

University of Groningen

GeneNetTools

Bernal, Victor; Soanatl-Aguilar, Venustiano; Bulthuis, Jonas; Guryev, Victor; Horvatovich, Peter; Grzegorzcyk, Marco

Published in:
Bioinformatics

DOI:
[10.1093/bioinformatics/btac657](https://doi.org/10.1093/bioinformatics/btac657)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2022

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Bernal, V., Soanatl-Aguilar, V., Bulthuis, J., Guryev, V., Horvatovich, P., & Grzegorzcyk, M. (2022). GeneNetTools: Tests for Gaussian graphical models with shrinkage. Unpublished. Advance online publication. <https://doi.org/10.1093/bioinformatics/btac657>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



Gene expression

GeneNetTools: tests for Gaussian graphical models with shrinkage

Victor Bernal ^{1,2}, Venustiano Soancatl-Aguilar¹, Jonas Bulthuis¹, Victor Guryev³, Peter Horvatovich ^{4,*†} and Marco Grzegorzczuk^{2,†}

¹Center of Information Technology, University of Groningen, Groningen 9747 AJ, The Netherlands, ²Department of Mathematics, Bernoulli Institute, University of Groningen, Groningen 9747 AG, The Netherlands, ³European Research Institute for the Biology of Ageing, University Medical Center Groningen, University of Groningen, Groningen 9713 AV, The Netherlands and ⁴Department of Analytical Biochemistry, Groningen Research Institute of Pharmacy, University of Groningen, Groningen 9713 AV, The Netherlands

*To whom correspondence should be addressed.

†The authors wish it to be known that, in their opinion, the last two authors should be regarded as Joint Last Authors.

Associate Editor: Inanc Birol

Received on June 22, 2022; revised on September 14, 2022; editorial decision on September 21, 2022; accepted on September 29, 2022

Abstract

Motivation: Gaussian graphical models (GGMs) are network representations of random variables (as nodes) and their partial correlations (as edges). GGMs overcome the challenges of high-dimensional data analysis by using shrinkage methodologies. Therefore, they have become useful to reconstruct gene regulatory networks from gene-expression profiles. However, it is often ignored that the partial correlations are ‘shrunk’ and that they cannot be compared/assessed directly. Therefore, accurate (differential) network analyses need to account for the number of variables, the sample size, and also the shrinkage value, otherwise, the analysis and its biological interpretation would turn biased. To date, there are no appropriate methods to account for these factors and address these issues.

Results: We derive the statistical properties of the partial correlation obtained with the Ledoit–Wolf shrinkage. Our result provides a toolbox for (differential) network analyses as (i) confidence intervals, (ii) a test for zero partial correlation (null-effects) and (iii) a test to compare partial correlations. Our novel (parametric) methods account for the number of variables, the sample size and the shrinkage values. Additionally, they are computationally fast, simple to implement and require only basic statistical knowledge. Our simulations show that the novel tests perform better than DiffNetFDR—a recently published alternative—in terms of the trade-off between true and false positives. The methods are demonstrated on synthetic data and two gene-expression datasets from *Escherichia coli* and *Mus musculus*.

Availability and implementation: The R package with the methods and the R script with the analysis are available in <https://github.com/V-Bernal/GeneNetTools>.

Contact: p.l.horvatovich@rug.nl

Supplementary information: [Supplementary data](#) are available at *Bioinformatics* online.

1 Introduction

A Gaussian graphical model (GGM) (Edwards, 2000) consists of a network structure where random variables are nodes and their partial correlations are edges. Partial correlations are a full-conditional (linear) measure of the association between pairs of random variables, thus the GGM characterizes conditional independences.

To compute the partial correlations, it is necessary to standardize the inverse of the covariance matrix (i.e. the precision matrix). While the covariance matrix can always be estimated from data, in this case, the estimated matrix must be invertible and well-conditioned. This requirement ensures that the inverse of the covariance matrix exists and that its computation is stable (not damaged

by numerical or estimation errors). For a dataset of n samples and p variables, the sample covariance estimator is invertible and well-conditioned only when n is greater than p . In other cases, it is invertible but ill-conditioned when n is comparable to p , or not even invertible when n is smaller than p (Ledoit and Wolf, 2004). These last two cases are common in large-scale applications, particularly in bioinformatics, where molecular information from a large set of genes are measured from few samples, such as transcripts, or proteins, and are often referred to as ‘high-dimensions’, ‘small n , large p ’ or ‘ $n \ll p$ ’ scenarios.

Shrinkage, a type of regularization, deals with the ‘high-dimensional problem’ by stabilizing the estimator. Among the shrinkage approaches, we find Glasso (Friedman *et al.*, 2008) and

the Ledoit–Wolf (LW) shrinkage (Ledoit and Wolf, 2003, 2004). Lasso estimates a sparse precision matrix with a L_1 penalty to force some of its entries to be zero. The LW-shrinkage estimates an invertible covariance (or correlation) matrix using a bias towards a sparser matrix structure. These methodologies have made GGMs popular for large-scale applications, e.g. bioinformatics and biomedicine (Beerenwinkel et al., 2007; Benedetti et al., 2017; Das et al., 2017; Imkamp et al., 2019; Keller et al., 2008; McNally et al., 2015), where gene-expression alteration and (condition-specific) gene networks would reflect the underlying biological mechanisms (Barabási et al., 2011).

However, one important challenge arises which is frequently ignored. These ‘shrunk’ partial correlations cannot be compared directly; different datasets imply distinct shrinkage values, and ultimately, different biases (Bernal et al., 2019). For L_1 penalty methods there are some methods to quantify pairs-wise changes in the precision and/or the partial correlation matrix (Liu, 2017; Yuan et al., 2017; Zhang et al., 2019), and consequently for differential network analysis (Class et al., 2018; Zhang et al., 2018, 2019). To the best of our knowledge, and despite its wide use, there is no approach for differential network analysis based on the LW-shrinkage GGM.

In this article, we adapt three classical (parametric) statistics to the LW-shrinkage (Schäfer and Strimmer, 2005). Our results provide a toolbox for (differential) network analysis that include (i) confidence intervals, (ii) a test for null effects and (iii) a test to compare ‘shrunk’ partial correlations. Each of these account for the number of variables, the sample size and shrinkage values.

2 Materials and methods

In this section, we present the LW shrinkage and GGMs (Ledoit and Wolf, 2003, 2004). We show how the shrinkage can be included in several test statistics. To this end, we will study a rescaled version of the partial correlation, which will prove to be easier to interpret and advantageous to develop statistical tests.

Throughout the manuscript, matrices are represented with uppercase bold letters, and estimators are denoted with a hat symbol (e.g. \mathbf{X} is a matrix, and $\hat{\mathbf{X}}$ is an estimator of \mathbf{X}).

2.1 The ‘shrunk’ partial correlation

GGMs are network models where random variables are represented with nodes and partial correlations with edges. The partial correlation is a full-conditional correlation; it measures the linear association between two Gaussian variables, while all the others are held constant.

For a dataset of p variables and n samples, there are $p(p-1)/2$ partial correlations in total, which can be computed via

$$P_{ij} = \frac{-\Omega_{ij}}{\sqrt{\Omega_{ii}\Omega_{jj}}}, \quad (1)$$

where P_{ij} denotes the partial correlation between the i -th and j -th variables, and Ω is the inverse of the $p \times p$ covariance matrix \mathbf{C} (or equivalently, the inverse of the correlation matrix \mathbf{R}).

The covariance matrix \mathbf{C} can be estimated from data, e.g. with the sample covariance matrix $\hat{\mathbf{C}}^{\text{SM}}$, however, this task turns challenging when n is comparable to, or smaller than, p . The reason is that $\hat{\mathbf{C}}^{\text{SM}}$ becomes ill-conditioned (numerically unstable) or singular (non-invertible). In other words, the number of parameters to estimate (all the partial correlations) is too large relative to the amount of information available in the dataset (the sample size). This same issue arises with the sample correlation matrix $\hat{\mathbf{R}}^{\text{SM}}$. For instance, the Pearson’s correlation coefficient has degrees of freedom $k = n - 1$, and the partial correlation coefficient $k = n - (p - 2) - 1$. In the well-conditioned case ($n \geq p + 1$), both degrees of freedom are positive. In the ill-conditioned case ($n < p + 1$), for the partial correlation k would turn negative (and meaningless).

The LW shrinkage overcomes this issue via a ‘shrunk’ covariance matrix $\hat{\mathbf{C}}^{[\lambda]}$ defined as,

$$\hat{\mathbf{C}}^{[\lambda]} := (1 - \lambda)\hat{\mathbf{C}}^{\text{SM}} + \lambda\mathbf{T}. \quad (2)$$

Here $\lambda \in (0, 1)$, also called the shrinkage value, represents the weight allocated to a target matrix \mathbf{T} . The shrinkage λ has an optimal value that is obtained by minimizing the mean square error (Schäfer and Strimmer, 2005). In this work, \mathbf{T} is taken as a diagonal matrix of variances, though other alternatives are possible. This choice shrinks the magnitudes of the covariances (off-diagonal), while the variances (diagonal) remain intact, which is equivalent to shrinking \mathbf{R}^{SM} towards the identity matrix.

Using the inverse of $\hat{\mathbf{C}}^{[\lambda]}$, denoted here by $\hat{\Omega}^{[\lambda]}$, Equation (1) becomes,

$$\hat{P}_{ij}^{[\lambda]} = \frac{-\hat{\Omega}_{ij}^{[\lambda]}}{\sqrt{\hat{\Omega}_{ii}^{[\lambda]}\hat{\Omega}_{jj}^{[\lambda]}}}, \quad (3)$$

which is the ‘shrunk’ partial correlation between the i -th and j -th variables (Schäfer and Strimmer, 2005).

2.2 The distribution of the partial correlation—GeneNet

In the classical scenario, the Pearson’s correlation coefficient r and the partial correlation coefficient ρ follow the same probability distribution f_0 , differing only in their degrees of freedom k (Fisher, 1924).

Under the null hypothesis $H_0 : \rho = 0$, the density of the partial correlation is

$$f_0(\rho) = \frac{(1 - \rho^2)^{(k-3)/2}}{\text{Beta}\left(\frac{1}{2}, \frac{k-1}{2}\right)}, \quad (4.1)$$

with $k = n - 1 - (p - 2)$, which turns negative whenever $n < p + 1$.

Equation (4.1) is used in the popular method GeneNet (Schäfer and Strimmer, 2005). This method computes ‘shrunk’ partial correlations with Equation (3) and (approximate) its P -values via empirical null-fitting. It relies in that Equation (4.1) is a reasonable approximation to the real ‘shrunk’ probability density, at least for small shrinkage values and sparse networks.

However, as analytical results about Equation (3) have been missing, appropriated discussions of effect size and significance have remained limited.

2.3 The distribution of the ‘shrunk’ partial correlation

In the ‘shrunk’ scenario, the same holds; the distribution of the ‘shrunk’ correlation $r^{[\lambda]}$ and ‘shrunk’ partial correlation $\rho^{[\lambda]}$ are the same, differing only in their degrees of freedom $k^{[\lambda]}$ (Bernal et al., 2019).

Under the null hypothesis of $H_0 : \rho^{[\lambda]} = 0$, the density of the ‘shrunk’ partial correlation is

$$f_0^{[\lambda]}(\rho^{[\lambda]}) = \frac{\left((1 - \lambda)^2 - \rho^{[\lambda]2}\right)^{(k^{[\lambda]}-3)/2}}{\text{Beta}\left(\frac{1}{2}, \frac{k^{[\lambda]}-1}{2}\right)(1 - \lambda)^{(k^{[\lambda]}-2)}, \quad (4.2)$$

where $k^{[\lambda]}$ has no closed form, but can be estimated via maximum likelihood.

At this point, we switch our attention to $y = (1 - \lambda)^{-1}\rho^{[\lambda]}$ which will prove convenient to derive some useful results. This transformation rescales the magnitude of $\rho^{[\lambda]}$, while the configuration space remains intact. The density in Equation (4.2) satisfies that

$$\int_{-(1-\lambda)}^{(1-\lambda)} f_0^{[\lambda]}(\rho^{[\lambda]}) d\rho^{[\lambda]} = 1.$$

Replacing $(1 - \lambda)y = \rho^{[\lambda]}$ and $(1 - \lambda)dy = d\rho^{[\lambda]}$ gives

$$\int_{-1}^1 \left(\frac{\left((1-\lambda)^2 - ((1-\lambda)y)^2 \right)^{\frac{k^{[2]}-3}{2}}}{\text{Beta}\left(\frac{1}{2}, \frac{k^{[2]}-1}{2}\right) (1-\lambda)^{k^{[2]}-2}} \right) dy = 1$$

and factoring the terms with $(1-\lambda)$, we have that

$$f_0^{[2]} \left(\frac{\rho^{[2]}}{(1-\lambda)} \right) = \frac{\left(1 - \left(\frac{\rho^{[2]}}{(1-\lambda)} \right)^2 \right)^{\frac{k^{[2]}-3}{2}}}{\text{Beta}\left(\frac{1}{2}, \frac{k^{[2]}-1}{2}\right)}. \quad (4.3)$$

Which is equal to Equation (4.1). While $\rho^{[2]} \in [-(1-\lambda), (1-\lambda)]$, the rescaled version $(1-\lambda)^{-1}\rho^{[2]} \in [-1, 1]$.

In other words, $(1-\lambda)^{-1}\rho^{[2]}$ is distributed as Pearson's correlation with degrees of freedom $k^{[2]}$. Therefore, methodologies developed for Pearson's correlation can be readily adapted to the rescaled 'shrunk' partial correlation $(1-\lambda)^{-1}\rho^{[2]}$.

2.4 Test for null-effects $H_0 : \rho^{[2]} = 0$

Pearson's and partial correlation coefficients are often tested using a t -statistic.

Let $\hat{\rho}$ be the estimated (partial) correlation from data, with population value ρ_{true} and degrees of freedom k . Then, under the null hypothesis of a zero-effect $H_0 : \rho_{\text{true}} = 0$,

$$t_{v=k-1} = \hat{\rho} \sqrt{\frac{k-1}{1-\hat{\rho}^2}}, \quad (5)$$

follows a Student's t -distribution with $v = k - 1$. Equation (5) holds for Pearson's and partial correlations each with their corresponding k (Cohen, 1988; Levy and Narula, 1978). An equivalent test for $\rho^{[2]}$ is straightforward due to Equations (4.1)–(4.3).

Under the null hypothesis $H_0 : \hat{\rho}_{\text{true}}^{[2]} = 0$,

$$t_{v=k^{[2]}-1}^{[2]} = \frac{\hat{\rho}^{[2]}}{(1-\lambda)} \sqrt{\frac{k^{[2]}-1}{1-\left(\frac{\hat{\rho}^{[2]}}{(1-\lambda)}\right)^2}}, \quad (6)$$

which has a Student's t -distribution with degrees of freedom $v = k^{[2]} - 1$, and tests the null hypothesis of a zero 'shrunk' partial correlation.

2.5 Confidence intervals of $\rho^{[2]}$

Confidence intervals for Pearson's (and partial) correlations are commonly computed using Fisher's transformation $F(*) = \text{arctanh}(*)$.

Let $\hat{\rho}$ be the estimated partial correlation with population value ρ_{true} . Then, the Fisher-transformed $\hat{\rho}$, denoted here by $F(\hat{\rho})$, is normally distributed with expectation $F(\rho_{\text{true}})$ and standard error $\sqrt{(k-2)^{-1}}$, or

$$F(\hat{\rho}) \sim N\left(F(\rho_{\text{true}}), \frac{1}{k-2}\right). \quad (7)$$

Analogously, the Fisher's transformed 'shrunk' partial correlation $F\left((1-\lambda)^{-1}\hat{\rho}^{[2]}\right)$ is normally distributed with expectation $F\left((1-\lambda)^{-1}\rho_{\text{true}}^{[2]}\right)$ and standard error $\sqrt{(k^{[2]}-2)^{-1}}$.

In other words,

$$F\left(\frac{\hat{\rho}^{[2]}}{(1-\lambda)}\right) \sim N\left(F\left(\frac{\rho_{\text{true}}^{[2]}}{(1-\lambda)}\right), \frac{1}{k^{[2]}-2}\right), \quad (8.1)$$

or

$$\left(\frac{F\left(\frac{\hat{\rho}^{[2]}}{(1-\lambda)}\right) - F\left(\frac{\rho_{\text{true}}^{[2]}}{(1-\lambda)}\right)}{\sqrt{\frac{1}{(k^{[2]}-2)}}} \right) \sim N(0, 1). \quad (8.2)$$

Therefore, confidence limits for $\rho_{\text{true}}^{[2]}$ can be computed as follows. Let $z \sim N(0, 1)$ then,

$$P\left(-q_{(1-\frac{\alpha}{2})} \leq z \leq q_{(1-\frac{\alpha}{2})}\right) = \left(1 - \frac{\alpha}{2}\right),$$

where $q_{(1-\alpha)}$ is the $(1-\alpha)$ -th quantile of a normal random variable and the following inequality holds,

$$-q_{(1-\frac{\alpha}{2})} \leq \frac{F\left(\frac{\hat{\rho}^{[2]}}{(1-\lambda)}\right) - F\left(\frac{\rho_{\text{true}}^{[2]}}{(1-\lambda)}\right)}{\sqrt{\frac{1}{(k^{[2]}-2)}}} \leq q_{(1-\frac{\alpha}{2})}.$$

Both sides can be multiplied by $\sqrt{(k-2)^{-1}}$ and by minus one (which reverses the inequality), turning it into

$$\begin{aligned} F\left(\frac{\hat{\rho}^{[2]}}{(1-\lambda)}\right) + \frac{q_{(1-\frac{\alpha}{2})}}{\sqrt{(k^{[2]}-2)}} &\leq F\left(\frac{\rho_{\text{true}}^{[2]}}{(1-\lambda)}\right) \\ &\leq F\left(\frac{\hat{\rho}^{[2]}}{(1-\lambda)}\right) - \frac{q_{(1-\frac{\alpha}{2})}}{\sqrt{(k^{[2]}-2)}} \end{aligned}$$

applying the inverse transformation $F^{-1}(*) = \text{tanh}(*)$, we have that

$$P\left(\tanh(l^{[2]}) \leq \frac{\rho_{\text{true}}^{[2]}}{(1-\lambda)} \leq \tanh(u^{[2]})\right) = \left(1 - \frac{\alpha}{2}\right), \quad (9.1)$$

with

$$\begin{aligned} l^{[2]} &= F\left(\frac{\hat{\rho}^{[2]}}{(1-\lambda)}\right) - \sqrt{\frac{1}{(k^{[2]}-2)}} q_{(1-\frac{\alpha}{2})} \\ u^{[2]} &= F\left(\frac{\hat{\rho}^{[2]}}{(1-\lambda)}\right) + \sqrt{\frac{1}{(k^{[2]}-2)}} q_{(1-\frac{\alpha}{2})}. \end{aligned} \quad (9.2)$$

Equations (9.1) and (9.2) define the confidence intervals for the rescaled (population) 'shrunk' partial correlation $(1-\lambda)^{-1}\rho_{\text{true}}^{[2]}$.

Confidence intervals for the 'shrunk' partial correlation $\rho_{\text{true}}^{[2]} \in [-(1-\lambda), (1-\lambda)]$ can also be obtained multiplying Equations (9.1) and (9.2) by $(1-\lambda)$. However, the rescaled version is superior in terms of its interpretability, as it is distributed as the classical Pearson's correlation coefficient between -1 and 1 .

2.6 Test to compare partial correlations $H_0 : \rho^{[\lambda_1]} = \rho^{[\lambda_2]}$

Two (partial) correlation coefficients can be compared using Fisher's transformation. The test is built by subtracting two Fisher-transformed coefficients [see Equation (7)] and assuming that their (unknown) population values $F(\rho_{\text{true}})$ are equal.

Two 'shrunk' (partial) correlation coefficients can be compared in the same way. Let us suppose that $\hat{\rho}^{[\lambda_1]}$ and $\hat{\rho}^{[\lambda_2]}$ are estimates from two datasets with population values $\rho_{\text{true}}^{[\lambda_1]}$ and $\rho_{\text{true}}^{[\lambda_2]}$, and degrees of freedom $k^{[\lambda_1]}$ and $k^{[\lambda_2]}$. As discussed before, these estimates are not comparable *per se*, as they have different shrinkages—and scales—and a test for $H_0 : \rho_{\text{true}}^{[\lambda_1]} - \rho_{\text{true}}^{[\lambda_2]} = 0$ would be meaningless. However, their appropriately rescaled versions $(1-\lambda_1)^{-1}\hat{\rho}^{[\lambda_1]}$ and $(1-\lambda_2)^{-1}\hat{\rho}^{[\lambda_2]}$ are comparable.

Therefore, a test for $H_0 : (1-\lambda_1)^{-1}\rho_{\text{true}}^{[\lambda_1]} - (1-\lambda_2)^{-1}\rho_{\text{true}}^{[\lambda_2]} = 0$ is

$$z^{[2]} = \frac{F\left(\frac{\hat{\rho}^{[\lambda_1]}}{(1-\lambda_1)}\right) - F\left(\frac{\hat{\rho}^{[\lambda_2]}}{(1-\lambda_2)}\right)}{\sqrt{\frac{1}{(k^{[\lambda_1]}-2)} + \frac{1}{(k^{[\lambda_2]}-2)}}} \sim N(0, 1). \quad (10)$$

This z -statistics is distributed as $N(0, 1)$. It tests whether two 'shrunk' partial correlations estimated from independent datasets are statistically different.

The practical implementation of the results in this section can be found in the [Supplementary Material S1](#).

2.6.1 Data

2.6.1.1 *Escherichia coli* microarray data. This dataset consists of *E. coli* microarray gene-expression measurements. The study

explores the temporal stress response upon expression of recombinant human superoxide dismutase (SOD) (Schmidt-Heck et al., 2004), induced by isopropyl β -D-1-thiogalactopyranoside (a lactose analogue inducer of the lac operon). The measurements were collected at 8, 15, 22, 45, 68, 90, 150 and 180 min. In total, 102 out of 4289 protein-coding genes are differentially expressed across the nine time points. A \log_2 -ratio transformation of transcript microarray intensity was applied with respect to the initial time points. The dataset was obtained from the R package *GeneNet* version 1.2.13. Accessed on April 7, 2022.

2.6.1.2 *Mus musculus* RNA-sequencing data. Data are from single-end RNA-Seq reads from 21 male mice from two strains (B6, $n = 10$ and D2, $n = 11$). The dataset was downloaded from ReCount: <http://bowtie-bio.sourceforge.net/recount/> under the PubMed Identifier 21455293 (Bottmly et al., 2011), on April 7, 2022. Low expressed genes (<5 reads on average) were excluded from the data before pre-processing. In total, 223 genes out of 9431 are differentially expressed (adjusted by strain) with the R package *limma* at Benjamini–Hochberg adjusted false discovery rate <0.05 (Benjamini and Hochberg, 1995; Ritchie et al., 2015). Before statistical analysis the transcript quantitative values were \log_2 -transformed and upper quartile normalized.

3 Results

3.1 Analysis of simulated data

Here, we evaluate the performance of the proposed ‘shrunk’ z -test [Equation (10)] in terms of the area under the receiver operator curve (AUROC) and the area under the precision-recall curve (AUPRC). AUROCs and AUPRCs are metrics defined between 0 and 1 that measure the trade-off between true and false positives. We also compare our method against DiffNetFDR (Zhang et al., 2019). DiffNetFDR is a computational approach for differential network analysis that uses the residuals obtained from lasso multivariate regression (Liu, 2017). The authors of DiffNetFDR recently reported a better performance compared to several other alternatives (Class et al., 2018; Zhang et al., 2018).

We simulate data from one fixed network with 100 nodes and sample sizes 30, 40, 50, ... and 100. For each pair of datasets, we reconstruct the partial correlations and compare them with the z -score [Equation (10)]. Figure 1a and b shows the AUROCs and AUPRCs varying the sample size. Figure 1c and d shows the difference in AUROCs and AUPRCs (in percentages) between the z -score [Equation (10)] and DiffNetFDR.

The proposed z -statistics shows higher AUROCs and AUPRCs (upper panels), and performs better than DiffNetFDR (lower panels). Supplementary Figures S1 and S2 show the AUROCs and AUPRCs for different networks sizes and for different proportion of edges.

3.2 Analysis of experimental data

3.2.1 Effects of human SOD protein expression on transcript expression in *E.coli*

Here, we compare the P -values obtained with (i) the ‘shrunk’ t -test [Equation (6)] and (ii) the ‘shrunk’ probability density [Equation (4.2)]. Following previous works (Bernal et al., 2019; Schäfer and Strimmer, 2005), the dataset is treated as static. The significance level is 0.05, and the optimal shrinkage is 0.18.

The differences between the P -values are in the order 10^{-07} . This is smaller than the tolerance value of the numerical integration (see previous method) and can thus be considered to be (numerically) zero. Both methods retrieved 238 edges in agreement previous analysis (Bernal et al., 2019), though the ‘shrunk’ t -test was more than 10 times faster. The confidence intervals for the 15 strongest edges are displayed in Figure 2b. These include transcripts of *lacA*, *lacY* and *lacZ* genes, involved in the induction of the lac operon.

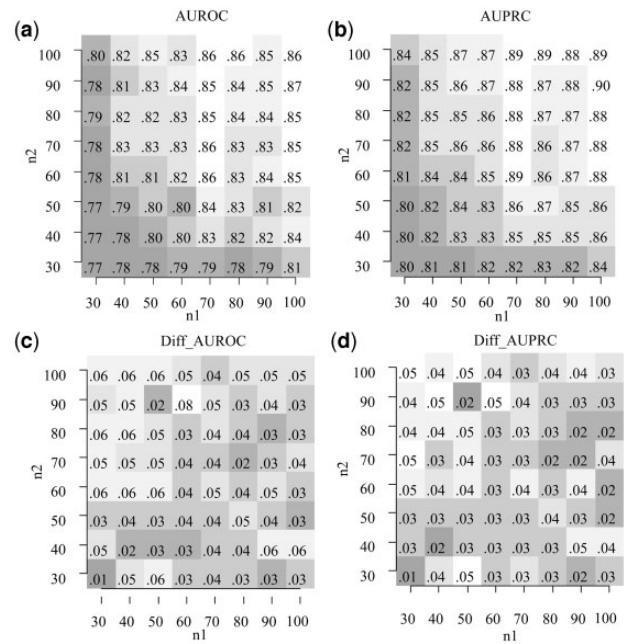


Fig. 1. AUROC and AUPRC. (a and b) AUROCs and AUPRCs of the z -score [Equation (10)]. The performance increases with the sample size. (c) AUROCs of the z -score minus the AUROCs of DiffNetFDR. (d) AUPRCs of the z -score minus the AUPRCs of DiffNetFDR. Positive values show a better trade-off of true and false positives for the proposed z -statistics. Data were simulated from networks with $p = 100$ nodes and sample sizes n_1 and n_2 between 30 and 100

The vertical lines show the 0.1 and 0.3 thresholds for weak and mild effects correlations (Cohen, 1988).

3.3 Analysis of *M.musculus* RNA-seq dataset

Here, we compare the GGMs of two mice strains B6 and D2. Figure 3 presents the $|z$ -score| that compares edge-wise the strains via Equation (10). The significance level is 0.05 (i.e. $|z$ -score| >1.96), and the optimal shrinkages are 0.80 (B6) and 0.72 (D2).

In one case, 463 partial correlations (edges) are stronger in B6 than in D2, and relate to 193 genes. In the opposite case, 385 partial correlations are stronger in D2 than in B6, and relate to 184 genes. A recent B6-D2 comparison of the striatal proteome found 160 differentially expressed proteins, among which eight are well-known functional sequence variants at protein level (Parks et al., 2019). The edges in our differential network analysis include four genes, which encode four of these eight proteins, namely; ALAD, GLO1, GABRA2 and COX7A2L.

4 Discussion

GGMs employ partial correlations to model interactions in the form of a network. The reconstruction of GGMs fall into the ‘high dimensional’ scenario in large-scale applications, which is a common case with gene-expression data. Shrinkage estimators solve the issues that arise in high dimensions, however, the ‘shrunk’ partial correlations depend on their shrinkage value, and cannot be compared directly.

In this work, we adapted some classical statistical tests to the partial correlations obtained with the LW shrinkage. We showed how an appropriate and simple transformation turns the ‘shrunk’ partial correlation into a (classical) Pearson’s correlation coefficient. Leveraging this fact, we derived (i) confidence intervals, (ii) a test for the hypothesis of a zero partial correlation and (iii) a test for the difference between pairs of partial correlations. These tests retrieve effect sizes (the strength and direction of the relationship), and

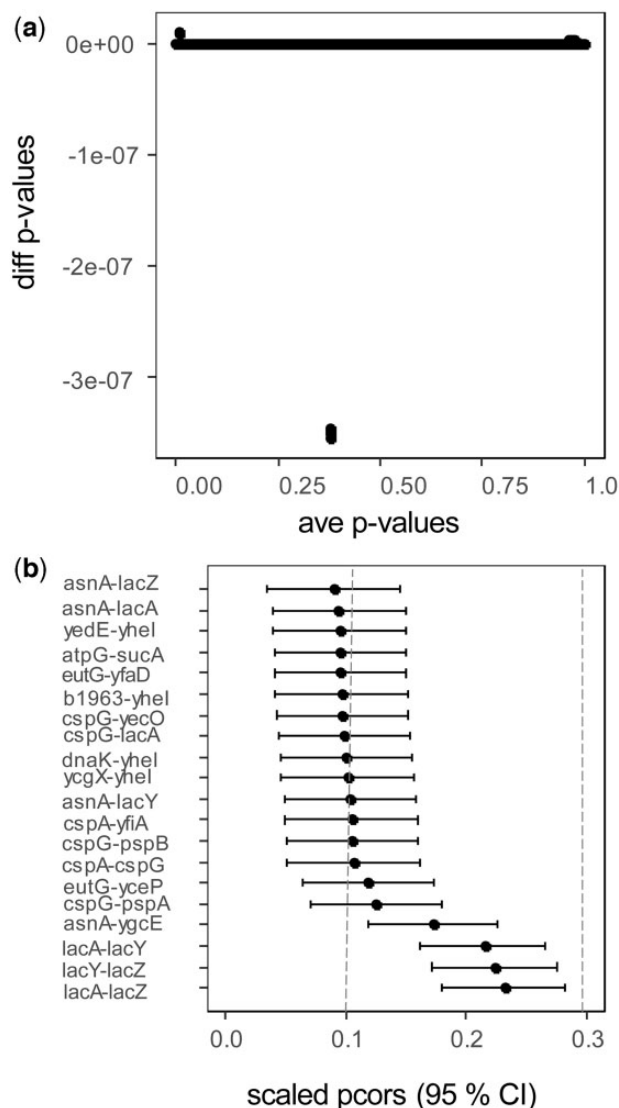


Fig. 2. *Escherichia coli* microarray network analysis. (a) Bland–Altman plot between the P -values obtained from the t -test [Equation (6)] and the ‘shrunk’ probability density [Equation (4.2)]. The methods are equivalent as the differences are in the order 10^{-7} . (b) Forest plot of partial correlations. The 15 strongest edges are displayed with their 95% confidence intervals. The vertical lines show the 0.1 and 0.3 thresholds for weak and mild correlations (Cohen, 1988)

P -values (the statistical significance), and account for the number of variables, the sample sizes and the shrinkage values.

Our simulations show that our test for differences between partial correlations (z -score) has an appropriate balance of false and true positives (Fig. 1a and b). Furthermore, its AUROC and AUPRC are higher than for the (recently published) state-of-the-art method DiffNetFDR (Fig. 1c and d). Our test of zero partial correlation (t -test) for the *E.coli* dataset is in agreement with previous results (Fig. 2a) (Bernal *et al.*, 2019), while being easier to implement and 10 times faster. Our differential network analysis (z -score) for the *M.musculus* dataset retrieves four well-known protein that might drive strain-specific regulation of gene expression (Parks *et al.*, 2019).

Finally, our methods are potentially useful to validate earlier network analyses, to compare gene regulatory networks from different experiments (differential networks analysis), and to further design of multi-omics/layer partial correlation methods. As most software includes efficient functions for a t -test, the arctanh and the quantiles of a normal density; our formulae are straightforward to implement, computationally fast even for large-scale applications and accessible to a broad audience with a basic background in Statistics.

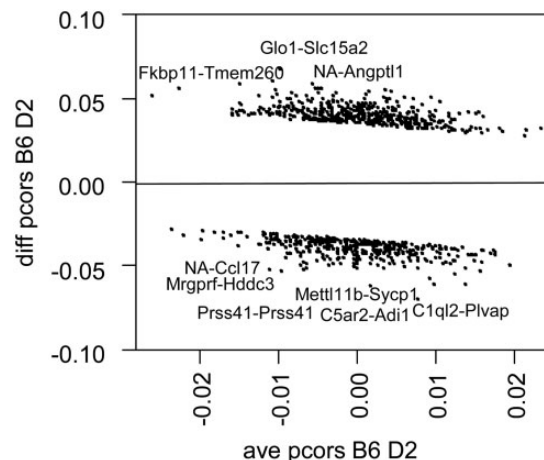


Fig. 3. *Mus musculus* RNA-seq differential network analysis. Bland–Altman plot between the partial correlations for strains B6 and D2. The figure shows only the significantly different partial correlations at the 0.05 level (i.e. $|z$ -score $>$ 1.96), and the names of the nine strongest gene pairs

Acknowledgements

We acknowledge the Centre of Data Science and System Complexity of the University of Groningen.

Funding

This work was supported by the Center of Information Technology of the University of Groningen. This research was part of the Netherlands X-omics Initiative and partially funded by NWO, project 184.034.019.

Conflict of Interest: none declared.

References

- Barabási, A.-L. *et al.* (2011) An integrative systems medicine approach to mapping human metabolic diseases. *Nat. Rev. Genet.*, **12**, 56–68.
- Beerenwinkel, N. *et al.* (2007) Genetic progression and the waiting time to cancer. *PLoS Comput. Biol.*, **3**, e225.
- Benedetti, E. *et al.* (2017) Network inference from glycoproteomics data reveals new reactions in the IgG glycosylation pathway. *Nat. Commun.*, **8**, 1–15.
- Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.*, **57**, 289–300.
- Bernal, V. *et al.* (2019) Exact hypothesis testing for shrinkage-based Gaussian graphical models. *Bioinformatics*, **35**, 5011–5017.
- Bottomly, D. *et al.* (2011) Evaluating gene expression in C57BL/6J and DBA/2J mouse striatum using RNA-Seq and microarrays. *PLoS One*, **6**, e17820.
- Class, C.A. *et al.* (2018) iDINGO—integrative differential network analysis in genomics with shiny application. *Bioinformatics*, **34**, 1243–1245.
- Cohen, J. (1988) *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Assoc., Hillsdale, NJ.
- Das, A. *et al.* (2017) Interpretation of the precision matrix and its application in estimating sparse brain connectivity during sleep spindles from human electrocorticography recordings. *Neural Comput.*, **29**, 603–642.
- Edwards, D. (2000) *Introduction to Graphical Modelling*. 2nd edn. Springer Science & Business Media, New York.
- Fisher, R.A. (1924) The distribution of the partial correlation coefficient. *Metron*, **3**, 329–332.
- Friedman, J. *et al.* (2008) Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, **9**, 432–441.
- Imkamp, K. *et al.* (2019) Gene network approach reveals co-expression patterns in nasal and bronchial epithelium. *Sci. Rep.*, **9**, 1–13.
- Keller, M.P. *et al.* (2008) A gene expression network model of type 2 diabetes links cell cycle regulation in islets with diabetes susceptibility. *Genome Res.*, **18**, 706–716.

- Ledoit, O. and Wolf, M. (2003) Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *J. Empir. Financ.*, **10**, 603–621.
- Ledoit, O. and Wolf, M. (2004) A well-conditioned estimator for large-dimensional covariance matrices. *J. Multivar. Anal.*, **88**, 365–411.
- Levy, K.J. and Narula, S.C. (1978) Testing hypotheses concerning partial correlations: some methods and discussion. *Int. Stat. Rev.*, **46**, 215–218.
- Liu, W. (2017) Structural similarity and difference testing on multiple sparse Gaussian graphical models. *Ann. Stat.*, **45**, 2680–2707.
- McNally, R.J. et al. (2015) Mental disorders as causal systems: a network approach to posttraumatic stress disorder. *Clin. Psychol. Sci.*, **3**, 836–849.
- Parks, C. et al. (2019) Comparison and functional genetic analysis of striatal protein expression among diverse inbred mouse strains. *Front. Mol. Neurosci.*, **12**, 128.
- Ritchie, M.E. et al. (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.*, **43**, e47.
- Schäfer, J. and Strimmer, K. (2005) A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Stat. Appl. Genet. Mol. Biol.*, **4**, 1175–1189.
- Schmidt-Heck, W. et al. (2004) Reverse engineering of the stress response during expression of a recombinant protein. In: *Proceedings of the EUNITE 2004 European Symposium on Intelligent Technologies, Hybrid Systems and their Implementation on Smart Adaptive Systems, June 10-12, 2004, Aachen, Germany, Verlag Mainz, Wissenschaftsverlag, Aachen*. pp. 407–441.
- Yuan, H. et al. (2017) Differential network analysis via lasso penalized D-trace loss. *Biometrika*, **104**, 755–770.
- Zhang, X.-F. et al. (2018) DiffGraph: an R package for identifying gene network rewiring using differential graphical models. *Bioinformatics*, **34**, 1571–1573.
- Zhang, X.-F. et al. (2019) DiffNetFDR: differential network analysis with false discovery rate control. *Bioinformatics*, **35**, 3184–3186.