Contents lists available at ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/ymeth

Node-based learning of differential networks from multi-platform gene expression data

Le Ou-Yang [a], Xiao-Fei Zhang [b,*], Min Wu [c], Xiao-Li Li [c]

[a] *College of Information Engineering & Shenzhen Key Laboratory of Media Security, Shenzhen University, Shenzhen, China*
[b] *School of Mathematics and Statistics & Hubei Key Laboratory of Mathematical Sciences, Central China Normal University, Wuhan, China*
[c] *Institute for Infocomm Research (I2R), A*STAR, 1 Fusionopolis Way, Singapore*

ARTICLE INFO

Article history:
Received 15 January 2017
Received in revised form 11 April 2017
Accepted 18 May 2017
Available online xxxx

Keywords:
Gaussian graphical model
Differential network analysis
Multi-view learning
Group lasso
Gene expression

ABSTRACT

Recovering gene regulatory networks and exploring the network rewiring between two different disease states are important for revealing the mechanisms behind disease progression. The advent of high-throughput experimental techniques has enabled the possibility of inferring gene regulatory networks and differential networks using computational methods. However, most of existing differential network analysis methods are designed for single-platform data analysis and assume that differences between networks are driven by individual edges. Therefore, they cannot take into account the common information shared across different data platforms and may fail in identifying driver genes that lead to the change of network. In this study, we develop a node-based multi-view differential network analysis model to simultaneously estimate multiple gene regulatory networks and their differences from multi-platform gene expression data. Our model can leverage the strength across multiple data platforms to improve the accuracy of network inference and differential network estimation. Simulation studies demonstrate that our model can obtain more accurate estimations of gene regulatory networks and differential networks than other existing state-of-the-art models. We apply our model on TCGA ovarian cancer samples to identify network rewiring associated with drug resistance. We observe from our experiments that the hub nodes of our identified differential networks include known drug resistance-related genes and potential targets that are useful to improve the treatment of drug resistant tumors.

1. Introduction

Biological processes often involve the interactions of genetic components such as mRNAs and proteins. Characterizing the regulatory interactions between genes is critical for elucidating the structural and functional architecture within cells [1–3]. Moreover, there is strong evidence that gene regulatory networks (GRN) undergo changes in response to different conditions such as cancer progression and drug resistance [4–6]. Therefore, inferring gene regulatory networks and exploring how theses networks change across different conditions are fundamental for understanding the biological mechanisms behind disease development [7].

With the accumulation of gene expression data, an increasing number of computational methods have been proposed for gene regulatory network estimation [8,9]. Gaussian graphical models (GGMs), which can identify conditional dependence (or direct dependence) relationships between genes, have been widely used for network inference [10]. Based on the assumption that the observed gene expression data are generated from a multivariate normal distribution, the gene regulatory network can be determined directly from the precision matrix (or inverse covariance matrix) of GGMs [11]. That is, two genes interact with each other if and only if the corresponding entry of the precision matrix is nonzero. Therefore, based on GGMs, the problem of gene regulatory network estimation can be turned into a problem of precision matrix estimation. However, traditional GGMs typically infer one network for a specific condition, and do not consider the network rewiring between different conditions.

In recent years, several differential network analysis methods have been developed for identifying altered dependencies between genes across different conditions [12–14]. Based on GGMs, the difference between two group-specific networks can be identified by calculating the difference between the two corresponding precision matrices [13]. Thus, most existing differential network analysis methods first estimate each group-specific network separately, and then calculate their difference [15]. However, estimating the group-specific networks separately may lose the global dependencies that preserve across all conditions. To exploit the similarity between the true group-specific networks, several methods have

[*] Corresponding author.
E-mail address: zhangxf@mail.ccnu.edu.cn (X.-F. Zhang).

http://dx.doi.org/10.1016/j.ymeth.2017.05.014
1046-2023/© 2017 Elsevier Inc. All rights reserved.

been proposed to jointly estimate multiple graphical models that share certain characteristics [13,16]. Most of these methods assume that the differences between networks are driven by individual edges. This is unrealistic in many real-world applications since the difference between gene regulatory networks might be driven by certain genes whose patterns of connectivity to other genes are disrupted across conditions. To provide a more intuitive interpretation of the network differences, Mohan et al. introduced a node-based learning approach to jointly estimate multiple GGMs [14].

Rapidly evolving technologies make it possible to collect gene expression data for same patients from different experimental platforms [17]. As gene expression data collected from different platforms (multi-platform gene expression data) describe the expression levels of genes for same patients from different views, they may share some consistent information. Therefore, integrating multi-platform gene expression data may improve the accuracy of gene regulatory network estimation and differential network analysis [18,13]. However, previous differential network analysis methods focus on analyzing the gene expression data collected from a single platform, which could not effectively leverage the common information provided by multi-platform gene expression data.

To address the above problems, we propose a novel node-based multi-view learning algorithm called co-perturbed node joint graphical lasso (CPJGL) model, to simultaneously infer multiple gene regulatory networks corresponding to different patient groups and the differential networks between these patient groups based on gene expression data collected from multiple data platforms (Fig. 1). Our model is an extension of the node-based learning approach proposed by Mohan et al. [14] to the case where gene expression data are characterized in terms of two aspect: patient groups and platform types. Instead of assuming that individual edges are shared or differed across disease states, we assume that the differences between networks are driven by certain perturbed regulatory genes. Based on the row-column overlap norm regularizer [14] and the group lasso penalty [19], our model can exploit the characteristics shared by gene expression data collected from different types of platforms. We propose an alternating direction method of multiplier (ADMM) algorithm to solve the optimization problem. In simulation studies, our proposed CPJGL demonstrated better performance than other competing methods in network inference and differential network analysis. To illustrate the effectiveness of CPJGL on real biological data, we apply CPJGL on TCGA ovarian cancer samples to identify network rewiring associated with platinum resistance. We identify three key regulator genes, namely TSC1, IRS1 and PDPK1, from mTOR signaling pathway and two perturbed genes (MYC and BMP7) from TGF-β signaling pathway. By literature search, we find that these five genes play important roles in drug resistance.

2. Methods

2.1. Gaussian graphical models

Gaussian graphical models can encode the conditional dependencies among a set of p genes, where the expression levels (denoted by a p-dimensional random vector $X = (X_1, \ldots, X_p)^T$) of these p genes are assumed to follow a multivariate Gaussian distribution $N(\mu, \Sigma)$ (here $\mu \in \mathbb{R}^p$ and Σ is a positive definite $p \times p$ matrix). Then two genes are conditionally independent if and only if the corresponding entry of the inverse covariance matrix (precision matrix) $\Theta = \Sigma^{-1}$ is zero [11], i.e., genes i and j are independent of each other given all of the other genes if and only if $\Theta_{ij} = 0$. These conditional dependence relationships can be described by a graph in which nodes denote genes and edges connect conditionally dependent pairs of genes. To estimate the conditional dependencies among p genes, it suffices to estimate the sparsity pattern of the corresponding precision matrix Θ. Suppose that we have n observations that are independently drawn from a multivariate Gaussian distribution $N(\mu, \Sigma)$. When $n > p$, we can estimate the precision matrix $\Theta = \Sigma^{-1}$ by maximum likelihood. However, when $p > n$, this approach fails since the empirical covariance matrix is singular and cannot be inverted to yield an estimate of Σ^{-1}. To deal with this problem, a number of studies [20–22] have instead taken a penalized log-likelihood:

$$\max_{\Theta} \frac{n}{2}(\log\det(\Theta) - \mathrm{tr}(S\Theta)) - \lambda\|\Theta\|_1, \tag{1}$$

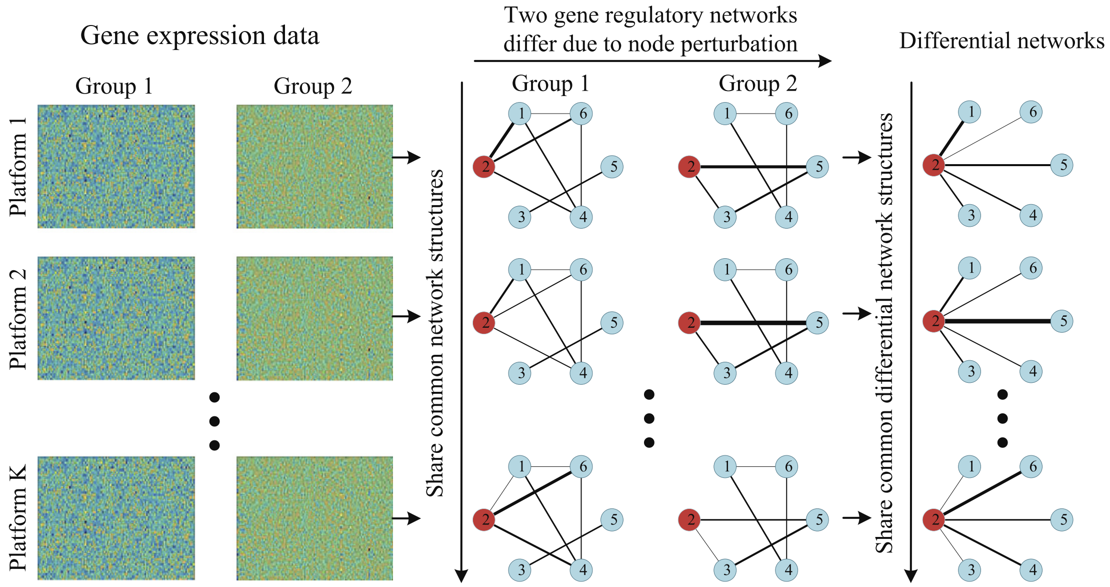

Gene expression data
Group 1
Group 2
Platform 1
Platform 2
Platform K
Share common network structures
Two gene regulatory networks differ due to node perturbation
Group 1
Group 2
1
2
3
4
5
6
Share common differential network structures
Differential networks

Fig. 1. Motivation and overview of our model. The input data are gene expression data for two different patient groups collected from K data platforms. CPJGL jointly estimates the corresponding $2K$ gene regulatory networks and the K differential networks between these two patient groups by drawing support from the K data platforms. CPJGL encourages the inferred networks and differential networks to share common network structures. It also imposes hub structures on the resulting differential networks. The red node denotes the driver gene that perturbs the network structure.

Please cite this article in press as: L. Ou-Yang et al., Methods (2017), http://dx.doi.org/10.1016/j.ymeth.2017.05.014

where S is the sample covariance matrix, $\det(\cdot)$ is the determinant of a matrix, $\mathrm{tr}(\cdot)$ is the trace of a matrix, $\|\Theta\|_1$denotes the sum of the absolute values of the elements of Θ and λ is a nonnegative tuning parameter. The solution to problem (1) provides a sparse estimate for Θ.

2.2. Problem statement and notations

Suppose that we have independent observations of p genes collected from K different data platforms for $n = n_1 + n_2$ subjects that can be divided into two groups (n_c denotes the number of subjects in the c-th group), i.e., $\mathbf{x}_i^{kc} = (x_{i1}^{kc}, \ldots, x_{ip}^{kc})^T \in \mathbb{R}^p$ for $i = 1, \ldots, n_c, c = 1, 2, k = 1, \ldots, K$. Furthermore, we assume that $\mathbf{x}_1^{kc}, \ldots, \mathbf{x}_{n_c}^{kc} \sim N(\mu_c, \Sigma^{kc})$ for $c = 1, 2, k = 1, \ldots, K$, where Σ^{kc} denotes the covariance matrices for c-th group, corresponding to the k-th data platform. Without loss of generality, we assume that the variables within each group are centred such that $\mu_c = 0$. The goal of this study is to construct group-specific gene regulatory networks and explore the changes of gene regulatory networks between two different patient groups, based on gene expression data collected from K different platforms. That is, we seek to estimate the $2K$ precision matrices $\Theta^{kc} = (\Sigma^{kc})^{-1}$ for $c = 1, 2, k = 1, \ldots, K$, and identify the differences between Θ^{k1} and Θ^{k2} for $k = 1, \ldots, K$.

For the sake of convenience, we denote $\{\Sigma^{kc}\}_{k=1,\ldots,K}^{c=1,2}$ and $\{\Theta^{kc}\}_{k=1,\ldots,K}^{c=1,2}$ as $\{\Sigma^{kc}\}$ and $\{\Theta^{kc}\}$ respectively. Suppose that $A \in \mathbb{R}^{p\times p}$ is a $p \times p$ matrix with element A_{ij}. Its Frobenius norm, L_1 norm and $L_{2,1}$ norm are defined as $\|A\|_F = (\sum_{i,j} A_{ij}^2)^{1/2}$, $\|A\|_1 = \sum_{i,j}|A_{ij}|$ and $\|A\|_{2,1} = \sum_{j=1}^{p}\sqrt{\sum_{i=1}^{p} A_{ij}^2} = \sum_{j=1}^{p}\|A_j\|_2$ (where A_j is the j-th column of the matrix A), respectively.

2.3. Co-perturbed node joint graphical lasso

In this section, we propose a co-perturbed node joint graphical lasso (CPJGL) model to jointly estimate multiple gene regulatory networks corresponding to distinct but related platform types and patient groups.

Let $S^{kc} = (1/n_c)\sum_{i=1}^{n_c} \mathbf{x}_i^{kc} (\mathbf{x}_i^{kc})^T$ be the sample covariance matrix for the k-th platform and the c-th patient group. The negative log-likelihood for the data can be written as [13]

$$L\left(\{\Theta^{kc}\}\right) = \sum_{k=1}^{K}\sum_{c=1}^{2} \frac{n_c}{2}\left(\mathrm{tr}\left(S^{kc}\Theta^{kc}\right) - \log\det\left(\Theta^{kc}\right)\right). \tag{2}$$

In order to provide a more intuitive interpretation of the network differences, we consider the following assumptions: (1) The gene regulatory networks of different patient groups are quite similar to each other and the network differences are arisen from certain genes (e.g., transcription factors or kinases) whose functional dependencies with other genes are disrupted across conditions; (2) The number of edges in a biological network may be far less than the full connected network. Therefore, we can require the resulting estimation of precision matrices to be sparse; (3) As gene expression data collected from different platforms share certain common information, the precision matrices Θ^{k1} and Θ^{k2} as well as their difference (i.e., $\Theta^{k1} - \Theta^{k2}$) estimated from each platform may be similar with each other or share some common structures. Therefore, jointly estimating the $2K$ precision matrices and their differences from $2K$ data sets may result in more accurate estimations.

Unlike previously developed node-based joint graphical lasso models that infer the precision matrices corresponding to each data platform separately, in addition to the loss function (2), we impose a group lasso penalty on the precision matrices and a row-column overlap norm penalty on the network difference, and develop a novel co-perturbed node joint graphical lasso (CPJGL) model:

$$\begin{aligned}
&\min_{\{\Theta^{kc}\}\in S_{++}^{p},\{V^k\}\in R^{p\times p}} \sum_{k=1}^{K}\sum_{c=1}^{2} n_c\left(\mathrm{tr}\left(S^{kc}\Theta^{kc}\right) - \log\det\left(\Theta^{kc}\right)\right)\\
&\quad + \lambda_0\sum_{k=1}^{K}\|V^k\|_1 + \lambda_1\sum_{c=1}^{2}\sum_{i\neq j}\left(\sum_{k=1}^{K}\left(\Theta_{ij}^{kc}\right)^2\right)^{1/2} + \lambda_2\sum_{j=1}^{p}\left\|\begin{bmatrix}V^1\\ \vdots\\ V^k\end{bmatrix}_j\right\|_2.\\
&\text{s.t. } \Theta^{k1} - \Theta^{k2} = V^k + \left(V^k\right)^T, \text{ for } k = 1, \ldots, K
\end{aligned} \tag{3}$$

where S_{++}^{p} denotes the sets of positive definite matrices of size p, and λ_0, λ_1 and λ_2 are non-negative tuning parameters. The $l_{2,1}$-norm regularization (group lasso penalty) $\sum_{i\neq j}\left(\sum_{k=1}^{K}\left(\Theta_{ij}^{kc}\right)^2\right)^{\frac{1}{2}}$ defined on $\{\Theta^{kc}\}$ plays an important role in our CPJGL method: it is the minimization of this penalty function that enforces the entries $\Theta_{ij}^{kc}, k = 1, 2, \ldots, K$, to have consistent magnitudes, all either zeros or nonzeros. Following the idea of row-column overlap norm penalty [14], we decompose the differential networks $\Theta^{k1} - \Theta^{k2}$ as $\Theta^{k1} - \Theta^{k2} = V^k + (V^k)^T$ for $k = 1, \ldots, K$, where V^k need not be symmetric. By penalizing the columns of V^k, we could find hub nodes that drive the difference between two precision matrices Θ^{k1} and Θ^{k2}. The sparsity of each hub node's connections to other nodes is controlled by λ_0. Furthermore, based on the group lasso penalty, a shared structure is encouraged among the K differential networks. The choices of λ_1 and λ_0 control the sparsity of resulting gene regulatory networks and differential networks respectively, while the choice of λ_2 controls the selection of hub nodes. We present our parameter selection strategy at the end of this section.

2.4. Algorithm for parameter estimation

In this section, we solve the optimization problem (3) by using an alternating direction method of multipliers (ADMM) [23]. We reformulate (3) by introducing new variables:

$$\begin{aligned}
&\min_{\{\Theta^{kc}\}\in S_{++}^{p},\{Z^{kc}\},\{V^k\},\{W^k\}} \sum_{k=1}^{K}\sum_{c=1}^{2} n_c\left(\mathrm{tr}\left(S^{kc}\Theta^{kc}\right) - \log\det\left(\Theta^{kc}\right)\right)\\
&\quad + \lambda_0\sum_{k=1}^{K}\|V^k\|_1 + \lambda_1\sum_{c=1}^{2}\sum_{i\neq j}\left(\sum_{k=1}^{K}\left(Z_{ij}^{kc}\right)^2\right)^{1/2} + \lambda_2\sum_{j=1}^{p}\left\|\begin{bmatrix}V^1\\ \vdots\\ V^k\end{bmatrix}_j\right\|_2.\\
&\text{s.t. } \Theta^{k1} - \Theta^{k2} = V^k + W^k,\ V^k = \left(W^k\right)^T,\\
&\Theta^{kc} = Z^{kc}, \text{ for } k = 1, \ldots, K \text{ and } c = 1, 2.
\end{aligned} \tag{4}$$

The augmented Lagrangian to (4) is given by

$$\begin{aligned}
&\sum_{k=1}^{K}\sum_{c=1}^{2} n_c\left(\mathrm{tr}(S^{kc}\Theta^{kc}) - \log\det(\Theta^{kc})\right) + \lambda_1\sum_{c=1}^{2}\sum_{i\neq j}\left(\sum_{k=1}^{K}\left(Z_{ij}^{kc}\right)^2\right)^{1/2}\\
&+ \lambda_0\sum_{k=1}^{K}\|V^k\|_1 + \lambda_2\sum_{j=1}^{p}\left\|\begin{bmatrix}V^1\\ \vdots\\ V^K\end{bmatrix}_j\right\|_2\\
&+ \sum_{k=1}^{K}\langle F^k, \Theta^{k1} - \Theta^{k2} - (V^k + W^k)\rangle + \sum_{k=1}^{K}\langle G^k, V^k - (W^k)^T\rangle\\
&+ \sum_{c=1}^{2}\sum_{k=1}^{K}\langle Q^{kc}, \Theta^{kc} - Z^{kc}\rangle\\
&+ \frac{\rho}{2}\sum_{k=1}^{K}\left(\|\Theta^{k1} - \Theta^{k2} - (V^k + W^k)\|_F^2 + \|V^k - (W^k)^T\|_F^2 + \sum_{c=1}^{2}\|\Theta^{kc} - Z^{kc}\|_F^2\right).
\end{aligned} \tag{5}$$

where $\{F^k\} = F^1, \ldots, F^K, \{G^k\} = G^1, \ldots, G^K$ and $\{Q^{kc}\} = Q^{11}, \ldots, Q^{K2}$ are dual variables and ρ serves as a penalty parameter. Based on this augmented Lagrangian, the computational algorithm for solving (3) is given in Algorithm 1, in which the operator Expand is given by

$$\begin{aligned}\text{Expand}(A, \rho, n_c) &= \text{argmin}_{\Theta \in S^p_{++}} \{-n_c \log\det(\Theta) + \rho\|\Theta - A\|_F^2\} \\ &= \tfrac{1}{2} U\Big(D + \sqrt{D^2 + \tfrac{2n_c}{\rho} I}\Big)U^T,\end{aligned} \tag{6}$$

where UDU^T is the eigenvalue decomposition of a symmetric matrix A and n_c is the number of subjects in the c-th group. The operator $\mathcal{T}_{1,2}$ is given by the following sparse group lasso problem which has closed solution [19,13]

$$\mathcal{T}_{1,2}(A, \lambda, \beta) = \text{argmin}_X \left\{\frac{1}{2}\|X - A\|_F^2 + \lambda\|X\|_1 + \beta\sum_{j=1}^{p}\|X_j\|_2\right\}, \tag{7}$$

In our implementation of this algorithm, the stopping criterion for the inner loop is

$$\max_{k\in\{1,2,\ldots,K\}, c\in\{1,2\}} \left\{\frac{\|(\Theta^{kc})^{(t+1)} - (\Theta^{kc})^{(t)}\|_F}{\|(\Theta^{kc})^{(t)}\|_F}\right\} \leqslant \epsilon, \tag{8}$$

where $(\Theta^{kc})^{(t)}$ denotes the estimate of Θ^{kc} in the t-th iteration of the ADMM algorithm and ϵ is a tolerance that is set to 10^{-4} in our experiments.

2.5. Tuning parameter selection

In this section, we propose the following Akaike information criterion (AIC)-type quantity for selecting the tuning parameters λ_0, λ_1 and λ_2:

$$\begin{aligned}AIC(\lambda_0, \lambda_1, \lambda_2) = &\sum_{k=1}^{K}\sum_{c=1}^{2}\Big(n_c trace(S^{kc}\hat{\Theta}^{kc}) - n_c log(det(\hat{\Theta}^{kc}))\Big) \\ &+ 2\sum_{k=1}^{K}\sum_{c=1}^{2}|\hat{\Theta}^{kc}| + 2\sum_{k=1}^{K}\Big(\upsilon^k + \alpha\cdot\big(|\widehat{V}^k| - \upsilon^k\big)\Big).\end{aligned} \tag{9}$$

where υ^k is the number of estimated hub nodes ($\upsilon^k = \sum_{j=1}^{p} 1_{\{\|\widehat{V}^k_j\|_0 > 0\}}$), $|\hat{\Theta}^{kc}|$ and $|\widehat{V}^k|$ are the cardinalities of $\hat{\Theta}^{kc}$ and $\widehat{V}^k$ for $k = 1, \ldots, K, c = 1, 2$ and α is a constant between zero and one. We select the tuning parameters $(\lambda_0, \lambda_1, \lambda_2)$ for which the quantity $AIC(\lambda_0, \lambda_1, \lambda_2)$ is minimized. Following the choice of [24], we take $\alpha = 0.2$ in this study.

Algorithm 1 ADMM algorithm for solving the CPJGL optimization problem (3).

```
Input:
sample covariance matrices $\{S^{kc}\}_{k=1,\ldots,K}^{c=1,2}$, parameters $\lambda_0, \lambda_1$ and $\lambda_2$.
Initialize:
$\Theta^{kc} = Z^{kc} = I, V^k = W^k = 0, F^k = 0, G^k = 0, Q^{kc} = 0$, for $k = 1, \ldots, K$ and $c = 1, 2, \rho = 0.5, \mu = 5, \epsilon = 10^{-4}, t_{max} = 1000$.
for $t = 1 : t_{max}$ do
  $\rho \leftarrow \mu\rho$
  while (not converged) do
  1: $\Theta^{k1} \leftarrow \text{Expand}\Big(\frac{1}{2}(\Theta^{k2} + V^k + W^k + Z^{k1}) - \frac{1}{2\rho}(Q^{k1} + n_1 S^{k1} + F^k), \rho, n_1\Big)$, for $k = 1, \ldots, K$;
  2: $\Theta^{k2} \leftarrow \text{Expand}\Big(\frac{1}{2}(\Theta^{k1} - (V^k + W^k) + Z^{k2}) - \frac{1}{2\rho}(Q^{k2} + n_2 S^{k2} - F^k), \rho, n_2\Big)$, for $k = 1, \ldots, K$;
  3: $Z^{kc}_{ij} = \max\left\{1 - \frac{\lambda_1}{\rho\{\sum_{k=1}^{K}(\Theta^{kc} + Q^{kc}/\rho)^2_{ij}\}^{\frac{1}{2}}}, 0\right\}(\Theta^{kc} + Q^{kc}/\rho)_{ij}$ for $i \neq j$ and $Z^{kc}_{ii} = (\Theta^{kc} + Q^{kc}/\rho)_{ii}, c = 1, 2$ and $k = 1, \ldots, K$;
  4: Let $H^k = \frac{1}{2}\Big(\Theta^{k1} - \Theta^{k2} - W^k + (W^k)^T\Big) + \frac{1}{2\rho}(F^k - G^k)$, for $k = 1, \ldots, K$; $\begin{bmatrix} V^1 \\ \vdots \\ V^K \end{bmatrix} \leftarrow \mathcal{T}_{1,2}\left(\begin{bmatrix} H^1 \\ \vdots \\ H^K \end{bmatrix}, \frac{\lambda_0}{2\rho}, \frac{\lambda_2}{2\rho}\right)$;
  5: $W^k \leftarrow \frac{1}{2}\Big((V^k)^T - V^k + (\Theta^{k1} - \Theta^{k2})\Big) + \frac{1}{2\rho}(F^k + (G^k)^T)$ for $k = 1, \ldots, K$;
  6: $F^k \leftarrow F^k + \rho(\Theta^{k1} - \Theta^{k2} - (V^k + W^k))$ for $k = 1, \ldots, K$;
  7: $G^k \leftarrow G^k + \rho(V^k - (W^k)^T)$ for $k = 1, \ldots, K$;
  8: $Q^{kc} \leftarrow Q^{kc} + \rho(\Theta^{kc} - Z^{kc})$ for $c = 1, 2$ and $k = 1, \ldots, K$;
  end while
end for
```

3. Simulation studies

In this section, we assess the performance of our proposed co-perturbed node joint graphical lasso (CPJGL) model by comparing it with other Gaussian graphical model-based algorithms. The competing algorithms are the perturbed-node joint graphical lasso (PNJGL) as proposed in [14], the joint graphical lasso with group lasso penalty (GGL) as proposed in [13] and the gene network reconstruction (GNR) method proposed by Wang et al. [18]. We use the MATLAB code provided by Mohan et al. [14] to implement PNJGL. For GGL, we use the JGL function with 'penalty = group' from the R package JGL. We use the software provided by Wang et al. [18] to implement GNR. As PNJGL is not designed for multi-platform data, when applying PNJGL, we fit networks for each platform separately. When applying GGL, precision matrices correspond to multiple platforms are fitted for each group separately, and then we estimate the differential network correspond to each platform by calculating the difference between the inferred group-specific precision matrices. Since GNR is designed to reconstruct a gene network using multiple datasets, when applying GNR, we first

Please cite this article in press as: L. Ou-Yang et al., Methods (2017), http://dx.doi.org/10.1016/j.ymeth.2017.05.014

utilize multiple datasets to infer the gene network corresponding to each group separately, and then estimate the differential network by calculating the difference between the inferred group-specific gene networks. PNJGL has two parameters λ_1 and λ_2 which control the sparsity of the estimated networks and differential networks respectively. GGL has two parameters λ_1 and λ_2 which control the sparsity of the estimated networks and the consistency across multiple precision matrices. To ease interpretation, similar to [13], we reparameterize the tuning parameters for GGL, where $\omega_1 = \lambda_1 + \frac{1}{\sqrt{2}}\lambda_2$ and $\omega_2 = \frac{1}{\sqrt{2}}\lambda_2/(\lambda_1 + \frac{1}{\sqrt{2}}\lambda_2)$. GNR has one parameter λ which controls the sparsity of the estimated network.

We use several metrics to evaluate the performance of various algorithms. If $\hat{\Theta}_{ij}^{kc}$ is the (i,j)th entry of an estimator $\hat{\Theta}^{kc}$ and Θ_{ij}^{kc} is the (i,j)th entry of the true Θ^{kc}, for $k = 1,\ldots,K$ and $c = 1,2$, the true positive (TP) edges, false positive (FP) edges, true positive differential (TPD) edges, false positive differential (FPD) edges, precision (Pre) and recall (Rec) are defined as

$$TP = \sum_{k=1}^{K}\sum_{c=1}^{2}\sum_{i<j} I\left\{\hat{\Theta}_{ij}^{kc} \neq 0 \quad \text{and} \quad \Theta_{ij}^{kc} \neq 0\right\},$$

$$FP = \sum_{k=1}^{K}\sum_{c=1}^{2}\sum_{i<j} I\left\{\hat{\Theta}_{ij}^{kc} \neq 0 \quad \text{and} \quad \Theta_{ij}^{kc} = 0\right\},$$

$$TPD = \sum_{k=1}^{K}\sum_{i<j} I\left\{\hat{\Theta}_{ij}^{k1} \neq \hat{\Theta}_{ij}^{k2} \quad \text{and} \quad \Theta_{ij}^{k1} \neq \Theta_{ij}^{k2}\right\},$$

$$FPD = \sum_{k=1}^{K}\sum_{i<j} I\left\{\hat{\Theta}_{ij}^{k1} \neq \hat{\Theta}_{ij}^{k2} \quad \text{and} \quad \Theta_{ij}^{k1} = \Theta_{ij}^{k2}\right\},$$

$$Pre = \frac{\sum_{k=1}^{K}\sum_{i<j} I\left(\hat{\Theta}_{ij}^{k1} \neq \hat{\Theta}_{ij}^{k2} \quad \text{and} \quad \Theta_{ij}^{k1} \neq \Theta_{ij}^{k2}\right)}{\sum_{k=1}^{K}\sum_{i<j} I\left(\hat{\Theta}_{ij}^{k1} \neq \hat{\Theta}_{ij}^{k2}\right)},$$

$$Rec = \frac{\sum_{k=1}^{K}\sum_{i<j} I(\hat{\Theta}_{ij}^{k1} \neq \hat{\Theta}_{ij}^{k2} \quad \text{and} \quad \Theta_{ij}^{k1} \neq \Theta_{ij}^{k2})}{\sum_{k=1}^{K}\sum_{i<j} I(\Theta_{ij}^{k1} \neq \Theta_{ij}^{k2})}.$$

3.1. Simulation set-up

We consider two groups of subjects and their observations of p genes collected from $K = 3$ data platforms. In particular, we generate 6 scale-free networks for the two groups of subjects and 3 data platforms, and all of them include the same set of p genes. We consider scale-free networks because many biological networks have been reported to be scale-free and it is more difficult to estimate power-law degree distributions [1]. Across all simulation setups, we set $p = 100$ and $n_1 = n_2 = n \in \{100, 200, 400\}$. Four of the p genes are modified to create the perturbed genes (hub nodes in the differential network) [14]. The structure of networks and differential networks are preserved across the 3 data platforms. In particular, we generate the data as follows:

1. We use the SFNG functions in Matlab with parameters *mlinks* $= 2$ and *seed* $= 1$ to generate a scale-free network with $p = 100$ nodes. We randomly select four nodes as perturbed nodes that drive the differential network.
2. For $k = 1,\ldots,K$, we repeat Steps 3–5 to generate data sets for each data platform.
3. We create a $p \times p$ symmetric matrix A to denote the adjacency matrix of the generated scale-free network. In particular, $A_{ij} = 0$ if there is no edge between nodes i and j, and a uniform distribution with support $[-0.6, -0.3] \bigcup [0.3, 0.6]$ is used to generate the nonzero entry of A. Then we duplicated A into two matrices S_1^k and S_2^k. For each selected perturbed node, we randomly selected 50 elements of corresponding row and column of either S_1^k or S_2^k and reset their values to be drawn from a uniform distribution with support $[-0.6, -0.3] \bigcup [0.3, 0.6]$. This results in four hub nodes with degree 50 in the differential network.
4. We let $d = min(\lambda_{min}(S_1^k), \lambda_{min}(S_2^k))$, where $\lambda_{min}(\cdot)$ denotes the smallest eigenvalue of the matrix. To ensure positive definiteness, we set $(\Sigma^{k1})^{-1} = S_1^k + (0.1 + |d|)I$ and $(\Sigma^{k2})^{-1} = S_2^k + (0.1 + |d|)I$, where I is a $p \times p$ identity matrix.
5. We generate $n_1 = n_2$ independent subjects for each group, i.e., $\mathbf{x}_1^{k1}, \ldots, \mathbf{x}_{n_1}^{k1} \sim N(0, \Sigma^{k1})$ and $\mathbf{x}_1^{k2}, \ldots, \mathbf{x}_{n_2}^{k2} \sim N(0, \Sigma^{k2})$.

3.2. Simulation results

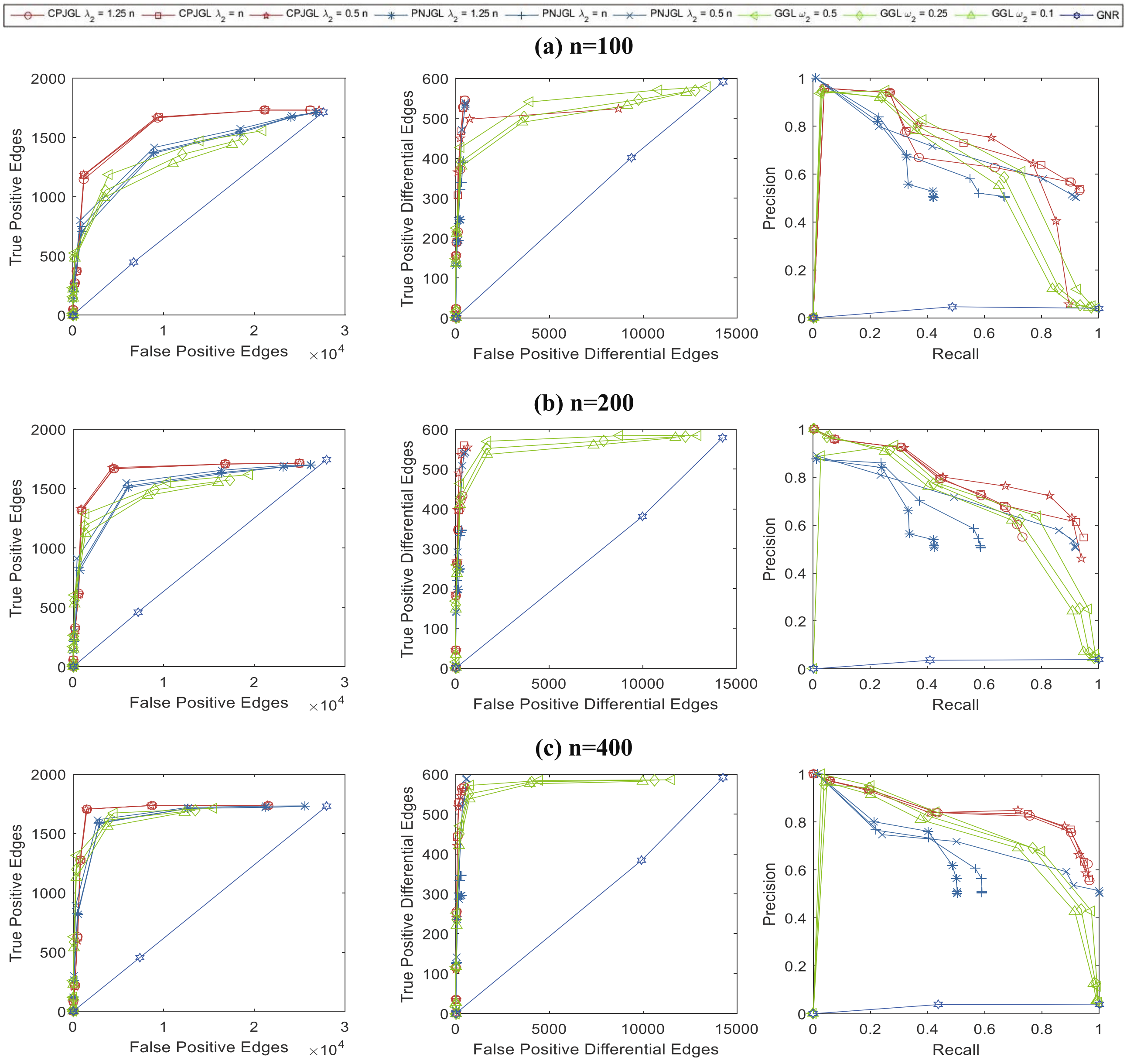
Fig. 2 displays the performance of the compared methods on simulation data with $p = 100$ and $n = 100, 200, 400$. In each plot, different colored lines correspond to the evaluation results of different methods with respect to different tuning parameter values. For example, each colored line for CPJGL corresponds to the results obtained with a fixed value of the tuning parameter λ_2, as the value of λ_1 varied (here we set $\lambda_0 = \lambda_1$). As shown in Fig. 2, CPJGL outperforms PNJGL for a suitable range of parameter λ_2. PNJGL exploits the similarity between two group-specific networks, but this method can only analyze each data platform separately. The comparison with PNJGL shows the gain of performance when CPJGL utilizes group sparse penalty to borrow strength across multiple data platforms. We can also observe from Fig. 2 that CPJGL performs better than GGL and GNR in both network inference and differential network analysis. GGL is designed to jointly estimate multiple precision matrices that share certain characteristics, but it can only infer each group-specific network separately and does not consider the hub structure of the differential network. GNR can borrows strength across multiple datasets to infer a group-specific network, but like GGL, it can only infer each group-specific network separately and does not consider the hub structure of the differential network. The comparison with GGL and GNR demonstrates the advantage of making use of the characteristics shared by different patient groups and different data platforms.

4. Identification of differential networks associated with ovarian cancer

In this section, we present the experimental results of applying CPJGL to real gene expression data sets.

4.1. Data sets

The development of drug resistance is the predominant cause of treatment failure and death in ovarian cancer [25]. In this experiment, we aim to reconstruct the gene regulatory networks of two groups of ovarian tumors (i.e., drug sensitivity patient group and drug resistance patient group), as well as to identify genes whose interactions with other genes vary significantly between the two patient groups. We download the gene expression profiles (level 3) for ovarian cancer patients from The Cancer Genome Atlas (TCGA) website (on February 2016). In this study, we consider three platforms, namely, Agilent 244 K Custom Gene Expression G450, Affymetrix HT Human Genome U133 Array Plate Set, and Affymetrix Human Exon 1.0 ST Array [26]. In particular, we have collected gene expression levels of 11,750 genes for 514 patients



(a) n=100

(b) n=200

(c) n=400

Fig. 2. Performance of different methods with $p = 100, K = 3$ and (a) $n = 100$, (b) $n = 200$, and (c) $n = 400$. Each colored curve corresponds to a fixed value of λ_2 (ω_2 for GGL), with λ_1 (ω_1 for GGL and λ for GNR) varied. Red line: CPJGL; blue line: PNJGL; green line: grouped graphical lasso (GGL); dark blue line: GNR. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

across all these three platforms. To make the data more normally distributed, we apply a logarithmic transformation on each data set. In the following experiments, for the sake of convenience, we refer to these three platforms as G450, U133 and HuEx, respectively.

According to the criterion used in [27], we divided the patients into two groups, namely platinum sensitivity patient group and platinum resistance patient group. A patient is defined as platinum sensitive if there is no evidence of disease progression within 6 months from the end of the last primary treatment, and the follow-up interval is at least 6 months from the date of last primary treatment. Patients with evidence of disease progression within 6 months from the end of primary treatment are defined as platinum resistant. Among the 514 patients, 340 patients have explicit platinum status, with 242 platinum sensitive and 98 platinum resistant. For each platform, we normalize the gene expression data to have mean zero and standard deviation one within each patient group.

The mTOR signaling pathway has been suggested to be frequently mutated in ovarian cancer [26], and is frequently implicated in resistance to anticancer therapies [28]. Besides mTOR signaling pathway, TGF-β signaling pathway also plays an important role in drug resistance [29]. Therefore, we take a pathway-based analysis in this study. We download the mTOR signaling pathway and TGF-β signaling pathway from the Kyoto Encyclopedia of Genes and Genomes database [30]. Out of the 60 genes in the mTOR signaling pathway and the 84 genes in the TGF-β signaling pathway, there are 51 and 75 genes included in our considered gene expression data sets, respectively. The standardized gene expression information for each sample can be downloaded from https://github.com/Oyl-CityU/CPJGL.

4.2. Differential network analysis

We first use CPJGL to infer the network difference among genes belonging to mTOR signaling pathway, based on gene expression

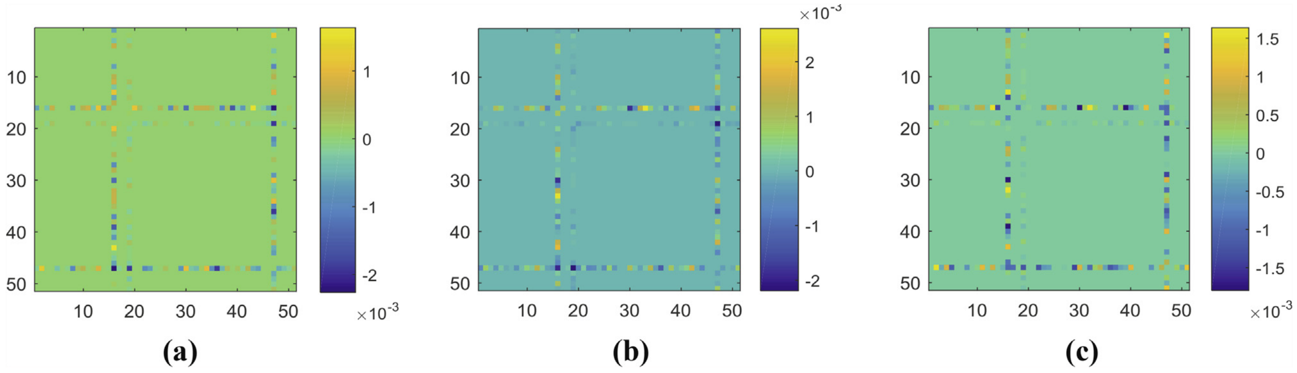


Fig. 3. The differential networks between drug sensitivity and drug resistance patient groups (mTOR signaling pathway) estimated by CPJGL from three data platforms (a) G450, (b) U133 and (c) HuEx, with $\lambda_0 = \lambda_1 = 7.84$ and $\lambda_2 = 171.5$. Three genes, namely TSC1, IRS1 and PDPK1 are identified as hub nodes in the three differential networks.

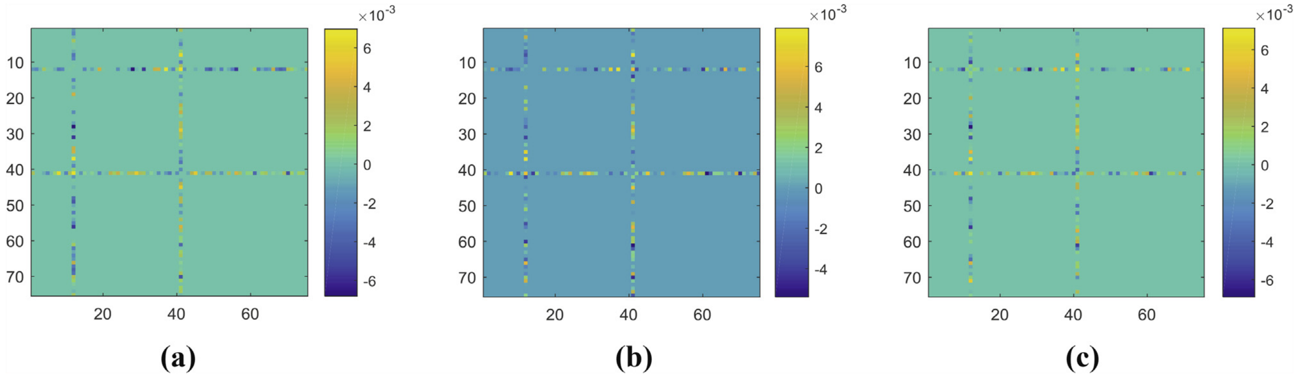


Fig. 4. The differential networks between drug sensitivity and drug resistance patient groups (TGF-β signaling pathway) estimated by CPJGL from three data platforms (a) G450, (b) U133 and (c) HuEx, with $\lambda_0 = \lambda_1 = 7.84$ and $\lambda_2 = 196$. Two genes, namely MYC and BMP7 are identified as hub nodes in the three differential networks.

data collected from the above three platforms. The tuning parameters $(\lambda_0, \lambda_1, \lambda_2)$ of CPJGL are selected by AIC. According to the selection result, we set $\lambda_0 = \lambda_1 = 7.84$ and $\lambda_2 = 171.5$. The estimated differential networks for three platforms are shown in Fig. 3.

As shown in Fig. 3, owing to the group lasso penalty, the differential networks identified by CPJGL from K data platforms have same structures. We can find from Fig. 3 that three hub nodes in the differential network, namely TSC1, IRS1 and PDPK1, are identified. TSC1 interacts with TSC2 to form a multi-protein complex (TSC complex) which acts as an important integrator of different signaling pathways controlling mTOR signaling [31]. The TSC complex has been implicated as a tumor suppressor [32]. IRS1 is an important growth-regulatory protein. Researches have found that the majority of malignant epithelial ovarian tumors show IRS1 overexpression when compared with normal ovarian tissue, which means IRS1 may be a possible target in ovarian cancer. As a signalling adapter protein, IRS1 is an important mediator of ovarian cancer cell growth suppression and can be a potential effective target for chemotherapeutic intervention [33]. Recently, Eckstein have reported that IRS1 has a functional role in cancer progress and platinum resistance [34]. PDPK1 is a master kinase, which plays an important role in the signalling pathways activated by several growth factors and hormones including insulin signaling. Lohneis et al. have found a negative correlation between PDPK1 expression and ovarian tumor grade, which indicates that PDPK1 might be a prognostic marker and a possible therapeutic target in ovarian serous carcinoma [35]. Moreover, Wu et al. have found that PDPK1 is associated with chemoresistance in ovarian cancer cells [36]. Therefore, our identified hubs TSC1, IRS1 and PDPK1, which are likely to be associated with platinum resistance in ovarian cancer, provide valuable insights for drug resistance analysis.

When applying CPJGL on the TGF-β signaling pathway, we set $\lambda_0 = \lambda_1 = 7.84$ and $\lambda_2 = 196$ according to AIC. The estimated differential networks for three platforms are shown in Fig. 4.

As can be seen from Fig. 4, there are two key hub nodes in the estimated differential networks: MYC and BMP7. MYC is a regulator gene that codes for a transcription factor. Amplification of the MYC gene has been found in a significant number of epithelial ovarian cancer cases [37]. Research has shown that MYC is associated with faster recurrence and poor overall survival of patients with high-grade serous ovarian cancer, and with cisplatin resistance in ovarian cancer cells [38]. In ovarian cancer, overexpression of BMP7 has been reported as dysregulated by microarray analysis [39]. Recently, BMP7 has been identified as a candidate gene that might play a role in secondary drug resistance [40]. Therefore, BMP7 might be potentially used as a molecular target to improve the treatment of platinum-resistant tumors.

5. Conclusions

In this study, a novel node-based multi-view learning algorithm called co-perturbed node joint graphical lasso (CPJGL) model is developed to jointly estimate multiple gene regulatory networks corresponding to different patient groups and the differential networks between these patient groups, based on gene expression data collected from multiple data platforms. Unlike previous differential network analysis methods that assume the differences between networks are driven by individual edges, our model made more reasonable assumption that the differences between networks are driven by certain regulator genes that are perturbed. Moreover, instead of inferring the differential network from each data platform separately, we jointly estimate multiple differential networks from multi-view gene expression data, which can exploit the common information provided by different views of data. Based on the idea of row-column overlap norm and group lasso penalty, our penalty function encourages the appearance of hub nodes in the estimated differential networks and encourages a similar pattern of sparsity across all precision matrices and precision matrix difference. Simulation studies demonstrate that our method consistently outperforms existing state-of-the-art network inference and differential network analysis algorithms. We apply our method to TCGA ovarian gene expression data collected from three platforms to study network rewiring associated with drug resistance. Once again, the experiment results demonstrate the potential of our method in detecting edges and genes that could provide insights into the molecular mechanisms of drug resistance.

When analyzing real data, due to the lack of gold standard, we are unable to evaluate the performance of network inference and differential network analysis methods. Therefore, it is difficult to compare different methods in terms of the accuracy of the estimated networks and differential networks. As a result, we only compare our method with state-of-the-art methods using simulated data and apply our method on real data to explore the network rewiring associated with drug resistance.

Besides graphical models, partial correlation and part mutual information have also been used to infer the direct interactions among genes [9,41,42]. However, when inferring differential networks, these correlation-based methods need to estimate each group-specific gene regulatory network separately, which totally ignore the similarity between the two group-specific gene regulatory networks. Based on graphical models, any structures of interest (such as the consistency across multiple data platforms) can be easily imposed on the resulting differential networks by utilizing suitable penalty functions. Thus, in this study, we use a graphical model to formulate the differential network estimation problem as a statistical learning problem.

Recently, a novel statistical method has been developed to infer individual-specific networks based on statistical perturbation analysis of a single sample against a group of given control samples [43]. Elucidating molecular mechanisms of individual-specific diseases may be beneficial to personalized diagnosis and individualized treatment. However, as a graphical model, our model can only infer an aggregated network for a group of samples. In the future, we will investigate how to utilize data collected from multiple platforms to identify sample-specific differential networks. Another limitation of our method is the assumption of the Gaussian distribution, since this assumption only holds for microarray-based gene expression data. In the future, we will consider how to extend our model to handle non-Gaussian data.

Competing interests

The authors declare that they have no competing interests.

Authors contributions

LOY, XFZ, MW and XLL conceived and designed the experiments. LOY and XFZ performed the experiments. LOY and XFZ analyzed the data. LOY and MW draft the manuscript under the guidance and supervision of XLL. All coauthors have seen a draft copy of the manuscript and agree with its publication.

Acknowledgements

This work is supported by the National Science Foundation of China (61402190, 61532008, 61602309), Self-determined Research Funds of CCNU from the colleges' basic research and operation of MOE (CCNU15ZD011).

References

[1] A.-L. Barabási, N. Gulbahce, J. Loscalzo, Network medicine: a network-based approach to human disease, Nat. Rev. Genet. 12 (1) (2011) 56–68.
[2] R. Küffner et al., Inferring gene regulatory networks by anova, Bioinf. 28 (10) (2012) 1376–1382.
[3] L. Ou-Yang et al., Detecting temporal protein complexes from dynamic protein-protein interaction networks, BMC Bioinf. 15 (1) (2014) 335.
[4] T. Ideker, N.J. Krogan, Differential network biology, Mol. Syst. Biol. 8 (1) (2012) 565.
[5] W. Kolch, M. Halasz, M. Granovskaya, B.N. Kholodenko, The dynamic control of signal transduction networks in cancer cells, Nat. Rev. Cancer 15 (9) (2015) 515–527.
[6] Y. Liu, X. Zeng, Z. He, Q. Zou, Inferring microrna-disease associations by random walk on a heterogeneous network with multiple data sources, IEEE/ACM Trans. Comput. Biol. Bioinf. PP (99) (2016). 1–1.
[7] R. De Smet, K. Marchal, Advantages and limitations of current network inference methods, Nat. Rev. Microbiol. 8 (10) (2010) 717–729.
[8] D. Marbach et al., Wisdom of crowds for robust gene network inference, Nat. Methods 9 (8) (2012) 796–804.
[9] X. Zhang, X.-M. Zhao, K. He, L. Lu, Y. Cao, J. Liu, J.-K. Hao, Z.-P. Liu, L. Chen, Inferring gene regulatory networks from gene expression data by path consistency algorithm based on conditional mutual information, Bioinformatics 28 (1) (2012) 98–104.
[10] B. Alipanahi, B.J. Frey, Network cleanup, Nat. Biotechnol. 31 (8) (2013) 714–715.
[11] S.L. Lauritzen, Graphical Models, Oxford Press, 1996.
[12] C.H. Zheng et al., Gene differential coexpression analysis based on biweight correlation and maximum clique, BMC Bioinf. 15 (Suppl 15) (2014) S3.
[13] P. Danaher, P. Wang, D.M. Witten, The joint graphical lasso for inverse covariance estimation across multiple classes, J. R. Stat. Soc. 76 (2) (2014) 373–397.
[14] K. Mohan, P. London, M. Fazel, D.M. Witten, S.-I. Lee, Node-based learning of multiple gaussian graphical models, J. Mach. Learn. Res. 15 (1) (2014) 445–488.
[15] J. Guo, E. Levina, G. Michailidis, J. Zhu, Joint estimation of multiple graphical models, Biometrika 98 (1) (2011) 1–15.
[16] W. Lee, Y. Liu, Joint estimation of multiple precision matrices with common structures, J. Mach Learn. Res. 16 (2015) 1035–1062.
[17] A.G. Deshwar, Q. Morris, Plida: cross-platform gene expression normalization using perturbed topic models, Bioinformatics 30 (7) (2014) 956–961.
[18] Y. Wang, T. Joshi, X.-S. Zhang, D. Xu, L. Chen, Inferring gene regulatory networks from multiple microarray datasets, Bioinformatics 22 (19) (2006) 2413–2420.
[19] J. Friedman, T. Hastie, R. Tibshirani, A note on the group lasso and a sparse group lasso, (2010), arXiv preprint <arXiv:1001.0736>.
[20] M. Yuan, Y. Lin, Model selection and estimation in the gaussian graphical model, Biometrika 94 (1) (2007) 19–35.
[21] J. Friedman, T. Hastie, R. Tibshirani, Sparse inverse covariance estimation with the graphical lasso, Biostatistics 9 (3) (2008) 432–441.
[22] A.J. Rothman, P.J. Bickel, E. Levina, J. Zhu, Sparse permutation invariant covariance estimation, Electron. J. Stat. 2 (2008) 494–515.
[23] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, Found. Trends Mach. Learn. 3 (1) (2011) 1–122.
[24] K.M. Tan, P. London, K. Mohan, S.-I. Lee, M. Fazel, D.M. Witten, Learning graphical models with hubs, J. Mach. Learn. Res. 15 (1) (2014) 3297–3331.
[25] D.D. Bowtell et al., Rethinking ovarian cancer ii: reducing mortality from high-grade serous ovarian cancer, Nat. Rev. Cancer 15 (11) (2015) 668–679.
[26] C.G.A.R. Network et al., Integrated genomic analyses of ovarian carcinoma, Nature 474 (7353) (2011) 609–615.
[27] S. Nabavi, D. Schmolze, M. Maitituoheti, S. Malladi, A.H. Beck, Emdomics: a robust and powerful method for the identification of genes differentially expressed between heterogeneous classes, Bioinformatics (2015), http://dx.doi.org/10.1093/bioinformatics/btv634. btv634.
[28] H.A. Burris III, Overcoming acquired resistance to anticancer therapy: focus on the pi3k/akt/mtor pathway, Cancer Chemother. Pharmacol. 71 (4) (2013) 829–842.
[29] D. Brunen, S. Willems, U. Kellner, R. Midgley, I. Simon, R. Bernards, Tgf-β: an emerging player in drug resistance, Cell Cycle 12 (18) (2013) 2960–2968.
[30] M. Kanehisa, S. Goto, Kegg: kyoto encyclopedia of genes and genomes, Nucleic Acids Res. 28 (1) (2000) 27–30.
[31] K. Inoki, M.N. Corradetti, K.-L. Guan, Dysregulation of the tsc-mtor pathway in human disease, Nat. Genet. 37 (1) (2005) 19–24.
[32] D.M. Sabatini, mtor and cancer: insights into a complex relationship, Nat. Rev. Cancer 6 (9) (2006) 729–734.

[33] S. Ravikumar et al., Insulin receptor substrate-1 is an important mediator of ovarian cancer cell growth suppression by all-trans retinoic acid, Cancer Res. 67 (19) (2007) 9266–9275.

[34] N. Eckstein, Platinum resistance in breast and ovarian cancer cell lines, J. Exp. Clin. Cancer Res. 30 (1) (2011) 1.

[35] P. Lohneis et al., Pdk1 is expressed in ovarian serous carcinoma and correlates with improved survival in high-grade tumors, Anticancer Res. 35 (11) (2015) 6329–6334.

[36] Y.-H. Wu, T.-H. Chang, Y.-F. Huang, C.-C. Chen, C.-Y. Chou, Col11a1 confers chemoresistance on ovarian cancer cells through the activation of akt/c/ebpβ pathway and pdk1 stabilization, Oncotarget 6 (27) (2015) 23748–23763.

[37] C.-H. Chen, J. Shen, W.-J. Lee, S.-N. Chow, Overexpression of cyclin d1 and c-myc gene products in human primary epithelial ovarian cancer, Int. J. Gynecol. Cancer 15 (5) (2005) 878–883.

[38] J.M. Reyes-González, G.N. Armaiz-Peña, L.S. Mangala, F. Valiyeva, C. Ivan, S. Pradeep, I.M. Echevarŕa-Vargas, A. Rivera-Reyes, A.K. Sood, P.E. Vivas-Mejía, Targeting c-myc in platinum-resistant ovarian cancer, Mol. Cancer Ther. 14 (10) (2015) 2260–2269.

[39] H. Donninger, T. Bonome, M. Radonovich, C.A. Pise-Masison, J. Brady, J.H. Shih, J.C. Barrett, M.J. Birrer, Whole genome expression profiling of advance stage papillary serous ovarian cancer reveals activated pathways, Oncogene 23 (49) (2004) 8065–8077.

[40] V. Camara-Clayette, S. Koscielny, S. Roux, T. Lamy, J. Bosq, M. Bernard, T. Fest, V. Lazar, G. Lenoir, V. Ribrag, Bmp7 expression correlates with secondary drug resistance in mantle cell lymphoma, PLoS One 8 (9) (2013) e73993.

[41] X. Zhang, J. Zhao, J.-K. Hao, X.-M. Zhao, L. Chen, Conditional mutual inclusive information enables accurate quantification of associations in gene regulatory networks, Nucleic Acids Res. 43 (5) (2015). e31–e31.

[42] J. Zhao, Y. Zhou, X. Zhang, L. Chen, Part mutual information for quantifying direct associations in networks, Proc. Nat. Acad. Sci. 113 (18) (2016) 5130–5135.

[43] X. Liu, Y. Wang, H. Ji, K. Aihara, L. Chen, Personalized characterization of diseases using sample-specific networks, Nucleic Acids Res. (2016), http://dx.doi.org/10.1093/nar/gkw772.