_\bullet(\boldsymbol{X})\,| > t_\bullet)]$. The edge thresholding mechanism is desired to achieve sparse graphs in rBGR due to the large number of parameters. For example, the ovarian PPI network in our application requires $qp(p-1)/2 = 3280$ parameters and results in a dense graph with inefficient inference. To solve the problem, we truncate edges with small magnitudes with an indicator function $\mathbb{I}(|\,\theta_{j,k}(\boldsymbol{X})\,| > t_j)$, where t_j is the threshold parameter specific to the node j. An edge is shrunk to 0 and removed when the magnitude is smaller than the threshold parameter, resulting in a sparse graph. One might consider the threshold parameter as $t_{j,k}$. However, $t_{j,k}$ is not fully identifiable when $\alpha_{j,k,h} = 0$ for all $h = 1, \ldots, q$ since when $\theta_{j,k}(\boldsymbol{X}_i) = 0$, the value of $t_{j,k}$ can be arbitrary. To alleviate the problem, we assume $t_{j,k} = t_j$ to improve the identifiability as long as one of $\theta_{j,k} \neq 0$.

3.4 Prior specification

rBGR contains 3 parameters: (1) random scales d_j, (2) threshold parameter t_j, and (3) covariate coefficients $\alpha_{j,k,h}$. To complete the model specification, we assign priors as follows:

$$d_j \sim (1-\pi_j)\delta_1(d_j) + \pi_j p_j(d_j);\;\; t_j \sim \text{Unif}(0, t_{\max})$$
$$\alpha_{j,k,h} \sim \gamma_{j,k,h} N(0, \nu_{j,k,h}) + (1-\gamma_{j,k,h}) N(0, \nu_0 \nu_{j,k,h}),$$

where ν_0 and $t_{\max}$ are pre-specified hyperparameters, π_j models the degree of non-normality with a beta prior of $\pi_j \sim Beta(a_\pi, b_\pi)$, p_j is a function to accommodate the non-normality, $\gamma_{j,k,h}$ is a binary variable with a Bernoulli prior of $\gamma_{j,k,h} \sim Ber(\rho_j)$, and $\nu_{j,k,h}$ decides the variance of $\alpha_{j,k,h}$ with an inverse gamma prior of $\nu_{j,k,h} \sim InvGa(a_\nu, b_\nu)$. Specifically, when $d_j = 1$, Y_j is normally distributed. When $d_j \sim p_j$, Y_j follows a non-normal distribution. We match the tail behavior of p_j and the marginal distribution of Y_j and allow each marginal distribution Y_j to have a different level of non-normality by specific π_j. For the current model, we focus on the Y_j with polynomial decay, as illustrated by the motivating data in Figure 1, and assign an inverse gamma distribution on $p_j(d_j) = \text{InvGa}(d_j^2 \mid a_d, b_d)$. For threshold parameter t_j, we assign a uniform prior on t_j to model the thresholding mechanism and control the graph sparsity. Intuitively, when $t_j \to 0$, no edge is truncated with a fully connected graph. When $t_j \to \infty$, all edges are shrunk to 0 with all nodes disconnected. For covariate coefficients $\alpha_{j,k,h}$, we assign a spike-and-slab prior to achieve the covariate sparsity with a small ν_0 because not all covariates necessarily contribute to the varying structure. These model specifications enable an efficient Gibbs sampler, and we show details of the posterior inference, including the conditional distributions and coefficient symmetrization in Supplementary Material Section S2.

4 SIMULATION STUDIES

We empirically demonstrate the performance of rBGR under a variety of non-normal contaminations and against other competing models in terms of edge and covariate selection. To the best of our knowledge, no other existing method estimates covariate-specific graphs for non-normal data. Therefore, we compare rBGR to 2 models that estimate the covariate-specific graph without addressing the violation of the normality assumption. Specifically, we consider Bayesian graphical regression (BGR; Ni et al., 2019) and the Gaussian graphical regression models with covariates (RegGMM; Zhang and Li, 2022) representative of a fully Bayesian and a frequentist penalization-based model for the covariate-specific graph under the normal assumption, respectively. Implementation details of all algorithms can be found in Supplementary Materials Section S3.1.

Data-generating mechanism. We generate the observed non-normal data by multiplying the random scale by the latent data $\boldsymbol{Y}_i^* = \left[Y_{i1}^*, \ldots, Y_{ip}^*\right]^{\mathrm{T}}$. Specifically, the latent data are normally distributed as $\boldsymbol{Y}_i^* \overset{iid}{\sim} \boldsymbol{N}_p(\boldsymbol{0}, \boldsymbol{\Omega}^{-1}(\boldsymbol{X}_i))$, where covariates follow a uniform distribution as $\boldsymbol{X}_i \overset{iid}{\sim} U(-1, 1)$ and $\boldsymbol{\Omega}(\boldsymbol{X}_i)$ is the true precision matrix representing the undirected graph. For $\boldsymbol{\Omega}(\boldsymbol{X}_i)$, we assign unit diagonal elements and randomly pick 2% of the off-diagonal to be non-zero. We let the non-zero precision depend on the covariates linearly and truncate the precision with a magnitude smaller than 0.15. We obtain the random scales from a mixture distribution of the point mass at 1 and an inverse gamma distribution and assign 3 different levels of non-normal contamination: $\pi \in \{0, 0.5, 0.8\}$. We multiply the random scales by $\boldsymbol{Y}_i^*$ to generate the observed data of $[Y_{i1}, \ldots, Y_{ip}] = [Y_{i1}^* d_{i1}, \ldots, Y_{ip}^* d_{ip}]$. For all simulations, we set the sample size and the dimensions of $\boldsymbol{Y}_i$ and $\boldsymbol{X}_i$ as $(n, p, q) = (250, 50, 3)$ based on our real data case studies. We show the results for 50 independent replicates.

Performance metrics. We evaluate the graph recovery through the edge and covariate selection. For covariate selection, we report the true positive rate (TPR), true negative rate (TNR), Matthew's correlation coefficient (MCC) with the cut-off for the posterior inclusion probability (PIP) at $c_0 = 0.5$, and the area under the receiver operating characteristic curve (AUC). For edge selection, we use AUC and 3 metrics of TPR, TNR, and MCC with the cut-off for the edge posterior probability (ePP) at $c_1 = 0.5$. We further investigate the sign consistency by examining the agreement between the posterior probability for the signs of CSIF $\text{sgn}(\widehat{\beta}_{j,k}(\boldsymbol{X}_i))$ and the true signs of $\text{sgn}(\beta_{j,k}(\boldsymbol{X}_i))$. Specifically, we exclude the zero CISF and focus on the subset of the data with both true and estimated non-zero CSIF to restrict the problem as 2-class classification (positive versus negative). We assess the sign consistency by MCC (referred to as sign-MCC).

Simulation results. Panel (A) of Figure 3 shows the simulation results for covariate selection. We observe that rBGR outperforms BGR and RegGMM across all non-normality levels, as indicated by higher MCC and AUC. The difference in MCC and AUC between rBGR and the other competing methods increases when the non-normality contamination level increases,

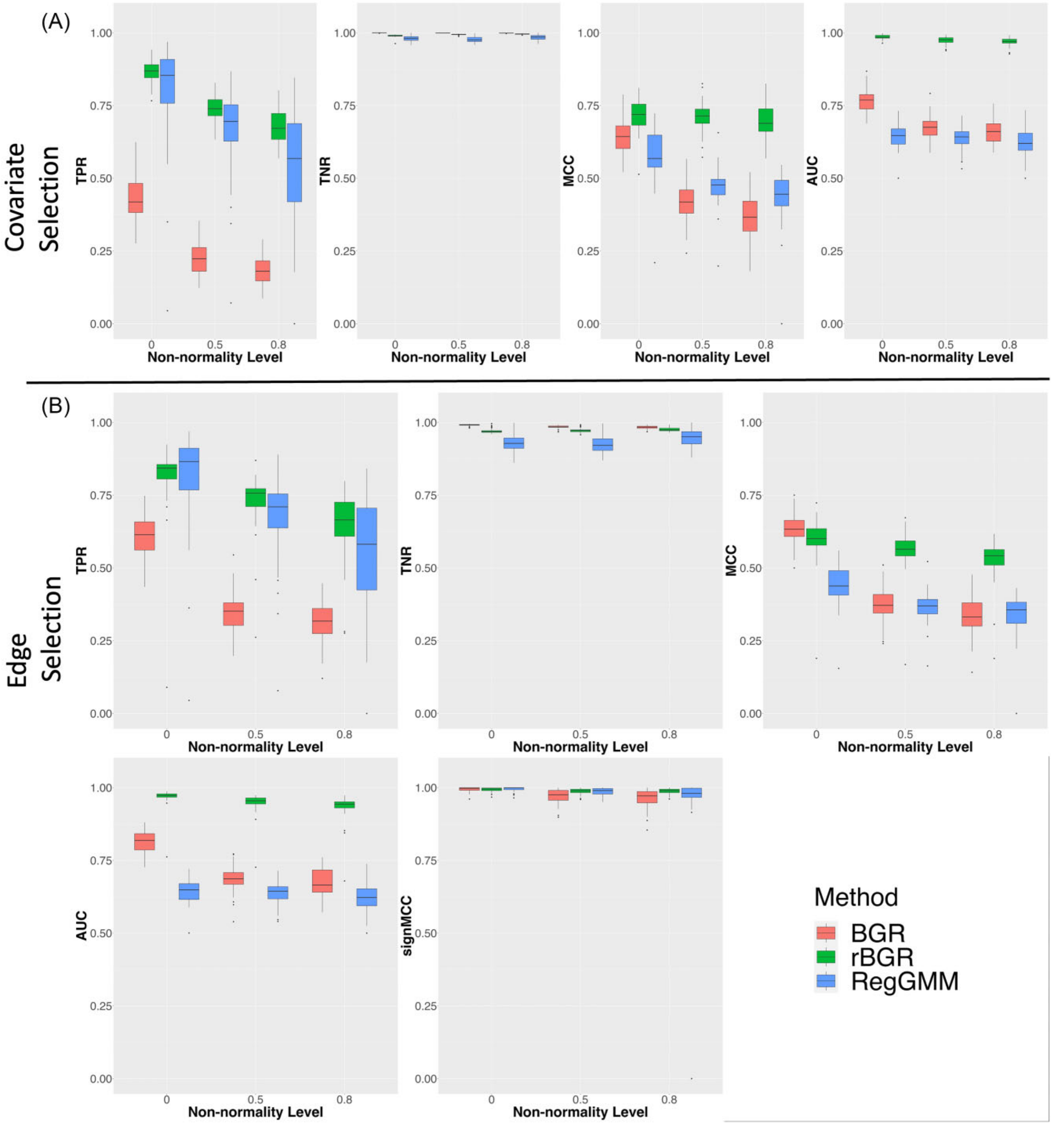

FIGURE 3 Simulation results: graph recovery for Bayesian graphical regression (BGR) (red), robust Bayesian graphical regression (rBGR) (green), and Gaussian graphical regression models with covariates (RegGMM) (blue) under different levels of non-normality in terms of (A) covariate selection (top row) and (B) edge selection (bottom 2 rows). Panel (A) shows the covariate selection through 4 metrics (from left to right: true positive rate [TPR], true negative rate [TNR], Matthew's correlation coefficient [MCC], and area under the receiver operating characteristic curve [AUC]) that are measured under 3 different levels of non-normality. Panel (B) demonstrates the edge selection by 4 criteria (from upper left to lower left: TPR, TNR, MCC, and AUC) and the sign consistency by sign-MCC (lower right) for non-zero edges. All values for TPR, TNR, and MCC are measured at a cut-off at $c_0 = c_1 = 0.5$.

which is expected. For TNR, rBGR performs slightly worse than BGR but better than RegGMM across all non-normality levels. However, all 3 methods select correct covariates ($> 93\%$) with small differences ($< 5\%$) in terms of TNR. For TPR, rBGR outperforms BGR under all levels of non-normality, and the advantage of rBGR becomes more prominent as the non-normality increases. Compared to RegGMM, rBGR's performance is comparable under a normal distribution in TPR, but rBGR is preferred when the level of non-normality increases. Overall, modeling non-normality using random scales in rBGR is favorable compared to models without random scales in terms of covariate selection.

We show the graph recovery for the edge selection in Panel (B) of Figure 3. For edge selection, rBGR outperforms BGR and RegGMM in AUC, and the advantage of rBGR increases with a larger discrepancy between rBGR and the competing methods when the non-normality level increases. For MCC, rBGR outperforms RegGMM under all levels of non-normality, but is slightly inferior compared to BGR under the normal distribution. However, rBGR is favored when the non-normality level increases. For TPR, rBGR is better than BGR under all levels of non-normality, and slightly worse than RegGMM under the normal assumption. However, when non-normality increases, rBGR starts to surpass the RegGMM. Both TNR and sign-MCC show excellent selection performance ($> 95\%$) for all 3 methods, with minimal differences ($< 5\%$) across the 3 non-normality levels.

In summary, modeling non-normality through random scales in rBGR results in equivalent (under the normal distribution) or better performances in all metrics for edge selection compared to the other methods.

Additional simulations and model evaluations. We provide additional simulation results in the Supplementary Material for (i) details of data-generating mechanisms (Section S3.2), (ii) convergence of the algorithm (Section S3.3), (iii) different cut-off of c_0 and c_1 controlling for false discovery rates (Section S3.4), (iv) exponential tails of $p_j(d_j) = \text{Exp}(d_j^2|a_d)$ (Section S3.5), (v) high-dimensional networks of $p = 100$ and $q = 1$ (Section S3.6), and (vi) comparison to homogeneous graph models (Section S3.7). Overall, rBGR generates equivalent or better performances compared to other methods under various non-normal tails and higher dimensional networks.

5 ANALYSES OF PROTEOMIC NETWORKS UNDER IMMUNOGENIC HETEROGENEITY

Key scientific questions and dataset overview. Aberrant PPIs are associated with various diseases, including cancer (Cheng et al., 2020), and immune cells around the tumor can modulate malfunctioning PPIs to influence tumor growth and progression (Joyce and Fearon, 2015). In cancer, cells around the tumor form the tumor microenvironment (TME), which closely interacts with the tumor (Whiteside, 2008). For example, dysregulated PPIs in the tumor suppress multiple immune cells in the TME to escape detection of the tumor by the immune system (Whiteside, 2008), while immune cells in the TME can alter the aberrant PPIs to eliminate cancerous cells (Joyce and Fearon, 2015). This demonstrates the connection between the dysregulated PPIs and the TME and shows the importance of immunogenic heterogeneity in tumor behavior. A better understanding of the association of immune cells with aberrant PPIs offers a foundational paradigm for potential targeted therapies in cancer (Cheng et al., 2020). To this end, our key scientific questions were as follows: (i) identify important PPIs across different cancer types and (ii) discover the effect of immunogenic heterogeneity on aberrant PPIs as potential targets for future investigation.

We exemplify the utility of rBGR by using data from TCGA to build patient-specific PPI networks and investigate the association of immunogenic heterogeneity across 2 different cancers. Specifically, we used a reverse-phase protein array for proteomic data ($\boldsymbol{Y}$) to build the PPI network of a CSIx graph, and we incorporated the immune cell transcriptome signatures as covariates ($\boldsymbol{X}$) as markers of immunogenic heterogeneity. Our analysis focuses on OV and LUAD as representative examples of 2 cancers that elicit distinct immune responses. OV represents an immunologically “cold” tumor, with a weaker immune response, while LUAD is considered an immunologically “hot” tumor, with a stronger immune response (Galon and Bruni, 2019). We focus on proteins in 12 important cancer-related pathways (Ha et al., 2018) and obtain $p = 41$ proteins with $n = 241$ and $n = 360$ patients for OV and LUAD, respectively. For covariates, we included mRNA-derived immune cell gene signatures and quantified the immune cell abundance corresponding to T cells and 2 crucial types of myeloid-derived suppressor cells, monocytes and neutrophils, for both OV and LUAD. Both T cells and myeloid-derived suppressor cells are essential in both OV and LUAD since T cells are the main immune component that kills cancer cells, while myeloid-derived suppressor cells regulate T cells (Whiteside, 2008). We ran rBGR on OV and LUAD with 20 000 iterations and discarded the first 19 000 iterations. The convergence diagnostics and the details of data preprocessing procedures are provided in Supplementary Material Section S4.1.

5.1 Population-level proteomic networks

We first focus on the covariate-dependent population-level networks for OV and LUAD that are estimated by $\widehat{\alpha}_{j,k,h}$. The corresponding networks are shown in Figure 4 (Panels (A) for LUAD and (B) for OV). We observed that the number of edges is much fewer in OV compared to LUAD for all immune components (T cells: [7, 15], monocytes: [5, 82], and neutrophils: [7, 260] for [OV, LUAD]). This is further evidenced in Panel (C), which shows the distribution of PIPs for OV and LUAD. Interestingly, we observe that the PIPs for LUAD are higher than those for OV for all immune components (medians of [OV, LUAD] for T cells: [0.123, 0.271], monocytes: [0.131, 0.307], and neutrophils: [0.127, 0.380]). The higher PIPs in LUAD imply that immune components have a greater correlation with PPIs in LUAD compared to OV. This finding is consistent with the existing biology (Galon and Bruni, 2019), as LUAD belongs to the immune hot tumors, with a stronger immune response.

Population graphs also confer specific information about the interaction between proteins. For example, we observe an edge between Akt and PTEN with the highest PIP regulated by T cells for LUAD (Figure 4A), suggesting an association of T cells with the PPI between Akt and PTEN. It is well known that PTEN downregulates Akt, and the loss of tumor suppressor PTEN often leads to a dysregulated PI3K pathway, including Akt and tumor growth in LUAD (Conciatori et al., 2020). For OV, despite the smaller number of PPIs, we still identify important PPIs. For example, rBGR suggests a PPI regulated by T cells between Caveolin-1 and PR. In OV, Caveolin-1 is regulated by progesterone, which is mediated by PR, and suggests a result consistent with the estimated PPI between Caveolin-1 and PR (Syed

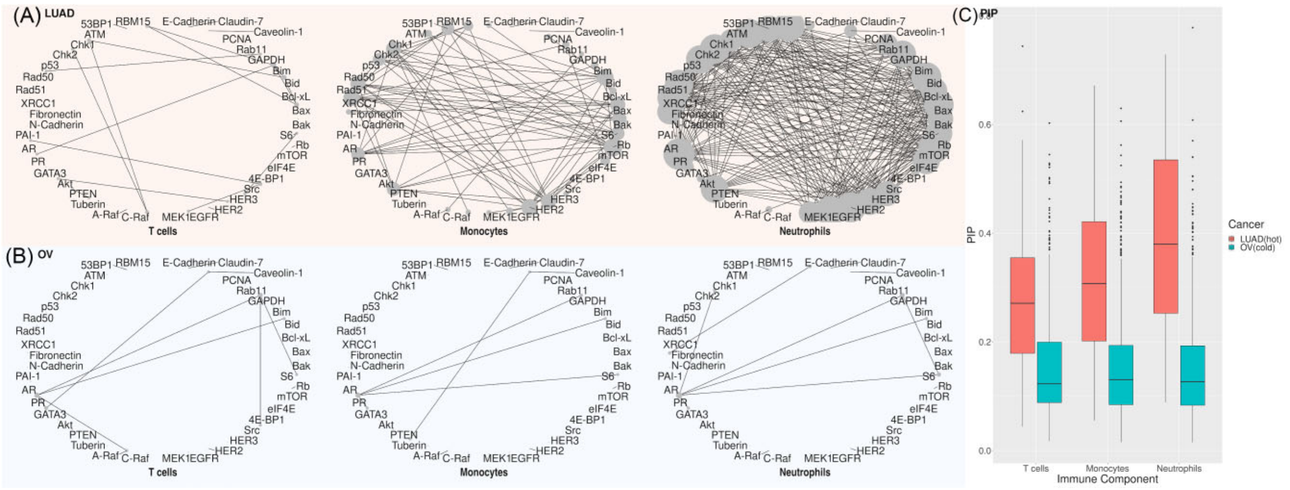

FIGURE 4 Population-level proteomic networks. The posterior inclusion probability (PIP) (Panel (C)) and the population networks of PPIs with the cut-off at $c_0 = 0.5$ for lung adenocarcinoma (LUAD) (Panel (A)) and and ovarian cancer (OV) (Panel (B)). For each panel of LUAD and OV, PPI networks of specific immune components are shown from left to right for T cells, monocytes, and neutrophils. For each protein, a bigger node represents a higher degree of connectivity.

et al., 2005). Overall, our analyses capture important hub proteins and characterize the cancer PPIs, and the results are highly concordant with the existing cancer literature.

5.2 Patient-specific networks: $\boldsymbol{\beta}_{jk}(\boldsymbol{X}_i)$

We next focus on patient-specific PPI networks to examine the effect of immune component abundance ($\boldsymbol{X}_i$) on PPIs. Specifically, we vary 1 immune component, with the rest of the components fixed at their mean, and generate networks of CSIx for different individuals at 5 percentiles (5th, 25th, 50th, 75th, and 95th percentiles) of the varying immune component. We set the cut-off for the ePP at $c_1 = 0.5$ and show the networks for LUAD in Figure 5 with the networks for OV in Supplementary Material S4.

We focus on PPIs of CSDx that show dependency on the abundance of a specific immune component. We present PPIs that change signs in the 5th and 95th percentiles, indicating specific PPIs that are correlated with an immune component, such as Akt-PTEN for T cells, Bid-PCNA for monocytes, and Bax-GATA3 for neutrophils. Interestingly, we discovered that the sign of Akt-PTEN is positively correlated with the T-cell abundance. Specifically, when T-cell abundance is higher, Akt-PTEN is positive, whereas Akt-PTEN is negative when T cells are scarcer. It is well established that PTEN suppresses Akt signaling and that the loss of PTEN results in the hyper-activation of Akt in cancer cells and low T-cell abundance in lung cancer (Conciatori et al., 2020). In addition, we find the Bid-PCNA edge is positively correlated with monocyte abundance. It has been shown that PCNA promotes Bid through caspase proteins and is crucial to immune evasion in cancers (Wang et al., 2021). Finally, we discover that the Bax-GATA3 edge is positively correlated with neutrophil abundance. Recently, GATA3 has been found to downregulate BCL-2 (Cohen et al., 2014), which inhibits the Bax protein (Cohen et al., 2014), and neutrophils promote the Bax to induce apoptosis (Li et al., 2020). These findings highlight specific PPIs that are influenced by the abundance of immune components and suggest potential targets for further investigation.

6 DISCUSSION

In this paper, we develop a flexible Bayesian framework called rBGR to construct heterogeneous networks that explicitly account for covariate-specific information for non-normally distributed data. By accommodating the non-normal marginal tail behaviors through random scales, we construct covariate-specific graphs through graphical regression-based approaches. This framework allows us to explicitly characterize, infer, and interpret covariate-specific edge dependencies through CSIFs. We also propose an efficient Gibbs sampler for posterior inference. Our simulations demonstrate that rBRG outperforms other existing Gaussian-based methods that construct the covariate-specific graphs under a variety of settings, which display non-normal marginal behavior such as heavy tails or skewness.

We employ rBGR on proteogenomic datasets in 2 cancers to build patient-specific PPI networks and identify PPIs that are impacted by tumor heterogeneity. Specifically, we quantify the immune cell abundance to elucidate the effects of immunogenic heterogeneity on aberrant PPIs for lung and OVs that are triggered by different levels of immune responses. Our analyses align with existing biology along 3 major axes: (i) immune responses, (ii) hub proteins, and (iii) PPIs. For example, higher connections in LUAD are consistent with existing biology since LUAD belongs to the class of the immunologically "hot" tumors. We identify a hub protein of HER2, which is associated with poor survival in LUAD. Another example is the PPI of Akt-PTEN, which is consistent with the knowledge that PTEN downregulates Akt. Our study further suggests PPIs that vary with specific immune components. For example, we discover that PPIs of Akt-PTEN, Bid-PCNA, and Bax-GATA3 vary positively with

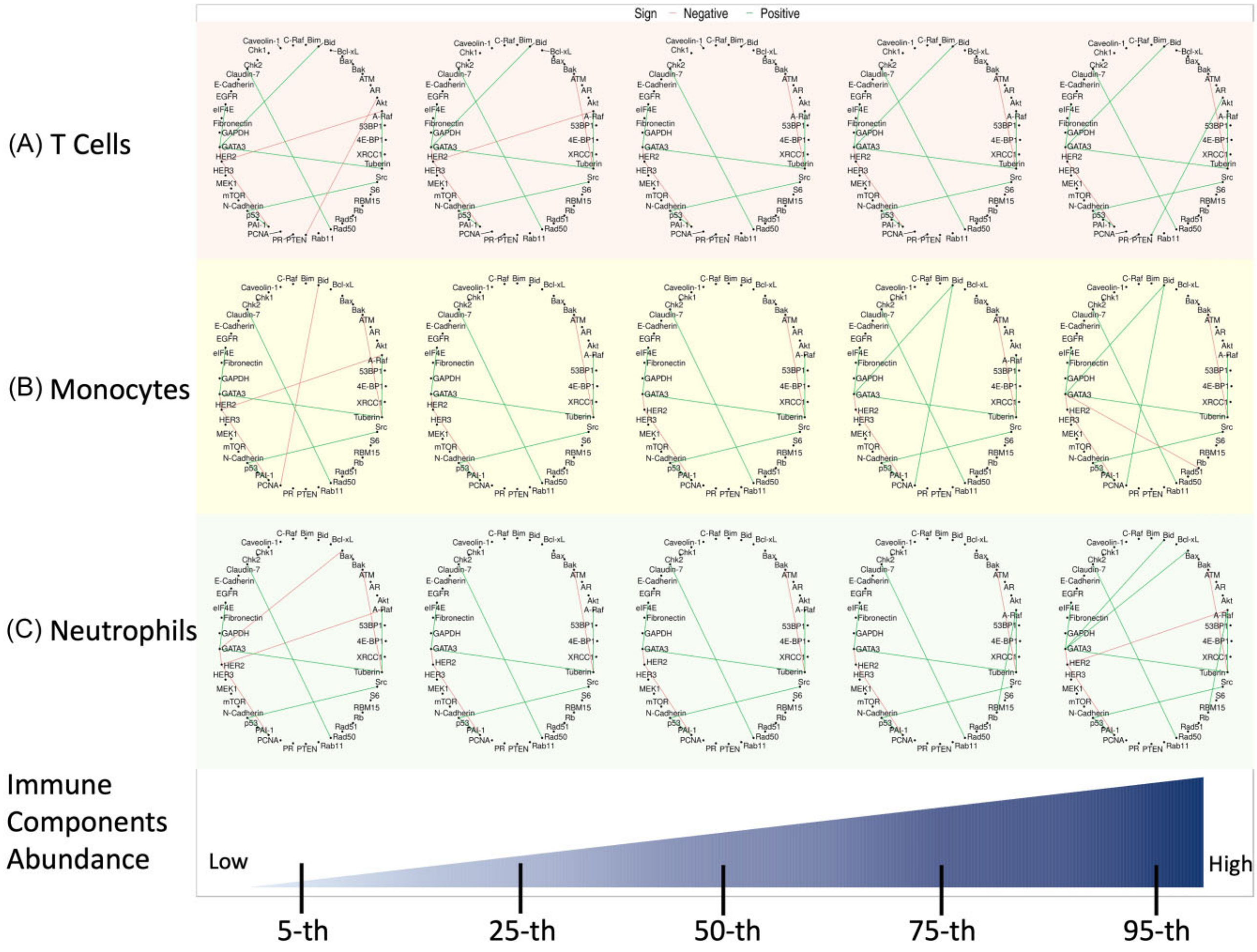

FIGURE 5 Covariate-specific networks of lung adenocarcinoma (LUAD) under 5 different percentiles of immune component abundances for (A) T cells, (B) monocytes, and (C) neutrophils with the other 2 components fixed at mean zero. The estimated networks for varying immune components are shown from left to right for 5th, 25th, 50th, 75th, and 95th percentiles. Edges are identified with signs (green: positive and red: negative) when the edge posterior probabilities (ePPs) are higher than $c_1 = 0.5$.

T cells, monocytes, and neutrophils, respectively. These findings suggest potential future targets for immunotherapy in lung cancer.

In the current implementation, we consider only the linear effect of continuous covariates to reduce the inferential and computation burden. One extension is to consider discrete covariates such as membership indicators that result in a robust multiple-group graphical model. It is also possible to include the nonlinear functionals through basis expansion techniques such as splines (Ni et al., 2019); however, this will increase the computational burden. Another possible extension is other types of graphical dependencies. For example, a chain graph considers an ordered multi-level structure via directed and undirected edges (eg, Chakraborty et al., 2024). By introducing random scales and generalizing the regression coefficients as functional coefficients, the model can include the covariates in the precision matrix to build the subject-specific chain graphs. Another direction can be to include discrete nodes and to extend the concept of CSIx for discrete data (Bhadra et al., 2018). All these directions are left for future investigations.

SUPPLEMENTARY MATERIALS

Supplementary material is available at *Biometrics* online.

Web appendices, Figures, and data and code referenced in Sections 2 to 5 are available with this paper at the Biometrics website on Oxford Academic. The code is also available on GitHub https://github.com/bayesrx/rBGR.

FUNDING

None declared.

CONFLICT OF INTEREST

None declared.

DATA AVAILABILITY

The cancer proteomics data that support the findings in this paper are available in the online Supplementary Material.

REFERENCES

Airoldi, E. M. (2007). Getting started in probabilistic graphical models. *PLoS Computational Biology*, 3, e252.

Baladandayuthapani, V., Talluri, R., Ji, Y., Coombes, K. R., Lu, Y., Hennessy, B. T. et al. (2014). Bayesian sparse graphical models for classification with application to protein expression data. *Annals of Applied Statistics*, 8, 1443–1468.

Bhadra, A., Rao, A. and Baladandayuthapani, V. (2018). Inferring network structure in non-normal and mixed discrete-continuous genomic data. *Biometrics*, 74, 185–195.

Chakraborty, M., Baladandayuthapani, V., Bhadra, A. and Ha, M. J. (2024). Bayesian robust learning in chain graph models for integrative pharmacogenomics. *The Annals of Applied Statistics*, 18, 3274–3296.

Cheng, S. S., Yang, G. J., Wang, W., Leung, C. H. and Ma, D. L. (2020). The design and development of covalent protein–protein interaction inhibitors for cancer treatment. *Journal of Hematology & Oncology*, 13, 26.

Chung, H. C., Gaynanova, I. and Ni, Y. (2022). Phylogenetically informed Bayesian truncated copula graphical models for microbial association networks. *Annals of Applied Statistics*, 16, 2437–2457.

Cohen, H., Ben-Hamo, R., Gidoni, M., Yitzhaki, I., Kozol, R., Zilberberg, A. et al. (2014). Shift in GATA3 functions, and GATA3 mutations, control progression and clinical presentation in breast cancer. *Breast Cancer Research*, 16, 464.

Conciatori, F., Bazzichetto, C., Falcone, I., Ciuffreda, L., Ferretti, G., Vari, S. et al. (2020). PTEN function at the interface between cancer and tumor microenvironment: implications for response to immunotherapy. *International Journal of Molecular Sciences*, 21, 5337.

Danaher, P., Wang, P. and Witten, D. M. (2014). The joint graphical lasso for inverse covariance estimation across multiple classes. *Journal of the Royal Statistical Society, Series B (Methodology)*, 76, 373–397.

Dobra, A. and Lenkoski, A. (2011). Copula Gaussian graphical models and their application to modeling functional disability data. *Annals of Applied Statistics*, 5, 969–993.

Finegold, M. and Drton, M. (2011). Robust graphical modeling of gene networks using classical and alternative *t*-distributions. *Annals of Applied Statistics*, 5, 1057–1080.

Finegold, M. and Drton, M. (2014). Robust Bayesian graphical modeling using dirichlet *t*-distributions. *Bayesian Analysis*, 9, 521–550.

Galon, J. and Bruni, D. (2019). Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. *Nature Reviews Drug Discovery*, 18, 197–218.

Ha, M. J., Banerjee, S., Akbani, R., Liang, H., Mills, G. B., Do, K.-A. et al. (2018). Personalized integrated network modeling of the cancer proteome atlas. *Scientific Reports*, 8, 14924.

Janku, F. (2014). Tumor heterogeneity in the clinic: is it a real problem?. *Therapeutic Advances in Medical Oncology*, 6, 43–51.

Joyce, J. A. and Fearon, D. T. (2015). T cell exclusion, immune privilege, and the tumor microenvironment. *Science*, 348, 74–80.

Lauritzen, S. L. (1996). *Graphical Models*. New York: Oxford University Press.

Li, R., Zou, X., Zhu, T., Xu, H., Li, X. and Zhu, L. (2020). Destruction of neutrophil extracellular traps promotes the apoptosis and inhibits the invasion of gastric cancer cells by regulating the expression of bcl-2, bax and nf-κb. *OncoTargets and Therapy*, 13, 5271–5281.

Liu, H., Han, F., Yuan, M., Lafferty, J. and Wasserman, L. (2012). High-dimensional semiparametric Gaussian copula graphical models. *Annals of Statistics*, 40, 2293–2326.

Meinshausen, N. and Bühlmann, P. (2006). High-dimensional graphs and variable selection with the Lasso. *Annals of Statistics*, 34, 1436–1462.

Ni, Y., Baladandayuthapani, V., Vannucci, M. and Stingo, F. C. (2022). Bayesian graphical models for modern biological applications. *Statistical Methods & Applications*, 31, 197–225.

Ni, Y., Stingo, F. C. and Baladandayuthapani, V. (2019). Bayesian graphical regression. *Journal of the American Statistical Association*, 114, 184–197.

Ni, Y., Stingo, F. C. and Baladandayuthapani, V. (2022). Bayesian covariate-dependent Gaussian graphical models with varying structure. *Journal of Machine Learning Research*, 23, 1–29.

Peng, J., Wang, P., Zhou, N. and Zhu, J. (2009). Partial correlation estimation by joint sparse regression models. *Journal of the American Statistical Association*, 104, 735–746.

Peterson, C. B., Stingo, F. C. and Vannucci, M. (2015). Bayesian inference of multiple Gaussian graphical models. *Journal of the American Statistical Association*, 110, 159–174.

Pitt, M., Chan, D. and Kohn, R. (2006). Efficient Bayesian inference for Gaussian copula regression models. *Biometrika*, 93, 537–554.

Syed, V., Mukherjee, K., Lyons-Weiler, J., Lau, K. M., Mashima, T., Tsuruo, T. et al. (2005). Identification of ATF-3, caveolin-1, DLC-1, and NM23-H2 as putative antitumorigenic, progesterone-regulated genes for ovarian cancer cells by gene profiling. *Oncogene*, 24, 1774–1787.

Wang, Y. L., Lee, C. C., Shen, Y. C., Lin, P. L., Wu, W. R., Lin, Y. Z. et al. (2021). Evading immune surveillance via tyrosine phosphorylation of nuclear PCNA. *Cell Reports*, 36, 109537.

Weinstein, J. N., Collisson, E. A., Mills, G. B., Shaw, K. R., Ozenberger, B. A., Ellrott, K. et al. (2013). The Cancer Genome Atlas Pan-Cancer analysis project. *Nature Genetics*, 45, 1113–1120.

Whiteside, T. L. (2008). The tumor microenvironment and its role in promoting tumor growth. *Oncogene*, 27, 5904–5912.

Zhang, J. and Li, Y. (2022). High-dimensional Gaussian graphical regression models with covariates. *Journal of the American Statistical Association*, 118, 2088–2100.

Received: October 24, 2023; Revised: November 25, 2024; Accepted: December 19, 2024