Learning Directed-Acyclic-Graphs from Large-Scale Double-Knockout Experiments

Fabio Nikolay, Marius Pesavento

Communication Systems Group, Technische Universität Darmstadt, Germany

Email: {fnikolay, pesavento}@nt.tu-darmstadt.de

Abstract—In this paper we consider the problem of learning the genetic-interaction-map, i.e., the topology of a directed acyclic graph (DAG) of genetic interactions from noisy double knockout (DK) data. Based on a set of well established biological interaction models we detect and classify the interactions between genes. Furthermore, we propose a novel linear integer optimization framework called Genetic-Interactions-Detector (GENIE) to identify the complex biological dependencies between genes and to compute the DAG topology that matches the DK measurements best where we make use of the well known branch-and-bound (BB) principle. Finally, we show via numeric simulations that the GENIE framework clearly outperforms the conventional techniques.

Index Terms—Genetic interactions analysis, large scale gene networks, discrete optimization, big data

I. Introduction

Genetic interaction analysis aims at uncovering the interactions among a set of genes with respect to a specified cell function of a biological system, e.g., bacteria. The interactions among the genes under study are well described by a directed-acyclic-graph (DAG) where the hierarchical relationship among two genes of a DAG describe their hierarchical interaction type. However DAGs cannot be observed directly but only the specified cell function under study which yields observable phenotypes. The role of the studied genes in the cell machinery, the hierarchical interaction types of the genes, as well as the DAG, which describes the latter ones, can only be learned by means of knock-out experiments where a gene or a set of genes is functionally switched off and the phenotype is observed. Traditionally, only single-knockout (SK) experiments haven been conducted but those mainly allow for a statement about the importance of a single gene for the investigated cell process and do not convey much information about the interaction among the genes under study. Recently with the technological advances in micro arrays and the development of the synthetic-genetic-array technologies [7] new approaches have been taken that are based on large scale knock-out experiments of pairs of genes. Such double knock-out (DK) experiments are much more powerful for exploring genetic interactions since a DK phenotype of an arbitrary pair of genes generally differs considerably from the superposition of the corresponding SK phenotypes of this pair of genes. According to [6], the gene pairs can be classified to one out of five hierarchical relationship classes based on their SK and DK phenotypes. Further, based on the hierarchical relationship classes the DAG underlying the observed SK and DK phenotypes can be computed which directly reflects the genetic interactions among the genes. In order to detect the DAG underlying the SK and DK phenotypes a variety of statistical methods based on scoring the measurements or on Pearson correlation of the genetic-interaction (GI)-profile, e.g. [1]-[5] respectively, have been developed. Although showing poor performance, methods as those presented in [1]-[5] are the most commonly used methods to detect the DAG underlying the measured DK data. In contrast to that, we propose an approach that is based on the biological system model of [6]. Since the hierarchical relationship classes are mutually dependent classifying each pair of genes to a specific hierarchical relationship class corresponds to a multi-hypotheses. Thus, we formulate this multi-hypotheses test as a linear integer optimization program in order to find the set of hierarchical relationship classes, best matching the observed SK and DK phenotypes. Based on the detected set of hierarchical relationship classes, the set of edges of the DAG which reflects the interactions among the genes can be computed.

This paper is organized as follows. We first summarize the biological system model of [6], present in Section 3 the Genetic-Interactions-Detector (GENIE)-framework of detecting the set of hierarchical relationship classes, that represents a valid DAG and matches the DK measurements best, and continue with presenting numerical results which demonstrate the performance of our algorithm. Finally, we summarize the key parts of this paper and give a brief outlook on future work.

II. System Model

In this section we first provide a mathematical description of a DAG as well as its biological implications. Furthermore, we introduce the common biological terms and follow with the model of [6] where we provide an intuitive interpretation of how to read and interpret a DAG of a genetic interaction map.

The functional dependencies among a set of genes $\mathcal{G} = \{g_1, ..., g_G\}$, with $G = |\mathcal{G}|$ elements, for a given cell process and specie can be well described by a genetic-interaction-map (GI-map,[9],[10]) which is essentially a DAG with a common root node, i.e., the reporter level R. In particular, an arbitrary DAG $\mathcal{D}$ can be described as a graph $\mathcal{D} = (\mathcal{G}_\mathcal{D}, \mathcal{E}_\mathcal{D})$ with the set of nodes $\mathcal{G}_\mathcal{D} = \{\mathcal{G} \cup R\}$ and the set of directed edges $\mathcal{E}_\mathcal{D} = \left\{ \{g_i, g_j\}, ..., \{g_j, g_l\} \right\}$ where edge $\{g_i, g_j\}$ denotes an edge from gene g_i to gene g_j. As the genetic interactions can

$k = 1:$ $k = 2:$ $k = 3:$ $k = 4:$ $k = 5:$

$\mu_1(i, j) = R(j)$ $\mu_2(i, j) = R(i)$ $\mu_3(i, j) = R(i)R(j)$ $\mu_4(i, j) = \frac{1}{2}[R(i)R(j) + R(j)]$ $\mu_5(i, j) = \frac{1}{2}[R(i)R(j) + R(i)]$
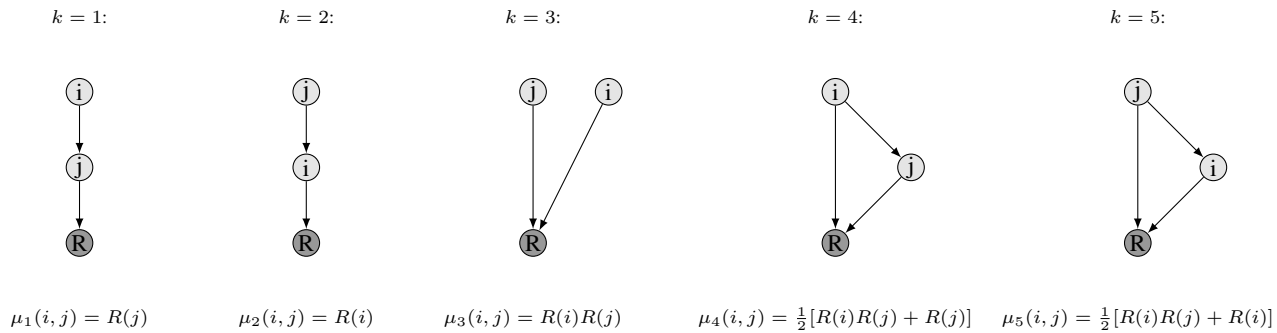
Fig. 2: Possible hierarchical relationship classes between two arbitrary genes i, j of DAG $\mathcal{D}$ according to [6]
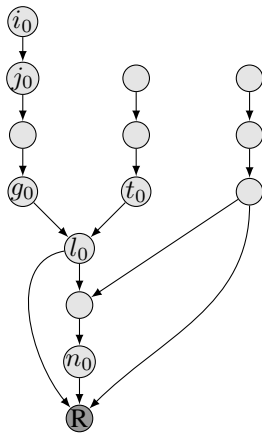
Fig. 1: DAG $\mathcal{D}_0$ of 13 genes and root node R

only be observed through the reporter all edges are always orientated in such a way that each path parting from any arbitrary gene $g_i \in \mathcal{G}$ always terminates in the root node R and any gene appears on the path at most once, i.e., there exist no cycles in the graph. Hence, the DAG $\mathcal{D}$ is always connected via its root node R. For the sake of notational convenience, in most cases we write gene i when addressing gene g_i. The reporter node R is an artifical node in the concept of a DAG and represents the measured phenotype of the specific cell process under study. The term phenotype generally describes the specific manifestation of a biological attribute of an organism which can be observed, e.g., for bacteria a common biological attribute is growth measured in colony size, where a specific size of the bacteria colony is a phenotype of this biological attribute. In order to provide a better understanding of the information encoded in a DAG we state a simple example based on DAG $\mathcal{D}_0$ displayed in Fig. 1. In $\mathcal{D}_0$ there exists a direct edge from gene i_0 to gene j_0, i.e. $\{i_0, j_0\} \in \mathcal{E}_{\mathcal{D}_0}$, which indicates that the activity of gene i_0 controls the activity of gene j_0. Hence, gene i_0 only affects the phenotype via gene j_0 and not directly. We emphasize that in this model the existence of edge $\{i_0, j_0\}$ in the DAG only describes the functional dependency between genes i_0 and j_0

and not the quantitative effect of gene i_0 on gene j_0.

Denote $R(i) \in \mathbb{R}$ as the phenotype for a single gene $i \in \mathcal{G}$ functionally switched off. In the same fashion we define the phenotype for the DK of genes $i, j \in \mathcal{G}$: $j > i$ as $R(i, j) \in \mathbb{R}$. According to [6] we assume that each pair of genes $i, j \in \mathcal{G}$:$j > i$ belongs to exactly one out of five hierarchical relationship classes that are characterized in Fig. 2. The hierarchical relationship classes $k \in \mathcal{K} = \{1, ..., 5\}$ are defined according to the model $\mu_k(i, j)$ in which the single knock-out phenotypes $R(i)$ and $R(j)$ are related with the DK phenotype $R(i, j)$. If the gene pair $i, j \in \mathcal{G}$: $j > i$ belongs to the hierarchical relationship class k then the observed DK phenotype $R(i, j)$ is described by the model $\mu_k(i, j)$ of Fig. 2. We remark that the five hierarchical dependency graphs in Fig. 2 do not reflect the absolute adjacency relations, but the hierarchical relations between genes i, j in DAG $\mathcal{D}$. To clarify this further, let us consider the example DAG $\mathcal{D}_0$ of Fig. 1. All paths from gene i_0 to node R pass through gene j_0, i.e., they are in a linear pathway with gene i_0 upwards of gene j_0. Thus the pair of genes i_0, j_0 belongs to class $k = 1$. Note that with the same line of argument, we conclude that also genes i_0 and l_0 belong to relationship class $k = 1$.

Since all paths from gene i_0 to the reporter level R do not pass through gene t_0 and vice versa, genes i_0 and t_0 belong to the hierarchical relationship class $k = 3$ as given in Fig. 2, which states that genes i_0 and t_0 are independent of each other and the DK phenotype amounts to $R(i_0, t_0) = \mu_3(i_0, t_0)$. Finally, let us inspect the structural relation between genes t_0 and n_0 in DAG $\mathcal{D}_0$. It is obvious that gene t_0 has (at least) one path to node R which does not pass through gene n_0, i.e., genes only having paths to R that do not pass through gene n_0 do not affect the activity of gene n_0. Since there is (at least) one other path from gene t_0 to R passing through gene n_0, we can conclude that genes t_0 and n_0 belong to class $k = 4$. Generally, there are strong implications among the hierarchical relationship classes of [6], i.e., if some pairs belong to a specific class then this has strong implications for all other pairs. Let us consider the case that DAG $\mathcal{D}_0$ was not known and only the hierarchical relationship classes for genes i_0 and j_0, i.e., genes i_0 and j_0 belong to class $k = 1$, as well as the hierarchical relationship class for genes i_0 and g_0, i.e.,

genes i_0 and g_0 belong to class $k = 1$ were available. By definition of the hierarchical dependency graphs in Fig. 2 and the assumptions, that genes i_0 and j_0 belong to class $k = 1$ as well as that genes i_0 and g_0 belong to class $k = 1$, we conclude that all paths from gene i_0 to R pass through genes j_0 and g_0. Thus, either all paths from gene g_0 to R pass through gene j_0, or vice versa. Consequently, genes j_0 and g_0 either belong to the hierarchical relationship class $k = 1$, or $k = 2$. As we have emphasized by the example above, generally if the hierarchical relationship class is known for two arbitrary genes $i, j \in \mathcal{G} : j > i$ as well as for another pair $i, l \in \mathcal{G} : l > i$, then this has strong logical implications on the hierarchical relationship classes genes $j, l \in \mathcal{G} : l > j$ can belong to. Since we can interpret the classification of the pairs of genes $i, j \in \mathcal{G} : j > i$, based on their observed SK and DK phenotypes $R(i), R(j)$ and $R(i, j)$, respectively, to exactly one out of the five hierarchical relationship classes as a coupled multi-hypotheses test, we address this problem in the next section by a linear integer optimization program. The proposed linear integer optimization program identifies the set of hierarchical relationship classes, that represents a valid DAG and matches the DK measurements best with respect to the logical coupling between the classes.

III. GENIE-FRAMEWORK

In this section, we formulate the problem of classifying the gene pairs $i, j \in \mathcal{G} : j > i$ into the classes of hierarchical relationships based on the observed SK and DK phenotype values as a linear integer optimization program. Furthermore, we translate the logical implications among the hierarchical relationship classes into constraints that ensure that the detected set of hierarchical relationship classes represents a valid graph. That is the detected set of hierarchical relationship classes represents a graph which is a DAG as defined in Section 2. In order to quantify the mismatch between the measured DK phenotypes $R(i, j)$ and the expected phenotype $\mu_k(j, j)$ under the hypothesis that the gene pairs $i, j \in \mathcal{G} : j > i$ belong to class $k \in \mathcal{K}$ given their respective SK values, we propose a simple quadratic score [6] as given in Eq. (1):

$$s_k(i, j) = \big(R(i, j) - \mu_k(i, j) \big)^2 \quad k \in \mathcal{K}, \quad \forall i, j :\in \mathcal{G} : j > i \tag{1}$$

Let us define the following selection variables

$$\alpha_k(i, j) = \begin{cases} 1 & \text{if } i, j \text{ are in class } k \\ 0 & \text{else} \end{cases}$$
$$k \in \mathcal{K}, \quad \forall i, j :\in \mathcal{G} : j > i \tag{2}$$

Then the GENIE problem of classifying the gene pairs $i, j \in \mathcal{G} : j > i$ into the set of hierarchical relationship classes, that represents a valid DAG and matches the DK measurements

best can be formulated as

O_{GENIE} :

$$\min_{\{\alpha_k(i,j)\}} \quad \sum_{i=1}^{G} \sum_{j=i+1}^{G} \Big(\sum_{k=1}^{|\mathcal{K}|} s_k(i, j) \alpha_k(i, j) \Big) \tag{3a}$$

$$\text{s. t.} \quad \alpha_k(i, j) \in \{0, 1\} \ \forall k \in \mathcal{K}, \ \forall i, j \in \mathcal{G} : j > i \tag{3b}$$

$$\sum_{k=1}^{|\mathcal{K}|} \alpha_k(i, j) = 1, \ \forall i, j \in \mathcal{G} : j > i \tag{3c}$$

$$\mathcal{L} \implies \text{additional topology constraints} \tag{3d}$$

where $A^{O_{\text{GENIE}}} = \bigcup_{\forall i, j \in \mathcal{G}_\mathcal{D} : j > i} \Big\{ \alpha_1^{O_{\text{GENIE}}}(i, j), ..., \alpha_5^{O_{\text{GENIE}}}(i, j) \Big\}$ denotes the solution to program O_{GENIE} and the set of best matching selection variables $A^{O_{\text{GENIE}}}$ corresponds to the most consistent pattern of hierarchical relationship classes. Program O_{GENIE} is a linear integer program which can be solved efficiently by BB-methods [11]. The objective of program O_{GENIE} is to minimize the overall mismatch in classifying each gene pair $i, j \in \mathcal{G} : j > i$ to one out of five hierarchical relationship classes. The constraints in (3b) reflect the binary nature of the selection variables, while (3c) represents a multiple choice constraint that enforces the gene pairs i, j to select only one hierarchical relationship class.

The set $\mathcal{L}$ in (3d) contains additional constraints to ensure that the detected set of selection variables $A^{O_{\text{GENIE}}}$ always represents a valid graph, i.e., a DAG. In the following we exemplarily derive topology constraints in set $\mathcal{L}$. In order to identify the numerous logical implications among the selection variables $\alpha_k(i, j), k \in \mathcal{K}$ for all $i, j \in \mathcal{G} : j > i$ we proceed in the following way. We first fix the assumption that genes $i, j \in \mathcal{G} : j > i$ belong to class $k = 1$, i.e., $\alpha_1(i, j) = 1$. Further we assume that genes $i, l \in \mathcal{G} : l > i$ belong to class k', i.e. $\alpha_{k'}(i, l) = 1$. Then we derive the set of classes $\mathcal{K}''$ that genes $j, l \in \mathcal{G} : l > j$ can belong to under the assumptions made. In the following, we have formulated the logical implications among the selection variables for $\alpha_1(i, j) = 1$ as linear integer inequalities defined in constraints (4a)-(4e) and summarize them as set $\mathcal{L}_1$

$$\mathcal{L}_1 = \Bigg\{$$
$$\alpha_1(j, l) + \alpha_2(j, l) \geq \alpha_1(i, j) + \alpha_1(i, l) - 1 \tag{4a}$$
$$\alpha_2(j, l) \geq \alpha_1(i, j) + \alpha_2(i, l) - 1 \tag{4b}$$
$$\alpha_2(j, l) + \alpha_3(j, l) + \alpha_5(j, l) \geq$$
$$\alpha_1(i, j) + \alpha_3(i, l) - 1 \tag{4c}$$
$$\alpha_2(j, l) + \alpha_4(j, l) \geq \alpha_1(i, j) + \alpha_4(i, l) - 1 \tag{4d}$$
$$\alpha_5(j, l) + \alpha_2(j, l) \geq \alpha_1(i, j) + \alpha_5(i, l) - 1 \tag{4e}$$
$$\Bigg\} \forall i, j, l \in \mathcal{G}_\mathcal{D} : l > j > i$$

where constraints (4a)-(4e) are convex after the continuous relaxation of the selection variables $\alpha_k(i, j), \forall i, j \in \mathcal{G} : j > i$. To explain the origin and the functionality of the constraints in $\mathcal{L}_1$, let us further define a sub-genetic-interactions-map

(SMAP) $\mathcal{S}$ as given in Fig. 3 according to the following definition where we adopt the graph notation of [8]:

Definition:

Given a non-empty set of edges $\mathcal{E}_{\text{in}}$ and a non-empty set of edges $\mathcal{E}_{\text{out}}$. Graph $\mathcal{S} = \left(\mathcal{G}_\mathcal{S}, \mathcal{E}_\mathcal{S} \right)$, with set of nodes $\mathcal{G}_\mathcal{S}$ and set of edges $\mathcal{E}_\mathcal{S}$, is a SMAP if the following conditions are fulfilled: $i)$ the graph $\mathcal{S}$ is acyclic and directed $ii)$ there $\exists e_{\text{in}} \in \mathcal{E}_{\text{in}}, e_{\text{out}} \in \mathcal{E}_{\text{out}}$ such that each path P through graph $\mathcal{S}$ incides $\mathcal{S}$ via egde e_{in} and leaves graph $\mathcal{S}$ via edge e_{out}.
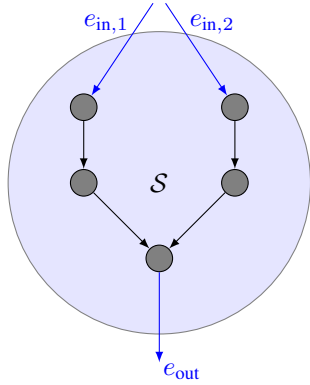

Fig. 3: Example sub-GI-map $\mathcal{S}$

The DAG $\mathcal{D}_1$, as displayed on the left-hand-side (LHS) Fig. 4, consists of genes i, j and SMAPs $\mathcal{S}_1$ and $\mathcal{S}_2$. It is obvious that genes i, j belong to class $k = 1$, i.e., $\alpha_1(i, j) = 1$. Furthermore, all genes $l \in \mathcal{G}_{\mathcal{D}_1} \setminus \{R\} : l > j > i$ for which $\alpha_1(i, l) = 1$ must be either located in SMAP $\mathcal{S}_1$ or $\mathcal{S}_2$. Hence it follows from DAG $\mathcal{D}_1$ in Fig. 4 that the gene pair j, l is either in hierarchical relationship class $k = 1$ or $k = 2$, i.e., $\alpha_1(j, l) = 1$ or $\alpha_2(j, l) = 1$.

This logical implication is directly reflected by constraint (4a). Given $\alpha_1(i, j) = 1$ and $\alpha_1(i, l) = 1$ the right-hand-side (RHS) of (4a) amounts to 1. In this case also the (LHS) of (4a) becomes 1 to fulfill the inequality (4a). Thus either $\alpha_1(j, l) = 1$ or $\alpha_2(j, l) = 1$. Reversely, assume that $\alpha_1(i, j) = 1$ and $\alpha_1(i, l) = 1$ does not hold, then the RHS of (4a) is less than 1, i.e., 0 or -1, while the LHS of (4a) is always greater than 0. Hence, constraint (4a) is fulfilled irrespectively of the choice of $\alpha_k(j, l)$, i.e., constraint (4a) enforces no logical implications.

Similarly for DAG $\mathcal{D}_2$ of the RHS of Fig. 4, it is obvious that genes i, j belong to the hierarchical relationship class $k = 1$, i.e., $\alpha_1(i, j) = 1$. All genes $l \in \mathcal{G}_{\mathcal{D}_2} \setminus \{R\} : l > j > i$ which are in a linear pathway upstream of gene i, i.e. $\alpha_2(i, l) = 1$, must be located in SMAP $\mathcal{S}_3$. Hence it directly follows from DAG $\mathcal{D}_2$ that also gene l must be in a linear pathway upstream of gene j, i.e., $\alpha_2(j, l) = 1$. This logical implication is compactly represented in constraint (4b). Under the assumption that $\alpha_1(i, j) = 1$ and $\alpha_2(i, l) = 1$, the RHS of (4b) amounts to 1 enforcing $\alpha_2(j, l) = 1$, so that the LHS of (4b) equals the RHS and the inequality in (4b) is fulfilled. Reversely, assume that $\alpha_2(i, l) = 0$, then the RHS of (4b) is less than 1 and

hence the LHS of (4b) is always bigger than or equal to the RHS irrespectively of the choice of $\alpha_k(j, l)$, i.e., constraint (4a) enforces no logical implications. We can proceed in the same fashion to explain constraints (4c)-(4e). Furthermore, with the same line of argument we can derive the sets $\mathcal{L}_k$ for $k \in \mathcal{K} \setminus 1$ which reflect the logical implications among the selection variables under the assumptions that $\alpha_k(i, j) = 1$ for $k \in \mathcal{K} \setminus 1$. However, due to space limitations we omit the derivation of the full set of logical implications and refer the interested reader to a future journal version of this paper. The
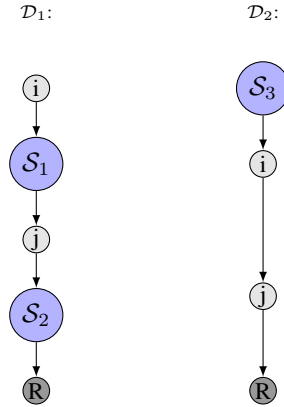

Fig. 4: Schematically reduced DAGs

full set of topology constraints $\mathcal{L}$ in (3d) can be computed as

$$\mathcal{L} = \bigcup_{k=1}^{|\mathcal{K}|} \{\mathcal{L}_k\}. \tag{5}$$

The totality of the computed selection variables, i.e., the set $A^{\text{O}_{\text{GENIE}}}$, represents DAG $\mathcal{D}$ underlying the observed SK and DK phenotpyes in a different domain. It can be theoretically proven that this representation is not unique, i.e., the computed selection variable pattern $A^{\text{O}_{\text{GENIE}}}$ represents not only the true DAG $\mathcal{D}$, but also a set of similar DAGs which are in line with the pattern $A^{\text{O}_{\text{GENIE}}}$, but have a minorly different set of edges as the true DAG $\mathcal{D}$. Based on the detected set of selection variables, an estimate $\hat{\mathcal{E}}_\mathcal{D}$ of the true set of edges $\mathcal{E}_\mathcal{D}$ of DAG $\mathcal{D}$ can be computed.

IV. SIMULATION RESULTS

We have generated the ideal SK phenotypes $R(i) \in \mathbb{R}$ for all $i \in \mathcal{G}$ as well as the ideal DK phenotypes $R(i, j) \in \mathbb{R}$ for all $i, j \in \mathcal{G} : j > i$ according to the model of [6]. We compare our method with the well known GI-profile similarity approach ([12],[6]), where the Pearson correlation between the GI-profiles of genes i and j is computed and an edge in the DAG is detected if the Pearson correlation is above a pre-defined threshold t_{corr}, where the directionality is inferred from the selection variable $\alpha_k(i, j)$ corresponding to the least mismatch model $\mu_k(i, j)$. Furthermore, we compare our method with the solution of program O_{GENIE} without considering set $\mathcal{L}$ as a constraint, which means simply classifying each pair i, j to the least mismatch scoring hierarchical relationship class based

on the SK and DK phenotypes $R(i)$ and $R(i,j)$ respectively. For the simulations, we consider a total of 10 genes in order to limit the Monte-Carlo simulation time. In Fig. 5 we display the percentage of erroneously detected edges D_{ed} in the detected DAG normalized to the true number of edges $|\mathcal{E}_\mathcal{D}|$ as defined in Eq. (6) versus the SNR.

$$D_{\text{ed}} = \frac{\left(\mathcal{E}_\mathcal{D} \bigcup \hat{\mathcal{E}}_\mathcal{D}\right) \setminus \mathcal{E}_\mathcal{D}}{|\mathcal{E}_\mathcal{D}|} \qquad (6)$$

In Fig. 5 we observe that in the low SNR regime, the Pearson correlation based method performs best in terms of erroneously detected edges, however it fails to improve performance with increasing SNR, because for correct directionality information of the edges this approach relies on the hierarchical relationship classes detected by method O$_{\text{GENIE}}$ without considering $\mathcal{L}$. In the high SNR regime, the proposed GENIE-approach clearly outperforms the Pearson correlation based method as well as the method which simply detects the highest scoring class regardless of their coupling with other classes. In Fig. 6 we display the percentage of missing edges D_{mis} in the detected DAG normalized to the true number of edges $|\mathcal{E}_\mathcal{D}|$ as defined in Eq. (7) versus the SNR, i.e.,

$$D_{\text{mis}} = \frac{\left(\mathcal{E}_\mathcal{D} \bigcup \hat{\mathcal{E}}_\mathcal{D}\right) \setminus \hat{\mathcal{E}}_\mathcal{D}}{|\mathcal{E}_\mathcal{D}|} \qquad (7)$$

In Fig. 6 we observe that our proposed GENIE-approach clearly outperforms the competing methods over the entire SNR region. Particularly, the Pearson correlation based method shows poor performance.
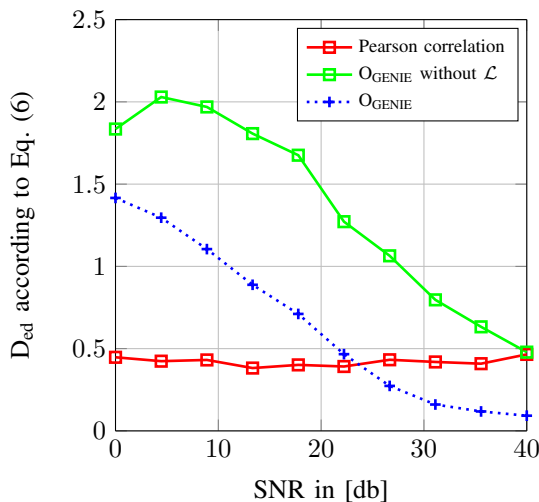

Fig. 5: D_{ed} versus SNR; $t_{\text{corr}} = .8$; 200 Monte Carlo runs

V. CONCLUSION AND OUTLOOK

In this paper, we have demonstrated the benefits of the GENIE-framework over conventional techniques for detecting the topology of the DAG which is underlying the SK and DK data measured. Especially in the high SNR-regime, our proposed algorithm outperforms the conventional techniques
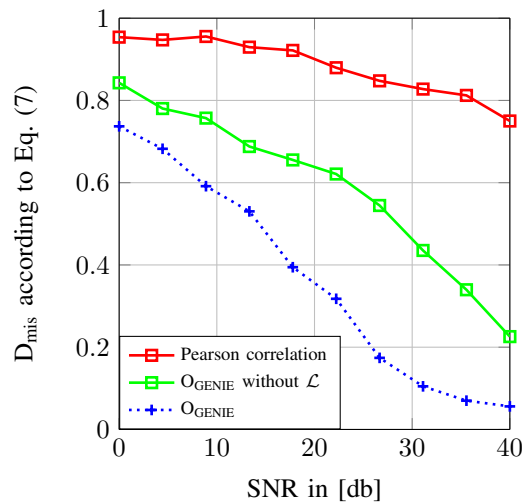

Fig. 6: D_{mis} versus SNR; $t_{\text{corr}} = .8$; 200 Monte Carlo runs

on simulated data. Future work may consider a robust modification of the presented GENIE-approach with respect to noise as well as an incorporation of correlation data into the GENIE-framework in order to enhance the reliability of the detection results even further.

REFERENCES

[1] B. Snijder, P. Liberali, M. Frechin, T. Stoeger and L. Pelkmans, *Predicting functional gene interactions with the hierarchical interaction score*, Nature Methods, 2013
[2] A. Baryshinkova et al. *Quantitative analysis of fitness and genetic interactions in yeast on a genome scale*, Nature Methods, 2010
[3] S.R. Collins, A. Roguev, N.J. Krogan *Quantitative Genetic Interaction Mapping Using the E-MAP Approach*, Methods Enzymol., 2010
[4] R.O. Linden, V.P. Eronen, T. Aittokallio *Quantitative maps of genetic interactions in yeast - Comparative evaluation and integrative analysis*, BMC Systems Biology, 2011
[5] S.J. Dixon, M. Constanzo, A. Baryshinkova, B. Andrews, C. Boone *Systematic Mapping of Genetic Interaction Networks*, The Annual Review of Genetics, 2009
[6] A. Battle, M.C. Jonikas, P. Walter, J.S. Weissman and D. Koller *Automated identification of pathways from quantitative genetic interaction data*, Molecular Systems Biology, 2010
[7] A.H.Y. Tong et al. *Systematic genetic analysis with ordered arrays of yeast deletion mutants*, Science, 2001
[8] R. Diestel *Graphentheorie*, Springer-Verlag, Heidelberg, ISBN 978-3-642-14911-5, 2012
[9] A. Jaimovich et al. *Modularity and directionality in genetic interaction maps*, Nature Methods, 2010
[10] A. Baryshinkova, M. Constanzo, C.L. Myers, B. Andrews, C. Boone *Genetic Interaction Networks: Toward an Understanding of Heritability*, Annual Review of Genomics and Human Genetics, 2013
[11] V. Balakrishnan, S. Boyd, S. Balemi *Branch and bound algorithm for computing the minimum stability degree of parameter-dependent linear systems*, International Journal of Robust and Nonlinear Control,1(4):295317, 1991
[12] M. Babu et al. *Quantitative Genome-Wide Genetic Interaction Screens Reveal Global Epistatic Relationships of Protein Complexes in Escherichia coli*, PLOS Genetics, 2014