Preprints of the 8th IFAC Symposium on Advanced Control of Chemical Processes
The International Federation of Automatic Control
Furama Riverfront, Singapore, July 10-13, 2012

Reconstructing Plant Connectivity using Directed Spectral Decomposition ⋆

S. Gigi * **A. K. Tangirala** **

* *Indian Institute of Technology Madras, India (e-mail: gigi.sebastian@yahoo.com)*
** *Indian Institute of Technology Madras, India (Tel: 91-44-22574181; e-mail: arunkt@iitm.ac.in)*

Abstract: Process connectivity is a key information that is sought in a diverse set of applications ranging from design to fault diagnosis of engineering and biological processes. The present work develops a methodology for reconstruction of plant connectivity from dynamic data using directional spectral analysis, a novel adaptation of ideas from neurosciences and econometrics. The method is based on the concept of Granger causality while the procedure rests on the directional decomposition of power spectrum into direct and indirect energy transfers. The quantification of effective connectivity is obtained using a structural vector auto-regressive (SVAR) representation of the process. Results from simulation studies demonstrate the potential of the proposed method.

Keywords: connectivity; causality; vector autoregressive model; frequency domain; spectral analysis; directed energy transfer

1. INTRODUCTION

Chemical processes are characterized by interconnected unit operations involving large flow of materials, pumping, mixing, heating and cooling with and without chemical reactions. The numerous interconnections and recycles in process plants, coupled with heat integration operations add to the complexity of plant topology. The knowledge of plant connectivity is crucial in several applications, particularly in performance monitoring and fault diagnosis. The process flow diagram (PFD) depicts the pipings and connections among the major equipments of the plant. Incorporation of PFD information in the analysis of multivariable data requires conversion of the diagram to a form of connectivity matrix. Several methods have been proposed to diagnose and isolate root cause of fault propagation. Construction of signed digraph (SDG) [Maurya et al., 2003a,b], computer aided engineering exchange (CAEX) software [Sim et al., 2006] and adjacency matrix [Jiang et al., 2009] are a few of them. Practical application of such methods can be quite complicated owing to the time-consuming nature and the magnitude of human intervention that is required. Furthermore, a process flow diagram can only provide the structural information but not the strength of the connectivities. It is not uncommon that the flow diagram shows the existence of physical connectivity between two units, while the actual amount of material/energy transfer may be weak to nil. Such ambiguities can only be resolved through a careful analysis of measurements. The connectivity strengths can also change with the operating conditions which is once again not provided by the PFD. Thus, from both automation and practical viewpoints there exists a strong need for reconstructing pathways from process data. Needless to state, inferences from data should be corroborated with information available through process flow diagrams.

Reconstructing pathways from data is closely related to the causality analysis of a process. The problem of determining causality from observed data dates back to the preliminary efforts by Wiener [1956]. Granger [1969] adopted Wiener's ideas to give rise to a practically implementable definition of causality, now known as Granger causality [Granger, 1969]. The idea is based on temporal effects. A variable X is said to Granger-cause another variable Y if the prediction of Y is improved with the incorporation of the past of X. If the reverse relationship exists *i.e.*, if the inclusion of past values of Y improves the prediction of X, there exists a feedback relation between X and Y, also known as bi-directional causality. Granger causality does not include instantaneous causality. If the inclusion of present value of X into the prediction of Y results in improved values, there exists an instantaneous causality from X to Y. Instantaneous causality is harder to detect from data in the sense that it is unresolvable when it exists in both directions. Fortunately in process systems, mostly the instantaneous causality is uni-directional, *i.e.*, it exists only in the feedback path.

Connectivity analysis using Granger causality [Granger, 1969] has emerged as a major tool for examining information flow between brain regions [Smith et al., 2011, Baccala and Sameshima, 2001]. Over the last decade, a few powerful methods have emerged Hlavackova-Schindler et al. [2007] in neuroscience for detection of connectivities. Notable among such methods are the concepts of directed transfer function (DTF) and partial directed coherence (PDC). A good review of the related methods appears in Blinowska [2011].

⋆ Correspondence concerning this work should be addressed to Arun K. Tangirala at arunkt@iitm.ac.in

Granger's concept of causality in its original form did not rest on any particular model for prediction. However, the presentation recommended the use of the vector-autoregressive (VAR) representation of the multivariable process due to its ease of estimation and implementation. Frequency-domain extensions of these concepts were largely propelled by the works of Geweke [1982] and Geweke [1984]. In neurosciences, the problem of determining causal relationships in frequency domain spectral factorization were brought forth by the works of Saito and Harashima [1981], Kaminski and Blinowska [1991] which gave birth to the concept of directed transfer function (DTF). The DTF measures how much a white-noise source $e_j[k]$ (in a variable $x_j[k]$) affects a variable $x_i[k]$. Baccala and Sameshima [2001] showed that the DTF measures the total influence, which includes influences on both direct and indirect pathways and therefore is not suited for structure determination; instead they proposed an alternative measure known as the partial directed coherence (PDC). PDC measures the direct influence between two variables and is therefore suited for structure determination. However, PDC and its variants [Baccala et al., 2007, Astolfi et al., 2006, Schelter et al., 2009] aid in structure determination but do not quantify the strength of connectivities in the sense of energy transfers as shown in the work by Gigi and Tangirala [2010]. The authors [Gigi and Tangirala, 2010] present exact expressions for the connectivity strengths in terms of energy transfers in the frequency domain by showing that the total energy (variance) transfer at any frequency between two variables in a multivariable process consists of *direct*, *indirect* and *interference* terms at that frequency.

The methods outlined in the foregoing discussion rest on the linear representations. Non-linear measures of causal dependence have also been reported in the literature [Hlavackova-Schindler et al., 2007]. Development of causal maps from plant data is reported in Bauer et al. [2007] where the authors detect the directionality of disturbance propagation using transfer entropy. The success of the method is demonstrated on two industrial case studies; however its implementation can be fairly cumbersome for large process and sensitive to the parameters of the algorithm. The estimation of transfer entropy is based on probability distribution function (PDF), which generally is considered as rigorous to estimate. Moreover, transfer entropy is a bivariate measure. As in several other applications, linear measures are preferable because of their ease of implementation and robustness to noise despite their restricted ability in handling highly non-linear processes [Blinowska, 2011]. The present work makes use of linear measures, specifically those mentioned above.

The main contributions of this work are, (i) an automated method to detect process connectivity under open-loop and closed-loop conditions and (ii) a measure of connectivity strength in the sense of energy transfer.

The remainder of this article is organized as follows. Section 2 provides the theoretical background of the developments in this work and a brief review of the concepts of DTF and PDC. Expressions for quantitative energy transfers and connectivity strengths are presented in Section 3. Simulation results are presented in Section 4. The article concludes with Section 5.

2. THEORETICAL BACKGROUND

2.1 Spectral Factorization Theorem

The basis for the proposed method of analysis for plant topology is the well-known spectral factorization theorem [Gevers and Anderson, 1981]. The cross power spectral density matrix of a jointly (stationary) process $\mathbf{\Phi}(\omega)$ can be factored as

$$\mathbf{\Phi}(\omega) = \mathbf{H}(\omega)\Sigma_e\mathbf{H}^*(\omega) \tag{1}$$

where Σ_e is the covariance matrix of the innovations (white noise) driving the multivariate process. The term

$$\mathbf{H}(\omega) = \begin{bmatrix} h_{11}(\omega) & h_{12}(\omega) & \dots & h_{1m}(\omega) \\ h_{21}(\omega) & h_{22}(\omega) & \dots & h_{2m}(\omega) \\ \vdots & \vdots & \ddots & \vdots \\ h_{m1}(\omega) & h_{m2}(\omega) & \dots & h_{mm}(\omega) \end{bmatrix} \tag{2}$$

is the transfer function matrix in the frequency domain. The superscript $(.)^*$ denotes the Hermitian conjugate of a matrix. The factorization giving rise to $\mathbf{H}(\omega)$ is the key in determining the directionality of the energy transfer leading to topology construction. The spectral factor $\mathbf{H}(\omega)$ forms the vector moving average (VMA) model coefficient matrix, the calculation of which is achieved through time series vector auto-regressive (VAR) modelling.

2.2 VAR / VMA Models

Consider a multivariate process with m measurements denoted by the vector $\mathbf{x} = [\,x_1\ x_2\ \cdots\ x_m\,]^T$. Each variable is treated as an input as well as an output simultaneously. The idea is to express the evolution of the j^{th} random variable, x_j as a result of two components (i) a (fictitious) random input e_j and (ii) the linear influence of its own past as well as the past states of the other variables. Mathematically, the representation of VAR model [Lutkepohl, 2005] is

$$\mathbf{x}[k] = \sum_{r=1}^{p} \mathbf{A}_r\,\mathbf{x}[k-r] + \mathbf{e}[k] \tag{3}$$

where $\mathbf{A}_r$ is the matrix of auto-regressive coefficients at lag r and $\mathbf{e}[k]$ is an m-dimensional vector of white noise sequences. An important feature of the VAR model to be estimated is the model order (p). The optimum model order is determined using standard theoretic information criteria such as the Akaike Information Criterion (AIC) and the Schwarz information criteria (SIC).

In VMA model the present value of the variable is expressed purely in terms of past and present shocks. An M^{th} order VMA process [Shumway and Stoffer, 2000, Priestley, 1981] is represented as,

$$\mathbf{x}[k] = \sum_{r=1}^{M} \mathbf{H}_r\mathbf{e}[k-r] + \mathbf{e}[k] \tag{4}$$

where $\mathbf{H}_r$ is the VMA coefficient matrix. The estimation of $\mathbf{H}(\omega)$ is done from its relationship with VAR model in frequency domain, represented as

$$\mathbf{H}(\omega) = \bar{\mathbf{A}}^{-1}(\omega) \quad \text{where} \tag{5}$$

$$\bar{\mathbf{A}}(\omega) = I - \mathbf{A}(\omega) \quad \text{and} \tag{6}$$

$$\mathbf{A}(\omega) = \sum_{r=1}^{p} A_r z^{-r}|_{z=e^{j\omega}} \tag{7}$$

is the Fourier transform [Smith, 1997] of the VAR coefficients, A_r.

The model in (3) represents a strictly causal one *i.e.*, it does not account for instantaneous causality. In feedback control systems, the instantaneous causality is inevitable due to the presence of controller with proportional action. To account for instantaneous causality and correlated innovations, a *structural* VAR (SVAR) model given by,

$$\mathbf{A}_0\mathbf{x}[k] = \sum_{r=1}^{p} \tilde{\mathbf{A}}_r \,\mathbf{x}[k-r] + \mathbf{B}\mathbf{w}[k] \tag{8}$$

is appropriate. The matrix $\mathbf{A}_0$ contains the instantaneous relationships between the variables. The quantity $\mathbf{w}[k]$ is an m-dimensional vector of white noise sequences, with covariance matrix $\Sigma_w = \mathbf{I}_{m\times m}$. The SVAR model form is not unique since the multiplication by any positive definite matrix will give a mathematically valid model. This model can be converted to a unique reduced form as:

$$\begin{aligned}\mathbf{x}[k] &= \sum_{r=1}^{p} \mathbf{A}_0^{-1}\tilde{\mathbf{A}}_r \,\mathbf{x}[k-r] + \mathbf{A}_0^{-1}\mathbf{B}\mathbf{w}[k] \\ &= \sum_{r=1}^{p} \mathbf{A}_r \,\mathbf{x}[k-r] + \mathbf{e}[k]\end{aligned} \tag{9}$$

From (3) and (8), with $\Sigma_w = \mathbf{I}_{m\times m}$ we get

$$\Sigma_e = \mathbf{A}^{-1}\mathbf{B}\mathbf{B}^T\mathbf{A}^{-T} \tag{10}$$

The model is generally evaluated in the reduced form and then converted to the structural form using suitable techniques [Lutkepohl, 2005, Pfaffe and Taunus, 2008, Tsay, 2010].

The quantities $\mathbf{H}(\omega)$ and $\bar{\mathbf{A}}(\omega)$ are used in estimating the linear relationships among the variables of the process.

2.3 DTF and PDC

The estimation of DTF and PDC is carried out by VAR modelling of the joint stationary representation of the process. The basic quantity required for estimating DTF and PDC is the white noise transfer function matrix $\mathbf{H}(\omega)$. The DTF, $\gamma_{ij}(\omega)$ from x_j to x_i is calculated as

$$\gamma_{ij}(\omega) = \frac{h_{ij}(\omega)}{\sqrt{\sum_{j=1}^{m} |h_{ij}(\omega)|^2}} \tag{11}$$

where m is the dimensionality of the multivariate time-series. The DTF, being a decomposition of coherence provides the total (*i.e.*, the sum of direct and indirect) directional influence of variables.

The PDC is defined as

$$\pi_{ij}(\omega) = \frac{\bar{a}_{ij}(\omega)}{\sqrt{\bar{\mathbf{a}}^*_{.j}(\omega)\bar{\mathbf{a}}_{.j}(\omega)}} = \frac{\bar{a}_{ij}(\omega)}{\sqrt{\sum_{i=1}^{m} |\bar{a}_{ij}(\omega)|^2}} \tag{12}$$

It is a directional decomposition of partial coherence indicating the direct exchange of information.

Due to the normalization used, the sum of squared PDC values in a column at each frequency is unity. The same is applicable to the sum of squared DTF values at each frequency in a row. The normalization in the definition of DTF and PDC is a major difference between the two analysis techniques. The DTF is a quantitative measure of the net energy transfer [Eichler, 2006] whereas PDC is only a qualitative measure of direct energy transfer [Gigi and Tangirala, 2010]. Thus PDC can be used to determine the structure of the process. Even if a normalization similar to DTF as suggested in Schelter et al. [2009] is used for PDC, the magnitude of direct and indirect energy transfers cannot be obtained from a comparison of DTF and PDC values. The following section provides the methodology to arrive at the quantification of direct and indirect effects.

3. QUANTIFICATION OF CONNECTIVITY STRENGTHS

Reconstruction of connectivity from data requires three steps namely (i) detecting directed connectivity (ii) quantification of the strength of the connectivity and (iii) representation in suitable form.

3.1 Quantification of energy transfers

The quantification of direct and indirect influences in terms of energy transfer is obtained by developing the corresponding direct/indirect transfer functions. In order to derive the expressions for transfer functions, Gigi and Tangirala [2010] use a method based on signal-flow graph representation of the process as VAR/VMA models supported by the mathematical derivations of the associated spectral relationships. The results are presented here with essential descriptions.

Total transfer function,

$$h_{ij}(\omega) = \frac{(\mathrm{adj}(\bar{A}(\omega)))_{ij}}{\det(\bar{A}(\omega))} \tag{13}$$

Direct transfer function,

$$\begin{aligned}h_{D,ij}(\omega) &= \frac{-\bar{a}_{ij}(\omega)\det(\bar{M}_{ji}(\omega))}{\det(\bar{A}(\omega))} \text{ for } i \neq j \\ &= h_{ij}(\omega) \qquad\qquad \text{for } \quad i = j\end{aligned} \tag{14}$$

Indirect transfer function,

$$h_{I,ij}(\omega) = h_{ij}(\omega) - h_{D,ij}(\omega) \tag{15}$$

where $\bar{M}_{ij}(\omega)$ is the minor of the matrix $\bar{A}(\omega)$ and of size $(m-2)\times(m-2)$, obtained by eliminating both i^{th} and j^{th} row *and* column from $\bar{A}(\omega)$.

The magnitudes of energy transfers occuring directly and indirectly (through another variable) are obtained from the transfer functions. When there exists both direct and indirect transfer functions, there will be an interference term in the energy transfers, occuring due to the phase difference between the direct and indirect transfer functions [Gigi and Tangirala, 2010]. Hence, we have the total energy transferred from x_j to x_i, given by

$$\begin{aligned}|h_{ij}(\omega)|^2 &= (h_{D,ij}(\omega) + h_{I,ij}(\omega))(h_{D,ij}(\omega) + h_{I,ij}(\omega))^* \\ &= |h_{D,ij}(\omega)|^2 + |h_{I,ij}(\omega)|^2 + 2\Re(h_{D,ij}(\omega)h^*_{I,ij}(\omega)) \\ &= |h_{D,ij}(\omega)|^2 + |h_{I,ij}(\omega)|^2 + h_{IF,ij}(\omega)\end{aligned} \tag{16}$$

$\Longrightarrow$Total energy transfer = Direct energy transfer + Indirect energy transfer + Interference effect where ϕ_D and ϕ_I represents the phase of direct and indirect transfer functions respectively and $h_{IF,ij}(\omega) = 2|h_{D,ij}(\omega)||h_{I,ij}(\omega)|cos(\phi_D - \phi_I)$. The direct energy transfer serves as a tool for structure determination of the process.

The direct energy transfer serves also as a tool for the determination of plant connectivity along with the strength. As mentioned in Section 2.2 in a joint representation of the process, each signal x_j is assumed to be driven by a distinct white noise driving force e_j. The source e_j affects x_j directly. Also the effect can be through some other signal(s) (*e.g.* $e_j \rightarrow x_i \rightarrow x_j$) termed as indirect effect. In addition, e_j affects signal other than x_j through direct and indirect paths. Once the information of direct connectivity is developed, the indirect pathways can be deduced.

3.2 Connectivity Strengths

The strength of a directed connectivity is quantified in terms of the direct energy transfer. In general, connectivities can be measured and quantified using other measures as discussed extensively in Muskulus et al. [2009]. However, variance (second-order measures) is a popular choice, both from physical interpretation as well as implementation viewpoints.

To facilitate the comparison of strengths, an index $\beta_{i,j}$ representing the strength of connection from source in x_j and sink in x_i is defined as

$$\beta_{i,j} = \frac{\displaystyle\int_0^{\pi} |h_{D,ij}(\omega)|^2 \, d\omega}{\max\left(1, \; \left(\displaystyle\int_0^{\pi} |h_{D,ij}(\omega)|^2 \, d\omega\right)_{\max, i\neq j}\right)} \quad (17)$$

The index is zero if and only if the connectivity is zero. It provides a relative strength of links within the process. In order to prevent small values of strength causing an abnormal rise of the index, (*e.g.*, in a diagonal system, all the off-diagonals will have small insignificant values for the numerator term $\int_0^{\pi} |h_{D,ij}(\omega)|^2 \, d\omega$ due to estimation errors) the minimum value of the denominator term is chosen as 1.

Detection of connectivity requires a significance limit on the index. A statistical development of significance limit is beyond the scope of this work. Hence, a threshold limit for the strength of the connectivity as 0.015 is estimated through Monte Carlo simulations. A low value, say, strength < 0.015 indicates the connection is weak enough to be treated as insignificant. The strength estimated is always positive and represents the *effective strength of connectivity* under the prevailing operating conditions. The effective connectivity permits one to allow due weighting to strong connections. The calculated values stored in a matrix form is suited for automated analysis. On the other hand, a graphical representation aids in easier understanding of the process nature.

3.3 Estimation

The developed model can be considered essentially as constituting of two parts, the coefficients representing the plant model *i.e.*, $\mathbf{H}$ matrix, and the innovation covariance matrix, Σ_e denoting the noise model. There can be three different situations.
(*i*) $\mathbf{A}_0 = I$ and Σ_e diagonal. In this case the plant model and the noise model are uniquely identifiable.
(*ii*) $\mathbf{A}_0 \neq I$ and Σ_e diagonal. This situations arises due to the presence of instantaneous relationships among the variables. If the positions of instantaneous causality are known, both plant and noise models are identifiable. In essence, the identifiability is provisional.
(*iii*) $\mathbf{A}_0 \neq I$ and $\Sigma_e \neq I$. This condition arises due to correlated innovations with or without the presence of instantaneous causality. There are identifiability constraints when the effects of instantaneous causality and correlated noise coincide. The direction of the causal relationships in the noise model can be resolved only if process knowledge is available. Also, it is not possible to distinguish between the effects due to instantaneous causality and noise correlation.

The instantaneous causality is basically an integral part of the plant model. Incorporation of instantaneous causality into the plant model is achieved by converting the RVAR model (3) to SVAR (8) form. A worthwhile point with reference to control systems is that, instantaneous causality can exist exclusively due to feedback control and recycle streams. As the feedback loops (due to proportional controller, which acts as a gain element) and recycle streams are known, the corresponding knowledge on the instantaneous causalities are available.

Once the instantaneous causality information, A_0 is extracted, the revised noise covariance matrix is calculated as $\tilde{\Sigma}_e = \mathbf{A}_0 \Sigma_e \mathbf{A}_0'$ where Σ_e is the innovation covariance matrix of the RVAR model. The significance of off-diagonal terms in $\tilde{\Sigma}_e$ provides the information on correlated innovations. If the correlated innovations are present, a further scaling of the SVAR model with the inverse of the Cholesky factor of $\tilde{\Sigma}_e$ (*i.e.*, B matrix) is used to account for them. This corresponds to the AB form of SVAR model.

The approach adopted in this work assumes that, the data has sufficient excitation. Under this condition, irrespective of the noise model, the plant model and noise model are estimated correctly. This means that the predictable part is captured well by the model. Hence, once the effects of instantaneous causality are accounted for, the model can be considered as one with a unit variance white noise driving force in all channels. This implies that the diagonals of both A and B matrices in SVAR model can be assigned as unity.

4. SIMULATION RESULTS

The example below shows an open loop VAR model.

$$\begin{bmatrix} x_1[k] \\ x_2[k] \\ x_3[k] \end{bmatrix} = -\begin{bmatrix} 0.5 & 0 & 0 \\ 0.2 & 0.3 & 0 \\ 0.1 & 0.3 & 0.4 \end{bmatrix} \begin{bmatrix} x_1[k-1] \\ x_2[k-1] \\ x_3[k-1] \end{bmatrix} - \begin{bmatrix} 0.3 & 0 & 0 \\ 0 & 0.2 & 0 \\ 0.2 & 0 & 0.4 \end{bmatrix} \begin{bmatrix} x_1[k-2] \\ x_2[k-2] \\ x_3[k-2] \end{bmatrix} + \begin{bmatrix} e_1[k] \\ e_2[k] \\ e_3[k] \end{bmatrix} \quad (18)$$

The direct effects for the process are calculated from the data of the process and is represented in figure 1. The direct energy transfer plots reflect the structure of the process. The strength of the connectivities are estimated. The

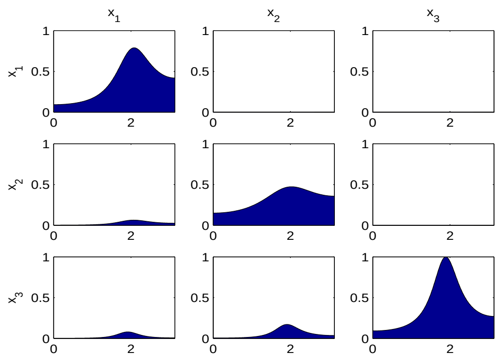

Fig. 1. Direct energy transfer among signals for example 1 x-axis: Frequency (0 to π); y-axis: Magnitude of direct energy transfer

topology reconstructed is shown along with the relative strength of connectivities (within the process) in Figure 2. The connectivities matches with the structure of the process. The values shown are normalized to provide a comparitive value within the process.

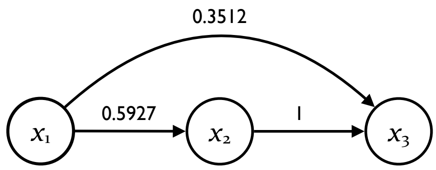

Fig. 2. Connectivity strength for example 1

Example 2

Now, the process in example 1 is modified to give a mutual lagged relationship between x_1 and x_2. This gives a VAR process with closed loop structure given by

$$\begin{bmatrix} x_1[k] \\ x_2[k] \\ x_3[k] \end{bmatrix} = -\begin{bmatrix} 0.5 & 0.3 & 0 \\ 0.2 & 0.3 & 0 \\ 0.1 & 0.3 & 0.4 \end{bmatrix} \begin{bmatrix} x_1[k-1] \\ x_2[k-1] \\ x_3[k-1] \end{bmatrix} - \begin{bmatrix} 0.3 & 0 & 0 \\ 0 & 0.2 & 0 \\ 0.2 & 0 & 0.4 \end{bmatrix} \begin{bmatrix} x_1[k-2] \\ x_2[k-2] \\ x_3[k-2] \end{bmatrix} + \begin{bmatrix} e_1[k] \\ e_2[k] \\ e_3[k] \end{bmatrix} \quad (19)$$

The relative strength of connectivities are shown in Figure 3. The results show the connection form x_2 to x_1 and the topology corresponds to the structure of the process.

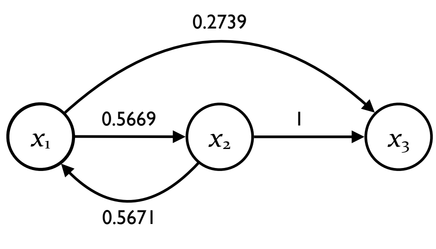

Fig. 3. Connectivity strength for example 2

Example 3

As a third example for topology reconstruction the data generated from a discrete domain multivariable non-VAR process shown in Table 1 is used. The process has three inputs u_j and four outputs y_i. The pairs (y_1, u_1) and (y_2, u_2) have feedback relationships with instantaneous causality. The nature of the transfer function in the feedback loop (proportional integral controller) introduces instantaneous causality from output to input (due to the proportional action.) The process is simulated in SIMULINK/MATLAB to create the data. The data is represented as

$$\mathbf{x} = [\, y_1 \; y_2 \; y_3 \; y_4 \; u_1 \; u_2 \; u_3 \,]^T$$

The index is calculated by (17). The relative connectivity strengths are shown in Figure 4. The connectivity values reflects the transfer function relationships.

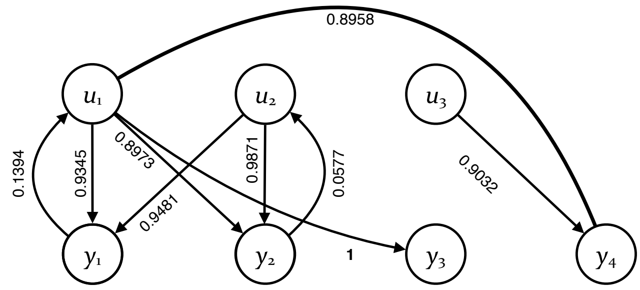

Fig. 4. Connectivity strength for example 3

This process has instantaneous causality from $y_1 \rightarrow u_1$ and $y_2 \rightarrow u_2$. If the connectivities are estimated from the RVAR model (strictly causal), without accounting for the instantaneous causality in A_0 matrix, these connections will remain undetected in the analysis. The results are not shown due to space constraints.

5. CONCLUSIONS

This work presents a topology reconstruction method based on causality analysis in frequency domain. The signal flow path is determined from a quantification of energy transfer into direct and indirect contributions among the signals. The direction and relative strength of the connectivity are derived from the magnitude of direct energy transfer. The methodology relies on spectral factorization result, while the estimation of the spectral factor uses a VAR modelling approach. The method is fairly robust to the order of the model, which is a desirable feature.

Simulation results demonstrate the promise that this method holds in reconstructing and quantifying plant connectivities from data in an automated manner. Future works include extensive testing of the proposed method to industrial processes and refinements of the quantification of connectivity strengths. A refinement of the threshold for connectivity through surrogate data analysis should provide better results. In conclusion, it may also be remarked that the methodology rests on second-order measures of connectivity. Therefore, extensions using other distance measures is also envisaged.

REFERENCES

L. Astolfi, F. Cincotti, D. Mattia, M. G. Marciani, L. Baccala, F. de Vico Fallani, S. Salinari, M. Ursino,

Table 1. Transfer function representation of process in example 3

Signal	y_1	y_2	y_3	y_4	u_1	u_2	u_3
y_1	1	0	0	0	$\frac{z^{-1}}{1-0.4*z^{-1}}$	$\frac{z^{-1}}{1-0.2*z^{-1}}$	0
y_2	0	1	0	0	$\frac{z^{-1}}{1-0.1*z^{-1}}$	$\frac{z^{-1}}{1-0.3*z^{-1}}$	0
y_3	0	0	1	0	$\frac{z^{-1}}{1-0.5*z^{-1}}$	0	0
y_4	0	0	0	1	$\frac{z^{-1}}{1-0.2*z^{-1}}$	0	$\frac{z^{-1}}{1-0.2*z^{-1}}$
u_1	$\frac{-(0.4-0.2*z^{-1})}{1-z^{-1}}$	0	0	0	1	0	0
u_2	0	$\frac{-(0.3-0.1*z^{-1})}{1-z^{-1}}$	0	0	0	1	0
u_3	0	0	0	0	0	0	1

M. Zavaglia, and F. Babiloni. Assessing cortical functional connectivity by partial directed coherence:simulations and application to real data. *IEEE transactions on Biomedical Engineering*, 53(9):1802–1812, 2006.

L. A. Baccala and K. Sameshima. Partial directed coherence: a new concept in neural structure determination. *Biological Cybernetics*, 84:463–474, 2001.

L. A. Baccala, K. Sameshima, and D.Y. Takahashi. Generalized partial directed coherence. *Proceedings of the 15th International Conference on Digital Signal Processing DSP 2007*, pages 163–166, 2007.

M. Bauer, J. W. Cox, M. H. Caveness, J. J. Downs, and N. F. Thornhill. Finding the direction of disturbance propagation in a chemical process using transfer entropy. *IEEE Transactions on Control Syatems Technology*, 15: 12–21, 2007.

K. J. Blinowska. Review of the methods of determination of directed connectivity from multichannel data. *Medical Biological Engineering and Computing*, 49:521–529, 2011.

Michael Eichler. On the evaluation of information flow in multivariate systems by the directed transfer function. *Biological Cybernetics*, 94:469–482, 2006.

M. R. Gevers and B. D. O. Anderson. Representations of jointly stationary stochastic feedback processes. *International Journal of Control*, 33(5):777–809, 1981.

J. Geweke. Measures of conditional linear dependence and feedback between multiple time series. *Journal of the American Statistical Association*, 77(378):304–313, 1982.

J. Geweke. Measurement of linear dependence and feedback between multiple time series. *Journal of the American Statistical Association*, 79(388):907–915, 1984.

S. Gigi and A. K. Tangirala. Quantitative analysis of directional strengths in multivariate processes. *Biological Cybernetics*, 103:119–133, 2010.

C. W. J. Granger. Investigating causal relationships by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438, 1969.

K. Hlavackova-Schindler, M. Palus, M. Vejmelka, and J. Bhattacharya. Causality detection based on information-theoretic approaches in time-series analysis. *Physics Reports*, 441:1–46, 2007.

H. Jiang, R. Patwardhan, and S. L. Shah. Root cause diagnosis of plant-wide oscillations using the concept of adjacency matrix. *Journal of Process Control*, 19:1347–1354, 2009.

M. J. Kaminski and K. J. Blinowska. A new method of the description of the information flow in the brain structures. *Biological Cybernetics*, 65:203–210, 1991.

H. Lutkepohl. *New Introduction to Multiple Time Series Analysis*. Springer, 2005.

M. R. Maurya, R. Rengaswamy, and V. Venkatasubramanian. A systematic framework for the development and analysis of signed digraphs for chemical processes. 1. algorithms and analysis. *Industrial & Engineering Chemistry Research*, 42:4789–4810, 2003a.

M. R. Maurya, R. Rengaswamy, and V. Venkatasubramanian. A systematic framework for the development and analysis of signed digraphs for chemical processes. 2. control loops and flow sheet analysis. *Industrial & Engineering Chemistry Research*, 42:4811–4827, 2003b.

M. Muskulus, S. Houweling, S. Verduyn-Lunel, and A. Daffertshofer. Functional similarities and distance properties. *Journal of Neuroscience Methods*, 183:31–41, 2009.

B. Pfaffe and K. im. Taunus. VAR, SVAR and SVEC models:implementation within R package vars. *Journal of Statistical Software*, 27(4):1–32, 2008.

M. B. Priestley. *Spectral Analysis and Time Series*. Academic Press, 1981.

Y. Saito and H. Harashima. Tracking of information within multichannel EEG record-causal analysis in EEG. in: N. Yamaguchi and K. Fujisawa (Eds), recent advances in EEG and MEG data processing. 1981.

B. Schelter, J. Timmer, and M. Eichler. Assessing the strength of directed influences among neural signals using renormalized partial directed coherence. *Journal of Neuroscience Methods*, 179(1):121–130, April 2009.

R. H. Shumway and D. S. Stoffer. *Time Series Analysis and Its Applications*. Springer, 2000.

S. Y. Sim, H. G. Ananthakumar, L. Benabbas, A. Horch, R. Drath, and N. F. Thornhill. Using process topology in plant-wide control loop performance assessment. *Computers and Chemical Engineering*, 31:86–99, 2006.

S. W. Smith. *Scientist and Engineer's Guide to Digital Signal Processing*. California Technical publishing, 1997.

S.M. Smith, K.L. Miller, C. Salimi-Khorshidi, M. Webster, C.F. Beckmann, T.E. Nichold, J.D. Ramsey, and M.W. Woolrich. Network modelling methods for FMRI. *NeuroImage*, 54:875–891, 2011.

R. S. Tsay. *Analysis of financial time series*. John Wiley & sons,Inc, 2010.

N. Wiener. The theory of prediction. In E. Beckenbach, editor, *Modern Mathematics for Engineers*, 1, pages 125–139. McGraw Hill, 1956.