Robust dynamical network reconstruction

Ye Yuan[1], Guy-Bart Stan[2], Sean Warnick[3] and Jorge Gonçalves[1].

***Abstract*— This paper addresses the problem of robustly reconstructing network structure from input-output data. Previous work identified necessary and sufficient conditions for network reconstruction of LTI systems, assuming perfect measurements (no noise) and perfect system identification. This paper assumes that the previously identified necessary and sufficient conditions for network reconstruction are satisfied but here we additionally take into account noise and unmodelled dynamics (including nonlinearities). In order to identify the network structure that generated the data, we compute the smallest distances between the measured data and the data that would have been generated by particular Boolean networks. By striking a compromise between such distance and network complexity, we provide methods for revealing the correct network structure from data despite the presence of noise and nonlinearities.**

I. Introduction

The challenges faced by the network reconstruction problem [1] come from the necessity to deal with noisy and partial measurements (in particular, the number of hidden/unobservable nodes and their position in the network is unknown) taken from a nonlinear and stochastic network. Even in the ideal situation where the underlying network is assumed to be linear and time-invariant (LTI) and the measurements are assumed to be non-noisy, it can be shown that, due to partial observability, this problem is unsolvable using classical system identification techniques [2]. In particular, identification of the system transfer function (obtained, for example, using system identification approaches) is useless to solve the network structure reconstruction problem since transfer functions do not contain sufficient information for that purpose.

Based on this latter observation, a new representation for LTI systems, called dynamical structure functions was introduced in [2]. Dynamical structure functions capture information at an intermediate level between transfer function and state space representation. Specifically, dynamical structure functions not only encode structural information at the measurement level, but also contain some information about hidden states. Based on the theoretical results presented in [2], we proposed an experimental guideline for the design of an experimental data-acquisition protocol which allows the collection of data containing sufficient information for the network structure reconstruction problem to become solvable. In the particular case of biological network reconstruction, we have shown that *if nothing is known about the network a priori* (i.e., if we consider a black box network structure reconstruction problem), then the data-collection experiments must be performed as follows:

1) for a network composed of p measured species, the same number of experiments p must be performed;
2) each experiment must independently control a measured species, i.e., control input i must first affect measured species i.

If the experiments are not performed in this way the network cannot be reconstructed based on the corresponding input-output measurements, and any network structure fits the data equally well (e.g. a fully decoupled network or a fully connected network). If a priori information about the network is available (i.e., if we consider a grey box network structure reconstruction problem), as is usually the case, then these conditions can be relaxed as explained in [2].

The main contribution of this paper is to propose an efficient algorithm for reconstructing the network structure best fitting noisy input-output data. In particular, the proposed network reconstruction method is designed to be robust with respect to noise and dynamic uncertainties (such as unmodelled nonlinearities). In the rest of the paper, we assume that the conditions for network reconstruction presented above in (1) and (2) are satisfied.

Our method uses the same type of information as system identification methods, i.e., input-output measurements. However, contrary to what can be done with a direct application of classical system identification methods, steady-state (resp. time-series data) can here be used to reconstruct the Boolean (resp. dynamical network) structure of the system.

The structure of the paper is as follows. In Section I-B, we give a simple example which shows that direct system identification from input-output data does not allow the reconstruction of the network structure when hidden/unobservable states are presented. In Section II, dynamical structure functions are defined and fundamental results concerning their usefulness in the network reconstruction problem are stated. Section III presents the main results of the paper, i.e., robust network reconstruction from input-output data in the presence of noise and nonlinearities. Finally, we conclude the paper with some discussions and future work in Section IV.

A. Notation

For a matrix A, A_{ij} denotes the element in the i^{th} row and j^{th} column while A_j denotes its j^{th} column. For a column vector α, $\alpha[i]$ denotes its i^{th} element. We define $e_r^T = [0, \ldots, 0, 1_{r^{th}}, 0, \ldots, 0] \in \mathbb{R}^{1\times N}$. I denotes the identity matrix of appropriate dimension.

[1]Control Group, Department of Engineering, University of Cambridge, Cambridge, U.K.. [2]Centre for Synthetic Biology and Innovation, Department of Bioengineering, Imperial College London. [3]Computer Science Department, Brigham Young University.

Corresponding author: jmg77@cam.ac.uk (Jorge Gonçalves).

ISBN 978-963-311-370-7

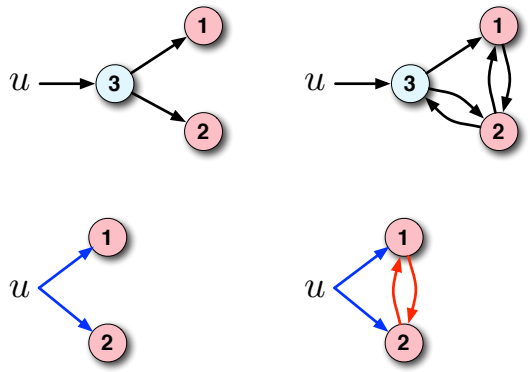

Fig. 1. The same transfer function yields two minimal realisations with very different network structures: decoupled internal structure (left) and coupled internal structure (right). Measured species are depicted by red circles while hidden species correspond to blue circles. The complete biochemical network, reflected in each state space realisation, is shown on top; on the bottom are the networks between measured species only. Blue and red arrows represent transfer functions and include the dynamics corresponding to hidden states.

B. Motivating example

Consider a linear time-invariant system from which partial, non-noisy input-output measurements are obtained. Using system identification, a transfer function describing the input-output behaviour of this system can be obtained. However, in the partial observation case, network reconstruction is not possible with no further information. To illustrate this, assume that the transfer function obtained from input-output data (e.g., using classical systems identification methods) is given by:

$$G(s) = \frac{1}{s+3}\begin{bmatrix} \frac{1}{s+1} \\ \frac{1}{s+2} \end{bmatrix}.$$

It can be shown that this transfer function is consistent with two state-space realisations $\dot{x} = Ax + Bu$, $y = Cx$ with very different internal structures, i.e.,

$$A_1 = \begin{bmatrix} -1 & 0 & 1 \\ 0 & -2 & 1 \\ 0 & 0 & -3 \end{bmatrix}, \quad A_2 = \begin{bmatrix} -2 & -1 & 1 \\ -1 & -3 & 1 \\ 0 & -1 & -1 \end{bmatrix},$$

$B_1 = B_2 = [0 \;\; 0 \;\; 1]^T$, and $C_1 = C_2 = [I \;\; 0] \in \mathbb{R}^{2\times 3}$ (i.e., the third state is hidden/non-observable). The networks in Figure 1 correspond to each of the indicated realisations of $G(s)$ (the (A_1, B_1, C_1) realisation (resp. the (A_2, B_2, C_2) realisation) corresponds to the left column (resp. right column) of Figure 1). Note that both realisations are minimal. This demonstrates that even in the idealised setting (LTI dynamics, non noisy data), network reconstruction in the presence of hidden/unobservable states is not possible without additional information about the system.

To ease the notation, we omit the explicit dependence of transfer functions on the Laplace variable s when this does not lead to confusion. Thus write G instead of $G(s)$ where the explicit depence on s is clear from the context.

II. Dynamical structure functions and network reconstruction

In [2] we introduced the notion of dynamical structure functions and showed how they can be used to obtain necessary and sufficient conditions for network reconstruction. For the sake of clarity and completeness, we state these previously obtained results here without proofs. We refer the interested reader to [2] for the corresponding proofs.

Consider a nonlinear system $\dot{\bar{x}} = f(\bar{x}, \bar{u}, w_1)$, $\bar{y} = h(\bar{x}, w_2)$ with p measured states $\bar{y}$, hidden states $\bar{z}$ (potentially a large number of them), m inputs $\bar{u}$, and noise w_1, w_2. The system is linearised around an equilibrium point (a point such that $f(\bar{x}^*, \bar{u}^*, 0) = 0$), and it is assumed that inputs and noise do not move the states too far from the equilibrium point so that the linearised system is a valid approximation of the original nonlinear system. The linearised system can be written as $\dot{x} = Ax + Bu$, $y = Cx$, where $x = \bar{x} - \bar{x}^*$, $u = \bar{u} - \bar{u}^*$ and $y = h(\bar{x}, 0) - h(\bar{x}^*, 0)$. The transfer function associated with this linearised system is given by $G(s) = C(sI - A)^{-1}B$.

Partition the linearised system as follows

$$\begin{aligned} \begin{bmatrix} \dot{y} \\ \dot{z} \end{bmatrix} &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} y \\ z \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u \\ y &= \begin{bmatrix} I & 0 \end{bmatrix}\begin{bmatrix} y \\ z \end{bmatrix} \end{aligned} \tag{1}$$

where $x = [y^T \;\; z^T]^T \in \mathbb{R}^n$, is the full state vector, $y \in \mathbb{R}^p$ is a partial measurement of the state (we assume $p > 1$), z are the $n - p$ "hidden" states, and $u \in \mathbb{R}^m$ is the control input. We restrict our attention to situations where output measurements constitute partial state information, i.e., $p < n$. Taking the Laplace transforms of the signals in (1), solving for Z, and substituting it into the Laplace transform of the first equation of yields $sY = WY + VU$, where $W = A_{11} + A_{12}\left(sI - A_{22}\right)^{-1} A_{21}$ and $V = A_{12}\left(sI - A_{22}\right)^{-1} B_2 + B_1$. Let D be the matrix composed of the diagonal elements of W and write $\left(sI - D\right)Y = \left(W - D\right)Y + VU$. Then $Y = QY + PU$ where

$$Q = \left(sI - D\right)^{-1}\left(W - D\right) \text{ and } P = \left(sI - D\right)^{-1} V \tag{2}$$

Given the system (1), we define the *dynamical structure function* of the system to be (Q, P), where Q and P are the *internal structure function* and *control structure function*, respectively, and given as in (2). If all the measured states are removed from the system except for Y_i and Y_j then the transfer function Q_{ij} corresponds to the exact 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.

It can be shown that $G = \left(I - Q\right)^{-1} P$ (see [2]). Based on this latter relation, it can be shown that the dynamical structure function of a system contains more information than the transfer function, and less information than the state-space representation. We can then conclude that, with no other information about the system, dynamical or Boolean reconstruction is not possible. Moreover, for *any* internal structure Q there is a dynamical structure function (Q, P)

that is consistent with G, i.e. that satisfies $G = (I - Q)^{-1} P$. In particular, this shows that the use of criteria such as sparsity or decoupledness to guide our selection of a proposal network structure can be misleading. If one were to optimise for decoupledness, for example, a dynamical structure $(0, G)$ could and would always be found, regardless of the true underlying structure. Thus, if we are to use these kinds of criteria, they must be firmly justified a priori.

Proposition 1: Given a $p \times m$ transfer function G, dynamical structure reconstruction is possible from partial structure information if and only if $p - 1$ elements in each column of $\begin{bmatrix} Q & P \end{bmatrix}^T$ are known that uniquely specify the component of (Q, P) in the nullspace of $\begin{bmatrix} G^T & I \end{bmatrix}$.

The importance of this result is that it identifies exactly what information about a system's structure, beyond knowledge of its transfer function, must be obtained to be able to recover the structure without appeal to a priori assumptions, such as sparsity, or parsimony, etc. This enables the design of experiments targeting precisely the extra information needed for reconstruction. In particular when $p = m$ and G is full rank, we observe that imposing that P is diagonal, i.e., that each input controls a measured state independently, is sufficient for reconstruction.

Corollary 1: If $m = p$, G is full rank, and there is no *a priori* information about the internal structure of the system, Q, then the dynamical structure can be reconstructed if each input controls a measured state independently, i.e., if, without loss of generality, the inputs can be numbered such that P is diagonal. Moreover, $H = G^{-1}$ characterises the dynamical structure as follows

$$Q_{ij} = -\frac{H_{ij}}{H_{ii}} \quad \text{and} \quad P_{ii} = \frac{1}{H_{ii}}. \tag{3}$$

III. Robust network structure reconstruction

In this section, we consider the problem of robustly reconstructing dynamical network structures. Data are obtained from input-output measurements of a noisy nonlinear system. From this type of data we aim to find the internal network structure Q associated with the linearised system (1).

For simplicity of exposition, we assume that there is no *a priori* information on the internal network structure Q. The results still follow if some *a priori* information about Q is available, and such information can typically be used to relax the experimental protocol according to Proposition 1. Hence, data are collected according to the measurement protocol described in the introduction:

(1) the number of distinct data-collection experiments is the same as the number of measured species. This in particular implies that $u(t), y(t) \in \mathbb{R}^p$;

(2) each input u_i controls first the measured state y_i so that P is a diagonal matrix ($p \times p$). To average out the noise, data-collection experiments are repeated N times.

In the following sections, we propose two approaches for estimating the dynamical structure function (Q, P) from measured input-output data. The first approach is indirect and involves estimating the transfer function G while the second approach relies on the solution of a direct optimisation problem. More precisely, in the first approach (see Figure 2 (a)), for each experiment i we first estimate $G_i(s)$ (i.e., the i^{th} column of $G(s)$) using standard system identification tools [6]. In a second step, the dynamical structure function $(Q(s), P(s))$ is computed from the estimated transfer function $G(s)$. Since information is lost in the process of estimating $G(s)$, we later we consider the case where $(Q(s), P(s))$ is directly estimated from data (without estimating first $G(s)$, see Figure 2 (b)).

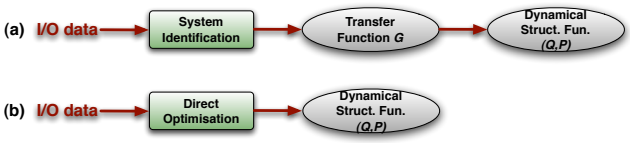

Fig. 2. Two approaches to obtain dynamical structure functions.

Concerning the type of input-output data collected, we first assume the case of time-series input-output data. We then consider the special case where only steady-state data are available.

A. Dynamical network reconstruction from identified transfer functions

In this section, we describe a first method relying on system identification. This method allow us to obtain dynamical structure functions from a transfer function identified using measured time-series data (see Figure 2 (a)).

Consider a transfer matrix $G(s)$ estimated from noisy data. According to Corollary 1, if G is full rank there is a unique Q and diagonal P satisfying $(I - Q)G = P$. Since G is an approximation of the actual system, Q and P will typically be mere approximations of the actual Q and P. Moreover, due to noise and unmodelled dynamics, it is likely that Q does not even have the correct Boolean structure. Typically, the internal structure function Q obtained from such a procedure will be fully connected, i.e., all non-diagonal elements of Q will be non-zero.

The main idea to solve the network reconstruction problem from noisy data is the following. For p measured states, Q has $p^2 - p$ unknowns. We want to quantify the *distance* from G (or directly from the measured data) to all possible Boolean structures (and there are 2^{p^2-p} of them). Some of such distances will be large revealing that the corresponding Boolean structures are likely not the correct structures while other will be small making them candidates for the correct structure.

To make the notion of distance rigorous, consider a Boolean mapping from a transfer matrix to a Boolean matrix in Definition 1 and a particular Boolean structure $\mathcal{B}_k$ in Definition 2.

Definition 1: A Boolean mapping $b : Q(s) \rightarrow b(Q)$, where $b(Q)$ is a Boolean matrix with the same dimension as transfer matrix Q and $\forall i, j \ \{b(Q)\}_{i,j} = 0$ if and only if $Q_{ij}(s) = 0$ for all s, otherwise, $\{b(Q)\}_{i,j} = 1$.

For a given p, there are 2^{p^2-p} possible Boolean networks B_k (remember that $Q(s)$ has zeros on the diagonal and therefore $b(Q)$ will always have zeros on the diagonal) which can be ordered using the index $k = 1, ..., 2^{p^2-p}$.

Definition 2: A Boolean structure $\mathcal{B}_k$ corresponding to a Boolean network B_k is defined as follows: $\{Q(s) : b(Q) = B_k\}$.

The distance from G to the Boolean structure $\mathcal{B}_k$ is defined as the smallest perturbation Δ to G (measured in some norm) so that the perturbed system G_Δ belongs to the set of transfer functions $\tilde{G}$ such that $Q \in \mathcal{B}_k$, where Q is obtained from $(I - Q)\tilde{G} = P$. Finding the distance from G to a Boolean structure $\mathcal{B}_k$, gives us a quantitative information about how much we would need to perturb G (or the data) to obtain a new system transfer function for which the associated Q corresponds to the considered Boolean structure, i.e., for which $Q \in \mathcal{B}_k$.

There are many possible approaches to define such "smallest perturbations", including several uncertainty models and norms to choose from. This choice is key to obtain a convex minimisation problem. For example, additive, multiplicative or uncertainty in the coprime factors all lead to non-convex minimisation problems. In order to obtain a convex minimisation problem, we consider the output (could also be input) feedback uncertainty model. In this framework, the "true" system is given by $(I + \Delta)^{-1}G$, where Δ represents unmodelled dynamics, including nonlinearities, and noise.

Based on this choice of dynamic uncertainty, the problem is defined as follows. Given a particular Boolean structure $\mathcal{B}_k$, the objective is to minimise $\|\Delta\|$, in some norm, such that Q obtained from $(I + \Delta)^{-1}G = (I - Q)^{-1}P$ has the desired Boolean structure, i.e. $Q \in \mathcal{B}_k$. All P_{ii} are also free (remember that, by assumption, P is diagonal).

We can rewrite the above equation as $\Delta = GP^{-1}(I - Q) - I$. So, we are looking to minimise $\|GP^{-1}(I-Q) - I\|$ over $Q \in \mathcal{B}_k$ and P diagonal. Since P is diagonal, its inverse P^{-1} is also diagonal.

Define a new matrix $X = P^{-1}(I - Q)$ whose diagonal is the diagonal of P^{-1} and for which the off diagonal elements are given by $P_{ii}^{-1}Q_{ij}$. Since $Q \in \mathcal{B}_k$ this imposes structural constraints on X, i.e., some off-diagonal $X_{ij} = 0$. These zero X_{ij} correspond to those Q_{ij} which are equal to zero (since $X_{ij} = P_{ii}^{-1}Q_{ij}$ for $i \neq j$).

Definition 3: For all k, define $\mathcal{X}_k \triangleq \{X(s) : b(X) = B_k + I_p\}$, where I_p is identity matrix of dimension p.

Remark 1: Definition 3 implies the following facts:

(i) when $i \neq j$, $X_{ij}(s) = 0$ for all the Boolean structures $\mathcal{B}_k$ in Definition 2 which are such that $B_k[i, j] = 0$; all other $X_{ij}(s)$ are free variables;

(ii) when $i = j$, $X_{ii}(s)$ is a free variable.

Using Defintion 3, the distance from G to a particular Boolean structure $\mathcal{B}_k$ can be written as

$$\alpha_k = \inf_{X \in \mathcal{X}_k} \|GX - I\|^2$$

which is a convex minimisation problem with a careful choice of a norm.

Remark 2: In this optimisation problem, $X(s) \in \mathcal{X}_k$ approximates the inverse of G as "close" as possible. If $\mathcal{X}_k$ corresponds to the fully connected Boolean network then the solution to this optimisation is exactly $X = G^{-1}$.

Next, we show that this problem can be casted as a least squares optimisation problem. If we use the norm defined by $\|\Delta\|^2 =$ sum of all $\|\Delta_{ij}\|_2^2$, where $\|\cdot\|_2$ stands as the $\mathcal{L}_2$-norm over $s = j\omega$, then using the projection theorem [9] the problem reduces to

$$\begin{aligned}\alpha_k = \inf_{X \in \mathcal{X}_k} \|GX - I\|^2 &= \inf_{X \in \mathcal{X}_k} \sum_i \|GX_i - e_i\|_2^2 \\ &= \sum_i \inf_{Y_i} \|A_i Y_i - e_i\|_2^2 \\ &= \sum_i \|A_i(A_i^* A_i)^{-1} A_i^* e_i - e_i\|_2^2,\end{aligned}$$

where X_i is the i^{th} column of $X \in \mathcal{X}_k$, Y_i is a column vector composed by the free (i.e., nonzero) elements of X_i, A_i is obtained by deleting the j^{th} column of G when the corresponding element $X_i(j)$ is 0 for all j, and $(\cdot)^*$ denotes transpose conjugate. The infimum is achieved by choosing $X_i = (A_i^* A_i)^{-1} A_i^* e_i$, and $A_i^* A_i$ is always invertible since G is full rank in Corollary 1. After obtaining all the α_k for all k, the optimal distance

$$\alpha = \min_k \alpha_k.$$

If experiments are repeated N times (as they should) and we obtain a transfer function G_i for each experiment, then the above analysis still follows simply by forming a higher dimensional matrix $G = [G_1^T \ \cdots \ G_N^T]^T$.

Penalising connections: The above methodology suffers from a crucial weakness: there are several Boolean structures with distances smaller or equal than the distance to the "true" network. Indeed, the extra degrees of freedom of the fully-connected network allow the corresponding distance α_k to be the smallest of all. This is similar to the noisy data overfitting problem encountered in system identification where the higher the order of the transfer function, the better the fit. Obviously, if we only focus on noisy data best fit, eventually we end up fitting noise and so a large system order is not typically a good choice. Therefore, a compromise has to be struck.

If the true network has l non-existent connections (l off-diagonal elements in Q are zero) and the data are non-noisy, then there are $2^l - 1$ different networks that have a smaller or equal distance (due to the additional degrees of freedom provided by the extra connections). When noise is present, then the "true" network will typically have an optimal distance similar to these other l networks. The question of how to find the "true" network thus arises. With repeated experiments, small enough noise (i.e., large enough signal-to-noise ratio) and negligible nonlinearities, the optimal distances of those l networks are comparable, and they are typically much smaller than those of the other networks. To try to reveal the "true" network, one can strike a compromise between network complexity (in terms

of number of connections) and data fitness by penalising extra connections. The are several methods to strike this compromise. Here, we introduce methods known as Akaike's information criterion (AIC) [4], or some of its variants such as AICc (which is AIC with a second order correction for small sample sizes), and the Bayesian information criterion (BIC).

The AIC-type approach is a test between models - a tool for model selection. Given a data set, several competing models may be ranked according to their AIC value, with the one having the lowest AIC being the best. From the AIC value one may typically infer that the best models are in a tie and the rest are far worse, but it would be arbitrary to assign a value above which a given model is rejected [10]. The AIC value in our case for a particular Boolean network B_k is defined as:

$$AIC_k = 2L_k - \ln \alpha_k, \tag{4}$$

where L_k is number of (non-zero) connections in the Boolean network B_k and α_k is the optimal distance based on this parameter constraint.

Although finding the optimal distance in the second term of eq. (4) can be done efficiently, the number of Boolean networks 2^{p^2-p} grows very fast with the number of measured states p. To find the network with the smallest distance it is thus not desirable to compute the optimal distance for each possible Boolean network. Fortunately, there are ways to reduce the number of networks that need to be considered. As we saw in the previous section $\inf_{X\in\mathcal{X}_k} \|GX - I\|^2 = \sum_i \inf_{Y_i} \|A_i Y_i - e_i\|_2^2$ meaning that we can solve each optimisation problem separately. Since each Y_i corresponds to $p-1$ unknowns in the i^{th} row of Q, this reduces the problem to solving $p2^{p-1}$ optimal distances. Finding a polynomial-time algorithm to compute the optimal distance through this method is a subject of current investigation.

B. Dynamical network reconstruction directly from time-series data

The previous sections used a two-step approach in which system identification was first used to estimate a transfer function from measured input-output data and then, in a second step, the identified transfer function was used to obtain a dynamical structure function representation of the system which is optimal in terms of a particular metric. This section proposes a method which allows identification of the optimal dynamical structure function representation directly from the measured input output data (see Figure 2 (b)). The advantage of this direct network structure reconstruction from data is that no information is lost during the initial transfer function identification stage.

Due to the equivalence between dynamical uncertainty perturbations, we are free to chose, without loss of generality, the type of uncertainty perturbation that best suits our needs. For the direct method, instead of a feedback uncertainty as was considered in the previous section, the uncertainty perturbation we are considering here is the additive dynamic uncertainty on the output, i.e., $Y = G_\Delta(U + \Delta)$. In this case, we think about the "distance" in terms of how much we need to change the input (data) to fit a particular Boolean structure. Since $G_\Delta = (I - Q)^{-1}P = X^{-1}$, the equality $Y = G_\Delta(U + \Delta)$ can be written as

$$\Delta = XY - U,$$

where $X \in \mathcal{X}_k$, for some particular Boolean network k. Recall that structural constraints in Q can be imposed directly on X from the equality $X = P^{-1}(I - Q)$. We can therefore use system identification theory for non-causal autoregression models under the structural constraints to identify X (which might be non-causal). In this case, the distance is defined as the maximum likelihood of the estimation problem.

Reconstruction with the zero norm: Taking the number of connections into account, we formulate the optimisation problem as follows:

$$\inf_{X\in\mathcal{X}_k} \left(\|XY - U\|^2 + \beta \|X\|_0 \right), \tag{5}$$

where β is a parameter balancing data-fitting and model complexity (i.e., the number of non-zero connections). In (5), $\|X\|_0$ denotes the number of nonzero element in the matrix X, and it is known as the zero norm. Notice that this minimisation problem can be equivalently written as:

$$\begin{aligned} &\inf_{X\in\mathcal{X}_k} \left(\|XY - U\|^2 + \beta \|X\|_0 \right) \\ &= \sum_i \inf_{X_i} \left(\|X_i^T Y - U_i^T\|_2^2 + \beta \|X_i^T\|_0 \right), \end{aligned} \tag{6}$$

where X_i^T is the i^{th} row of $X \in \mathcal{X}_k$ and U_i^T is the i^{th} row of U. Directly solving this problem is in general NP-hard. A frequently discussed approximation is to replace the zero norm with the 1-norm to obtain a convex problem. Moreover, since there are p independent optimisations in eq. (6), we can choose different β_i for each i. Alas, there is no clear rule for selecting β_i to balance optimally the two terms in eq. (6) [5]. The choice of β_i is currently under investigation.

C. Boolean network reconstruction from steady-state data

So far we have assumed that time-series data are available. Frequently, however, experimentation costs and limited resources only permit steady-state measurements. In addition, with steady-state measurements it is typically possible to perform a larger number of experiments for the same time, effort and cost. As shown below, most of the connectivity of the network together with the associated steady-state gains (and the associated positive or negative sign) can still be reconstructed from steady-state date. However, no dynamical information will be obtainable. In other words, for most cases we can still recover the Boolean network from steady-state data.

Assume that after some time of maintaining the control input concentrations at a constant value, the measured outputs y have converged to a steady-state value. This is equivalent (if the system is stable or quasi-stable [1]) to assume that we can obtain $G_0 = G(0)$, i.e., $G(s)$ evaluated at $s = 0$. If $Q_0 = Q(0)$ and $P_0 = P(0)$, then $(I - Q(s))G(s) = P(s)$ evaluated

at $s = 0$ becomes $(I - Q_0)G_0 = P_0$. From this equation, all of the results given in Section III-A and III-B follow provided that no element of $G(s)$ has a system zero [3] at 0. In that case, a nonzero element in the obtained Boolean network indicates the existence of a causal relationship between the corresponding pair of nodes while a zero element indicates the absence of such relationship.

IV. Discussion, conclusion and future work

The main contribution of this paper is to propose an efficient method for reconstructing from noisy input-output data the network structure responsible for the generation of these data. In particular, the proposed network reconstruction method is designed to be robust with respect to noise and dynamic uncertainties (such as unmodelled nonlinearities). The proposed network structure reconstruction method explicitly takes into account the dynamics of hidden states without assuming knowledge of their number or location in the network. The key idea underlying our results is to find minimal distances between the existent data and the data that would have been required to obtain particular network structures. The smallest of these distances then allows to identify the (or a set of) network candidate(s) that most likely would be responsible for these data. The results are obtained using the newly-defined concept of *dynamical structure functions*.

As a first step, the proposed method has been designed for Linear Time-Invariant systems and therefore has limitations when applied to nonlinear or time-variant systems. However, when applied to the reconstruction of various linearisations of time-invariant nonlinear models given in the literature ([7],[8]), we observed that, using our method, network reconstruction was always possible provided the signal-to-noise ratio of the measured data was sufficiently large. In future works, we plan to extend the proposed network reconstruction approach to nonlinear, time-invariant, and delayed networked systems.

V. Acknowledgement

This work was supported in part by EPSRC grant number EP/G066477/1 and by Microsoft Research through the PhD Scholarship Program.

References

[1] E.D. Sontag, "Network reconstruction based on steady-state data", Essays in Biochemistry, 45:161-176, 2008.

[2] J. Gonçalves and S. Warnick, "Necessary and sufficient conditions for dynamical structure reconstruction of LTI networks", IEEE Transactions on Automatic Control, vol. 53, 2008.

[3] K. Zhou, J. Doyle and K. Glover, "*Robust and Optimal Control*," Prentice Hall, 1996.

[4] A. Hirotugu, "A new look at the statistical model identification," IEEE Transactions on Automatic Control 19 (6): 716-723, 1974.

[5] A. Julius, M. Zavlanos, S. Boyd, and G. Pappas, "Genetic Network Identification using Convex Programming," IET Systems Biology, 3(3):155-166, May 2009.

[6] L. Ljung. *System Identification–Theory for the User*. Prentice Hall, 1999.

[7] M. Xiong, J. Li and X. Fang, "Identification of genetic networks," Genetics, 2004.

[8] D. Di Bernardo, T. Gardner and J. Collins, "Robust identification of large genetic networks." Pac Symp Biocomput; 486-497, 2004.

[9] N. Young, *An introduction to Hilbert space*.Cambridge university press, 1988.

[10] K. Burnham and D. Anderson, *Model Selection and Inference - A practical information-theoretic approach*.Springer-Verlag, 1998.