Proceedings of the
46th IEEE Conference on Decision and Control
New Orleans, LA, USA, Dec. 12-14, 2007



# Dynamical Structure Functions for the Reverse Engineering of LTI Networks

Jorge Gonçalves*
Control Group
Department of Engineering
Cambridge University
Cambridge CB2 1PZ, UK
jmg77@cam.ac.uk

Russell Howes, Sean Warnick*
Information and Decision Algorithm Laboratories
Department of Computer Science
Brigham Young University
Provo, UT 84602, USA
sean@cs.byu.edu

*Abstract*— **This research explores the role and representation of network structure for LTI systems with partial state observations. We demonstrate that input-output representations, i.e. transfer functions, contain no internal structural information of the system. We further show that neither the additional knowledge of system order nor minimality of the true realization is generally sufficient to characterize network structure. We then introduce dynamical structure functions as an alternative, graphical-model based representation of LTI systems that contain both dynamical and structural information of the system. The main result uses dynamical structure to precisely characterize the additional information required to obtain network structure from the transfer function of the system.**

## I. INTRODUCTION

One of the fundamental issues for modeling, identifying, and controlling complex networked systems is inferring system structure, or reverse engineering, from input-output data. Structure is often the key for understanding a variety of complex systems because it enables a decomposition of the complete system into an interconnection of subsystems. When analysis of the subsystems is comparatively simple, and the interconnection structure is well understood, then the behavior of the complex system can be deduced from an understanding of its components. Moreover, exploiting structural information can tremendously reduce the conservatism of robust solutions designed to compensate for system uncertainty. This impact on complexity and uncertainty makes structural information extremely important in the analysis of complex networked systems.

Examples of scientists working on identifying or exploiting network structure arise in a variety of disciplines. Social scientists have developed a rich literature on the use of network models to describe interpersonal associations, perhaps one of the most famous works being Milgram's "small world" experiment in the 1960's in which letters passed from person to person were able to reach a particular target individual in only about six steps [16]. More recently, attention has focused on networks of business communities [5], [14], [17], internet-enabled virtual communities [11], citation networks in scientific communities [19], [21], preference networks for product recommender systems [7], [20], distribution networks [1], and the detection and destabilization of terrorist networks [3]. Epidemiologists have developed models for the dynamics of both epidemic and endemic diseases spreading through population networks [10], computer scientists have developed algorithms for searching over networks that are deployed in a number of popular applications [2], and biologists use microarray and other data sources to infer the regulation structure in genomic, proteomic, and metabolic networks [9], [15], [22], [23] .

Discovering structure from data, however, can be difficult. Typical identification methods do not emphasize structure estimation, but rather focus on behavior generalization by selecting a model that accurately predicts system outputs for unobserved inputs. As long as the dynamic behavior of the system is accurately described, the question of structure is often avoided altogether. For many applications, various model structures for the same input-output map are equally useful for forecasting and control. Nevertheless, sometimes it is important not only to describe the system dynamics accurately, but to do so with a model that correctly represents the structure of the original system.

In contrast with these identification methods that emphasis system dynamics over structure, inference methods have been developed that emphasis structure over dynamics. These methods employ graphical models to describe network structure. Nodes represent system states, understood to be random variables, and edges indicate conditional dependence between variables. Using Bayes rule, measurements can then be used to update prior distributions believed to characterize relationships throughout the network. A rich literature has grown in this area, and even issues of inferring causality from correlation have been addressed at some level [12], [13], [18].

Nevertheless, although these Bayesian Networks provide an efficient way to parameterize the joint probability distribution characterizing the entire system, conditional probabilities do not capture system dynamics, and the most successful inference techniques only work on directed acyclic graphs [4]. For some applications, such as modeling the citation network for a particular body of research, assuming the network is acyclic is reasonable since papers generally only

*These authors contributed equally to this project.

cite previously published work. There are many applications, however, such as modeling biological or social or economic networks, where such an assumption insisting on the absence of feedback dependencies between system states would be entirely unreasonable. Moreover, often an accurate representation of system dynamics is as important as that of system structure. In these situations, new methods are needed.

This paper introduces dynamical structure functions as a representation of LTI systems that captures both the dynamics and graphical structure of a system or network. The next section motivates these representations by demonstrating that transfer functions contain no internal structural information of a system, and they are generally not sufficient to obtain structure even when system order or minimality of the true system realization is known. We then introduce the dynamical structure function of an LTI system in Section III and discuss its properties. Section IV then uses this representation to precisely characterize the additional information required to obtain dynamical structure from the transfer function of the system. A discussion illustrating various corollaries and implications of the result concludes the paper.

## II. MOTIVATION

We begin our discussion by introducing the concept of the network structure of a dynamical system and illustrating that input-output representations of LTI systems, i.e. transfer functions, contain no internal structural information of the system. We go on to show that even when system order is known, or when the system is assumed to be minimal, the transfer function is still not generally sufficient information to obtain the network structure of the system. Although these concepts will be made precise in later sections of the paper, we illustrate the ideas here with some motivating examples.

### A. Network structure of a dynamical system

The network structure of a dynamical system is a description of the causal dependencies between system variables. These dependencies are typically represented by a directed graph where variables of the system are nodes, and an arrow between nodes indicates a causal relationship between variables.

In an input-output setting, the structure of a dynamical system is a graph from inputs to outputs (the "control structure") superimposed on a graph relating output variables to each other (the "internal structure of the measured states"). Although these concepts will be made precise in the next section, let us illustrate these ideas with a simple example. Consider the following system:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} -1 & 0 & -1 \\ 1 & -2 & 0 \\ 0 & 3 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 1 & 3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix},$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

This system has the "ring" structure shown in Figure 1. Observe that the state-space description completely characterizes the network structure.

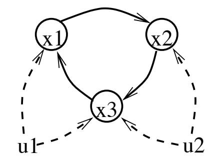


Fig. 1. System structure is composed of the internal structure (solid) and the control structure (dashed).

### B. Transfer functions do not imply internal structure

Since transfer functions are input-output representations of a system, it is not surprising that they do not fully characterize the internal structure of a system. One might think, however, that some structural information could be derived from the transfer function of a system. Nevertheless, as will be proven later, it turns out that every transfer function $G$ has a state-space realization that is consistent with any possible internal structure. Moreover, this is true even while preserving the same output equations of the system.

We illustrate this fact with a simple example. Consider a system with the following transfer function:

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

It can be shown that this transfer function is consistent with two systems with very different internal structures, given by

$$A_1 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & -1 \end{bmatrix}, \; B_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \; C_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

and

$$A_2 = \begin{bmatrix} -1 & 0 \\ 1 & -2 \end{bmatrix}, \; B_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \; C_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

The networks in Figure 2 correspond to each of the indicated realizations of $G$. It is easy to show that the remaining pair of distinct internal structures ($y_1 \leftarrow y_2$ and $y_1 \rightleftharpoons y_2$) can likewise be obtained from suitable realizations of $G$.

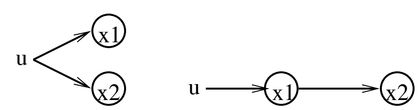


Fig. 2. Two possible networks given the transfer function: decoupled internal structure (left) and coupled internal structure (right).

### C. Neither order nor minimality determine structure

Although network structure can not be recovered from a transfer function without additional information or assumptions, it seems plausible that specification of the system order may lead to network reconstruction. Certainly there is a well-developed body of literature exploring techniques for estimating the order of a system from input-output data, suggesting that one may obtain a reasonable estimate of the order of the system when identifying the system.

Nevertheless, order in and of itself is not enough information to reconstruct the system network from its transfer function. On the other hand, if, in addition, one also knows a bijective relationship between the measured outputs and the system states, then the network structure is determined precisely by the transfer function. Such special conditions enabling measurement of the entire state vector, however, appear to be unreasonable in most situations.

Another plausible assertion is that knowledge of the minimality of the system is sufficient to recover system structure from its transfer function. Certainly in the absence of concrete information to the contrary, Occam's Razor would lead us to presume that the actual realization of the system would be minimal. Nevertheless, as the following example illustrates, even minimal realizations of simple systems with known output equations can have wildly different network structures. Consider a system with the following transfer function:

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

It can be shown that this transfer function is consistent with two systems with very different internal structures, given by

$$A_1 = \begin{bmatrix} -1 & 0 & 1 \\ 0 & -2 & 1 \\ 0 & 0 & -3 \end{bmatrix}, \; B_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

$$A_2 = \begin{bmatrix} -2 & -1 & 1 \\ -1 & -3 & 1 \\ 0 & -1 & -1 \end{bmatrix}, \; B_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

and

$$C_1 = C_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

The networks in Figure 3 correspond to each of the indicated realizations of $G$. Note that both realizations are minimal.

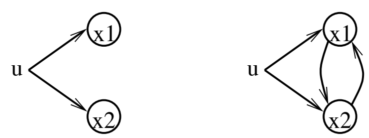


Fig. 3. Two possible networks corresponding to minimal realisations of the transfer function: decoupled (left) and coupled (right) internal structure.

### D. Problem statement

The fact that network structure can not be specified from a given transfer function, even when the order of the system is specified or an assumption of minimality is made, motivates the question as to exactly what information is needed to reconstruct an LTI network from its transfer function. The remainder of the paper addresses this issue.

## III. Dynamical structure

To formulate the network reconstruction problem, we consider the system given by

$$\begin{aligned} \begin{bmatrix} \dot{y} \\ \dot{x}_h \end{bmatrix} &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} y \\ x_h \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u \\ y &= \begin{bmatrix} I & 0 \end{bmatrix} \begin{bmatrix} y \\ x_h \end{bmatrix} \end{aligned} \quad (2)$$

where $x = \begin{bmatrix} y' & x_h' \end{bmatrix}' \in \mathbb{R}^n$, is the full state vector, $y \in \mathbb{R}^p$ is a partial measurement of the state, $x_h$ are the $n-p$ "hidden" states, and $u \in \mathbb{R}^m$ is the control input. In this work we restrict our attention to situations where output measurements constitute partial state information. [8] Note that in many applications the most sensible description of the system is in terms of the measured outputs as states, although rarely can we measure all the states of the system.

We have seen that the state-space realization of a system completely determines structure. Nevertheless, this refined description of the structure is too detailed–it will be as hard to recover from the transfer function as the state-space description itself. We would like a notion of network structure at the resolution of our measurements, something that suppresses information about the hidden states but accurately captures the interaction structure between measured states (internal structure) and the inputs and measured states (control structure). In the sequel we derive expressions for these structural representations.

Taking Laplace Transforms of the signals in (2), we find

$$\begin{bmatrix} sY \\ sX_h \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} Y \\ X_h \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} U \quad (3)$$

Solving for $X_h$, gives $X_h = (sI - A_{22})^{-1} A_{21} Y + (sI - A_{22})^{-1} B_2 U$. Substituting into the first equation of (3) then yields $sY = WY + VU$, where $W = A_{11} + A_{12}(sI - A_{22})^{-1} A_{21}$ and $V = A_{12}(sI - A_{22})^{-1} B_2 + B_1$. Let $D$ be a matrix with the diagonal term of $W$, i.e. $D = \text{diag}(W_{11}, W_{22}, ..., W_{pp})$. Then, $(sI - D)Y = (W - D)Y + VU$. Note that $W - D$ is a matrix with zeros on its diagonal. We then have

$$Y = QY + PU \quad (4)$$

where

$$Q = (sI - D)^{-1}(W - D) \quad (5)$$

and

$$P = (sI - D)^{-1} V \quad (6)$$

The matrix $Q$ is a matrix of transfer functions from $Y_i$ to $Y_j$, $i \neq j$, or relating each measured signal to all other measured signals (note that $Q$ is zero on the diagonal). Likewise, $P$ is a matrix of transfer functions from each input to each output without depending on any additional measured state $Y_i$.

These matrices, $Q$ and $P$, are precisely the structural representations we were seeking. Compare and contrast the entries in $Q$ or $P$ with those in $G$. While the elements of these matrices are all rational polynomials and have interpretations as transfer functions, they are, in fact quite different. For example, $G_{ij}$ is indeed the transfer function describing how the $j^{th}$ input impacts the $i^{th}$ output, nevertheless this effect may couple together influences of many of the states of the system, including other measured states. On the other hand, $P_{ij}$ denotes the direct impact of the $j^{th}$ on the $i^{th}$ output, where direct is understood to mean exclusive of the other measured states. This element of $P$ is, of course, a rational function in $s$ and subsequently has dynamics, but these dynamics represent the action of some hidden states or that of the $i^{th}$ measured state, never the interaction with another measured state. From this we consider the following definitions:

*Definition 1:* Given the system (2), we define the *Dynamical Structure Function* of the system to be $(Q, P)$, where

$Q$ and $P$ are the *Internal Structure* and *Control Structure*, respectively, and given as in (5) and (6).

*Definition 2:* Consider the system (2) with associated dynamical structure function, $(Q,P)$, and transfer function, $G$. We say the system's dynamical structure can be *reconstructed* if $(Q,P)$ can be derived from $G$.

This definition of dynamical structure and its reconstruction enable a precise exploration of its properties. The following lemmata demonstrate that $(Q,P)$ does, in fact, have the interpretation of structure in a sense that is generally less detailed than the complete state-space realization of the system, but more detailed than the transfer function of the system.

*Lemma 1:* The dynamical structure function of any system (2) exists and is unique.

This fact is true by construction of $Q$ and $P$ and ensures that the dynamical structure function of a system is well defined.

*Lemma 2:* Consider the system (2). If $p = n$ then the dynamical structure can be reconstructed.

This follows the fact that if $p = n$ in (2), then the associated transfer function $G$ has a unique realization and thus, a unique dynamical structure.

*Lemma 3:* Consider the system (2) with $p < n$. The dynamical structure function $(Q,P)$ is invariant to changes of coordinates on the hidden states, $z_h = Tx_h$ with $T$ invertible.

*Proof:* The change of coordinates yields a new system given by

$$\begin{aligned} \begin{bmatrix} \dot{y} \\ \dot{z}_h \end{bmatrix} &= \begin{bmatrix} A_{11} & A_{12}T^{-1} \\ TA_{21} & TA_{22}T^{-1} \end{bmatrix} \begin{bmatrix} y \\ z_h \end{bmatrix} + \begin{bmatrix} B_1 \\ TB_2 \end{bmatrix} u \\ y &= \begin{bmatrix} I & 0 \end{bmatrix} \begin{bmatrix} y \\ z_h \end{bmatrix}. \end{aligned}$$

Construction of the dynamical structure function for this system reveals it to be precisely that of the untransformed system (2). ∎

From this lemma we see that when hidden states are present, the dynamical structure representation of a system contains strictly less information than its state-space description. In particular, the dynamical structure function is invariant to any change of coordinates (and corresponding change of structure) only involving the hidden states, showing that this information is suppressed in this description of the system.

*Lemma 4:* The transfer function, $G$, of the system (2), is related to its dynamical structure, $(Q,P)$, by

$$G = (I-Q)^{-1}P. \tag{7}$$

This fact follows directly from (4) and $Y = GU$. The lemma demonstrates that dynamical structure can be interpreted as a factorization of the system transfer function, and that dynamical structure has more information about the system than $G$ alone. The main result in the next section precisely characterizes the additional information contained in the dynamical structure representation of a system beyond that of its associated transfer function, thereby providing conditions when dynamical structure can be reconstructed.

*Lemma 5:* $\text{Rank}(P) = \text{rank}(G)$.

This follows immediately from the fact $I-Q$ is invertible.

*Lemma 6:* Every entry of the dynamical structure function, $Q_{ij}$ or $P_{kl}$, is a strictly proper function.

Strict properness follows from the fact that $(sI-D)^{-1}$ (which is strictly proper) is multiplying transfer functions in $W$ or $V$ that are at most proper (never *improper*). This fact is important for the interpretation of $Q$ and $P$ as network *structures*. The strict properness of non-zero entries of these matrix functions imply causal relations, consistent with a directed edge interpretation between nodes on a graph. Moreover, zeros of the matrix functions indicate the absence of a direct causal relationship between nodes, even when considering hidden states of the system; only indirect paths through other measured states $y$ are possible. In this way, $Q$ and $P$ can be seen to capture precisely the structure of the system at the resolution of the given measurements without imposing any kind of structure on the hidden part of the system.

An equivalent representation of equation (7) is

$$(I-Q)G = P \tag{8}$$

which can also be written as as $\mathscr{A}\vec{x} = \vec{g}$, where $\mathscr{A}$ is a functions of the elements of $G$ only, $\vec{x} = (\vec{q}, \vec{p})$ is a vector stacked with the rows of $Q$ (excluding the diagonal of $Q$) and $P$, respectively, and $\vec{g}$ is a vector of the elements of $G$ stacked similarly. The operator $\mathscr{A}$ has $pm$ rows and $pm + p^2 - p$ columns.

*Lemma 7:* The dimension of the range of $\mathscr{A}$ is $pm$.

The result follows from the fact that $\mathscr{A}$ can be written as $\mathscr{A} = [F(G) \quad I]$, where $F(G)$ is a linear matrix function of $G$, and therefore $\mathscr{A}$ has full row rank. This lemma says that every $G$ is in the range of the operator $\mathscr{A}$, guaranteeing that there will always be at least one $(Q,P)$ pair consistent with $G$. Note that this result agrees with Lemma 4.

*Lemma 8:* The operator $\mathscr{A}$ has a null space of dimension $p^2 - p$. Moreover, $(Q,P)$ is an element of the null space of $\mathscr{A}$ if and only if $P = -QG$.

Since the operator $\mathscr{A}$ has $pm$ rows and $pm + p^2 - p$ columns and is rank $pm$, it has a null space of dimension $p^2 - p$. The second result is an immediate consequence of $G = P + QG$. This null space completely characterizes all dynamical structures $(Q,P)$ which are consistent with a given transfer function $G$.

Equipped with this representation of an LTI system and a precise notion of network reconstruction, we are now prepared to fully characterize the conditions when a system's dynamical structure can be reconstructed. We focus our attention on this task in the next section.

## IV. Dynamical Structure Reconstruction

Dynamical structure reconstruction deals with finding the "true" dynamical structure of a system given its transfer function. We begin with a complete state-space system as in (2). Presumably this system is unknown, but it generated enough data for us to estimate its transfer function $G$. Ultimately we would like to know the dynamical structure function of this system, so the question becomes exactly what information do we need above and beyond the transfer function to uniquely specify it.

Lemma 8 makes the problem of uniquely specifying dynamic structure given a transfer function clear. Essentially, one simply needs to specify which element of the null space characterizes the correct structure.

*Theorem 1:* Consider a system (2) and assume the only available information is its associated transfer function $G$. The dynamical structure can be reconstructed if and only if the component of $(Q,P)$ in the null-space of $\mathscr{A}$ is known.

*Proof:* The "if" part follows from the fact that $\vec{g}$ completely specifies the element of $(Q,P)$ in the row space of $\mathscr{A}$ (see Figure 4). Thus, knowledge of the null space component of $(Q,P)$ uniquely species the dynamical structure of (2).

For the "only if"part, let $\vec{x}_1$ satisfy $\mathscr{A}\vec{x}_1 = \vec{g}$ and $\vec{x}_N \neq 0$ be an element of the null space of $\mathscr{A}$. Then, there exists a large enough nonnegative integer $k$ such that

$$\vec{x}_2 = \vec{x}_1 + \vec{x}_N \frac{1}{(s+1)^k}$$

which is also a solution of $\mathscr{A}\vec{x}_2 = \vec{g}$ and all the elements in $\vec{x}_2$ are strictly causal. We have then found another dynamical structure, i.e., another set of strictly proper $Q_1 \neq Q_2$ and $P_1 \neq P_2$ consistent with $G$. Thus, the dynamical structure cannot be reconstructed uniquely, which is a contradiction. ■

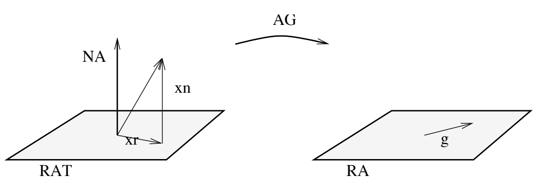


Fig. 4. The component of dynamical structure in the row space of $\mathscr{A}$, $\vec{x}_R$, is completely specified by the transfer function $\vec{g}$. Thus, $\vec{x}_N$ represents the additional information in dynamical structure.

The implications of the theorem are that we must know $p^2 - p$ transfer functions in order to reconstruct dynamical structure from $G$. Which $p^2 - p$ transfer functions need to be known, however, depends on the basis of the null space of $\mathscr{A}$. The following corollary, illustrates an important special case of this result when a transfer function will not have an entire row or column of zeros. This situation is typical for many systems.

*Corollary 1:* Consider a system (2) and assume the only available information is its associated transfer function $G$, which does not have any rows or columns that are entirely zero. The dynamical structure can be reconstructed if and only if at least $p^2 - p$ transfer functions of $Q$ or $P$ are known.

*Proof:* From Lemma 8, the knowledge of $p^2 - p$ elements between $Q$ and $P$ completely specifies the component of $(Q,P)$ in the null space from the equation $P = -QG$. The proof then follows from Theorem 1. ■

These results provide necessary and sufficient conditions for the dynamical structure reconstruction of LTI networks. Next we consider some important special cases of this theorem.

## V. Discussion

Having identified precisely what must be known to specify dynamical structure from a transfer function (which, in turn, is assumed to have been identified from data), we now turn our attention to some important special cases or implications of these results.

### *A. Every transfer function admits any internal structure*

This fact was illustrated by example in Section II-B. Now the result follows trivially from previous results.

*Corollary 2:* Given $G$ and $Q$ there is a $P$ such that $(Q,P)$ is consistent with $G$.

This follows from (8) and it shows that without additional information besides $G$, any internal structure is possible. In particular, any transfer function has a realization with an internal structure that is completely decoupled (i.e. $Q = 0$) or fully connected.

### *B. Transfer functions with zero rows or columns*

Zero column vectors in $G$ can be removed without loss of generality since the corresponding inputs have no impact on the system. To see this, after rearranging the terms in $G$ this can be decomposed into $G = \begin{bmatrix} G_1 & 0 \end{bmatrix}$ where $G_1$ has no zero column vectors. Decompose $P$ in a similar way $P = \begin{bmatrix} P_1 & P_2 \end{bmatrix}$. From (8) we have $(I-Q)\begin{bmatrix} G_1 & 0 \end{bmatrix} = \begin{bmatrix} P_1 & P_2 \end{bmatrix}$ from which we conclude that $(I-Q)G_1 = P_1$ and $P_2 = 0$. This basically says that the inputs corresponding to the zero vector entries in $G$ are vacuous and can be removed from the system. For reminder of the paper, we assume $G$ has no zero column vectors.

Assume now $G$ has zero row vectors. After rearranging the terms in $G$ this can be decomposed into $G = [G_1; 0]$ where $G_1$ has no zero row vectors. Decompose also $Q$ and $P$ as $Q = [Q_{11}\ Q_{12}; Q_{21}\ Q_{22}]$ and $P = [P_1; P_2]$. From (8) we get two equations: $(I - Q_{11})G_{11} = P_1$ and $-Q_{21}G_{11} = P_2$ with $Q_{12}$ and $Q_{22}$ are undetermined. Therefore network reconstruction demands *a priori* knowledge of $Q_{12}$ and $Q_{22}$.

### *C. Unknown internal network structure*

This sections assumes there is no information on the internal structure (i.e. no information on $Q$). Thus, we assume that in addition to $G$ having no column vectors, $G$ must also have no row vectors or otherwise the internal structure cannot be reconstructed (in particular, $Q_{12}$ and $Q_{22}$ can note be recovered as from the discussion in Section V-B). We consider the following four cases: $m < p-1$ (there are 2 or less inputs than measured states), $m = p-1$ (1 less input than measured states), $m = p$ (same number of inputs and measured states) and $m > p$ (more inputs than measured states).

*1) $m < p-1$: two or less inputs than measured states:* If $m < p-1$ and there is no information on the internal structure, then the dynamical structure cannot be recovered. In this case, the total number of known transfer functions (all from $P$) is upper bounded by $mp < p^2 - p$, which does not satisfy the condition of Corollary 1 that is necessary for reconstruction.

*2) $m = p-1$: one less input than measured states:* If $m = p-1$ and there is no information on the internal structure, then the dynamical structure can be recovered if and only if $P$ is precisely known. To see this, the total number of known transfer functions (all from $P$) is $mp = p^2 - p$ which satisfies the condition of Corollary 1.

For example, consider the system from Section II-B with $p = 2$ measured states, $m = 1$ control input and $G = (G_{11}, G_{21})$ given by (1). With no knowledge of $P$, several different networks satisfy (8), which has two equations and four unknowns:

$$\begin{cases} G_{11} - Q_{12}G_{21} &= P_{11} \\ G_{21} - Q_{21}G_{11} &= P_{21} \end{cases}$$

giving 2 degrees of freedom between the internal structure ($Q_{12}$ and $Q_{21}$) and the control structure ($P_{11}$ and $P_{21}$). For example, a possible solution is to set $Q_{12} = Q_{21} = 0$, i.e. decoupled network between $y_1$ and $y_2$. In that case, $P_{11} = G_{11}$ and $P_{21} = G_{21}$ (left of Figure 2). Note the hidden state in $P_{21}$ ($P_{21}$ is second-order) and the system is not controllable (due to the common pole at $-1$), which explains why $G$ is second-order and there are three states. An alternative is $P_{21} = 0$ (which fixes $Q_{21} = G_{21}/G_{11}$) and $Q_{12} = 0$ (which fixes $P_{11} = G_{11}$), which can be seen on the right of Figure 2. Note that in this case $P_{11} \neq 0$ or otherwise $Q_{12}$ would be non-proper.

*3) $m = p$: same number of inputs and measured states:* With no knowledge of $Q$, to satisfy the condition of Corollary 1, we must know $p^2 - p$ transfer functions in $P$. Knowing *any* nonzero transfer function in $P$ requires much more information than knowing that an entry in $P$ equals zero, since specifying a nonzero transfer function demands characterising all the poles, zeros, and gain of that particular transfer function. As a result, requiring knowledge of nonzero transfer functions to satisfy the conditions of Corollary 1 may be an unreasonable assumption. However, knowing some structure in $P$, such as the location of some of its zeros may be more reasonable.

Consider the case where $P$ is all zero except for $p$ nonzero transfer functions. Here we restrict our attention to the case where $G$ is invertible, as this results in a very useful closed-form solution for $(Q, P)$. The next result assumes no knowledge of $Q$ or any of the $p$ nonzero transfer functions in $P$.

*Corollary 3:* If $m = p$, $G$ is full rank, and there is no information on the internal structure and on nonzero transfer functions in $P$, then the dynamical structure can be reconstructed if and only if each input controls a measured state independently, i.e. $P_{ij} = 0$ for $i \neq j$. Moreover, $H = G^{-1}$ characterizes the dynamical structure as follows

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

*Proof:* The "if" part of the proof follows from the fact that there are $p^2 - p$ known transfer functions (equal to zero) in $P$ and using Corollary 1. Multiplying $(I - Q)G = P$ on the right by $H = G^{-1}$ yields $I - Q = PH$. Since $Q$ has zeros on its diagonal and $P$ is diagonal, we have $1 = P_{ii}H_{ii}$ or $P_{ii} = 1/H_{ii}$. Finally we can now solve for $Q = I - PH$ and the result follows.

For the "only if" part, using Corollary 1 again shows that $p^2 - p$ entries in $P$ must be known. Because there is no knowledge of nonzero transfer functions in $P$, then $p^2 - p$ entries in $P$ must be zero. By Lemma 5, $\text{rank}(P) = \text{rank}(G) = p$. Thus, there must be exactly $p$ unknown and nonzero entries in $P$. Since $P$ is full rank, each row and column must have exactly one of these entries. Without loss of generality the inputs can be renamed and reordered so that the diagonal of $P$ contains the unknown and nonzero entries. ■

This result says that in addition to having a square and full rank $G$ it is necessary and sufficient to know that each control $i$ affects first state $i$ before it affects any other measurable state to reconstruct the dynamical structure. That allows us to reduce the number of unknowns to $p^2 - p + p = p^2$ which can now be solved.

However, if there is some *a priori* information about the internal structure (such as some of the $Q_{ij} = 0$) then there is more flexibility and less information and constraints are required for $P_{ij}$. As long as there are only $p^2$ nonzero elements between $Q_{ij}$ and $P_{ij}$ then the dynamical structure can be reconstructed by solving the linear system of equations (8).

If $P$ is not diagonal and there is additional information on how the inputs affect the measured states, there may be a change of basis in the control vector that allows it to be converted to a diagonal matrix that can then be used in Theorem 1. For example, if $x_1$ is controlled by $u_1 + u_2$ and $x_2$ by $u_1 - u_2$ then one could define two new input vectors $v_1 = u_1 + u_2$ and $v_2 = u_1 - u_2$.

*4) $m > p$: more inputs than measured states:* From Theorem 1, we need to know $p^2 - p$ from the $pm > p^2$ elements in $P$. It may seem intuitive that if there are more inputs then there should be more information. However, the extra inputs are in a way redundant. The reason is the fact that although $G$ is $p \times m$, $\text{rank}(G) = p$, which means that the inputs really only have $p$ degrees of freedom. Thus, the problem reduces to having the same number of inputs as measured states. The difference here is that we may be able to choose from the $m$ inputs $p$ that are known to control directly each measurable state.

### D. The danger of steady-state measurements

This section clarifies some misconceptions that arise from time to time in some scientific communities concerning network reconstruction using steady-state data. For instance, in [6] the authors proposed a method to estimate networks based on full state measurement and control (i.e. $n = p$ and $B = I$). In this special case, for $i \neq j$, $H_{ij} = -a_{ij}$, since $G(s) = (sI - A)^{-1}$ or $H(s) = sI - A$. Thus, $a_{ij} = 0$ $(i \neq j)$ if and only if $H_{ij} = 0$, which means the network structure can be obtained from $H$.

However, in the realistic case there are less measurements and control available than states. If instead of estimating $G$ from time-series data we were to use only steady-state data, this could lead to mistakes as the following example shows.

Consider a third order system with measurements and control on the first 2 states $x_1$ and $x_2$ and the following dynamics

$$A = \begin{bmatrix} -1 & 1 & -1 \\ 1 & -1 & -1 \\ 1 & 1 & -1 \end{bmatrix}$$

This is a fully connected network, and so we expect the reduced network consisting on $x_1$ and $x_2$ to be fully connected as well. In this case,

$$H(s) = \begin{bmatrix} \frac{s^2+2s+2}{s+1} & -\frac{s}{s+1} \\ -\frac{s}{s+1} & \frac{s^2+2s+2}{s+1} \end{bmatrix}$$

When $s \to 0$,

$$H(0) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

which could lead one to think the reduced order network is not connected at all, i.e. $x_1$ does not affect $x_2$ and vice versa. In general, for third order systems this is always true if and only if $a_{12}a_{33} + a_{13}a_{32} = 0$ for the connection from $x_2$ to $x_1$ and $a_{21}a_{33} + a_{23}a_{31} = 0$ for the connection from $x_1$ to $x_2$. Note that even when these equalities are not exactly zero but near zero, the presence of noise may again lead to wrong decisions.

## VI. Conclusions

This paper discussed the role of network structure for LTI systems. In particular, it was shown that transfer functions alone contain no information about the internal structure of an LTI system. We then introduced a new representation for such systems, a factorization of the system's transfer function that we call dynamical structure. Dynamical structure functions contain more information about the system than the transfer function because they also describe the network structure between inputs and outputs. Nevertheless, dynamical structure contains less information about the system than its state-space description because no attempt is made to realize the network structure relating the non-measured, hidden state variables to the rest of the system. In this way, dynamical structure is a convenient analysis tool representing system information at a resolution consistent with its number of measured states, somewhere between a system's full state space realization (full structural information) and its transfer function (no structural information).

We then used dynamical structure to explore the network reconstruction problem. In this problem, one would like to estimate network structure given only a transfer function obtained from input-output data. This problem is extremely important for a variety of fields, such as biology or counter-terrorism, that attempt to draw structural conclusions from data. Necessary and sufficient conditions were presented that indicate that network reconstruction demands careful experiment design. Moreover, various examples were provided throughout the paper that demonstrate how failure to respect the necessary conditions may lead to incorrect conclusions about the network structure.

## VII. Acknowledgements

The authors would like to thank Glenn Vinnicombe for his valuable input and comments, especially regarding the notions of uniqueness and properness of $Q$ and $P$. Moreover, the authors would like to thank their sponsoring organizations: the NSF (grant DMS0639328), the Rollins Center for eBusiness, the BYU Bookstore, Park City Group, the Office of Research and Creative Activities at BYU, The Nuffield Foundation (grant NAL/01062/G), EPSRC (grant EP/E02761X/1), the College of Physical and Mathematical Sciences at BYU and The Nuffield Foundation for generously supporting Dr. Warnick's leave at Cambridge University.

## References

[1] L. A. N. Amaral, A. Scala, M. Barthelemy, and H. E. Stanley. Classes of small-world networks. *Proc. Natl. Acad. Sci. USA*, 97:11149–11152, 2000.

[2] S. Brin and L. Page. The anatomy of a large-scale hypertextual web search engine. *Computer Networks*, 30:107–117, 1998.

[3] K. M. Carley. Destabilization of covert networks. *Computational and Mathematical Organization Theory*, 12(1):51–66, April 2006.

[4] R. G. Cowell, A. P. Dawid, S. L. Lauritzen, and D. J. Spiegelhalter. *Probabilistic Networks and Expert Systems*. Springer-Verlag, 1999.

[5] J. Galaskiewicz and P. V. Marsden. Interorganizational resource networks: Formal patterns of overlap. *Social Science Research*, 7:89–107, 1978.

[6] T. Gardner, D. Bernardo, D. Lorenz, and J. Collins. Inferring genetic networks and identifying compound mode of action via expression profiling. *Science*, 301:102–105, July 2003.

[7] D. Goldberg, D. Nichols, B. M. Oki, and D. Terry. Using collaborative filtering to weave an information tapestry. *Comm. ACM*, 35:61–70, 1992.

[8] J. Gonçalves and S. Warnick. Necessary and sufficient conditions for dynamical structure reconstruction of LTI networks. *Submitted to IEEE Transactions of Automatic Control*, 2007.

[9] N. Guelzim, S. Bottani, P. Bourgine, and F. Kepes. Topological and causal structure of the yeast transcriptional regulatory network. *Nature Genetics*, 31:60–63, 2002.

[10] H. W. Hethcote. The mathematics of infectious diseases. *SIAM Review*, 42:599–653, 2000.

[11] P. Holme, C. R. Edling, and F. Liljeros. Structure and time-evolution of an internet dating community. *Social Networks*, 26:155–174, 2004.

[12] F. V. Jensen. *Bayesian Networks and Decision Diagrams*. Springer, 2001.

[13] M. I. Jordan. *Learning in Graphical Models*. MIT Press, 1998.

[14] P. Mariolis. Interlocking directorates and control of corporations: The theory of bank control. *Social Science Quarterly*, 56:425–439, 1975.

[15] S. Maslov and K. Sneppen. Specificity and stability in topology of protein networks. *Science*, 296:910–913, 2002.

[16] S. Milgram. The small world problem. *Psychology Today*, 2:60–67, 1967.

[17] M. S. Mizruchi. *The American Corporate Network, 1904-1974*. Sage, 1982.

[18] J. Pearl. *Causality*. Cambridge, 2000.

[19] S. Redner. How popular is your paper? an empirical study of the citation distribution. *Eur. Phys. J. B*, 4:131–134, 1998.

[20] P. Resnick and H. R. Varian. Recommender systems. *comm. ACM*, 40:56–58, 1997.

[21] P. O. Seglen. The skewness of science. *J. Amer. Soc. Inform. Sci.*, 43:628–638, 1992.

[22] J. Stelling, S. Klamt, K. Bettenbrock, S. Schuster, and E. D. Gilles. Metabolic network structure determines key aspects of functionality and regulation. *Nature*, 420:190–193, 2002.

[23] P. Uetz, L. Giot, G. Cagney, T. A. Mansfield, R. S. Judson, J. R. Knight, D. Lockshon, V. Narayan, M. Srinivasan, P. Pochart, A. Qureshi-Emili, Y. Li, B. Godwin, D. Conover, T. Kalbfleisch, G. Vijayadamodar, M. Yang, M. Johnston, S. Fields, and J. M. Rothberg. A comprehensive analysis of protein-protein interactions in saccharomyces cerevisiae. *Nature*, 403:623–627, 2000.