Biometrics 000, 000–000 DOI: 000
000 0000

Inference for Nonparanormal Partial Correlation via Regularized Rank-based Nodewise Regression

Haoyan Hu, and Yumou Qiu*

Department of Statistics, Iowa State University, Ames, Iowa, 50010, U.S.A.

**email:* yumouqiu@iastate.edu (corresponding author)

Summary:

Partial correlation is a common tool in studying conditional dependence for Gaussian distributed data. However, partial correlation being zero may not be equivalent to conditional independence under non-Gaussian distributions. In this paper, we propose a statistical inference procedure for partial correlations under the high-dimensional nonparanormal (NPN) model where the observed data are normally distributed after certain monotone transformations. The nonparanormal partial correlation is the partial correlation of the normal transformed data under the NPN model, which is a more general measure of conditional dependence. We estimate the NPN partial correlations by regularized nodewise regression based on the empirical ranks of the original data. A multiple testing procedure is proposed to identify the nonzero NPN partial correlations. The proposed method can be carried out by a simple coordinate descent algorithm for lasso optimization. It is easy-to-implement and computationally more efficient compared to the existing methods for estimating NPN graphical models. Theoretical results are developed to show the asymptotic normality of the proposed estimator and to justify the proposed multiple testing procedure. Numerical simulations and a case study on brain imaging data demonstrate the utility of the proposed procedure and evaluate its performance compared to the existing methods. Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database.

Key words: FDR control; high dimensionality; nonparanormal model; partial correlation; regularised regression.

This paper has been submitted for consideration for publication in *Biometrics*

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/biom.13624

1. Introduction

Measures of dependence are commonly used to understand the interactions among variables and the data generation mechanism of a complex system. Studying variable dependence is an essential problem in many biological studies, especially in the experiments that collect data from a large number of variables, for example, studying gene expression network (Wang and Huang, 2014) and brain connectivity (Huang et al., 2010; Qiu and Zhou, 2020, 2021). It is widely believed that different brain regions work together in performing our daily activities (Bullmore and Sporns, 2009). However, neurodegenerative diseases, such as Alzheimer's disease, may alter the interactions among brain regions (Supekar et al., 2008; Qiu and Zhou, 2020). Therefore, understanding the brain functional connectivity can help the diagnosis and treatment of such diseases.

Partial correlation measures conditional dependence after controlling the effects of other variables. Compared to marginal correlation, it provides a direct association between two variables after adjusting the confounding effects of other variables. When the number of variables is greater than the sample size, estimates for partial correlation cannot be directly computed as the sample covariance matrix is not invertible. Over the past decade, estimating a high-dimensional precision matrix has gained increasing attention. Under a suitable sparsity condition on the population precision matrix, it can be consistently estimated by the neighborhood selection method (Meinshausen and Bühlmann, 2006), graphical lasso (Friedman et al., 2008), and CLIME (Cai et al., 2011, 2016) by penalized estimation or constraint ℓ_1 norm optimization. See other penalized estimators in Ming and Yi (2007); Banerjee et al. (2008); Peng et al. (2009). Statistical inference procedures for high-dimensional precision or partial correlation matrices were constructed in Liu (2013); Ren et al. (2015); Wang et al. (2016); Chang et al. (2018); Qiu and Zhou (2020) using residuals from either pairwise or nodewise regressions.

Despite the popularity, the equivalence between zero partial correlation and conditional independence relies on the Gaussian assumption. To construct more general dependence measures under non-Gaussian distributions, Liu et al. (2009) introduced the nonparanormal (NPN) model where the data are normally distributed after certain unknown monotone transformations. Under the NPN model, the NPN partial correlations of the normal transformed data can be used to infer conditional independence. Liu et al. (2012); Xue and Zou (2012) proposed regularized estimates for precision matrices under the NPN graphical model. However, statistical inference procedures were not considered.

Gu et al. (2015); Xu et al. (2016); He et al. (2017); Barber and Kolar (2018) constructed novel statistical inference procedures for each entry of the precision matrix under either the NPN model or a more general transelliptical model with theoretical guarantees. In particular, He et al. (2017) proposed a multiple testing procedure for recovering nonzero NPN precision coefficients. All the aforementioned approaches estimate the precision coefficients by a regularized inverse of the covariance estimate, obtained by the sine transformation of the sample Kendall's τ from the original data. However, such a covariance estimate may not be positive semi-definite, which may incur non-convexity issues in estimating the precision coefficients and bring heavy computation. Moreover, those inference procedures require estimating the variance of a debiased estimator of the regularized inverse, which is computationally intensive for data sets with many variables. A more detailed discussion on the computation of those methods is presented in Section 3.3.

In this paper, we propose a computationally efficient and easy-to-implement approach via regularized rank-based nodewise regression (RRNR) to estimate and test the partial correlations under the NPN model. The proposed approach applies nodewise regression on the normal quantile transformation of the empirical ranks of the original data. A multiple testing procedure is constructed to recover the nonzero partial correlations. The proposed procedure

can be implemented by a simple coordinate descent algorithm for lasso optimization and does not suffer from non-convexity issues in computation. Theoretical results show that the RRNR estimator of partial correlation is asymptotically normal under a high-dimensional setting, and the proposed multiple testing procedure can control the false discovery rate (FDR) at the nominal level. Simulation studies are conducted to evaluate the performance of the proposed procedure, which shows that it reaches the highest power among the existing methods while controls the FDR. Empirical analysis on a brain imaging data set demonstrates the utility of the proposed procedure in practice.

The rest of the paper is organized as follows. The nonparanormal model and the hypotheses of interest are introduced in Section 2. The proposed RRNR estimator, the multiple testing procedure, and the connection between the proposed estimator and the existing methods are given in Section 3. Section 4 provides theoretical properties for the proposed method. Simulation studies and a real data analysis are reported in Sections 5 and 6, respectively. Discussion on the applications and future extensions of the proposed method is presented in Section 7. All technical proofs are relegated to the Supporting Information.

2. Preliminary and Model

In this section, we introduce the nonparanormal model and the target partial correlations under this model. The notations to be used throughout the paper are summarized in the following. For an $n \times p$ matrix $\mathbf{A} = (A_{ij})_{n\times p}$, let $\mathbf{A}_{(j)}$ denote the jth column of $\mathbf{A}$, and $\mathbf{A}_{-j}$ be the sub-matrix of $\mathbf{A}$ without the jth column. Let $|\mathbf{A}|_\infty = \max_{1\leqslant i\leqslant n, 1\leqslant j\leqslant p} |A_{ij}|$ be the matrix element-wise maximum norm. For a vector $\mathbf{a} = (a_1, \ldots, a_d)^{\mathrm{T}} \in \mathbb{R}^d$, let $|\mathbf{a}|_q = (\sum_{1\leqslant j\leqslant d} |a_j|^q)^{1/q}$ denote the ℓ_q norm of $\mathbf{a}$, and $|\mathbf{a}|_0 = \sum_{j=1}^d \mathbb{I}(a_j \neq 0)$ and $|\mathbf{a}|_\infty = \max_{1\leqslant j\leqslant d} |a_j|$ be the ℓ_0 and ℓ_∞ norms of $\mathbf{a}$, where $\mathbb{I}(\cdot)$ is the indicator function. Let $\mathbf{a}_{-j}$ be the sub-vector without the jth component.

Let $\mathbf{Z} = (\mathbf{Z}_1, \ldots, \mathbf{Z}_n)^{\mathrm{T}}$ be the observed data matrix, where $\mathbf{Z}_i = (Z_{i1}, \ldots, Z_{ip})^{\mathrm{T}}$ for

$i = 1, \ldots, n$, and the dimension p could be much larger than the sample size n. Suppose the observations $\{\mathbf{Z}_i\}_{i=1}^n$ are independent and identically distributed (i.i.d.) random vectors drawn from a p-dimensional distribution $\mathbf{F}_{\mathrm{Z}}$ with mean $\boldsymbol{\mu}_{\mathrm{Z}} = (\mu_{\mathrm{Z},1}, \ldots, \mu_{\mathrm{Z},p})^{\mathrm{T}}$ and covariance $\boldsymbol{\Sigma}_{\mathrm{Z}} = (\sigma_{\mathrm{Z},j_1j_2})$. Let $\boldsymbol{\Omega}_{\mathrm{Z}} = (\omega_{\mathrm{Z},j_1j_2}) = \boldsymbol{\Sigma}_{\mathrm{Z}}^{-1}$ be the precision matrix of $\mathbf{Z}_i$. We assume $\mathbf{Z}_i$ follows the nonparanormal distribution $\mathrm{NPN}(\boldsymbol{\Sigma}_{\mathrm{X}}, \mathbf{H})$, where $\boldsymbol{\Sigma}_{\mathrm{X}} = (\sigma_{\mathrm{X},j_1j_2})$, and $\mathbf{H}(\cdot) = (h_1(\cdot), \ldots, h_p(\cdot))^{\mathrm{T}}$ is a p-dimensional transformation with strictly monotone univariate functions $h_j(\cdot)$ for $j = 1, \ldots, p$. The nonparanormal model assumes that the transformed data $\mathbf{X}_i = \mathbf{H}(\mathbf{Z}_i) = (h_1(Z_{i1}), \ldots, h_p(Z_{ip}))^{\mathrm{T}}$ follows a p-dimensional multivariate normal distribution with mean $\mathbf{0}$ and covariance $\boldsymbol{\Sigma}_{\mathrm{X}}$ with diagonal elements $\sigma_{\mathrm{X},jj} = 1$ for all $j = 1, \ldots, p$. Let $\boldsymbol{\Omega}_{\mathrm{X}} = (\omega_{\mathrm{X},j_1j_2}) = \boldsymbol{\Sigma}_{\mathrm{X}}^{-1}$ be the precision matrix of $\mathbf{X}_i$, and $\mathbf{X} = (\mathbf{X}_1, \ldots, \mathbf{X}_n)^{\mathrm{T}}$ be the transformed data matrix. Let $\boldsymbol{\Psi}_{\mathrm{Z}} = (\rho_{\mathrm{Z},j_1j_2})$ and $\boldsymbol{\Psi}_{\mathrm{X}} = (\rho_{\mathrm{X},j_1j_2})$ be the partial correlation matrices of the observed observation $\mathbf{Z}_i$ and the normal transformed data $\mathbf{X}_i$, respectively, where $\rho_{\mathrm{Z},j_1j_2} = -\omega_{\mathrm{Z},j_1j_2}(\omega_{\mathrm{Z},j_1j_1}\omega_{\mathrm{Z},j_2j_2})^{-1/2}$ and $\rho_{\mathrm{X},j_1j_2} = -\omega_{\mathrm{X},j_1j_2}(\omega_{\mathrm{X},j_1j_1}\omega_{\mathrm{X},j_2j_2})^{-1/2}$ from the relationship between partial correlations and precision coefficients (Peng et al., 2009, Lemma 1).

Notice that, under the NPN model, $\sigma_{\mathrm{X},j_1j_2} = 0$ is equivalent to the j_1th and j_2th variables being marginally independent which implies $\sigma_{\mathrm{Z},j_1j_2} = 0$. But the other direction of this conclusion may not hold as $\sigma_{\mathrm{Z},j_1j_2}$ being zero does not guarantee the independence between the two variables. This indicates nonzero $\sigma_{\mathrm{Z},j_1j_2}$ implies $\sigma_{\mathrm{X},j_1j_2} \neq 0$, and the support of $\boldsymbol{\Sigma}_{\mathrm{X}}$ includes that of $\boldsymbol{\Sigma}_{\mathrm{Z}}$. Similarly, under the NPN model, $\omega_{\mathrm{X},j_1j_2} = 0$ is equivalent to the conditional independence of Z_{ij_1} and Z_{ij_2} given the rest of variables. However, this does not necessarily indicate $\omega_{\mathrm{Z},j_1j_2} = 0$ since the partial correlation between Z_{ij_1} and Z_{ij_2} only takes off the linear effects of all other variables. Nonlinear effects of $\mathbf{Z}_{-(j_1,j_2)}$ may still exist and make $\omega_{\mathrm{Z},j_1j_2}$ nonzero even if the j_1th and j_2th variables are conditionally independent. See Example 4 in Baba et al. (2004) as an illustration of this point by the lognormal distribution.

This means that conditional independence has no close ties with zero partial correlation of the original data $\{\mathbf{Z}_i\}$ except in the case of multivariate normal distributed $\mathbf{Z}_i$, and the partial correlation $\mathbf{\Psi}_{\mathrm{X}}$ of the transformed data is a more accurate measure for the conditional independence of $\mathbf{Z}_i$ since $\rho_{\mathrm{X},j_1j_2} = 0$ is equivalent to the conditional independence between those two variables under the NPN model. Therefore, the target parameters in this paper are all the elements ρ_{X,j_1j_2} in the high-dimensional NPN partial correlation matrix $\mathbf{\Psi}_{\mathrm{X}}$ of the latent variables $\mathbf{X}_i$. Particularly, we are interested in the multiple hypotheses

$$H_{0,j_1j_2} : \rho_{\mathrm{X},j_1j_2} = 0 \quad \text{vs.} \quad H_{a,j_1j_2} : \rho_{\mathrm{X},j_1j_2} \neq 0 \tag{2.1}$$

to identify the nonzero NPN partial correlations. Compared to the precision coefficient $\omega_{\mathrm{X},j_1j_2}$ under the NPN model, not only can the partial correlation ρ_{X,j_1j_2} show whether two variables are conditionally dependent or not, but also it provides the strength of the conditional dependence.

From Lemma 1 in Peng et al. (2009), the partial correlation ρ_{X,j_1j_2} can be obtained from p nodewise regressions on the transformed data $\mathbf{X}_i$ as

$$X_{ij_1} = \beta_{j_1,0} + \sum_{j_2 \neq j_1} \beta_{j_1,j_2} X_{ij_2} + \epsilon_{ij_1}, \tag{2.2}$$

where ϵ_{ij_1} and $\mathbf{X}_{i,-j_1}$ are uncorrelated for $j_1 = 1, \ldots, p$. Let $\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \ldots, \epsilon_{ip})^{\mathrm{T}}$ and $\mathbf{V} = (v_{j_1j_2}) = \mathrm{cov}(\boldsymbol{\epsilon}_i)$ be the covariance matrix of the regression error $\boldsymbol{\epsilon}_i$. It can be shown that

$$\rho_{\mathrm{X},j_1j_2} = -\frac{\omega_{\mathrm{X},j_1j_2}}{\sqrt{\omega_{\mathrm{X},j_1j_1}\omega_{\mathrm{X},j_2j_2}}} = -\frac{v_{j_1j_2}}{\sqrt{v_{j_1j_1}v_{j_2j_2}}} \quad \text{for} \quad j_1 \neq j_2. \tag{2.3}$$

In the following section, we propose an estimator for ρ_{X,j_1j_2} with an inference procedure based on the ranks of $\{Z_{ij}\}_{i=1}^n$ and the relationship (2.3).

3. Method

In this section, we introduce the regularized rank-based nodewise regression (RRNR) procedure for estimating the nonparanormal partial correlations and a multiple testing procedure for the hypotheses (2.1) with FDR control based on the proposed estimator.

3.1 *RRNR procedure*

The proposed RRNR procedure first estimates $\{X_{ij}\}_{i=1}^n$ by the ranks of $\{Z_{ij}\}_{i=1}^n$ and then applies a regularized regression method on (2.2) with the estimated $\mathbf{X}_i$ to estimate the nonparanormal partial correlation ρ_{X,j_1j_2}.

Let $\Phi(\cdot)$ and $\Phi^{-1}(\cdot)$ be the cumulative distribution function (cdf) and the quantile function of the standard normal distribution, respectively. It can be shown that $h_j(\cdot) = \Phi^{-1}\{F_j(\cdot)\}$ for $j = 1, \ldots, p$, where $F_j(\cdot)$ is the cdf of Z_{ij}. Note that $F_j(Z_{ij})$ follows the Uniform$(0, 1)$ distribution as Z_{ij} is a continuous random variable under the NPN model. Let r_{ij} denote the rank of Z_{ij} among the n observations $\{Z_{kj}\}_{k=1}^n$ for $i = 1, \ldots, n$ and $j = 1, \ldots, p$. It is natural to estimate $h_j(\cdot)$ by plugging in the winsorized empirical cdf $\widehat{F}_{n,j}(x)$ of $\{Z_{ij}\}_{i=1}^n$, where $\widehat{F}_{n,j}(x) = \min\{\widetilde{F}_{n,j}(x), 1 - 1/n^2\}$ is an estimate of $F_j(\cdot)$, and $\widetilde{F}_{n,j}(x) = \frac{1}{n}\sum_{k=1}^n \mathbb{I}(Z_{kj} \leqslant x)$ is the empirical cdf of $\{Z_{kj}\}_{k=1}^n$. Let

$$\widehat{X}_{ij} = \widehat{h}_j(Z_{ij}) = \Phi^{-1}\{\widehat{F}_{n,j}(Z_{ij})\} = \begin{cases} \Phi^{-1}(r_{ij}/n) & , \quad r_{ij} < n; \\ \Phi^{-1}(1 - 1/n^2), & r_{ij} = n \end{cases} \tag{3.4}$$

be the estimated X_{ij} based on the rank of Z_{ij}, where the last equality is due to $\widehat{F}_{n,j}(Z_{ij}) = r_{ij}/n$. The use of the winsorization $\Phi^{-1}(1-1/n^2)$ is to avoid the transformation being infinity for the maximum value of $\{Z_{ij}\}_{i=1}^n$. Similar estimator of X_{ij} was considered in Liu et al. (2009) with a winsorization of $\{r_{ij}/n\}$ at the level $n^{-1/4}(\log n)^{-1/2}$ and $1 - n^{-1/4}(\log n)^{-1/2}$ for the lower and upper empirical quantiles, respectively. Mai and Zou (2015) also used the winsorized estimator in discriminant analysis under the NPN model.

Let $\widehat{\mathbf{X}}_i = (\widehat{X}_{i1}, \ldots, \widehat{X}_{ip})^{\mathrm{T}}$ for $i = 1, \ldots, n$, and $\widehat{\mathbf{X}} = (\widehat{\mathbf{X}}_1, \ldots, \widehat{\mathbf{X}}_n)^{\mathrm{T}}$. We fit the nodewise regressions (2.2) with the empirically transformed data $\{\widehat{\mathbf{X}}_i\}$ by lasso (Tibshirani, 1996). Note that $\sum_{i=1}^n \widehat{X}_{ij} = \Phi^{-1}(1 - 1/n^2)$ for all $1 \leqslant j \leqslant p$. Let $d_n = \Phi^{-1}(1 - 1/n^2)/n$, which is the average of the empirically transformed data $\{\widehat{X}_{ij}\}_{i=1}^n$ for each variable. Let $\mathbf{D}_n = d_n\mathbf{1}_p$,

where $\mathbf{1}_p$ is a p-dimensional vector of 1. Since d_n is at the order $\sqrt{\log(n)}n^{-1}$, $\mathbf{D}_n$ is negligible in our theoretical analysis.

Let $\boldsymbol{\beta}_j = (\beta_{j,1}, \ldots, \beta_{j,j-1}, -1, \beta_{j,j+1}, \ldots, \beta_{j,p})^{\mathrm{T}}$ be the coefficients of regressing X_{ij} on the rest of variables in (2.2) with $\beta_{j,j} = -1$. We estimate $\boldsymbol{\beta}_j$ by lasso as

$$\widehat{\boldsymbol{\beta}}_j = \underset{\boldsymbol{\beta}_j,\ \beta_{j,j}=-1}{\arg\min} \left[\frac{1}{n} \sum_{i=1}^{n} \{\boldsymbol{\beta}_j^{\mathrm{T}} (\widehat{\mathbf{X}}_i - \mathbf{D}_n)\}^2 + \lambda_j \sum_{1 \leqslant k \neq j \leqslant p} |\beta_{j,k}| \right] \tag{3.5}$$

for a penalty parameter λ_j. Let $\widehat{\boldsymbol{\epsilon}}_i = (\widehat{\epsilon}_{i1}, \ldots, \widehat{\epsilon}_{ip})^{\mathrm{T}}$ be the residuals of the ith observation from the p nodewise regressions for $i = 1, \ldots, n$, where $\widehat{\epsilon}_{ij} = -\widehat{\boldsymbol{\beta}}_j^{\mathrm{T}} (\widehat{\mathbf{X}}_i - \mathbf{D}_n)$. Similar to Liu (2013); Qiu and Zhou (2020), we construct the bias corrected estimator

$$\widehat{v}_{j_1 j_2} = -\frac{1}{n} \sum_{i=1}^{n} (\widehat{\epsilon}_{ij_1} \widehat{\epsilon}_{ij_2} + \widehat{\beta}_{j_1,j_2} \widehat{\epsilon}_{ij_2}^{\,2} + \widehat{\beta}_{j_2,j_1} \widehat{\epsilon}_{ij_1}^{\,2}) \tag{3.6}$$

for $1 \leqslant j_1, j_2 \leqslant p$ to estimate the covariance matrix $\mathbf{V} = \mathrm{cov}(\boldsymbol{\epsilon}_i)$. From the relationship (2.3), the proposed RRNR estimator for the nonparanormal partial correlation $\rho_{\mathrm{X},j_1 j_2}$ is

$$\widehat{\rho}_{\mathrm{X},j_1 j_2} = -\frac{\widehat{v}_{j_1 j_2}}{(\widehat{v}_{j_1 j_1} \widehat{v}_{j_2 j_2})^{1/2}} \quad \text{for } j_1 \neq j_2. \tag{3.7}$$

The first advantage of the proposed estimator is the convexity and the normally distributed covariates $\{\widehat{X}_{ij}\}$ of the lasso program (3.5) for the nodewise regressions, which can be efficiently solved by existing algorithms like coordinate descent and least angle regression. Once the nodewise regressions are fitted, the computation of $\widehat{\rho}_{\mathrm{X},j_1 j_2}$ is only based on simple arithmetic calculations. The computation complexity of estimating all partial correlations is at the order np^2. Secondly, from Theorem 1 in Section 4, we have that $\sqrt{n}\widehat{\rho}_{\mathrm{X},j_1 j_2}$ is asymptotically normal with mean 0 and variance 1 under some regularity conditions and the null hypothesis $\rho_{\mathrm{X},j_1 j_2} = 0$. Therefore, a multiple testing procedure for the hypotheses (2.1) can be constructed only based on $\{\widehat{\rho}_{\mathrm{X},j_1 j_2}\}$, which does not require the estimation of additional parameters. This is introduced in the next subsection.

3.2 *Multiple testing procedure*

We construct a multiple testing procedure for ρ_{x,j_1j_2} based on the asymptotic normality of $\widehat{\rho}_{\mathrm{x},j_1j_2}$. Specifically, let

$$\mathcal{T}_0(s) \;=\; \big\{(j_1,j_2) : |\widehat{\rho}_{\mathrm{x},j_1j_2}| > s(1-\breve{\rho}^{\,2}_{\mathrm{x},j_1j_2})\{\log(p)/n\}^{1/2}\big\} \tag{3.8}$$

be the set to recover the nonzero ρ_{x,j_1j_2}, where $\breve{\rho}_{\mathrm{x},j_1j_2} = \widehat{\rho}_{\mathrm{x},j_1j_2}\mathbb{I}[|\widehat{\rho}_{\mathrm{x},j_1j_2}| > 2\{\log(p)/n\}^{1/2}]$ is the thresholding estimator of ρ_{x,j_1j_2}.

To choose the cut-off value s in $\mathcal{T}_0(s)$, we control the FDR of the hypotheses (2.1) at a pre-specified value α. Note that FDR is the expectation of false discovery proportion (FDP), which is defined as the number of false positives over the number of discoveries. Let $\mathcal{S}_0 = \{(j_1,j_2) : \rho_{\mathrm{x},j_1j_2} = 0\}$ be the set of zero nonparanormal partial correlations. For any set $\mathcal{A}$, let $|\mathcal{A}|$ be the cardinality of $\mathcal{A}$. The FDP of $\mathcal{T}_0(s)$ in (3.8) can be expressed as

$$\mathrm{FDP}_0(s) = \frac{|\mathcal{T}_0(s)\cap\mathcal{S}_0|}{\max\{1,|\mathcal{T}_0(s)|\}},$$

where the numerator can be approximated by its expectation $\sum_{(j_1,j_2)\in\mathcal{S}_0}\mathrm{E}(\mathbb{I}[|\widehat{\rho}_{\mathrm{x},j_1j_2}| > s(1-\breve{\rho}^{\,2}_{\mathrm{x},j_1j_2})\{\log(p)/n\}^{1/2}])$ which is bounded by $2\{1-\Phi(s\sqrt{\log p})\}(p^2-p)$ asymptotically. Therefore, to control the FDR of $\mathcal{T}_0(s)$ at $\alpha\in(0,1)$, we choose

$$s_0^* = \inf\left\{s\in(0,2] : \frac{2\{1-\Phi(s\sqrt{\log p})\}(p^2-p)}{\max\{1,|\mathcal{T}_0(s)|\}} \leqslant \alpha\right\}. \tag{3.9}$$

If s_0^* does not exist, set $s_0^* = 2$. The proposed multiple testing procedure rejects the null hypothesis H_{0,j_1j_2} in (2.1) if $(j_1,j_2)\in\mathcal{T}_0(s_0^*)$. Since $\{1-\Phi(2\sqrt{\log p})\}(p^2-p)\to 0$ as $p\to\infty$, there is no need to search the cut-off value s in (3.9) beyond 2.

3.3 *Connections to the existing methods*

Under the high-dimensional NPN model, the existing methods (Liu et al., 2012; Xue and Zou, 2012) estimate the precision matrix $\mathbf{\Omega}_{\mathrm{x}}$ via the trigonometric relationship $\sigma_{\mathrm{x},j_1j_2} = \sin(\pi\tau_{\mathrm{z},j_1j_2}/2)$ between the covariance of the transformed data and the Kendall's τ of the original data, where τ_{z,j_1j_2} denotes the population Kendall's τ coefficient between Z_{ij_1} and

Z_{ij_2}. Let $\widetilde{\tau}_{\mathrm{z},j_1j_2}$ be the sample Kendall's τ of τ_{z,j_1j_2}, $\widetilde{\sigma}_{\mathrm{x},j_1j_2} = \sin(\pi\widetilde{\tau}_{\mathrm{z},j_1j_2}/2)$ be the trigonometric estimator of $\sigma_{\mathrm{x},j_1j_2}$, and $\widetilde{\mathbf{\Sigma}}_{\mathrm{x}} = (\widetilde{\sigma}_{\mathrm{x},j_1j_2})$. Particularly, recent works (Gu et al., 2015; Xu et al., 2016; He et al., 2017; Barber and Kolar, 2018) developed inference procedures for each of the precision coefficient $\omega_{\mathrm{x},j_1j_2}$ based on $\widetilde{\mathbf{\Sigma}}_{\mathrm{x}}$.

Since $\widetilde{\mathbf{\Sigma}}_{\mathrm{x}}$ may not be positive semi-definite which leads to non-convex optimization, either the restricted lasso estimator (Barber and Kolar, 2018) or the Dantzig selector type estimator (Gu et al., 2015; Xu et al., 2016; He et al., 2017) is used to estimate $\mathbf{\Omega}_{\mathrm{x}}$ based on $\widetilde{\mathbf{\Sigma}}_{\mathrm{x}}$. The restricted lasso estimation requires an additional tuning parameter that bounds the ℓ_1 norm of the estimated $\mathbf{\Omega}_{\mathrm{x}}$. However, there is no practical guideline to choose this constraint parameter and an inappropriate value may lead to unstable results due to the non-convex optimization. Although the Dantzig selector does not have the non-convexity issue, it is computationally slower than Lasso type methods, especially for large p cases. As a comparison, the proposed method only requires the classical convex lasso estimation (3.5) on the nodewise regressions of the rank-transformed data $\widehat{\mathbf{X}}_i$, which can be solved efficiently.

Moreover, the debiased estimators of $\omega_{\mathrm{x},j_1j_2}$ in Gu et al. (2015); Xu et al. (2016); Barber and Kolar (2018) are constructed by formulations related to the matrix product $\widetilde{\mathbf{\Omega}}_{\mathrm{x}}\widetilde{\mathbf{\Sigma}}_{\mathrm{x}}\widetilde{\mathbf{\Omega}}_{\mathrm{x}}$, where $\widetilde{\mathbf{\Omega}}_{\mathrm{x}}$ is the regularized inverse of $\widetilde{\mathbf{\Sigma}}_{\mathrm{x}}$ by the restricted lasso method or the Dantzig selector. But the variances of such debiased estimators are much more complex than that of the proposed partial correlation estimator. Testing for each $\omega_{\mathrm{x},j_1j_2}$ requires calculating quadratic products of a $p \times p$ matrix that is constructed by using all pairs $\{\mathbf{Z}_{i_1}, \mathbf{Z}_{i_2}\}$ of the original data. Given a regularized estimate of $\mathbf{\Omega}_{\mathrm{x}}$, the computation complexity of those methods for the inference of all precision coefficients is at least at the order n^2p^4, which is higher than the computation complexity np^2 of the proposed method. The computation times of those methods in the simulation study are reported in Table 1, which verifies this point.

Despite of the fast computation of the proposed method, it requires a slightly more restrictive sparsity condition on the maximum number of nonzero elements in each row of $\mathbf{\Omega}_{\mathrm{X}}$ by a factor of $n^{-2\delta_0}(\log p)^{-1}$ compared with that in Gu et al. (2015); Xu et al. (2016); Barber and Kolar (2018), where δ_0 is an arbitrarily small positive constant. The extra condition is due to controlling the variation of estimating the distribution function $F_j(x)$ by the empirical cdf $\widetilde{F}_{n,j}(x)$. More detailed discussion on the theoretical results is provided in the next section.

4. Theory

In this section, we show the consistency of the rank-based lasso estimator $\widehat{\boldsymbol{\beta}}_j$ in (3.5), and derive the asymptotic distribution of the proposed RRNR estimator $\widehat{\rho}_{\mathrm{X},j_1j_2}$ in (3.7). Based on these results, we show the FDR control of the proposed multiple testing procedure (3.9) for the hypotheses (2.1).

Let $\lambda_p \leqslant \ldots \leqslant \lambda_1$ be the eigenvalues of the NPN precision matrix $\mathbf{\Omega}_{\mathrm{X}}$. Recall that $\mathbf{\Omega}_{\mathrm{X},(j)}$ denotes the jth column of $\mathbf{\Omega}_{\mathrm{X}}$. Let $s_0 = \max_{1\leqslant j\leqslant p}|\mathbf{\Omega}_{\mathrm{X},(j)}|_0$ and $s_1 = \max_{1\leqslant j\leqslant p}|\mathbf{\Omega}_{\mathrm{X},(j)}|_1$ be the maximum number of nonzero elements in each column of $\mathbf{\Omega}_{\mathrm{X}}$ and its matrix ℓ_1 norm, respectively. Let C be a positive constant which may change under different occasions. We make the following assumptions on the sparsity and eigenvalues of $\mathbf{\Omega}_{\mathrm{X}}$.

CONDITION 1: Assume $s_0 = o\{n^{1/2-\delta_0}(\log p)^{-1}\}$ for an arbitrarily small constant $\delta_0 > 0$.

CONDITION 2: There exists a positive constant C, such that $C^{-1} \leqslant \lambda_p \leqslant \lambda_1 \leqslant C$.

Condition 1 regulates the sparsity of $\mathbf{\Omega}_{\mathrm{X}}$, which is needed for establishing the consistency of $\widehat{\boldsymbol{\beta}}_j$ from the nodewise regression (3.5) of $\widehat{\mathbf{X}}$. A slightly weaker condition $\max_{1\leqslant j\leqslant p}|\mathbf{\Omega}_{\mathrm{Z},(j)}|_0 = o\{n^{1/2}(\log p)^{-3/2}\}$ is required for the high-dimensional nodewise regression of the observed data $\mathbf{Z}$ and the statistical inference for its precision matrix $\mathbf{\Omega}_{\mathrm{Z}}$ and partial correlation matrix $\mathbf{\Psi}_{\mathrm{Z}}$ (Liu, 2013; Chang et al., 2018; Qiu and Zhou, 2020). We need a more restrictive sparsity condition for testing each element of $\mathbf{\Psi}_{\mathrm{X}}$ due to the unobservable $\mathbf{X}_i$ under the NPN model

and the estimation of the distribution function $F_j(x)$ for each variable. Condition 2 requires the eigenvalues of $\mathbf{\Omega}_{\mathrm{X}}$ being bounded away from 0 and infinity, which is commonly assumed in the literature on estimating high-dimensional covariance and precision matrices (Bickel and Levina, 2008). This condition also implies that $\max_{1\leqslant j_1<j_2\leqslant p}|\rho_{\mathrm{X},j_1j_2}|\leqslant 1-c_0$ for a small positive constant c_0.

Let $s_{0,j}=|\boldsymbol{\beta}_j|_0$ and $s_{1,j}=|\boldsymbol{\beta}_j|_1$. Notice that $|\boldsymbol{\beta}_j|_0=|\mathbf{\Omega}_{\mathrm{X},(j)}|_0$, $s_{0,j}\leqslant s_0$ and $s_{1,j}\leqslant Cs_1$ for $j=1,\ldots,p$. We first establish the consistency of the rank-based lasso estimator. The following proposition gives a uniform bound for the ℓ_1 and ℓ_2 distance between $\boldsymbol{\beta}_j$ and its estimator $\widehat{\boldsymbol{\beta}}_j$ for all j. The statistical property of the partial correlation estimator $\widehat{\rho}_{\mathrm{X},j_1j_2}$ can be derived based on this result.

PROPOSITION 1: Under Conditions 1 and 2, and a penalty parameter λ_j in (3.5) at the order $s_{1,j}(\log p)n^{-1/2+\delta}$ for any $\delta\in(0,1/3)$, we have

$$\max_{1\leqslant j\leqslant p}|\widehat{\boldsymbol{\beta}}_j-\boldsymbol{\beta}_j|_1\leqslant C\lambda_j s_{0,j}\quad\text{and}\quad\max_{1\leqslant j\leqslant p}|\widehat{\boldsymbol{\beta}}_j-\boldsymbol{\beta}_j|_2\leqslant C\lambda_j s_{0,j}^{1/2}$$

with probability $1-p^{-c}$ for a positive constant c.

Proposition 1 provides the convergence rates of $\widehat{\boldsymbol{\beta}}_j$ to $\boldsymbol{\beta}_j$. It extends the classical theoretical results of lasso (Bühlmann and Van De Geer, 2011) with observed covariates and response to the case with estimated covariates and response by the empirical ranks of the data. Compared to the classical lasso with ℓ_1 and ℓ_2 convergence rates at $s_{0,j}\{\log(p)/n\}^{1/2}$ and $\{s_{0,j}\log(p)/n\}^{1/2}$, the proposed rank-based estimator $\widehat{\boldsymbol{\beta}}_j$ has slower convergence rates due to the use of a larger penalty level λ_j to control the additional variation caused by estimating the distribution functions $\{F_j(x)\}$. Notice that the bound for $|(\widehat{\mathbf{X}}_{(j)}-\widehat{\mathbf{X}}_{-j}\boldsymbol{\beta}_j)^{\mathrm{T}}\widehat{\mathbf{X}}_{-j}/n|_\infty$ is increased by replacing $\mathbf{X}$ with $\widehat{\mathbf{X}}$, even though $|\boldsymbol{\epsilon}_{(j)}^{\mathrm{T}}\mathbf{X}_{-j}/n|_\infty=O_p(\{\log(p)/n\}^{1/2})$, where $\boldsymbol{\epsilon}_{(j)}=(\epsilon_{1j},\ldots,\epsilon_{nj})^{\mathrm{T}}$ are the regression errors from the jth nodewise regression in (2.2). Under the condition of bounded ℓ_1 norm of $\mathbf{\Omega}_{\mathrm{X}}$ which is assumed in Gu et al. (2015); Xu

et al. (2016); He et al. (2017) for testing coefficients of $\mathbf{\Omega}_{\mathrm{X}}$, the ℓ_1 and ℓ_2 convergence rates of $\widehat{\boldsymbol{\beta}}_j$ are $s_{0,j}\log(p)n^{-1/2+\delta}$ and $s_{0,j}^{1/2}\log(p)n^{-1/2+\delta}$, respectively. Those rates are inferior to the classical lasso by a factor $(\log p)^{1/2}n^{\delta}$, where δ is an arbitrarily small positive constant. Barber and Kolar (2018) also considered testing for each component of $\mathbf{\Omega}_{\mathrm{X}}$ via an approach related to pairwise regression for precision matrices (Ren et al., 2015). The estimated coefficients from the pairwise regression enjoy similar convergence rates of the classical lasso. However, as discussed in Section 3.3, their approach encounters non-convexity issues in the regularized regression and requires heavier computation.

Recall that $\mathcal{S}_0 = \{(j_1, j_2) : \rho_{\mathrm{X},j_1j_2} = 0\}$ is the set of zero nonparanormal partial correlations. The following theorem establishes the asymptotic expansion of the estimator $\widehat{v}_{j_1j_2}$ for the covariance between the nodewise regression errors $\{\epsilon_{ij}\}$ over $\mathcal{S}_0$, which implies the asymptotic normality of $\widehat{\rho}_{\mathrm{X},j_1j_2}$.

THEOREM 1: *Under the conditions in Proposition 1,* $p \leqslant n^{\xi}$ *for a positive constant* ξ*, and* $s_0 s_1^2 = o\{n^{1/2-2\delta_0}(\log p)^{-2}\}$ *for an arbitrarily small positive constant* δ_0*, we have* $\max_{j_1\neq j_2}|\widehat{\rho}_{X,j_1j_2} - \rho_{X,j_1j_2}| \to 0$ *as* $n, p \to \infty$*. Particularly,* $\widehat{v}_{j_1j_2} = -\sum_{i=1}^{n}\epsilon_{ij_1}\epsilon_{ij_2}/n + o_p(n^{-1/2})$ *for all* $(j_1, j_2) \in \mathcal{S}_0$*, where the small order term* $o_p(n^{-1/2})$ *is uniform over* $\mathcal{S}_0$*. Moreover,* $\sqrt{n}\widehat{\rho}_{X,j_1j_2} \to N(0,1)$ *as* $n, p \to \infty$ *for all* $(j_1, j_2) \in \mathcal{S}_0$*.*

This asymptotic normality result in Theorem 1 is the foundation of the proposed multiple testing procedure. As discussed after Condition 2, Theorem 1 requires a more strict sparsity condition on s_0 than $\max_{1\leqslant j\leqslant p}|\mathbf{\Omega}_{\mathrm{Z},(j)}|_0 = o\{n^{1/2}(\log p)^{-3/2}\}$ in estimation and inference of $\omega_{\mathrm{Z},j_1j_2}$ (Liu, 2013). If the ℓ_1 norm of $\mathbf{\Omega}_{\mathrm{X}}$ is bounded, Theorem 1 prescribes s_0 being at a smaller order of $n^{1/2-2\delta_0}(\log p)^{-2}$ for a small positive constant δ_0, which is milder than the sparsity conditions ($s_0 = o\{n^{1/4}(\log p)^{-1/2}\}$) made in Theorem 4.5 in He et al. (2017). However, our sparsity condition is stronger than that ($s_0 = o\{n^{1/2}(\log p)^{-1}\}$) made in Gu et al. (2015); Barber and Kolar (2018) for the inference of $\omega_{\mathrm{X},j_1j_2}$. The more restrictive condition on s_0 by

a factor $n^{-2\delta_0}(\log p)^{-1}$ is caused by controlling the difference between $\mathbf{X}_i$ and its estimate $\hat{\mathbf{X}}_i$ from the empirical ranks. In general, as $s_1 \leqslant C\sqrt{s_0}$ for a positive constant C, a sufficient condition for $s_0 s_1^2 = o\{n^{1/2-2\delta_0}(\log p)^{-2}\}$ is $s_0 = o\{n^{1/4-\delta_0}(\log p)^{-1}\}$. This requires the number of nonzero elements in each row of $\mathbf{\Omega}_{\mathrm{x}}$ being at a smaller order of $n^{1/4}$.

Let $\overline{\mathcal{S}}_0$ be the compliment of $\mathcal{S}_0$, which is the set of signals in $\mathbf{\Psi}_{\mathrm{x}}$. Let $\mathrm{FDP}_0(s_0^*) = |\mathcal{T}_0(s_0^*) \cap \mathcal{S}_0| / \max\{1, |\mathcal{T}_0(s_0^*)|\}$ be the FDP of the proposed multiple testing procedure for the hypotheses (2.1), where $\mathcal{T}_0(s)$ and s_0^* are given in (3.8) and (3.9). To study the type I error of the proposed multiple testing procedure, we make an additional technical assumption.

CONDITION 3: Let $\mathcal{M} = \{(j_1, j_2) : |\rho_{\mathrm{x},j_1 j_2}| > \widetilde{C}_1(\log p)n^{-1/2+\delta}\}$ for a small positive constant δ. Assume $|\mathcal{M}| > \widetilde{C}_2\sqrt{\log p}$ for positive constants $\widetilde{C}_1$ and $\widetilde{C}_2$.

Condition 3 is a mild condition on the NPN partial correlation matrix $\mathbf{\Psi}_{\mathrm{x}}$, which requires at least $\widetilde{C}_2\sqrt{\log p}$ elements in $\mathbf{\Psi}_{\mathrm{x}}$ with absolute values larger than $\widetilde{C}_1(\log p)n^{-1/2+\delta}$, where δ is a small positive constant. It guarantees the number of significant nonzero $\rho_{\mathrm{x},j_1 j_2}$ by the proposed procedure (3.9) is at least of the order $\sqrt{\log p}$. The following theorem presents the FDR control of the significant set $\mathcal{T}_0(s_0^*)$ for the hypotheses (2.1).

THEOREM 2: *Under the conditions in Theorem 1, Condition 3,* $s_0 \leqslant Cp(\log p)^{-1-\eta_1}$ *for positive constants* C *and* η_1*, we have that* $\lim_{n,p\to\infty} \mathbb{P}(\mathrm{FDP}_0(s_0^*) \leqslant \alpha + \varepsilon) = 1$ *for any* $\varepsilon > 0$.

Theorem 2 shows that the proposed multiple testing procedure can control the FDR at the nominal level α. The condition $s_0 \leqslant Cp(\log p)^{-1-\eta_1}$ implies that the number of nonzero partial correlations in each row of $\mathbf{\Psi}_{\mathrm{x}}$ is less than the order $p/(\log p)^{1+\eta_1}$ for a positive constant η_1. It is satisfied if $\sqrt{n} \leqslant Cp$ under Condition 1. Let $\mathcal{N} = \{(j_1, j_2) : \rho_{\mathrm{x},j_1 j_2} \geqslant (\log p)^{-1-\eta_2}\}$ be the set of relatively large NPN partial correlations (larger than a multi-$\log p$ term that converges to zero), where η_2 is a positive constant. This condition leads to a bound on the size of $\mathcal{N}$ such that $|\mathcal{N}| \leqslant Cp^2/(\log p)^{1+\eta_1}$, which is required in Cai and Liu (2016) for multiple

testing of marginal correlations. Since $\text{corr}(\epsilon_{ij_1}\epsilon_{ij_2}, \epsilon_{ik_1}\epsilon_{ik_2}) = \rho_{\text{X},j_1k_1}\rho_{\text{X},j_2k_2} + \rho_{\text{X},j_1k_2}\rho_{\text{X},j_2k_1}$ for the pairs with $\rho_{\text{X},j_1j_2} = \rho_{\text{X},k_1k_2} = 0$, the bound on $|\mathcal{N}|$ is used to control the number of relatively large correlations among pairs of the estimated partial correlations $\{\widehat{\rho}_{\text{X},j_1j_2}\}$.

5. Simulation

In this section, we evaluate the performance of the proposed method (RRNR) in terms of the ROC curve of the true positive rate versus the false positive rate, and its empirical FDR and power for testing the hypotheses (2.1). We also compare the proposed method with the existing methods in Liu (2013) (Liu), Gu et al. (2015) (GCNL), Xu et al. (2016) (XTG), He et al. (2017) (CFC), Barber and Kolar (2018) (ROCKET) and Qiu and Zhou (2020) (QZ). Note that GCNL, XTG, CFC and ROCKET are designed for the precision coefficients $\{\omega_{\text{X},j_1j_2}\}$ under either the nonparanormal model or the transelliptical model, and Liu and QZ target on testing for the precision coefficients $\{\omega_{\text{Z},j_1j_2}\}$ and the partial correlations $\{\rho_{\text{Z},j_1j_2}\}$ of the original data, respectively. Particularly, XTG builds a distributed algorithm for estimating $\mathbf{\Omega}_{\text{X}}$. In our simulation, we only adopt the inference method of XTG without using its distributed computation algorithm. As $\rho_{\text{X},j_1j_2} = 0$ is equivalent to $\omega_{\text{X},j_1j_2} = 0$, we consider the hypotheses (2.1) to compare those seven methods in simulation. Also note that multiple testing procedures are proposed in Liu, CFC and QZ approaches, but GCNL, XTG and ROCKET only consider the inference for each element of the precision matrix $\mathbf{\Omega}_{\text{X}}$ without a FDR control procedure. Therefore, we first present the ROC curves for comparing all seven methods and then report the empirical sizes and powers of RRNR, Liu, QZ and CFC when controlling the FDR at the nominal level 5%.

To generate the NPN distributed data $\mathbf{Z}_i$, we first generate i.i.d. $\mathbf{X}_i = (X_{i1}, \ldots, X_{ip})^{\text{T}}$ from the multivariate normal distribution with mean $\mathbf{0}$ and covariance $\mathbf{\Sigma}_{\text{X}} = (\sigma_{\text{X},j_1j_2})$. Then, let $Z_{ij} = h_j^{-1}(X_{ij})$ by applying the inverse functions of the transformation $\mathbf{H} = (h_1(\cdot), \ldots, h_p(\cdot))^{\text{T}}$ on $\mathbf{X}_i$. Two covariance structures for $\mathbf{\Sigma}_{\text{X}}$ are considered: (i) AR covariance

with $\sigma_{\mathrm{x},j_1j_2} = 0.6^{|j_1-j_2|}$, and (ii) block-diagonal (BD) covariance with block size 4×4 and $\sigma_{\mathrm{x},j_1j_2} = \mathbb{I}(j_1 = j_2) + \sum_{k=1}^{p/4} \rho_k \mathbb{I}(4(k-1)+1 \leqslant j_1 \neq j_2 \leqslant 4k)$, where ρ_k is randomly chosen from the uniform distribution on $(0.3, 0.9)$ for $k = 1, \ldots, p/4$. Notice that under the AR covariance, ρ_{x,j_1j_2} is nonzero only on the main diagonal and the first off-diagonal. Under the BD covariance, the partial correlation matrix $\mathbf{\Psi}_{\mathrm{x}}$ is also block diagonal with size 4. Once a covariance is generated, it is kept fixed throughout all 1000 repetitions of each simulation setting. There are 40 different settings in our simulation, which are composed of different sample sizes, dimensions, covariance structures and data transformations. We consider log and cubic root (CR) transformations on all variables and on the odd variables only, as well as the identity transformation, which leads to 5 types of data transformation $\mathbf{H}$ listed as follows.

(1) Identity: $h_j(x) = x$ such that $Z_{ij} = X_{ij}$ for all $j = 1, \ldots, p$;

(2) CR–All: $h_j(x) = x^{1/3}$ such that $Z_{ij} = X_{ij}^3$ for all $j = 1, \ldots, p$;

(3) CR–Odd: $Z_{ij} = X_{ij}^3$ for odd j and $Z_{ij} = X_{ij}$ for even j;

(4) Log–All: $h_j(x) = \log(x)$ such that $Z_{ij} = \exp(X_{ij})$ for all $j = 1, \ldots, p$;

(5) Log–Odd: $Z_{ij} = \exp(X_{ij})$ for odd j and $Z_{ij} = X_{ij}$ for even j.

We set $n = 60, 100$ and $p = 100, 400$. All the scenarios are repeated 1000 times.

Let the true positive rate (TPR) and the false positive rate (FPR) of a multiple testing procedure for the hypotheses (2.1) be the number of true positives over the number of nonzero ρ_{x,j_1j_2} and the number of false positives over the number of zero ρ_{x,j_1j_2}, respectively. To obtain the ROC curve for testing the hypotheses (2.1), in each repetition, we first calculate the TPR and FPR at a series of cut-off values on the p values of testing $\rho_{\mathrm{x},j_1j_2} = 0$ for the proposed method RRNR, on the p values of testing $\omega_{\mathrm{x},j_1j_2} = 0$ for the GCNL, XTG, ROCKET and CFC procedures, and on the p values of testing $\omega_{\mathrm{z},j_1j_2} = 0$ for the Liu and QZ methods. Then, to summarize the ROC curves over all repetitions, we bin FPR from 0 to 0.2 by an

increment 0.01 and calculate the average maximum TPR with FPR smaller than $k \times 0.01$ over the 1000 repetitions for $k = 1, \ldots, 20$. Finally, we plot the average TPR versus the sequence $\{k \times 0.01 : k = 1, \ldots, 20\}$ of FPR. We also calculate the ROC curve of the nodewise regression approach (3.5)–(3.7) applied on the true transformation, denoted by True, which serves as an oracle benchmark. The average ROC curves under $p = 100$ are shown in Figures 1. The ROC curves under $p = 400$ are similar to those in the case of $p = 100$, which are reported in Web Figure 1 in the Supporting Information.

[Figure 1 about here.]

From those two figures, we observe that the proposed RRNR method has the highest ROC curve among the other six methods, CFC, ROCKET, QZ, Liu, GCNL and XTG, under almost all the scenarios. It is comparable to the oracle method constructed based on the true transformation. This shows the effectiveness of the RRNR procedure on estimating the nonparanormal partial correlations without knowing the true transformation. The performance of RRNR is consistent over the five transformations considered in the simulation. This demonstrates the robustness of the proposed method to monotone transformations. Particularly, RRNR is comparable to Liu and QZ under the identity transformation, where the original data follow the Gaussian distribution. Note that Liu and QZ estimate the precision coefficient and the partial correlation of the original data which provide valid tests for the hypotheses (2.1) under this case. Compared to the procedures designed for Gaussian data, the proposed rank-based approach does not suffer power loss in testing for conditional independence under the Gaussian distribution. The performances of CFC, GCNL and XTG are comparable to RRNR under the AR covariance, but they suffer some power loss under the block-diagonal covariance. Also, the difference between RRNR and GCNL and that between RRNR and XTG in the ROC curve are not large. However, the latter two methods are much more computationally intensive. Their computation times are listed in Table 1. Notice that

Liu and QZ have higher ROC curves under the settings with transformations applied on odd variables only than those with transformations on all variables. This indicates Liu and QZ lose power in detecting nonlinear associations. ROCKET has the lowest ROC curve, but its performance improves when the sample size is increased to 100, which implies that ROCKET may need a much larger sample size to obtain an accurate estimate for $\mathbf{\Omega}_{\mathrm{X}}$.

The computation times of the aforementioned methods for the ROC curves are reported in Table 1, which are assessed by a SuperMicro server with two 8-core Intel Haswell processors (2.6 GHz). We see that the proposed method is much faster than the other four methods, CFC, ROCKET, XTG and GCNL, for the inference of nonparanormal graphical models. This is due to the easy implementation of the lasso estimation on the rank transformed data compared to the Dantzig type estimators. This result verifies the computational efficiency of the proposed method discussed in Section 3.3. ROCKET is faster than XTG and GCNL, since it uses the restricted lasso program to estimate the NPN precision matrix. GCNL under $p = 400$ is not computed as it requires calculating a matrix with dimensions at the order $p^2 \times p^2$, which demands a large memory. Also note that QZ has a longer computation time than Liu because QZ uses the scaled lasso program, which is slower than the classical lasso program.

[Table 1 about here.]

To evaluate the FDR control of the proposed multiple testing procedure, Table 2 reports the empirical FDR and average power (TPR) of RRNR, QZ, Liu and CFC for the hypotheses (2.1) under the nominal FDR level 0.05. ROCKET, GCNL and XTG are not considered here as they do not offer a multiple testing procedure. From Table 2, we see that the proposed procedure RRNR can control the FDR around the nominal level with high power under all the cases, although it is slightly liberal under the AR covariance structure. The power of RRNR increases as the increase of n, and it almost reaches 1 under the AR covariance and

$n = 100$. Similar to the ROC curve, the sizes and powers of RRNR are quite consistent across different transformations under each combination of covariance structure, sample size and dimension. This is due to the ranks of the data remain unchanged under monotone transformations. CFC is also able to control the FDR at 5%. Although the power of CFC is slightly inferior to that of RRNR under the AR covariance, their overall performances are comparable under this setting. However, RRNR has a much higher power than CFC under the block-diagonal covariance. Liu and QZ can control the FDR with a high power for the identity transformation. However, both methods fail for the non-identity transformations, as they are not designed for estimating nonparanormal partial correlations. Meanwhile, for the identity transformation, the power of RRNR is comparable to those of Liu and QZ under the block-diagonal covariance and is slightly lower than those of Liu and QZ under the AR covariance. This manifests that the proposed method does not suffer power loss if the original data are Gaussian distributed.

[Table 2 about here.]

6. Real data analysis

In this section, we demonstrate the utility of the proposed procedure on an FDG-PET brain image data set (Huang et al., 2010), which was initially obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI). The data contain the readings of 42 brain anatomical volumes of interest (AVOI) from 49 Alzheimer's disease patients (AD) and 67 normal control subjects (NC). These 42 selected AVOIs are distributed in 4 brain regions that are identified as the most affected ones by Alzheimer's disease: prefrontal, parietal, occipital and temporal lobes. The Shapiro-Wilk normality test is applied on each of the 42 variables for the AD and NC groups, and the p value histograms of the two groups are reported in Web Figure 4 in the Supporting Information. From the histograms, we see that there are many small

p values for both groups, indicating non-normal distributions for the readings from some brain AVOIs. As the data may not follow the Guassian distribution, we apply the proposed RRNR procedure to recover the nonzero NPN partial correlations ρ_{x,j_1j_2}, which estimates the graph of conditional dependence among the 42 brain AVOIs. As a comparison, we also apply the CFC method for testing ρ_{x,j_1j_2} being zero and the QZ method for testing the partial correlation ρ_{z,j_1j_2} of the original data being zero. We control the FDRs of all three methods at 0.05.

Figure 2 reports the estimated graphs of conditional dependence by the RRNR, CFC and QZ procedures. The nodes in the graphs represent the 42 AVOIs, colored by the four brain regions they belong to. This figure appears in color in the electronic version of this article, and any mention of color refers to that version. The pairs of AVOIs with significant nonzero partial correlations (ρ_{x,j_1j_2} for RRNR and CFC, and ρ_{z,j_1j_2} for QZ) are connected by lines, which are called edges of the graph. The first and second rows in Figure 2 present the results for the AD and NC groups, respectively. The results of RRNR, CFC and QZ are reported in the first, second and third columns of Figure 2, respectively. For the graphs of RRNR, all edges are colored in red. For the graphs of CFC and QZ, the common connections shared by RRNR are colored in grey, while their distinct connections are colored in blue. Those blue edges connect the pairs identified as conditionally dependent by either the CFC or QZ method but not by the proposed method.

From Figure 2, we observe that RRNR discovers 213 edges for the AD group, much more than the 35 and 86 edges made by CFC and QZ, respectively. There is only one edge, connecting two AVOIs within the temporal lobe, discovered by CFC but not by RRNR, while there are 179 edges discovered by RRNR but not by CFC. Comparing RRNR with QZ, there is one edge discovered by QZ but not by RRNR, while there are 128 edges only discovered by RRNR. For the NC group, RRNR discovers 190 edges, while CFC and QZ discover 50

and 84 edges, respectively. Comparing RRNR with CFC, there are 7 edges only found by CFC and 147 edges only discovered by RRNR. Comparing RRNR with QZ, there are 4 edges exclusively detected by QZ and 110 edges only revealed by RRNR. Those findings coincide with the results from the simulation study, which imply the proposed method could be more powerful than the existing methods in recovering the conditional dependence structure among variables.

[Figure 2 about here.]

7. Discussion

This paper considers testing for partial correlations under the nonparanormal model. The NPN partial correlation can be viewed as a rank-based measure for conditional dependence, which extends the classical partial correlation for linear dependence to nonlinear dependence. Compared to the existing methods on estimating the NPN graphical model, the proposed RRNR procedure is easy to implement and can be efficiently computed for data sets with a large number of variables. Since studying the interactions among variables and exploring their relationships is one of the critical questions in biomedical research, the proposed method has a wide range of applications. Besides the brain connectivity study introduced in the real data analysis, it can also be applied to study gene regulatory networks and microbiome networks by recovering the conditional dependence structure among variables. Another potential application is to infer the conditional association among multi-omics data, for example, the impact of plant gene expression levels on plant phenotypes conditional on soil microbiomes.

Estimating partial correlation is advantageous over precision coefficient, as not only does zero partial correlation imply conditional independence, but its value also provides the strength of conditional dependence. Besides the hypotheses (2.1) for detecting conditional

dependence among variables, we may also be interested in the hypotheses

$$H_{0,j_1j_2}^{(c)} : |\rho_{\mathrm{X},j_1j_2}| \leqslant c \quad \text{vs.} \quad H_{a,j_1j_2}^{(c)} : |\rho_{\mathrm{X},j_1j_2}| > c \tag{7.10}$$

to recover the pairs with absolute NPN partial correlation greater than a pre-specified constant $c \in (0,1)$. Similar to the rejection set $\mathcal{T}_0(s)$ in (3.8) for the hypotheses (2.1), we set

$$\mathcal{T}_c(s) \;=\; \big\{(j_1,j_2) : |\widehat{\rho}_{\mathrm{X},j_1j_2}| > c + s(1-\breve{\rho}^2_{\mathrm{X},j_1j_2})\{\log(p)/n\}^{1/2}\big\}, \tag{7.11}$$

and reject $H_{0,j_1j_2}^{(c)}$ in (7.10) if $(j_1,j_2) \in \mathcal{T}_c(s_c^*)$ where

$$s_c^* = \inf\left\{ s \in (0,2] : \frac{2\{1-\Phi(s\sqrt{\log p})\}(p^2-p)}{\max\{1, |\mathcal{T}_c(s)|\}} \leqslant \alpha \right\}. \tag{7.12}$$

If s_c^* does not exist, set $s_c^* = 2$.

In real applications, there may exist covariates that affect the response variables. For example, the patient's age, pre-existing conditions and risk factors may impact the brain image readings. Let $\mathbf{W}_i = (W_{i1}, \ldots, W_{im})^{\mathrm{T}}$ be the m dimensional covariates of the ith observation, where m could be much larger than n. Consider the linear model

$$Z_{ij} = \phi_{j,0} + \sum_{k=1}^{m} W_{ik}\phi_{j,k} + U_{ij} \tag{7.13}$$

for each response variable, where $\mathrm{E}(U_{ij}) = 0$ and $\mathrm{E}(U_{ij}W_{ik}) = 0$ for all $j = 1, \ldots, p$ and $k = 1, \ldots, m$. Let $\mathbf{U}_i = (U_{i1}, \ldots, U_{ip})^{\mathrm{T}}$ be the regression errors of the ith observation after adjusting the covariates effects. We assume a nonparanormal model on $\mathbf{U}_i$ such that $\mathbf{X}_i = \mathbf{H}(\mathbf{U}_i) = (h_1(U_{i1}), \ldots, h_p(U_{ip}))^{\mathrm{T}}$ follows a p-dimensional multivariate normal distribution with mean $\mathbf{0}$ and covariance $\mathbf{\Sigma}_{\mathrm{X}}$ with diagonal elements being one, where $\mathbf{H}(\cdot) = (h_1(\cdot), \ldots, h_p(\cdot))^{\mathrm{T}}$, and each univariate transformation function $h_j(\cdot)$ is strictly monotone. Our focus is still on the NPN partial correlation matrix $\mathbf{\Psi}_{\mathrm{X}} = (\rho_{\mathrm{X},j_1j_2})$ of the normal transformed data $\mathbf{X}_i$. Here, $\rho_{\mathrm{X},j_1j_2} = 0$ implies that Z_{ij_1} and Z_{ij_2} are conditionally independent after adjusting the covariates $\mathbf{W}_i$. Regularized estimation can be applied for the regression (7.13), and let $\{\widehat{U}_{ij}\}$ be the residuals. The proposed RRNR procedure can be applied on the ranks of $\{\widehat{U}_{ij}\}$ for

testing the conditional dependence among the variables of $\mathbf{U}_i$. We will investigate this topic in a future work.

Acknowledgments

The authors thank the AE and two reviewers for comments and suggestions which led to improvement in the presentation of the paper, and Professor Shuai Huang of The University of Washington for preprocessing the FDG-PET brain image data set. The authors acknowledge the support from the United States Department of Energy grant DE-SC0020355. Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (`adni.loni.usc.edu`). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: `http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf`. Data collection and sharing for this project was funded by ADNI (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012).

Data availability statement

The data that support the findings in this paper were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (`adni.loni.usc.edu`). The preprocessed data are available from Huang et al. (2010). Restrictions may apply to the availability of these data.

References

Baba, K., Shibata, R., and Sibuya, M. (2004). Partial correlation and conditional correlation as measures of conditional independence. *Australian & New Zealand Journal of Statistics* **46,** 657–664.

Banerjee, O., Ghaoui, L. E., and d'Aspremont, A. (2008). Model selection through sparse maximum likelihood estimation for multivariate gaussian or binary data. *Journal of Machine Learning Research* **9,** 485–516.

Barber, R. F. and Kolar, M. (2018). Rocket: Robust confidence intervals via kendall's tau for transelliptical graphical models. *The Annals of Statistics* **46,** 3422–3450.

Bickel, P. J. and Levina, E. (2008). Covariance regularization by thresholding. *The Annals of Statistics* **36,** 2577–2604.

Bühlmann, P. and Van De Geer, S. (2011). *Statistics for high-dimensional data: methods, theory and applications*. Springer Science & Business Media.

Bullmore, E. and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature reviews neuroscience* **10,** 186–198.

Cai, T., Liu, W., and Luo, X. (2011). A constrained l1 minimization approach to sparse precision matrix estimation. *Journal of the American Statistical Association* **106,** 594–607.

Cai, T. T. and Liu, W. (2016). Large-scale multiple testing of correlations. *Journal of the American Statistical Association* **111,** 229–240. PMID: 27284211.

Cai, T. T., Liu, W., and Zhou, H. H. (2016). Estimating sparse precision matrix: Optimal rates of convergence and adaptive estimation. *The Annals of Statistics* **44,** 455–488.

Chang, J., Qiu, Y., Yao, Q., and Zou, T. (2018). Confidence regions for entries of a large precision matrix. *Journal of Econometrics* **206,** 57–82.

Friedman, J., Hastie, T., and Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* **9,** 432–441.

Gu, Q., Cao, Y., Ning, Y., and Liu, H. (2015). Local and global inference for high dimensional nonparanormal graphical models.

He, Y., Zhang, X., Wang, P., and Zhang, L. (2017). High dimensional gaussian copula

graphical model with fdr control. *Computational Statistics & Data Analysis* **113,** 457–474.

Huang, S., Li, J., Sun, L., Ye, J., Fleisher, A., Wu, T., et al. (2010). Learning brain connectivity of alzheimer's disease by sparse inverse covariance estimation. *NeuroImage* **50,** 935–949.

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

Liu, H., Lafferty, J., and Wasserman, L. (2009). The nonparanormal: Semiparametric estimation of high dimensional undirected graphs. *Journal of Machine Learning Research* **10,** 2295–2328.

Liu, W. (2013). Gaussian graphical model estimation with false discovery rate control. *The Annals of Statistics* **41,** 2948–2978.

Mai, Q. and Zou, H. (2015). Sparse semiparametric discriminant analysis. *Journal of Multivariate Analysis* **135,** 175–188.

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

Ming, Y. and Yi, L. (2007). Model selection and estimation in the gaussian graphical model. *Biometrika* **94,** 19–35.

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.

Qiu, Y. and Zhou, X.-H. (2020). Estimating c-level partial correlation graphs with application to brain imaging. *Biostatistics* **21,** 641–658.

Qiu, Y. and Zhou, X.-H. (2021). Inference on multi-level partial correlations based on multi-subject time series data. *Journal of the American Statistical Association, Accepted* .

Ren, Z., Sun, T., Zhang, C.-H., and Zhou, H. H. (2015). Asymptotic normality and optimalities in estimation of large gaussian graphical models. *The Annals of Statistics* **43,** 991–1026.

Supekar, K., Menon, V., Rubin, D., Musen, M., and Greicius, M. D. (2008). Network analysis of intrinsic functional brain connectivity in alzheimer's disease. *PLoS computational biology* **4,** e1000100.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)* **58,** 267–288.

Wang, L., Ren, X., and Gu, Q. (2016). Precision matrix estimation in high dimensional gaussian graphical models with faster rates. In Gretton, A. and Robert, C. C., editors, *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, volume 51 of *Proceedings of Machine Learning Research*, pages 177–185, Cadiz, Spain. PMLR.

Wang, Y. R. and Huang, H. (2014). Review on statistical methods for gene network reconstruction using expression data. *Journal of Theoretical Biology* **362,** 53–61.

Xu, P., Tian, L., and Gu, Q. (2016). Communication-efficient distributed estimation and inference for transelliptical graphical models.

Xue, L. and Zou, H. (2012). Regularized rank-based estimation of high-dimensional nonparanormal graphical models. *The Annals of Statistics* **40,** 2541–2571.

Supporting Information

All the technical proofs and some additional figures referenced in Sections 5 and 6 are available in the Web Appendices with this paper at the Biometrics website on Wiley Online Library. The R code of the proposed method is available on Github at `https://github.com/Haoyan-Hu/RRNR`.

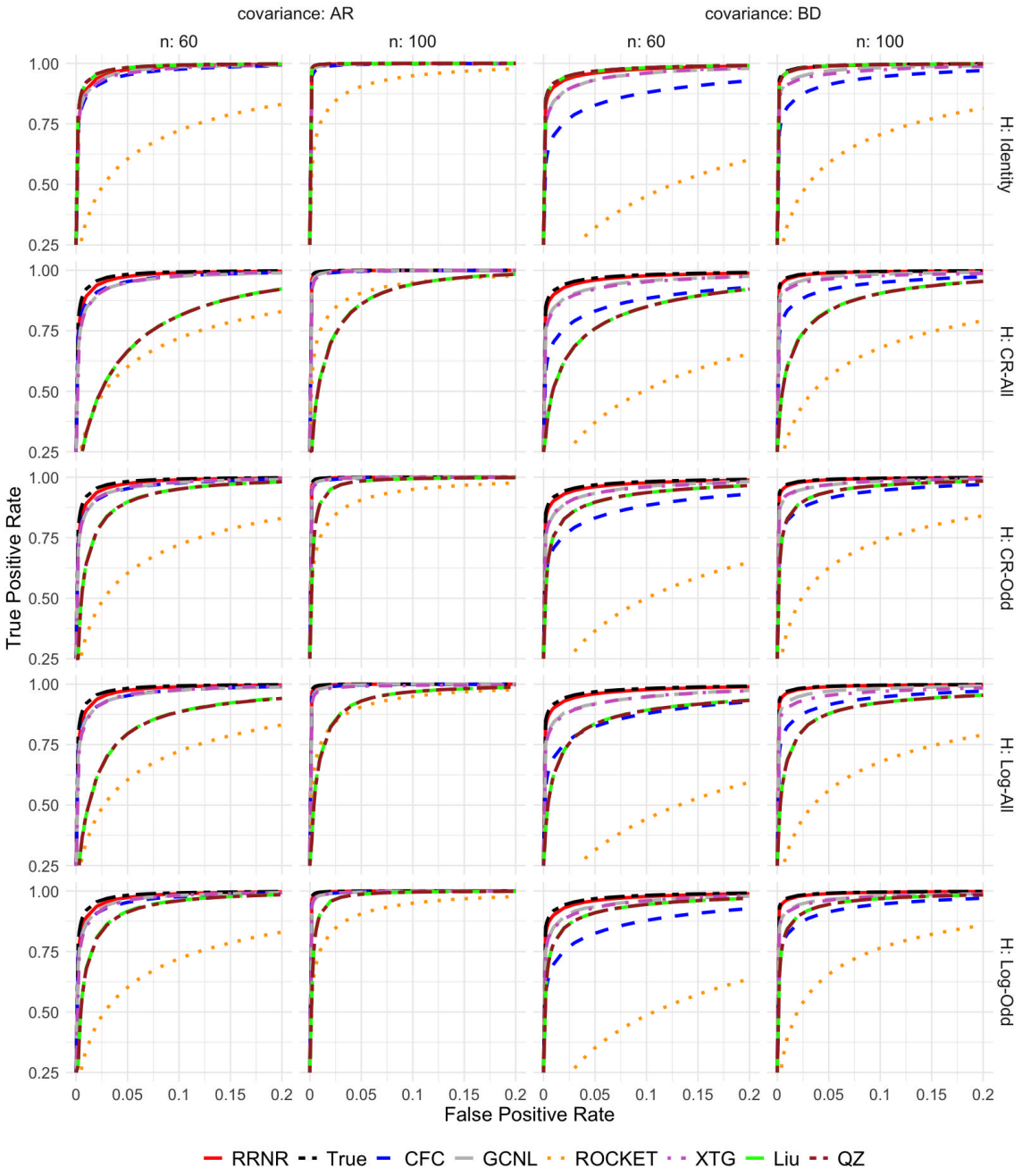

Figure 1. ROC curves ($p = 100$) of the proposed RRNR, CFC (He et al., 2017), GCNL (Gu et al., 2015), ROCKET (Barber and Kolar, 2018), XTG (Xu et al., 2016), Liu (Liu, 2013) and QZ (Qiu and Zhou, 2020) procedures for testing the hypotheses (2.1) of NPN partial correlations under five transformations: identity, cubic root on all variables (CR-All), cubic root on odd variables (CR-Odd), log on all variables (Log-All) and log on odd variables (Log-Odd). The nodewise regression approach applied on the true transformations, denoted by 'True', is included as the oracle benchmark. This figure appears in color in the electronic version of this article, and any mention of color refers to that version.

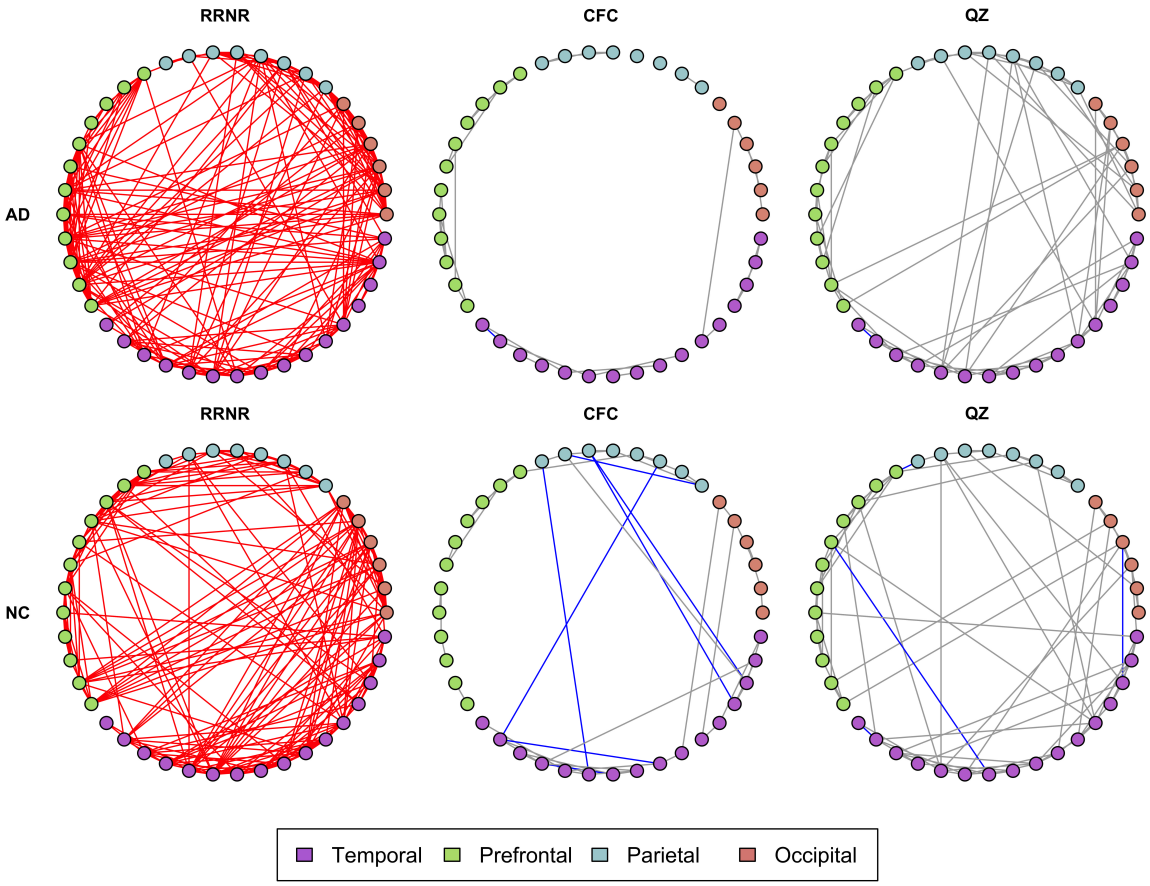

Figure 2. Graphs of the brain connectivity of the AD group (top three panels) and the NC group (bottom three panels). The three columns show the identified nonzero NPN partial correlation ρ_{x,j_1j_2} by the proposed RRNR procedure (left) and the CFC method (center), and the identified nonzero partial correlation ρ_{z,j_1j_2} of the original data by the QZ method (right), respectively. Each node represents an AVOI, which is colored by the brain region it belongs to. Red edges indicate the discovered connections by the RRNR procedure. For the graphs of CFC and QZ, the common connections shared by RRNR are marked in grey, and their distinct connections are marked in blue. This figure appears in color in the electronic version of this article, and any mention of color refers to that version.

Table 1

Computation times (in seconds per repetition) of the proposed RRNR, CFC (He et al., 2017), GCNL (Gu et al., 2015), ROCKET (Barber and Kolar, 2018), XTG (Xu et al., 2016), Liu (Liu, 2013) and QZ (Qiu and Zhou, 2020) procedures for testing the hypotheses (2.1) under the AR covariance and the log transformation on all variables (Log-All).

n	p	RRNR	CFC	GCNL	ROCKET	XTG	Liu	QZ
60	100	0.37	2.23	1504.55	25.08	53.77	0.32	4.79
60	400	5.90	109.24	NA	2542.62	4146.73	4.02	88.83
100	100	0.39	2.71	1608.09	45.18	138.28	0.30	9.02
100	400	6.24	116.35	NA	4617.30	7168.55	4.46	142.82

Table 2

Empirical FDRs and powers of the proposed RRNR, QZ (Qiu and Zhou, 2020), Liu (Liu, 2013) and CFC (He et al., 2017) procedures for testing the hypotheses (2.1) of nonparanormal partial correlations at 5% nominal FDR level.

AR Covariance									
n	p	Transformation	FDR				Power			
			RRNR	QZ	Liu	CFC	RRNR	QZ	Liu	CFC
60	100	Identity	0.075	0.103	0.085	0.043	0.681	0.776	0.753	0.598
		CR–All	0.075	0.376	0.585	0.041	0.680	0.086	0.317	0.600
		CR–Odd	0.074	0.390	0.288	0.044	0.675	0.042	0.358	0.595
		Log–All	0.075	0.287	0.476	0.041	0.678	0.135	0.416	0.597
		Log–Odd	0.077	0.337	0.270	0.042	0.683	0.065	0.380	0.602
60	400	Identity	0.051	0.067	0.052	0.045	0.454	0.587	0.545	0.464
		CR–All	0.051	0.630	0.832	0.046	0.456	0.049	0.282	0.461
		CR–Odd	0.050	0.826	0.558	0.042	0.455	0.009	0.197	0.464
		Log–All	0.051	0.534	0.734	0.043	0.456	0.092	0.355	0.465
		Log–Odd	0.050	0.802	0.498	0.043	0.455	0.009	0.189	0.462
100	100	Identity	0.091	0.087	0.077	0.037	0.970	0.980	0.978	0.908
		CR–All	0.092	0.174	0.465	0.037	0.970	0.089	0.558	0.908
		CR–Odd	0.094	0.142	0.244	0.036	0.969	0.105	0.806	0.908
		Log–All	0.091	0.141	0.391	0.037	0.969	0.163	0.683	0.907
		Log–Odd	0.092	0.131	0.239	0.037	0.970	0.159	0.820	0.909
100	400	Identity	0.071	0.089	0.078	0.033	0.929	0.954	0.949	0.856
		CR–All	0.072	0.368	0.724	0.034	0.929	0.067	0.497	0.858
		CR–Odd	0.072	0.411	0.361	0.034	0.929	0.038	0.682	0.857
		Log–All	0.072	0.306	0.629	0.033	0.929	0.122	0.619	0.857
		Log–Odd	0.071	0.370	0.334	0.033	0.929	0.057	0.698	0.858
BD Covariance										
n	p	Transformation	FDR				Power			
			RRNR	QZ	Liu	CFC	RRNR	QZ	Liu	CFC
60	100	Identity	0.036	0.044	0.035	0.029	0.784	0.811	0.794	0.471
		CR–All	0.038	0.107	0.369	0.030	0.785	0.238	0.508	0.476
		CR–Odd	0.036	0.058	0.140	0.030	0.784	0.275	0.590	0.481
		Log–All	0.037	0.083	0.275	0.029	0.784	0.305	0.566	0.464
		Log–Odd	0.037	0.045	0.112	0.030	0.784	0.289	0.602	0.464
60	400	Identity	0.029	0.036	0.027	0.027	0.727	0.774	0.761	0.420
		CR–All	0.028	0.233	0.641	0.027	0.726	0.197	0.517	0.403
		CR–Odd	0.029	0.115	0.274	0.026	0.726	0.186	0.577	0.411
		Log–All	0.029	0.178	0.521	0.027	0.726	0.257	0.581	0.419
		Log–Odd	0.029	0.086	0.214	0.027	0.727	0.200	0.592	0.411
100	100	Identity	0.041	0.045	0.041	0.035	0.904	0.892	0.886	0.636
		CR–All	0.041	0.052	0.331	0.034	0.906	0.181	0.597	0.659
		CR–Odd	0.040	0.032	0.130	0.034	0.906	0.253	0.695	0.638
		Log–All	0.042	0.050	0.264	0.035	0.905	0.242	0.642	0.642
		Log–Odd	0.041	0.029	0.113	0.034	0.906	0.292	0.701	0.639
100	400	Identity	0.037	0.042	0.036	0.031	0.893	0.909	0.901	0.610
		CR–All	0.037	0.107	0.581	0.030	0.893	0.197	0.614	0.596
		CR–Odd	0.037	0.073	0.247	0.030	0.893	0.240	0.713	0.605
		Log–All	0.037	0.106	0.493	0.031	0.893	0.257	0.669	0.603
		Log–Odd	0.038	0.072	0.211	0.030	0.893	0.278	0.719	0.607