HAL
archives-ouvertes.fr

Graphical Lasso Granger Method with 2-Levels-Thresholding for Recovering Causality Networks

Sergiy Pereverzyev Jr., Kateřina Hlaváčková-Schindler

To cite this version:

Sergiy Pereverzyev Jr., Kateřina Hlaváčková-Schindler. Graphical Lasso Granger Method with 2-Levels-Thresholding for Recovering Causality Networks. Christian Pötzsche; Clemens Heuberger; Barbara Kaltenbacher; Franz Rendl. 26th Conference on System Modeling and Optimization (CSMO), Sep 2013, Klagenfurt, Austria. Springer Berlin Heidelberg, IFIP Advances in Information and Communication Technology, AICT-443, pp.220-229, 2014, System Modeling and Optimization. <10.1007/978-3-662-45504-3_21>. <hal-01286461>

HAL Id: hal-01286461

https://hal.inria.fr/hal-01286461

Submitted on 10 Mar 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution| 4.0 International License

Graphical Lasso Granger method with 2-levels-thresholding for recovering causality networks

Sergiy Pereverzyev Jr.[1] and Kateřina Hlaváčková-Schindler[2]

[1] Applied Mathematics Group, Department of Mathematics, University of Innsbruck, Austria
sergiy.pereverzyev@uibk.ac.at

[2] Chair of Bioinformatics, University of Natural Resources and Life Sciences Vienna, Austria
katerina.schindler@gmail.com

Abstract. The recovery of the causality networks with a number of variables is an important problem that arises in various scientific contexts. For detecting the causal relationships in the network with a big number of variables, the so called Graphical Lasso Granger (GLG) method was proposed. It is widely believed that the GLG-method tends to overselect causal relationships. In this paper, we propose a thresholding strategy for the GLG-method, which we call 2-levels-thresholding, and we show that with this strategy the variable overselection of the GLG-method may be overcomed. Moreover, we demonstrate that the GLG-method with the proposed thresholding strategy may become superior to other methods that were proposed for the recovery of the causality networks.
Keywords: causality network, gene causality network, Granger causality, graphical Lasso method, 2-levels-thresholding.

1 Introduction

Causality is a relationship between a cause and its effect (its consequence). One can say that *inverse problems* solving, where one would like to discover unobservable features of the cause from the observable features of an effect [4], i.e., searching for the cause of an effect, can in general be seen as a causality problem.

A *causality network* is a directed graph with nodes, which are variables $\{x^j,\ j = 1, \ldots, p\}$, and directed edges, which are the causal influences between the variables. We write $x^i \leftarrow x^j$ if the variable x^j has a causal influence on the variable x^i. Causality networks arise in various scientific contexts.

For example, In Cell Biology one considers causality networks which involve sets of active genes of a cell. An active gene produces a protein. It has been observed that the amount of the protein which is produced by a given gene may depend on, or may be *causally* influenced by, the amount of the proteins which are produced by other genes. In this way, causal relationships between genes and the corresponding causality network arise. These causality networks are

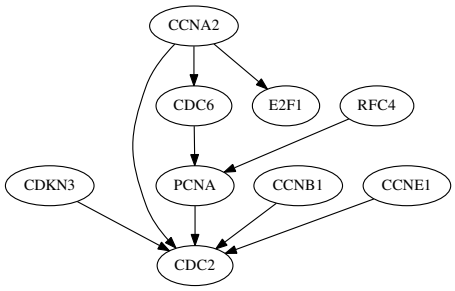

Fig. 1. Causality network of the human cancer cell HeLa genes from the BioGRID database (www.thebiogrid.org).

also called gene regulatory networks. An example of such a network is presented in Fig. 1. This network is achieved from the biological experiments in [9], and it can be found in the BioGRID database. This network has been used in several works [14, 11, 15] as a test network.

Knowledge of the correct causality networks is important for changing them. In Cell Biology, these networks are used in the research of the causes of genetic diseases. For example, the network in Fig. 1 consists of genes that are active in the human cancer cell HeLa [18]. If one wants to suppress the genes expression in this network, then the primary focus of the suppression therapy should be on the causing genes. For the use of the causality networks in other sciences see, for example, [12].

How can causality network be recovered? In practice, the first information that can be known about the network is the time evolution (time series) of the involved variables $\{x_t^j,\ t = 1, \ldots, T\}$. How can this information be used for inferring causal relationships between the variables?

The statistical approach to the derivation of the causal relationships between a variable y and variables $\{z^j,\ j = 1, \ldots, p\}$ using the known time evolution of their values $\{y_t, z_t^j,\ t = 1, \ldots, T,\ j = 1, \ldots, p\}$ consists in considering a model of the relationship between y and $\{z^j,\ j = 1, \ldots, p\}$. As a first step, one can consider a linear model of this relationship: $y_t \approx \sum_{j=1}^{p} \beta^j z_t^j, \quad t = 1, \ldots, T$. The coefficients $\{\beta^j,\ j = 1, \ldots, p\}$ can be specified using the least-squares method. Then, in Statistics [19] by fixing the value of a threshold parameter $\beta_{\mathrm{tr}} > 0$, one says that there is a causal relationship $y \leftarrow z^j$ if $|\beta^j| > \beta_{\mathrm{tr}}$.

For detecting causal relationships between variables $\{x^j,\ j = 1, \ldots, p\}$ the concept of the so called *multivariate Granger causality* has been proposed. This concept originated in the work of Clive Granger [6], who was awarded the Nobel Prize in Economic Sciences in 2003. Based on the intuition that the cause should precede its effect, in Granger causality one says that a variable x^i can be potentially caused by the past versions of the involved variables $\{x^j,\ j = 1, \ldots, p\}$.

Then, in the spirit of the statistical approach and using a linear model for the causal relationship, we consider the following approximation problem:

$$x_t^i \approx \sum_{j=1}^{p} \sum_{l=1}^{L} \beta_l^j x_{t-l}^j, \quad t = L+1, \ldots, T, \tag{1}$$

where L is the so called *maximal lag*, which is the maximal number of the considered past versions of the variables. The coefficients $\{\beta_l^j\}$ can be determined by the least-squares method. As in the statistical approach, one can now fix the value of the threshold parameter $\beta_{\mathrm{tr}} > 0$ and say that

$$x^i \leftarrow x^j \quad \text{if} \quad \sum_{l=1}^{L} |\beta_l^j| > \beta_{\mathrm{tr}}. \tag{2}$$

It is well known that for a big number of genes p, as it is pointed out for example in [11], the causality network, which is obtained from the approximation problem (1), is not satisfactory. First of all, it cannot be guaranteed that the solution of the corresponding minimization problem is unique. Another issue is connected with the number of the causality relationships that is obtained from (1). This number is typically very big, while one expects to have a few causality relationships with a given gene. To address this issue, various *variable selection procedures* can be employed. The Lasso [16] is a well known example of such a procedure. In the regularization theory, this approach is known as the l_1-Tikhonov regularization. It has been extensively used for reconstructing the sparse structure of an unknown signal. We refer the interesting reader to [3, 13, 7, 5, 10] and the references therein.

The causality concept that is based on the Lasso was proposed in [1] and is named *Graphical Lasso Granger (GLG)* method. However, it is stated in the literature that the Lasso suffers from the variable overselection. And therefore, in the context of the gene causality networks several Lasso modifications were proposed. In [11], the so called *group Lasso* method was considered for recovering gene causality networks using the multivariate Granger causality. The corresponding method can be named Graphical group Lasso Granger (GgrLG) method. And in [15], the *truncating Lasso* method was proposed. The resulting method can be named Graphical truncating Lasso Granger (GtrLG) method.

Nevertheless, it seems that an important tuning possibility of the Lasso, namely an appropriate choice of the *threshold parameter* β_{tr}, has been overlooked in the literature devoted to the recovery of the gene causality networks. In this paper, we are going to show that the GLG-method, which is equipped with an appropriate *thresholding strategy* and an appropriate *regularization parameter choice rule*, may become a superior method in comparison to other methods that were proposed for the recovery of the gene causality networks.

The paper is organized as follows. In Sect. 2, we recall the GLG-method. The quality measures of the graphical methods are presented in Sect. 3. In Sect. 4, we use the network from Fig. 1 to compare the performance of the known graphical methods with the ideal version of the GLG-method, which we call the *optimal*

GLG-estimator. Such a comparison demonstrates the potential of the GLG-approach. In Sect. 5, we propose a thresholding strategy for the GLG-method that allows its automatic realization, which we describe in Sect. 6. Then again we use the network from Fig. 1 to compare the performance of the proposed version of the GLG-method with other graphical methods. It turns out that the proposed method has a superior quality compared to the known methods. The paper is finished with the conclusion and outlook in Sect. 7.

2 Graphical Lasso Granger method

Let us specify the application of the least-squares method to the approximation problem (1). For this purpose, let us define the vectors $Y^i = (x^i_{L+1}, x^i_{L+2}, \ldots, x^i_T)'$, $\boldsymbol{\beta} = (\beta^1_1, \ldots, \beta^1_L, \beta^2_1, \ldots, \beta^2_L, \ldots, \beta^p_1, \ldots, \beta^p_L)'$, and the matrix

$$X = \big((x^1_{t-1}, \ldots, x^1_{t-L}, x^2_{t-1}, \ldots, x^2_{t-L}, \ldots, x^p_{t-1}, \ldots, x^p_{t-L});\ t = L+1, \ldots, T \big).$$

Then, in the least-squares method, one considers the following minimization problem:

$$\|Y^i - X\boldsymbol{\beta}\|^2 \to \min_{\boldsymbol{\beta}}, \tag{3}$$

where $\|\cdot\|$ denotes the l_2-norm.

As it was mentioned in the introduction, the solution of (3) defines unsatisfactory causal relationships and various variable selection procedures should be employed instead. A well-known example of such procedures is the Lasso [16]. In this procedure, one considers the following minimization problem:

$$\|Y^i - X\boldsymbol{\beta}\|^2 + \lambda\|\boldsymbol{\beta}\|_1 \to \min_{\boldsymbol{\beta}}. \tag{4}$$

Solution of (4) for each variable $\{x^i,\ i = 1, \ldots, p\}$ with the causality rule (2) defines an estimator of the causality network between the variables $\{x^i\}$, and in this way one obtains the Graphical Lasso Granger (GLG) method [1].

3 Quality measures of the graphical methods

A *graphical method* is a method that reconstructs the causality network, which is a directed graph, with the variables $\{x^j\}$. The *quality* of a graphical method can be estimated from its performance on a *known* causality network. The network in Fig. 1 has been used for testing methods' quality in several publications [14, 11, 15]. What measures can be used for estimating the quality of a graphical method?

First of all, let us note that a causality network can be characterized by the so called *adjacency matrix* $A = \{A_{i,j} \mid \{i,j\} \subset \{1, \ldots, p\}\ \}$ with the following elements:

$$A_{i,j} = 1 \quad \text{if} \quad x^i \leftarrow x^j; \quad A_{i,j} = 0 \quad \text{otherwise.}$$

The adjacency matrix A^{true} for the causality network in Fig. 1 is presented in Fig. 2. There, the white squares correspond to $A_{i,j} = 1$, and the black squares — to the zero-elements. The genes are numbered in the following order: CDC2, CDC6, CDKN3, E2F1, PCNA, RFC4, CCNA2, CCNB1, CCNE1.

Now, imagine that there is a *true* adjacency matrix A^{true} of the true causality network, and there is its *estimator* A^{estim}, which is produced by a graphical method. The quality of the estimator A^{estim} can be characterized by the following quality measures: *precision (P)*, *recall (R)*, F_1*-score* (F_1). See, for example, [12] for the detailed definition of these measures.

As it was already mentioned, the causality network in Fig. 1 has been used for testing quality of graphical methods. In particular, in [15] one finds the above mentioned quality measures for the following methods: GgrLG, GtrLG and CNET. CNET is a graph search-based algorithm that was introduced in [14]. The data $\{x_t^j\}$ is taken from the third experiment of [18] consisting of 47 time points, and the maximal lag L is taken to be equal to 3. The quality measures from [15] are presented in Table 1.

Table 1. Quality measures of the known graphical methods.

	P	R	F_1
GgrLG	0.24	0.44	0.3
GtrLG	0.3	0.33	0.32
CNET	0.36	0.44	0.4

As it is seen from the table, CNET has the highest F_1-score. However, CNET is the most computationally expensive among the considered methods that does not allow its application to large networks. GgrLG has a good recall but a poor precision, and thus, GtrLG can be considered as a better method among the considered methods.

4 Optimal GLG-estimator

As we have seen in the previous section, the graphical methods, which are based on the Lasso modifications, were tested on the network in Fig. 1 (Table 1). However, the application of the graphical method that is based on the pure Lasso (GLG) to the network in Fig. 1 has not been reported. Moreover, it seems that the possibility of varying the threshold parameter β_{tr} in GLG also has not been considered in the literature devoted to the reconstruction of the causality networks.

Assume that the true causality network with the variables $\{x^j\}$ is given by the adjacency matrix A^{true}. Assume further that the observation data $\{x_t^j\}$ is given. What is the *best* reconstruction of A^{true} that can be achieved by the GLG-method? The answer to this question is given by, what we call, the *optimal* GLG-estimator. Let us specify its construction.

First of all, let us define the following quality measure, which we call Fs-measure: $Fs = \frac{1}{p^2}\|A^{\text{true}} - A^{\text{estim}}\|_1,\ 0 \le Fs \le 1$. Fs-measure represents the number of *false* elements in the estimator A^{estim} that is scaled with the total number of elements in A^{estim}.

Now, let $\boldsymbol{\beta}_i(\lambda)$ denote the solution of the minimization problem (4) in the GLG-method, and $\boldsymbol{\beta}_i^j(\lambda) = (\beta_{1,i}^j, \ldots, \beta_{L,i}^j)$. Then, the GLG-estimator $A^{\text{GLG}}(\lambda, \beta_{\text{tr}})$ of the adjacency matrix A^{true} is defined as follows:

$$A_{i,j}^{\text{GLG}}(\lambda, \beta_{\text{tr}}) = 1 \quad \text{if} \quad \|\boldsymbol{\beta}_i^j(\lambda)\|_1 > \beta_{\text{tr}}; \quad A_{i,j}^{\text{GLG}}(\lambda, \beta_{\text{tr}}) = 0 \quad \text{otherwise.}$$

The optimal GLG-estimator $A^{\text{GLG,opt}}$ of the true adjacency matrix A^{true} is the GLG-estimator $A^{\text{GLG}}(\lambda, \beta_{\text{tr}})$ with the parameters λ, β_{tr} such that the corresponding Fs-measure is minimal.

The optimal GLG-estimator of the adjacency matrix for the causality network in Fig. 1 is presented in Fig. 2. Its quality measures can be found in Table 2. We used the same data $\{x_t^j\}$ as in [14, 11, 15]. Also, as in [11, 15], we take the maximal lag $L = 3$. As one can see, the optimal GLG-estimator reconstructs almost completely the causing genes of the most caused gene in the network. The recall of $A^{\text{GLG,opt}}$ is equal to the highest recall in Table 1, but precision and F_1-score are considerably higher.

Of course, $A^{\text{GLG,opt}}$ is given by the ideal version of the GLG-method, where we essentially use the knowledge of A^{true}. How close can we come to $A^{\text{GLG,opt}}$ without such a knowledge? To answer this question, let us first decide about the choice of the threshold parameter β_{tr}.

5 Thresholding strategy

The purpose of the threshold parameter β_{tr} is to cancel the causal relationships $x^i \leftarrow x^j$ with *small* $\|\boldsymbol{\beta}_i^j(\lambda)\|_1$. When can we say that $\|\boldsymbol{\beta}_i^j(\lambda)\|_1$ is small? We propose to consider the following *guideindicators* of smallness:

$$\begin{aligned} \beta_{\min}^i(\lambda) &= \min\{\ \|\boldsymbol{\beta}_i^j(\lambda)\|_1,\ j = 1, \ldots, p \mid \|\boldsymbol{\beta}_i^j(\lambda)\|_1 \neq 0\ \}, \\ \beta_{\max}^i(\lambda) &= \max\{\ \|\boldsymbol{\beta}_i^j(\lambda)\|_1,\ j = 1, \ldots, p\ \}. \end{aligned} \tag{5}$$

In particular, we propose to consider the threshold parameter of the following form:

$$\beta_{\text{tr},\alpha}^i(\lambda) = \beta_{\min}^i(\lambda) + \alpha(\ \beta_{\max}^i(\lambda) - \beta_{\min}^i(\lambda)\ ). \tag{6}$$

As a default value we take $\alpha = 1/2$.

In the optimal GLG-estimator $A_{\text{tr},1/2}^{\text{GLG,opt}}$ with the threshold parameter $\beta_{\text{tr},1/2}^i$ we choose λ such that the corresponding Fs-measure is minimal. For the causality network in Fig. 1, this estimator is presented in Fig. 2. Its quality measures can be found in Table 2. One observes that although there is some quality decrease in comparison to $A^{\text{GLG,opt}}$, the quality measures are still higher than for the methods in Table 1. However, can this quality be improved?

The choice of the threshold parameter $\beta^i_{\mathrm{tr},1/2}$ rises the following issue. With such a choice we always assign a causal relationship, unless the solution of (4) $\boldsymbol{\beta}_i(\lambda)$ is identically zero. But how strong are these causal relationships compared to each other? The norm $\|\boldsymbol{\beta}^j_i(\lambda)\|_1$ can be seen as a *strongness indicator* of the causal relationship $x^i \leftarrow x^j$.

Let us now construct a matrix $A^{\mathrm{GLG,opt};\boldsymbol{\beta}}_{\mathrm{tr},1/2}$, similarly to the adjacency matrix $A^{\mathrm{GLG,opt}}_{\mathrm{tr},1/2}$, where instead of the element 1 we put the norm $\|\boldsymbol{\beta}^j_i(\lambda)\|_1$, i.e.

$$\begin{aligned} A^{\mathrm{GLG,opt};\boldsymbol{\beta}}_{\mathrm{tr},1/2}(i,j) =& \|\boldsymbol{\beta}^j_i(\lambda^{\mathrm{tr},1/2}_{\mathrm{opt},i})\|_1 \quad \text{if} \quad \|\boldsymbol{\beta}^j_i(\lambda^{\mathrm{tr},1/2}_{\mathrm{opt},i})\|_1 > \beta^i_{\mathrm{tr},1/2}, \\ A^{\mathrm{GLG,opt};\boldsymbol{\beta}}_{\mathrm{tr},1/2}(i,j) =& 0 \quad \text{otherwise.} \end{aligned}$$

This matrix is presented in Fig. 2. One observes that the false causal relationships of the estimator $A^{\mathrm{GLG,opt}}_{\mathrm{tr},1/2}$ are actually weak. This observation suggests to use a second thresholding that is done on the network, or adjacency matrix, level.

We propose to do the thresholding on the network level similarly to the thresholding on the gene level. Namely, let us define the guideindicators of smallness on the network level similarly to (5):

$$\begin{aligned} A_{\min} =& \min\{\ A^{\mathrm{GLG,opt};\boldsymbol{\beta}}_{\mathrm{tr},1/2}(i,j) \neq 0\ \}, \\ A_{\max} =& \max\{\ A^{\mathrm{GLG,opt};\boldsymbol{\beta}}_{\mathrm{tr},1/2}(i,j)\ \}. \end{aligned}$$

And similarly to (6), define the threshold on the network level as follows:

$$A_{\mathrm{tr},\alpha_1} = A_{\min} + \alpha_1(\ A_{\max} - A_{\min}\ ). \tag{7}$$

We find it suitable to call the described combination of the two thresholdings on the gene and network levels as *2-levels-thresholding*. The adjacency matrix obtained by this thresholding strategy is the following:

$$\begin{aligned} A^{\mathrm{GLG,opt}}_{\mathrm{tr},1/2;\alpha_1}(i,j) =& 1 \quad \text{if} \quad A^{\mathrm{GLG,opt};\boldsymbol{\beta}}_{\mathrm{tr},1/2}(i,j) > A_{\mathrm{tr},\alpha_1}, \\ A^{\mathrm{GLG,opt}}_{\mathrm{tr},1/2;\alpha_1}(i,j) =& 0 \quad \text{otherwise.} \end{aligned}$$

It turns out that with $\alpha_1 = 1/4$ in (7) the optimal GLG-estimator can be fully recovered.

6 An automatic realization of the GLG-method

For an automatic realization of the GLG-method, i.e. a realization that does not rely on the knowledge of the true adjacency matrix A^{true}, in addition to a thresholding strategy one needs a choice rule for the regularization parameter λ in (4). For such a choice, we propose to use the so called quasi-optimality criterion [17, 2, 8]. Some details of the application of this criterion can be found in [12].

The reconstruction obtained by the GLG-method with the 2-levels-thresholding and quasi-optimality criterion $A^{\mathrm{GLG,qo}}_{\mathrm{tr},1/2;1/4}$ is presented in Fig. 2. Its quality measures can be found in Table 2. One observes that there is a little decrease in recall in comparison to the optimal GLG-method; however, this recall is the same as for the GtrLG-method (Table 1). But due to the highest precision, the F_1-score remains to be higher than for the methods in Table 1. Thus, one may say that the proposed realization of the GLG-method outperforms the methods in Table 1.

Nevertheless, one may still wonder, why the proposed realization of the GLG-method captures only the causal relationships of the most caused gene. It appears that the value of the maximal lag L plays an important role in the selection of the causal relationships.

In the modifications of the GLG-method the authors of [11, 15] considered $L = 3$. All results presented so far were also obtained with $L = 3$. It turns out that for $L = 4$ the optimal GLG-estimator (see Fig. 2) delivers a much better reconstruction of the causality network. In particular, two more caused genes are recovered.

The proposed automatic realization of the GLG-method with $L = 4$ (Fig. 2) recovers an additional caused gene in comparison to the realization with $L = 3$. Also, all considered quality measures for our automatic realization of the GLG-method with $L = 4$ (Table 2) are considerably higher than for the methods in Table 1. We would like to stress that no use of the knowledge of A^{true} is needed for obtaining $A^{\mathrm{GLG,qo}}_{\mathrm{tr},1/2;1/4}$, and no readjustment of the design parameters α, α_1 is necessary.

Table 2. Quality measures of the various GLG-estimators.

	Fs	P	R	F_1
GLG-opt	6.2%	1	0.44	0.62
GLG-opt; $\mathrm{tr}, 1/2$	14.8%	0.38	0.56	0.45
GLG-qo; $\mathrm{tr}, 1/2; 1/4$	7.4%	1	0.33	0.5
GLG-opt, $L = 4$	3.7%	0.88	0.78	0.82
GLG-qo;$L = 4$;$\mathrm{tr}, 1/2; 1/4$	7.4%	0.71	0.56	0.63

7 Conclusion and outlook

The proposed realization of the Graphical Lasso Granger method with 2-levels-thresholding and quasi-optimality criterion for the choice of the regularization parameter shows a considerable improvement of the reconstruction quality in comparison to other graphical methods. So, the proposed realization is a very promising method for recovering causality networks. Further tests and developments of the proposed realization are worthwhile. In particular, applications to larger causality networks are of interest.

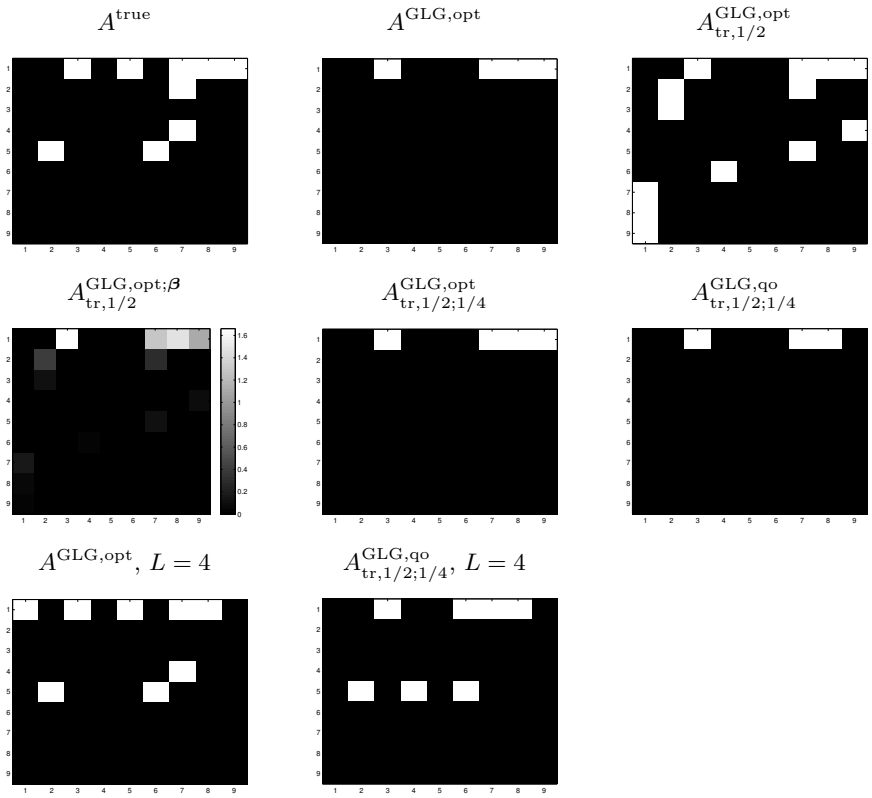

Fig. 2. The adjacency matrix A^{true} for the causality network in Fig. 1 and its various GLG-estimators.

As an open problem for the future, one could consider a study of the choice of the maximal lag and its possible variation with respect to the caused and causing genes.

References

1. A. Arnold, Y. Liu, and N. Abe. Temporal causal modeling with graphical Granger methods. In *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 66–75, USA, 2007. ACM New York.
2. F. Bauer and M. Reiß. Regularization independent of the noise level: an analysis of quasi-optimality. *Inverse Probl.*, 24(5):16 p., 2008.
3. I. Daubechies, M. Defrise, and C. De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Commun. Pure Appl. Math.*, 57(11):1413–1457, 2004.
4. H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of inverse problems.* Kluwer Academic Publishers, Dordrecht, 1996.

5. M. Fornasier, editor. *Theoretical foundations and numerical methods for sparse recovery.* Berlin: de Gruyter, 2010.
6. C. Granger. Investigating causal relations by econometric models and crossspectral methods. *Econometrica*, 37:424–438, 1969.
7. M. Grasmair, M. Haltmeier, and O. Scherzer. Sparse regularization with l^q penalty term. *Inverse Probl.*, 24(5):13, 2008.
8. S. Kindermann and A. Neubauer. On the convergence of the quasioptimality criterion for (iterated) Tikhonov regularization. *Inverse Probl. Imaging*, 2(2):291–299, 2008.
9. X. Li and et al. Discovery of time-delayed gene regulatory networks based on temporal gene expression profiling. *BMC Bioinformatics*, 7(26), 2006.
10. D. A. Lorenz, P. Maass, and Q. M. Pham. Gradient descent for Tikhonov functionals with sparsity constraints: theory and numerical comparison of step size rules. *Electronic Transactions on Numerical Analysis*, 39:437–463, 2012.
11. A. C. Lozano, N. Abe, Y. Liu, and S. Rosset. Grouped graphical Granger modeling for gene expression regulatory networks discovery. *Bioinformatics*, 25:110–118, 2009.
12. S. Pereverzyev Jr and K. Hlaváčková-Schindler. Graphical Lasso Granger method with 2-levels-thresholding for recovering causality networks. Technical report, University of Innsbruck, Department of Mathematics, Applied Mathematics Group, 2013.
13. R. Ramlau and G. Teschke. A Tikhonov-based projection iteration for nonlinear ill-posed problems with sparsity constraints. *Numer. Math.*, 104(2):177–203, 2006.
14. F. Sambo, B. D. Camillo, and G. Toffolo. CNET: an algorithm for reverse engineering of causal gene networks. In *NETTAB2008*, Varenna, Italy, 2008.
15. A. Shojaie and G. Michailidis. Discovering graphical Granger causality using the truncating lasso penalty. *Bioinformatics*, 26:i517–i523, 2010.
16. R. Tibshirani. Regression shrinkage and selection via the Lasso. *J. R. Stat. Soc. B*, 58:267–288, 1996.
17. A. N. Tikhonov and V. B. Glasko. Use of the regularization method in non-linear problems. *USSR Comp. Math. Math. Phys.*, 5:93–107, 1965.
18. M. L. Whitfield and et al. Identification of genes periodically expressed in the human cell cycle and their expression in tumors. *Mol. Biol. Cell*, 13:1977–2000, 2002.
19. Wikipedia. Causality — Wikipedia, The Free Encyclopedia, 2013. [Online; accessed 11-October-2013].