Detection of Direct Causality Based on Process Data

Ping Duan, Fan Yang, Tongwen Chen, and Sirish L. Shah

Abstract—Direct causality detection is an important and challenging problem in root cause and hazard propagation analysis. Several methods provide effective solutions to this problem for linear relationships. For nonlinear situations, currently only causality analysis can be conducted, but the direct causality cannot be identified based on process data. In this paper, we describe a direct causality detection approach suitable for both linear and nonlinear connections. Based on an extension of the transfer entropy approach, a direct transfer entropy (DTE) concept is proposed to detect whether there is a direct information and/or material flow pathway from one variable to another. A discrete DTE and a differential DTE are defined for discrete and continuous random variables, respectively; and the relationship between them is discussed. The effectiveness of the proposed method is illustrated by two examples and an experimental case study.

I. INTRODUCTION

Detection and diagnosis of plant-wide abnormalities and disturbances are major problems in the process industry. Because of the high degree of interconnections among different parts in a large scale complex system, a simple fault may propagate along information and material flow pathways and affect other parts of the system. To determine the root cause(s) of certain abnormality, it is important to capture the process connectivity and find the connecting pathways.

A qualitative process model in the form of a digraph has been widely used in root cause and hazard propagation analysis [1]. Digraph-based models usually define the fault propagation pathways by incorporating expert knowledge of the process. A drawback is that extracting expert knowledge is very time consuming and the knowledge is not always available [2]. The modeling of digraphs can also be based on mathematical equations [3], yet for large scale complex processes it is difficult to establish practical and precise mathematical models.

Data driven methods provide another way to find the causal relationships. A few data-based methods are capable of detecting the causal relationships for linear processes. In the frequency domain, directed transfer functions (DTF) [4] and partial directed coherence (PDC) [5] are widely used in brain connectivity analysis. Other methods such as Granger causality [6], predictability improvement [7], path analysis [8], and cross-correlation analysis by looking for time delays [9] are also commonly used. Information theory provides a wide variety of approaches for measuring causal influence among multivariate time series [10]. Based on transition probabilities containing all information on causality between two variables, transfer entropy is proposed to distinguish driving and responding elements [11] and is suitable for both linear and nonlinear relationships; it has been successfully used in chemical processes [12] and neurosciences [13].

Since information flow means how variation propagates from one variable to another, it is valuable to detect whether the causal influence between a pair of signals is along a direct pathway without any intermediate variables or indirect pathways through some intermediate variables. For linear relationships, in the frequency domain, a DTF/PDC-based method for quantification of direct and indirect energy flow in a multivariate process was recently proposed [14]; in the time domain, a path analysis method was used to calculate the direct effect coefficients [15].

For both linear and nonlinear relationships, the partial transfer entropy (PTE) was proposed to quantify the total amount of indirect coupling mediated by the environment and was successfully used in neurosciences [16]. On one hand, in the definition of PTE, all the environmental variables are considered as intermediate variables, which is not true in most cases and will greatly increase the computational burden. On the other hand, the utility of the PTE is to detect unidirectional causalities, which is suitable for neurosciences; however, in industrial processes, feedback and bidirectional causalities are very common. Thus, the PTE method cannot be directly used for direct causality detection in the process industry. This paper proposes a data-based direct causality detection method via the transfer entropy approach, which can be used for both linear and nonlinear multivariate relationships among process variables.

II. DETECTION OF DIRECT CAUSALITY

In this section, an extension of the transfer entropy—direct transfer entropy (DTE)—is proposed to detect the direct causality between two variables.

A. Direct Transfer Entropy

In order to determine the information and material flow pathways to construct a precise topology of a process, it is important to determine whether the influence between a pair of process variables is along direct or indirect pathways. A direct pathway means direct influence without any intermediate or confounding variables.
The transfer entropy measures the amount of information transferred from one variable \( x \) to another variable \( y \). This extracted transfer information represents the total causal influence from \( x \) to \( y \). It is difficult to distinguish whether this influence is along a direct pathway or indirect pathways through some intermediate variables. In order to detect the direct and indirect pathways of the information transfer, the definition of a DTE is introduced as follows.

Given three discrete random variables \( x, y, \) and \( z \), let them be sampled at time instants \( i \) and denoted by \( x_i, y_i, \) and \( z_i \) with \( i = 1, 2, \ldots, N \), where \( N \) is the number of samples. The causal relationships between each pair of them can be estimated by calculating transfer entropies [11].

Let \( y_{i+h_1} \) denote the value of \( y \) at time instant \( i + h_1 \), that is, \( h_1 \) steps in the future from \( i \), and \( h_1 \) is referred to as the prediction horizon; \( y_i^{(k_1)} = [y_i, y_{i-\tau_1}, \ldots, y_{i-(k_1-1)\tau_1}] \) and \( x_i^{(l_1)} = [x_i, x_{i-\tau_1}, \ldots, x_{i-(l_1-1)\tau_1}] \) denote embedding vectors with elements from the past values of \( y \) and \( x \), respectively. \((k_1, l_1)\) is the embedding dimension of \( y \) and \( l_1 \) is the embedding dimension of \( x \); \( \tau_1 \) is the time interval that allows the scaling in time of the embedded vector. These parameters can be set to be \( k_1 \leq 3, \ l_1 \leq 3, \) and \( h_1 = \tau_1 \leq 4 \) as a rule of thumb [12]; \( p(y_{i+h_1}, y_i^{(k_1)}, x_i^{(l_1)}) \) denotes the joint probability distribution, and \( p(\cdot|\cdot) \) denotes the conditional probabilities, and thus \( p(y_{i+h_1} | y_i^{(k_1)}, x_i^{(l_1)}) \) denotes the probability that \( y_{i+h_1} \) has a certain value when past values \( y_i^{(k_1)} \) and \( x_i^{(l_1)} \) are known and \( p(y_{i+h_1} | y_i^{(k_1)}) \) denotes the probability that \( y_{i+h_1} \) has a certain value when past values \( y_i^{(k_1)} \) are known. The transfer entropy from \( x \) to \( y \) is calculated as follows:

\[
t_{x\rightarrow y} = \sum p(y_{i+h_1}, y_i^{(k_1)}, x_i^{(l_1)}) \cdot \log \frac{p(y_{i+h_1} | y_i^{(k_1)}, x_i^{(l_1)})}{p(y_{i+h_1} | y_i^{(k_1)})},
\]

(1)

where the sum symbol represents \( k_1 + l_1 + 1 \) sums over all amplitude bins of the joint probability distribution and conditional probabilities, and the base of the logarithm is 2.

Similarly, the transfer entropy from \( x \) to \( z \) is calculated as follows:

\[
t_{x\rightarrow z} = \sum p(z_{i+h_2}, z_i^{(m_1)}, x_i^{(l_2)}) \cdot \log \frac{p(z_{i+h_2} | z_i^{(m_1)}, x_i^{(l_2)})}{p(z_{i+h_2} | z_i^{(m_1)})},
\]

(2)

where \( h_2 \) is the prediction horizon, \( z_i^{(m_1)} = [z_i, z_{i-\tau_2}, \ldots, z_{i-(m_1-1)\tau_2}] \) and \( x_i^{(l_2)} = [x_i, x_{i-\tau_1}, \ldots, x_{i-(l_2-1)\tau_2}] \) are embedding vectors with time interval \( \tau_2 \).

The transfer entropy from \( z \) to \( y \) is calculated as follows:

\[
t_{z\rightarrow y} = \sum p(y_{i+h_3}, y_i^{(k_2)}, z_i^{(m_2)}) \cdot \log \frac{p(y_{i+h_3} | y_i^{(k_2)}, z_i^{(m_2)})}{p(y_{i+h_3} | y_i^{(k_2)})},
\]

(3)

where \( h_3 \) is the prediction horizon, \( y_i^{(k_2)} = [y_i, y_{i-\tau_1}, \ldots, y_{i-(k_2-1)\tau_1}] \) and \( z_i^{(m_2)} = [z_i, z_{i-\tau_2}, \ldots, z_{i-(m_2-1)\tau_2}] \) are embedding vectors with time interval \( \tau_3 \).

If \( t_{x\rightarrow y}, t_{x\rightarrow z}, \) and \( t_{z\rightarrow y} \) are all larger than zero, then we conclude that \( x \) causes \( y \), \( x \) causes \( z \), and \( z \) causes \( y \). In this case, we need to distinguish whether the causal influence from \( x \) to \( y \) is only via the indirect pathway through the intermediate variable \( z \), or in addition to this, there is another direct pathway from \( x \) to \( y \), as shown in Fig. 1. We define a direct causality from \( x \) to \( y \) as \( x \) directly causing \( y \), which means there is a direct information and/or material flow pathway from \( x \) to \( y \) without any intermediate variables.

In order to detect whether there is a direct causality from \( x \) to \( y \), we define a direct transfer entropy (DTE) from \( x \) to \( y \) as follows:

\[
d_{x\rightarrow y} = \sum p(y_{i+h_3}, z_i^{(m_2)} | x_i^{(l_1)}, y_i^{(k_1)}, x_i^{(l_1)}) \cdot \log \frac{p(y_{i+h_3} | z_i^{(m_2)}, x_i^{(l_1)}, y_i^{(k_1)}, x_i^{(l_1)})}{p(y_{i+h_3} | z_i^{(m_2)}, x_i^{(l_1)}, y_i^{(k_1)})},
\]

(4)

where the prediction horizon \( h_3 \) is set to be \( h = \max(h_1, h_3) \); the embedding vector \( z_i^{(m_2)} = [z_{i+h_3}, z_{i+h_3-\tau_3}, \ldots, z_{i+h_3-(m_2-1)\tau_3}] \) denotes the past values of \( z \) which can provide useful information for predicting the future \( y \) at time instant \( i + h \), where \( m_2 \) and \( \tau_3 \) are determined by (3); the embedding vector \( x_i^{(l_1)} = [x_i, x_{i+h_1-\tau_1}, \ldots, x_i-(l_1-1)\tau_2] \) denotes the past values of \( x \) which can provide useful information to predict the future \( y \) at time instant \( i + h \), where \( l_1 \) and \( \tau_1 \) are determined by (1). Note that the parameters in DTE are all determined by the calculation of the transfer entropies for consistency.

The DTE represents the information about a future observation of \( y \) obtained from the simultaneous observation of past values of both \( x \) and \( z \), after discarding the information about the future \( y \) obtained from the past \( z \) alone. This can be understood as follows: if the pathway from \( z \) to \( y \) is cut off, will the history of \( x \) still provide some helpful information to predict the future \( y \)? Obviously, if this information is non-zero (greater than zero), then there is a direct pathway from \( x \) to \( y \). Otherwise there is no direct pathway from \( x \) to \( y \), and the causal influence from \( x \) to \( y \) is all along the indirect pathway via the intermediate variable \( z \).

Note that the direct causality here is a relative concept; since the measured process variables are limited, the direct causality analysis is only based on these variables. In other words, even if there are intermediate variables in the connecting pathway between two measured variables, as long as none of these intermediate variables is measured, we still state that the causality is direct between the pair of measured variables.

After the calculation of \( d_{x\rightarrow y} \), if there is direct causality from \( x \) to \( y \), we need to further judge whether the causality from \( z \) to \( y \) is true or spurious, because it is possible that \( z \) is not a cause of \( y \) and the spurious causality from \( z \) to \( y \) is generated by \( x \), i.e., \( x \) is the common source of both \( z \) and \( y \). As shown in Fig. 2, there are still two cases of
the information and material flow pathways between x, y, and z, and the difference is whether there is true and direct causality from z to y.

Thus, the DTE from z to y needs to be calculated:

\[ d_{z \rightarrow y} = \sum p(y_{i+h}, z_{i+h-h_{1}}, x_{i+h-h_{2}}) \cdot \log \frac{p(y_{i+h}|z_{i+h-h_{1}}, x_{i+h-h_{2}})}{p(y_{i+h}|z_{i+h-h_{1}})}, \]

(5)

where the parameters are the same as in (4). If \( d_{z \rightarrow y} > 0 \), then there is true and direct causality from z to y, as shown in Fig. 2(a). Otherwise, the causality from z to y is spurious, which is generated by the common source x, as shown in Fig. 2(b).

The transfer entropy and DTE mentioned above are restricted to discrete random variables. To extend these ideas to handle continuous random variables, the differential transfer entropy and the differential DTE can be defined in a similar way. The probability distribution function (PDF) \( f \), and the sum should be changed to the integration.

The differential transfer entropy from x to y, for continuous variables, is calculated as follows:

\[ T_{x \rightarrow y} = \int f(y_{i+h_{1}}, y_{i}^{(k_{1})}, x_{i}^{(l_{1})}) \cdot \log \frac{f(y_{i+h_{1}}, y_{i}^{(k_{1})}, x_{i}^{(l_{1})})}{f(y_{i+h_{1}}|y_{i}^{(k_{1})}, x_{i}^{(l_{1})})} dw, \]

(6)

where \( f(y_{i+h_{1}}, y_{i}^{(k_{1})}, x_{i}^{(l_{1})}) \) denotes the joint PDF; \( f(\cdot) \) denotes the conditional PDF; \( w \) denotes the random vector \([y_{i+h_{1}}, y_{i}^{(k_{1})}, x_{i}^{(l_{1})}]\). By assuming that the elements of \( w \) are \( w_{1}, w_{2}, \ldots, w_{s} \), \( \int (\cdot) dw \) denotes \( \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (\cdot) dw_{1} \cdots dw_{s} \) for simplicity, and the following similar notations have the same meaning as this one. The meaning of other parameters remains unchanged.

The differential transfer entropies from x to z, \( T_{x \rightarrow z} \), and from z to y, \( T_{z \rightarrow y} \), are also calculated in the same way.

Similarly, a differential DTE is defined as follows:

\[ D_{x \rightarrow y} = \int f(y_{i+h_{1}}, z_{i+h-h_{2}}^{(m_{2})}, x_{i+h-h_{1}}^{(l_{1})}) \cdot \log \frac{f(y_{i+h_{1}}, z_{i+h-h_{2}}^{(m_{2})}, x_{i+h-h_{1}}^{(l_{1})})}{f(y_{i+h_{1}}|z_{i+h-h_{2}}^{(m_{2})}, x_{i+h-h_{1}}^{(l_{1})})} dv, \]

(7)

where \( v \) denotes the random vector \([y_{i+h_{1}}, z_{i+h-h_{2}}^{(m_{2})}, x_{i+h-h_{1}}^{(l_{1})}]\). The definitions of other quantities are similar to that in (4). Note that the calculation of differential DTE from z to y, \( D_{z \rightarrow y} \), is also similar to \( D_{x \rightarrow y} \).

B. Relationships Between DTE and Differential DTE

We establish a connection between the discrete transfer entropy and the differential transfer entropy, and between the discrete DTE and the differential DTE.

First, we establish the relationship between the discrete transfer entropy and the differential transfer entropy. From (1) we can express the discrete transfer entropy using conditional Shannon entropies.

\[ D_{x \rightarrow y} = \lim_{\Delta x, \Delta y \rightarrow 0} \frac{\Delta x}{\Delta y} \int f(y_{i+h_{1}}, y_{i}^{(k_{1})}, x_{i}^{(l_{1})}) \cdot \log \frac{f(y_{i+h_{1}}, y_{i}^{(k_{1})}, x_{i}^{(l_{1})})}{f(y_{i+h_{1}}|y_{i}^{(k_{1})}, x_{i}^{(l_{1})})} dw, \]

(8)

where \( H(y_{i+h_{1}|y_{i}^{(k_{1})}, x_{i}^{(l_{1})}}) \) and \( H(y_{i+h_{1}|y_{i}^{(k_{1})}, x_{i}^{(l_{1})}}) \) are the conditional Shannon entropies.

For a discrete random variable, the bin size is the implicit width of each of the \( n \) bins whose probabilities are denoted by \( p_{n} \). When we generalize it to the continuous domain, we assume the bin sizes of the continuous random variables x, y, and z are \( \Delta x, \Delta y, \) and \( \Delta z \), respectively. Then, as the bin sizes approach to zero, the probability \( p(y_{i+h_{1}}, y_{i}^{(k_{1})}, x_{i}^{(l_{1})}) \) in (8) is approximated by \( \Delta y \Delta k_{y} \Delta y^{k_{1}} f(y_{i+h_{1}}, y_{i}^{(k_{1})}, x_{i}^{(l_{1})}) \) and the transfer entropy from x to y can be approximated by the discrete transfer entropy:

\[ \lim_{\Delta x, \Delta y \rightarrow 0} \frac{\Delta x}{\Delta y} \int f(y_{i+h_{1}}, y_{i}^{(k_{1})}, x_{i}^{(l_{1})}) \cdot \log \frac{f(y_{i+h_{1}}, y_{i}^{(k_{1})}, x_{i}^{(l_{1})})}{f(y_{i+h_{1}}|y_{i}^{(k_{1})}, x_{i}^{(l_{1})})} dw, \]

(9)

As \( \Delta x, \Delta y \rightarrow 0 \), we have

\[ \sum \Delta y \Delta y^{k_{1}} f(y_{i+h_{1}}, y_{i}^{(k_{1})}, x_{i}^{(l_{1})}) \rightarrow \int f(y_{i+h_{1}}, y_{i}^{(k_{1})}, x_{i}^{(l_{1})}) dw = 1, \]

\[ \sum \Delta y \Delta y^{k_{1}} f(y_{i+h_{1}}, y_{i}^{(k_{1})}, x_{i}^{(l_{1})}) \rightarrow \int f(y_{i+h_{1}}, y_{i}^{(k_{1})}, x_{i}^{(l_{1})}) du = 1, \]

and the integral of the function \( f(\cdot) \log f(\cdot) \) can be approx-
imated in the Riemannian sense by
\[ \sum \Delta y \Delta x \mathcal{D}_g \mathcal{D}_f \frac{d}{dx} f(y_{i+h}, y_{i}, x_{i+1}) \log f(y_{i+h}, y_{i}, x_{i+1}) \] 
\[ \rightarrow \int f(y_{i+h}, y_{i}, x_{i+1}) \log f(y_{i+h}, y_{i}, x_{i+1}) d\mathbf{w}, \]
\[ \sum \Delta y \Delta x \mathcal{D}_g \mathcal{D}_f \frac{d}{dy} f(y_{i+h}, y_{i}, x_{i+1}) \log f(y_{i+h}, y_{i}, x_{i+1}) \] 
\[ \rightarrow \int f(y_{i+h}, y_{i}, x_{i+1}) \log f(y_{i+h}, y_{i}, x_{i+1}) d\mathbf{u}. \]
Thus,
\[ \lim_{\Delta y, \Delta x \to 0} \frac{d}{dx} f \xrightarrow{\Delta x \to 0} \frac{\int f(y_{i+h}, y_{i}, x_{i+1}) \log f(y_{i+h}, y_{i}, x_{i+1}) d\mathbf{w}}{\int f(y_{i+h}, y_{i}, x_{i+1}) d\mathbf{w}} \]
\[ - \lim_{\Delta y, \Delta x \to 0} \frac{\int f(y_{i+h}, y_{i}, x_{i+1}) \log f(y_{i+h}, y_{i}, x_{i+1}) d\mathbf{u}}{\int f(y_{i+h}, y_{i}, x_{i+1}) d\mathbf{u}} \]
\[ = \int f(y_{i+h}, y_{i}, x_{i+1}) \log f(y_{i+h}, y_{i}, x_{i+1}) d\mathbf{w} \]
\[ - \int f(y_{i+h}, y_{i}, x_{i+1}) \log f(y_{i+h}, y_{i}, x_{i+1}) d\mathbf{u} \]
\[ = \int f(y_{i+h}, y_{i}, x_{i+1}) \log f(y_{i+h}, y_{i}, x_{i+1}) d\mathbf{w} \]
\[ = T_{x \to y}. \]  
(10)

This means that the differential transfer entropy is the limit of the discrete transfer entropy as the bin sizes of both \( x \) and \( y \) approach to zero.

For the differential DTE and the discrete DTE, using the same proof procedure with the transfer entropy, we can obtain
\[ \lim_{\Delta x, \Delta y, \Delta z \to 0} d\mathcal{D}_x = D_{x \to y}, \]
which means that the differential DTE is the limit of the discrete DTE as the bin sizes of \( x, y, \) and \( z \) approach to zero.

C. Calculation Method

1) Estimation of the Differential Transfer Entropy and the Differential DTE: Since the data analyzed here is uniformly sampled data, as obtained from processes that are continuous, the proposed differential transfer entropy and the differential DTE are used in this paper.

For the transfer entropy from \( x \) to \( y \), since (6) can be written as:
\[ T_{x \to y} = E \left\{ \log \frac{f(y_{i+h}, y_{i}, x_{i+1})}{f(y_{i+h}, y_{i})} \right\}, \]
it can be approximated by
\[ T_{x \to y} = \frac{1}{N - h - 1 - r + 1} \sum_{i=r}^{N-h-1} \log \frac{f(y_{i+h}, y_{i}, x_{i+1})}{f(y_{i+h}, y_{i})}, \]  
(11)
where \( N \) is the number of samples and \( r = \max\{(k_i - 1)\tau_{1} + 1, (l_i - 1)\tau_{1} + 1\} \).

Just as with transfer entropy, the differential DTE (7) can be approximated by
\[ D_{x \to y} = \frac{1}{N - h - j + 1} \sum_{i=j}^{N-h} \log \frac{f(y_{i+h}, y_{i}, x_{i+1})}{f(y_{i+h}, y_{i})}, \]  
(12)
where \( j = \max\{-h + h_3 + (m_2 - 1)\tau_3 + 1, -h + h_1 + (l_1 - 1)\tau_1 + 1\} \).

2) Kernel Estimation of PDFs: In (11) and (12), the conditional PDFs are expressed by joint PDFs and then obtained by the kernel estimation method [17]. Here the following Gaussian kernel function is used:
\[ k(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2}. \]
Then a univariate PDF can be estimated by
\[ \hat{f}(x) = \frac{1}{Nh} \sum_{i=1}^{N} k\left( \frac{x - X_i}{h} \right), \]
where \( N \) is the number of samples, and \( h \) is the bandwidth chosen to minimize the mean integrated squared error of the PDF estimation and calculated by \( h = 1.06N^{-1/5} \) according to the “normal reference rule-of-thumb” [18], where \( \sigma \) is the standard deviation of the sampled data \( \{X_i\}_{i=1}^{N} \).

For \( q \) dimensional multivariate data, we use the Fukunaga method [17] to estimate the joint PDF. Suppose that \( X_1, \ldots, X_N \) constitute a \( q \) dimensional i.i.d. vector \( (X_i)_{i \in \mathbb{R}^q} \) with a common PDF \( f(x_1, x_2, \ldots, x_q) \). Let \( x \) denote the \( q \) dimensional vector \( (x_1, x_2, \ldots, x_q) \), then the kernel estimation of the joint PDF is
\[ \hat{f}(x) = (\det(S))^{-1/2} \frac{1}{Nh^q} \sum_{i=1}^{N} K\left\{ H^{-1/2}(x - X_i)\right\} S^{-1}(x - X_i), \]

where \( H = 1.06N^{-1/(4+q)} \), \( S \) is the covariance matrix of the sampled data, and \( K(u) = (2\pi)^{-q/2} e^{-\frac{1}{2}u^2} \).

3) Normalization of the DTE: Similar to the transfer entropy, it can be shown that the DTE represents the conditional mutual information; thus it is always non-negative. However, small values of the DTE suggest no direct causality while large values do. In order to quantify the strength of the direct causality, normalization is necessary.

Since the differential DTE in (7) represents the information directly provided from the past \( x \) to the future \( y \), a normalized differential DTE is defined as
\[ NDT_E^{x \to y} = \frac{D_{x \to y}}{T_{x \to y}} \in [0, 1]. \]  
(13)
This represents the percentage of direct causality in the total causality from \( x \) to \( y \).

D. Extension to Multiple Intermediate Variables

The definition of the DTE from \( x \) to \( y \) can be easily extended to multiple intermediate variables \( z_1, z_2, \ldots, z_q \):
\[ d_{x \to y} = \sum p(y_{i+h}, z_{1,(i+1)}, \ldots, z_{q,(i+1)}, x_{i+h}) \cdot \log \frac{p(y_{i+h}, z_{1,(i+1)}, \ldots, z_{q,(i+1)}, x_{i+h})}{p(y_{i+h}, z_{1,i}, \ldots, z_{q,i}, x_{i+h})}, \]  
(14)
III. Examples and an Experimental Case Study

In this section, we give two examples and an experimental case study to show the usefulness of the proposed methods.

A. Examples

Example 1: Assume three linear correlated continuous random variables \( x, y, \) and \( z \) satisfying:

\[
\begin{align*}
    z_{k+1} &= 0.8x_k + 0.2z_k + v_{1k}, \\
y_{k+1} &= 0.6z_k + v_{2k},
\end{align*}
\]

where \( x_k \sim N(0,1), \) \( v_{1k}, v_{2k} \sim N(0,0.1), \) and \( z(0) = 3.2. \) The simulation data consists of 6000 samples. To assure stationarity, the initial 3000 data points were discarded.

According to (6), we chose \( h_1 = h_2 = h_3 = 1, \) \( \tau_1 = \tau_2 = \tau_3 = 1, \) \( k_1 = m_1 = k_2 = 0, \) and \( l_1 = l_2 = m_2 = 2; \) these parameters are determined using the same procedure as in [12], and we will also use these values later. The calculated transfer entropies between each pair of \( x, z, \) and \( y \) are shown in Table I. From this table, we can see that \( x \) causes \( z, z \) causes \( y, \) and \( x \) causes \( y \) because \( T_{x \rightarrow z} = 1.5563, \) \( T_{z \rightarrow y} = 1.0083, \) and \( T_{x \rightarrow y} = 0.7722 \) are relatively large. Thus we need to first determine whether there is direct causality from \( x \) to \( y. \) According to (7), we obtain \( D_{x \rightarrow y} = 0.3430. \) From (13), the normalized DTE from \( x \) to \( y \) is \( NDTE_{x \rightarrow y}^z = 0.3430/0.6905 = 0.4967, \) which is much larger than zero. Thus, we conclude that there is direct causality from \( x \) to \( y. \) Second, we need to detect whether there is true and direct causality from \( z \) to \( y. \) According to (5), we obtain \( D_{z \rightarrow y} = 0.5082, \) and thus the normalized DTE from \( z \) to \( y \) is \( NDTE_{z \rightarrow y}^z = 0.5082/0.8557 = 0.5939, \) which is much larger than zero. Hence, we conclude that there is true and direct causality from \( z \) to \( y. \) The information flow pathways for Example 2 are shown in Fig. 3(b).

This conclusion is consistent with the mathematical function, from which we can see that there are direct information flow pathways both from \( x \) to \( y \) and from \( z \) to \( y. \)

No matter whether the relationships of variables are linear or nonlinear, the DTE can detect the direct causality and the normalized DTE can quantify the strength of the direct causality.

B. Experimental Case Study

In order to show the effectiveness of the proposed methods, a 3-tank experiment is implemented. The schematic of the 3-tank system is shown in Fig. 4. Water is drawn from a reservoir to tanks 1 and 2 by a gear pump and a three way valve. The water in tank 2 can flow down into tank 3. The water in tanks 1 and 3 eventually flows down into the reservoir. The experiment is conducted under open-loop conditions.

The water levels are measured by level transmitters. We denote the water levels of tanks 1, 2, and 3 by \( x_1, x_2, \) and \( x_3, \) respectively. The flow rate of the water out of the pump is measured by a flow meter; we denote this flow rate by \( x_4. \) In this experiment, the normal flow rate of the water out of the pump is 10 L/min. However, the flow rate varies randomly with a mean value of 10 L/min because of the noise in the sensor and minor fluctuations in the pump. The

---

**TABLE I**

Calculated transfer entropies for Example 1.

<table>
<thead>
<tr>
<th>( T_{\text{row} \rightarrow \text{column}} )</th>
<th>( x )</th>
<th>( z )</th>
<th>( y )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x )</td>
<td>NA</td>
<td>1.5563</td>
<td>0.7722</td>
</tr>
<tr>
<td>( z )</td>
<td>0.0710</td>
<td>NA</td>
<td>1.0083</td>
</tr>
<tr>
<td>( y )</td>
<td>0.0669</td>
<td>0.0654</td>
<td>NA</td>
</tr>
</tbody>
</table>

**TABLE II**

Calculated transfer entropies for Example 2.

<table>
<thead>
<tr>
<th>( T_{\text{row} \rightarrow \text{column}} )</th>
<th>( x )</th>
<th>( z )</th>
<th>( y )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x )</td>
<td>NA</td>
<td>1.1416</td>
<td>0.6905</td>
</tr>
<tr>
<td>( z )</td>
<td>0</td>
<td>NA</td>
<td>0.8557</td>
</tr>
<tr>
<td>( y )</td>
<td>0</td>
<td>0.0245</td>
<td>NA</td>
</tr>
</tbody>
</table>

---

Fig. 3. Information and material flow pathways for (a) Example 1, (b) Example 2, and (c) 3-tank system.

causality from \( x \) to \( y, \) and the causal effects from \( x \) to \( y \) are all along the indirect pathways via the intermediate variables \( z_1, z_2, \ldots, z_q. \) If \( d_{x \rightarrow y} \) is larger than zero, then there is direct causality from \( x \) to \( y. \)
stationary sampled data with length of 3000 are analyzed. The sampling time is one second.

The calculated transfer entropies between each pair of \( x_1 \), \( x_2 \), \( x_3 \), and \( x_4 \) are shown in Table III. From this table, we can see that \( x_2 \) causes \( x_3 \), and \( x_4 \) causes \( x_1 \), \( x_2 \), and \( x_3 \). Since \( x_4 \) causes \( x_2 \), \( x_2 \) causes \( x_3 \), and \( x_4 \) causes \( x_3 \), we need to first detect whether there is direct causality from \( x_4 \) to \( x_3 \). According to (7), we obtain \( D_{x_4 \rightarrow x_3} = 0.0062 \). According to (13), the normalized DTE from \( x_4 \) to \( x_3 \) is \( NDTEx_4 \rightarrow x_3 = 0.0062/0.1391 = 0.0446 \), which is very small. Thus, we conclude that there is almost no direct causality from \( x_4 \) to \( x_3 \). The corresponding information and material flow pathways according to these calculation results are shown in Fig. 3(c), which are consistent with the calculated information and material flow pathways of the physical 3-tank system.

IV. CONCLUSION AND FUTURE WORK

In industrial processes, abnormalities often spread from one process variable to neighboring variables. It is important to determine the fault propagation pathways to find the root cause of the abnormalities and the corresponding fault propagation routes. A data-based direct causality detection method has been proposed to detect whether there is a direct information and/or material flow pathway between each pair of variables. The discrete DTE for discrete random variables and the differential DTE for continuous random variables have been defined. The differential DTE has been shown to be equivalent to the limit of the discrete DTE as the bin sizes approach to zero. The normalized differential DTE has been defined to measure the connectivity strength of the direct causality. The proposed method is suitable for both linear and nonlinear relationships and has been validated by two examples and an experimental case study.

Since the DTE is employed based on the calculation results of transfer entropies, it is important to determine whether there is causality from one variable to the other based on the calculation results of the transfer entropies. However, we can only make qualitative decisions based on observations and comparisons. For the conclusion of the direct causality based on the normalized DTE, the same question remains. Thus, our ongoing study is related to thresholds determination of the calculated transfer entropy and the normalized DTE.

REFERENCES