Supplementary Materials: Comprehensive molecular characterisation of HPV-induced transformation by longitudinal statistical modelling

The work in this thesis was funded by a grant from the VU University Medical Center-Cancer Center Amsterdam (VUMC-CCA, project CCA2011-5-02) and carried out under the auspices of the VU University Amsterdam.

The printing of this thesis was kindly supported by the VU University Medical Center Amsterdam.

vrije Universiteit amsterdam

ISBN: 978-90-9028117-9

Document prepared with LaTeX 2$_\varepsilon$ and typeset by pdfTeX (font)
Cover design by Livija Blanozan
Printed by: Fine Graf, Belgrade, Serbia (www.finegraf.rs/)

VRIJE UNIVERSITEIT

Comprehensive molecular characterization of HPV-induced transformation by longitudinal statistical modelling

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor
aan de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. V. Subramaniam,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Geneeskunde
op maandag 10 september 2018 om 13.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Viktorian Miok

geboren te Zrenjanin, Servië

promotoren:	prof.dr.ir. M.A. van de Wiel
	dr. R.D.M. Steenbergen
copromotoren:	dr. W.N. van Wieringen
	dr. S.M. Wilting

Contents

Chapter 1

Introduction

1.1 No Supplementary Materials available

Chapter 2

tigaR: integrative significance analysis of temporal differential gene expression induced by genomic abnormalities

(***Miok, V., Wilting, S. M., van de Wiel, M. A., Jaspers, A., van Noort, P. I., Brakenhoff, R. H., Snijders, P. J. F., Steenbergen, R. D. M. and van Wieringen, W. N., BMC Bioinformatics (2014), 15: 327.***)

2.1 Effect of shrinkage

The proposed empirical Bayes [5] procedure shrinks the dispersion-related parameters to constrain flexibility of the spline(s) and protect against overfitting. Shrinkage on the random effects $\gamma_{k,j}$, which control the smoothness of the fit, is achieved via the choice of the hyperprior. Here we investigate the effect of this choice. Two different hyperpriors are considered: the Gamma distribution and a mixture distribution of the Gamma and a point mass at zero. The point mass at zero in the latter hyperprior is expected to yield more shrinkage. We compare the fit (of the splines) with both hyperpriors. Figure 2.1 illustrates the difference between the fits (using the two hyperpriors). The difference between the fits is not dramatic, but it shows (as expected) that the mixture of Gamma and point mass at zero yields splines with

decreased smoothness.

The decreased smoothness as an effect of including point mass at zero, which enforce stronger penalization of non-zero random effects, is considered in the sensitivity and specificity analysis. Figure 2.2 shows the sensitivity and specificity compared among the methods tigaR, EDGE, timecourse, BATS and tigaR (now employing a 'Gamma plus point mass at zero' hyperprior). From Figure 2.2 we conclude that less smooth splines does not substantially affect the sensitivity and specificity of the method. Thus, the Gamma hyperprior guards (sufficiently) against over-smoothing and potential over-fitting.

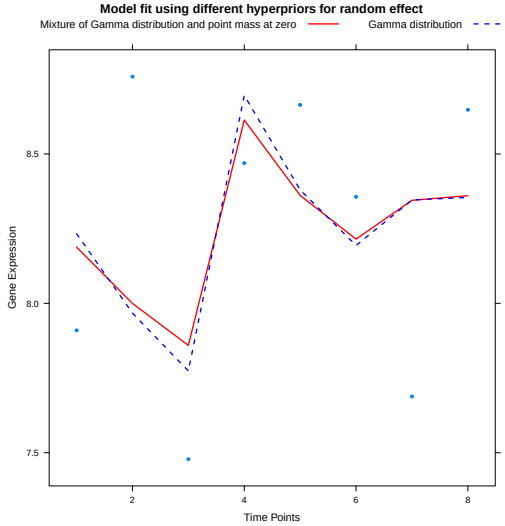

Fig. 2.1 The effect of using different hyperpriors in a single cell line. Gene expression is plotted against time (single cell line only). The solid red line is the fit of the model with a standard hyperprior, while the dashed blue line is that of the model with the alternative mixture hyperprior.

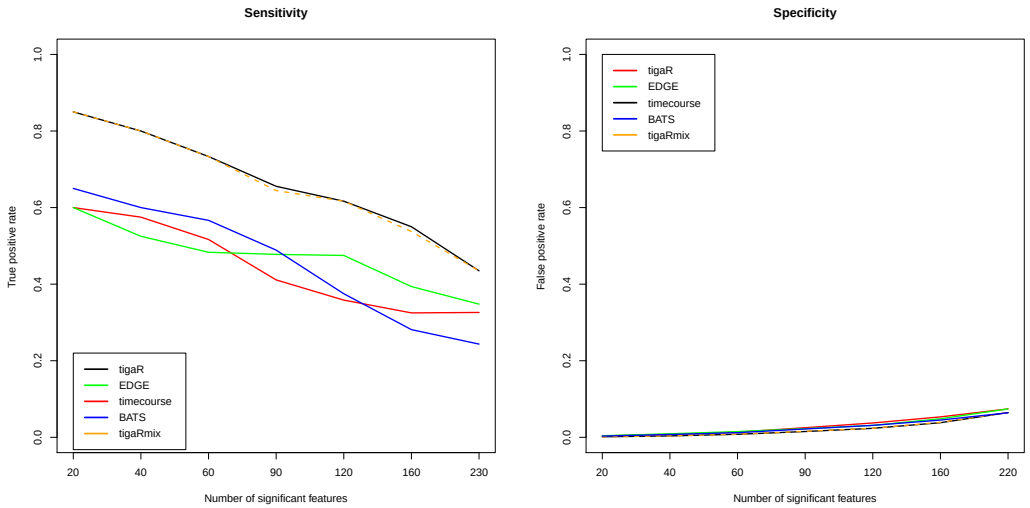

Fig. 2.2 Plots represent comparison of packages tiagR, EDGE, BATS, timecourse and tigaR which employ hyperprior composed of Gamma and point mass at zero for the sensitivity and specificity on HPV-induced transformation data. The left panel represents the sensitivity (of each model), where the true positive rate is plotted against the number of significant genes. The right panel illustrates the specificity (of each method), where the false positive rate is plotted against the number of significant genes.

2.2 Effect of a gene specific covariate on shrinkage

DNA copy number data varies from one chromosome to another, but within a single chromosome these data exhibit genomic regions that share the same aberration patterns over the samples. We assume a common hyperprior for the gene dosage effect within and between such regions. This is justified within a region, but between regions it may be more appropriate to assume a different dirac-Gaussian hyperprior. By simulation we show that the effect of assuming a common hyperprior for all regions is not unreasonable, in the sense that it does not yield results that differ much from those obtained when employing a separate hyperprior for each region.

The simulation involves the data set described in Section 2.1. For 15 regions (comprising features with a common DNA copy number pattern over the samples) we estimate parameters of the gene dosage effect prior distribution with both separate Dirac-Gaussian hyperpriors and a common one, employing the procedure described in Section 1.2. First, we estimate the hyperparameters of the DNA copy number effect using a different Dirac-Gaussian hyperprior for each region. Then, we also estimate hyperparameters of the gene dosage effect using a dirac-Gaussian hyperprior common to all regions. Figure 2.3 shows the histograms of the sampling distribution from the

averaged (over the regions) prior and the common prior of the gene dosage effect. At first glance the histograms are similar. A better comparison is provided by the Quantile-Quantile plot (right panel of Figure 2.3). The Q-Q plot reveals that these prior distributions (with separate and common Dirac-Gaussian hyperpriors) are alike around their modes, but differ somewhat in their tails. This difference in the tails will decrease as the number of regions increases. In most data sets the number of regions easily exceeds a hundred yielding a much better approximation in the tails. On the other hand, a mixture of many Gaussians with common mean but different variances may be approximated by a t-distribution with the degrees of freedom equal to the number of Gaussians in the mixture. Furthermore, as the degrees of freedom of a t-distribution increase it approaches a Gaussian. Hence, a better approximation in the tails is expected with more regions.

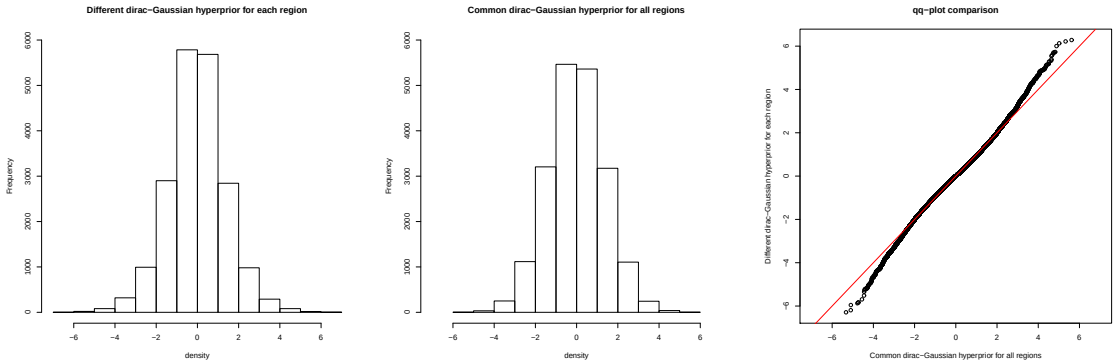

Fig. 2.3 Comparison of common and separate Dirac-Gaussian hyperpriors. The left panel depicts the prior distribution of the gene dosage effect using a common Dirac-Gaussian hyperpriors for all regions. The middle panel represents the prior distribution for different Dirac-Gaussian hyperpriors per region. The right panel compares these distributions using a Q-Q plot.

2.3 Degree of freedom estimation for penalized splines

Within the context of generalized linear mixed models the effective degrees of freedom are (usually) defined by the trace of the hat matrix. The fit of the mixed Model (2) per gene j is obtained by $\widehat{\mathbf{Y}}_j = \mathbf{H}_{\lambda,j}\mathbf{Y}_j$, where $\mathbf{H}_{\lambda,j}$ is the hat matrix, which projects the data onto the space spanned by the design matrix. The matrix $\mathbf{Y}_{*,j,*}$ is transformed into a vector by the vec-operator: $\mathbf{Y}_j = \text{vec}(\mathbf{Y}_{*,j,*})$, where the measurements of gene j are first ordered by cell lines i and within a cell line by time. The hat matrix for Model (2) is defined using smoothing parameter $\lambda = \sigma^2_{\gamma,j}/\sigma^2_{\varepsilon,j}$, which controls the smoothness of the fit. Let $\mathbf{C}_j$ be the joint design matrix comprising covariate information for both the fixed and random effects. The hat matrix is then:

$$\mathbf{H}_{\lambda,j} = \mathbf{C}_j \left(\mathbf{C}_j^{\mathrm{T}}\mathbf{C}_j + \lambda\mathbf{I}\right)^{-1} \mathbf{C}_j^{\mathrm{T}}.$$

We exemplify the use of the matrix $\mathbf{C}_j$. Consider the data set from Section 2.1 using a model with a common spline. Then, $\mathbf{C}_j = [\mathbf{W}_{i,t}|\widetilde{\mathbf{Z}}_t]$, where $\mathbf{W}_{i,t}$ is a matrix with cell line and DNA copy number covariate information, while $\widetilde{\mathbf{Z}}_t$ contains that for the random (time) effect. Matrix C_j has dimensions 32×7, where 32 corresponds to the number of observations on gene j, while 7 represents the number of covariates: four cell line effects, one gene dosage effect, and two for the spline (a common spline with two knots). The rows of $\mathbf{C}_j$ are ordered as those of $\mathbf{Y}_j$.

Generalizing the definition of degrees of freedom, it is defined as the trace of the hat matrix: $\text{df} = \text{tr}(\mathbf{H}_{\lambda,j})$. This is thus the degree of freedom of mixed Model (2), depending on the smoothing parameter λ. It can be interpreted as the number of fitted parameters. Thus, df indicates the 'amount of fitting' done by $\mathbf{H}_{\lambda,j}$, which is strictly monotone in λ.

2.4 Parameter estimation using the spatial prior

In this section we present how parameters of Model (2) are re-estimated when imposing a spatial prior on the gene dosage effect. First, all hyperparameters of Model (2) are estimated univariate as described in Section 1.2. Then, hyperparameter ρ (of the spatial prior) is estimated by regressing β_j on β_{j-1} (justified by the assumption that β_j follows a first-order autoregressive process). Combining the univariate estimated hyperparameters $\sigma^2_{j-1}, \sigma^2_j, \sigma^2_{j+1}$ with the estimated spatial hyperprior parameter ρ, yields the covariance matrix of the spatial hyperprior of a contiguous triplet. With this estimated spatial prior and the other (univariate) priors at hand, model parameters are re-estimated trivariately (per triplet of contiguous genes). Due to the fact that INLA does not allow to estimate a trivariate model with an AR(1) covariance structure on a fixed effect, estimators of model parameters are derived analytically.

Analytical estimators are derived separately for fixed and random effects. For convenience of notation we write $\mathbf{W}_{i,j}$, the joint design matrix of the cell line and DNA copy number information, and $\boldsymbol{\theta}_j = (\boldsymbol{\alpha}_{i,j}, \beta_j)$, the vector of corresponding parameters. Further, the prior distributions for fixed and random effects can be written as $\boldsymbol{\theta}_j|\sigma^2_{\theta,j} \sim \mathcal{N}(\mathbf{0}, \sigma^2_{\theta,j}\boldsymbol{\Sigma})$ and$\boldsymbol{\gamma}_j|\sigma^2_{\gamma,j} \sim \mathcal{N}(\mathbf{0}, \sigma^2_{\gamma,j}\boldsymbol{\Psi})$, for appropriate covariance matrices $\boldsymbol{\Sigma}$ and $\boldsymbol{\Psi}$. The part of covariance matrix $\boldsymbol{\Sigma}$ corresponding to parameters β_j incorporates the AR(1) correlation structure.

From [**?**] we obtain the estimator of the fixed effect in a trivariate linear mixed model, given the hyperparameters:

$$\widehat{\boldsymbol{\theta}}_j = \left(\mathbf{W}^{\mathrm{T}}_{i,j}\mathbf{V}^{-1}_j\mathbf{W}_{i,j} + \boldsymbol{\Sigma}^{-1}\frac{\sigma^2_{\varepsilon,j}}{\sigma^2_{\theta,j}}\right)^{-1}\mathbf{W}^{\mathrm{T}}_{i,j}\mathbf{V}^{-1}_j\mathbf{Y}_{*,j,*},$$

where

$$\mathbf{V}^{-1}_j = \left(\widetilde{\mathbf{Z}}_t\boldsymbol{\Psi}\widetilde{\mathbf{Z}}^{\mathrm{T}}_t\frac{\sigma^2_{\varepsilon,j}}{\sigma^2_{\gamma,j}} + \mathbf{I}\right)^{-1}.$$

Using $\widehat{\boldsymbol{\theta}}$ and previously estimated hyperparameters, the random effect estimator from a trivariate linear mixed model is:

$$\widehat{\boldsymbol{\gamma}}_j(\widehat{\boldsymbol{\theta}}_j) = \left(\widetilde{\mathbf{Z}}^{\mathrm{T}}_t\widetilde{\mathbf{Z}}_t + \boldsymbol{\Psi}^{-1}\frac{\sigma^2_{\varepsilon,j}}{\sigma^2_{\gamma,j}}\right)^{-1}\widetilde{\mathbf{Z}}^{\mathrm{T}}_t\left(\mathbf{Y}_{*,j,*} - \mathbf{W}_{i,j}\widehat{\boldsymbol{\theta}}_j\right).$$

Full derivation of these estimators can be found in Chapter 6 of [**?**].

With these analytic estimators we obtain the trivariate fit of the model. From

this trivariate fit, only the parameter estimate of the middle gene in the triplet is preserved and considered its final parameter estimate. Sliding the triplet window over the genome yields final parameter estimates for each gene.

2.5 Optimal number of knots for penalized splines

DIC is a measure that combines model fit with model complexity (the latter is represented by the number of effective parameters) ([162]). When denoting by $\boldsymbol{\theta}_j$ the vector containing all parameters of Model (2), then the goodness-of-fit (for gene j) may measured by the deviance $D(\boldsymbol{\theta}_j) = -2\log[L(Y_{*,j,*}|\boldsymbol{\theta}_j)]$. Write $D(\bar{\boldsymbol{\theta}}_j) = D(E_{\boldsymbol{\theta}_j|Y_{*,j,*}}[\boldsymbol{\theta}_j])$ for the deviance of the posterior mean, where $\bar{\boldsymbol{\theta}}_j$ is the expectation of $\boldsymbol{\theta}_j$ and $\bar{D}(\boldsymbol{\theta}_j) = E_{\boldsymbol{\theta}_j|Y_{*,j,*}}[D(\boldsymbol{\theta}_j)]$ for the posterior mean of the deviance. The effective number of parameters $p_{D,j}$ is then calculated per gene as:

$$p_{D,j} \quad = \quad \bar{D}(\boldsymbol{\theta}_j) - D(\bar{\boldsymbol{\theta}}_j),$$

The effective number of parameters $p_{D,j}$ may also be defined directly through the effective degrees of freedom. For the normal hierarchical model, $p_{D,j} = \text{tr}(\mathbf{H}_{\lambda,j})$, with $\mathbf{H}_{\lambda,j}$ the hat matrix ([162]). The DIC is now defined as:

$$DIC_j \quad = \quad D(\bar{\boldsymbol{\theta}}_j) + 2p_{D,j} \quad = \quad \bar{D} + p_{Dj}.$$

To decide on the optimal number of knots, we determine the DIC per gene for different numbers of knots $k = 2, \ldots, 5$. The value of k with the smallest DIC represents the model which is the best supported by the data (for that particular gene). The optimal number of knots for the whole data set (i.e. all genes) is that k which is the most frequent optimal k over the genes. For the data set from Section 2.1, Figure 2.4 indicates that $k = 2$ is the optimal number of knots.

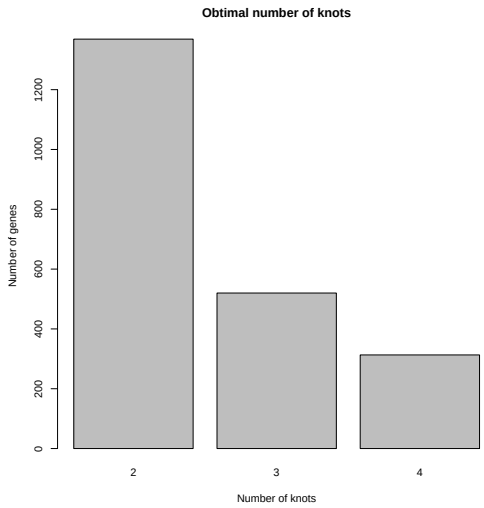

Fig. 2.4 The bar plot illustrates the frequency of genes where the number of knots k is identified with the smallest DIC.

2.6 Comparison using the data set from the EDGE package

The proposed method for temporal differential expression was developed with the data set from Section 2.1 in mind. This may accidentally favor our method in the comparison with its competitors EDGE, timecourse, BATS and the reference (univariate fitted) method. Therefore the comparison is repeated using the data set from EDGE-package ([171]), thus possibly favoring the EDGE. Again sensitivity, specificity and reproducibility are compared among packages EDGE, timecourse, tigaR, BATS and the reference method. Figure 2.5 shows the number of true and false positive rates of the four methods. Qualitatively the same picture emerges as from the comparison using the data from Section 2.1. Similarly, when comparing the reproducibility of the methods (confer Figure 2.6), again tigaR and BATS outperform timecourse and EDGE.

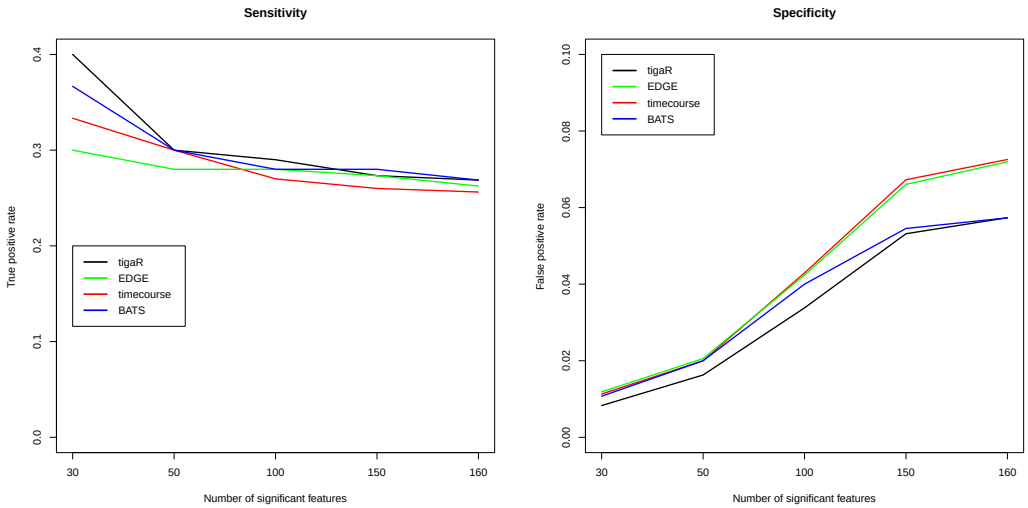

Fig. 2.5 Plots represent comparison of tigaR, EDGE, timecourse and BATS package for the sensitivity and specificity on the data set from the EDGE-package. The left panel represents the sensitivity (of each model), where the true positive rate is plotted against the number of significant genes. The right panel illustrates the specificity (of each method), where the false positive rate is plotted against the number of significant genes.

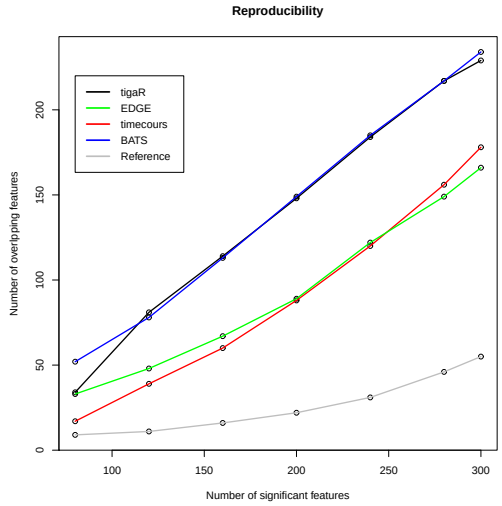

Fig. 2.6 The plot illustrates the comparison of the reproducibility between tigaR, EDGE, timecourse, BATS and reference model. The number of significant genes is plotted against number of overlapping genes identified between two groups.

2.7 Additional plots

In the main manuscript Figure 2, Figure 6 and Figure 7 show only a single cell line. This is done for reasons of clarity: to make a certain effect more visible. Here, in Figure 2.7, Figure 2.8 and Figure 2.9 these effects are shown in all four cell lines.

Additionally, a Venn diagram (Figure 2.10) is presented. It illustrates the various intersections of the sets of features exhibiting temporal differential expression as declared by tigaR, EDGE, timecourse and BATS. The timecourse packages only provides a rank-ordered list of genes. For the comparison the top genes from this list are selected, their number being equal to the number of features selected by tigaR (with a common spline).

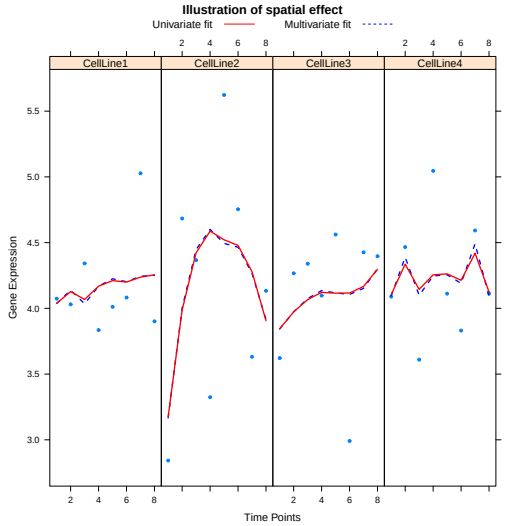

Fig. 2.7 Illustration of the effect of the spatial prior of one gene in four cell lines. Gene expression is plotted against time (in four cell lines). The lines represent the univariate (red, solid line) and multivariate fit (blue, dashed line).

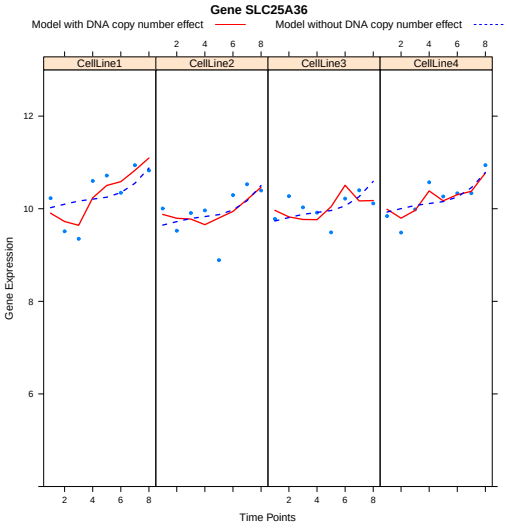

Fig. 2.8 Dots represent expression levels of gene SLC25A36 plotted against time in four cell lines. The solid (red) line represents the full model while the dashed (blue) line illustrates the fit of the model without copy number effect.

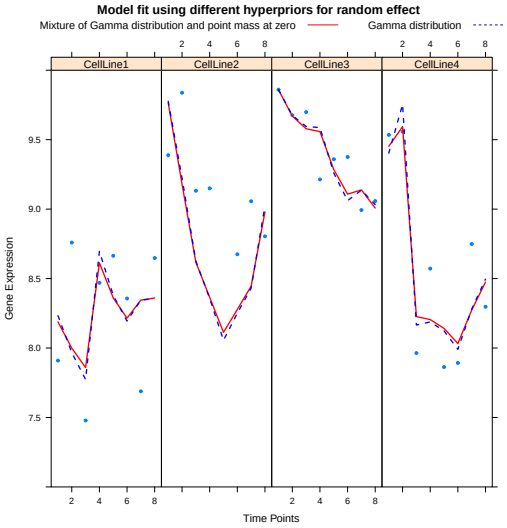

Fig. 2.9 Effect of using different priors in four cell lines. Gene expression is plotted against time (in four cell lines). The solid red line is the fit of the model with a standard prior, while the dashed blue line is that of the model with an alternative prior.

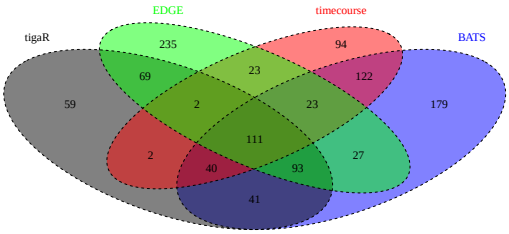

Fig. 2.10 Venn diagram of methods tigaR, EDGE, timecourse and BATS.

Chapter 3

Ridge estimation of the VAR(1) model and its time series chain graph from multivariate time-course omics data

(*Miok, V., Wilting, S. M. and van Wieringen, W. N., Biometrical Journal (2017), 59(1), 172-191*)

3.1 Pseudo-code of ridge ML estimation

Box 1: Ridge ML estimation of the parameters of the VAR(1) model.

Initiate

Obtain an initial estimate $\widehat{\mathbf{A}}^{(0)}(\lambda_a)$ of $\mathbf{A}$ from the ridge OLS estimator.

Iterate

For $k = 1, \ldots, k_{\max}$ iterate the following steps:

1) Use the latest estimate of $\mathbf{A}$, $\widehat{\mathbf{A}}^{(k-1)}(\lambda_a)$, to calculate the sample error covariance matrix $\mathbf{S}_\varepsilon^{(k)}$. Given $\mathbf{S}_\varepsilon^{(k)}$, the ridge ML precision estimator yields the updated estimate of $\boldsymbol{\Omega}_\varepsilon$, $\widehat{\boldsymbol{\Omega}}_\varepsilon^{(k)}(\lambda_\omega)$.
2) Get $\widehat{\mathbf{A}}^{(k)}(\lambda_a)$ from the ridge ML estimator of $\mathbf{A}$ with $\widehat{\boldsymbol{\Omega}}_\varepsilon^{(k)}(\lambda_\omega)$ substituted for $\boldsymbol{\Omega}_\varepsilon$.

Terminate

Iterate the previous step, until convergence.

3.2 Comparison of computation time

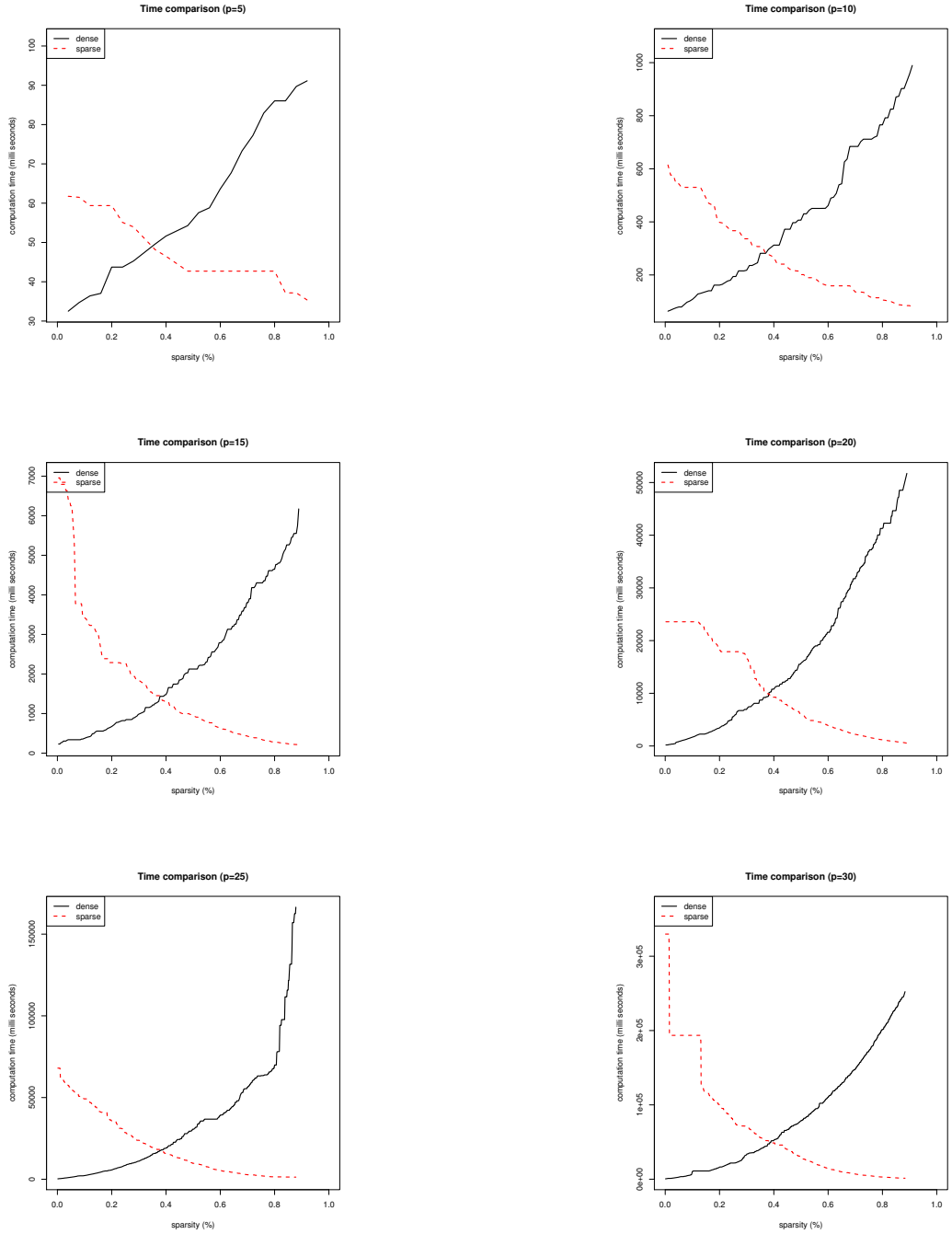

Fig. 3.1 The computational costs of evaluating the estimator of $\mathbf{A}$ with known support for two plotted against the sparsity percentage of the support. Two different implementations are evaluated: for a dense (few zeros; black, solid line) and sparse (many zeros; red, dashed line) support. Computation times were evaluated using the `microbenchmark`-package ([108]).

3.3 Decomposable graph illustration

Figure 3.2 shows an example of a decomposable graph and its decomposition. Clearly, every path between the two cliques runs via the separator. Furthermore, both cliques and separator are complete subgraphs. Hence, they form a decomposition of the graph. Figure 3.3 containing two additional examples of decomposable graph: the tree and the chain graph.

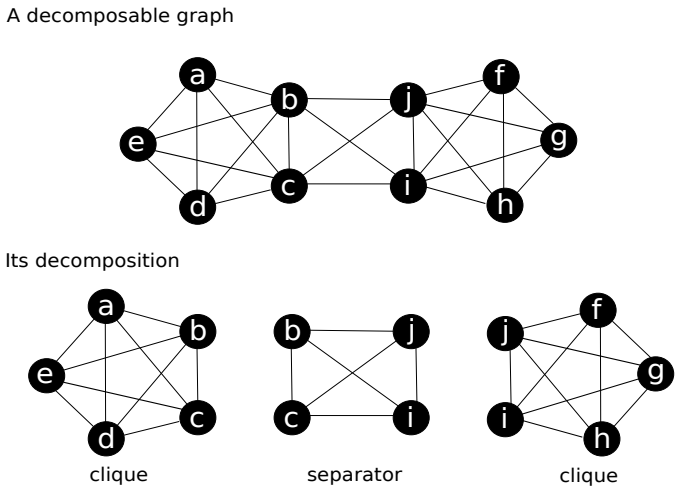

Fig. 3.2 The top row shows a decomposable graph. The bottom row displays its decomposition in terms of the two cliques, comprising vertices $\{a, b, c, d, e\}$ and $\{f, g, h, i, j\}$ and the edges connecting them, and a separator with vertices $\{b, c, i, j\}$ and the edges connecting them.

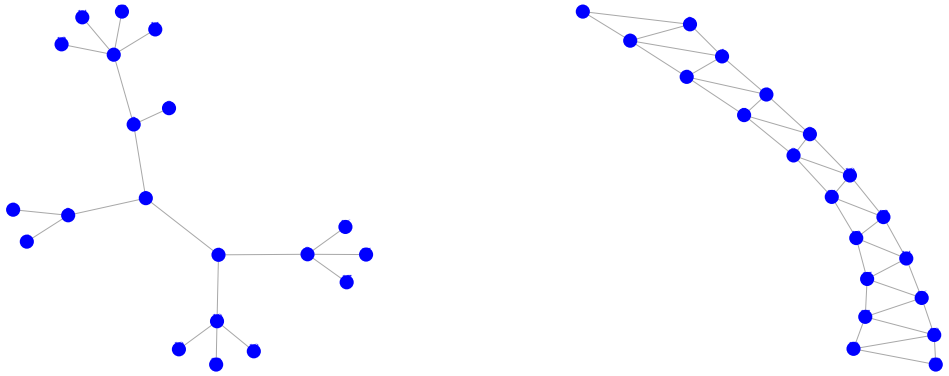

Fig. 3.3 Two examples of decomposable graphs: the tree and a chain graph.

3.4 Initial estimate of $\boldsymbol{\Omega}$ with chordal support

Under the assumption of a decomposable conditional independence graph $\mathcal{G}$, the ridge ML estimator of $\boldsymbol{\Omega}_\varepsilon$ can be expressed in limiting cases ($\lambda_\omega = 0$, $\lambda_\omega = \infty$ and $\mathcal{S} = \emptyset$) as a combination of the ridge ML estimators of the cliques and separators that decompose $\mathcal{G}$. Fully analogous to unpenalized case specified by Proposition 5.9 of [91] but with the unpenalized marginal ML covariance estimates of cliques and separators replaced by their penalized counterparts.

To derive the ridge ML estimator for the aforementioned limiting cases, note by direct observation that the decomposability of the support of $\boldsymbol{\Omega}_\varepsilon$ implies:

$$\begin{aligned} \operatorname{tr}[(\boldsymbol{\Omega}_\varepsilon - \boldsymbol{\Omega}_0)^\top(\boldsymbol{\Omega}_\varepsilon - \boldsymbol{\Omega}_0)] &= \sum_{C\in\mathcal{C}} \operatorname{tr}\{[(\boldsymbol{\Omega}_\varepsilon)_{C,C} - (\boldsymbol{\Omega}_0)_{C,C}]^\top[(\boldsymbol{\Omega}_\varepsilon)_{C,C} - (\boldsymbol{\Omega}_0)_{C,C}]\} \\ &\quad - \sum_{S\in\mathcal{S}} \operatorname{tr}\{[(\boldsymbol{\Omega}_\varepsilon)_{S,S} - (\boldsymbol{\Omega}_0)_{S,S}]^\top[(\boldsymbol{\Omega}_\varepsilon)_{S,S} - (\boldsymbol{\Omega}_0)_{S,S}]\}, \end{aligned} \tag{3.1}$$

in which $\boldsymbol{\Omega}_0$ is assumed to share its support with $\boldsymbol{\Omega}_\varepsilon$. For square, symmetric $\mathbf{S}_\varepsilon \succeq$, $\boldsymbol{\Omega}_\varepsilon \succ 0$, and $\boldsymbol{\Omega}_0 \succ 0$ of arbitrary but equal dimensions define

$$\mathcal{L}_\omega^{\text{pen}}(\mathbf{S}_\varepsilon, \boldsymbol{\Omega}_\varepsilon, \boldsymbol{\Omega}_0, \lambda_\omega) = \frac{1}{2}n(\mathcal{T}-1)\Big\{\log|\boldsymbol{\Omega}_\varepsilon| - \operatorname{tr}(\mathbf{S}\boldsymbol{\Omega}_\varepsilon) - \lambda_\omega \operatorname{tr}[(\boldsymbol{\Omega}_\varepsilon - \boldsymbol{\Omega}_0)^\top(\boldsymbol{\Omega}_\varepsilon - \boldsymbol{\Omega}_0)]\Big\}.$$

Using the reformulated ridge penalty (3.1) together with iterative application of Lemma 5.5. of [91], which decomposes the unpenalized log-likelihood in a linear combination of those of the cliques and separators, the loss function for $\boldsymbol{\Omega}_\varepsilon$ can then

be rewritten to:

$$\sum_{\mathcal{C}\in\mathcal{C}} \mathcal{L}_\omega^{\text{pen}}[(\mathbf{S}_\varepsilon)_{\mathcal{C},\mathcal{C}}, (\boldsymbol{\Omega}_\varepsilon)_{\mathcal{C},\mathcal{C}}, (\boldsymbol{\Omega}_0)_{\mathcal{C},\mathcal{C}}, \lambda_\omega] \\ - \sum_{\mathcal{S}\in\mathcal{S}} \nu(\mathcal{S})\mathcal{L}_\omega^{\text{pen}}[(\mathbf{S}_\varepsilon)_{\mathcal{S},\mathcal{S}}, (\boldsymbol{\Omega}_\varepsilon)_{\mathcal{S},\mathcal{S}}, (\boldsymbol{\Omega}_0)_{\mathcal{S},\mathcal{S}}, \lambda_\omega], \tag{3.2}$$

where $\nu(\mathcal{S})$ is the number of times separator $\mathcal{S}$ forms the intersection of two neighboring cliques.

Recall from the main text that, when $\mathcal{G}$ decomposes into cliques $\mathcal{C}_1$, $\mathcal{C}_2$ and $\mathcal{S}$, the proposed ridge ML estimator is:

$$\widehat{\boldsymbol{\Omega}}_\varepsilon(\lambda_\omega) = \begin{pmatrix} [\widehat{\boldsymbol{\Omega}}_\varepsilon^{(\mathcal{C}_1)}(\lambda_\omega)]_{\mathcal{C}_1\setminus\mathcal{S},\mathcal{C}_1\setminus\mathcal{S}} & [\widehat{\boldsymbol{\Omega}}_\varepsilon^{(\mathcal{C}_1)}(\lambda_\omega)]_{\mathcal{C}_1\setminus\mathcal{S},\mathcal{S}} & \mathbf{0}_{|\mathcal{C}_1\setminus\mathcal{S}|\times|\mathcal{C}_2\setminus\mathcal{S}|} \\ [\widehat{\boldsymbol{\Omega}}_\varepsilon^{(\mathcal{C}_1)}(\lambda_\omega)]_{\mathcal{S},\mathcal{C}_1\setminus\mathcal{S}} & [\widehat{\boldsymbol{\Omega}}_\varepsilon^{(\mathcal{C}_1)}(\lambda_\omega)]_{\mathcal{S},\mathcal{S}} + [\widehat{\boldsymbol{\Omega}}_\varepsilon^{(\mathcal{C}_2)}(\lambda_\omega)]_{\mathcal{S},\mathcal{S}} - \widehat{\boldsymbol{\Omega}}_\varepsilon^{(\mathcal{S})}(\lambda_\omega) & [\widehat{\boldsymbol{\Omega}}_\varepsilon^{(\mathcal{C}_2)}(\lambda_\omega)]_{\mathcal{S},\mathcal{C}_2\setminus\mathcal{S}} \\ \mathbf{0}_{|\mathcal{C}_2\setminus\mathcal{S}|\times|\mathcal{C}_1\setminus\mathcal{S}|} & [\widehat{\boldsymbol{\Omega}}_\varepsilon^{(\mathcal{C}_2)}(\lambda_\omega)]_{\mathcal{C}_2\setminus\mathcal{S},\mathcal{S}} & [\widehat{\boldsymbol{\Omega}}_\varepsilon^{(\mathcal{C}_2)}(\lambda_\omega)]_{\mathcal{C}_2\setminus\mathcal{S},\mathcal{C}_2\setminus\mathcal{S}} \end{pmatrix}$$

where $\widehat{\boldsymbol{\Omega}}_\varepsilon^{(\mathcal{C}_1)}(\lambda_\omega)$, $\widehat{\boldsymbol{\Omega}}_\varepsilon^{(\mathcal{C}_1)}(\lambda_\omega)$ and $\widehat{\boldsymbol{\Omega}}_\varepsilon^{(\mathcal{S})}(\lambda_\omega)$ are the marginal ridge ML covariance estimators for cliques $\mathcal{C}_1$, $\mathcal{C}_2$ and separator $\mathcal{S}$. Note that each of these marginal estimator satisfy the marginal estimating equation, e.g.:

$$[\widehat{\boldsymbol{\Omega}}_\varepsilon^{(\mathcal{C}_1)}(\lambda_\omega)]^{-1} - \lambda_\omega \widehat{\boldsymbol{\Omega}}_\varepsilon^{(\mathcal{C}_1)}(\lambda_\omega) = (\mathbf{S}_\varepsilon)_{\mathcal{C}_1,\mathcal{C}_1} - \lambda_\omega(\boldsymbol{\Omega}_0)_{\mathcal{C}_1,\mathcal{C}_1},$$

confer [189] for details.

To show that this estimator maximizes (in the three limiting cases) the ridge penalized log-likelihood under the assumption of a decomposable conditional independence graph, we study this penalized log-likelihood in the vicinity of the proposed estimator. Let a $\boldsymbol{\Delta}$ be a $p \times p$ dimensional matrix and γ a constant close to but unequal to zero. Consider a single summand of the penalized log-likelihood (3.2), say, that of the first clique. It may then, after noting that $\log(\mathbf{A} + \varepsilon\mathbf{B}) = \log(\mathbf{A}) + \varepsilon\mathbf{A}^{-1}\mathbf{B} + \mathcal{O}(\varepsilon^2)$ and $\log(|\mathbf{A}|) = \text{tr}[\log(\mathbf{A})]$, be approximated by:

$$\begin{aligned} &\mathcal{L}_\omega^{\text{pen}}\Big\{(\mathbf{S}_\varepsilon)_{\mathcal{C}_1,\mathcal{C}_1}, [\widehat{\boldsymbol{\Omega}}_\varepsilon(\lambda_\omega)]_{\mathcal{C}_1,\mathcal{C}_1} + \gamma\,(\boldsymbol{\Delta})_{\mathcal{C}_1,\mathcal{C}_1}, (\boldsymbol{\Omega}_0)_{\mathcal{C}_1,\mathcal{C}_1}, \lambda_\omega\Big\} \\ &\quad\approx \log\Big\{|[\widehat{\boldsymbol{\Omega}}_\varepsilon(\lambda_\omega)]_{\mathcal{C}_1,\mathcal{C}_1}|\Big\} - \text{tr}\Big\{(\mathbf{S}_\varepsilon)_{\mathcal{C}_1,\mathcal{C}_1}[\widehat{\boldsymbol{\Omega}}_\varepsilon(\lambda_\omega)]_{\mathcal{C}_1,\mathcal{C}_1}\Big\} \\ &\qquad -\frac{1}{2}\lambda_\omega \text{tr}\Big(\Big\{[\widehat{\boldsymbol{\Omega}}_\varepsilon(\lambda_\omega)]_{\mathcal{C}_1,\mathcal{C}_1} - (\boldsymbol{\Omega}_0)_{\mathcal{C}_1,\mathcal{C}_1}\Big\}^\top \Big\{[\widehat{\boldsymbol{\Omega}}_\varepsilon(\lambda_\omega)]_{\mathcal{C}_1,\mathcal{C}_1} - (\boldsymbol{\Omega}_0)_{\mathcal{C}_1,\mathcal{C}_1}\Big\}\Big) \\ &\qquad +\gamma\,\text{tr}\Big[\Big(\Big\{[\widehat{\boldsymbol{\Omega}}_\varepsilon(\lambda_\omega)]_{\mathcal{C}_1,\mathcal{C}_1}\Big\}^{-1} - \Big\{[\widehat{\boldsymbol{\Omega}}_\varepsilon(\lambda_\omega)]^{-1}\Big\}_{\mathcal{C}_1,\mathcal{C}_1}\Big)(\boldsymbol{\Delta})_{\mathcal{C}_1,\mathcal{C}_1}\Big)\Big] \\ &\qquad -\frac{1}{2}\gamma^2\lambda_\omega \text{tr}\Big\{[(\boldsymbol{\Delta})_{\mathcal{C}_1,\mathcal{C}_1}]^\top(\boldsymbol{\Delta})_{\mathcal{C}_1,\mathcal{C}_1}\Big\}, \end{aligned}$$

in which we also used the estimating equation.

Only the latter two summands of this approximation of the penalized log-likelihood of the clique involve γ. Now consider the three boundary cases separately:

- When $\lambda_\omega = 0$ the summands in the approximation stemming from the penalty vanish and the claim is warranted by Lemma 5.5 of [91].
- When $\mathcal{S} = \emptyset$, the precision matrix is block diagonal and the blocks can be estimated marginally. This implies $[(\widehat{\boldsymbol{\Omega}}_\varepsilon)_{C_1,C_1}]^{-1} = [(\widehat{\boldsymbol{\Omega}}_\varepsilon)^{-1}]_{C_1,C_1}$ and ensures that the first summand involving γ vanishes. As the second summand can only contribute negatively to the log-likelihood for any $\gamma \neq 0$, the proposed estimator maximizes the log-likelihood.
- When λ_ω tends to infinity, the second summand involving γ dominates the other. For these dominant summands:
$$\operatorname{tr}(\boldsymbol{\Delta}^\top \boldsymbol{\Delta}) \;=\; \sum_{C \in \mathcal{C}} \operatorname{tr}\{[(\boldsymbol{\Delta})_{C,C}]^\top (\boldsymbol{\Delta})_{C,C}\} \;-\; \sum_{S \in \mathcal{S}} \operatorname{tr}\{[(\boldsymbol{\Delta})_{S,S}]^\top (\boldsymbol{\Delta})_{S,S}\}.$$
The total contribution of these terms to the log-likelihood is thus negative for any choice of γ other than zero.

Hence, the proposed estimator indeed maximizes the penalized log-likelihood in the three limiting cases.

3.5 Path analysis illustration

The path analysis of autocovariance is illustrated by means of a toy example. Consider a VAR(1) model parametrized by:

$$\nu = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} 1/2 & 1/2 & 1/2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{\Sigma}_\varepsilon = \begin{pmatrix} 1 & 1/5 & 1/5 \\ 1/5 & 1 & 0 \\ 1/5 & 0 & 1 \end{pmatrix}^{-1}.$$

The top row of Figure 3.4 displays the corresponding time series chain graph.

The autocovariance between $Y_{1,t}$ and $Y_{1,t+1}$ (dropping the sample index for notational clarity) is: $\text{Cov}(Y_{1,t}, Y_{1,t+1}) = (\mathbf{\Sigma}_\varepsilon \mathbf{A}^\top)_{1,1} = 0.326087$. From the time series chain graph it is immediate that $Y_{1,t}$ and $Y_{1,t+1}$ are connected by three paths (visualized by the bottom row of Figure 3.4): *i)* $Y_{1,t} \to Y_{1,t+1}$, *ii)* $Y_{1,t} \to Y_{2,t} \to Y_{1,t+1}$, and *iii)* $Y_{1,t} \to Y_{3,t} \to Y_{1,t+1}$. The covariance between $Y_{1,t}$ and $Y_{1,t+1}$ can now be decomposed into the contributions of each of these paths. Using the decomposition of the autocovariance in terms of the paths' contributions from the main text one obtains: 0.5434783 (for $Y_{1,t} \to Y_{1,t+1}$), -0.1086957 (for $Y_{1,t} \to Y_{2,t} \to Y_{1,t+1}$) and -0.1086957 (for $Y_{1,t} \to Y_{3,t} \to Y_{1,t+1}$). These contributions indeed sum to 0.326087, which equals $\text{Cov}(Y_{1,t}, Y_{1,t+1})$.

3.6 Design of time-course studies in GEO

[51] observed that more than 80% of time-course gene expression microarray data sets have 8 or fewer number of time points. This reference being somewhat out of date, we conducted our own survey using the GEO data base. Using the search terms “time course”, “cancer”, “human”, “cell line”, “gene expression”, “microarray”, 85 studies where selected (at June 23, 2015). From the resulting list we excluded studies involving hematological tumors and data set GSE762 with a rather intricate experimental design (among other a time component). Finally, 63 studies satisfy these criteria, involving cell lines of various cancer types (breast, colon, prostate, ovarian, kidney and lung) and were mainly conducted using Affymetrix gene expression arrays (confer Table 3.1 Figure 3.5 shows histograms of the number of time-points $\mathcal{T}$ and individuals n (left panels) and the scatterplot of these study characteristics. From the provided table and histograms we conclude that a reasonable optimistic and simultaneously empirically maximum number of time-points and individuals are 10 and 5, respectively.

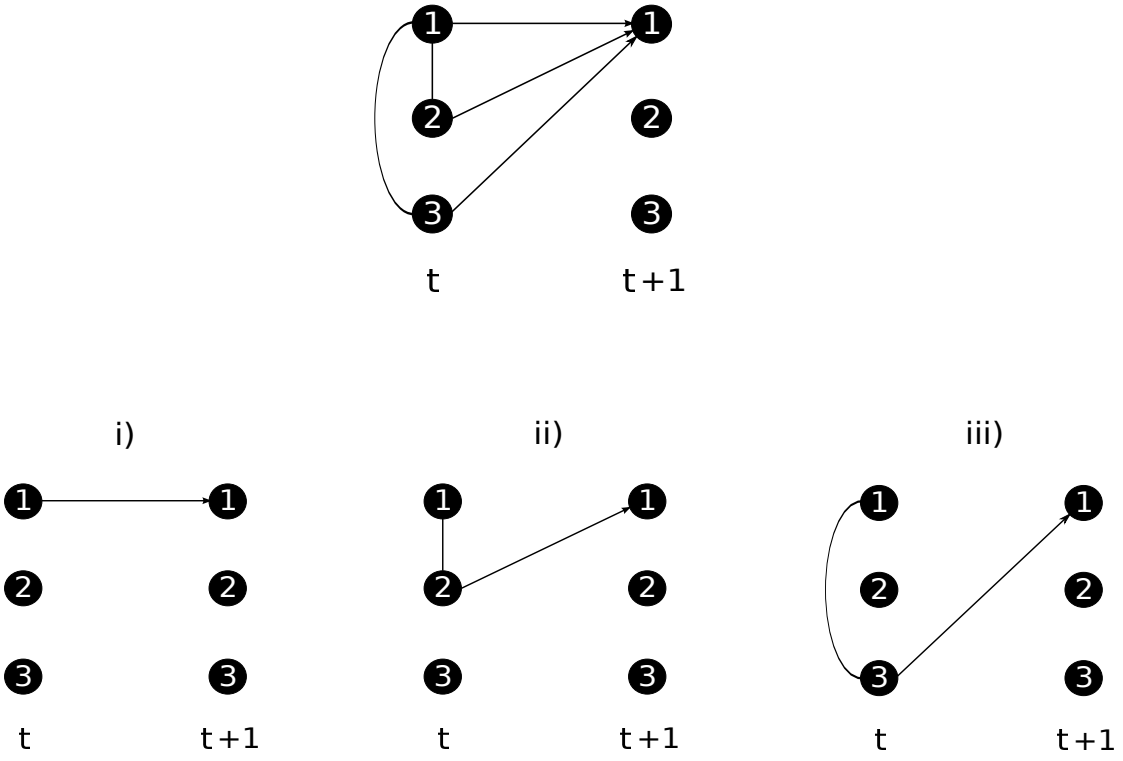

Fig. 3.4 The top row shows the time series chain graph of the toy model. The bottom row displays the three paths in the time series chain graph connecting the first variate between to consecutive time points.

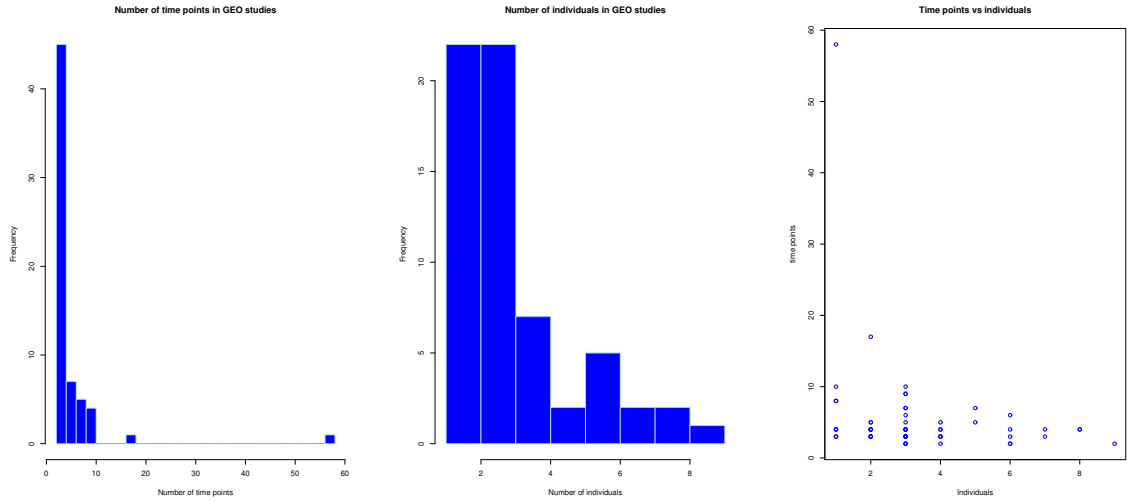

Fig. 3.5 Number of time-points and observations in microarray time-course gene expression studies from GEO database. Left panel presents the histogram of number of time-points, while the middle panel shows the histogram of the number of observations (cell lines) in GEO studies. The right panel plots the number of time points against the number of individuals.

data set ID	# time points	# individuals	# treatments	cancer type	organism	platform
GSE7868	3	3	1	prostate	human	Affymetrix
GDS2626	4	7	2	breast	human	Affymetrix
GSE29917	7	5	3	breast	human	Agilent
GSE13009	17	2	2	breast	human	Affymetrix
GDS4367	5	3	1	colon	mouse	Affymetrix
GDS4121	3	1	1	prostate	human	Affymetrix
GDS4110	9	3	1	prostate	human	Affymetrix
GDS4063	2	9	3	breast	human	Affymetrix
GDS4052	2	6	3	breast	human	Affymetrix
GDS3910	3	2	3	breast	human	Affymetrix
GDS3887	4	4	2	bladder	mouse	Affymetrix
GDS3872	3	3	2	ovarian	human	Affymetrix
GDS3761	3	4	2	prostate	human	Affymetrix
GDS3710	9	3	1	epithelial	human	Affymetrix
GDS3517	4	3	4	melanoma	human	Affymetrix
GDS3484	2	3	2	breast	human	Affymetrix
GDS3792	3	2	2	ovarian	human	Affymetrix
GDS3358	5	2	2	prostate	human	Affymetrix
GDS3319	3	1	3	thyroid	human	Affymetrix
GDS3285	4	3	1	breast	human	Affymetrix
GDS3283	2	6	2	breast	human	Affymetrix
GDS3116	58	1	2	breast	human	Affymetrix
GDS3111	3	3	2	prostate	human	Affymetrix
GDS2975	4	2	1	ovarian	human	Affymetrix
GDS2810	3	2	2	breast	human	Affymetrix
GDS2604	6	6		lung	human	Affymetrix
GDS2323	3	3	1	breast	human	Affymetrix
GDS2097	4	1	2	breast	human	Affymetrix
GDS2096	4	1	2	breast	human	Affymetrix
GDS1942	8	1	2	colon	human	Affymetrix
GDS1736	2	4	2	prostate	human	Affymetrix
GDS1627	3	2		breast	human	oligo
GDS848	3	7	3	breast	human	oligo
GDS725	10	1	1	prostate	human	Affymetrix
GDS719	4	2	3	prostate	human	Affymetrix
GDS528	3	3	1	colon	human	Affymetrix
GSE41828	5	5	2	prostate	human	Affymetrix
GSE41680	4	1	2	melanoma	human	Affymetrix
GSE31128	4	3	4	breast	human	Illumina
GSE28305	4	2	1	breast	human	Affymetrix
GDS4853	4	3	2	stomach	human	Affymetrix
GSE26834	2	3	6	breast	human	Affymetrix
GSE23616	3	3	2	ovarian	human	Affymetrix
GSE20361	8	1	1	breast	human	Affymetrix
GSE23845	4	4	1	bladder	human	Affymetrix
GSE15947	3	4	3	prostate	human	Affymetrix
GSE18494	4	3	3	breast	human	Affymetrix
GSE17708	10	3	2	lung	human	Affymetrix
GSE11428	3	3	6	prostate	human	Affymetrix
GSE16424	4	6	2	ovarian	human	Affymetrix
GSE13525	3	4	2	ovarian	human	Affymetrix
GSE13059	4	8	1	colon	human	Affymetrix
GSE8772	3	2	1	melanoma	human	Affymetrix
GSE8702	7	3	1	prostate	human	Affymetrix
GSE9936	2	3	16	breast	human	Affymetrix
GSE9826	6	3	3	ovarian	human	Affymetrix
GSE8057	5	4	8	ovarian	human	Affymetrix
GSE6013	5	2	3	lung	human	Affymetrix
GSE6930	3	6	4	ewing sarcoma	human	Affymetrix
GSE6653	4	2	1	ovarian	human	Affymetrix
GSE5345	7	3	2	prostate	human	IQUSP
GDS2034	4	2	1	ovarian	human	Affymetrix
GSE4917	4	8	3	breast	human	Affymetrix

Table 3.1 GEO data sets with a time-course design.

3.7 Comparison by simulation

The proposed ridge estimator is compared by means of simulation to (as far as we are aware) the only other method that estimates the VAR(1) model and its associated time series chain graph from high-dimensional data: SparseTSCGM proposed by [1]. SparseTSCGM too estimates parameters of the VAR(1) model in penalized fashion, but employs the smoothly clipped absolute deviation (SCAD, [55]) penalty on the autoregressive coefficient matrix $\mathbf{A}$ and precision matrix $\mathbf{\Omega}_\varepsilon$. As the estimation methods estimate the same VAR(1) model and only differ in the employed penalties, they can readily compared in terms of loss of the estimates and selection of the edges of the time series chain graph.

The ridge and 'SCAD' estimators of the VAR(1) model are compared under various choices of the model parameters $\mathbf{A}$ and $\mathbf{\Omega}_\varepsilon$, while the number of variates p, time points $\mathcal{T}$ and the number of samples n are varied in accordance with a factorial design. We first describe the levels of the factors of the factorial design. The number of variates equals either $p = 25$ or $p = 50$[1], representing the size of non-trivial but well-defined pathways. Simultaneously, we set $n \in \{5, 15\}$ and $\mathcal{T} \in \{10, 20\}$. We stress that the particular case of $(n = 5, \mathcal{T} = 10)$ is the most challenging but also most relevant as its design is closest to that which is most prevalent in practice. [51] observed that more than 80% of time-course gene expression microarray data sets have 8 or fewer number of time points. This is confirmed by a more recent survey of our own (confer SM III).

We now outline parameter choices of the various VAR(1) models employed in the simulation. Autoregression matrices $\mathbf{A}$ with a hub, cluster, random (each with two sparsity levels: approximately 5% and 25% nonzero elements), and full network structure choices are used. These matrices $\mathbf{A}$ are used in combination with a banded and full precision matrix $\mathbf{\Omega}_\varepsilon$ and two data-driven variants. We first detail the various choices of $\mathbf{\Omega}_\varepsilon$:

- Full $\mathbf{\Omega}_\varepsilon$: In the precision matrix $\mathbf{\Omega}_\varepsilon$ existing contemporaneous conditional independencies is represented with $\rho = 0.5$.
- Banded $\mathbf{\Omega}_\varepsilon$: A banded precision matrix is generated with $(\mathbf{\Omega}_\varepsilon)_{j_1,j_2} = \rho^{|j_1-j_2|}$ with $\rho = 0.5$ for $|j_1 - j_2| < 3$ and $(\mathbf{\Omega}_\varepsilon)_{j_1,j_2} = 0$ otherwise.
- Full but data-driven $\mathbf{\Omega}_\varepsilon$: The precision matrix is obtained from a VAR(1) model fitted to data from a time-course gene experiment experiment. The data of the

[1] Initially, we also included $p = 100$, but the estimation with the SCAD procedure fails to converge for $n = 5$ and $T = 10$, as is often the case for $p = 100$ with larger n and $\mathcal{T}$.

application detailed in the main document are used. The data are limited to genes mapping to the p53 signaling pathway as defined by KEGG. The VAR(1) model is fitted to these pathway data by means of the proposed ridge estimation procedure using penalty parameter values $\lambda_a = 0.3, \lambda_\omega = 0.1$. The thus estimated $\mathbf{\Omega}_\varepsilon$ is used in the simulation. For the various choice of p, the first p genes from the pathway are used in the comparison.

For the regression coefficient matrix $\mathbf{A}$ the following variants are employed:

- Hub $\mathbf{A}$: All elements of $\mathbf{A}$ are set to zero, except for $(\mathbf{A})_{j_1,j_2} = 0.3$ for $j_1 = \{d, 2d, 3d, ...\}$ with $j_1 > j_2$ and $d = 10$ and $d = 2$ for the less sparse case). Hence, only a few variates affect (in varying degree) the temporal variation of the others.
- Cluster $\mathbf{A}$: Eight (most sparse) or two (less sparse) equally-sized lower-triangular blocks filled with the value 0.3 are aligned along the diagonal. The remaining elements of $\mathbf{A}$ are zero.
- Random $\mathbf{A}$: A random network is generated, selecting 10% or 50% of the total possible edges, after which the lower triangle is retains, i.e. setting $(\mathbf{A})_{j_1,j_2} = 0$ if $j_1 < j_2$. Elements of $\mathbf{A}$ are set equal to 0.3 if the corresponding element in the adjacency matrix of the random network is nonzero.
- Full data-driven $\mathbf{A}$: The same data set as for the generation of the data-driven $\mathbf{\Omega}_\varepsilon$ is used to estimate $\mathbf{A}$. Using these data $\mathbf{A}$ is estimated by the ridge procedure with penalty parameters $\lambda_a = 0.3$ and $\lambda_\omega = 0.1$. The resulting $\mathbf{A}$ is used in the simulation.
- Sparse data-driven $\mathbf{A}$ is obtained by sparsifying (as outlined in Section 4.1 of the main text) the full data-driven $\mathbf{A}$.

Note that the majority of choices for $\mathbf{A}$ and $\mathbf{\Omega}_\varepsilon$ is sparse. This in principle favors the 'SCAD' estimation procedure SparseTSCGM.

For (combination of) these choices for the regression parameter matrix $\mathbf{A}$ and precision matrix $\mathbf{\Omega}_\varepsilon$, data are simulated in accordance with the VAR(1) model. Data for the first time point $t = 1$ are drawn from $\mathcal{N}(\mathbf{0}_{p\times 1}, \mathbf{\Omega}_\varepsilon^{-1})$ and for the following time points $t = 2, ..., T$ data are sampled from $\mathcal{N}(\mathbf{A}\mathbf{Y}_{*,i,t-1}, \mathbf{\Sigma}_\varepsilon)$. In total for each employed combination of the model parameters and design parameters p, n, $\mathcal{T}$, fifty time series data sets are generated.

For each simulated data set, the VAR(1) model is fitted using both the proposed ridge procedure and the 'SCAD'-based method SparseTSCGM. The parameters estimates obtained from both methods are chosen in an identical fashion: using maximization of the leave K-fold-out cross-validated log-likelihood. Instead of removing a

single element from the set of design points (as discussed in Section 3.4 of the main document), an individual (cell line) with all its time points is removed at each fold. This is due to the fact that the SparseTSCGM cannot deal with an unbalanced design. Formally, from the set of design points $\mathcal{D} = \{(t,i) : t = 1,\ldots,\mathcal{T}, i = 1,\ldots,n\}$ we remove a single sample $\mathcal{D} = \{(i) : i = 1,\ldots,n\}$ at the time. Resulting parameter estimates of $\mathbf{A}$ and $\boldsymbol{\Omega}_\varepsilon$ obtained by SparseTSCGM are sparse, whereas their ridge counterparts are not. This is irrelevant for the loss comparison (for which both a sparse and full matrix can be used). Only comparing the methods with respect to the ability to reconstruct the time series chain graph, the ridge estimates undergo post-estimation sparsification (discussed below).

The ridge and SCAD estimators of $\widehat{\mathbf{A}}(\lambda_a)$ and $\widehat{\boldsymbol{\Omega}}_\varepsilon(\lambda_a,\lambda_\omega)$ are evaluated in terms of their Frobenius loss and their ability to select edges of time series chain graph. The Frobenius loss for (e.g.) a precision matrix estimate is:

$$\left\|\widehat{\boldsymbol{\Omega}}_\varepsilon(\lambda_a,\lambda_\omega) - \boldsymbol{\Omega}_\varepsilon\right\|_F^2 \;=\; \sum_{j_1,j_2}\left\{\left[\widehat{\boldsymbol{\Omega}}_\varepsilon(\lambda_a,\lambda_\omega)\right]_{j_1,j_2} - \left(\boldsymbol{\Omega}_\varepsilon\right)_{j_1,j_2}\right\}^2.$$

The Frobenius loss for the estimate of $\mathbf{A}$ is defined similarly.

The edge selection ability of the ridge and SCAD estimation methods are compared by means of sensitivity and specificity. When the ridge estimation is augmented with the post-estimation empirical Bayes procedure, the two methods may yield differing number of edges. This may hamper the comparison of their sensitivity and specificity. To facilitate a better comparison, also adhering to the *ceteris paribus* principle, of the two methods in this respect, a simple but different post-estimation edge selection procedure is applied to the ridge estimate. It comprises of selecting the same number of edges as the SCAD method, thus favouring the latter. Having fixed the number of to-be-selected edges, they are selected on the basis of the (absolute) size of the statistic derived from the elements of $\mathbf{A}$ (as pointed out in Section 4.1 the main document) and (standardized) $\widehat{\boldsymbol{\Omega}}_\varepsilon(\lambda_a,\lambda_\omega)$. This thus means that in each data set for the number of edges selected by the SCAD estimator, we select the same number of largest elements of from the estimate of $\mathbf{A}$.

We first discuss the Frobenius loss comparison. Figures 3.6 and 3.14 show the Frobenius loss (as boxplots) of the ridge and SCAD estimates of $\mathbf{A}$ (upper panel) with roughly 5% and 25% nonzero elements, respectively, and $\boldsymbol{\Omega}_\varepsilon$ (lower panel) from fifty data sets generated with $p = 25$, $T = 10$, $n = 5$ (the most relevant case empirically), and the various combinations of $\mathbf{A}$ and $\boldsymbol{\Omega}_\varepsilon$. The panels of Figures 3.6 and 3.14 reveal that the Frobenius loss of the SCAD estimates of both $\widehat{\mathbf{A}}(\lambda_a)$ and $\widehat{\boldsymbol{\Omega}}_\varepsilon(\lambda_\omega)$ exceeds that

of its ridge counterpart. This observation is consistent over the data sets generated from different choices of the regression coefficient (with all levels of sparsity) and precision matrices. This is confirmed when the VAR(1) model is increased to include $p = 50$ variates. The difference in Frobenius loss between the ridge and SCAD estimators grows substantially in favour of the former (confer Figures 3.7 and 3.15). When the number of time points $\mathcal{T}$ is increased, the ridge procedure outperforms its SCAD counterpart (as can be witnessed from Figures 3.8 and 3.16 and Figures 3.9 and 3.17 representing the '$T = 20$, $n = 5$'-case with $p = 25$ and $p = 50$, respectively.

Increasing the number of samples to the practically non-observed (confer **??**) $n = 15$ benefits the SCAD estimator in the cases with $p = 25$ and the both sparse choices of $\mathbf{A}$ (confer Figures 3.10, 3.18, 3.12, and 3.20). When the number of variates increases to $p = 50$ the SCAD still outperforms the ridge estimator in for sparsest $\mathbf{A}$ (Figure 3.13) but roles are reversed when $\mathbf{A}$ contains roughly 25% nonzero elements (Figure 3.20).

In the latter the superior Frobenius loss of the SCAD estimator is consequence of the sparse $\widehat{\mathbf{A}}(\lambda_a)$ and $\widehat{\mathbf{\Omega}}_\varepsilon(\lambda_\omega)$ which favours the SCAD estimator due to the sparsity of the estimates. This is verified by comparing the ridge and SCAD estimates separately for zero and non-zero elements of $\mathbf{A}$ (not shown). The SCAD estimator performs better for the zero elements, while the ridge estimator outperforms on non-zero elements of $\mathbf{A}$. Only for the hub network does the SCAD estimator generally outperforms the ridge estimator.

We now turn to the sensitivity and specificity of the ridge procedure and that of the SCAD procedure (Figures 3.22, 3.23, 3.24, 3.25, 3.26, 3.27, 3.28, and 3.29). Recall, to make the sensitivity and specificity more comparable, the ridge procedure selects (from the largest elements elements of its estimate of $\mathbf{A}$) the same number of edges as found by the SCAD method, thus possibly favouring the latter. In the general picture that then emerges the methods are on a par. For the lower sample sized (and thereby more prevalent) settings a difference is virtually absent. More discriminative power is present when $n = 15$ and $mathcalT$, but even then no clear winner appear (although then the hub case appears to favour the SCAD method).

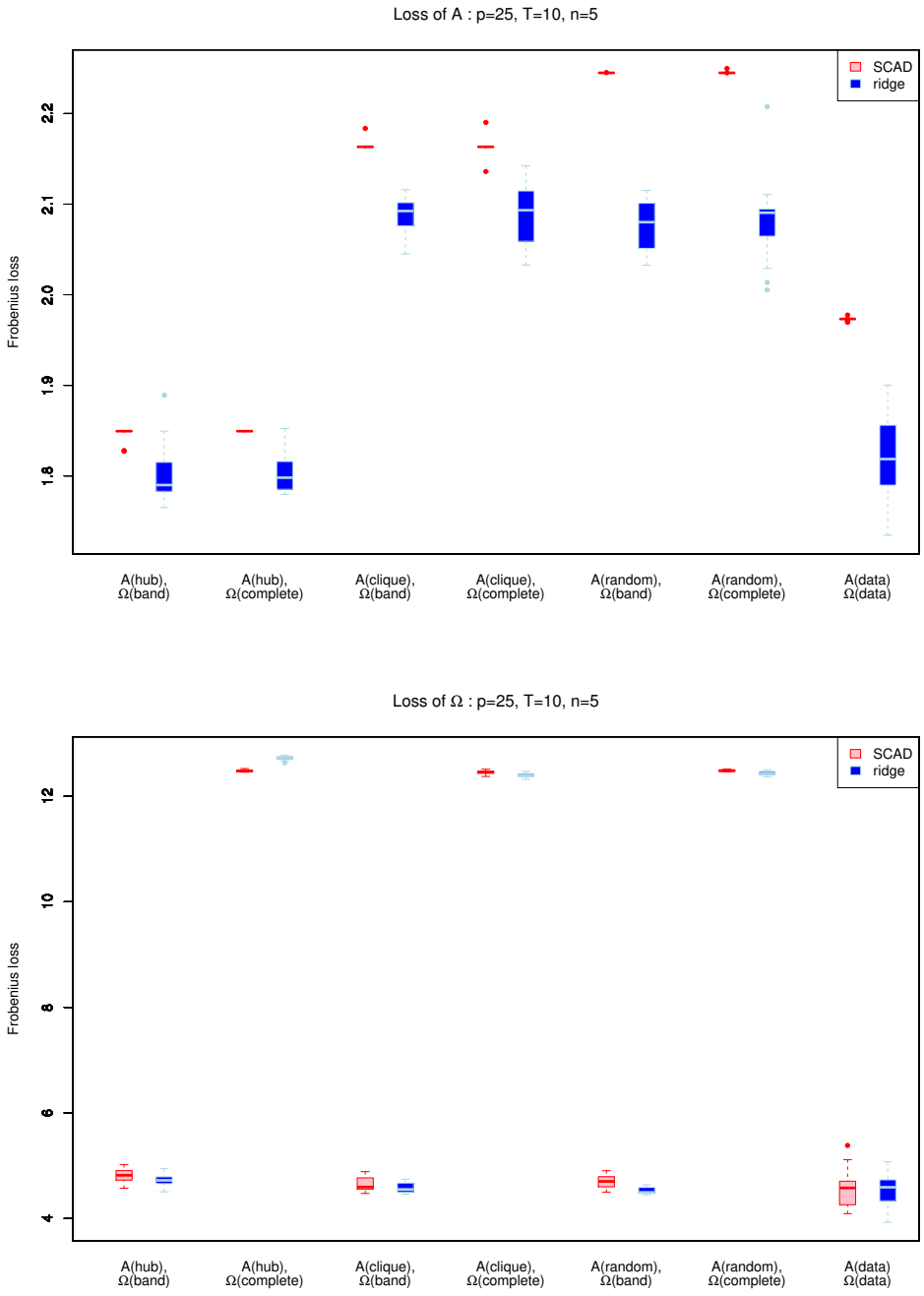

Fig. 3.6 Frobenius loss comparison between SCAD and ridge estimators for precision and autoregressive coefficient matrix on simulated data set where p=25, T=10, n=5 and **A** with roughly 5% nonzero elements.

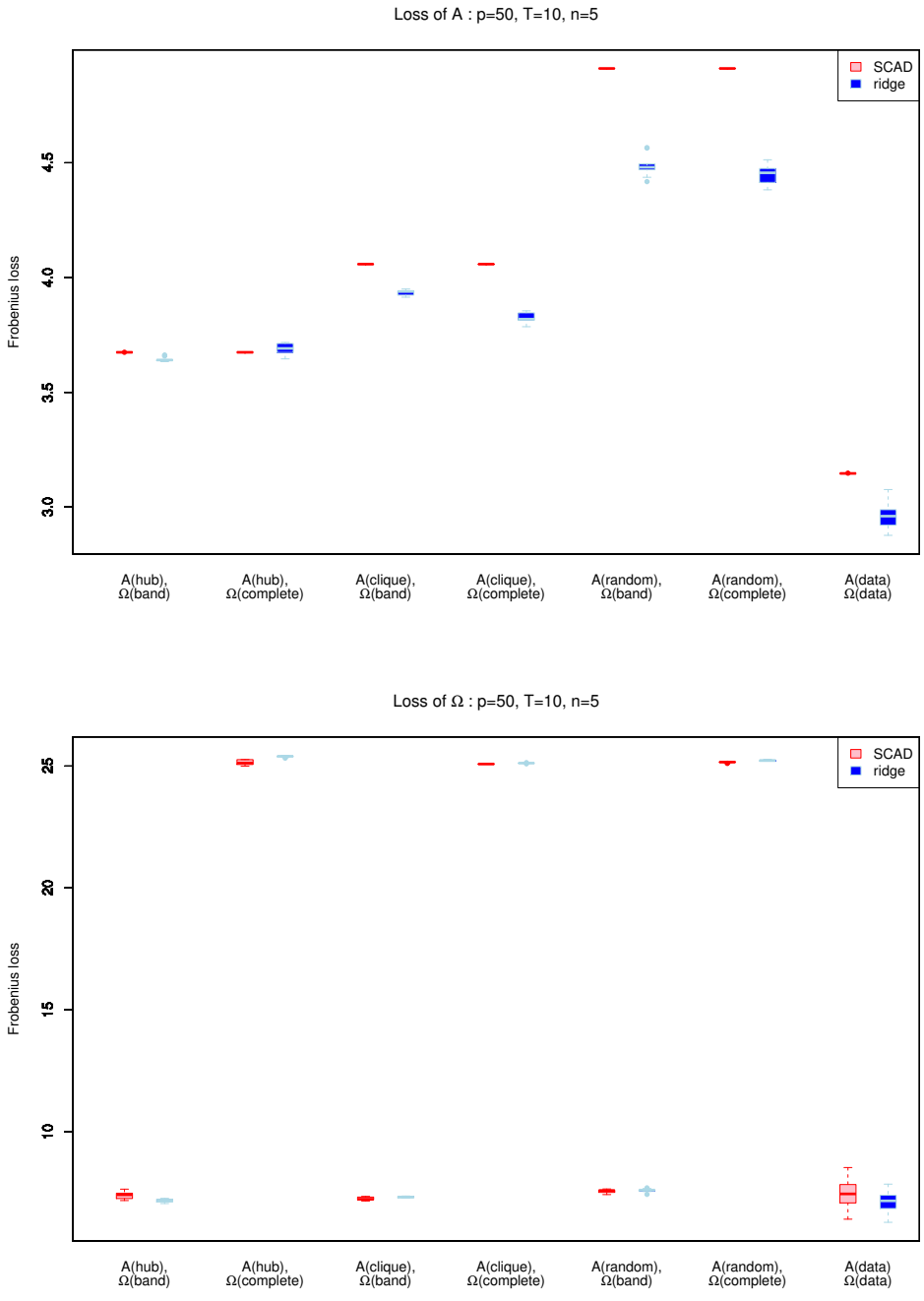

Fig. 3.7 Frobenius loss comparison between SCAD and ridge estimators for precision and autoregressive coefficient matrix on simulated data set where p=50, T=10, n=5 and **A** with roughly 5% nonzero elements.

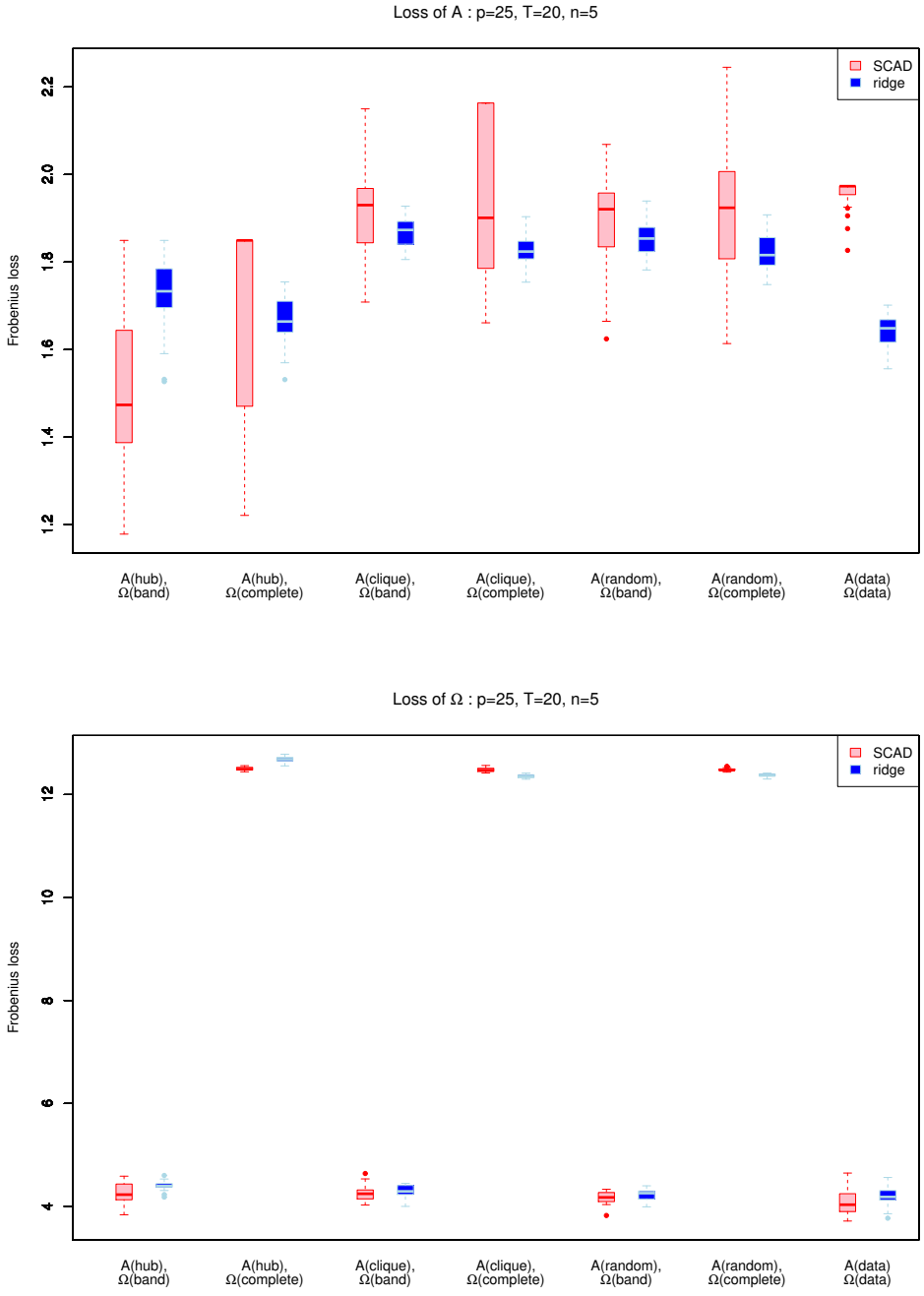

Fig. 3.8 Frobenius loss comparison between SCAD and ridge estimators for precision and autoregressive coefficient matrix on simulated data set where p=25, T=20, n=5 and **A** with roughly 5% nonzero elements.

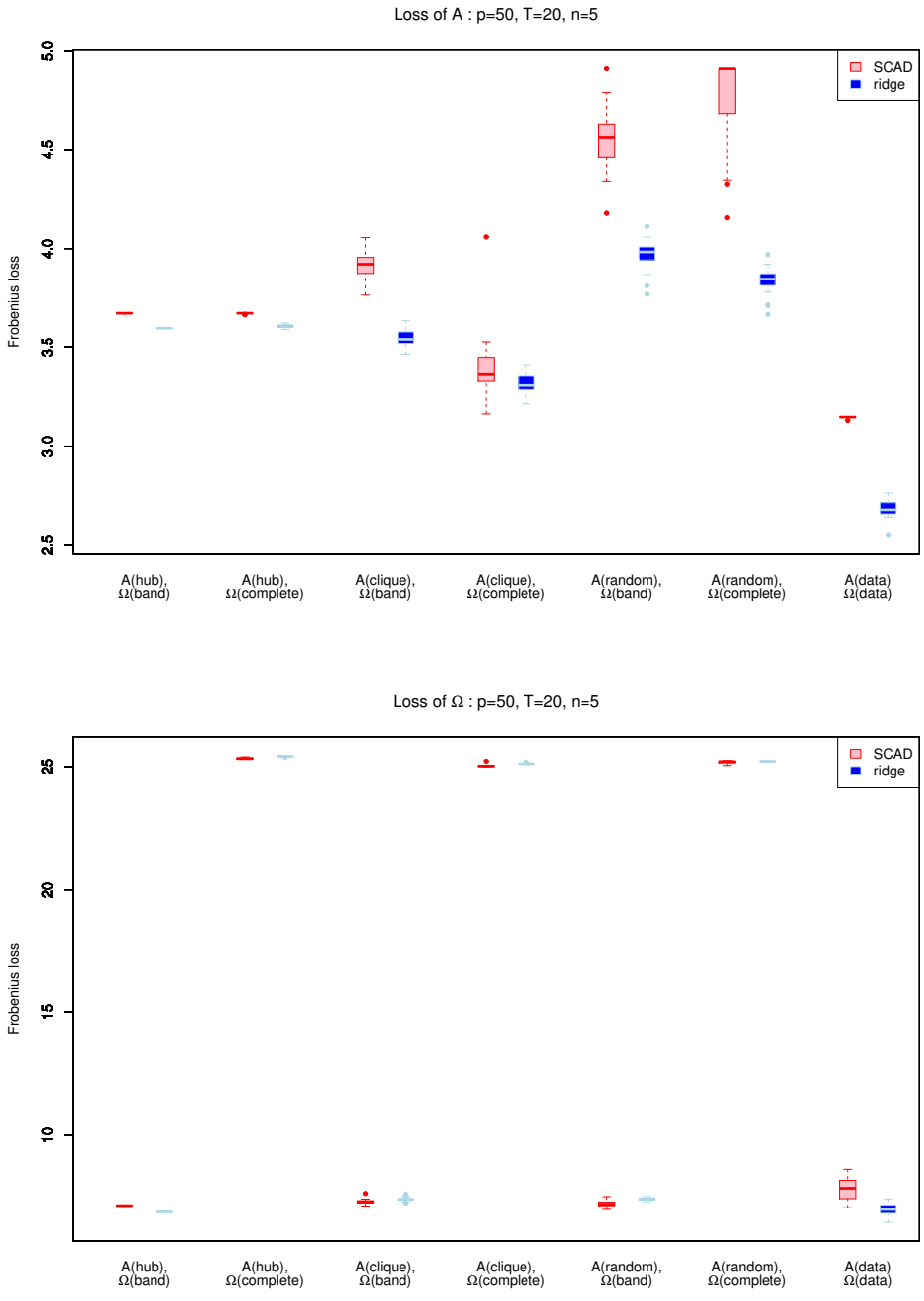

Fig. 3.9 Frobenius loss comparison between SCAD and ridge estimators for precision and autoregressive coefficient matrix on simulated data set where p=50, T=20, n=5 and **A** with roughly 5% nonzero elements.

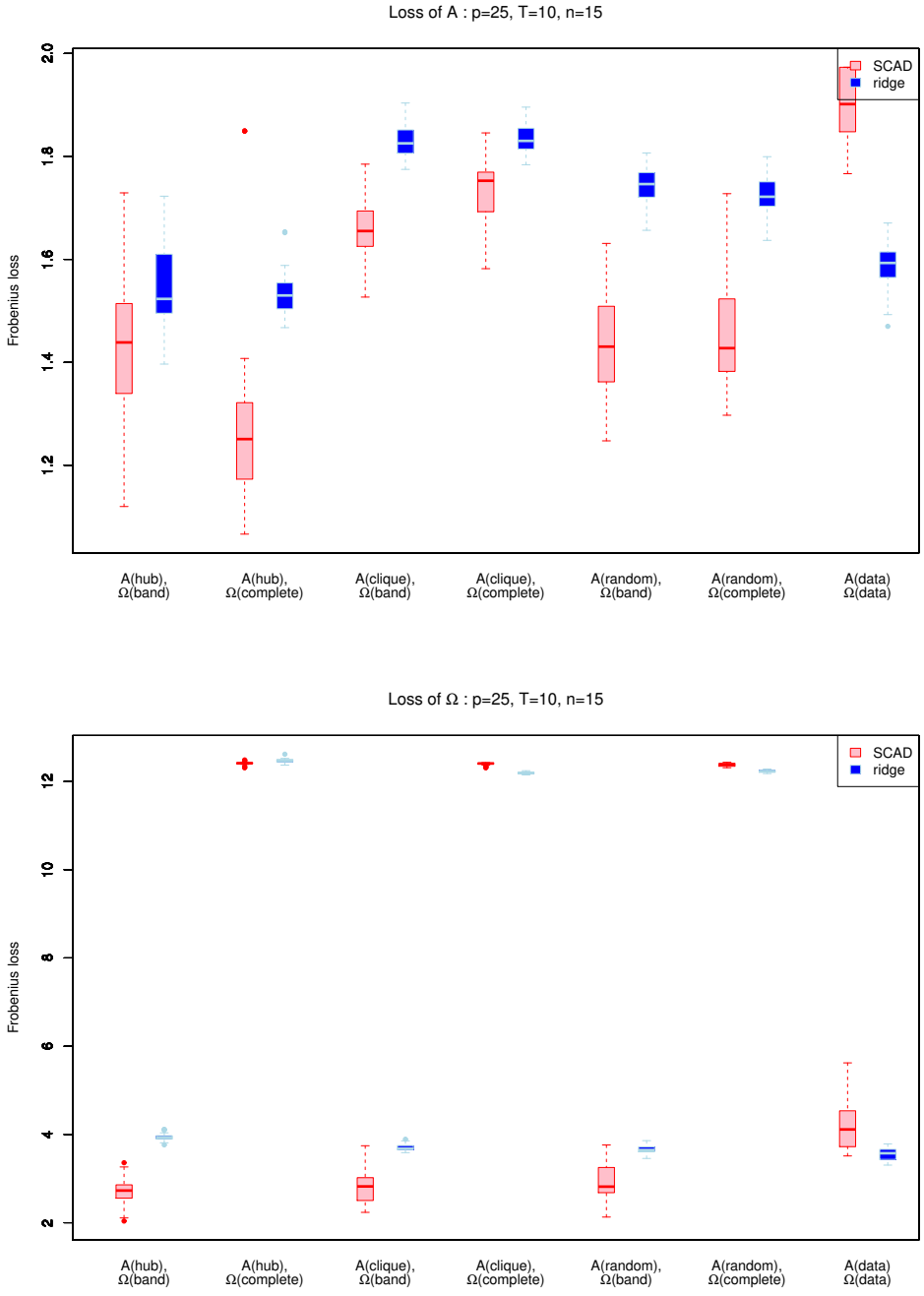

Fig. 3.10 Frobenius loss comparison between SCAD and ridge estimators for precision and autoregressive coefficient matrix on simulated data set where p=25, T=10, n=15 and **A** with roughly 5% nonzero elements.

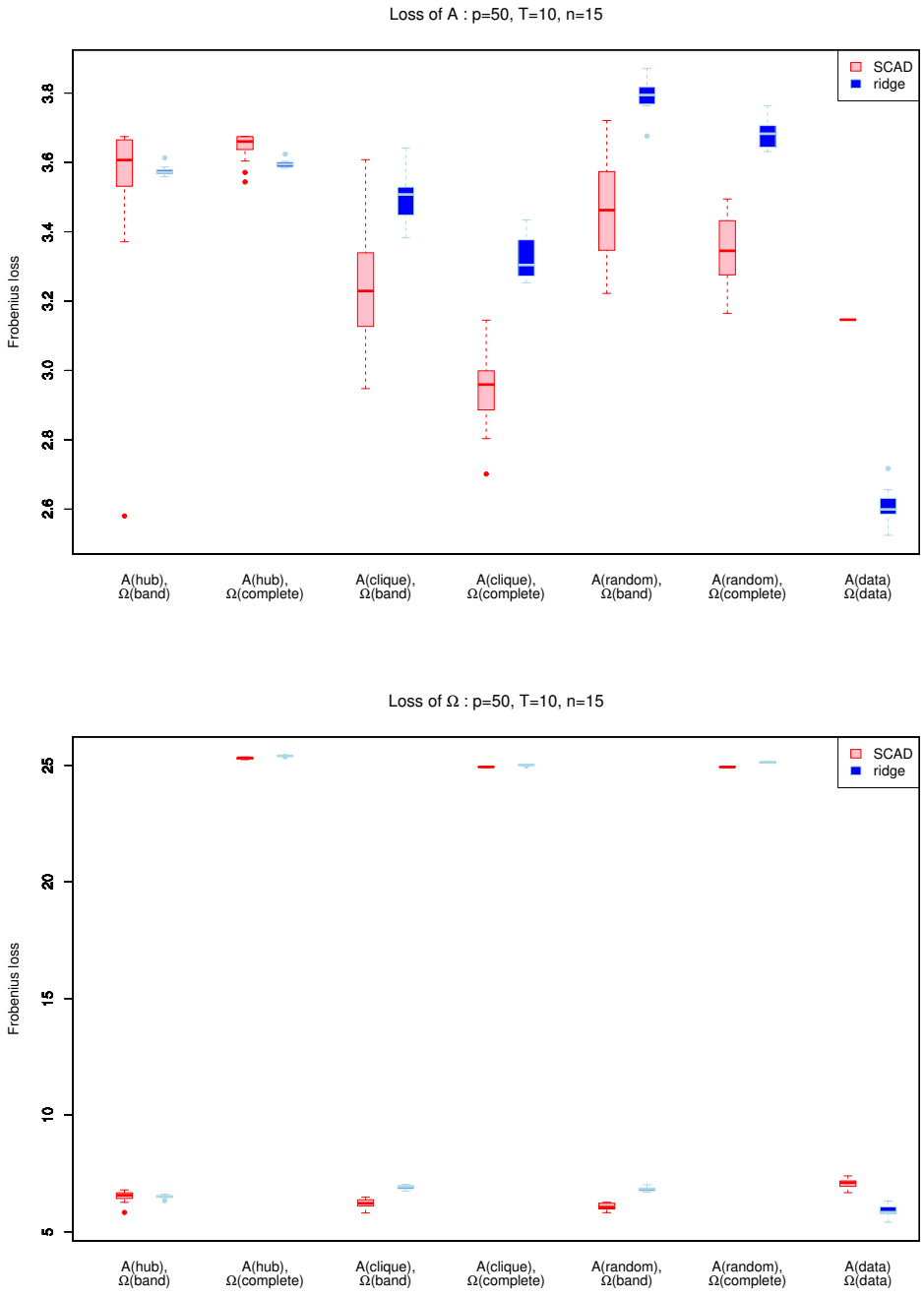

Fig. 3.11 Frobenius loss comparison between SCAD and ridge estimators for precision and autoregressive coefficient matrix on simulated data set where p=50, T=10, n=15 and **A** with roughly 5% nonzero elements.

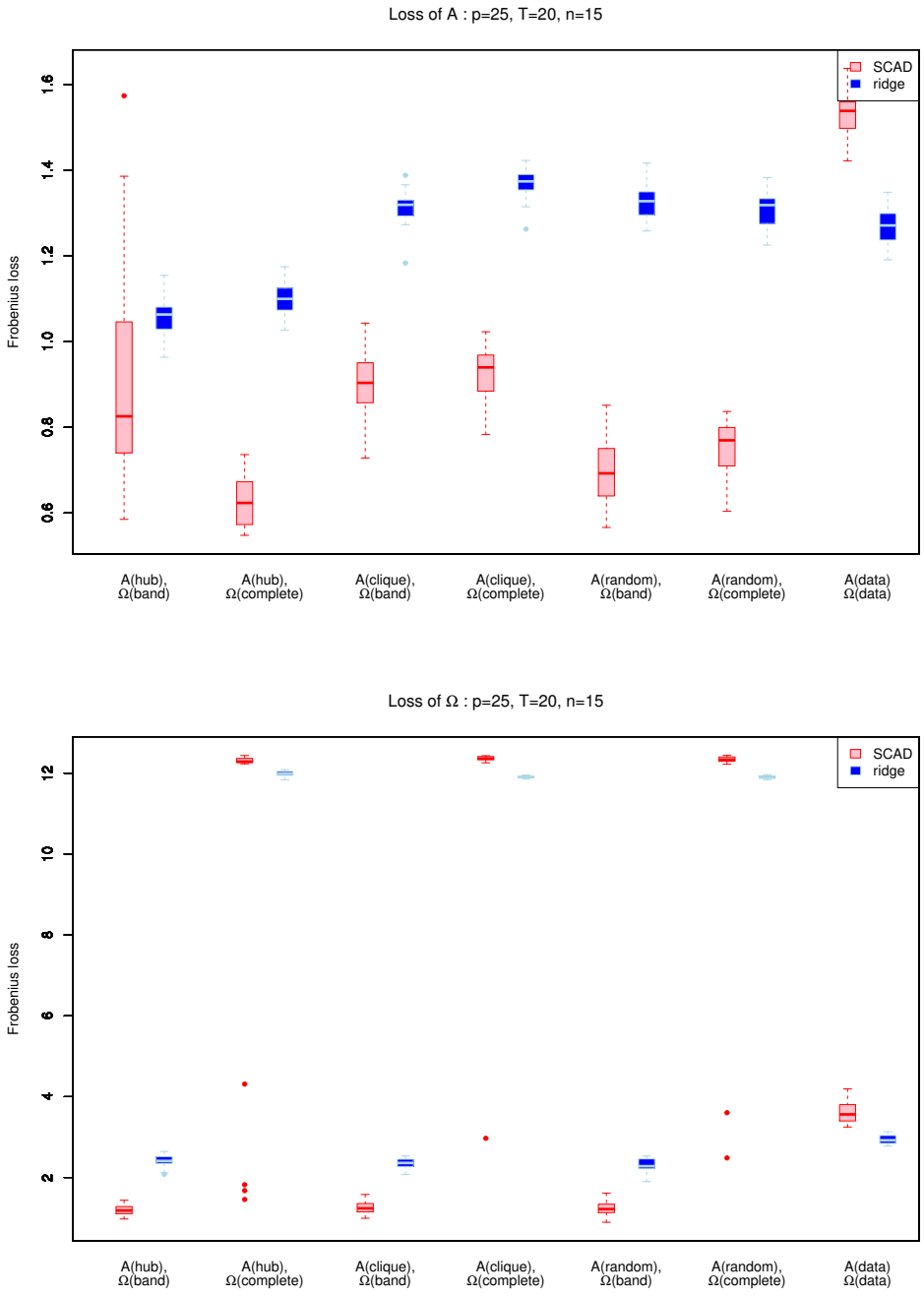

Fig. 3.12 Frobenius loss comparison between SCAD and ridge estimators for precision and autoregressive coefficient matrix on simulated data set where p=25, T=20, n=15 and $\mathbf{A}$ with roughly 5% nonzero elements.

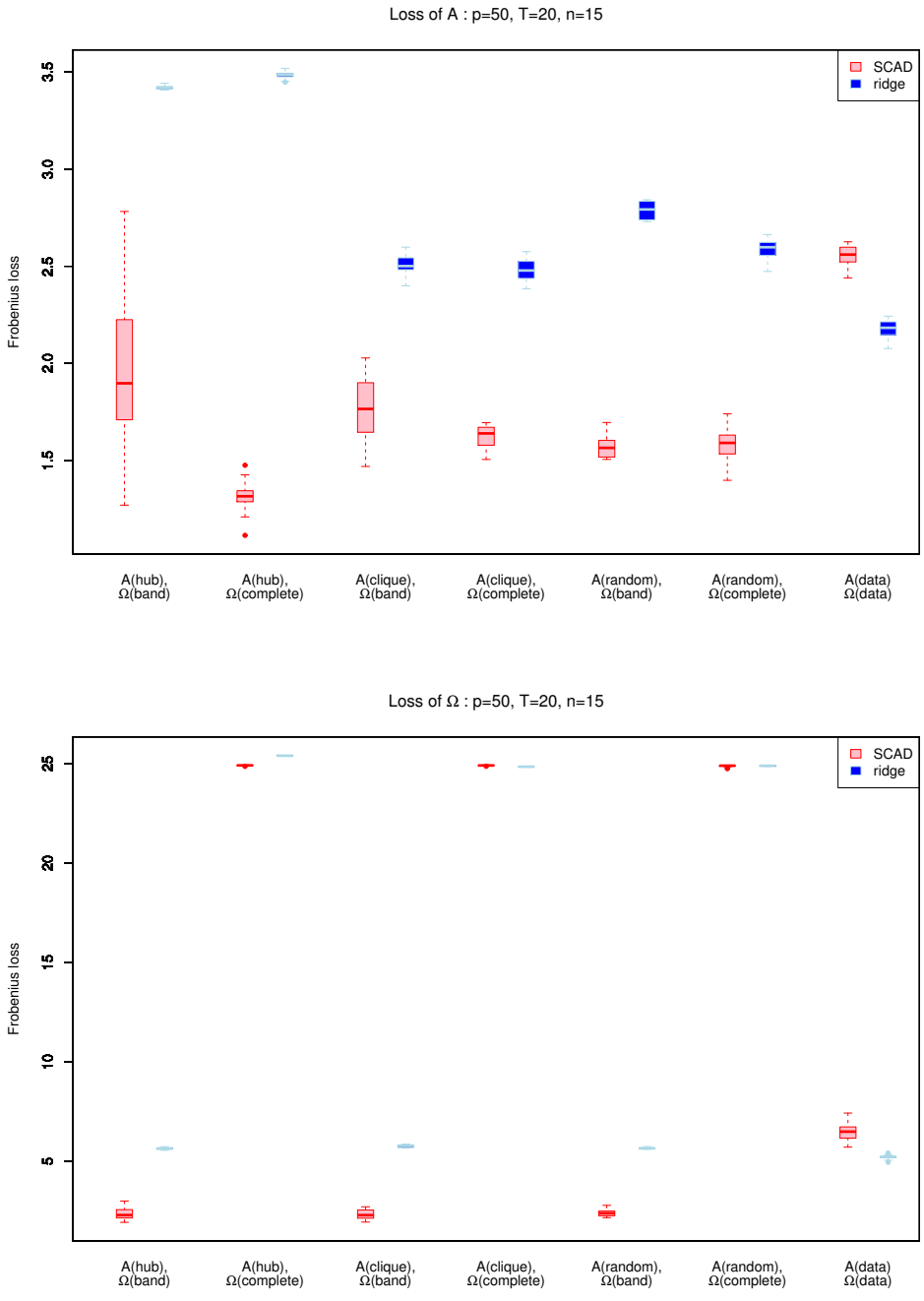

Fig. 3.13 Frobenius loss comparison between SCAD and ridge estimators for precision and autoregressive coefficient matrix on simulated data set where p=50, T=20, n=15 and $\mathbf{A}$ with roughly 5% nonzero elements.

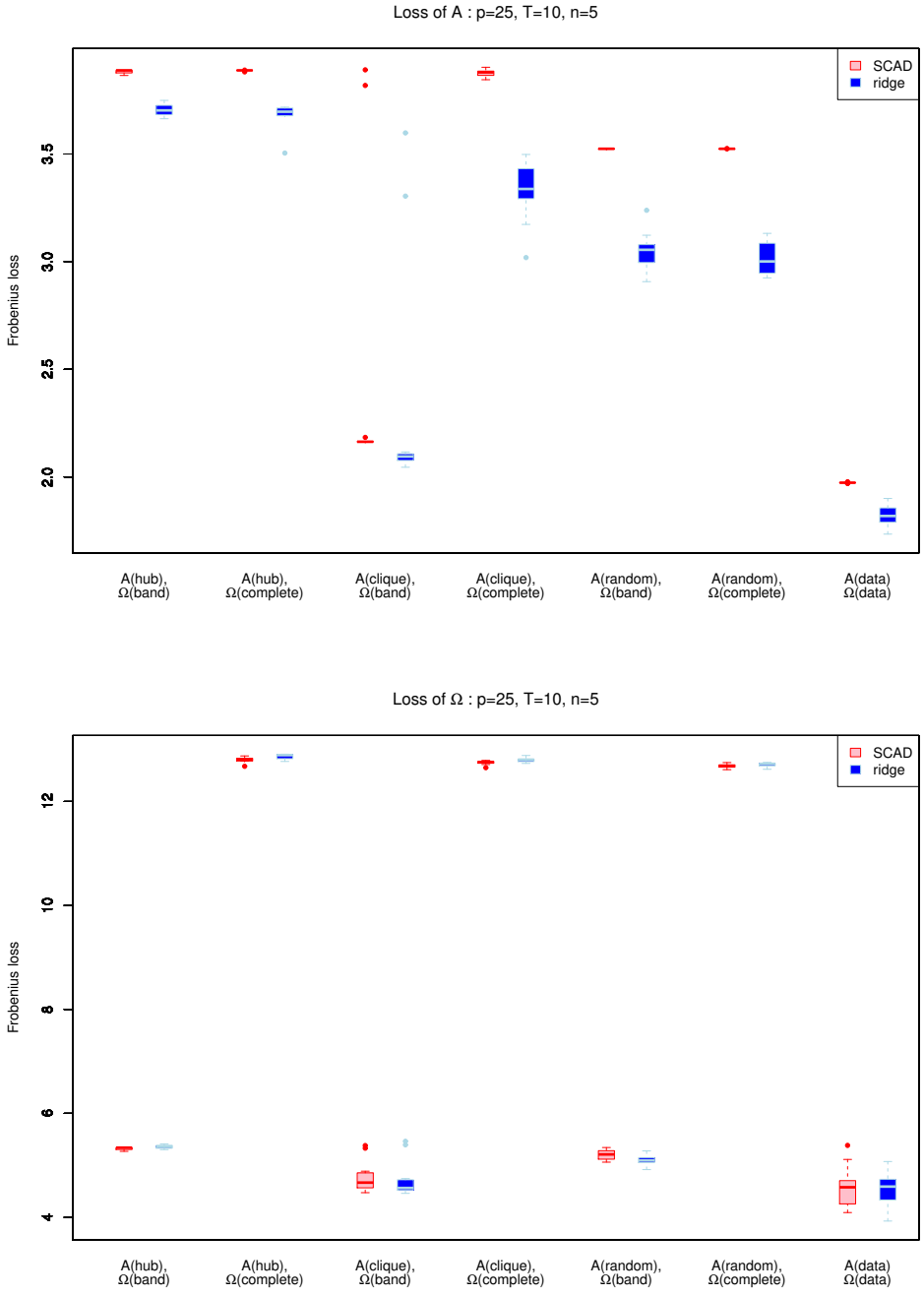

Fig. 3.14 Frobenius loss comparison between SCAD and ridge estimators for precision and autoregressive coefficient matrix on simulated data set where p=25, T=10, n=5 and **A** with roughly 25% nonzero elements.

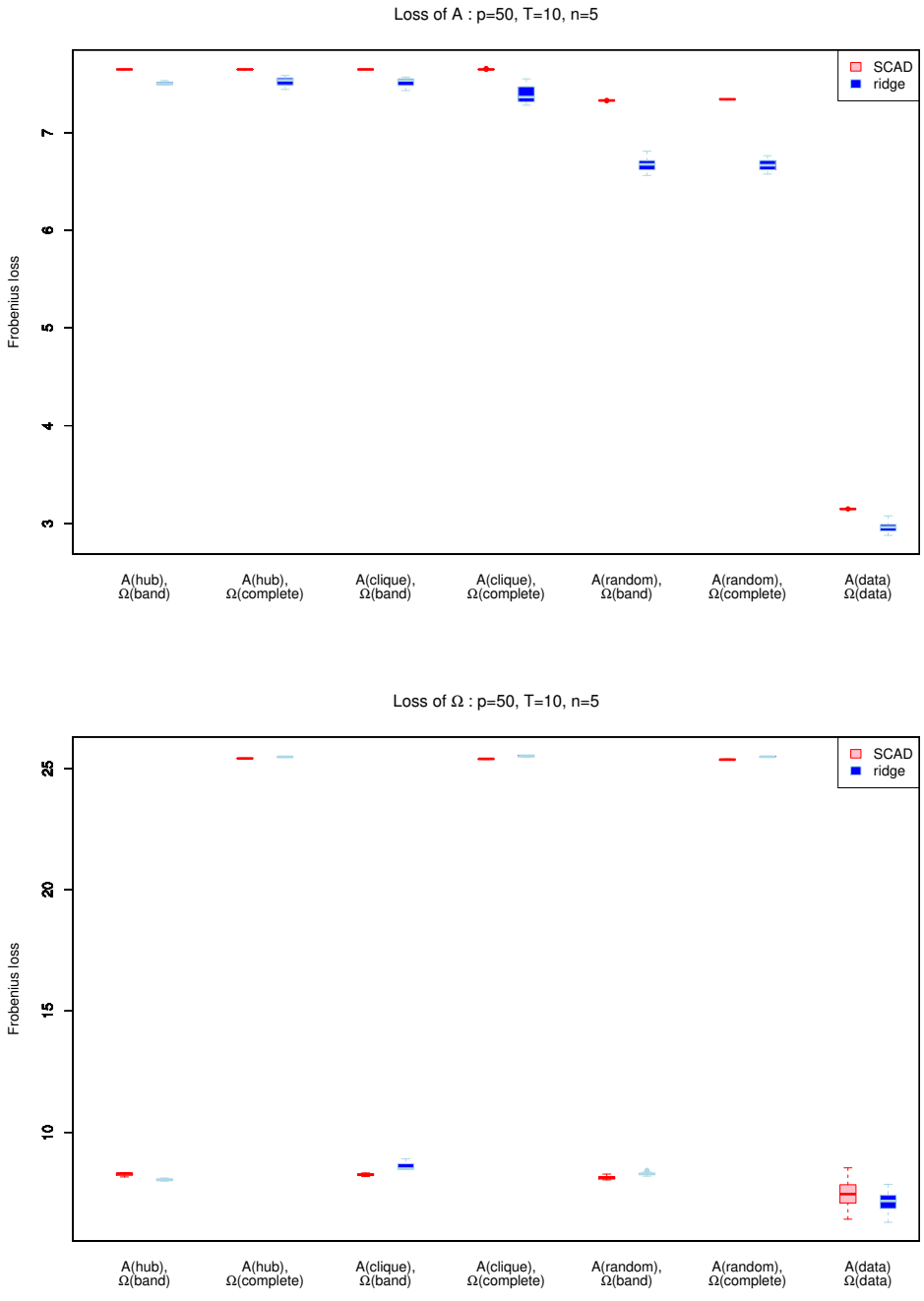

Fig. 3.15 Frobenius loss comparison between SCAD and ridge estimators for precision and autoregressive coefficient matrix on simulated data set where p=50, T=10, n=5 and **A** with roughly 25% nonzero elements.

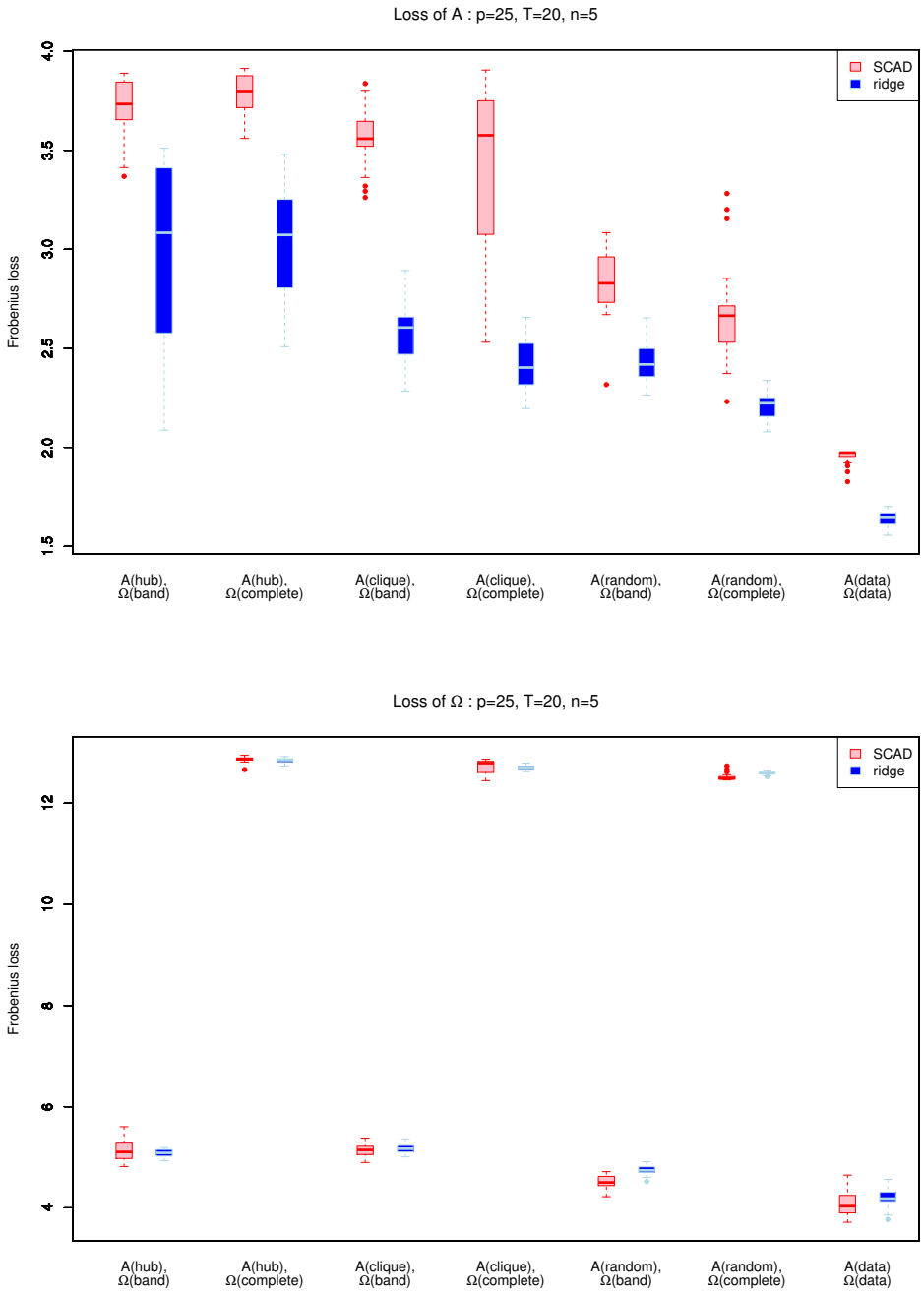

Fig. 3.16 Frobenius loss comparison between SCAD and ridge estimators for precision and autoregressive coefficient matrix on simulated data set where p=25, T=20, n=5 and **A** with roughly 25% nonzero elements.

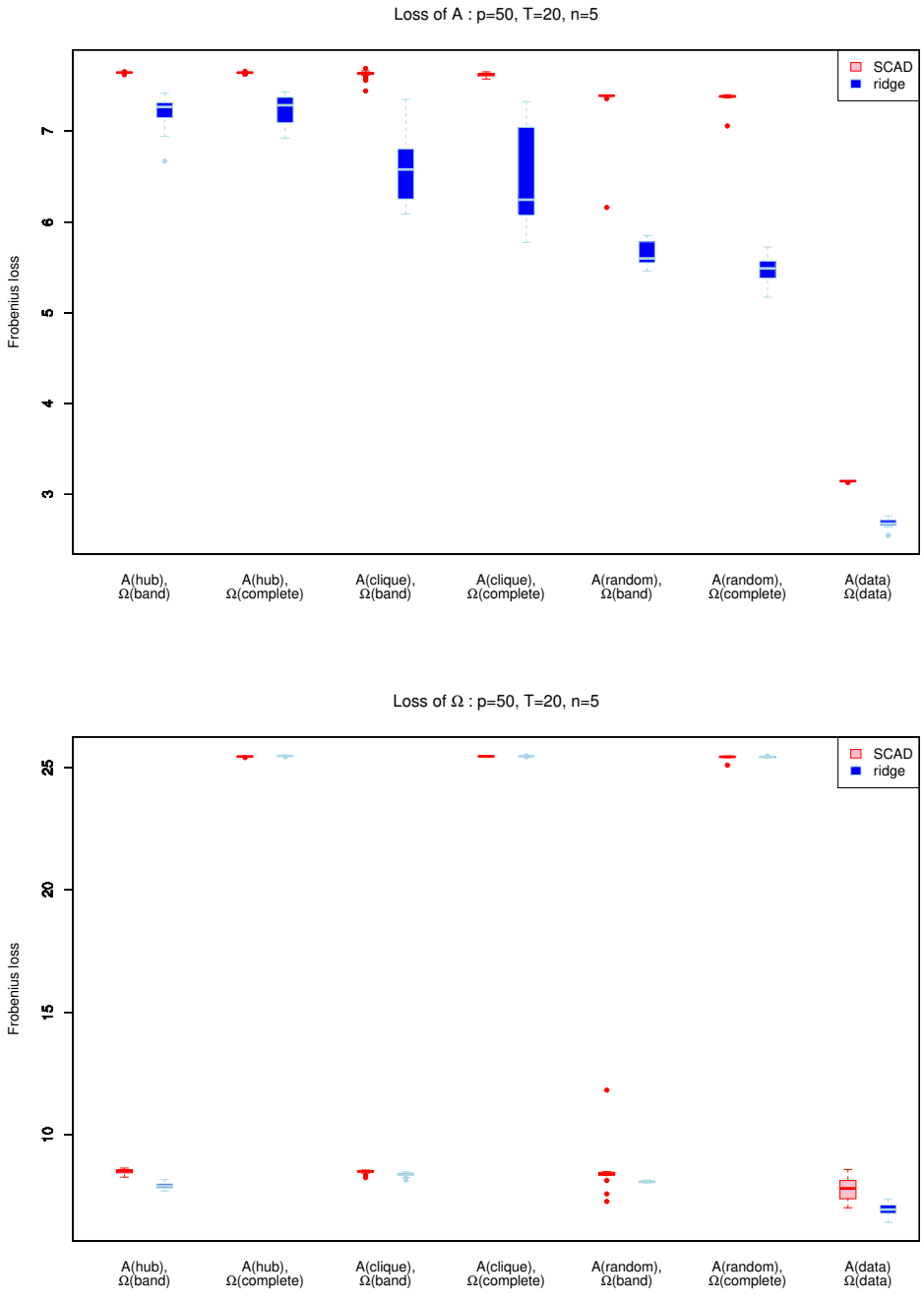

Fig. 3.17 Frobenius loss comparison between SCAD and ridge estimators for precision and autoregressive coefficient matrix on simulated data set where p=50, T=20, n=5 and **A** with roughly 25% nonzero elements.

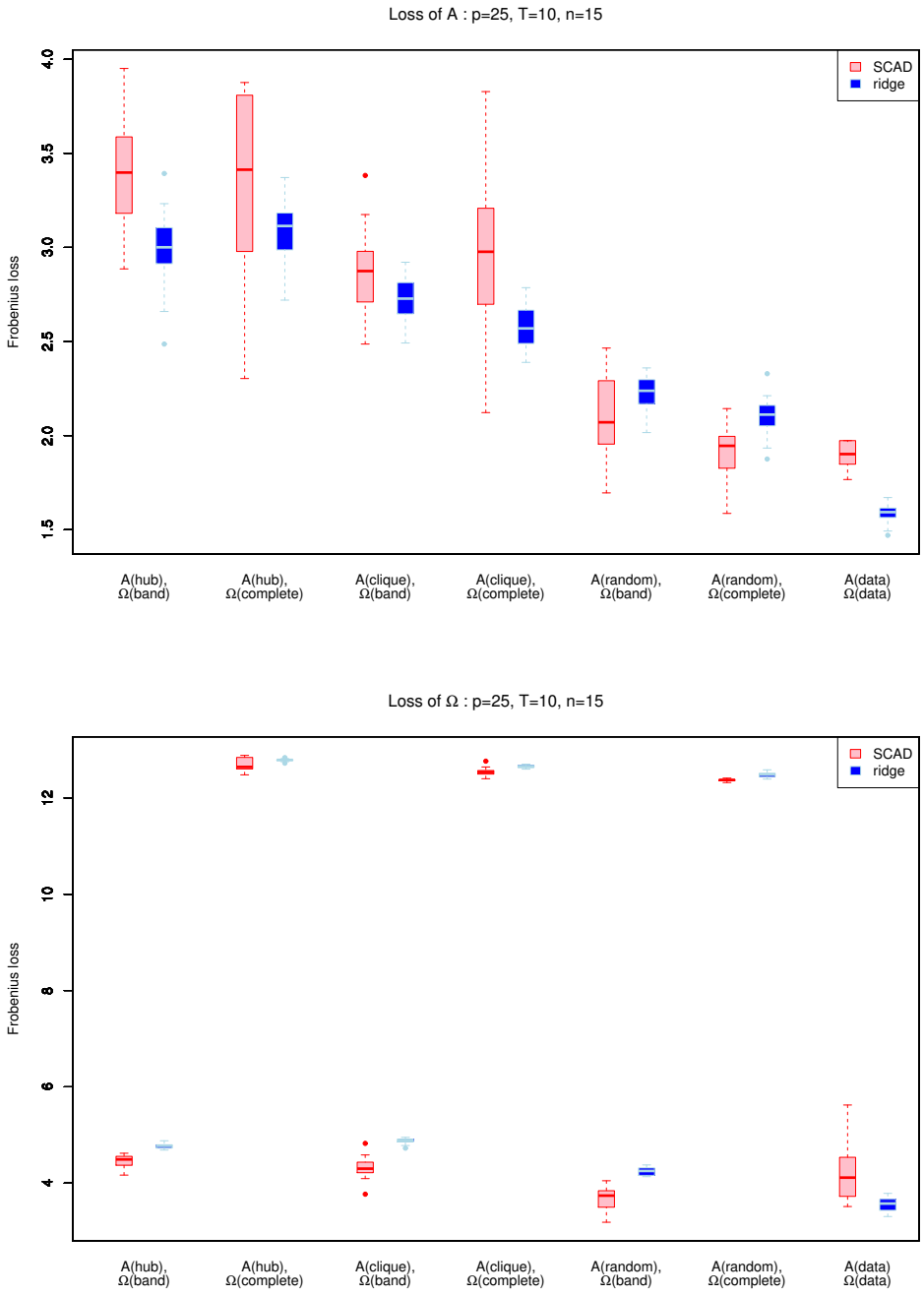

Fig. 3.18 Frobenius loss comparison between SCAD and ridge estimators for precision and autoregressive coefficient matrix on simulated data set where p=25, T=10, n=15 and **A** with roughly 25% nonzero elements.

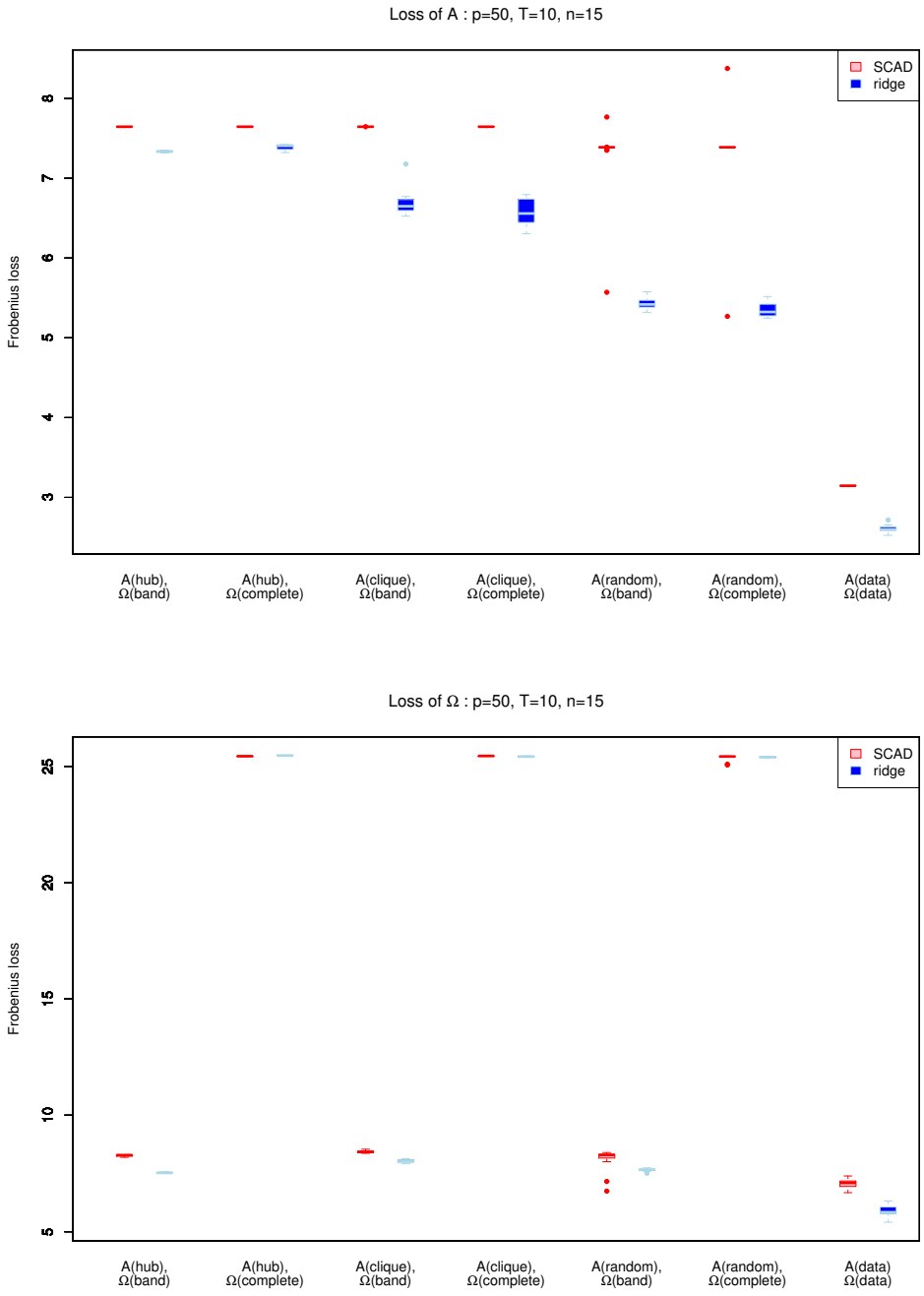

Fig. 3.19 Frobenius loss comparison between SCAD and ridge estimators for precision and autoregressive coefficient matrix on simulated data set where p=50, T=10, n=15 and $\mathbf{A}$ with roughly 25% nonzero elements.

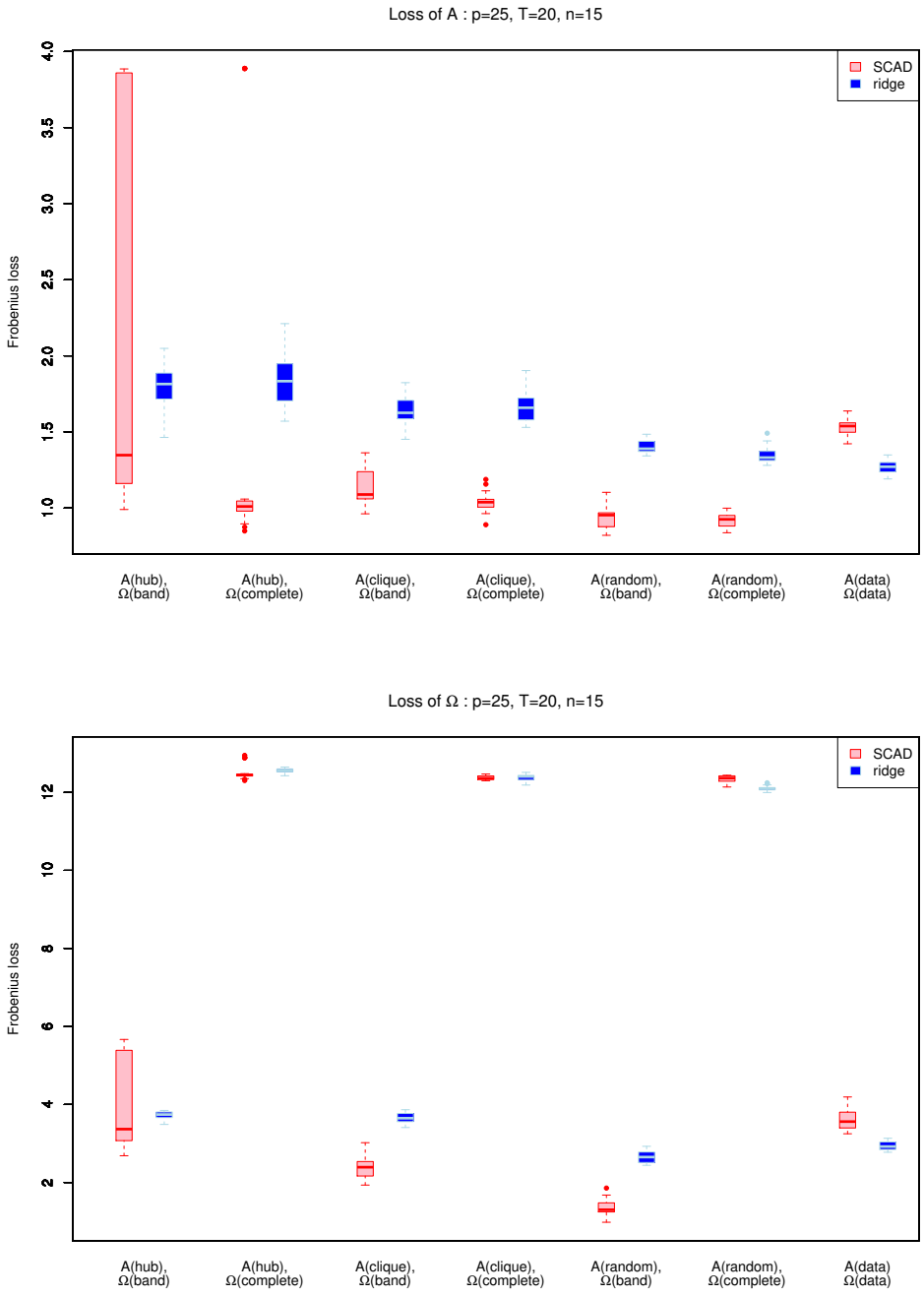

Fig. 3.20 Frobenius loss comparison between SCAD and ridge estimators for precision and autoregressive coefficient matrix on simulated data set where p=25, T=20, n=15 and $\mathbf{A}$ with roughly 25% nonzero elements.

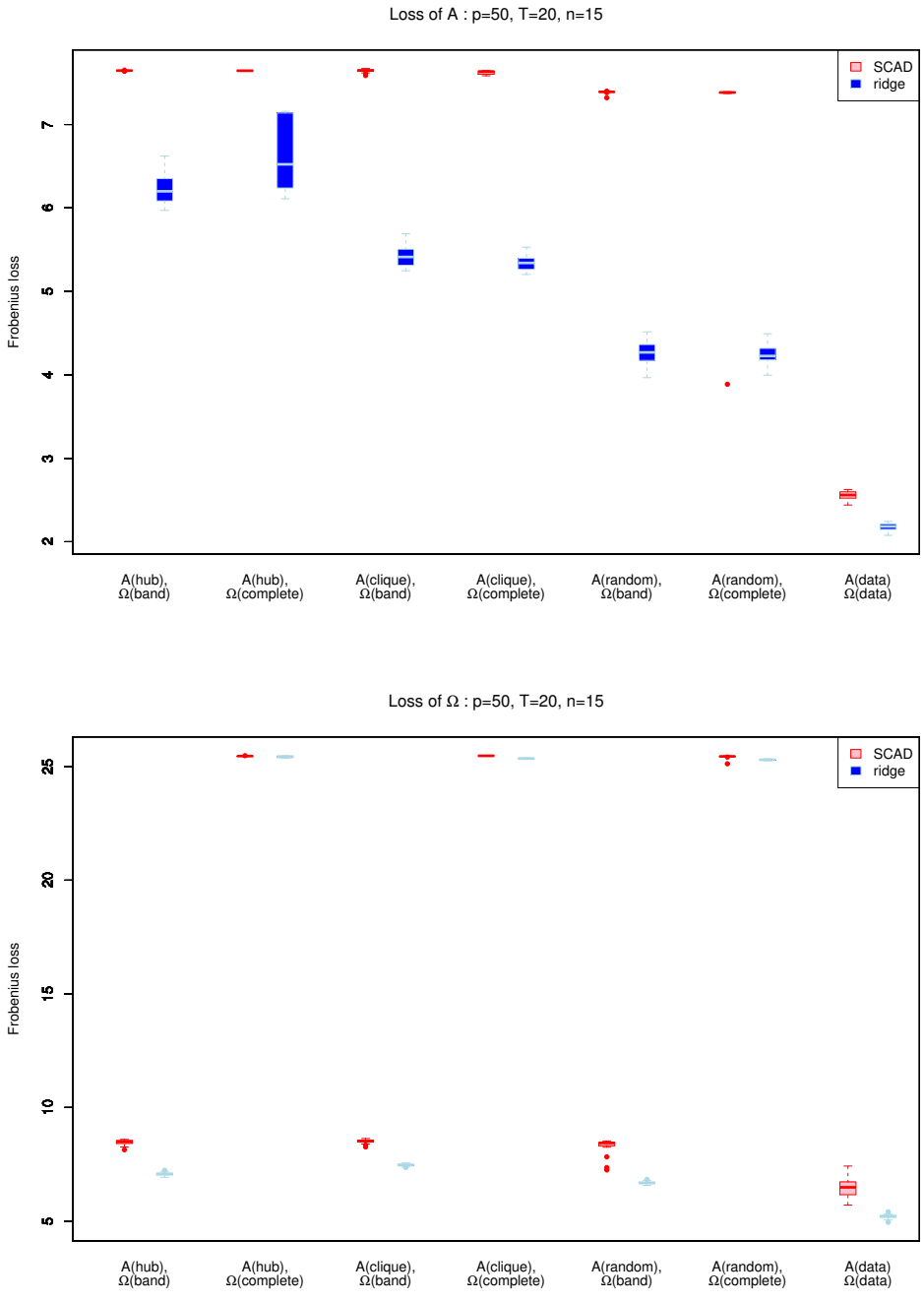

Fig. 3.21 Frobenius loss comparison between SCAD and ridge estimators for precision and autoregressive coefficient matrix on simulated data set where p=50, T=20, n=15 and **A** with roughly 25% nonzero elements.

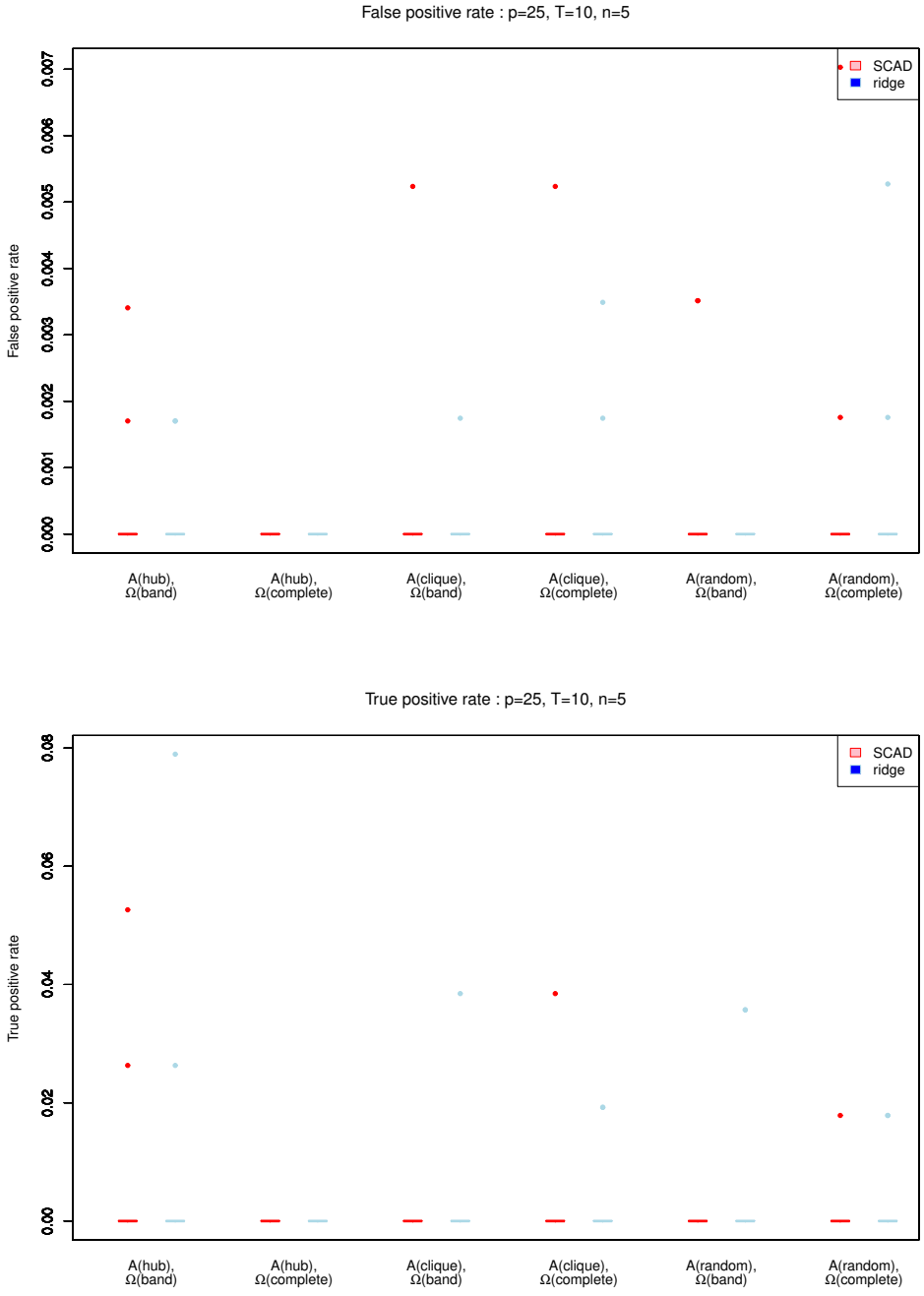

Fig. 3.22 Box plot of specificity (false positive rate) and sensitivity (true positive rate) of the ridge and SCAD methods on simulated data where p=25, T=10, n=5 and $\mathbf{A}$ with roughly 5% nonzero elements.

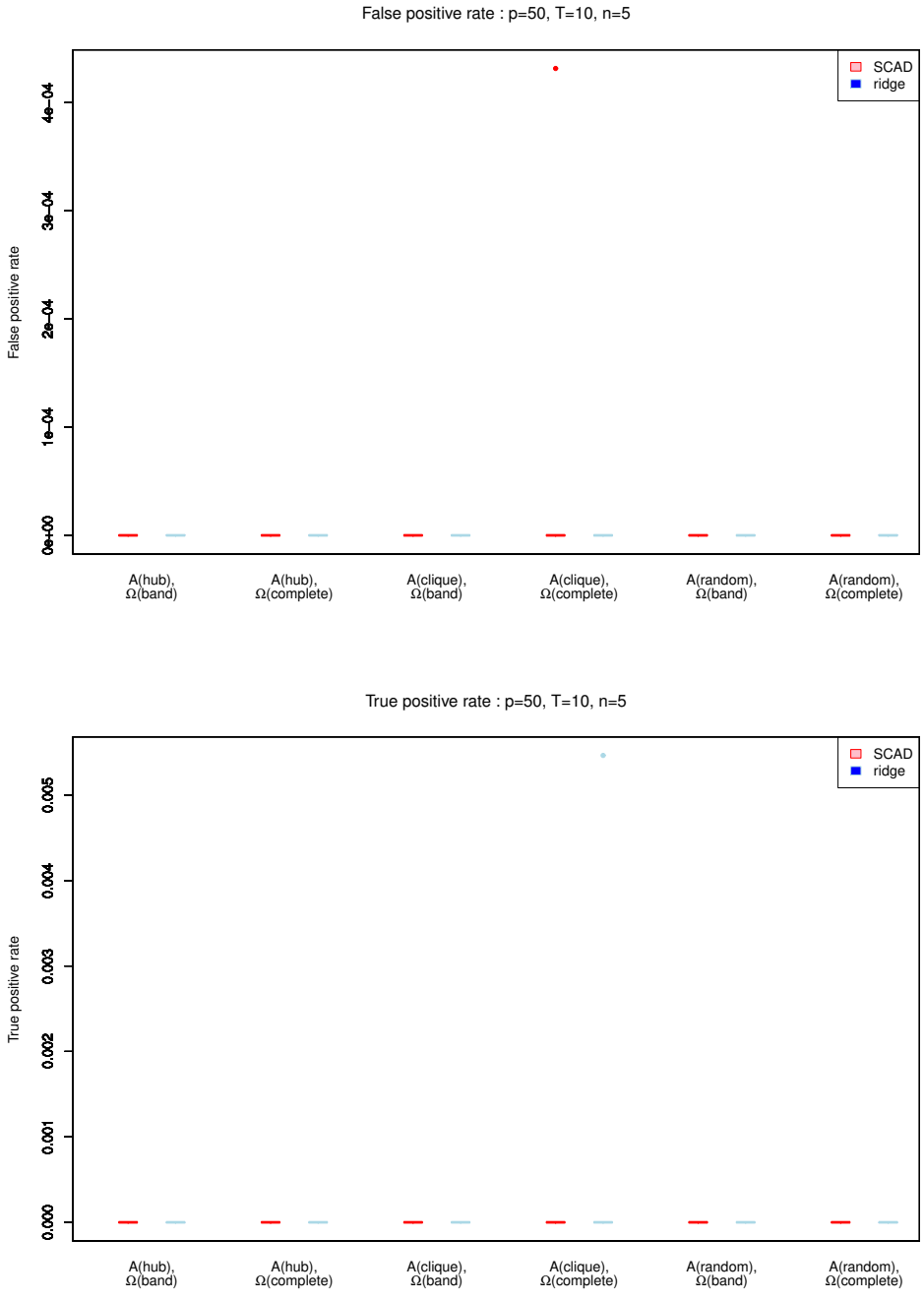

Fig. 3.23 Box plot of specificity (false positive rate) and sensitivity (true positive rate) of the ridge and SCAD methods on simulated data where p=50, T=10, n=5 and **A** with roughly 5% nonzero elements.

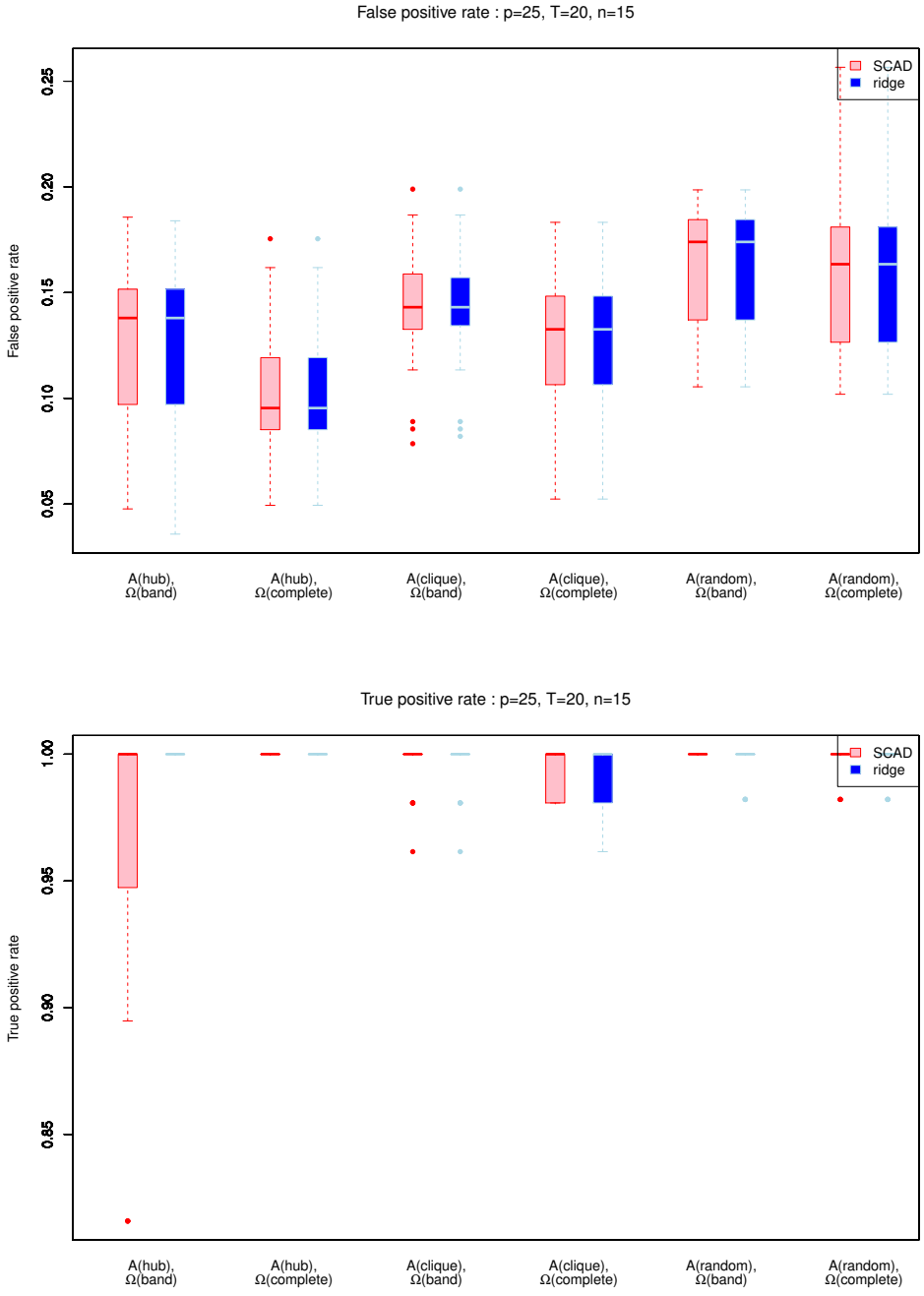

Fig. 3.24 Box plot of specificity (false positive rate) and sensitivity (true positive rate) of the ridge and SCAD methods on simulated data where p=25, T=20, n=15 and $\mathbf{A}$ with roughly 5% nonzero elements.

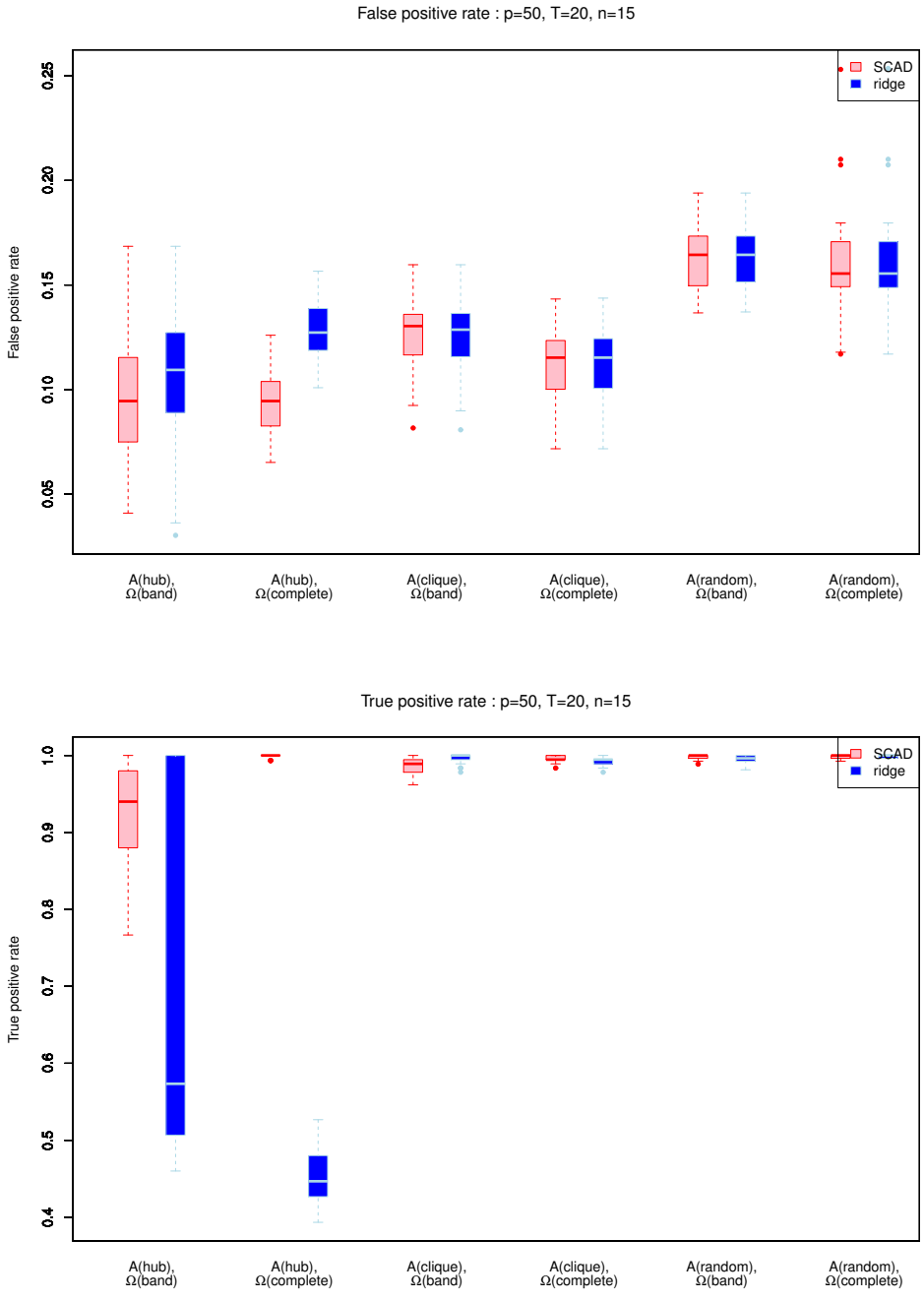

Fig. 3.25 Box plot of specificity (false positive rate) and sensitivity (true positive rate) of the ridge and SCAD methods on simulated data where p=50, T=20, n=15 and **A** with roughly 5% nonzero elements.

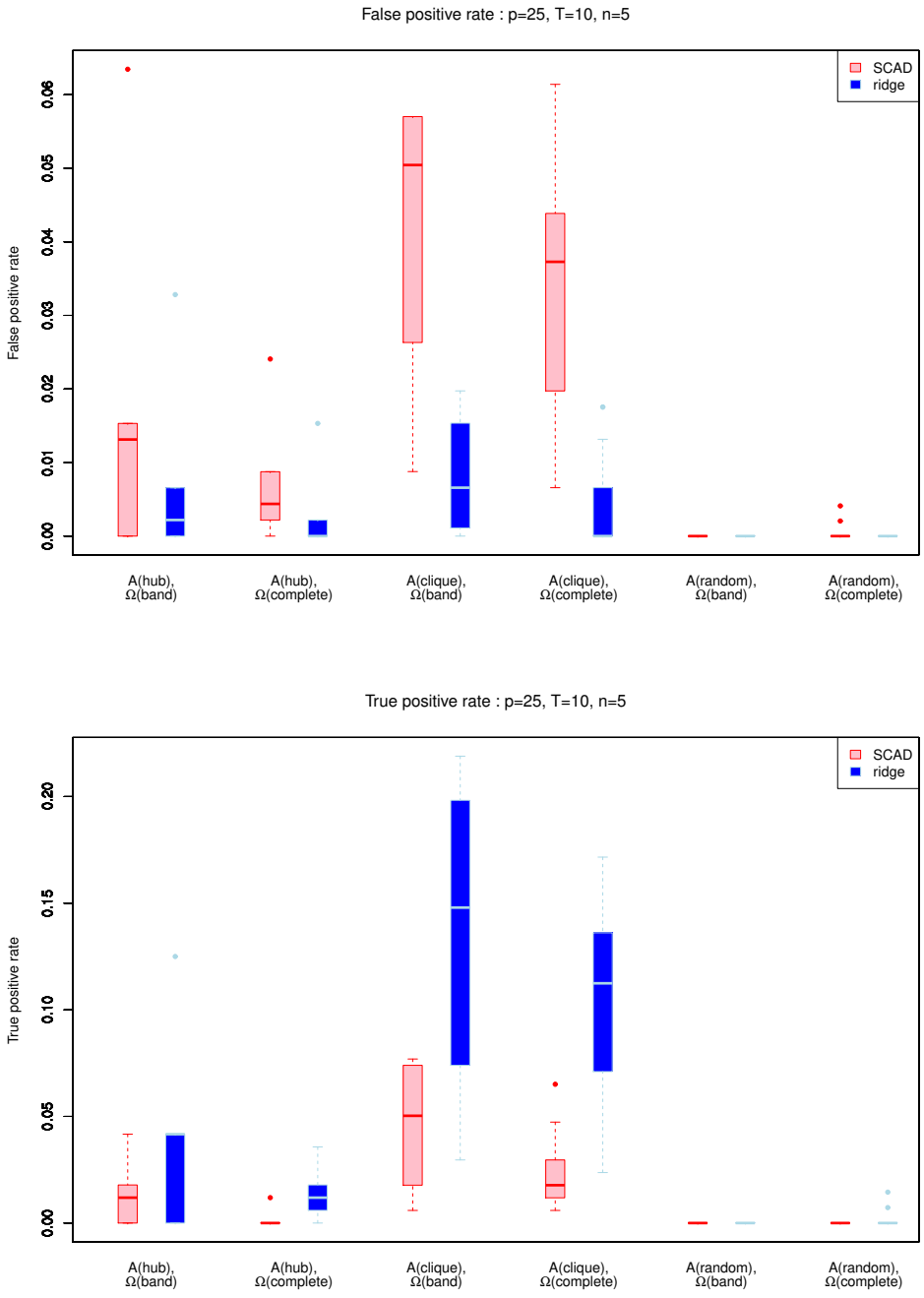

Fig. 3.26 Box plot of specificity (false positive rate) and sensitivity (true positive rate) of the ridge and SCAD methods on simulated data where p=25, T=10, n=5 and $\mathbf{A}$ with roughly 25% nonzero elements.

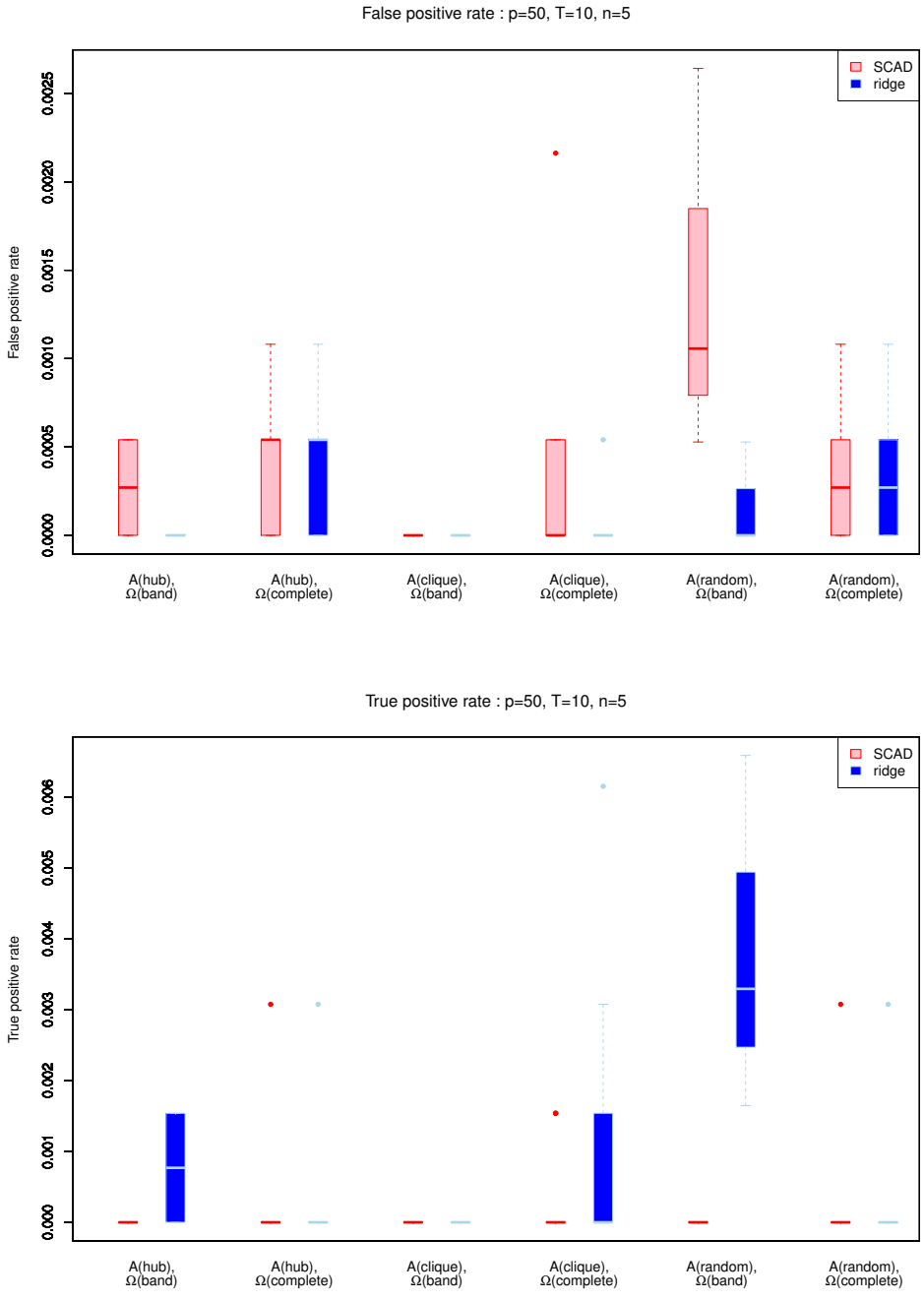

Fig. 3.27 Box plot of specificity (false positive rate) and sensitivity (true positive rate) of the ridge and SCAD methods on simulated data where p=50, T=10, n=5 and **A** with roughly 25% nonzero elements.

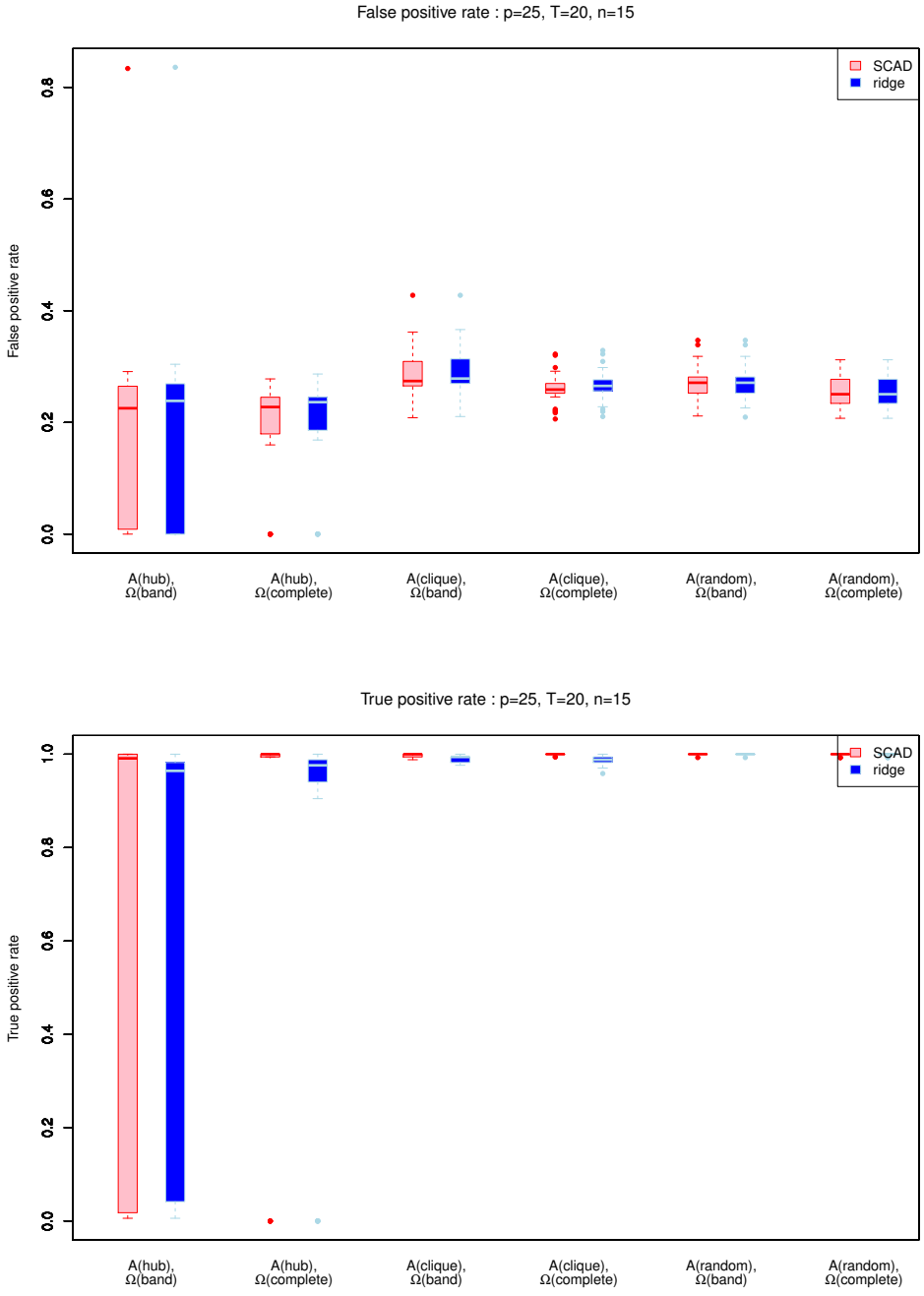

Fig. 3.28 Box plot of specificity (false positive rate) and sensitivity (true positive rate) of the ridge and SCAD methods on simulated data where p=25, T=20, n=15 and $\mathbf{A}$ with roughly 25% nonzero elements.

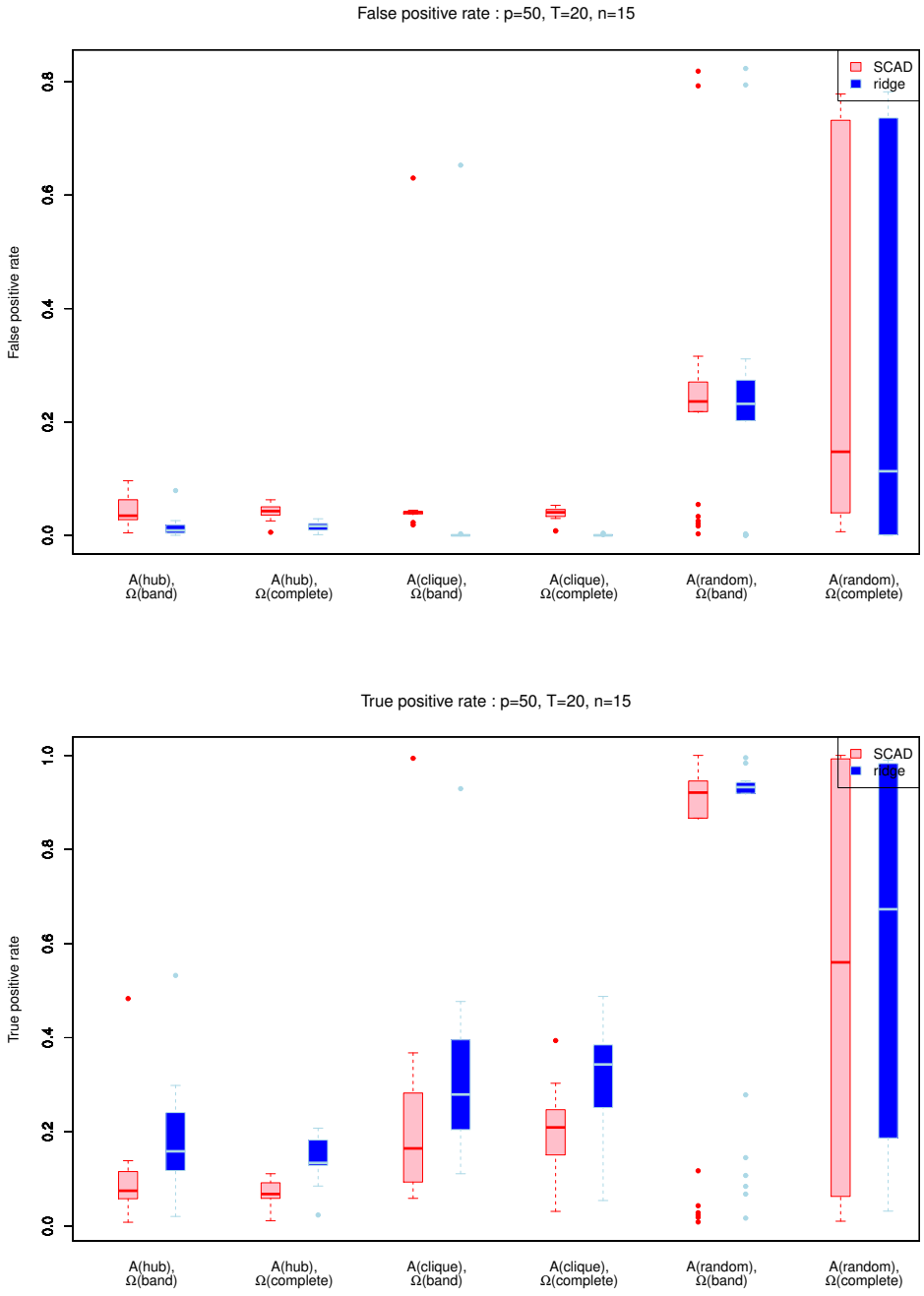

Fig. 3.29 Box plot of specificity (false positive rate) and sensitivity (true positive rate) of the ridge and SCAD methods on simulated data where p=50, T=20, n=15 and $\mathbf{A}$ with roughly 25% nonzero elements.

3.8 Illustration, additional material

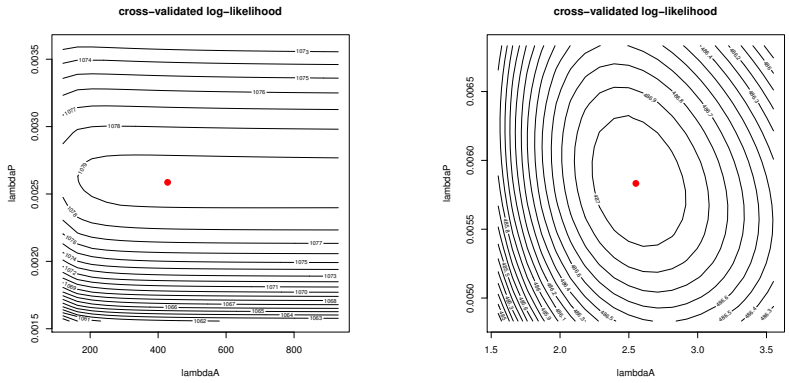

Fig. 3.30 Contour plots of LOOCV log-likelihood of VAR(1) model for the p53 signalling pathway data. Left panel represents the contour plot of the penalty parameters orginal estimation (without knowledge on the support), while in the right panel that with inferred support supplied.

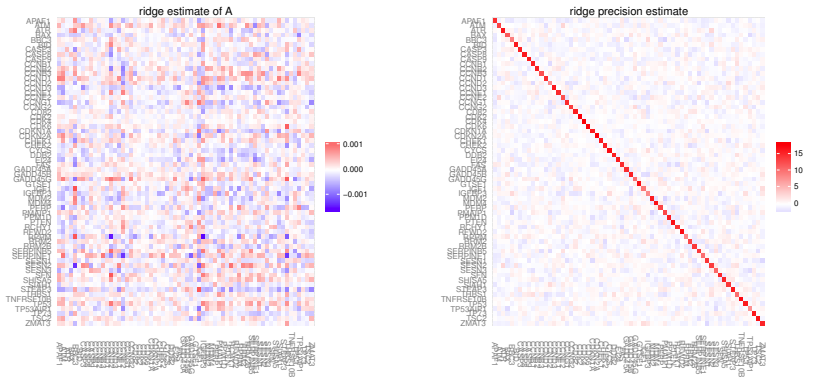

Fig. 3.31 Heatmaps of the ridge estimated parameters for p53 signalling pathway (without knowledge of the support). The left panel contains the estimate of $\mathbf{A}$, while the right panel that of $\mathbf{\Omega}_{\varepsilon}$. The rows and columns correspond to the 64 (alphabetically ordered) genes of the pathway. Each element (i.e. a small square) of the heatmap represents an element of the estimate. The color intensity represents the size of the estimated element.

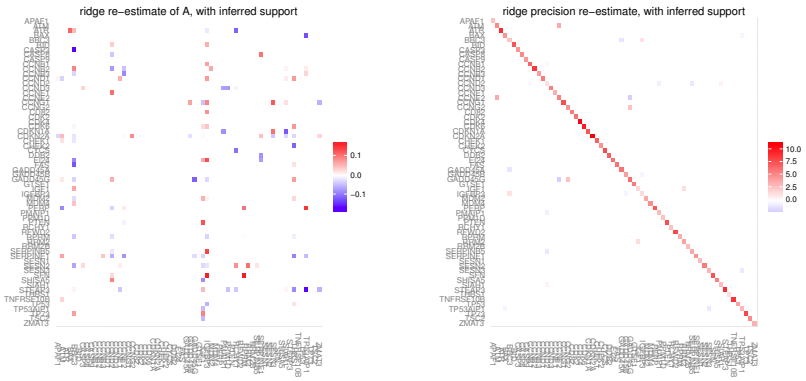

Fig. 3.32 Heatmaps of the ridge estimated parameters for p53 signalling pathway (with inferred support supplied). The left panel contains the estimate of $\mathbf{A}$, while the right panel that of $\mathbf{\Omega}_\varepsilon$. The rows and columns correspond to the 64 (alphabetically ordered) genes of the pathway. Each element (i.e. a small square) of the heatmap represents an element of the estimate. The color intensity represents the size of the estimated element.

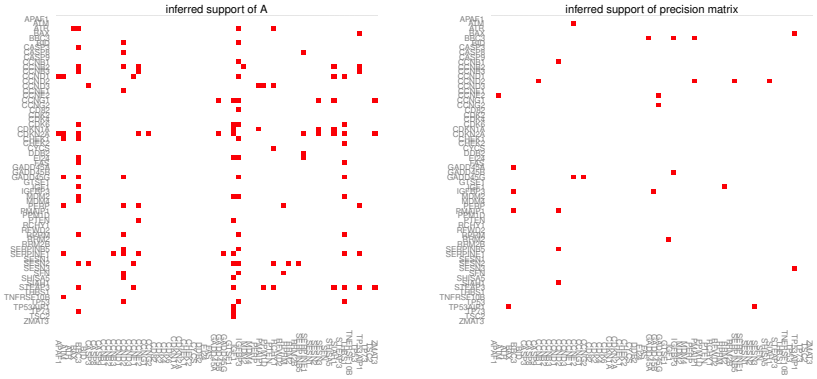

Fig. 3.33 Heatmaps of the support of matrices $\mathbf{A}$ and $\mathbf{\Omega}_\varepsilon$. The rows and columns correspond to the 64 (alphabetically ordered) genes of the pathway. Each element (i.e. a small square) of the heatmap represents an element of the estimate. A red element indicates the presence ('nonzero-ness') of the edge, while white indicates its absence.

	$\deg^-(\mathbf{A})$	$\deg^+(\mathbf{A})$	between.	close.	eigen centr.	mutual info.	impulse resp.
BBC3	0	17	17	0.00192	1.00000	0.05605	0.01497
CCND2	0	12	18	0.00187	0.68783	0.01054	0.00874
IGF1	1	14	0	0.00191	0.97635	0.01472	0.01086
IGFBP3	0	16	7	0.00190	0.87513	0.04498	0.01561
THBS1	0	11	0	0.00188	0.87717	0.01172	0.00965
CCNG1	6	0	0	0.00177	0.25154	0.00000	0.00000
CDKN2A	12	0	0	0.00181	0.49508	0.00000	0.00000
SERPINE1	8	4	0	0.00185	0.70869	0.01796	0.00458
SESN2	8	0	0	0.00180	0.26759	0.00000	0.00000
STEAP3	9	0	0	0.00179	0.36285	0.00000	0.00000

Table 3.2 Node statistics of the inferred time series chain graph of the p53 signalling pathway. The columns contain (from left to right): the gene corresponing to the node; the in-degree of lag-one temporal dependencies; the out-degree of lag-one temporal dependencies; the betweenness in the (global) partial correlation graph; the closeness in the (global) partial correlation graph; the eigen-centrality in the (global) partial correlation graph; the mutual information between the node's expression level at time t and that of all pathway nodes at $t+1$; the impulse response on the node's expression level at time t on that of all pathway nodes at $t+1$.

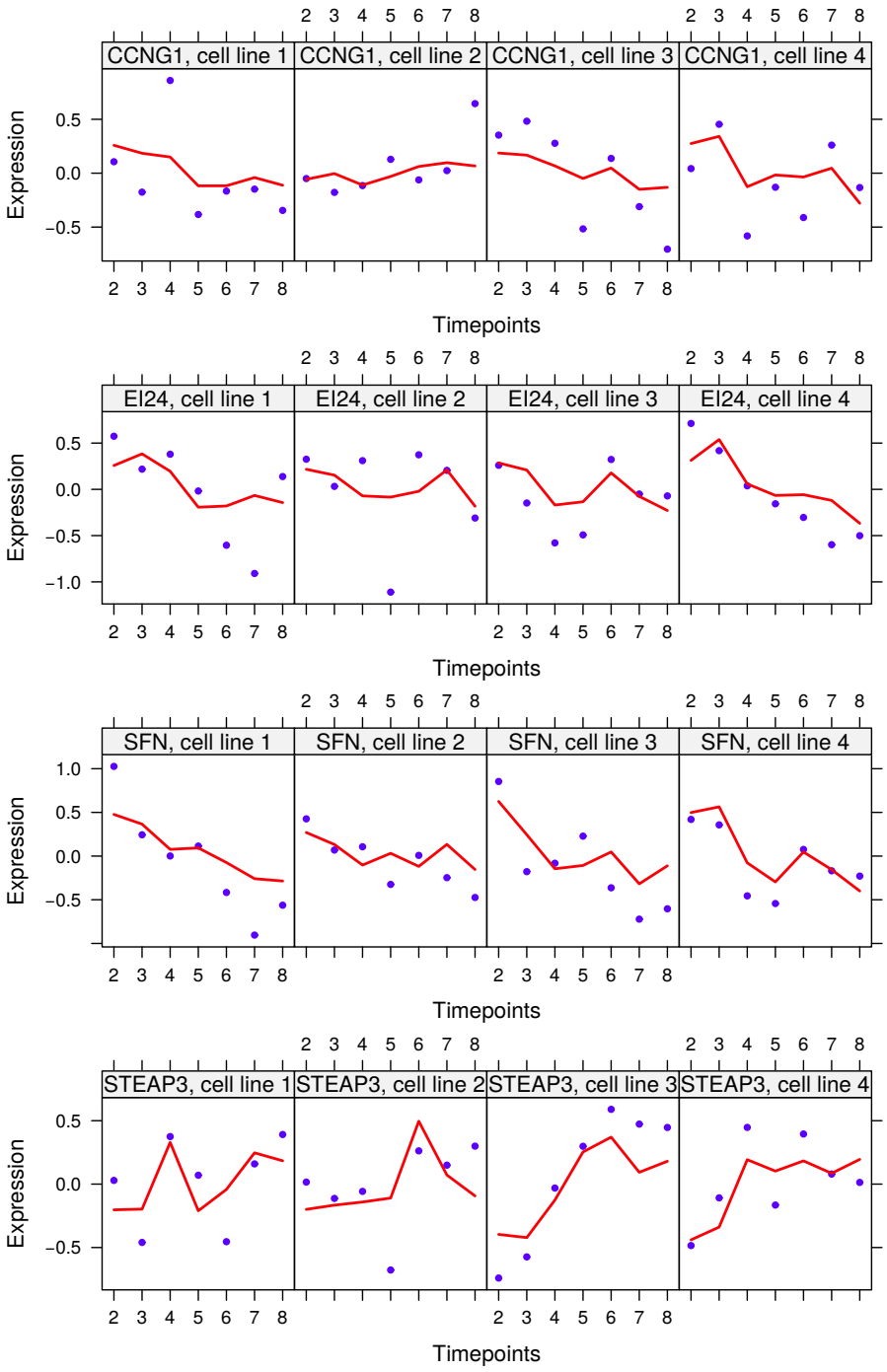

Fig. 3.34 Fit of the expression levels of the CCNG1, EI24, SFN and STEAP3 genes over time in all cell lines. Dots represent the expression levels, while solid red lines indicated the fit of the model over time.

Chapter 4

Ridge estimation of network models from time-course omics data

(***Miok, V., Wilting, S. M. and van Wieringen, W. N., Submitted for publication***)

4.1 The VAR(2) model

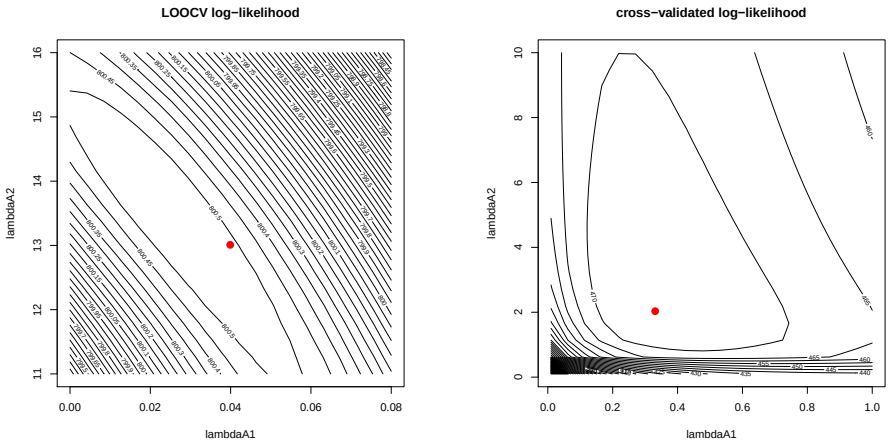

Fig. 4.1 Contourplot of the VAR(2) LOOCV log-likelihood vs. the penalty parameters. Left and right panels show these contour plots of the situation with full and sparsified support, respectively.

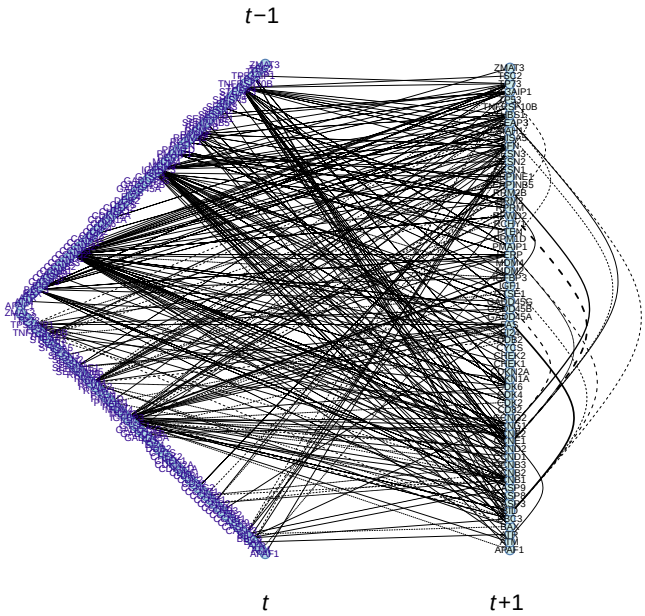

Fig. 4.2 Inferred time-series chain graph from the VAR(2) model. Solid and dashed lines represent a positive and negative relations, respectively. The thickness of the lines indicates the strength of the relation. Unconnected nodes have been pruned from the graph.

The mutual information between $Y_{j,i,t}$ and $\mathbf{Y}_{*,i,t+\tau}$ (given $\mathbf{Y}_{*,i,t-\tau}$ for $\tau \in \mathbb{N}$) is defined as:

$$\begin{aligned}\mathcal{I}(\mathbf{Y}_{*,i,t+\tau}, Y_{j,i,t} | \mathbf{Y}_{*,i,t-1}, \mathbf{Y}_{*,i,t-2}) \;&=\\ \mathcal{H}(\mathbf{Y}_{*,i,t+\tau} \,|\, \mathbf{Y}_{*,i,t-1}, \mathbf{Y}_{*,i,t-2}) &- \mathcal{H}(\mathbf{Y}_{*,i,t+\tau} \,|\, Y_{j,i,t}, \mathbf{Y}_{*,i,t-1}, \mathbf{Y}_{*,i,t-2}),\end{aligned}$$

where $\mathcal{H}(\cdot|\cdot)$ is the (conditional) entropy. Under normality the entropy equals, e.g.:

$$H(\mathbf{Y}_{*,t+\tau,i} \,|\, \mathbf{Y}_{*,t-1,i}, \mathbf{Y}_{*,i,t-2}) \;=\; \log(|\mathbb{V}(\mathbf{Y}_{*,t+\tau,i} \,|\, \mathbf{Y}_{*,t-1,i}, \mathbf{Y}_{*,t-2,i})|).$$

For the VAR(2) model this variance is given by the following recursive relations:

$$\begin{aligned}\mathbb{V}(\mathbf{Y}_{*,t+\tau,i} \,|\, \mathbf{Y}_{*,t-1,i}, \mathbf{Y}_{*,t-2,i}) \;&=\; \boldsymbol{\Sigma}_\varepsilon\\ &+\mathbf{A}_1 \mathbb{V}(\mathbf{Y}_{*,t+\tau-1,i} \,|\, \mathbf{Y}_{*,t-1,i}, \mathbf{Y}_{*,t-2,i}) \mathbf{A}_1^\top\\ &+\mathbf{A}_2 \mathbb{V}(\mathbf{Y}_{*,t+\tau-2,i} \,|\, \mathbf{Y}_{*,t-1,i}, \mathbf{Y}_{*,t-2,i}) \mathbf{A}_2^\top\end{aligned}$$

$$
\begin{aligned}
&+\mathbf{A}_1\mathrm{Cov}(\mathbf{Y}_{*,t+\tau-1,i}, \mathbf{Y}_{*,t+\tau-2,i} \,|\, \mathbf{Y}_{*,t-1,i}, \mathbf{Y}_{*,t-2,i})\mathbf{A}_2^\top \\
&+\mathbf{A}_2\mathrm{Cov}(\mathbf{Y}_{*,t+\tau-2,i}, \mathbf{Y}_{*,t+\tau-1,i} \,|\, \mathbf{Y}_{*,t-1,i}, \mathbf{Y}_{*,t-2,i})\mathbf{A}_1^\top, \\
\mathrm{Cov}(\mathbf{Y}_{*,t+\tau,i}, \mathbf{Y}_{*,t+\tau-1,i} \,|\, \mathbf{Y}_{*,t-1,i}, \mathbf{Y}_{*,t-2,i}) \;&= \\
&\mathbf{A}_1\mathbb{V}(\mathbf{Y}_{*,t+\tau-1,i} \,|\, \mathbf{Y}_{*,t-1,i}, \mathbf{Y}_{*,t-2,i}) \\
&+\mathbf{A}_2\mathbb{V}(\mathbf{Y}_{*,t+\tau-1,i} \,|\, \mathbf{Y}_{*,t-1,i}, \mathbf{Y}_{*,t-2,i})\mathbf{A}_1^\top \\
&+\mathbf{A}_2\mathrm{Cov}(\mathbf{Y}_{*,t+\tau-2,i}, \mathbf{Y}_{*,t+\tau-3,i} \,|\, \mathbf{Y}_{*,t-1,i}, \mathbf{Y}_{*,t-2,i})\mathbf{A}_2^\top.
\end{aligned}
$$

The following initiations complete these recursive relations:

$$
\begin{aligned}
\mathbb{V}(\mathbf{Y}_{*,t,i} \,|\, \mathbf{Y}_{*,t-1,i}, \mathbf{Y}_{*,t-2,i}) &= \boldsymbol{\Sigma}_\varepsilon \\
\mathbb{V}(\mathbf{Y}_{*,t+1,i} \,|\, \mathbf{Y}_{*,t-1,i}, \mathbf{Y}_{*,t-2,i}) &= \boldsymbol{\Sigma}_\varepsilon + \mathbf{A}_1\boldsymbol{\Sigma}_\varepsilon\mathbf{A}_1^\top \\
\mathrm{Cov}(\mathbf{Y}_{*,t,i}, \mathbf{Y}_{*,t-1,i} \,|\, \mathbf{Y}_{*,t-1,i}, \mathbf{Y}_{*,t-2,i}) &= \mathbf{0}_{pp}, \\
\mathrm{Cov}(\mathbf{Y}_{*,t+1,i}, \mathbf{Y}_{*,t,i} \,|\, \mathbf{Y}_{*,t-1,i}, \mathbf{Y}_{*,t-2,i}) &= \mathbf{A}_1\boldsymbol{\Sigma}_\varepsilon, \\
\mathrm{Cov}(\mathbf{Y}_{*,t+2,i}, \mathbf{Y}_{*,t+1,i} \,|\, \mathbf{Y}_{*,t-1,i}, \mathbf{Y}_{*,t-2,i}) &= \mathbf{A}_2\boldsymbol{\Sigma}_\varepsilon\mathbf{A}_1^\top + \mathbf{A}_1\boldsymbol{\Sigma}_\varepsilon + \mathbf{A}_1^2\boldsymbol{\Sigma}_\varepsilon\mathbf{A}_1^\top.
\end{aligned}
$$

Rests that of the other entropy term, the evaluation of $\mathbb{V}(\mathbf{Y}_{*,t+\tau,i} \,|\, Y_{j,t,i}, \mathbf{Y}_{*,t-1,i}, \mathbf{Y}_{*,t-2,i})$ employs the same recursive relations (also conditioning on $Y_{j,t,i}$ leaves them unaffected for $\tau > 2$). The initiations however change:

$$
\begin{aligned}
\mathbb{V}(\mathbf{Y}_{*,t,i} \,|\, Y_{j,t,i}, \mathbf{Y}_{*,t-1,i}, \mathbf{Y}_{*,t-2,i}) &= \boldsymbol{\Sigma}_{\varepsilon|j} \\
\mathbb{V}(\mathbf{Y}_{*,t+1,i} \,|\, Y_{j,t,i}, \mathbf{Y}_{*,t-1,i}, \mathbf{Y}_{*,t-2,i}) &= \boldsymbol{\Sigma}_\varepsilon + \mathbf{A}_1\boldsymbol{\Sigma}_{\varepsilon|j}\mathbf{A}_1^\top \\
\mathbb{V}(\mathbf{Y}_{*,t+2,i} \,|\, Y_{j,t,i}, \mathbf{Y}_{*,t-1,i}, \mathbf{Y}_{*,t-2,i}) &= \boldsymbol{\Sigma}_\varepsilon + \mathbf{A}_1\boldsymbol{\Sigma}_\varepsilon\mathbf{A}_1^\top + \mathbf{A}_2\boldsymbol{\Sigma}_{\varepsilon|j}\mathbf{A}_2^\top + \mathbf{A}_1^2\boldsymbol{\Sigma}_{\varepsilon|j}(\mathbf{A}_1^\top)^2 \\
\mathrm{Cov}(\mathbf{Y}_{*,t,i}, \mathbf{Y}_{*,t-1,i} \,|\, Y_{j,t,i}, \mathbf{Y}_{*,t-1,i}, \mathbf{Y}_{*,t-2,i}) &= \mathbf{0}_{pp}, \\
\mathrm{Cov}(\mathbf{Y}_{*,t+1,i}, \mathbf{Y}_{*,t,i} \,|\, Y_{j,t,i}, \mathbf{Y}_{*,t-1,i}, \mathbf{Y}_{*,t-2,i}) &= \mathbf{A}_1\boldsymbol{\Sigma}_{\varepsilon|j}, \\
\mathrm{Cov}(\mathbf{Y}_{*,t+2,i}, \mathbf{Y}_{*,t+1,i} \,|\, Y_{j,t,i}, \mathbf{Y}_{*,t-1,i}, \mathbf{Y}_{*,t-2,i}) &= \mathbf{A}_2\boldsymbol{\Sigma}_{\varepsilon|j}\mathbf{A}_1^\top + \mathbf{A}_1\boldsymbol{\Sigma}_\varepsilon + \mathbf{A}_1^2\boldsymbol{\Sigma}_{\varepsilon|j}\mathbf{A}_1^\top,
\end{aligned}
$$

in which $\boldsymbol{\Sigma}_{\varepsilon|j} = (\boldsymbol{\Sigma}_\varepsilon)_{\setminus j,\setminus j} - (\boldsymbol{\Sigma}_\varepsilon)_{\setminus j,j}[(\boldsymbol{\Sigma}_\varepsilon)_{j,j}]^{-1}(\boldsymbol{\Sigma}_\varepsilon)_{j,\setminus j}$, the covariance of the all but the j-th innovations conditional in the j-th one.

Table 4.1 Node statistics for the top five 'regulators' and 'regulatees' (as derived from the time-series chain graph of the p53 signaling pathway) identified from the HPV16 affected cell line data using the VAR(2) model. For these genes, names in the left most column, the following node statistics are evaluated: the in- and out-degree of the lag-one temporal dependencies; the in- and out-degree of the lag-two temporal dependencies; the mutual information between the expression level of the gene at time t and all other genes at time $t+2$ from the pathway and impulse response on the the expression level of the gene at time t and all other genes at time $t+2$.

	$\deg^-(\mathbf{A}_1)$	$\deg^+(\mathbf{A}_1)$	$\deg^-(\mathbf{A}_2)$	$\deg^+(\mathbf{A}_2)$	mutual info.	impulse resp.
IGFBP3	2	18	5	27	0.54358	0.02826
IGF1	1	18	0	10	0.05431	0.01213
RPRM	4	17	2	10	0.19450	0.00799
CCND2	2	4	5	32	0.05297	0.01623
THBS1	3	8	6	20	0.17556	0.01143
SFN	8	0	5	0	0.00000	0.00000
CCNB1	8	0	4	0	0.00000	0.00000
TP73	3	0	8	0	0.00000	0.00000
SESN2	3	0	4	0	0.00000	0.00000
DDB2	2	0	4	0	0.00000	0.00000

4.2 Multiple VAR(1) models

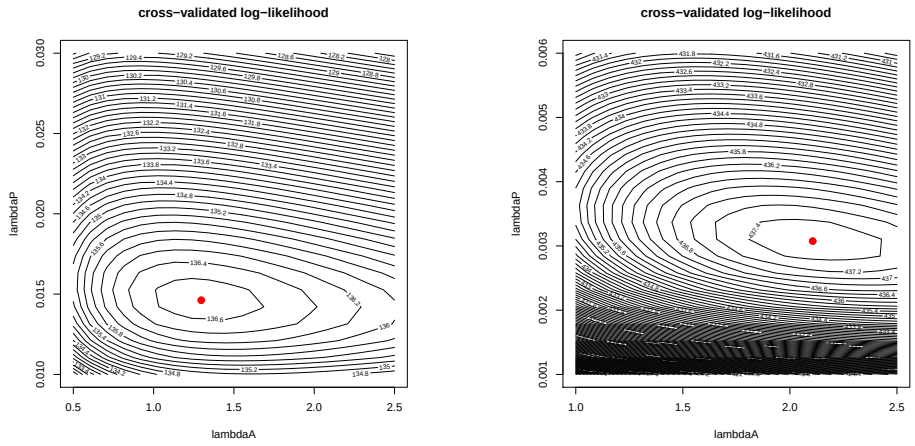

Fig. 4.3 Contourplot of the multiple VAR(1) LOOCV log-likelihood vs. the penalty parameters. Left and right panels show these contour plots of the data from cell lines affected with HPV16 and HPV18, respectively.

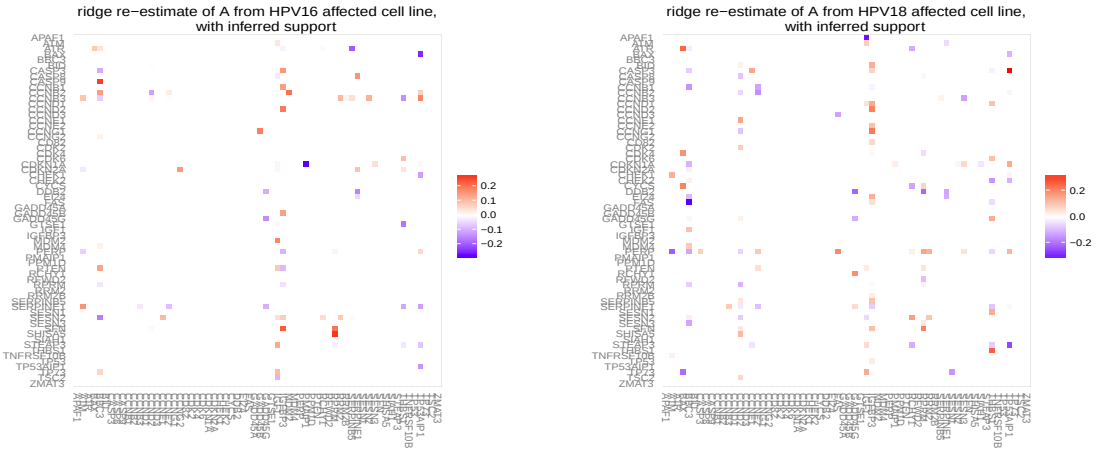

Fig. 4.4 Heatmaps of the sparsified re-estimated multiple VAR(1) model parameters. Left and right panel represent temporal interactions among mRNAs in the cell lines affected with HPV16 and HPV18, respectively.

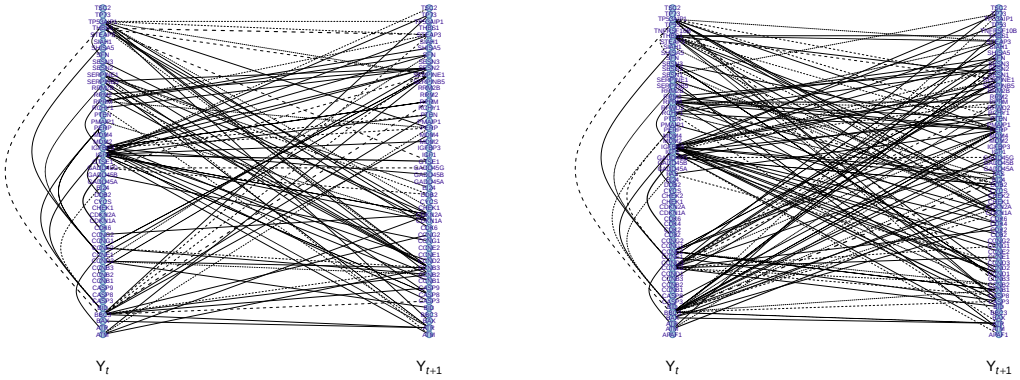

Fig. 4.5 Left and right panels depict the inferred time-series chain graph from the cell lines affected with HPV16 and HPV18, respectively. Solid and dashed lines represent positive and negative relations, respectively. The thickness of the lines corresponds to the strength of the relation. Unconnected nodes have been pruned from the graph.

4.3 The VARX(1) model

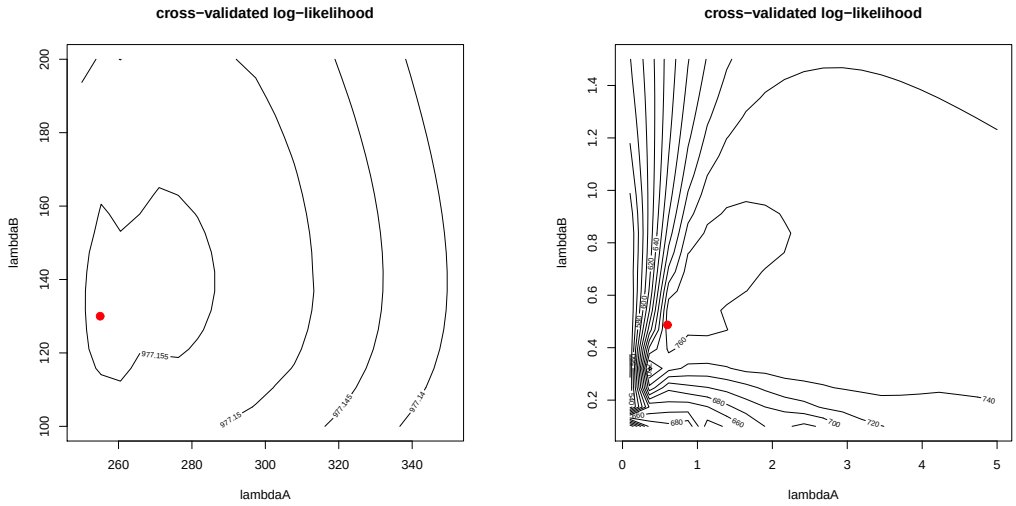

Fig. 4.6 Contour plots of LOOCV log-likelihood vs. penalty parameters of the VARX(1) model. Left and right panels display the contour plots without and with inferred support, respectively.

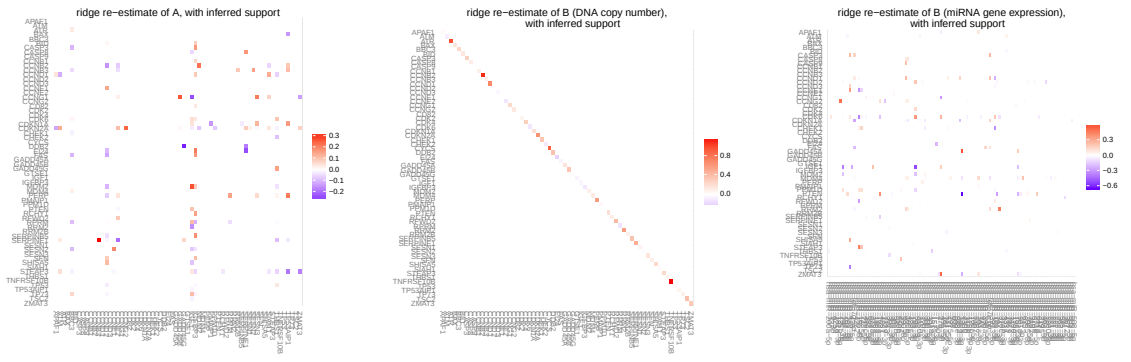

Fig. 4.7 Heatmaps of the sparse re-estimated VARX(1) model parameters. Left panel: Estimates of $\mathbf{A}$, representing the temporal relations among mRNAs. Central panel and right panel show the estimate of $\mathbf{B}$, partitioned by molecular level (DNA copy number and miRNA, respectively).

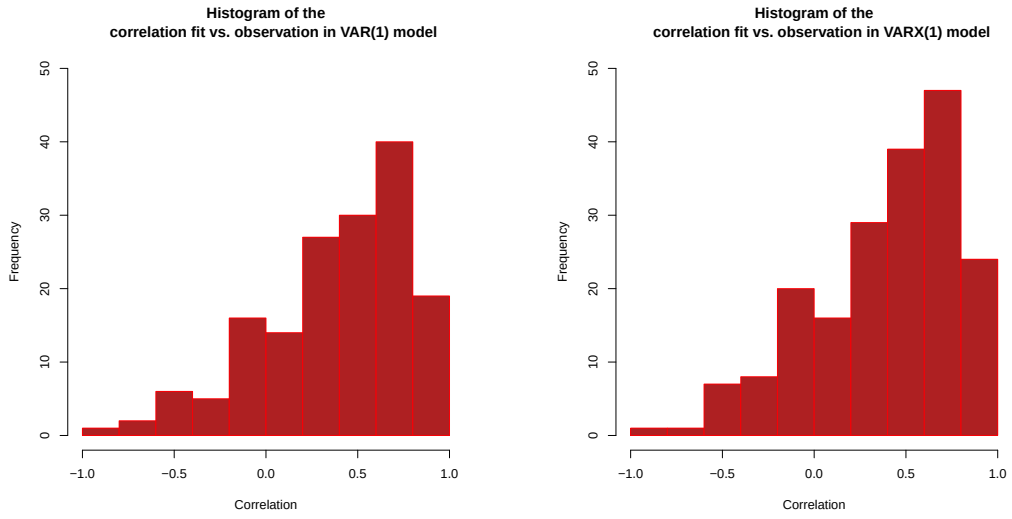

Fig. 4.8 Histograms of Spearman correlations between the observations and VAR(1) and VARX(1) model fits (left and right panel, respectively).

Chapter 5

Comprehensive molecular profiling of HPV-induced transformation over time

(***Miok, V., Babion, I., Jaspers, A., Meijer, C. J. L. M., Snijders, P. J. F. and Steenbergen, R. D. M. , van Wieringen, W. N., Wilting, S. M., In preparation***)

5.1 Un published material

Due to the fact that paper is still not published the Supplementary Materials from this chapter cannot be published online.

Chapter 6

Aberrant methylation-mediated silencing of microRNAs contributes to HPV-induced anchorage independence

(*Wilting, S. M., Miok, V., Jaspers, A., Boon, D., Sorgad, H., Lando, M., Snoek, B. C., van Wieringen, W. N., Meijer, C. J. L. M., Lyng H., Snijders, P. J. F. and Steenbergen, R. D. M., Oncotarget (2016), 7(28): 43805-43819*)

6.1 Supplementary Tables

nr	miRNA gene (CpG island associated)	cell line	mature miRNA
1	hsa-mir-1237	FK16A	hsa-miR-1237
2	hsa-miR-1292	FK16A	hsa-miR-1292
3	hsa-mir-1538	FK16B	hsa-miR-1538
4	hsa-mir-2861	FK16B	hsa-miR-2861
5	hsa-mir-3178	FK16A	hsa-miR-3178
6	hsa-mir-3180-1/-3	FK16A	hsa-miR-3180-5p
7	hsa-mir-3613	FK18A	hsa-miR-3613-3p
8	hsa-mir-3615	FK16B	hsa-miR-3615
9	hsa-mir-3652	FK16A	hsa-miR-3652
10	hsa-mir-564	FK18B	hsa-miR-564
11	hsa-mir-572	FK16B	hsa-miR-572
12	hsa-mir-632	FK18A	hsa-miR-632
13	hsa-mir-638	FK16B	hsa-miR-638
14	hsa-mir-1908	FK16B	hsa-miR-1908
15	hsa-mir-611	FK16A	hsa-miR-611
16	hsa-miR-943	FK16B	hsa-miR-943

Table 6.1 Supplementary Table S1: miRNA genes potentially silenced by methylation that were only identified in 1 out of 4 cell lines investigated

Table 6.2 Sequences for bisulfite sequencing primers. Size (bp) refers to the amplicon size in basepairs.

Gene	**Oligo**	**Sequence**	**Tm (°C)**	**size (bp)**
mir-129-2	Forward	GAGGGAGGAGTTTTATTTAGTTG	52.3	318
	Reverse	TCACCCACCAATCAAAAAA	53.4	
	Forward	TAATATGTTTTAGTTTAGGTTTTGTTG	52	345
	Reverse	ACCTATCCTATCCCTCTATCTCC	52.5	
	Forward	GGAGATAGAGGGATAGGATAGGT	52.5	474
	Reverse	ACTATTAAATTATATACAACAAACCCAAAC	54.5	
mir-137	Forward	GTTATTTGGATTTGGGTAGGAAGTAG	56.7	226
	Reverse	AAAAAAATCAAAAAACCAAACTACC	55.3	
mir-615	Forward	TGTTTATTAAAAGGTTTAGAGGTTGTGT	56.3	336
	Reverse	TAAAAAAAAAAAAAAAAACCCAAACTC	56.6	
	Forward	AGAGGGATTTGAAGGGTGAG	54.4	271
	Reverse	CCCAAACAAAAAACTTTCCAC	55	
mir-675	Forward	TAATTTATTTAGTAGTAGGTATAGGGGTAGT	52.8	334
	Reverse	AACTCCCCTAAATACTATACTATCTACC	52.1	
mir-935	Forward	TTGGGTTAGTTAGATTTAGG	45.5	260
	Reverse	AATAATAATTTTTCTTCTCCTC	45.4	
	Forward	GAGGAGAAGAAAAATTATTATT	45.4	158
	Reverse	CCAACCTTAAACAAATCC	46.5	
mir-2277	Forward	TTGAGTTAGGGGGTAAGTGAT	51.8	344
	Reverse	ACTTCAAAAAACTAATCTTAACCTCTA	52.2	
	Forward	GAGTTAATTTTTTGGAAGAGTTTTT	52.9	317
	Reverse	CCCATATTACTTTAACAATTACACTC	51.8	
mir-3663	Forward	TGTAAGGTTTAGTATTTAGTGGTTATTTATTA	54.1	375
	Reverse	AAAACCCTCCTCCCCTTC	54.1	
mir-3665	Forward	GGGGTTAGGTAGGTAGAAGGTGT	55.3	282
	Reverse	(G)AAACTAAAAAAACAAAAC(G/A)AACAACTA	55.3/53.6	
mir-4281	Forward	TTTAGGTTTTAAGGACGTTAGGA	53.3	381
	Reverse	AAAAACCCCCAACCCTAC	52.5	
	Forward	GGGTTTAGGGTTTAGAAGGG	53.3	210
	Reverse	ATAACCCCCTACTTAATTTTCTC	51.1	
	Forward	GGGGTTGTAGTTTGTAGGGTT	53.3	308
	Reverse	CCCTCCTTAACCATAAACAAAC	53.3	
mir-4323	Forward	GGATTTTTTTTGTGAAATGGTT	53.8	328
	Reverse	TCCCATATACTAATACCTCCTCTACTC	53.9	
	Forward	GTTTTTGTGTTTATGGGTTTATGT	53.2	288
	Reverse	CCAAACACCCCACATACC	52.8	

Table 6.3 Sequences for (q)MSP primers and probes. Size (bp) refers to the amplicon size in basepairs.

Gene	Oligo	Sequence	Tm (°C)	size (bp)
mir-129-2	Forward	GCGGAGTGGTGAGATTGAGTC	58.3	120
	Reverse	AAAATATACCGACTTCTTCGATTCG	58	
	Probe	DFO-CCTAAAACCGAACAAACTAAATCTCCCCAACG-BHQ2	69.2	
mir-137	Forward	CGGGTTTAGCGAGTAGTAAGAGTTTTG	60.7	115
	Reverse	GAAAAAAATCAAAAAACCAAACTACCG	60.3	
	Probe	JOE- GCTACTACTACCGCCGCCGCCG -BHQ1	69.3	
mir-615	Forward	TTTTTTGCGGAGTCGGTTTC	58.8	108
	Reverse	CAATAAACACCCTCGAAATCCG	59.3	
mir-935	Forward	GAGGTGATAGGCGTGTTGGTC	58.1	88
	Reverse	CAACCTTAAACAAATCCGAACG	57.4	
	Probe	DFO- GCCTCGCGACTACGCTCGATATAAATATTAAC -BHQ2	66.6	
mir-3663	Forward	GCGGGGAGGGGTTGTTC	59.8	110
	Reverse	AAAAAAACCAATTAAAAAATCACAATCG	59.8	
	Probe	FAM-CGAAAAAACAATAAAAAACGAAAAACACGAAACGA-BHQ1	59.8	
mir-3665	Forward	GAGTTATCGTCGTTGTTATTATCGTTGTC	60.3	107
	Reverse	CCCCGACCGCCACG	60.1	
	Probe	JOE- CGACCTCAAAAAACCTAAAACTCGAACTAACGCT -BHQ1	69.1	
mir-4281	Forward	GTTTTTTTTAGGTCGTTAGGATGGAC	58.6	115
	Reverse	TTCTCCGCCGCCTCG	59.1	
	Probe	5?-FAM-ATAACCCCCTACTTAATTTTCTCCGCGACTACC-BHQ1-3'	68.1	

Table 6.4 Sequences for RT-PCR primers.

Gene	Oligo	Sequence	Tm (°C)
mir-3665	looped RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACCGCCGC	
	Forward	GTGAGCAAGCAGGTGCGG	60.1
	Reverse	GTGCAGGGTCCGAGGT	53.4
mir-4281	looped RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACCCCCCC	60.7
	Forward	TAGCTAGGGTCCCGGGGA	59
	Reverse	GTGCAGGGTCCGAGGT	53.4

References

[1] Abegaz, F. and Wit, E. (2013). Sparse time series chain graphical models for reconstructing genetic networks. *Biostatistics*, 1:1–14.

[2] Abramovich, F. and Angelini, C. (2006). Bayesian Maximum a posteriori Multiple Testing Procedure. *The Indian Journal of Statistics*, 68:436–460.

[3] Agarwal, V., Bell, G. W., Nam, J. W., and Bartel, D. P. (2015). Predicting effective microRNA target sites in mammalian mRNAs. *eLife*, 4:e05005.

[4] Alix-Panabières, C. and Pantel, K. (2013). Circulating tumor cells: liquid biopsy of cancer. *Clinical Chemistry*, 59(1):110–118.

[5] Allen, G. I. and Liu, Z. (2013). A local Poisson graphical model for inferring networks from sequencing data. *IEEE Transactions on NanoBioscience*, 12(3):189–198.

[6] Alon, U. (2003). Biological networks: the tinkerer as an engineer. *Science*, 301:1866–1867.

[7] Alon, U. (2007). Network motifs: theory and experimental approaches. *Nature Reviews Genetics*, 8(6):450–461.

[8] Angelini, C., De Canditiis, D., Mutarelli, M., and Pensky, M. (2007). A Bayesian approach to estimation and testing in time-course microarray experiments. *Statistical Applications in Genetics and Molecular Biology*, 6:1–30.

[9] Anwar, S. L., Albat, C., Krech, T., Hasemeier, B., Schipper, E., Schweitzer, N., Vogel, A., Kreipe, H., and Lehmann, U. (2013). Concordant hypermethylation of intergenic microRNA genes in human hepatocellular carcinoma as new diagnostic and prognostic marker. *International Journal of Cancer*, 133(3):660–670.

[10] Balaguer, F., Link, A., Lozano, J. J., Cuatrecasas, M., Nagasaka, T., Boland, C. R., and Goel, A. (2010). Epigenetic silencing of miR-137 is an early event in colorectal carcinogenesis. *Cancer Research*, 70(16):6609–6618.

[11] Band, V., Zajchowski, D., Kulesa, V., and Sager, R. (1990). Human papilloma virus DNAs immortalize normal human mammary epithelial cells and reduce their growth factor requirements. *Proceedings of the National Academy of Sciences*, 87(1):463–467.

[12] Bandres, E., Agirre, X., Bitarte, N., Ramirez, N., Zarate, R., Roman-Gomez, J., Prosper, F., and Garcia-Foncillas, J. (2009). Epigenetic regulation of microRNA

expression in colorectal cancer. *International Journal of Cancer*, 125(11):2737–2743.

[13] Bar-Joseph, Z., Gerber, G. K., Lee, T. I., Rinaldi, N. J., Yoo, J. Y., Robert, F., Gordon, D. B., Fraenkel, E., Jaakkola, T. S., Young, R. A., and Gifford, D. K. (2003). Computational discovery of gene modules and regulatory networks. *Nature Biotechnology*, 21:1337–1324.

[14] Baxter, R. C. (2003). Insulin-like growth factor (IGF)-binding proteins: interactions with IGFs and intrinsic bioactivities. *American Journal of Physiology-Endocrinology and Metabolism*, 278:967–976.

[15] Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57:289–300.

[16] Bickel, P. J. and Doksum, K. A. (2001). *Mathematical Statistics*. Prentice Hall, New Jersey.

[17] Bier, A., Giladi, N., Kronfeld, N., Lee, H. K., Cazacu, S., Finniss, S., Xiang, C., Poisson, L., et al. (2013). MicroRNA-137 is downregulated in glioblastoma and inhibits the stemness of glioma stem cells by targeting RTVP-1. *Oncotarget*, 4(5):665.

[18] Bierkens, M., Krijgsman, O., Wilting, S. M., Bosch, L., Jaspers, A., Meijer, G. A., Meijer, C. J. L. M., Snijders, P. J. F., Ylstra, B., and Steenbergen, R. D. M. (2013). Focal aberrations indicate EYA2 and hsa-miR-375 as oncogene and tumor suppressor in cervical carcinogenesis. *Genes, Chromosomes and Cancer*, 52:56–68.

[19] Bilgrau, A. E., Peeters, C. F. W., Eriksen, P. S., Bøgsted, M., and Van Wieringen, W. N. (2015). Targeted fused ridge estimation of inverse covariance matrices from multiple high-dimensional data classes. *http://arxiv.org/pdf/1509.07982.pdf*, 1:1.

[20] Boldstad, B. M., Irizarry, R. A., Astrand, M., and Speed, T. P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on bias and variance. *Bioinformatics*, 19(2):185–193.

[21] Botezatu, A., Goia-Rusanu, C. D., Iancu, I. V., Huica, I., Plesa, A., Socolov, D., Ungureanu, C., and Anton, G. (2011). Quantitative analysis of the relationship between microRNA-124a,-34b and-203 gene methylation and cervical oncogenesis. *Molecular Medicine Reports*, 4:121–8.

[22] Brenet, F., Moh, M., Funk, P., Feierstein, E., Viale, A. J., Socci, N. D., and Scandura, J. M. (2011). DNA methylation of the first exon is tightly linked to transcriptional silencing. *PloS One*, 6(1):e14524.

[23] Buffart, T. E., Overmeer, R. M., Steenbergen, R. D. M., Tijssen, M., van Grieken, N. C. T., Snijders, P. J. F., Grabsch, H. I., van de Velde, C. J. H., Carvalho, B., and Meijer, G. A. (2008). MAL promoter hypermethylation as a novel prognostic marker in gastric cancer. *British Journal of Cancer*, 99(11):1802.

[24] Buttitta, F., Pellegrini, C., Marchetti, A., Gadducci, A., Cosio, S., Felicioni, L., Barassi, F., Salvatore, S., Martella, C., Coggi, G., et al. (2003). Human telomerase reverse transcriptase mRNA expression assessed by real-time reverse transcription polymerase chain reaction predicts chemosensitivity in patients with ovarian carcinoma. *Journal of Clinical Oncology*, 21(7):1320–1325.

[25] Charbonnier, C., Chiquet, J., and Ambroise, C. (2010). Weighted-LASSO for structured network inference from time course data. *Statistical Applications in Genetics and Molecular Biology*, 9:1–27.

[26] Chen, C., Ridzon, D. A., Broomer, A. J., Zhou, Z., Lee, D. H., Nguyen, J. T., Barbisin, M., Xu, N. L., Mahuvakar, V. R., Andersen, M. R., et al. (2005). Real-time quantification of microRNAs by stem–loop RT–PCR. *Nucleic Acids Research*, 33(20):e179–e179.

[27] Chen, Q., Chen, X., Zhang, M., Fan, Q., Luo, S., and Cao, X. (2011). miR-137 is frequently down-regulated in gastric cancer and is a negative regulator of Cdc42. *Digestive Diseases and Sciences*, 56(7):2009–2016.

[28] Chen, T. M., Pecoraro, G., and Defendi, V. (1993). Genetic analysis of in vitro progression of human papillomavirus-transfected human cervical cells. *Cancer Research*, 53:1167–1171.

[29] Chen, X., Hu, H., Guan, X., Xiong, G., Wang, Y., Wang, K., Li, J., Xu, X., Yang, K., and Bai, Y. (2012). CpG island methylation status of miRNAs in esophageal squamous cell carcinoma. *International Journal of Cancer*, 130(7):1607–1613.

[30] Chen, X., Ruan, A., Wang, X., Han, W., Wang, R., Lou, N., Ruan, H., Qiu, B., Yang, H., and Zhang, X. (2014). miR-129-3p, as a diagnostic and prognostic biomarker for renal cell carcinoma, attenuates cell migration and invasion via downregulating multiple metastasis-related genes. *Journal of Cancer Research and Clinical Oncology*, 140(8):1295–1304.

[31] Chen, X., Zhang, L., Zhang, T., Hao, M., Zhang, X., Zhang, J., Xie, Q., Wang, Y., Guo, M., Zhuang, H., et al. (2013). Methylation-mediated repression of microRNA 129-2 enhances oncogenic SOX4 expression in HCC. *Liver International*, 33(3):476–486.

[32] Chow, L. T., Broker, T. R., and Steinberg, B. M. (2010). The natural history of human papillomavirus infections of the mucosal epithelia. *Apmis*, 118(6-7):422–449.

[33] Colak, S. and ten Dijke, P. (2017). Targeting TGF-β signaling in cancer. *Trends in Cancer*.

[34] Cover, T. M. and Thomas, J. A. (2006). *Elements of Information Theory*. John Wiley, New York.

[35] Crainiceanu, C. M., Ruppet, D., and Wand, M. P. (2005). Bayesian Analysis for Penalized Spline Regression using WinBUGS. *Journal of Statistical Software*, 14:1–47.

[36] Dahl, J., Roychowdhury, V., and Vandenberghe, L. (2005). Maximum likelihood estimation of Gaussian graphical models: numerical implementation and topology selection. *Technical Report*, 1:1.

[37] Dahlhaus, R. (2000). Graphical interaction models for multivariate time series. *Metrika*, 51(2):157–172.

[38] Dahlhaus, R. and Eichler, M. (2003). Causality and graphical models in time series analysis. *Oxford Statistical Science Series*, 1:115–137.

[39] Den Boon, J. A., Pyeon, D., Wang, S. S., Horswill, M., Schiffman, M., Sherman, M., Zuna, R. E., Wang, Z., Hewitt, S. M., Pearson, R., et al. (2015). Molecular transitions from papillomavirus infection to cervical precancer and cancer: Role of stromal estrogen receptor signaling. *Proceedings of the National Academy of Sciences*, 112(25):3255–3264.

[40] D'Ercole, A. J. and Calikoglu, A. S. (2001). The case of local versus endocrine IGF-I actions: the jury is still out. *Growth Hormone & IGF Research*, 11:261–265.

[41] Dobra, A., Lenkoski, A., et al. (2011). Copula Gaussian graphical models and their application to modeling functional disability data. *The Annals of Applied Statistics*, 5(2A):969–993.

[42] Doorbar, J. (2006). Molecular biology of human papillomavirus infection and cervical cancer. *Clinical Science*, 110:525–541.

[43] Du, Y., Liu, Z., Gu, L., Zhou, J., Zhu, B.-d., Ji, J., and Deng, D. (2012). Characterization of human gastric carcinoma-related methylation of 9 miR CpG islands and repression of their expressions in vitro and in vivo. *BMC Cancer*, 12(1):1.

[44] Duensing, S. and Münger, K. (2002). Human papillomaviruses and centrosome duplication errors: modeling the origins of genomic instability. *Oncogene*, 21(40):6241.

[45] Duensing, S. and Münger, K. (2004). Mechanisms of genomic instability in human cancer: insights from studies with human papillomavirus oncoproteins. *International Journal of Cancer*, 109:157–162.

[46] Dürst, M., Dzarlieva-Petrusevska, R., Boukamp, P., Fusenig, N., and Gissmann, L. (1987). Molecular and cytogenetic analysis of immortalized human primary keratinocytes obtained after transfection with human papillomavirus type 16 DNA. *Oncogene*, 1(3):251–256.

[47] Eads, C. A., Danenberg, K. D., Kawakami, K., Saltz, L. B., Blake, C., Shibata, D., Danenberg, P. V., and Laird, P. W. (2000). MethyLight: a high-throughput assay to measure DNA methylation. *Nucleic Acids Research*, 28(8):e32–00.

[48] Eckel, J. E., Gennings, C., Chinchilli, V. M., Burgoon, L. D., and Zacharewski, T. R. (2004). Empirical Bayes gene screening tool for time-course or dose-response microarray data. *Journal of Biopharmaceutical Statistics*, 14:647–670.

[49] Efron, B. (2004). Large-scale simultaneous hypothesis testing: the choice of a null hypothesis. *Journal of the American Statistical Association*, 99(465):96–104.

[50] Efron, B., Tibshirani, R., Storey, J. D., and Tusher, V. (2001). Empirical Bayes analysis of a microarray experiment. *Journal of the American Statistical Association*, 96:1151–1160.

[51] Ernst, J., Nau, G. J., and Bar-Joseph, Z. (2005). Clustering short time series gene expression data. *Bioinformatics*, 21:i159–i168.

[52] Evans, A. J., Russell, R. C., Roche, O., Burry, T. N., Fish, J. E., Chow, V. W. K., Kim, W. Y., Saravanan, A., Maynard, M. A., Gervais, M. L., et al. (2007). VHL promotes E2 box-dependent E-cadherin transcription by HIF-mediated regulation of SIP1 and snail. *Molecular and Cellular Biology*, 27(1):157–169.

[53] Faith, J. J., Hayete, B., Thaden, J. T., Mogno, I., Wierzbowski, J., Cottarel, G., Kasif, S., Collins, J. J., and Gardner, T. S. (2007). Large-scale mapping and validation of escherichia coli transcriptional regulation from a compendium of expression profiles. *PLoS Biology*, 5(1):e8.

[54] Fan, D. M., Tian, X. Y., Wang, R. F., and Yu, J. J. (2014). The prognosis significance of TGF-β1 and ER protein in cervical adenocarcinoma patients with stage Ib˜IIa. *Tumor Biology*, 35(11):11237–11242.

[55] Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, 96:1348–1360.

[56] Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D., Forman, D., and Bray, F. (2015). Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer*, 136:359–386.

[57] Freedman, V. H. and Shin, S. (1974). Cellular tumorigenicity in nude mice: correlation with cell growth in semi-solid medium. *Cell*, 3:355–359.

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

[59] Frisch, S. M., Schaller, M., and Cieply, B. (2013). Mechanisms that link the oncogenic epithelial–mesenchymal transition to suppression of anoikis. *Journal of Cell Science*, 126(1):21–29.

[60] Gardiner-Garden, M. and Frommer, M. (1987). CpG islands in vertebrate genomes. *Journal of Molecular Biology*, 196(2):261–282.

[61] Garofalo, M., Quintavalle, C., Romano, G., M Croce, C., and Condorelli, G. (2012). miR221/222 in cancer: their role in tumor progression and response to therapy. *Current Molecular Medicine*, 12(1):27–33.

[62] Geiger, T. and Levitzki, A. (2007). Loss of robustness and addiction to IGF1 during early keratinocyte transformation by human papilloma virus 16. *PloS One*, 2:605.

[63] Genovesi, L. A., Carter, K. W., Gottardo, N. G., Giles, K. M., and Dallas, P. B. (2011). Integrated analysis of miRNA and mRNA expression in childhood medulloblastoma compared with neural stem cells. *PloS One*, 6(9):e23935.

[64] Guadamillas, M. C., Cerezo, A., and del Pozo, M. A. (2011). Overcoming anoikis pathways to anchorage-independent growth in cancer. *Journal of Cell Science*, 124:3189–3197.

[65] Hamilton, J. D. (1994). *Time series analysis*. Princeton University Press.

[66] Hanash, S. (2004). Integrated global profile of cancer. *Nature Reviews Cancer*, 4:638–644.

[67] Harden, S. V., Guo, Z., Epstein, J. I., and Sidransky, D. (2003). Quantitative GSTP1 methylation clearly distinguishes benign prostatic tissue and limited prostate adenocarcinoma. *The Journal of Urology*, 169(3):1138–1142.

[68] Harville, D. A. (2008). *Matrix Algebra From a Statistician's Perspective*. Springer, New York.

[69] Heinen, A., Rengifo, E., et al. (2003). *Multivariate modelling of time series count data: an autoregressive conditional Poisson model*. Universite Catholique de Louvain.

[70] Henken, F. E., Arce, J. D., Rösl, F., Bosch, L., Meijer, C. J. L. M., Snijders, P. J. F., and Steenbergen, R. D. M. (2012). The functional role of notch signaling in HPV-mediated transformation is dose-dependent and linked to AP-1 alterations. *Cellular Oncology*, 35(2):77–84.

[71] Henken, F. E., Wilting, S. M., Overmeer, R. M., Van Rietschoten, J. G., Nygren, A. O. H., Errami, A., Schouten, J. P., Meijer, C. J. L. M., Snijders, P. J. F., and Steenbergen, R. D. M. (2007). Sequential gene promoter methylation during HPV-induced cervical carcinogenesis. *British Journal of Cancer*, 97:1457–1464.

[72] Hong, F. and Li, H. (2006). Functional Hierarchical Models for Identifying Genes with Different Time-Course Expression Profiles. *Biometrics*, 62:534–544.

[73] Hosseini, E. S., Kashani, H. H., Sabzalipoor, H., Nikzad, H., Meryet-Figuiere, M., and Asemi, Z. (2017). Dysregulated expression of long noncoding RNAs in gynecologic cancers. *Molecular Cancer*, 16(1):107.

[74] Hou, F., Li, Z., Ma, D., Zhang, W., Zhang, Y., Zhang, T., Kong, B., and Cui, B. (2012). Distribution of Th17 cells and Foxp3-expressing T cells in tumor-infiltrating lymphocytes in patients with uterine cervical cancer. *Clinica Chimica Acta*, 413(23):1848–1854.

[75] Huang, Y. W., Liu, J. C., Deatherage, D. E., Luo, J., Mutch, D. G., Goodfellow, P. J., Miller, D. S., and Huang, T. H. M. (2009). Epigenetic repression of microRNA-129-2 leads to overexpression of SOX4 oncogene in endometrial cancer. *Cancer Research*, 69(23):9038–9046.

[76] Huber, W., Von Heydebreck, A., Sültmann, H., Poustka, A., and Vingron, M. (2002). Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics*, 18(suppl 1):S96–S104.

[77] Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U., Speed, T. P., et al. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*, 4(2):249–264.

[78] Jacobs, M. V., Snijders, P. J., Van Den Brule, A. J., Helmerhorst, T. J., Meijer, C. J., and Walboomers, J. M. (1997). A general primer GP5+/GP6 (+)-mediated PCR-enzyme immunoassay method for rapid detection of 14 high-risk and 6 low-risk human papillomavirus genotypes in cervical scrapings. *Journal of Clinical Microbiology*, 35(3):791–795.

[79] Jones, B. and West, M. (2005). Covariance decomposition in undirected Gaussian graphical models. *Biometrika*, 92(4):779–786.

[80] Kamio, M., Yoshida, T., Ogata, H., Douchi, T., Nagata, Y., Inoue, M., Hasegawa, M., Yonemitsu, Y., and Yoshimura, A. (2004). SOC1 inhibits HPV-E7-mediated transformation by inducing degradation of E7 protein. *Oncogene*, 23:3107–3115.

[81] Kanehisa, M. and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, 28(1):27–30.

[82] Kantz, H. and Schreiber, T. (2004). *Nonlinear time series analysis*, volume 7. Cambridge University Press.

[83] Khalifa, Y. B., Teissier, S., Tan, M. K. M., Phan, Q. T., Daynac, M., Wong, W. Q., and Thierry, F. (2011). The human papillomavirus E6 oncogene represses a cell adhesion pathway and disrupts focal adhesion through degradation of TAp63β upon transformation. *PLoS Pathology*, 7:e1002256.

[84] Klingelhutz, A. J. and Roman, A. (2012). Cellular transformation by human papillomaviruses: lessons learned by comparing high-and low-risk viruses. *Virology*, 424(2):77–98.

[85] Kodama, J., Hashimoto, I., Seki, N., Hongo, A., Yoshinouchi, M., Okuda, H., and Kudo, T. (2001). Thrombospondin-1 and-2 messenger RNA expression in invasive cervical cancer correlation with angiogenesis and prognosis. *Clinical Cancer Research*, 7:2826–2831.

[86] Korzeniewski, N., Spardy, N., Duensing, A., and Duensing, S. (2011). Genomic instability and cancer: lessons learned from human papillomaviruses. *Cancer Letters*, 305(2):113–122.

[87] Kozaki, K., Imoto, I., Mogi, S., Omura, K., and Inazawa, J. (2008). Exploration of tumor-suppressive microRNAs silenced by DNA hypermethylation in oral cancer. *Cancer Research*, 68(7):2094–2105.

[88] Kristensen, L. S. and Hansen, L. L. (2009). PCR-based methods for detecting single-locus DNA methylation biomarkers in cancer diagnostics, prognostics, and response to treatment. *Clinical Chemistry*, 55(8):1471–1483.

[89] Langevin, S. M., Stone, R. A., Bunker, C. H., Grandis, J. R., Sobol, R. W., and Taioli, E. (2010). MicroRNA-137 promoter methylation in oral rinses from patients with squamous cell carcinoma of the head and neck is associated with gender and body mass index. *Carcinogenesis*, 31(5):864–870.

[90] Langevin, S. M., Stone, R. A., Bunker, C. H., Lyons-Weiler, M. A., LaFramboise, W. A., Kelly, L., Seethala, R. R., Grandis, J. R., Sobol, R. W., and Taioli, E. (2011). MicroRNA-137 promoter methylation is associated with poorer overall survival in patients with squamous cell carcinoma of the head and neck. *Cancer*, 117(7):1454–1462.

[91] Lauritzen, S. L. (1996). *Graphical models*. Oxford University Press, Oxford.

[92] Lebre, S., Becq, J., Devaux, F., Stumpf, M. P. H., and Lelandais, G. (2010). Statistical inference of the time-varying structure of gene-regulation networks. *BMC Systems Biology*, 4(1):130.

[93] Lindenbergh-Van der Plas, M., Martens-de Kemp, S. R., de Maaker, M., Van Wieringen, W. N., Ylstra, B., Agami, R., Cerisoli, F., Leemans, C. R., Braakhuis, B. J. M., and Brakenhoff, R. H. (2013). Identification of lethal microRNAs specific for head and neck cancer. *Clinical Cancer Research*, 19:5647–5657.

[94] Liu, H., Han, F., Yuan, M., Lafferty, J., Wasserman, L., et al. (2012). High-dimensional semiparametric Gaussian copula graphical models. *The Annals of Statistics*, 40(4):2293–2326.

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

[96] Livak, K. J. and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\delta\delta$CT method. *Methods*, 25(4):402–408.

[97] Long, X. H., Zhou, Y. F., Peng, A. F., Zhang, Z. H., Chen, X. Y., Chen, W. Z., Liu, J. M., Huang, S. H., and Liu, Z. L. (2015). Demethylation-mediated miR-129-5p up-regulation inhibits malignant phenotype of osteogenic osteosarcoma by targeting homo sapiens valosin-containing protein (VCP). *Tumor Biology*, 36(5):3799–3806.

[98] Lönnstedt, I. and Speed, T. P. (2002). Replicated microarray data. *Statistica Sinica*, 12:31–46.

[99] López, J. A. and Alvarez-Salas, L. M. (2011). Differential effects of miR-34c-3p and miR-34c-5p on SiHa cells proliferation apoptosis, migration and invasion. *Biochemical and Biophysical Research Communications*, 409:513–519.

[100] Lu, C. Y., Lin, K. Y., Tien, M. T., Wu, C. T., Uen, Y. H., and Tseng, T. L. (2013). Frequent DNA methylation of miR-129-2 and its potential clinical implication in hepatocellular carcinoma. *Genes, Chromosomes and Cancer*, 52(7):636–643.

[101] Lujambio, A., Calin, G. A., Villanueva, A., Ropero, S., Sánchez-Céspedes, M., Blanco, D., Montuenga, L. M., Rossi, S., Nicoloso, M. S., Faller, W. J., et al. (2008). A microRNA DNA methylation signature for human cancer metastasis. *Proceedings of the National Academy of Sciences*, 105(36):13556–13561.

[102] Luo, D., Wang, J., Li, J., and Post, M. (2011). Mouse snail is a target gene for HIF. *Molecular Cancer Research*, 9(2):234–245.

[103] Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer.

[104] Mack, H. I. D. and Munger, K. (2013). The LKB1 tumor suppressor differentially affects anchorage independent growth of HPV positive cervical cancer cell lines. *Virology*, 446:9–16.

[105] Massagué, J. (2008). TGF-β in cancer. *Cell*, 134(2):215–230.

[106] McCormack, S. J., Brazinski, S. E., Moore Jr, J. L., Werness, B. A., and Goldstein, D. J. (1997). Activation of the focal adhesion kinase signal transduction pathway in cervical carcinoma cell lines and human genital epithelial cells immortalized with human papillomavirus type 18. *Oncogene*, 15(3).

[107] Medina-Martinez, I., Barrón, V., Roman-Bassaure, E., Juárez-Torres, E., Guardado-Estrada, M., Espinosa, A., Bermudez, M., Fernández, F., Venegas-Vega, C., Orozco, L., et al. (2014). Impact of gene dosage on gene expression, biological processes and survival in cervical cancer: a genome-wide follow-up study. *PloS One*, 9(5):e97842.

[108] Mersmann, O. (2014). microbenchmark: Accurate timing functions. *CRAN*, 1.4-2:12.

[109] Meyer, P. E., Kontos, K., Lafitte, F., and Bontempi, G. (2007). Information-theoretic inference of large transcriptional regulatory networks. *EURASIP Journal on Bioinformatics and Systems Biology*, 2007(1):79879.

[110] Miok, V., Wilting, S. M., van de Wiel, M. A., Jaspers, A., van Noort, P. I., Brakenhoff, R. H., Snijders, P. J. F., Steenbergen, R., and van Wieringen, W. N. (2014). tigaR: integrative significance analysis of temporal differential gene expression induced by genomic abnormalities. *BMC Bioinformatics*, 15(1):327.

[111] Miok, V., Wilting, S. M., and van Wieringen, W. N. (2017). Ridge estimation of the VAR(1) model and its time series chain graph from multivariate time-course omics data. *Biometrical Journal*, 59(1):172–191.

[112] Miok, V., Wilting, S. M., and van Wieringen, W. N. (2018). Ridge estimation of network models from time-course omics data. *Submitted for publication*.

[113] Miranda, K. C., Huynh, T., Tay, Y., Ang, Y. S., Tam, W. L., Thomson, A. M., Lim, B., and Rigoutsos, I. (2006). A pattern-based method for the identification of microRNA binding sites and their corresponding heteroduplexes. *Cell*, 126(6):1203–1217.

[114] Monti, S., Tamayo, P., Mesirov, J., and Golub, T. (2003). Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. *Machine Learning*, 52(1-2):91–118.

[115] Moody, C. A. and Laimins, L. A. (2010). Human papillomavirus oncoproteins: pathways to transformation. *Nature Reviews Cancer*, 10:550–560.

[116] Mori, S., Chang, J. T., Andrechek, E. R., Matsumura, N., Baba, T., Yao, G., Kim, J. W., Gatza, M., Murphy, S., and Nevins, J. R. (2009). Anchorage-independent cell growth signature identifies tumors with metastatic potential. *Oncogene*, 28:2796–2805.

[117] Mutarelli, M., Cicatiello, L., Ferraro, L., Grober, O. M. V., Ravo, M., Facchiano, A. M., Angelini, C., and Weisz, A. (2008). Time-course analysis of genome-wide gene expression data from hormone-responisve human breast cancer cells. *BMC Bioinformatics*, 9:S12.

[118] Nau, G. J., Richmond, J. F. L., Schlesinger, A., Jennings, E. G., Lander, E. S., and Young, R. A. (2002). Human macrophage activation programs induced by bacterial pathogens. *Proceedings of the National Academy of Sciences*, 99:1503–1508.

[119] Neuzillet, C., Tijeras-Raballand, A., Cohen, R., Cros, J., Faivre, S., Raymond, E., and de Gramont, A. (2015). Targeting the TGFβ pathway for cancer therapy. *Pharmacology & Therapeutics*, 147:22–31.

[120] Newman, M. (2010). *Networks: an introduction*. Oxford University Press.

[121] Nicholls, D. F. and Pope, A. L. (1988). Bias in the estimation of multivariate autoregressions. *Australian Journal of Statistics*, 30A:296–309.

[122] Novianti, P. W., Snoek, B. C., Wilting, S. M., and van de Wiel, M. A. (2017). Better diagnostic signatures from RNAseq data through use of auxiliary co-data. *Bioinformatics*, 33(10):1572–1574.

[123] Oberhofer, W. and Kmenta, J. (1974). A general procedure for obtaining maximum likelihood estimates in generalized regression models. *Econometrica*, 42:579–590.

[124] Olshen, A. B., Venkatraman, E. S., Lucito, R., and Wigler, M. (2004). Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics*, 5(4):557–572.

[125] Oshlack, A., Robinson, M. D., and Young, M. D. (2010). From RNA-seq reads to differential expression results. *Genome Biology*, 11:220.

[126] Overmeer, R. M., Henken, F. E., Bierkens, M., Wilting, S. M., Timmerman, I., Meijer, C. J. L. M., Snijders, P. J. F., and Steenbergen, R. D. M. (2009). Repression of mal tumour suppressor activity by promoter methylation during cervical carcinogenesis. *The Journal of Pathology*, 219(3):327–336.

[127] Overmeer, R. M., Henken, F. E., Snijders, P. J. F., Claassen-Kramer, D., Berkhof, J., Helmerhorst, T. J. M., Heideman, D. A. M., Wilting, S. M., Murakami, Y., Ito, A., Meijer, C. J. L. M., and Steenbergen, R. D. M. (2008). Association between dense CADM1 promoter methylation and reduced protein expression in high-grade CIN and cervical SCC. *The Journal of Pathology*, 215:388–397.

[128] Paoli, P., Giannoni, E., and Chiarugi, P. (2013). Anoikis molecular pathways and its role in cancer progression. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, 1833(12):3481–3498.

[129] Park, N. H., Min, B. M., Li, S. L., Huang, M. Z., and Doniger, J. (1991). Immortalization of normal human oral keratinocytes with type 16 human papillomavirus. *Carcinogenesis*, 12(9):1627–1631.

[130] Pecoraro, G., Morgan, D., and Defendi, V. (1989). Differential effects of human papillomavirus type 6, 16, and 18 DNAs on immortalization and transformation of human cervical epithelial cells. *Proceedings of the National Academy of Sciences*, 86(2):563–567.

[131] Peel, D. and McLachlan, G. J. (2000). Robust mixture modeling using the t distribution. *Statistics and Computing*, 10:339–348.

[132] Pett, M. R., Alazawi, W. O. F., Roberts, I., Dowen, S., Smith, D. I., Stanley, M. A., and Coleman, N. (2004). Acquisition of high-level chromosomal instability is associated with integration of human papillomavirus type 16 in cervical keratinocytes. *Cancer research*, 64(4):1359–1368.

[133] Pinzón, N., Li, B., Martinez, L., Sergeeva, A., Presumey, J., Apparailly, F., and Seitz, H. (2017). microRNA target prediction programs predict many false positives. *Genome Research*, 27(2):234–245.

[134] Pirisi, L., Yasumoto, S., Feller, M., Doniger, J., and DiPaolo, J. A. (1987). Transformation of human fibroblasts and keratinocytes with human papillomavirus type 16 DNA. *Journal of Virology*, 61(4):1061–1066.

[135] Rau, A., Jaffrézic, F., Foulley, J. L., and Doerge, R. W. (2010). An empirical Bayesian method for estimating biological networks from temporal microarray data. *Statistical Applications in Genetics and Molecular Biology*, 9(1).

[136] Razani, B., Altschuler, Y., Zhu, L., Pestell, R. G., Mostov, K. E., and Lisanti, M. P. (2000). Caveolin-1 expression is down-regulated in cells transformed by the human papilloma virus in a p53-dependent manner. Replacement of caveolin-1 expression suppresses HPV-mediated cell transformation. *Biochemistry*, 39:13916–13924.

[137] Ritchie, M. E., Silver, J., Oshlack, A., Holmes, M., Diyagama, D., Holloway, A., and Smyth, G. K. (2007). A comparison of background correction methods for two-colour microarrays. *Bioinformatics*, 23:2700–2707.

[138] Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26:139–140.

[139] Robinson, M. D. and Smyth, G. K. (2007). Moderated statistical tests for assessing differences in tag abundance. *Bioinformatics*, 23:2881–2887.

[140] Rue, H., Matrino, S., and Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximation. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71:319–392.

[141] Ruiz-Lafuente, N., Alcaraz-García, M. J., García-Serna, A. M., Sebastián-Ruiz, S., Moya-Quiles, M. R., García-Alonso, A. M., and Parrado, A. (2015). Dock10, a Cdc42 and Rac1 GEF, induces loss of elongation, filopodia, and ruffles in cervical cancer epithelial HeLa cells. *Biology Open*, 4(5):627–635.

[142] Ruppert, D., Wand, M. P., and Carroll, R. J. (2009). Semiparametric regression during 2003-2007. *Electronic Journal of Statistics*, 3:1193.

[143] Sand, M., Skrygan, M., Sand, D., Georgas, D., Gambichler, T., Hahn, S. A., Altmeyer, P., and Bechara, F. G. (2013). Comparative microarray analysis of microRNA expression profiles in primary cutaneous malignant melanoma, cutaneous malignant melanoma metastases, and benign melanocytic nevi. *Cell and Tissue Research*, 351(1):85–98.

[144] Schena, M., Shalon, D., Davis, R. W., and Brown, P. O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, 270:467.

[145] Schmittgen, T. D. and Livak, K. J. (2008). Analyzing real-time PCR data by the comparative CT method. *Nature Protocols*, 3(6):1101–1108.

[146] Schütze, D. M., Krijgsman, O., Snijders, P. J. F., Ylstra, B., Weischenfeldt, J., Mardin, B., Stütz, A. M., Korbel, J. O., de Winter, J. P., Meijer, C. J. L. M., et al. (2016). Immortalization capacity of HPV types is inversely related to chromosomal instability. *Oncotarget*, 7(25):37608.

[147] Schütze, D. M., Snijders, P. J. F., Bosch, L., Kramer, D., Meijer, C. J. L. M., and Steenbergen, R. D. M. (2014). Differential in vitro immortalization capacity of eleven, probable high-risk human papillomavirus types. *Journal of Virology*, 88(3):1714–1724.

[148] Schwock, J., Dhani, N., Cao, M. P. J., Zheng, J., Clarkson, R., Radulovich, N., Navab, R., Horn, L.-C., and Hedley, D. W. (2009). Targeting focal adhesion kinase with dominant-negative FRNK or Hsp90 inhibitor 17-DMAG suppresses tumor growth and metastasis of SiHa cervical xenografts. *Cancer Research*, 69(11):4750–4759.

[149] Shapira, S. D., Gat-Viks, I., Shum, B. O. V., Dricot, A., de Grace, M. M., Wu, L., Gupta, P. B., Hao, T., Silver, S. J., Root, D. E., et al. (2009). A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection. *Cell*, 139(7):1255–1267.

[150] Shen, R., Pan, S., Qi, S., Lin, X., and Cheng, S. (2010). Epigenetic repression of microRNA-129-2 leads to overexpression of SOX4 in gastric cancer. *Biochemical and Biophysical Research Communications*, 394(4):1047–1052.

[151] Shen, W., Tao, G. G., Zhang, Y., Cai, B., Sun, J., and Tian, Z. G. (2017). TGF-β in pancreatic cancer initiation and progression: two sides of the same coin. *Cell & Bioscience*, 7(1):39.

[152] Shimizu, T., Suzuki, H., Nojima, M., Kitamura, H., Yamamoto, E., Maruyama, R., Ashida, M., Hatahira, T., Kai, M., Masumori, N., et al. (2013). Methylation of a panel of microRNA genes is a novel biomarker for detection of bladder cancer. *European Urology*, 63(6):1091–1100.

[153] Silber, J., Lim, D. A., Petritsch, C., Persson, A. I., Maunakea, A. K., Yu, M., Vandenberg, S. R., Ginzinger, D. G., James, C. D., Costello, J. F., et al. (2008).

miR-124 and miR-137 inhibit proliferation of glioblastoma multiforme cells and induce differentiation of brain tumor stem cells. *BMC Medicine*, 6(1):1.

[154] Silver, J. D., Ritchie, M. E., and Smyth, G. K. (2008). Microarray background correction: maximum likelihood estimation for the normal exponential convolution. *Biostatistics*, 10(2):352–363.

[155] Smeets, S. J., van der Plas, M., Schaaij-Visser, T., van Veen, E. A. M., van Meerloo, J., Braakhuis, B. J. M., Steenbergen, R. D. M., and Brakenhoff, R. H. (2011). Immortalization of oral keratinocytes by functional inactivation of the p53 and pRb pathways. *International Journal of Cancer*, 128(7):1596–1605.

[156] Smyth, G. K. (2004). Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology*, 3(1):1–25.

[157] Snellenberg, S., Cillessen, S. A. G. M., Van Criekinge, W., Bosch, L., Meijer, C. J. L. M., Snijders, P. J. F., and Steenbergen, R. D. M. (2014). Methylation-mediated repression of PRDM14 contributes to apoptosis evasion in HPV-positive cancers. *Carcinogenesis*, 35(11):2611–2618.

[158] Snellenberg, S., De Strooper, L. M. A., Hesselink, A. T., Meijer, C. J. L. M., Snijders, P. J. F., Heideman, D. A. M., and Steenbergen, R. D. M. (2012). Development of a multiplex methylation-specific PCR as candidate triage test for women with an HPV-positive cervical scrape. *BMC Cancer*, 12(1):1.

[159] Song, L., Kolar, M., and Xing, E. P. (2009). Time-varying dynamic Bayesian networks. pages 1732–1740. Elsevier.

[160] Sopov, I., Sörensen, T., Magbagbeolu, M., Jansen, L., Beer, K., Kühne-Heid, R., Kirchmayr, R., Schneider, A., and Dürst, M. (2004). Detection of cancer-related gene expression profiles in severe cervical neoplasia. *International Journal of Cancer*, 112(1):33–43.

[161] Soto-Reyes, E., González-Barrios, R., Cisneros-Soberanis, F., Herrera-Goepfert, R., Pérez, V., Cantú, D., Prada, D., Castro, C., Recillas-Targa, F., and Herrera, L. A. (2012). Disruption of CTCF at the miR-125b1 locus in gynecological cancers. *BMC Cancer*, 12(1):1.

[162] Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and van der Liner, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64:583–639.

[163] Srivastava, K., Pickard, A., McDade, S., and McCance, D. J. (2015). p63 drives invasion in keratinocytes expressing HPV16 E6/E7 genes through regulation of Src-FAK signalling. *Oncotarget*.

[164] Srivastava, S. K., Ahmad, A., Zubair, H., Miree, O., Singh, S., Rocconi, R. P., Scalici, J., and Singh, A. P. (2017). MicroRNAs in gynecological cancers: Small molecules with big implications. *Cancer Letters*.

[165] Steenbergen, R. D. M., de Wilde, J., Wilting, S. M., Brink, A. A. T. P., Snijders, P. J. F., and Meijer, C. J. L. M. (2005). HPV-mediated transformation of the anogenital tract. *Journal of Clinical Virology*, 32:25–33.

[166] Steenbergen, R. D. M., Kramer, D., Braakhuis, B. J. M., Stern, P. L., Verheijen, R. H. M., Meijer, C. J. L. M., and Snijders, P. J. F. (2004). TSLC1 gene silencing in cervical cancer cell lines and cervical neoplasia. *Journal of the National Cancer Institute*, 96:294–305.

[167] Steenbergen, R. D. M., Kramer, D., Meijer, C. J. L. M., Walboomers, J. M. M., Trott, D. A., Cuthbert, A. P., Newbold, R. F., Overkamp, W. J. I., Zdzienicka, M. Z., and Snijders, P. J. F. (2001). Telomerase suppression by chromosome 6 in a human papillomavirus type 16-immortalized keratinocyte cell line and in a cervical cancer cell line. *Journal of the National Cancer Institute*, 93(11):865–872.

[168] Steenbergen, R. D. M., Snijders, P. J. F., Heideman, D. A. M., and Meijer, C. J. L. M. (2014). Clinical implications of (epi) genetic changes in HPV-induced cervical precancerous lesions. *Nature Reviews Cancer*, 14:395–405.

[169] Steenbergen, R. D. M., Walboomers, J. M. M., Meijer, C. J. L. M., Van der Raaij-Helmer, E. M. H., Parker, J. N., Chow, L. T., Broker, T. R., and Snijders, P. J. F. (1996). Transition of human papillomavirus type 16 and 18 transfected human foreskin keratinocytes towards immortality: activation of telomerase and allele loss at 3p, 10p, 11q and/or 18q. *Oncogene*, 13:1249–1257.

[170] Steponaitiene, R., Kupcinskas, J., Langner, C., Balaguer, F., Venclauskas, L., Pauzas, H., Tamelis, A., Skieceviciene, J., Kupcinskas, L., Malfertheiner, P., et al. (2015). Epigenetic silencing of miR-137 is a frequent event in gastric carcinogenesis. *Molecular Carcinogenesis*.

[171] Storey, J. D., Xiao, W., Leek, J. T., Tompkins, R. G., and Davis, R. W. (2005). Significance analysis of time course microarray experiments. *Proceedings of the National Academy of Sciences of the United States of America*, 102:12837–12842.

[172] Strimmer, K. (2008). A unified approach to false discovery rate estimation. *BMC Bioinformatics*, 9(1):303.

[173] Sun, L., Li, T., Wei, Q., Zhang, Y., Jia, X., Wan, Z., and Han, L. (2014). Upregulation of BNIP3 mediated by ERK/HIF-1α pathway induces autophagy and contributes to anoikis resistance of hepatocellular carcinoma cells. *Future Oncology*, 10(8):1387–1398.

[174] Szalmás, A. and Kónya, J. (2009). Epigenetic alterations in cervical carcinogenesis. 19:144–152.

[175] Tabas-Madrid, D., Muniategui, A., Sánchez-Caballero, I., Martínez-Herrera, D. J., Sorzano, C. O. S., Rubio, A., and Pascual-Montano, A. (2014). Improving miRNA-mRNA interaction predictions. *BMC Genomics*, 15(10):S2.

[176] Tai, Y. C. and Speed, T. P. (2006). A multivariate empirical Bayes statistics for replicated microarray time course data. *The Annals of Statistics*, 34:2387–2412.

[177] Takwi, A. A., Wang, Y. M., Wu, J., Michaelis, M., Cinatl, J., and Chen, T. (2014). miR-137 regulates the constitutive androstane receptor and modulates doxorubicin sensitivity in parental and doxorubicin-resistant neuroblastoma cells. *Oncogene*, 33(28):3717–3729.

[178] Tang, W., Feng, X., Zhang, S., Ren, Z., Liu, Y., Yang, B., Cai, Y., Xia, J., Ge, N., et al. (2015). Caveolin-1 confers resistance of hepatoma cells to anoikis by activating IGF-1 pathway. *Cellular Physiology and Biochemistry*, 36(3):1223–1236.

[179] Thomas, L. K., Bermejo, J. L., Vinokurova, S., Jensen, K., Bierkens, M., Steenbergen, R., Bergmann, M., von Knebel Doeberitz, M., and Reuschenbach, M. (2014). Chromosomal gains and losses in human papillomavirus-associated neoplasia of the lower genital tract–a systematic review and meta-analysis. *European Journal of Cancer*, 50(1):85–98.

[180] Touleimat, N. and Tost, J. (2012). Complete pipeline for Infinium® Human Methylation 450K BeadChip data processing using subset quantile normalization for accurate DNA methylation estimation. *Epigenomics*, 4(3):325–341.

[181] Toussaint-Smith, E, D. D. B. and Roman, A. (2004). Expression of human papillomavirus type 16 E6 and E7 oncoproteins in primary foreskin keratinocytes is sufficient to alter the expression of angiogenic factors. *Oncogene*, 23:2988–2995.

[182] Tsai, K. W., Wu, C. W., Hu, L. Y., Li, S. C., Liao, Y. L., Lai, C. H., Kao, H. W., Fang, W. L., Huang, K. H., Chan, W. C., et al. (2011). Epigenetic regulation of miR-34b and miR-129 expression in gastric cancer. *International Journal of Cancer*, 129(11):2600–2610.

[183] Tusher, V. G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences*, 98(9):5116–5121.

[184] Van de Wiel, M. A., Kim, K. I., Vosse, S. J., Van Wieringen, W. N., Wilting, S. M., and Ylstra, B. (2007). CGHcall: calling aberrations for array CGH tumor profiles. *Bioinformatics*, 23:892–894.

[185] Van de Wiel, M. A., Leday, G. G. R., Pardo, L., Rue, H., Van der Vaart, A. W., and Van Wieringen, W. N. (2013). Bayesian analysis of RNA sequencing data by estimating multiple shrinkage priors. *Biostatistics*, 14:113–128.

[186] Van de Wiel, M. A., Picard, F., Van Wieringen, W. N., and Ylstra, B. (2010). Preprocessing and downstream analysis of microarray DNA copy number profiles. *Briefings in Bioinformatics*, 12(1):10–21.

[187] Van Loan, C. F. (2000). The ubiquitous kronecker product. *Journal of Computational and Applied Mathematics*, 123:85–100.

[188] Van Wieringen, W. N., Berkhof, J., and Van de Wiel, M. A. (2010). A random coefficients model for regional co-expression associated with DNA copy number. *Statistical Applications in Genetics and Molecular Biology*, 9:1–30.

[189] Van Wieringen, W. N. and Peeters, C. F. W. (2016). Ridge estimation of inverse covariance matrices from high-dimensional data. *Computational Statistics & Data Analysis*, 103:284–303.

[190] Van Wieringen, W. N., Unger, K., Leday, G. G. R., Krijgsman, O., de Menezes, R., Ylstra, B., and Van de Wiel, M. A. (2012). Matching of array CGH and gene expression microarray features for the purpose of integrative genomic analysis. *BMC Bioinformatics*, 13:80.

[191] Van Wieringen, W. N., Van de Wiel, M. A., and Ylstra, B. (2007). Weighted clustering of called array CGH data. *Biostatistics*, 9(3):484–500.

[192] Van Wieringen, W. N. and van der Vaart, A. W. (2015). Transcriptomic heterogeneity in cancer as a consequence of dysregulation of the gene-gene interaction network. *Bulletin of Mathematical Biology*, 77(9):1768–1786.

[193] Van Zeeburg, H. J. T., Snijders, P. J. F., Pals, G., Hermsen, M. A., Rooimans, M. A., Bagby, G., Soulier, J., Gluckman, E., Wennerberg, J., Leemans, C. R., et al. (2005). Generation and molecular characterization of head and neck squamous cell lines of fanconi anemia patients. *Cancer Research*, 65(4):1271–1276.

[194] Veigaard, C. and Kjeldsen, E. (2014). Exploring the genome-wide relation between copy number status and microRNA expression. *Genomics*, 104(4):271–278.

[195] Vogelstein, B. and Kinzler, K. W. (2004). Cancer genes and the pathways they control. *Nature Medicine*, 10:789–799.

[196] Vrba, L., Muñoz-Rodríguez, J. L., Stampfer, M. R., and Futscher, B. W. (2013). miRNA gene promoters are frequent targets of aberrant DNA methylation in human breast cancer. *PloS One*, 8(1):e54398.

[197] Walboomers, J. M. M., Jacobs, M. V., Manos, M. M., Bosch, F. X., Kummer, J. A., Shah, K. V., Snijders, P. J. F., Peto, J., Meijer, C. J. L. M., and Muñoz, N. (1999). Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *The Journal of Pathology*, 189(1):12–19.

[198] Wang, F., Liu, M., Li, X., and Tang, H. (2013). MiR-214 reduces cell survival and enhances cisplatin-induced cytotoxicity via down-regulation of Bcl2l2 in cervical cancer cells. *FEBS Letters*, 587(5):488–495.

[199] Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, 10:57–63.

[200] Weinberg, R. A. (2006). *The Biology of Cancer*. Garland Science, New York.

[201] Wentzensen, N., Sherman, M. E., Schiffman, M., and Wang, S. S. (2009). Utility of methylation markers in cervical cancer early detection: appraisal of the state-of-the-science. *Gynecologic Oncology*, 112:293–299.

[202] Whittaker, J. (1990). *Graphical Models in Applied Multivariate Statistics*. John Wiley & Sons, Chichester, England.

[203] Wiklund, E. D., Gao, S., Hulf, T., Sibbritt, T., Nair, S., Costea, D. E., Villadsen, S. B., Bakholdt, V., Bramsen, J. B., Sørensen, J. A., et al. (2011). MicroRNA alterations and associated aberrant DNA methylation patterns across multiple sample types in oral squamous cell carcinoma. *PLoS One*, 6(11):e27840.

[204] Willey, J. C., Broussoud, A., Sleemi, A., Bennett, W. P., Cerutti, P., and Harris, C. C. (1991). Immortalization of normal human bronchial epithelial cells by human papillomaviruses 16 or 18. *Cancer Research*, 51(19):5370–5377.

[205] Wilting, S. M., de Wilde, J., Meijer, C. J. L., Berkhof, J., Yi, Y., Van Wieringen, W. N., Braakhuis, B. J. M., Meijer, G. A., Ylstra, B., Snijders, P. J. F., and Steenbergen, R. D. M. (2008). Integrated genomic and transcriptional profiling identifies chromosomal loci with altered gene expression in cervical cancer. *Genes, Chromosomes and Cancer*, 47:890–905.

[206] Wilting, S. M., Miok, V., Jaspers, A., Boon, D., Sørgård, H., Lando, M., Snoek, B. C., van Wieringen, W. N., Meijer, C. J. L. M., Lyng, H., et al. (2016). Aberrant methylation-mediated silencing of microRNAs contributes to HPV-induced anchorage independence. *Oncotarget*, 7(28):43805.

[207] Wilting, S. M., Snijders, P. J. F., Meijer, G. A., Ylstra, B., Van den Ijssel, P. R., Snijders, A. M., Albertson, D. G., Coffa, J., Schouten, J. P., Van de Wiel, M. A., Meijer, C. J. L. M., and Steenbergen, R. D. M. (2006). Increased gene copy number at chromosome 20q are frequent in both squamous cell carcinomas and adenocarcinomas of the cervix. *The Journal of Pathology*, 209:220–230.

[208] Wilting, S. M. and Steenbergen, R. D. M. (2016). Molecular events leading to HPV-induced high grade neoplasia. *Papillomavirus Research*, 2:85–88.

[209] Wilting, S. M., van Boerdonk, R. A. A., Henken, F. E., Meijer, C. J. L. M., Diosdado, B., Meijer, G. A., le Sage, C., Agami, R., Snijders, P. J. F., and Steenbergen, R. D. M. (2010). Methylation-mediated silencing and tumour suppressive function of hsa-miR-124 in cervical cancer. *Molecular Cancer*, 9(1):1.

[210] Wilting, S. M., Verlaat, W., Jaspers, A., Makazaji, N. A., Agami, R., Meijer, C. J. L. M., Snijders, P. J. F., and Steenbergen, R. D. M. (2013). Methylation-mediated transcriptional repression of microRNAs during cervical carcinogenesis. *Epigenetics*, 8:220–228.

[211] Wong, N. and Wang, X. (2014). miRDB: an online resource for microRNA target prediction and functional annotations. *Nucleic Acids Research*, 43(D1):D146–D152.

[212] Wong, Y. F., Cheung, T. H., Tsao, G. S. W., Lo, K. W. K., Yim, S. F., Wang, V. W., Heung, M., Chan, S., Chan, L. K. Y., Ho, T. W. F., et al. (2006). Genome-wide gene expression profiling of cervical cancer in Hong Kong women by oligonucleotide microarray. *International Journal of Cancer*, 118(10):2461–2469.

[213] Wu, H. S., Li, Y. F., Chou, C. I., Yuan, C. C., Hung, M. W., and Tsai, L. C. (2002). The concentration of serum Transforming Growth Factor beta-1 (TGF-β1) is decreased in cervical carcinoma patients. *Cancer Investigation*, 20(1):55–59.

[214] Yang, M. H., Wu, M. Z., Chiou, S. H., Chen, P. M., Chang, S. Y., Liu, C. J., Teng, S. C., and Wu, K. J. (2008). Direct regulation of TWIST by HIF-1α promotes metastasis. *Nature Cell Biology*, 10(3):295–305.

[215] Yao, T., Rao, Q., Liu, L., Zheng, C., Xie, Q., Liang, J., and Lin, Z. (2013). Exploration of tumor-suppressive microRNAs silenced by DNA hypermethylation in cervical cancer. *Virology Journal*, 10(1):1.

[216] Ye, H., Zhang, Y., Geng, L., and Li, Z. (2015). Cdc42 expression in cervical cancer and its effects on cervical tumor invasion and migration. *International Journal of Oncology*, 46(2):757–763.

[217] Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association*, 57:977–992.

[218] Zhang, J., Li, S., Yan, Q., Chen, X., Yang, Y., Liu, X., and Wan, X. (2013). Interferon-β induced microRNA-129-5p down-regulates HPV-18 E6 and E7 viral gene expression by targeting SP1 in cervical cancer cells. *PLoS One*, 8(12):e81366.

[219] Zhang, L., Huang, J., Yang, N., Greshock, J., Megraw, M., Giannakakis, A., Liang, S., Naylor, T., Barchetti, A., Ward, M., et al. (2006). microRNAs exhibit high frequency genomic alterations in human cancer. *Proceedings of the National Academy of Sciences*, 103(24):9136–9141.

[220] Zhang, L., Volinia, S., Bonome, T., Calin, G. A., Greshock, J., Yang, N., Liu, C. G., Giannakakis, A., Alexiou, P., Hasegawa, K., et al. (2008). Genomic and epigenetic alterations deregulate microRNA expression in human epithelial ovarian cancer. *Proceedings of the National Academy of Sciences*, 105(19):7004–7009.

[221] Zhu, H., Luo, H., Shen, Z., Hu, X., Sun, L., and Zhu, X. (2016). Transforming growth factor-β1 in carcinogenesis, progression, and therapy in cervical cancer. *Tumor Biology*, 37(6):7075–7083.

[222] Zhu, X., Er, K., Mao, C., Yan, Q., Xu, H., Zhang, Y., Zhu, J., Cui, F., Zhao, W., and Shi, H. (2013). miR-203 suppresses tumor growth and angiogenesis by targeting VEGFA in cervical cancer. *Cellular Physiology and Biochemistry*, 32(1):64–73.

[223] Zoppoli, P., Morganella, S., and Ceccarelli, M. (2010). Timedelay-ARACNE: Reverse engineering of gene networks from time-course data by an information theoretic approach. *BMC Bioinformatics*, 11(1):154.