2013 European Control Conference (ECC)
July 17-19, 2013, Zürich, Switzerland.

Network reconstruction using knock-out and over-expression data

David Hayden, Ye Yuan† and Jorge Gonçalves

Abstract— **This paper outlines necessary and sufficient conditions for network reconstruction of linear, time-invariant systems using data from either knock-out or over-expression experiments. These structural system perturbations, which are common in biological experiments, can be formulated as unknown system inputs, allowing the network topology and dynamics to be found. We assume that only partial state measurements are available and propose an algorithm that can reconstruct the network at the level of the measured states using either time-series or steady-state data. A simulated example illustrates how the algorithm successfully reconstructs a network from data.**

I. Introduction

For the vast majority of systems, it is not possible to measure every single state. Hence, while keeping in mind the existence of hidden states, we focus our attention on identifying networks between measured states. Here we would like to know exactly how measured states affect each other in a causal way. For example, if A affects B which in turn affects C, the causal network structure is given by $A \to B \to C$ and not $A \to B$ and $A \to C$ (one may be tempted to write it this way since A does affect both B and C, though A affects first B, which then affects C).

A particular focus on this problem has grown most recently in the systems biology community. Molecules inside a cell form a complex network with multiple feedback loops. Once an organism's DNA has been sequenced, the next step is to understand the functionality of the identified genes: what regulates what and how? This complex network is critical for understanding how the system works and hence for constructing a map from genotype to phenotype [1].

Motivated by this problem, there now exist a large number of network reconstruction tools [2]. These include information-theoretical (e.g. ARACNE [3]), Bayesian inference (e.g. Banjo [4]) and ODE-based methods (e.g. [5]–[9]). However, the first cannot identify causality, while the second assumes that the time-series measurements are independent and identically distributed (iid). Many methods based on ODE models assume that all states are measured or impose restrictions on the network topology, such as sparsity, in order to obtain a solution. If these assumptions are incorrect, there is no guarantee that the correct network will be found or even that the network obtained will be close to the true one.

Recent work [10] characterised the exact amount of information required for network reconstruction, for networks of Linear, Time-Invariant (LTI) systems driven by known external inputs. This work was extended to include systems with noise and nonlinearities [11], and with algorithms solved in polynomial-time [12]. However, this framework did not include some of the most common types of biological experiments currently performed: gene knock-out and over-expression, which is the topic of this present paper.

In a gene knock-out, the gene is either deleted from the DNA or its protein is rendered non-functional, either through genetic modification or by treatment with a reagent such as a short DNA or RNA oligonucleotide with a sequence complementary to the mRNA transcript or gene. In gene over-expression, the gene is constitutively expressed, i.e. it is constantly being produced. In both gene knock-out and over-expression, no transcription factor regulates the gene, effectively removing its inputs from the rest of the network.

A similar problem was considered recently in [8] for consensus networks, where a sufficient condition for network reconstruction was given. This required every state and every pair of states to be knocked-out, which is more restrictive than the conditions we impose here. Thus, the major theoretical contribution of this paper is to provide necessary and sufficient conditions for network reconstruction using such mutant data.

Finally, it is important to point out that the scope of this paper is not limited to biological systems. Applications include distributed network control of multi-agent systems and dynamics in social networks [13]–[15]. In some biologically-inspired phenomena, such as flocking [16] and synchronisation [17], to achieve a desired collective behaviour, agents transmit their state to their neighbours and update their own values accordingly. Knowledge of the dynamical network structure is key to understanding both the network topology and dynamic interactions between agents.

Notation

For a matrix A, $A(i,j)$ denotes the element in the i^{th} row and j^{th} column, $A(i,:)$ denotes its i^{th} row and $A(:,j)$ its j^{th} column. Similarly, for a vector x, the i^{th} element is denoted $x(i)$.

II. Preliminaries

We consider networks characterised by the following linear, time-invariant (LTI) system representation, which may be obtained by linearization of a nonlinear system:

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= \begin{bmatrix} I & 0 \end{bmatrix} x \end{aligned} \tag{1}$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$ and $y \in \mathbb{R}^p$. The network *structure* is completely specified by the matrices A and B but is in

Control Group, Department of Engineering, University of Cambridge, UK. †For correspondence, yy311@cam.ac.uk.

general not obtainable from input-output data alone. Given the transfer function, G, from inputs to measured states, there are many possible state-space realisations.

From measured data it is possible to consistently estimate transfer functions and Laplace transforms using system identification techniques [18], [19]. Whilst there are significant practical considerations and requirements here (some of which are discussed in Section IV), this aspect of the problem is *identifiable* and we seek to determine further conditions on the identifiability of structural network information from this unstructured data. Specifically, we seek to infer the direct causal dependencies between the p measured states, y, where typically $p < n$ and n itself is unknown.

A. Dynamical Structure Functions

Partition (1) as follows:

$$\begin{aligned} \begin{bmatrix} \dot{y} \\ \dot{z} \end{bmatrix} &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u \\ y &= \begin{bmatrix} I & 0 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} \end{aligned} \tag{2}$$

where $x = [y^T \ z^T]^T$ is the full state vector and $z \in \mathbb{R}^{n-p}$ are the $n-p$ "hidden" states. Taking the Laplace transform of (2) and eliminating Z (where uppercase letters denote Laplace transforms of their lowercase counterparts) yields $sY = WY + VU$, where $W = A_{11} + A_{12}\left(sI - A_{22}\right)^{-1} A_{21}$ and $V = A_{12}\left(sI - A_{22}\right)^{-1} B_2 + B_1$. Now, letting D be the diagonal matrix composed of the diagonal elements of W, we write $\left(sI - D\right)Y = \left(W - D\right)Y + VU$ and rearrange:

$$Y = QY + PU \tag{3}$$

where $Q = \left(sI - D\right)^{-1}\left(W - D\right)$ and $P = \left(sI - D\right)^{-1} V$.

Given the system (2), the *dynamical structure function* is defined as (Q, P). If all the measured states are removed from the system except for $Y(i)$ and $Y(j)$ then $Q(i,j)$ corresponds to the transfer function between $Y(j)$ (considered as input) and $Y(i)$ (considered as output). The same holds for P in terms of $U(j)$ and $Y(i)$. Note that the effect of subtracting DY above is to fix the diagonal elements of Q to be zero.

It is straightforward to show that Q and P are related to G as follows:

$$G = (I - Q)^{-1} P \tag{4}$$

and that the solution (Q, P) for a given G is in general not unique. The problem of inferring (Q, P) from G was the subject of [10] and the main results are summarised in the following subsection.

B. Network Reconstruction

Under the assumption that G is full rank, the following theorem gives necessary and sufficient conditions for network reconstruction:

Theorem 1 ([10]). *The dynamical structure function, (Q, P), can be obtained uniquely from the transfer matrix, G, if and only if p elements in each row of $[Q \ P]$ are known.*

Note that if G is not full rank, then at least one of the input experiments is redundant and hence provides no new information. Redundant columns of G may simply be removed until the remaining columns are full rank.

Since Q has zeros on the diagonal, this constitutes one known element for each row, hence a further $p - 1$ known elements are needed. If nothing is known about the network structure, Q, then these known elements must come from the input structure, P. By designing targeted input experiments, we can choose the structure of P and therefore satisfy this condition. However, assuming that we do not know the nonzero transfer functions $P(i,j)$, we must rely on zero elements of P. The following corollary outlines a protocol for reconstruction in this manner:

Corollary 1 ([10]). *If $m = p$, G is full rank and nothing is known about the internal structure, Q, then the dynamical structure function, (Q, P), can be obtained uniquely from G if P is square and diagonal. In this case, the elements of Q and P are given by:*

$$Q(i,j) = -\frac{H(i,j)}{H(i,i)} \qquad P(i,i) = \frac{1}{H(i,i)} \tag{5}$$

where $H = G^{-1}$.

The knowledge that each input directly and uniquely affects a single measured state provides sufficient knowledge of the zero elements of P to satisfy the condition of Theorem 1.

C. Special Case: Full State Measurement

It is common in the network reconstruction literature to assume that all the system states are available for measurement and this greatly simplifies the problem. For completeness here we demonstrate this special case within our framework. The system equations are given as in (2) with $p = n$:

$$\begin{aligned} \dot{x} &= A_{11} x + B_1 u \\ y &= Ix \end{aligned} \tag{6}$$

Given the transfer function, G, from U to Y, the state space realisation with $C = I$ is unique. To see this, note that any state transformation (eg. $q = Tx$) would change the output matrix to IT^{-1}, hence the only permissible transformation to retain $C = I$ is $T = I$. From G, obtain any realisation:

$$\begin{aligned} \dot{q} &= A_q q + B_q u \\ y &= C_q q \end{aligned} \tag{7}$$

The true state-space matrices can then be computed:

$$A_{11} = C_q A_q C_q^{-1}, \qquad B_1 = C_q B_q \tag{8}$$

Hence if there are no hidden states, knowledge of G is sufficient to recover the entire network topology. However, the presence of a single hidden state nullifies this result [10]. Since in practice one may not even know the number of states in a system, let alone be able to measure all of them, we return to the general case.

III. Main Results

This section addresses a number of issues with the approach outlined in Section II; firstly, knowledge of G requires knowledge of the inputs, U, which we may not have; secondly, although we are applying targeted inputs, there may be extraneous inputs to the system, also unknown and beyond our control; thirdly, we may not be able to apply inputs to the system as such, but rather targeted perturbations (such as a knock-out), which do not fit directly into the existing framework.

A. Extraneous and Unmeasured Inputs

Consider first the ideal case for network reconstruction under the conditions of Corollary 1, where we have performed p input experiments and have perfect knowledge of the Laplace transforms of the measured states and the inputs. Let Y_i and U_i denote the Laplace transforms of the measured states and inputs respectively for the i^{th} experiment, which are related by: $Y_i = QY_i + PU_i$. Combine the data for all p experiments as follows:

$$\left[Y_1 Y_2 \dots Y_p\right] = Q\left[Y_1 Y_2 \dots Y_p\right] + P\left[U_1 U_2 \dots U_p\right] \quad (9)$$

and note that P and also $\mathbf{U} = [U_1\ U_2\ \dots\ U_p]$ must be diagonal. By forming $G = \mathbf{Y}\mathbf{U}^{-1}$, where $\mathbf{Y} = [Y_1\ Y_2\ \dots\ Y_p]$, we may solve for (Q, P) from Corollary 1.

Now suppose that the inputs U_i are in fact unknown (but still diagonal) and, in addition, there is an unknown input U_0 applied at every experiment. The former restriction is realistic in gene knockdown or silencing operations in biological systems, where it may be known that the expression of a certain gene has been disrupted, but not by exactly how much. The latter may represent an unknown initial condition, such as that caused by transferring cells to a different medium. By performing one additional experiment, it is possible to reconstruct Q, but not P, as follows.

Let Y_0 denote the Laplace transform of the measured states in a *wild-type* experiment, that is, where the only input is the extraneous input, U_0:

$$Y_0 = QY_0 + PU_0 \quad (10)$$

Now perform the p experiments as before:

$$\mathbf{Y} = Q\mathbf{Y} + P(\mathbf{U} + \mathbf{U_0}) \quad (11)$$

where $\mathbf{U_0} = [U_0\ U_0\ \dots\ U_0]$. Subtracting $\mathbf{Y_0} = [Y_0\ Y_0\ \dots\ Y_0]$ from both sides gives (via (10)):

$$\mathbf{Y} - \mathbf{Y_0} = Q(\mathbf{Y} - \mathbf{Y_0}) + P\mathbf{U} \quad (12)$$

This is equivalent to (9) with an alternative state matrix $\tilde{\mathbf{Y}} = \mathbf{Y} - \mathbf{Y_0}$, hence the effect of the extraneous input has been removed.

Lemma 1. *Under the conditions of Corollary 1, if the applied inputs are not measured, but $\tilde{\mathbf{Y}} = \mathbf{Y} - \mathbf{Y_0}$ is full rank, then the dynamical structure function $(Q, P\mathbf{U})$ can be obtained uniquely from $\tilde{\mathbf{Y}}$.*

Proof. The system equations are given by:

$$\begin{aligned}\tilde{\mathbf{Y}} &= Q\tilde{\mathbf{Y}} + P\mathbf{U} \\ &= Q\tilde{\mathbf{Y}} + \tilde{\mathbf{P}}\tilde{\mathbf{U}}\end{aligned} \quad (13)$$

where $\tilde{\mathbf{Y}} = \mathbf{Y} - \mathbf{Y_0}$ is known, $\tilde{\mathbf{P}} = P\mathbf{U}$ and $\tilde{\mathbf{U}} = I$. The matrix $[Q\ \tilde{\mathbf{P}}]$ now satisfies the condition of Theorem 1 and we can solve for Q and $\tilde{\mathbf{P}}$ by Corollary 1 with $G = \tilde{\mathbf{Y}}$. ■

Thus, if the values of the targeted inputs are unknown, we can still solve for Q but not P, and, in the presence of an unknown extraneous input, one additional wild-type experiment allows us to distinguish the effect of this input and remove it.

B. Structural Perturbations

We now consider perturbing the system by making targeted changes to the network structure. Motivated by cell biology, we consider two operations: node knock-out and node over-expression. The former represents an effective removal of a node from the network and can be modelled by deleting all incoming links to the node and initialising it at zero. In the latter, a dominant input is applied to a node, effectively overpowering all other inputs, and this can be modelled in a similar way by again deleting all incoming links then adding such an input. Figure 1 illustrates the concepts.

These perturbations can be replaced by certain (unknown) inputs and the reconstruction problem can then be solved using the results of Section III-A. First, note that in the case of only node knock-out data, an extraneous input is required in order to perturb the system. In fact, this input must ensure that every measured state is perturbed, either directly or indirectly, in the wild-type experiment.

Therefore, we assume the presence of an unknown extraneous input, U_0, as before and perform a wild-type experiment to obtain: $Y_0 = QY_0 + PU_0$. Now perturb the first state and partition the measured states as: $Y_1 = [Y_{11}^T\ Y_{12}^T]^T$, where Y_{11} is the scalar Laplace transform of the first measured state. The system equations are:

$$\begin{bmatrix}Y_{11}\\ Y_{12}\end{bmatrix} = \begin{bmatrix}Q_{11} & 0\\ Q_{21} & Q_{22}\end{bmatrix}\begin{bmatrix}Y_{11}\\ Y_{12}\end{bmatrix} + \begin{bmatrix}0\\ P_2\end{bmatrix}U_0 + \begin{bmatrix}\zeta_{11}\\ 0\end{bmatrix} \quad (14)$$

where $Q_{11} = 0$ by definition, since Q has zeros on the diagonal, and ζ_{11} is equal to zero in the case of a knock-out perturbation or is an unknown input signal in the case of over-expression.

In order to restore the original network topology that we wish to obtain, replace the structural changes by a fictitious input signal, V_{11} as follows:

$$\begin{bmatrix}Y_{11}\\ Y_{12}\end{bmatrix} = \begin{bmatrix}Q_{11} & Q_{12}\\ Q_{21} & Q_{22}\end{bmatrix}\begin{bmatrix}Y_{11}\\ Y_{12}\end{bmatrix} + \begin{bmatrix}P_1\\ P_2\end{bmatrix}U_0 + \begin{bmatrix}V_{11}\\ 0\end{bmatrix} \quad (15)$$

where $V_{11} \triangleq -Q_{12}Y_{12} - P_1U_0 + \zeta_{11}$. The input V_{11} is simply cancelling the effect of all other inputs to the knocked out state, forcing it to be equal to ζ_{11}. In general, we can replace

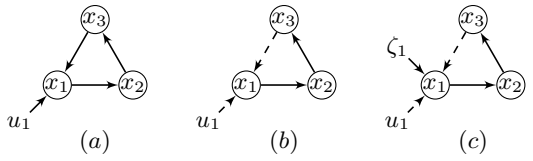


Fig. 1: Examples of input experiment (a), node knock-out (b) and node over-expression (c); circles denote measured states, solid arrows denote causal transfer functions and dashed arrows denote deleted connections. In (a), the correct network is shown and an input is applied to node 1; in (b), the knock-out of node 1 is modelled by deleting all inputs to that node; in (c), node 1 is over-expressed by deleting all its inputs and adding a new input, ζ_1. For clarity, extraneous inputs are omitted.

the perturbation of the i^{th} state by applying the following input to it:

$$V_{ii} = -Q(i,:)Y_i - P(i,:)U_0 + \zeta_{ii} \tag{16}$$

Combining the results for m experiments yields:

$$\mathbf{Y} = Q\mathbf{Y} + P\mathbf{U_0} + V \tag{17}$$

where as before $\mathbf{Y} = [Y_1\ Y_2\ \dots\ Y_m]$, $\mathbf{U_0} = [U_0\ U_0\ \dots\ U_0]$ and V is the $p \times m$ matrix with entries $V(i,j) = V_{ii}$ if state i is perturbed in the j^{th} experiment. Subtracting $\mathbf{Y_0} = [Y_0\ Y_0\ \dots\ Y_0]$ from (17) allows the extraneous input to be removed, as in Section III-A, resulting in:

$$\tilde{\mathbf{Y}} = Q\tilde{\mathbf{Y}} + V \tag{18}$$

where $\tilde{\mathbf{Y}} = \mathbf{Y} - \mathbf{Y_0}$. We can now extend the results of [10] as follows:

Lemma 2. *If $\tilde{\mathbf{Y}} = \mathbf{Y} - \mathbf{Y_0}$ is full rank, the dynamical structure function (Q, V) can be obtained uniquely from $\tilde{\mathbf{Y}}$ if and only if p elements in each row of $[Q\ V]$ are known.*

Proof. The problem can be expressed as:

$$\tilde{\mathbf{Y}} = Q\tilde{\mathbf{Y}} + VI \tag{19}$$

from (18), which is equivalent to (3) with $Y \leftarrow \tilde{\mathbf{Y}}$, $P \leftarrow V$, $U \leftarrow I$ and hence $G = \tilde{\mathbf{Y}}$. The proof follows by direct analogy with Theorem 1. ■

We can now, analogous to Lemma 1 where the inputs are unknown, state conditions under which the reconstruction problem can be solved with system perturbations.

Corollary 2. *If nothing is known about Q and $\tilde{\mathbf{Y}}$ is full rank, the dynamical structure function (Q, V) can be recovered uniquely from $\tilde{\mathbf{Y}}$ if p perturbation experiments are performed, where a unique measured state is perturbed in each experiment.*

Proof. The conditions of the corollary require that V can be written as the following square, diagonal matrix:

$$V = \begin{bmatrix} V_{11} & 0 & \dots & 0 \\ 0 & V_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & V_{pp} \end{bmatrix} \tag{20}$$

The problem as given in (19) now satisfies the conditions of Corollary 1, hence we can solve uniquely for V and, most importantly, Q. Specifically, letting $\mathbf{H} = \tilde{\mathbf{Y}}^{-1}$:

$$Q(i,j) = -\frac{\mathbf{H}(i,j)}{\mathbf{H}(i,i)} \qquad V(i,i) = \frac{1}{\mathbf{H}(i,i)} \tag{21}$$

■

Note that the perturbations applied to satisfy the condition of Corollary 2 need not all be of the same type, for example some may be knock-outs and some may be over-expressions. The method is summarized as follows:

Algorithm 1 Network reconstruction using structural perturbation data

Perform wild-type experiment to obtain Y_0.
for $i = 1 : p$ **do**
  Perturb state i to obtain Y_i.
  $\tilde{\mathbf{Y}}(:,i) = Y_i - Y_0$.
end for
$\mathbf{H} = \tilde{\mathbf{Y}}^{-1}$
for $i = 1 : p$ **do**
  for $j = 1 : p$ **do**
    $Q(i,j) = -\frac{\mathbf{H}(i,j)}{\mathbf{H}(i,i)} \qquad i \neq j$
  end for
end for

IV. Practical Considerations

The preceding section provided necessary and sufficient conditions to obtain the network structure and dynamics (Q) from identifiable quantities ($\mathbf{Y}$ and $\mathbf{Y_0}$). In practice, the estimation stage may be a significant source of error, particularly in biological applications where data availability is typically sparse. Here we show how previous results on robust and steady-state identification may be applied in the case of structural perturbations.

A. Robust Reconstruction

Previous work [11], [12] addressed this issue for the case of input experiments by modelling the error in the estimated transfer matrix, G, as a feedback uncertainty as follows:

$$G_t = (I + \Lambda)^{-1}G \tag{22}$$

where G_t is the true transfer matrix and Λ is a dynamic perturbation. Since G is no longer known exactly, the network topology (the Boolean structure of Q) can no longer be found uniquely and we must solve a subset selection problem to obtain the topology that is most likely, given G. A certain transfer matrix G_i is said to be consistent with a given (Q_i, P_i) if it satisfies (4). This matrix, G_i, can be related to G by a perturbation Δ_i as follows: $G_i = (I + \Delta_i)^{-1}G$, which, by eliminating G_i using (4), allows us to relate the estimated transfer matrix, G, to any (Q_i, P_i). Therefore, by minimising:

$$\|\Delta_i\| = \|GP_i^{-1}(I - Q_i) - I\| \tag{23}$$

subject to the constraints of the Boolean structure of Q_i, we can define a measure of distance between the topology of Q_i and the estimated transfer function G. With an appropriate choice of norm, (23) is convex, hence we can obtain a residual, $\delta_i = \min_{(Q_i,P_i)} \|\Delta_i\|$ for every possible topology and use a model selection procedure to select the most likely solution. In [11], this was done by exhaustive search and in [12] by a polynomial-time algorithm.

These existing results are directly transferable to the structural perturbation case. Under the conditions of Corollary 2, the system equations are given by (18). Post-multiply (22) by $P^{-1}V$ to give:

$$\tilde{\mathbf{Y}}_{\mathbf{t}} = (I + \Lambda)^{-1}\tilde{\mathbf{Y}} \tag{24}$$

where the subscript t denotes the true, exact, transfer functions. The method now proceeds as before using $\tilde{\mathbf{Y}}$ instead of G and V instead of P.

B. Steady-State Reconstruction

Theorem 1 and Corollary 1 provide conditions under which the dynamical structure function $(Q(s), P(s))$ can be reconstructed exactly from $G(s)$. Section III-B extends these results to obtain $(Q(s), V(s))$ from $\tilde{\mathbf{Y}}(\mathbf{s})$. If $G(s)$ (or $\tilde{\mathbf{Y}}(\mathbf{s})$) is only known at a particular frequency, then we can obtain the dynamical structure function at that frequency. For example, if $\tilde{\mathbf{Y}}(j\omega)$ is known, then under Corollary 2 we can obtain $(Q(j\omega), V(j\omega))$.

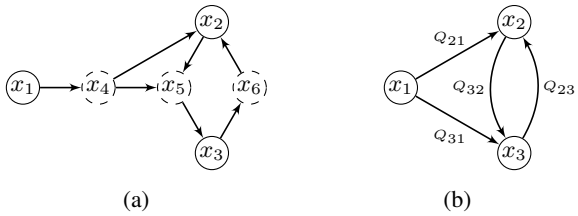
In particular, we can obtain $\tilde{\mathbf{Y}}(0)$ from steady-state data, allowing reconstruction of $(Q(0), V(0))$. Providing no element $Q(i,j) \neq 0$ has a zero at $s = 0$, the Boolean structure of $Q(0)$ will be the same as that of $Q(s)$, revealing the network topology.

V. Example

Here we demonstrate the approach with a simulated network reconstruction example. Fig. 2(a) shows a network of six states, three of which are measured, where arrows between states represent first-order linear dynamics. Fig. 2(b) shows the representation of the network between only the measured states, where arrows denote direct relations containing all the dynamics of the hidden states. The dynamical structure function describes both the topology and dynamics of these interactions and it is this which we aim to reconstruct.

This example contains network features which are typically difficult to identify such as feedback between states x_2 and x_3, co-regulation of states x_2 and x_3 by x_1 and connections of different orders (the connection from $x_1 \to x_3$ is third-order, while all the others are second-order). The arrangement of measured states interacting via hidden states may be representative of gene RNA (measured) interacting via translated proteins (unmeasured).

The network is fully characterised by the following state-space matrices, where numerical values were chosen at random (subject to A being stable) and, for simplicity, inputs directly affect all states:



(a) (b)

Fig. 2: An example network topology of six states, three of which are measured. Solid circles denote measured states, dashed circles denote hidden (unmeasured) states and arrows denote direct causal connections. In (a) all states are shown, whereas (b) is the representation of the network between measured states, which we seek to obtain. The labels Q_{ij} refer to the elements of (26).

$$A = \begin{bmatrix} -4.3 & 0 & 0 & 0 & 0 & 0 \\ 0 & -2.5 & 0 & 1.9 & 0 & -6.6 \\ 0 & 0 & -4.1 & 0 & 9.6 & 0 \\ -6.8 & 0 & 0 & -7.7 & 0 & 0 \\ 0 & 4.8 & 0 & 8.5 & -7.4 & 0 \\ 0 & 0 & 1.5 & 0 & 0 & -7.1 \end{bmatrix} \tag{25}$$

$$B = \begin{bmatrix} 2.6 & -4.7 & 1.9 & 7.9 & 5.9 & -3.9 \end{bmatrix}^T$$

with $C = \begin{bmatrix} I & 0 \end{bmatrix}$. The resulting Q is then:

$$Q = \begin{bmatrix} 0 & 0 & 0 \\ Q_{21} & 0 & Q_{23} \\ Q_{31} & Q_{32} & 0 \end{bmatrix} \tag{26}$$

where the nonzero elements are the transfer functions of the links of Fig. 2(b) with the following dynamics:

$$\begin{aligned} Q_{21} &= \frac{-12.9}{(s+10.4)(s+2.5)} \\ Q_{23} &= \frac{46.1}{(s+8.6)(s+2.5)} \\ Q_{31} &= \frac{-554.9}{(s+8.9+j3.0)(s+8.9-j3.0)(s+2.0)} \\ Q_{32} &= \frac{-9.9}{(s+8.7)(s+2.0)} \end{aligned} \tag{27}$$

We now simulate the above network and aim to recover (26) directly from data by knocking-out each of the measured states.

First, a unit step input was applied to the network to generate a wild-type experiment. This could represent the cells being placed in a different medium or undergoing a particular treatment. A step was used for convenience and neither this nor B need be known. Time-series data was generated by sampling the wild-type response at regular intervals and three experimental replicates were then obtained by adding zero-mean Gaussian noise to these samples (see Fig. 3a).

Next, each of the measured states was knocked-out in turn, the same step input applied and three replicates generated again by adding Gaussian noise. This data is shown in Figs. 3b-d. We now use only this noisy data to reconstruct Q; for simplicity here we consider only steady-state data (of the last three time points) and estimate $Q(0)$.

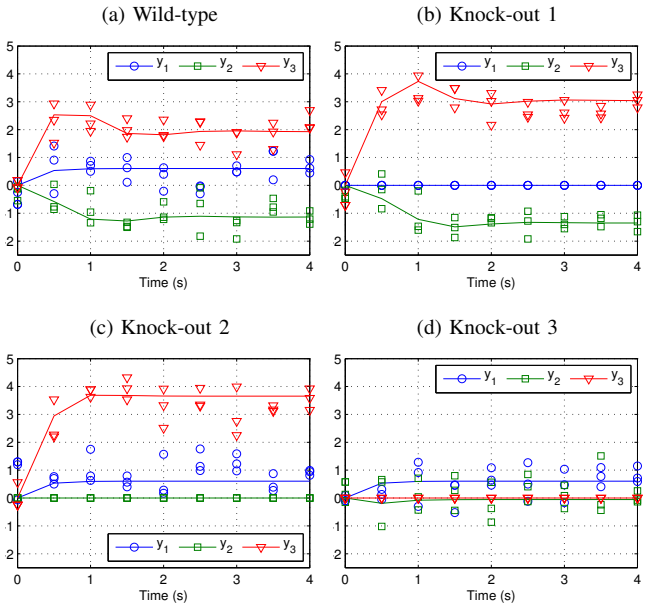


Fig. 3: Simulated time-series data for a) wild-type and b)-d) knock-out cases. The solid lines show the actual trajectories for each of the measured states; the data points were obtained by adding zero-mean Gaussian noise with variance 0.25 to these and represent three experimental repeats.

Using the procedure of Section IV-A, we computed a residual δ_i for each of the 64 possible topologies with three measured states. We then applied AIC model selection which selected the correct structure for Q with the following numerical values for $Q(0)$:

$$Q(0) \approx \begin{bmatrix} 0 & 0 & 0 \\ -0.82 & 0 & -0.84 \\ -0.61 & 0.89 & 0 \end{bmatrix} \tag{28}$$

which is a reasonable estimate of the true $Q(0)$:

$$Q(0) = \begin{bmatrix} 0 & 0 & 0 \\ -0.67 & 0 & -0.56 \\ -2.38 & 1.52 & 0 \end{bmatrix} \tag{29}$$

Most importantly, the topology is correct, but, in addition, the sign and approximate magnitude of the steady-state interactions are also obtained.

VI. Conclusion

This paper addressed the network reconstruction problem for partially observed LTI systems from mutant data. Rather than probing the system with inputs, each experiment represents a mutation of the network structure. Specifically, data is obtained from knock-out and over-expression experiments, which are typical in biology (e.g. in the identification of gene regulatory networks). We provide necessary and sufficient conditions under which exact reconstruction is possible, leading to an algorithm for guaranteed reconstruction from noiseless data. Previous results are applicable to make this procedure robust to the presence of noise in the data and to nonlinearities in the system.

We recognise that a major limitation of this approach is the restriction to LTI systems. Whilst for some applications, a linearised model may be reasonable, gene regulatory networks are likely to contain significant nonlinear components. In addition, gene mutations may cause the expression of other genes to saturate, potentially moving the states beyond their linearised range. Nevertheless, our theoretical contribution is important and the extension to nonlinear systems is significantly more challenging. The limited amount of data typically available in biological experiments also restricts the complexity of models that can reasonably be considered.

VII. Acknowledgements

We would like to thank the reviewers for their helpful comments and Miss Ni Ji for initiating this discussion. This work was supported by EPSRC grants EP/I03210X/1 and EP/P505445/1.

References

[1] Marbach et al., "Wisdom of crowds for robust gene network inference," *Nature Review*, vol. 9, no. 8, 2012.
[2] R. De Smet and K. Marchal, "Advantages and limitations of current network inference methods," *Nature Reviews Microbiology*, 8, 2010.
[3] A. Margolin, K. Wang, W. Lim, M. Kustagi, I. Nemenman and A. Califano, " Reverse engineering cellular networks," *Nature Protocol*, 1, pp. 662 - 671, 2006.
[4] J. Yu, V. Smith, P. Wang, A. Hartemink and E. Jarvis, "Advances to Bayesian network inference for generating causal networks from observational biological data," *Bioinformatics*, 12;20(18), 2004.
[5] B.N. Kholodenko, A. Kiyatkin, F. Bruggeman, E.D. Sontag, H. Westerhoff, and J. Hoek. Untangling the wires: a novel strategy to trace functional interactions in signaling and gene networks. *Proceedings of the National Academy of Sciences USA*, 99:12841–12846, 2002.
[6] Michael M. Zavlanos, Agung A. Julius, Stephen P. Boyd and George J. Pappas, Inferring stable genetic networks from steady-state data, *Automatica*, vol. 47, 2011, pp. 1113-1122.
[7] T. Gardner, D. Bernardo, D. Lorenz and J. Colins, "Inferring genetic networks and identifying compound mode of action via expression profiling", *Science*, 301 (5629), pp. 102-105, 2003.
[8] M. Nabi-Abdolyousefi and M. Mesbahi, "Network identification via node knock-outs," *IEEE Transactions on Automatic Control*, vol. 57, no. 12, pp. 3214-3219, 2012.
[9] D. Materassi and M. V. Salapaka, "On the problem of reconstructing an unknown topology via locality properties of the weiner filter", *IEEE Transactions on Automatic Control*, vol. 57, no. 7, 2012.
[10] J. Gonçalves and S. Warnick, "Necessary and sufficient conditions for dynamical structure reconstruction of LTI networks", *IEEE Transactions on Automatic Control*, vol. 53, pp. 1670-1674, 2008.
[11] Y. Yuan, G. Stan, S. Warnick and J. Gonçalves, "Robust dynamical network structure reconstruction," *Automatica*, vol. 47, 6, 2011.
[12] D. Hayden, Y. Yuan and J. Gonçalves, "Robust network reconstruction in polynomial time," *Proceedings of 51st IEEE Conference on Decision and Control*, 2012.
[13] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbour rules," *IEEE Transactions on Automatic Control*, vol. 48, no. 6, 2003.
[14] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, 2004.
[15] Ye Yuan, Decentralised network prediction and reconstruction algorithms, *Ph.D. Thesis, Cambridge University,* 2012.
[16] R. Olfati-Saber. "Flocking for multi-agent dynamic systems: algorithms and theory," *IEEE Trans. on Automatic Control*, vol. 51(3), pp. 401-420, Mar. 2006.
[17] L. Pecora and M. Barahona, "Synchronisation of oscillators in complex networks," *Chaos and Complexity Letters*, 1, 61, 2005.
[18] L. Ljung. *System Identification–Theory for the User*. Prentice Hall, 1999.
[19] K. Astrom, "Transient response analysis," *Technical Report, Department of Automatic Control, Lund University*, 1985.