Directed networks with underlying time structures from multivariate time series

Toshihiro Tanizawa, 1 Tomomichi Nakamura, 2 and Fumihiko Taya 3

1 Kochi National College of Technology, Monobe-Otsu 200-1, Nankoku, Kochi 783-8508, Japan
2 Graduate School of Simulation Studies, University of Hyogo, 7-1-28 Minatojima-minamimachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan
3 Singapore Institute for Neurotechnology, Centre for Life Sciences, National University of Singapore, 28 Medical Drive, #05-COR, Singapore 117456

(Dated: June 23, 2014)

Abstract

In this paper we propose a method of constructing directed networks of time-dependent phenomena from multivariate time series. As the construction method is based on the linear model, the network fully reflects dynamical features of the system such as time structures of periodicities. Furthermore, this method can construct networks even if these time series show no similarity: situations in which common methods fail. We explicitly introduce a case where common methods do not work. This fact indicates the importance of constructing networks based on dynamical perspective, when we consider time-dependent phenomena. We apply the method to multichannel electroencephalography (EEG) data and the result reveals underlying interdependency among the components in the brain system.

PACS numbers: 05.45.Tp, 89.20.Ff, 89.75.Fb, 89.75.Hc
I. INTRODUCTION

In the real world, many phenomena are time-dependent and time series usually show irregular fluctuations. When we investigate time series, we assume that (i) the systems generating the time series are not isolated from their surroundings and (ii) the periodicities in the time series contain important information. There are many cases in which several systems are interconnected or interrelated in some ways to varying degrees and the time series generated by these systems may have time structure due to various periodic (time delay or feedback) effects. Hence, we have to deal with multivariate time series to identify relevant components in these systems, taking into account the time structure. To understand the details of the interaction between components, network visualization with linking structure containing directional information is useful. In this paper, we introduce a method for constructing directed networks based on linear models from multivariate time series.

There are several methods to construct a network from multivariate time series \([1, 5]\), where all of these are based on the concept of linear correlation (correlation coefficient or cross correlation). The linear correlation is a statistic to estimate similarities between two data. The basic strategy of existing methods for constructing a network from multivariate time series is as follows. Each time series is considered as a basic node of a network. To investigate the relationship among multivariate time series, the linear correlation between each pair of time series taken from the time series is estimated. A pair is considered to be connected with an undirected link when the value of the linear correlation is larger than an appropriately chosen threshold. The basic idea underlying this method is that two time series are related to each other if they behave similarly. When time series are similar and take large values of linear correlation, it seems logical to expect some sort of similar activities between them or relationships between the activities. On the other hand, when time series are not similar, we may have the impression that these systems are independent or have no relationship.

However, it should be emphasized that “no similarity” is not equivalent to “no relationship.” That is to say, even when there is no similarity between two, it is still possible that they have some kind of relationship (that is, the two systems might be interconnected or interrelated). We will show an example in the next section, where linear correlation cannot identify interrelated components in a system.
Also, to treat a time-dependent phenomenon connected to their surroundings without detracting the nature as much as possible, we should regard it generated by a system composed of a certain number of appropriately chosen distinct elements. While the linear correlation is effective in measuring the correlation between two data, it cannot treat correlations among many concurrently. Even if correlations between two data are collected, the aggregation is not always identical to the correlations among many (that is, the patchwork of many two-body correlations might not be the same as the many-body correlation as a whole). To understand phenomena it is necessary and important to know the entire correlation among many components more accurately.

In addition, there are many cases in which the relationship is directional. In other words, even if system $\alpha$ has an effect on system $\beta$, system $\beta$ might have no effect on system $\alpha$. This implies that the edges in the networks constructed from multivariate time series should have directions. The networks constructed using the values of linear correlation are undirected. Therefore, even though the common methods relied on the concept of linear correlation have been proved to be effective \cite{1, 5}, it is very restrictive because of the above mentioned insufficiencies.

Nakamura and Tanizawa proposed a method of constructing a weighted directed network based on a sophisticated linear model from a single (univariate) time series \cite{6}. The networks constructed by this method incorporate the entire time structure of the time series. The links on the network are assigned weights based on a transformation of the coefficients of the linear model. Walker et al. applied this method to the system containing an enormous number of variables for identifying critical components of a grain material controlling the collapse due to a strong shear strain \cite{7}. Although multivariate time series modeling is used to construct directed networks in their work, time structure in time series is not reflected in their networks. The purpose of this paper is to construct directed networks reflecting time structures of periodicities contained in time-dependent phenomena from multivariate time series.

This paper is organized as follows. In the next section we discuss the shortcomings of the common methods in which the relationship between time series is defined by the values of linear correlation. In Section III we explain our method to construct directed networks based on linear models from multivariate time series. In Section IV we apply the method to electroencephalography (EEG) time series. We summarize this paper in Section V.
II. NETWORK CONSTRUCTION METHOD USING LINEAR CORRELATION

We show an example where the common methods for constructing undirected networks based on linear correlation (correlation coefficient or cross correlation) fails to capture explicit relationship between relevant components in a dynamical phenomena.

In these methods, a pair of nodes is connected with an undirected link when the value of linear correlation is larger than an appropriately chosen threshold. A simple example reveals a fundamental defect of this idea. The example consists of four dynamical variables, $x_1(t)$, $x_2(t)$, $x_3(t)$ and $x_4(t)$, and the time dependency is described by the following expressions:

\[
x_1(t) = 0.7x_1(t-1) - 0.4x_1(t-3) + 0.3x_2(t-4) + 0.2x_4(t-7) + \varepsilon_1(t),
\]
\[
x_2(t) = 3.0 + 0.6x_2(t-1) - 0.2x_2(t-6) + \varepsilon_2(t),
\]
\[
x_3(t) = 0.5x_1(t-2) + 0.3x_4(t-9) + \varepsilon_3(t),
\]
\[
x_4(t) = 0.2x_1(t-2) + 0.5x_4(t-1) - 0.3x_4(t-3) + \varepsilon_4(t),
\]

where $\varepsilon_i(t)$ ($i = 1, 2, 3, 4$) are independent and identically distributed Gaussian random variables with standard deviation 1.0. The terms in Eq. (1) are $x_1(t-1)$, $x_1(t-3)$, $x_2(t-4)$ and $x_4(t-7)$, those in Eq. (2) are $x_2(t-1)$ and $x_2(t-6)$, those in Eq. (3) are $x_1(t-2)$ and $x_4(t-9)$, and those in Eq. (4) are $x_1(t-2)$, $x_4(t-1)$ and $x_4(t-3)$. We distinguish “component” and “variable” as different terms in this paper. The term “component” is used for representing $x_i$, and the term “variable” for representing $x_i(t-l)$ with a time delay effect. That is, Eq. (1) has 3 components ($x_1$, $x_2$ and $x_4$) and 4 variables ($x_1(t-1)$, $x_1(t-3)$, $x_2(t-4)$ and $x_4(t-7)$). Equation (2) shows that the model for $x_2(t)$ is composed of only one component, $x_2$. In this regard, the dynamics of the component $x_2$ is independent of others. The dynamics of three other variables, $x_1(t)$, $x_3(t)$, and $x_4(t)$, are dependent on other components, as seen in Eqs. (1), (3), and (4).

Figure 1 show four time series generated by these models. At the first glance we cannot find clear similarity among them, although three of four variables are dependent on others. Table II shows the values of the cross correlation (CC) of all pairs. If we set the threshold for CC as 0.5 for example, the nodes, $x_1(t)$ and $x_3(t)$, are connected. This connection seems to be reasonable, since Eq. (3) representing the dynamics of $x_3(t)$ includes $x_1(t-2)$. However, Eq. (1) representing the dynamics of $x_1(t)$ does not include the component $x_3$. The undirectedness of the connection cannot thus capture this one-way relationship between $x_1$
and $x_3$. Furthermore, we would conclude that $x_1$ and $x_4$ are independent, since the value of CC between these components, 0.4452 is below the threshold 0.5. However, it is clearly untrue, because Eq. (4) representing the dynamics of $x_4(t)$ includes the variable $x_1(t - 2)$. It should be noted that one of the major defects of using CC to construct networks cannot deal with asymmetric interactions, as this simple example shows. The other defect is the arbitrariness of the threshold. If we set a smaller value, say 0.4, as the threshold, the components $x_1$, $x_3$, and $x_4$ are connected, and the resulting network is completely different from that corresponding to the threshold 0.5. If the dynamics of the component of $x_i$ is dependent on the component $x_j$, it seems much simpler and more straightforward to connect these two components with a directed link from $x_j$ to $x_i$. This is the main idea of the new method described in the next section.

III. DIRECTED NETWORK CONSTRUCTION FROM LINEAR MODELS

In the method proposed in this paper, directed networks are constructed based on linear models from multivariate time series. After we review the method of building linear models we use in this paper, we introduce an idea to construct directed networks based on linear models.
TABLE I. The largest absolute values of the cross correlation of all possible pairs. The data are generated by Eqs. (1) – (4) and the values are estimated using 1000 data point.

|     | $x_1$ | $x_2$ | $x_3$ | $x_4$ |
|-----|------|------|------|------|
| $x_1$ | 1.0000 | — | — | — |
| $x_2$ | 0.3608 | 1.0000 | — | — |
| $x_3$ | 0.6692 | 0.1931 | 1.0000 | — |
| $x_4$ | 0.4452 | 0.1790 | 0.4900 | 1.0000 |

A. Building linear models from multivariate time series

Periodic or nearly periodic behavior is an important feature for many time-dependent phenomena in the real world. There are several traditional or widely used techniques to estimate the period of a behavior, for example, spectral estimation [8], auto-correlation [8], wavelet transforms [9] and so on. All of these standard techniques either employ, are related to, or are a generalization of, Fourier series. Although the importance of taking nonlinearity into consideration is widely known when we treat time series, linear analyses remain attractive and widely applied [10]. Small and Judd have proposed a sophisticated method to detect periodicities more reliably from time series [11]. The technique is based on an information theoretic reduction of linear auto-regressive (AR) models so that only the essential features of an AR model are retained. These models are called reduced auto-regressive (RAR) models [12]. The essential features of RAR models include any periodicity present in the data. There are strong information theoretic arguments to support the statement that RAR model can detect any periodicity if they are present [11].

The formula of RAR model is as follows. Given a scalar time series $\{x(t)\}_{t=1}^{n}$ of $n$ observations, an RAR model with the largest time delay $w$ can be expressed by

$$x(t) = a_0 + a_1 x(t - l_1) + a_2 x(t - l_2) + \cdots + a_w x(t - l_w) + \varepsilon(t)$$

$$= a_0 + \sum_{i=1}^{w} a_i x(t - l_i) + \varepsilon(t),$$

(5)

where $1 \leq l_1 < l_2 \cdots < l_w$, $a_i \ (i = 0, 1, 2, \ldots, w)$ are unknown parameters to be determined, and $\varepsilon(t)$ is assumed to be independent and identically distributed Gaussian random variables, which are interpreted as a fitting error. The parameters $a_i$ are chosen to minimize the sum
of squares of the fitting error. We can easily apply the methodology to multivariate time series. An multivariate RAR model is expressed by

\[ x_i(t) = a_{i,0} + \sum_{j=1}^{N} \left( \sum_{k=1}^{w_i} a_{i,j,k} x_i(t-l_k) \right) + \varepsilon_i(t), \]  

where \( N \) is the number of components, \( w_i (\geq 0) \) is the largest time delay of the \( i \)-th component. The multivariate RAR model tells us periodicities reflecting the relationship between multivariate time series.

The RAR modeling approach we follow here also attempts to determine a smaller subspace in which to describe the data; however, the incorporation of the information theory description length introduces a more objective framework to decide what constitutes the “best” subspace. The description length of a model of observed data consists of a term capturing how well the model fits the data, and terms which penalize for the number of parameters needed for the model to achieve a prescribed accuracy. Specifically, an approximation to description length takes the form

\[ DL(k) = \left( \frac{n}{2} - 1 \right) \ln \frac{\mathbf{e}^T \mathbf{e}}{n} + (k + 1) \left( \frac{1}{2} + \ln \gamma \right) - \sum_{i=1}^{k} \ln \delta_i \]  

where \( n \) is the length of the time series to be fitted, \( \mathbf{e} \) are the fitting errors, \( k \) is the number of parameters (or model size), \( \gamma \) is related to the scale of the data, and the variables \( \delta \) can be interpreted as the relative precision to which the parameters are specified. The factor \( \gamma \) is a constant and typically fixed to be \( \gamma = 32 \) \[12\]. See more details in \[12, 13\].

We note that selecting the optimal subset from a large dictionary of basis functions is typically an NP-hard problem which usually has to be solved by heuristic methods \[12\]. The models obtained by selection algorithms can offer near-optimal models. In this paper, we use a selection method using the total error \[14\]. The algorithm has proven to be effective in modeling dynamics \[15\].

We apply the selection method to the data represented in Fig. \[1\]. In this case, we have four time series (that is, \( x_1(t), x_2(t), x_3(t) \) and \( x_4(t) \)) and 1000 data points, and we build four multivariate RAR models. Choosing a time delay up to 10 for time series of each component and the constant function give 41 candidate basis functions in the dictionary \[16\]. Using the dictionary we build the multivariate RAR model for each channel (that is, we build four multivariate RAR models). The variables included in the best models are the same as Eqs. \((1)-(4)\), which also verifies the consistency of multivariate RAR modeling.
Our procedure is to find the best subset of time series which best predicts a given components history. The best model is taken to be that model which minimizes the description length. We do this for time series of all components in the assembly. When we apply the selection method to multivariate time series only the correct terms are selected in each best model. That is, these models have only the same terms as Eqs. (1)–(4).

B. Directed network construction

After building the multivariate RAR model corresponding to the systems under consideration, we use the information of these best models to construct a directed network representing the system. Suppose that the \( i \)-th variable \( x_i(t) \) is expressed as

\[
x_i(t) = a_{i,0} + a_{i,i,1}x_i(t - l_1) + a_{i,i,2}x_i(t - l_2) + a_{i,j,3}x_j(t - l_3) + a_{i,k,4}x_k(t - l_4) + \varepsilon_i(t). \tag{8}
\]

To determine the value of \( x_i \) at time \( t \), we need the information of the values of \( x_i, x_j, \) and \( x_k \) at some previous times. We pack this information in the notation

\[
x_i = f_i(x_i, x_j, x_k), \tag{9}
\]

representing that \( x_i \) is a function of \( x_i, x_j \) and \( x_k \). When we translate the relationship to the network, we treat each component of \( x_i \) as a node; a node of the network represents each component of the multivariate time series. Next, we draw directed arrows from \( x_j \) to \( x_i \) and from \( x_k \) to \( x_i \), if \( x_j \) and \( x_k \) are selected in the model of \( x_i \). The constant term \( a_{i,0} \) in Eq. (8) is not used, because the term have no time structure. This basic idea enables us to construct a directed network embodying the essential relationship between components represented in a multivariate RAR model.

Let us apply the idea to the previously mentioned example, Eqs. (1)–(4). Eq. (1) shows that \( x_1(t) \) is expressed by a linear combination of \( x_1(t - 1), x_1(t - 3), x_2(t - 4) \) and \( x_4(t - 7) \). Hence, we summarize that the model is composed of three components \( x_1, x_2 \) and \( x_4 \). In the same fashion, we can summarize Eqs. (2), (3) and (4). As a result of this operation, we
obtain the following expression for Eqs. (1) to (4),

\[ x_1 = f_1(x_1, x_2, x_4), \]  
\[ x_2 = f_2(x_2), \]  
\[ x_3 = f_3(x_1, x_4), \]  
\[ x_4 = f_4(x_1, x_4). \]  

(10) \quad (11) \quad (12) \quad (13)

Even these reduced expressions give us interesting information. Eqs. (10) and (11) tell that \( x_2 \) is autonomous and has influence on the dynamics of \( x_1 \), although \( x_1 \) does not influence the dynamics of \( x_2 \). Eq. (12) tells that \( x_3 \) completely dependent on \( x_1 \) and \( x_4 \), and cannot drive itself at all. From Eqs. (10) and (13), we can understand that \( x_1 \) and \( x_4 \) are interconnected because both the components are mutually included in the models.

Using this summarized information we construct a directed network. Since the information relating to the parameters of linear models is erased in the reduced expressions, such as Eqs. (10)–(13), the links in the constructed network are unweighted. Figure 2 shows the relationship between interdependency of all components. A directed arrow from node \( x_j \) to node \( x_i \) represents that the component \( x_j \) is included in the model of \( x_i \). If the component \( x_i \) itself is included in the model of \( x_i \), the node \( x_i \) has a directed self-loop from \( x_i \) to \( x_i \). Since these self-loops make the network representation of the linear models rather cumbersome, we represent the nodes with self-loops as circles and the node without self-loop as a square in Fig. 2.

From Fig. 2 we can gain valuable information for the relationship between the dynamics of components. For example, although the component \( x_1 \) includes \( x_2 \), the component \( x_2 \) does not include \( x_1 \). The arrow between \( x_1 \) and \( x_2 \) has, therefore, a single direction from \( x_2 \) to \( x_1 \). However, as the components \( x_1 \) and \( x_4 \) are mutually included in each expression of the model, the arrow is bi-directed between \( x_1 \) and \( x_4 \). The directed network constructed by this idea reflects full dynamical features of the multivariate time series such as time structures of periodicities. It should be noted that we cannot obtain this relationship from the network construction method using the linear correlation (correlation coefficient or cross correlation).

To elucidate the feature of a directed network constructed by this method, the numbers of in-degree and out-degree of each node are useful. When the number of in-degree of a node is large, the node is influenced by many other nodes. When the number of out-degree of a node is large, the node has a large influence on many other nodes. Let us examine the
FIG. 2. (Color online) The directed network of linear models of Eqs. (1)–(4). Notation □ means that the model for a component includes the component itself, and notation ◯ means that the component is not included in the model.

TABLE II. The number of in-degree and out-degree of the directed network shown in Fig. 2.

|       | $x_1$ | $x_2$ | $x_3$ | $x_4$ |
|-------|-------|-------|-------|-------|
| in-degree | 3     | 1     | 2     | 2     |
| out-degree | 3     | 2     | 0     | 3     |

numbers of in-degree and out-degree of the directed network shown in Fig. 2. The number of in-degree is identical to the number of components included in each model. For example, Eq. (10) indicates that there are three components. Hence, the number of in-degree for node $x_1$ is 3. The number of in-degree can also be obtained from a different approach. As node $x_1$ has two incoming arrows and the form of the node is ◯ as shown in Fig. 2 we can know that the number of in-degree is 3. Also, as node $x_1$ has two outgoing arrows, the number of out-degree is 3. In the same approach we can investigate the number of in-degree and out-degree for nodes $x_2$, $x_3$ and $x_4$. Table II shows the numbers of in-degree and out-degree of all nodes.

As mentioned in Section III A, it should be emphasized that when we apply the selection method, only the same variables in Eqs. (1)–(4) are selected in the best model. This indicates that the identical directed network shown in Fig. 2 is always reproduced using the multivariate models with the selection method.
IV. APPLICATION

We apply the proposed method to multichannel electroencephalography (EEG) time series. The EEG data are measured simultaneously from a healthy human under the stable condition (eyes closed and resting). The measurement points follow the unipolar 10-20 Jasper registration scheme [17]. See Fig. 3 for the placement of the electrodes. Here, the capital letters, F, C, P, and O, stand for the frontal, central, parietal, and occipital lobes in the brain, respectively. We note that there is no “central lobe” and these electrodes partially cover the frontal and parietal lobes. The suffix z stands for zero and means that the corresponding electrodes are placed on the midline of the skull. The odd or even number following the capital characters means that the electrodes are placed on the left side or the right side, respectively. The frontal lobe is involved in highly organized functions of the human body, such as planning, reasoning, judgement, and impulse control, and the most uniquely human of all the brain structures. The parietal lobe integrates sensory information from various parts of the human body. The occipital lobe contains several functional visual areas and is related to the sense of sight.

The data were digitized at 1024 Hz using a 12-bit A/D converter. The EEG impedances were less than 5 KΩ. The data were amplified, gain = 18 000, and amplifier frequency cut-off settings of 0.03 Hz and 200 Hz were used [18]. We use time series obtained from 10 channels of Fz, Cz, Pz, Oz, F3, F4, C3, C4, P3 and P4. We use 1000 data points (around 1 second) to build multivariate RAR models. Figure 3 shows the placement of the channels and Fig. 4 shows the time series of each channel.

As there are 10 channels, choosing a time delay up to 25 for time series of each channel and the constant function give 251 candidate basis functions in the dictionary. Using the dictionary we build the multivariate RAR model for each channel. The reduced expressions
FIG. 3. (Color online) The International 10-20 system of electrode placement. We use data from 10 channels, Fz, Cz, Pz, Oz, F3, F4, C3, C4, P3 and P4.

FIG. 4. Multichannel electroencephalography (EEG) time series of Fz, Cz, Pz, Oz, F3, F4, C3, C4, P3 and P4 used for building multivariate RAR models.
of the obtained 10 multivariate RAR models are

\[ F_z = f_{F_z}(F_z, C_z), \quad (14) \]
\[ C_z = f_{C_z}(C_z, P_z, O_z, F_3), \quad (15) \]
\[ P_z = f_{P_z}(F_z, C_z, P_z, O_z, C_3, P_3, P_4), \quad (16) \]
\[ O_z = f_{O_z}(F_z, C_z, O_z), \quad (17) \]
\[ F_3 = f_{F_3}(F_3), \quad (18) \]
\[ F_4 = f_{F_4}(F_z, F_4), \quad (19) \]
\[ C_3 = f_{C_3}(F_3, C_3, P_3), \quad (20) \]
\[ C_4 = f_{C_4}(F_z, C_4, P_4), \quad (21) \]
\[ P_3 = f_{P_3}(P_z, O_z, P_3), \quad (22) \]
\[ P_4 = f_{P_4}(F_z, C_z, P_z, P_4). \quad (23) \]

In Fig. 5, we show the directed network constructed from these models representing the relationship of interdependency among these 10 channels. Note that all models contain their own components, which means that all nodes in Fig. 5 have self-loops.

Let us examine the information we can extract from the network structure. The direction of an arrow from node 1 to node 2 in a network represents that the channel corresponding to node 1 (channel 1) appears in the RAR model that describes the time series of the channel corresponding to node 2 (channel 2). In other words, the dynamics of channel 2 is partially controlled by channel 1. In Fig. 5 the channels, Cz, Pz, P3, P4, and Oz, are mutually connected by bi-directional arrows. Recall that the electrode Cz is attached to the frontal lobe, that the electrodes Pz, P3, and P4 are attached to the parietal lobe, and that the electrode OZ is attached to the occipital lobe. The frontal lobe relates to the highly organized human functions, such as planning, reasoning, judgement, and impulse control. The parietal lobe is integrating sensory information from various parts of the human body. The occipital lobe contains several functional visual areas and is related to the sense of sight. The bi-directional arrows indicate that the pair of Cz and Pz and the pair of Cz and OZ are mutually but separately interacting with each other and that the the triplet, Pz, P3, and P4, which belong to the same parietal lobe, is mutually connected indeed, but the interaction between P3 and P4, which belong to different hemispheres, is indirect mediated by the midline part of Pz. It is suggestive that the electrode OZ corresponding to
FIG. 5. (Color online) The directed network constructed by multivariate RAR models from Eq. (14) to Eq. (23). All nodes are represented by ○, as all models contain their own components. For the explanation of the notation, see Fig. 2.

TABLE III. The numbers of in-degree and out-degree for each node of the directed network shown in Fig. 5.

|       | Fz | Cz | Pz | Oz | F3 | F4 | C3 | C4 | P3 | P4 |
|-------|----|----|----|----|----|----|----|----|----|----|
| in-degree | 2  | 4  | 7  | 3  | 1  | 2  | 3  | 3  | 3  | 4  |
| out-degree  | 6  | 5  | 4  | 4  | 3  | 1  | 2  | 1  | 3  | 3  |

The numbers of in-degree and out-degree for each node in Fig. 5 are shown in Table III. The occipital lobe governing the sense of sight controls part Pz corresponding to the parietal lobe integrating sensory information but not vice versa. The directed arrows from part Fz to all the other midline parts Pz and Oz and to all the right hemisphere parts F4, C4, and P4 are also remarkable. Since Fz is related to the frontal lobe of the brain that governs totally coordinated human behavior and the right hemisphere of the brain is related to emotion and image handling, these directed arrows might indicate the hierarchical relationship between the different functions of the human brain.

The numbers of in-degree and out-degree for each node in Fig. 5 are shown in Table III. In Fig. 6 we show the directed networks, the same as that showed in Fig. 5 with each
FIG. 6. (Color online) Directed networks represented in Fig. 5 with the adjusted node size: (a) the node size is adjusted by the number of in-degree and (b) by the number of out-degree. Each node size corresponds to the number of in-degree and out-degree. An outgoing arrow from a node implies that the node partially controls the dynamics of the target node and an incoming arrow into a node implies that the dynamics of the node is partially governed by the source node of the arrow.

node size is proportional to the numbers of incoming arrows (Fig. 6(a)) and outgoing arrows (Fig. 6(b)). In Fig. 6(a), the size of the node Pz is the largest. Since an incoming arrow into a node implies that the dynamics of the node is partially governed by the source node of the arrow, the dynamics of the node Pz that belongs to the parietal lobe is determined under the influences of many other parts of the brain. On the other hand, the size of node Fz is the largest in Fig. 6(b). Since an outgoing arrow from a node implies that the node partially controls the dynamics of the target node, it implies that the node Fz that belongs to the frontal lobe is the most influential on the dynamics of the other nodes. It is striking that we can extract such amount of information about the brain functioning from the combination of the RAR modeling and the network visualization scheme described in this paper without any physiological knowledge.
V. DISCUSSION AND SUMMARY

The rich amount of information we can extract from the analysis of EEG data described in the previous section suggests the effectiveness of the new method of multivariate time series data analysis described in this paper. The strong point of this method is that it can be applicable to any kind of multivariate time series without any knowledge of underlying relationship between the components of the dynamical system under consideration. The relationship can be extracted after the construction of the network. One direction of the future investigation is surely to apply this method to many complex dynamical systems and investigate its effectiveness.

The other direction is to include the information of parameters and time delay expressed in the RAR model. There is at least one attempt to include this information for univariate time series. The extension of this theory to multivariate time series is an open question, but it is surely worthwhile.

In summary, we have proposed a method of constructing directed networks for time-dependent phenomena from multivariate time series. As the network is based on the linear model, the network reflects the dynamical features such as the time structures of periodicities and the mutual relationships between components of the dynamical system generating the time series. Furthermore, this method can construct the network even if any pair of the time series do not have a sufficient value of linear correlation; the situations for which the common methods using the values of linear correlation cannot handle. We have introduced a case where the common method does not work and the importance of constructing networks based on the dynamical system perspective. We have applied the proposed method to multichannel electroencephalography (EEG) data and have extracted a rich amount of physiological information of the functions and the relationships between various parts of the human brain from the constructed network.

VI. ACKNOWLEDGMENTS

Toshi Tanizawa would like to acknowledge the support of a Grant-in-Aid for Scientific Research (C) (No. 24540419) from the Japan Society for the Promotion of Science (JSPS). Tomo Nakamura also would like to acknowledge the partial support of a Grant-in-Aid for
Tomo Nakamura and the other authors would like to acknowledge Professor Paul E. Rapp (Uniformed Services University of the Health Sciences) for providing us with the EEG data.

[1] R. N. Mantegna, The European Physical Journal B, 11, 193 (1999).
[2] I. Farkas, H. Jeong, T. Vicsek, A.-L. Barabási, and Z. N. Oltvai, Physica A, 318, 601 (2003).
[3] K. Yamasaki, A. Gozolchiani, and S. Havlin, Physical Review Letters, 100, 228501 (2008).
[4] A. A. Tsonis and K. L. Swanson, Physical Review Letters, 100, 228502 (2008).
[5] C. K. Tse and F. C. M. L. Jing Liua, Journal of Empirical Finance, 17, 659 (2010).
[6] T. Nakamura and T. Tanizawa, Physica A, 391, 4704 (2012).
[7] D. M. Walker, A. Tordesillas, T. Nakamura, and T. Tanizawa, Physical Review E, 87, 032203 (2013).
[8] M. B. Priestly, Non-linear and Non-stationary Time Series Analysis (Academic Press, London, 1988).
[9] C. S. Burrus, R. A. Gopinath, and H. Guo, Introduction to Wavelets and Wavelet Transforms: A Primer (Prentice Hall, Upper Saddle River, N.J., 1998).
[10] L. Astolfi, F. Cincotti, D. Mattia, M. G. Marciani, L. A. Baccala, F. de Vico Fallani, S. Salinari, M. Ursino, M. Zavaglia, L. Ding, J. C. Edgar, G. A. Miller, B. He, and F. Babiloni, Human brain mapping, 28, 143 (2007).
[11] M. Small and K. Judd, Physical Review E, 59, 1379 (1999).
[12] K. Judd and A. Mees, Physica D, 82, 426 (1995).
[13] K. Judd and A. Mees, Physica D, 120, 273 (1998).
[14] We first apply the bottom-up method using the total error [15]. The best model might be trapped in one of the local minima. To avoid this situation and to reduce the problem to a manageable size, a model is built from the complete dictionary using the top-down method, but starting from the model whose size is 10 larger than that of the best model obtained by the bottom-up method. We select the model as the best model whose description length is the smallest. It should be noted that this method do no worse than the bottom-up method.
[15] T. Nakamura, D. Kilminster, K. Judd, and A. I. Mees, International Journal of Bifurcation and Chaos, 14, 1129 (2004).
[16] These are the constant function and $x_1(t-1)$, $x_1(t-2)$, ... , $x_1(t-10)$, $x_2(t-1)$, ... , $x_2(t-10)$, $x_3(t-1)$, ... , $x_3(t-10)$, $x_4(t-1)$, ... , $x_4(t-10)$.

[17] H. H. Jasper, Electroencephalography and Clinical Neurophysiology, 10, 371 (1958).

[18] T. A. A. Watanabe, C. J. Cellucci, E. Kohegyi, T. R. Bashore, R. C. Josiassen, N. N. Greenbaum, and P. E. Rapp, Psychophysiology, 40, 77 (2003).