Causal Structure Identification from Corrupt Data-Streams

Venkat Ram Subramanian, Andrew Lamperski, and Murti V. Salapaka

Abstract—Complex networked systems can be modeled and represented as graphs, with nodes representing the agents and the links describing the dynamic coupling between them. The fundamental objective of network identification for dynamic systems is to identify causal influence pathways. However, dynamically related data-streams that originate from different sources are prone to corruption caused by asynchronous time-stamps, packet drops, and noise. In this article, we show that identifying causal structure using corrupt measurements results in the inference of spurious links. A necessary and sufficient condition that delineates the effects of corruption on a set of nodes is obtained. Our theory applies to nonlinear systems, and systems with feedback loops. Our results are obtained by the analysis of conditional directed information in dynamic Bayesian networks. We provide consistency results for the conditional directed information estimator that we use by showing almost-sure convergence.

I. INTRODUCTION

Models of systems as networks of interacting systems are central to many domains such as climate science [1], geoscience [2], biological systems [3] [4], quantitative finance [5], social sciences [6], and in many engineered systems like the Internet of Things [7] and wireless sensor networks [8]. In many scenarios such as the power grid [9] and metabolic pathways in cells [10] it is impractical or impermissible to externally influence the system. Here causal structure identification via passive means is to be accomplished. With advancements in measurement technology, data processing and communication systems coupled with sensors and measurement devices becoming inexpensive, passive identification of causal graphs of dynamically related agents is becoming more tenable. Often, the data-streams in such large systems are not immune to effects of noise [11], asynchronous sensor clocks [12], and packet drops [13]. When dealing with problems of identifying structural and functional connectivity of a large network, there is a pressing need to rigorously study such uncertainties and address detrimental effects of corrupt data-streams on network reconstruction.

A. Related Work

Network identification for linear systems is extensively studied. Methods for identifying transfer functions that dynamically link nodes from time-series data are provided in [14], [15], and [16]. However, these works assume that the time-series are perfect.

Authors in [17] leveraged multivariate Wiener filters to reconstruct the undirected topology of the generative network model. Moreover, assuming that the interaction dynamics are strictly causal and using multivariate estimation based on a Granger filter, it was shown that the interaction structure can be accurately recovered with directions, and without any spurious links. Here too, results assume data to be uncorrupted with the interaction between agents governed via Linear time-invariant (LTI) dynamics.

For a network of interacting agents with nonlinear, dynamic dependencies and strictly causal interactions, the authors in [18] proposed the use of directed information to determine the directed structure of the network. Sufficient conditions to recover the directed structure are provided. More recently, [19], [20] defined and used information transfer to determine underlying causal interactions in a power network. Here too it is assumed that the data-streams are ideal with no distortions.

The authors in [21], [22] use dynamical structure functions (DSF) for network reconstruction [23] and consider measurement noise and non-linearities in the network dynamics. The proposed method first finds optimal DSF for all possible Boolean structures and then adopt a model selection procedure to determine the best estimate. The authors concluded that the performance of their algorithms degrades as noise, network size and non-linearities increase. However, a precise characterization of drawing spurious inferences in structure is not provided. In this article, we provide exact location of spurious links that arise during network reconstruction from corrupt data-streams.

Inspite its significance, little is known on the effects of uncertainties in the data-streams on network identification. Recently in [24], the issues of observation noise and undersampling on causal discovery from time-series data has been addressed. Although authors concluded that spurious links can be inferred, a rigorous characterization of such links was not proven nor a generalization of corruption models was provided. In [25] focusing on networks with linear time-invariant (LTI) interactions, authors provided characterization of the extent of spurious links that can appear due to data-corruption. However, the analysis is restricted to LTI systems. Moreover, in [25] the objective is to determine the topology of the networked system and not to deduce the directions.

B. Our Contribution

In this article, we focus our study to determine the directed structure of a network thereby inferring the causal structure of the network, using non-invasive means from corrupt data-streams. We consider networks admitting non-linear and strictly causal dynamical interactions.

We provide necessary and sufficient conditions to determine the directed network structure from corrupt data-streams. We
present tight characterization for the spurious links that arise due to corruption of data-streams by determining their location and orientation.

In [26], preliminary results that characterized the spurious links, in the framework of this article are provided. However, the analysis was limited to dynamical interactions such that every node was dependent dynamically on the entire history (strict) of its parent nodes. In this article, we consider general class of non-linear systems by relaxing the above assumption on dynamics. Moreover, we provide detailed and rigorous proofs to generalize the results obtained in [26] wherein only a proof sketch was provided. In addition, we establish convergence results for the estimator that we use to determine conditional directed information.

C. Paper Organization

We review needed graph theory notions and describe the framework for generative models in Section II. In Section III we provide models to characterize corruption of data streams that captures time uncertainty, packet loss and measurement noise. The methods to infer directed network structure for non-linear dynamical systems are described in Section IV. Our directed information estimator and simulation results are described in Section V. Finally, a conclusion is provided in Section VI.

II. PRELIMINARIES

A. Notations

\( y[\cdot] \) denotes a sequence and \( y(t) \) denotes the sequence \( y[0], y[1], \ldots, y[t] \).

\( P_X \) represents the probability density function of a random variable \( X \).

\( X \perp Y \) denotes that the random variables \( X \) and \( Y \) are independent.

\( \mathbb{E}[\cdot] \) denotes the expectation operator.

B. Graph Theory Background

In this subsection, few terminologies from graph theory that will be extensively used for network structure inference are reviewed. For further reference, see [27].

Definition 1 (Directed Graph). A directed graph \( G \) is a pair \( (V, A) \) where \( V \) is a set of vertices or nodes and \( A \) is a set of edges given by ordered pairs \( (i, j) \) where \( i, j \in V \). If \( (i, j) \in A \), then we say that there is an edge from \( i \) to \( j \).

We shall use \( i \rightarrow j \) indicates an arc or edge or link from node \( i \) to node \( j \) in a directed graph. \( i \sim j \) denotes one of \( i \rightarrow j \) or \( j \rightarrow i \).

Definition 2 (Children and Parents). Given a directed graph \( G = (V, A) \) and a node \( j \in V \), the children of \( j \) are defined as \( C(j) := \{ i \mid i \rightarrow j \in A \} \) and the parents of \( j \) as \( P(j) := \{ i \mid i \rightarrow j \in A \} \).

Definition 3 (Trail/Path). Nodes \( v_1, v_2, \ldots, v_k \in V \) forms a trail or a path in a directed graph, \( G \), if for every \( i = 1, 2, \ldots, k-1 \) we have \( v_i \rightarrow v_{i+1} \).

\[
\begin{array}{c}
1 \rightarrow 2 \\
\end{array}
\]

(a) Trail connecting 1 and 4 is active given \( Z = \{ \} \).

\[
\begin{array}{c}
1 \rightarrow 2 \\
\end{array}
\]

(b) Trail connecting 1 and 4 is active given \( Z = \{ 2 \} \).

Fig. 1: This figure shows when the trail connecting nodes 1 and 4 is active given \( Z \).

Definition 4 (Chain). In a directed graph \( G \), a chain from node \( v_i \) to node \( v_j \) comprises of a sequence of \( k \) nodes such that \( v_i \rightarrow w_1 \rightarrow \cdots \rightarrow w_{k-2} \rightarrow v_j \) holds in \( G \).

Definition 5 (Descendants and Ancestors). Suppose there exists a chain from a node \( v_j \) to \( v_k \) in a directed graph, \( G \). Then, \( v_k \) is called a descendant of node \( v_j \) and \( v_j \) is called an ancestor of \( v_k \).

Definition 6 (Fork). A node \( v_k \) is a fork in a directed graph \( G \), if there are two other nodes \( v_i, v_j \) such that \( v_i \leftarrow v_k \rightarrow v_j \) holds.

Definition 7 (Collider). A node \( v_k \) is a collider in a directed graph, \( G \), if there are two other nodes \( v_i, v_j \) such that \( v_i \rightarrow v_k \leftarrow v_j \) holds.

Definition 8 (Active Trail). In a directed graph \( G \), a trail \( v_1 \rightarrow v_2 \rightarrow \cdots \rightarrow v_n \) is active given a set of nodes \( Z \) if one of the following statements holds for every triple \( v_{m-1} \rightarrow v_m \rightarrow v_{m+1} \) along the trail:

a) If \( v_m \) is not a collider, then \( v_m \notin Z \).

b) If \( v_m \) is a collider, then \( v_m \) or one of its descendants is in \( Z \).

where \( m \in \{2, \ldots, n-1\} \).

See Figure 1 for an illustration.

Definition 9 (d-separation). Let \( X, Y \) and \( Z \) be a set of nodes in a directed graph, \( G \). In \( G \), \( X \) and \( Y \) are d-separated by \( Z \) if and only if there is no active trail between any \( x \in X \) and any \( y \in Y \) given \( Z \). It is denoted as \( d-sep(X, Y \mid Z) \).

Definition 10 (Directed Cycle). A directed cycle from a node \( v_i \) to \( v_i \) in a directed graph, \( G \), has the form \( v_i \rightarrow w_1 \rightarrow \cdots \rightarrow w_k \rightarrow v_i \) for some set of nodes \( \{ w_n \}_{n=1}^k \) in \( G \).

Definition 11 (Directed Acyclic Graph). A directed graph with no directed cycles is called a directed acyclic graph (DAG).

Definition 12 (Bayesian Network). Suppose \( G = (V, A) \) is a DAG whose \( N \) nodes represent random variables \( a_1, \ldots, a_N \). \( G \) is called a Bayesian network (BN) if for any three subsets \( X, Y \) and \( Z \) of \( V \), \( d-sep(X, Y \mid Z) \) implies \( X \) is independent of \( Y \) given \( Z \).

Definition 13 (Faithful Bayesian network). Suppose \( G = (V, A) \) is a DAG whose \( N \) nodes represent random variables \( a_1, \ldots, a_N \). \( G \) is called a Faithful Bayesian network if for any three subsets \( X, Y \) and \( Z \) of \( V \), it holds that \( X \) and \( Y \) are independent given \( Z \), if and only if \( d-sep(X, Y \mid Z) \) is true.
C. Generative Model

In this subsection, the generative model that is assumed to generate the measured data is described. Consider $N$ agents that interact over a network. For each agent $i$, we associate a discrete time sequence $y_i[t]$ and a sequence $e_i[t]$. We assume $e_i$ and $y_i$ to be random processes. The process $e_i[t]$ is considered innate to agent $i$ and thus $e_i$ is independent of $e_j$ if $i \neq j$. Moreover, $e_i$ is considered to be uncorrelated across time. Let $Y$ denote the set of all random process \( \{y_1, \ldots, y_N\} \) with a parent set $P^g(i)$ defined for $i = 1, \ldots, N$. We consider strictly causal non-linear dynamical relations. Here, $i$ can belong to the parent set $P^g(i)$. The generative model takes the form:

\[
y_i[t] = f_i \left( y_i^{(t-1)}, \bigcup_{j \in P^g(i)} y_j^{(t-1)}, e_i[t] \right), \tag{1}
\]

where $f_i$’s can be any finite valued non-linear function such that $|f_i| < \infty$.

For an illustration, consider the dynamics of a generative model described by:

\[
\begin{align*}
y_1[t] &= y_1[t-1]y_2[t-2] + e_1[t], \\
y_2[t] &= \sin(y_1[t-1] \cdot y_2[t-1] + e_2[t]), \\
y_3[t] &= (y_1[t-1] + y_2[t-1]) \cdot e_3[t], \\
y_4[t] &= y_2[t-1]^2 + y_3[t-2] + y_4[t-1] + e_4[t], \\
y_5[t] &= y_5[t-1] \cdot y_4[t-1] + e_5[t].
\end{align*}
\]

We remark that if $y_j$ appears on the right hand side of $y_i$ for any time instant $t$, then $j \in P^g(i)$; the parent set is thus not dependent on time.

D. Graphical Representation

Here we describe how networks of dynamical systems are represented by graphs.

**Generative Graph:** The structural description of (1) induces a generative graph $G = (V, A)$ formed by identifying each vertex $v_i$ in $V$ with random process $y_i$ and the set of directed links, $A$, obtained by introducing a directed link from every element in the parent set $P^g(i)$ of agent $i$ to $i$. Note that we do not show $i \to i$ in the generative graph and neither do we show the processes $e_i$.

The generative graph associated with the example described in (2) is given by Fig. 3(a). Note that the generative graph describes the relationships between the stochastic processes $y_i$. When the time variable is unraveled we obtain the Dynamic Bayesian Network as defined below.

**Dynamic Bayesian Network (DBN):** Let $G = (V, A)$ be a generative graph. Let $y_i$ be as defined in (1) for all $i \in V$. Suppose all discrete time sequences have a finite horizon assumed to be $T$. Let $S_{ij}[t] = \{t' : y_j[t'] \in y_i[t_{i-1}] \}$ as an argument of $f_i$ in expression of $y_i[t]$ in (1) for all $j \in P^g(i) \cup \{i\}$ and for all $t$. Consider the graph $G' = (V', A')$ where $V' = \bigcup_{t \in \{0, \ldots, T\}} y_i[t]$ and

\[
A' = \bigcup_{t \in \{0, \ldots, T\}} \bigcup_{i \in V} \bigcup_{j \in P^g(i) \cup \{i\}} \left( \bigcup_{k \in S_{ij}[t]} \left(y_j[k] \to y_i[t]\right) \right).
\]

The joint distribution of $Y^{(T)}$ is given by:

\[
P_{Y^{(T)}} = P_{y_1[0]} \cdots P_{y_N[0]} \prod_{t=1}^T \prod_{i=1}^N P_{y_i[t]|P(y_i[t])}, \tag{3}
\]

where the parents of $y_i[t]$ are obtained from $G'$. It can be shown that $G'$ is the Bayesian network for the random variables $\{y_i[t] : t = 0, 1, 2, \ldots, T, i = 1, 2, \ldots, N\}$ and is considered the Dynamic Bayesian Network for $\{y_i : i = 1, 2, \ldots, N\}$ (see [27]). Figure 2(b) represents the DBN for the system in (2) for three time steps.

III. UNCERTAINTY DESCRIPTION

In this section we provide a description for how uncertainty affects the time-series $y_i$. We interchangeably use corruption or perturbation to denote uncertainties in data-streams.

A. General Perturbation Models

Consider the $i$th node in a generative graph and it’s associated unperturbed time-series $y_i$. The corrupt data-stream $u_i$ associated with $i$ follows:

\[
u_i[t] = g_i(y_i^{(t)}, u_i^{(t-1)}, \zeta_i[t]), \tag{4}
\]

where $u_i$ can depend dynamically on $y_i$ till time $t$, its own values in the strict past, and $\zeta_i[t]$ which represents a stochastic process that is independent across time. We highlight a few important perturbation models that are practically relevant. See [26] for more details.

**Temporal Uncertainty:** Consider a node $i$ in a generative graph. Suppose $t$ is the true clock index but the node $i$ measures a noisy clock index which is given by a random process, $\zeta_i[t]$. One such probabilistic model is given by the following IID Bernoulli process:

\[
\zeta_i[t] = \begin{cases} d_1, & \text{with probability } p_i, \\
d_2, & \text{with probability } (1 - p_i),
\end{cases}
\]

where $d_1$ and $d_2$ are any non-positive integers such that at least one of $d_1$ and $d_2$ are not equal to 0. Randomized delays
in information transmission can be modeled as a convolution operation with the impulse function $\delta[t]$ shifted by $\zeta_i[t]$ as follows:

$$ u_i[t] = \delta[t + \zeta_i[t]] \ast y_i[t], $$

where,

$$ \delta[t] = \begin{cases} 
1, & t = 0 \\
0, & t \neq 0. 
\end{cases} $$

**Noisy Filtering:** Given a node $i$ in a generative graph, the data-stream $y_i$ is causally filtered and corrupted with independent measurement noise $\zeta_i[\cdot]$. This perturbation model is described by:

$$ u_i[t] = (L_i \ast y_i)[t] + \zeta_i[t], $$

where $L_i$ is a stable causal linear time-invariant filter.

**Packet Drops:** The measurement $u_i[t]$ corresponding to an ideal data-point $y_i[t]$ packet reception at time $t$ can be stochastically modeled as:

$$ u_i[t] = \begin{cases} 
y_i[t], & \text{with probability } p_i \\
y_i[t-1], & \text{with probability } (1 - p_i). 
\end{cases} $$

Consider an IID Bernoulli process $\zeta_i$ described by,

$$ \zeta_i[t] = \begin{cases} 
1, & \text{with probability } p_i \\
0, & \text{with probability } (1 - p_i). 
\end{cases} $$

The corruption model in (4) takes the form:

$$ u_i[t] = \zeta_i[t]y_i[t] + (1 - \zeta_i[t])u_i[t - 1]. $$

**B. Perturbed Dynamic Bayesian Network**

Here, we provide a discussion on how the dynamic Bayesian network associated with the measured data-streams gets altered when the data-streams are subject to corruption. Consider a generative graph $G = (V, A)$. Let $y_i$ be as defined in (4) for all $i \in V$. Suppose all discrete time sequences have a finite horizon assumed to be $T$. Let $G' = (V', A')$ be the associated dynamic Bayesian network. Suppose $Z \subset V$ is the set of perturbed nodes with perturbation model described in (4). For $i \in Z$, the measured (corrupt) data-stream corresponding to agent $i$, $u_i$, is related to $y_i$ via (4). Let $U_Z = \{u_i\}_{i \in Z}$ and $Y_Z = \{y_j\}_{j \in Z}$ where $Z = V \setminus Z$. Due to corruption only $U_Z$ and $Y_Z$ are measured and observed. Denote the measured data-streams by $W = U_Z \cup Y_Z$. For all $j \in Z$ let $SU_j[t] = \{t' : y_j[t'] \in u_j[t-1]\}$ as an argument of $g_i$ in expression of $u_j[t]$ in (4) and let $SY_j[t] = \{t' : y_j[t'] \in y_j[t] \text{ as an argument of } g_i \text{ in expression of } u_j[t] \text{ in (4)}\}$ for all $t$. Consider the graph $G'_Z = (V'_Z, A'_Z)$

where $V'_Z = V' \cup \left( \bigcup_{k \in Z} u_k[t] \right)$ and $A'_Z = A' \cup \left( \bigcup_{k \in Z} y_k[t] \rightarrow u_k[t] \right) \cup \left( \bigcup_{k \in S_y[k]} u_k[t] \rightarrow u_k[t] \right)$ for all $t \in \{0, 1, 2, \ldots, T\}$. Note that the vertex set $V'_Z$ consists of all measurements given by the set $W$, and the uncorrupted versions $y_k$ of the corrupted versions $u_k$ for $k \in Z$.

Consider the set of random variables, $R = \{y_i[t] : i \in \{1, 2, \ldots, N\} \text{ and } t \in \{0, 1, 2, 3, \ldots, T\}\} \cup \{u_i[t] : i \in \{1, 2, \ldots, N\} \text{ and } t \in \{0, 1, 2, 3, \ldots, T\}\}$. The joint distribution $P_R$ is given by:

$$ P_R = \left( \prod_{i \in V} P_{u_i[0]} \right) \left( \prod_{j \in Z} P_{y_j[0]} \right) \left( \prod_{i=1}^{T} \prod_{j=1}^{N} P_{u_i[t]|P(u_i[t])]} \right) \cdot \left( \prod_{i=1}^{T} \prod_{j=1}^{N} P_{y_j[t]|P(y_j[t])]} \right), $$

where the parents of $u_i[t], y_j[t]$ are obtained from $G'_Z$. $G'_Z$ is the Bayesian Network for the random variables $R$ and is considered as the perturbed DBN (PDBN) associated with $U_Z \cup Y'_Z$.

Fig. 3 shows an example of a perturbed DBN corresponding to the generative graph in Fig. 2(a) for three time slices when node 1 data-streams are corrupt following a noisy filtering model described in (4).

**IV. STRUCTURE IDENTIFICATION**

**A. STRUCTURE INFERENCE FROM IDEAL DATA-STREAMS**

First, we recall how the structure of a generative graph can be inferred using directed information in the case of ideal data-streams. Consider a generative graph $G$ with $N$ nodes and let $Y$ denote the collection of $N$ data-streams that are measured. The authors in [13] defined and applied directed information (DI) in a network of dynamically interacting agents, to determine if a process causally influences another. A slightly modified definition of DI as defined in [13] is:

**Definition 14** (Directed Information). The directed information (DI) from data-stream $y_j$ to $y_i$ is given by:

$$ I(y_j \rightarrow y_i || Y_{ij}) = \mathbb{E} \left[ \log \frac{P_{y_i | y_j, Y_{ij}}}{P_{y_i | Y_{ij}}} \right], $$

(10)
where \( P_{y_i, Y_{ij}} = \prod_{t=1}^{T} P_{y_i[t] | y_i[t-1], y_j[t-1], Y_{ij}^{(t-1)}} \), \( P_{y_i, Y_{ij}} = \prod_{t=1}^{T} P_{y_i[t] | y_i[t-1], Y_{ij}^{(t-1)}} \) and \( Y_{ij} = Y \setminus \{y_i, y_j\} \).

Note that DI is always non-negative. So, if there is no directed edge from \( j \) to \( i \) in \( G \), then we must have that \( I(y_j \rightarrow y_i \mid Y_{ij}^{*}) = 0 \).

The following theorem was proved in [18] that specifies a necessary and sufficient condition to detect a presence of link in the generative graph.

**Theorem 1.** A directed edge from \( j \) to \( i \) exists in the directed graph \( G \) if and only if \( I(y_j \rightarrow y_i \mid Y_{ij}^{*}) > 0 \).

**Remark 1.** In [18], the authors assume positive distribution for the random processes in \( Y \). Under this assumption the result in Theorem 1 is both necessary and sufficient.

**B. Main Result: Inferring Directed Graphs from Corrupt Data-streams**

In this subsection, we will describe how data uncertainty will lead to spurious probabilistic relationships between nodes that are not connected in the original graph.

To present the main result in Theorem 2, some definitions are required.

**Definition 15 (Perturbed Graph).** Let \( G = (V, A) \) be a generative graph. Suppose \( Z \subseteq V \) is the set of perturbed nodes with each perturbation model admitting a description provided in [4]. The perturbed graph, \( G_Z = (V, AZ) \), is a directed graph where there is an edge \( i \rightarrow j \) in \( AZ \) if and only if there is a trail, \( trlG : i = v_1 - v_2 - \cdots - v_k = j \) in \( G \) such that the following conditions hold:

**P1** If \( j \notin Z \), then \( v_k - j \rightarrow j \in A \).

**P2** For \( m \in \{2, 3, \ldots, k - 1\} \), if \( v_m - v_m+1 \rightarrow v_m \), then \( v_m+1 \in Z \).

**P3** If \( v_m \) is a node such that \( v_m - v_m+1 \) is a sub-path of the path \( v_1 - \cdots - v_k \) and \( v_m \) is not a collider, then \( v_m \notin Z \).

**Remark 2.** Note that the existence of a trail that does not meet the ‘if’ conditions in [P1], [P2] and [P3] guarantees that \( i \rightarrow j \in AZ \). For example, if \( i \rightarrow j \in A \) then \( i \rightarrow j \in AZ \).

Indeed, if \( j \notin Z \) then \( i \rightarrow j \in AZ \) by condition [P1]. Conditions [P2] and [P3] are not applicable. On the other hand, if \( j \in Z \), then none of the conditions [P1], [P2] or [P3] are applicable to the trail \( i \rightarrow j \). So, \( i \rightarrow j \in AZ \).

**Definition 16 (Spurious Links).** Let \( G = (V, A) \) be a generative graph, \( Z \subseteq V \) be the set of perturbed nodes and \( G_Z = (V, AZ) \) be the perturbed graph. Spurious links are those links \( i \rightarrow j \in AZ \) that do not belong to \( A \).

The following lemma will be used to prove our main result in Theorem 2.

**Lemma 1.** Consider a generative graph, \( G = (V, A) \), consisting of \( N \) nodes. Let \( Z = \{v_1, \ldots, v_N\} \subseteq V \) be the set of \( N \) perturbed nodes where each perturbation is described by [4]. Denote the data-streams as follows: \( U_Z := \{u_t\}_{t \in Z} \) and \( Y_Z := \{y_t\}_{t \in Z} \) where \( Z = V \setminus Z \). Let the measured data-streams be \( W = U_Z \cup Y_Z = \{w_1, w_2, \ldots, w_N\} \). Let \( G' = (V', A') \) be the dynamic Bayesian network (DBN) associated with \( G \) and \( G_Z = (V, AZ) \) be the perturbed DBN. If \( i \rightarrow j \notin A \) and if a trail in \( G_Z \) between \( w_i(t-1) \) and \( w_j(t) \) contains a node \( \alpha_{b_m}[t_m] \) such that \( t_m \geq t_b \) and \( b_m \in V \), then for all \( t > 0 \), the trail is not active given \( \{w_j(t-1), W_{ij}^{(t-1)}\} \).

Proof. Consider any trail from a node in \( w_i(t-1) \) to \( w_j(t) \) in \( G_Z \). Denote this by \( trlG_Z : w_i(t) = \alpha_{b_m}[t] = \alpha_{b_{m+1}}[t] = \cdots = \alpha_{b_{m-1}}[t-1] = \alpha_{b_m}[t] \) where \( 0 \leq t_1 < t \). Here, \( b_k \) denotes the corresponding vertex in \( V \) for \( k = \{1, 2, \ldots, r\} \). Also, \( \alpha_{b_m}[k] = u_{b_k}[k] \) if \( b_k \in Z \) or \( \alpha_{b_m}[t] = y_m[t] \) otherwise. For compact notation, set \( \theta := \{w_j(t-1), W_{ij}^{(t-1)}\} \).

The trail has length at least 3. As \( i \rightarrow j \notin A \) and \( j \notin Z \), then \( y_j[t] \) does not dynamically depend on process \( y_i \) and clearly not on \( u_i \). If \( j \in Z \), then by [4], \( u_j[t] \) does not dynamically depend on \( y_j \) or \( u_j \). Thus, there is no direct link of the form \( \alpha_{b_m}[t'] \rightarrow \alpha_{b_m}[t''] \) in \( G_Z \), for any \( t', t'' \). In particular, \( w_i(t) \rightarrow w_j(t) \notin G_Z \). Thus, there are at least 3 nodes in the trail, \( trlG_Z \).

**Unobserved collider in trail.** Without loss of generality, choose \( t_m = \max\{1, \ldots, t_r\} \geq t \). Consider the sub-trail \( subl' := \alpha_{b_m}[t_m-1] \rightarrow \alpha_{b_m}[t_m] - \alpha_{b_{m+1}}[t_m+1] \) of \( trlG_Z \). By maximality of \( t_m \), \( t_m \geq t_{m-1} \) and \( t_m \geq t_{m+1} \). We will show that one of \( \alpha_{b_m}[t_m-1], \alpha_{b_m}[t_m], \) and \( \alpha_{b_{m+1}}[t_m+1] \) is a collider not in \( \theta \) and therefore the trail \( trlG_Z \) cannot be active given \( \theta \).

Suppose \( t_m > t_{m-1} \) and \( t_m > t_{m+1} \). Then, \( subl' \) is of the form \( \alpha_{b_m}[t_m-1] \rightarrow \alpha_{b_m}[t_m] \leftarrow \alpha_{b_{m+1}}[t_m+1] \). Note that, as \( t_m \geq t \), it follows that neither \( \alpha_{b_m}[t_m] \) nor any of its descendants can be in \( \theta \) and hence not observed.

Now, consider \( t_m > t_{m-1} \) and \( t_m = t_{m+1} \). (The case of \( t_m > t_{m+1} \) and \( t_m = t_{m-1} \) can be proven similarly).

By the generative model in [1], by strict causality, for any node \( p \in V \), \( y_p[t] \) does not dynamically depend on any \( y_q[t] \) for \( q \in \{p, P^p(p)\} \). By the perturbation model described by [4], for any \( q \in V \), \( u_q[t] \) dynamically depends only on \( \{w_{q(t-1)}, y_q[t]\} \). As \( t_m = t_{m+1} \), we therefore have \( b_m = b_{m+1} \) such that \( b_m \in Z \) and, one of \( \alpha_{b_m}[t_m] \) and \( \alpha_{b_{m+1}}[t_m+1] \) is actually a perturbed measurement \( u_{b_m}[t_m] \) while the other being \( y_{b_m}[t_m] \).

Suppose \( \alpha_{b_m}[t_m] = u_{b_m}[t_m] \). Then, \( \alpha_{b_{m+1}}[t_m+1] = y_{b_m}[t_m] \). As \( t_m > t_{m-1} \), \( subl' \) is in fact \( \alpha_{b_m}[t_m-1] \rightarrow \alpha_{b_m}[t_m] \leftarrow \alpha_{b_{m+1}}[t_m+1] \). Note that, \( \alpha_{b_m}[t_m] \) is a collider and as \( t_m \geq t \), this node is not observed in \( \theta \).

Suppose instead that \( \alpha_{b_m}[t_m] = y_{b_m}[t_m] \). Then, \( \alpha_{b_{m+1}}[t_m+1] = u_{b_m}[t_m] \). As \( b_m \in Z \) and maximality of \( t_m \) implies \( \alpha_{b_{m+2}}[t_m+2] \in \{u_{b_m}[t_m-1], y_{b_m}[t_m-1]\} \). Thus, we have \( \alpha_{b_m}[t_m] = \alpha_{b_{m+1}}[t_m+1] = u_{b_m}[t_m] \leftarrow \alpha_{b_{m+2}}[t_m+2] \) in \( trlG_Z \). Therefore, \( \alpha_{b_{m+1}}[t_m+1] \) is a collider not observed in \( \theta \).

The following theorem states that the perturbed graph precisely characterizes the spurious links which arise from probabilistic relationships that are spuriously introduced due to corruption.
Theorem 2. Consider a generative graph, $G = (V, A)$, consisting of $N$ nodes. Let $Z = \{v_1, \ldots, v_n\} \subseteq V$ be the set of $n$ perturbed nodes where each perturbation is described by $\{\alpha\}$. Denote the data-streams as follows: $U_Z := \{u_{ij} \mid (i, j) \in Z\}$ and $Y_Z := \{y_{ij} \mid (i, j) \in Z\}$ where $Z = V \setminus Z$. Let the measured data-streams be $W = U_Z \cup Y_Z = \{w_{11}, w_{22}, \ldots, w_{nn}\}$. Let the perturbed graph be $G' = (V, A_Z)$ and its associated perturbed DBN be $G'_Z = (V_Z, A'_Z)$. If $i \rightarrow j \notin A_Z$, then $d$-sep($w_{ij}[t], w_{ij}^{t-1}) \mid \{w_{ij}^{t-1}, W_{ji}^{t-1}\})$ holds in $G'_Z$ for all $t > 0$.

Proof. We will show that if $i \rightarrow j \notin A_Z$, then there is no trail between $w_{ij}^{t-1}$ and $w_{ij}[t]$ that is active given $\{w_{ij}^{t-1}, W_{ji}^{t-1}\}$ in $G'_Z$, for all $t > 0$. Let us consider the above procedure from the trail $G'_Z$: $w_{ij}[t] = \alpha_1[t_1] - \alpha_{b_2}[t_2] - \cdots - \alpha_{b_{m-1}}[t_{m-1}]+1 - \alpha_{b_m}[t_m] = w_{ij}[t]$ where $0 \leq t < t_m$ and $b_k$ denotes the corresponding vertex in $V$ for $k = \{1, 2, \ldots, m\}$. Here, $\alpha_i[t_v] = u_{ij}[t]$ if $v \in Z$ or $\alpha_i[t_v] = y_{ij}[t_v]$ otherwise. Using Lemma 1 if any $t_i$ in $\{t_2, \ldots, t_m\}$ is such that $t' > t$, then $G'_Z$ is not active. Now, consider $0 \leq t_1, t_2, t_3, \ldots, t_m < t$. Construct a trail in $G = (V, A)$, $tr_lG := i = v_1 - v_2 - v_3 \ldots - v_{k-1} - v_k = j$ from the trail $G'_Z$: $w_{ij}[t] = \alpha_{b_0}[t_1] - \alpha_{b_2}[t_2] - \cdots - \alpha_{b_{r-1}}[t_{r-1}]+1 - \alpha_{b_r}[t_r] = w_{ij}[t]$ as follows:

**Initialize:** $k = 1$ and $v_1 = b_1$.

for $i = 1 : r - 1$ do

if $b_i \neq b_{i+1}$ in $\alpha_{b_0}[t_i] - \alpha_{b_{i+1}}[t_{i+1}]$ along $tr_lG$ then 

Set $w_{k+1} = b_{i+1}$.

Add edge $v_k \rightarrow v_{k+1}$ with the same direction as $\alpha_{b_0}[t_i] - \alpha_{b_{i+1}}[t_{i+1}]$.

Set $s_k = t_i$ and $\tau_{k+1} = t_{i+1}$

end if

end for

Additionally, note that $v_k \rightarrow v_{k+1}$ corresponds to an edge $\alpha_{b_0}[s_k] - \alpha_{b_{i+1}}[s_{k+1}]$ in $G'_Z$.

Now, let us reason out why such a construction is always feasible. To this, we claim that for any successive pair $\alpha_{b_0}[t_i] - \alpha_{b_{i+1}}[t_{i+1}]$, either $b_i = b_{i+1}$ or $b_i \neq b_{i+1}$ and $b_i 

Finally consider the case that $\{\alpha\}$ is violated. Then along the trail, $tr_lG$, in $G$, there must be a sub-path $v_{m-1} \rightarrow v_m$ such that the intermediate node, $v_m$, is not a collider and $v_m \notin Z$. As $v_m$ is not a collider, there is one outgoing directed edge from $v_m$ in the trail $tr_lG$ to either $v_{m-1}$ or $v_{m+1}$. By construction, there must be a corresponding node $\alpha_{v_m}[t_f]$ in the trail $tr_lG'_Z$ such that it has an outgoing edge to either $\alpha_{v_{m-1}}[t_p]$ or $\alpha_{v_{m+1}}[t_q]$ for some $t_p > t_m$ or $t_q > t_m$ respectively. Clearly, there is one $\alpha_{v_m}[t_m]$ in $tr_lG'_Z$ which is a non-collider. Then, as $v_m \notin Z$, we must have that $\alpha_{v_m}[t_m] = \alpha_{v_m}[t_m] = y_{v_m}[t_m]$. Note that $v_m \neq i$ by condition loopi. As
with the perturbed graph, the perturbed dynamic Bayesian network (DBN) associated with the perturbed graph be $G$.

Corollary 1. Consider a generative graph, $G = (V, A)$, consisting of $N$ nodes. Let $Z = \{v_1, \ldots, v_n\} \subseteq V$ be the set of $n$ perturbed nodes where each perturbation is described by $(4)$. Denote the data-streams as follows: $U_Z := \{u_i\}_{i \in Z}$ and $Y_Z := \{y_j\}_{j \in \bar{Z}}$, where $\bar{Z} = V \setminus Z$. Let the measured data-streams be $W = U_Z \cup Y_Z = \{w_1, w_2, \ldots, w_N\}$. Let the perturbed graph be $G_Z = (V, A_Z)$. If $I(w_i \rightarrow w_j \mid W_{ji}) > 0$, then $i \rightarrow j \in A_Z$.

Proof. We will show that if $i \rightarrow j \notin A_Z$, then $I(w_i \rightarrow w_j \mid W_{ji}) = 0$. Suppose, $i \rightarrow j \notin A_Z$. Let $G'_{Z} = (V', A'_{Z})$ be the perturbed dynamic Bayesian network (DBN) associated with the perturbed graph, $G_Z$. Then, using Theorem 2 for all $t > 0$, $d$-sep$(w_j[t], w_i[t], \cdot \mid \cdot)$ holds in $G_Z$. In other words, this implies $P_{w_j[t] \mid w_i[t]}(w_j[t-1], w_i[t-1]) W_{ji}(t-1)$ will hold true for all $t$ and thus, $I(w_i \rightarrow w_j \mid \cdot W_{ji}) = 0$.

The following example illustrates the intuition behind the presence of active trails and hence, the spurious links in the perturbed graph.

Example 1. Consider a generative graph as shown in Figure 4a. Suppose node 3 is subject to data-corruption and let $u_3$ be its measured data-stream. Denote the measured data-streams at nodes 1 and 2 as $y_1$ and $y_2$, $u_3$ is related to its ideal counterpart $y_3$ via $(4)$. The measured data streams are $\{w_1 = y_1, w_2 = y_2, w_3 = y_3\}$. The perturbed graph $G_Z$, is constructed as defined in definition 13 and is shown in figure 4b. The corresponding perturbed DBN, $G'_{Z}$, is shown for 3 time steps in figure 4c. We will reason out the presence and absence of an edge in $G_Z$ by identifying the presence and absence of active trails in the perturbed DBN.

Consider $1 \rightarrow 3 \in A_Z$. There is a trail $w_1[0] = y_1[0] \rightarrow w_2[1] = y_2[1] \leftarrow y_3[0] \rightarrow y_3[1] \rightarrow w_3[2] \in A'_Z$. Note that the collider $w_2[1]$ is observed. Therefore, the trail is active given $\{w_3[1], w_2[1]\}$.

Take the edge $2 \rightarrow 3 \in A_Z$. There exists a trail $w_2[1] \leftarrow y_3[0] \rightarrow w_3[2]$, in $G_Z$. Note that the node $y_3[0]$ is not a collider and is not observed. Thus, the trail is active given $\{w_3[1], w_2[1]\}$.

The edges $3 \rightarrow 1$ and $2 \rightarrow 1$ are absent in $G_Z$. Ideally, we look for a trail from $w_3[t]$, to $w_1[t] = y_1[t]$ that is active given $\{w_1[t-1], w_2[t-1]\}$ and a trail from $w_2[t]$ to $w_1[t] = y_1[t]$ that is active given $\{w_1[t-1], w_2[t-1]\}$. Note that every trail from $w_3[t]$ to $w_1[t]$ traverses through a node in $w_2[t]$ which is in the observed set and this holds for all $t$. This blocks the information flow along the trail. Therefore, all these trails are inactive.

Remark 3. The results in Theorem 2 and Corollary 1 respectively shows that existence of active trails is the PDBN and non-zero conditional directed information is sufficient to infer the presence of a directed link in the perturbed graph. However, under a mild assumption on the generative and the perturbation model, it can be shown that the respective conditions are also necessary to detect a directed link in the perturbed graph.

Assumption 1. Let the following conditions on the generative and the perturbation model hold:

C1) In the generative model $A$, for all agents $i \in \{1, 2, \ldots, N\}$, and all $j \in P(i)$, there is a number $k_{ij} \geq 1$ such that $y_j[t - k_{ij}]$ is an argument of $f_i$.

C2) For all perturbed nodes $i \in Z$, in the perturbation model $A_{Z}$, there is a number $k_i \geq 1$ such that $y_i[t - k_i]$ as it’s argument.

In addition, let at least one of the following conditions on corruption model hold:

B1) If a node $i \in Z$, then there is a number $k'_i \geq 1$ such that $y_i[t - k'_i]$ is an argument of $f_i$ in $A$.

B2) If a node $i \in Z$, then $y_i[t]$ is an argument of $g_i$ in $A$.

The following theorem asserts that if $i \rightarrow j \in A_Z$ then there exists a corresponding active trail in perturbed DBN.

Theorem 3. Consider a generative graph, $G = (V, A)$, consisting of $N$ nodes. Let $Z = \{v_1, \ldots, v_n\} \subseteq V$ be the set of $n$ perturbed nodes where each perturbation is described by $(4)$. Denote the data-streams as follows: $U_Z := \{u_i\}_{i \in Z}$ and $Y_Z := \{y_j\}_{j \in \bar{Z}}$, where $\bar{Z} = V \setminus Z$. Let the measured data-streams be $W = U_Z \cup Y_Z = \{w_1, w_2, \ldots, w_N\}$. Suppose, the generative model and the perturbation model satisfies the conditions for dynamics that is mentioned in Assumption 1.

If there is a directed edge from $i$ to $j$ in perturbed graph, $G_Z = (V, A_Z)$, then there exists a trail between a node in $w_i(t-1)$ and $w_j[t]$ that is active given $\{w_j[t-1], W_{ji}(t-1)\}$ in $G'_Z$, for some $t > 0$.

Proof. The proof is given in appendix A.
Under the following assumption we can in fact show that
\[ I(w_i \to w_j \mid w_j^{(i)}) > 0 \] is also a necessary condition for
\( i \to j \in A_Z \) as shown in Corollary 2.

**Assumption 2.** We assume that the generative model in \([1]\) and the perturbation model in \([4]\) are such that the corresponding DBN and PDBN are faithful Bayesian networks. Moreover, we consider positive joint distributions for the random processes \( Y \) and \( U \).

**Corollary 2.** Under assumption 2 and dynamics as described in Assumption 7 if \( i \to j \in A_Z \), then \( I(w_i \to w_j \mid w_j^{(i)}) > 0 \).

**Proof.** By theorem 3 if \( i \to j \in A_Z \), then there exists an a trail in PDBN between \( w_i^{(t-1)} \) and \( w_j[t] \) that is active given \( \{w_j^{(t-1)}, W_j^{(t-1)}\} \) in \( G_Z^t \), for some \( t > 0 \). Under faithfulness assumption, this implies \( P_{w_j[t]} \mid w_j^{(t-1)}, w_i^{(t-1)}, W_i^{(t-1)} \neq P_{w_j[t]} \mid w_i^{(t-1)}, W_i^{(t-1)} \). Thus, \( I(w_i \to w_j \mid w_j^{(i)}) > 0 \).

---

**V. ESTIMATION OF DIRECTED INFORMATION**

In \([28]\), consistency results for estimating directed information (DI) between a pair of random processes from data was proposed. However, in this article we extend the methods to determine the directed information between two processes conditioned on a set of other random processes. We provide consistency results of the estimator by showing convergence in almost sure sense (denoted as P-a.s).

**A. Pairwise Estimation of Directed Information**

Here, we present the definition used for directed information estimator proposed in \([28]\). Before that, the following notion of universal probability assignment is needed.

1) **Universal Probability Assignment:** Let \( Q \) denote a sequence probability assignment for a sequence \( x \). That is, the conditional probability mass function (pmf) for \( x[i] \) given \( x^{(i-1)} \) is given by \( Q(x[i] \mid x^{(i-1)}) \). The joint pmf for \( x^{(n)} \) is given by \( Q(x^{(n)}) = Q(x[0])Q(x[1] \mid x[0])Q(x[2] \mid x[1]) \cdots Q(x[n] \mid x[n-1]) \).

**Definition 17** (Universal Probability Assignment). Let \( P \) be the true joint pmf for \( x^{(n)} \). Then, a probability assignment \( Q \) is called as **universal** if the following holds:

\[
\lim_{n \to \infty} \frac{1}{n} \mathbb{E} \left[ \log \frac{P(x^{(n)})}{Q(x^{(n)})} \right] = 0. \tag{11}
\]

Context tree weighting (CTW) algorithm developed by \([29]\) will be used for computing sequential probability assignment.

2) **DI estimation:** Let \( X \) and \( Y \) be jointly stationary and ergodic processes. The directed information from \( X \) to \( Y \) can be expressed in terms of the entropy as follows:

\[
I(X \to Y) = H(Y) - H(Y \mid X) \tag{12}
\]

where \( H(Y) = \mathbb{E}[-\log P(Y)] \) and \( H(Y \mid X) = \mathbb{E}[-\log P(Y \mid X)] \) denotes the entropy of \( Y \) and the causally conditioned entropy \([30]\) respectively.

The directed information rate (DIR) from \( X \) to \( Y \) is defined as:

\[
I_r(X \to Y) := \lim_{n \to \infty} \frac{1}{n} I(X^{(n)} \to Y^{(n)}). \tag{13}
\]

Let \( H_r(Y) := \lim_{n \to \infty} \frac{1}{n} H(Y^{(n)}) \) and let \( H_r(Y \mid X) := \lim_{n \to \infty} \frac{1}{n} H(Y^{(n)} \mid X^{(n)}) \). Thus, if \( H_r(Y) \) and \( H_r(Y \mid X) \) converge, then \( I_r \) is convergent. That is,

\[
I_r = H_r(Y) - H_r(Y \mid X). \tag{14}
\]

In \([28]\), the following DIR estimator was defined:

\[
\hat{I}(X^{(n)} \to Y^{(n)}) = \frac{1}{n} \sum_{i=1}^{n} Q(y[i] \mid X^{(i-1)}, Y^{(i-1)}) \cdot \frac{1}{Q(y[i] \mid X^{(i-1)}, Y^{(i-1)})} - \frac{1}{n} \sum_{i=1}^{n} Q(y[i] \mid X^{(i-1)}, Y^{(i-1)}) \cdot \frac{1}{Q(y[i] \mid X^{(i-1)}, Y^{(i-1)})} \tag{15}
\]

We will extend the above to define conditional directed information as described in the following subsection.

**B. Estimation of Conditional Directed Information**

Let \( X, Y, Z \) be jointly stationary and ergodic processes. The conditional directed information from \( X \) to \( Y \) conditioned on \( Z \) can be expressed in terms of the entropy as follows:

\[
I(X \to Y \mid Z) = H(Y) - H(Y \mid X, Z). \tag{16}
\]

The causally conditional directed information rate (DIR) from \( X \) to \( Y \) now is defined as:

\[
I_r(X \to Y \mid Z) := \lim_{n \to \infty} \frac{1}{n} I(X^{(n)} \to Y^{(n)} \mid Z^{(n)}). \tag{17}
\]

Let \( H_r(Y \mid X, Z) := \lim_{n \to \infty} \frac{1}{n} H(Y^{(n)} \mid X^{(n)}, Z^{(n)}) \). Thus, if \( H_r(Y \mid X, Z) \) and \( H_r(Y \mid X, Z) \) converge, then \( I_r \) is convergent. That is,

\[
I_r = H_r(Y \mid Z) - H_r(Y \mid X, Z). \tag{18}
\]

The conditional directed information estimator \( \hat{I}(X^{(n)} \to Y^{(n)} \mid Z^{(n)}) \) is defined as under:

\[
\hat{I}(X^{(n)} \to Y^{(n)} \mid Z^{(n)}) = \frac{1}{n} \sum_{i=1}^{n} Q(y[i] \mid X^{(i-1)}, Y^{(i-1)}, Z^{(i-1)}) \cdot \frac{1}{Q(y[i] \mid X^{(i-1)}, Y^{(i-1)}, Z^{(i-1)})} - \frac{1}{n} \sum_{i=1}^{n} Q(y[i] \mid X^{(i-1)}, Y^{(i-1)}, Z^{(i-1)}) \cdot \frac{1}{Q(y[i] \mid X^{(i-1)}, Y^{(i-1)}, Z^{(i-1)})} \tag{19}
\]

The following theorem establishes the consistency result in estimating conditional DIR as defined in \([19]\). The proof is given in appendix \([8]\).

**Theorem 4.** Let \( Q \) be the probability assignment in the CTW algorithm. Suppose, \( X, Y, Z \) are jointly stationary irreducible aperiodic finite-alphabet Markov processes whose order is
Fig. 5: This figure shows how unreliable measurements at node 3 can result in spuriously inferring a direct dynamic influence of node 1 on the third and a spurious influence of node 2 on node 3.

bounded by the prescribed tree depth of the CTW algorithm. Then,

$$\lim_{n \to \infty} \hat{I}(X^{(n)} \to Y^{(n)} \parallel Z^{(n)}) = I_r(X \to Y \parallel Z) \quad P.a.s.$$ (20)

C. Simulation Results

To verify the predictions of Theorem 2, we first performed a simulation on a network consisting of 3 nodes with a single node being perturbed and on a network consisting of 6 nodes, of which 2 are corrupt. We estimate the directed information rates (DIR), which are DI estimates that are averaged along the sequence length till the horizon. We used the estimator described in [19] to compute DIR. For both the networks, the horizon length is chosen as $10^7$. The DIR estimates were then averaged over 50 trials.

1) Single node Perturbation: Consider a network consisting of 2 nodes with a common child as shown in Fig. 5(a). The dynamic interactions in the true generative model are as follows:

$$y_1[t] = e_1[t],$$
$$y_2[t] = y_1[t - 1] + y_3[t - 1] + e_2[t],$$
$$y_3[t] = e_3[t]$$

where $e_1[t] \sim \text{Bernoulli}(0.7)$, $e_2[t] \sim \text{Bernoulli}(0.4)$ and $e_3[t] \sim \text{Bernoulli}(0.6)$ and ‘+’ is logical ‘OR’ operation. Each of $y_1[t], y_2[t]$ and $y_3[t]$ has a finite alphabet $\{0, 1\}$.

The perturbation considered here is the time-origin uncertainty at node 3. The corruption model takes the form:

$$u_3[t] = \begin{cases} y_3[t] - 2, & \text{with probability 0.5} \\ y_3[t], & \text{with probability 0.5}. \end{cases}$$

The perturbed graph predicted by Theorem 2 is shown in Fig. 5(b). The DIR estimates from ideal ($Y$) and unreliable measurements ($U$) are shown in Fig. 5(c). We observe non-zero DIR estimates and add edges to $G_Z$ respectively. In particular, note the substantial rise in $I(u_1 \to u_3 \parallel u_2)$ and in $I(u_2 \to u_3 \parallel u_1)$. This indicates the presence of spurious links $1 \to 3$ and $2 \to 3$ in the inferred perturbed graph.

2) Multiple Perturbation: Consider a network of 6 nodes as shown in Fig. 6(a). The dynamic interactions in the true generative model are as follows:

$$y_1[t] = e_1[t],$$
$$y_2[t] = y_1[t - 1] + e_2[t],$$
$$y_3[t] = y_2[t - 1] + e_3[t],$$
$$y_4[t] = e_4[t],$$
$$y_5[t] = (y_2[t - 1] + y_4[t - 1] \cdot e_5[t],$$
$$y_6[t] = y_5[t - 1] + e_6[t]$$

where $e_1[t] \sim \text{Bernoulli}(0.55), e_2[t] \sim \text{Bernoulli}(0.5), e_3[t] \sim \text{Bernoulli}(0.2), e_4[t] \sim \text{Bernoulli}(0.4), e_5[t] \sim \text{Bernoulli}(0.3)$ and ‘+’ is logical `OR` operation while ‘$\cdot$’ is logical `AND` operation. Each of $y_1[t], y_2[t], ..., y_6[t]$ has a finite alphabet $\{0, 1\}$. The perturbations considered here are time-origin uncertainties at nodes 2 and 5. The corruption models take the form:

$$u_2[t] = \begin{cases} y_2[t] - 2, & \text{with probability 0.5} \\ y_2[t], & \text{with probability 0.5}. \end{cases}$$
and

\[ u_5[t] = \begin{cases} 
  y_5[t] - 2, & \text{with probability } 0.5 \\
  y_5[t], & \text{with probability } 0.5.
\end{cases} \]

The perturbed graph predicted by Theorem 2 is shown in figure 6b. The DIR estimates from ideal (Y) and unreliable measurements (U) are shown in figures 6c and 7. We observe non-zero DIR estimates and add edges to \( G_Z \) respectively. For clarity of visualization, only non-zero DIR estimates that would be predicted by Theorem 2 are shown.

VI. CONCLUSION

We studied the problem of inferring directed graphs for a large class of networks that admit non-linear and strictly causal interactions between several agents. We provided necessary and sufficient conditions to determine the directed structure from corrupt data-streams. Doing so, we particularly established that inferring causal structure from corrupt data-streams results in spurious edges and we precisely characterized the directionality of such spurious edges. Finally, we provided convergence results for the estimation of conditional directed information that was used to determine the directed structure. Simulation results were provided to verify the theoretical predictions.

Future Work

Currently, the emphasis was on characterizing the effects of data corruption on network inference and determining how spurious probabilistic relations are introduced. Future work will focus on quantifying the amount of data that is needed to detect network inter-relationships using directed information. Moreover, the problem of removing spurious edges in the network reconstructed from corrupt data streams will be addressed in future. This will play an integral role preceding system identification on networked systems.

APPENDIX

Proof for Theorem 3

Suppose \( i \to j \) is in \( A_Z \). Then there is a trail, \( trl_G \), described by \( i = v_1 - v_2 - \cdots - v_m = j \) in \( G \) satisfying conditions in Definition 15. We will first construct a trail in the perturbed DBN, \( G_Z \), from a node in \( w_i[t^{(1)}] \) to \( w_j[t] \) for some \( t > 0 \). We can construct a trail in \( G_Z \) as follows: for all \( l \in \{1, 2, \ldots, k-1 \} \), set \( t_l = t_{l+1} + k_{v_{l+1}, v_l} \) if \( v_l \to v_{l+1} \) holds in \( trl_G \). Otherwise, set \( t_l = t_{l+1} + k_{v_{l+1}, v_l} \) if \( v_l \leftarrow v_{l+1} \) holds in \( trl_G \). Such a construction is feasible because by condition C1, numbers \( k_{v_{l+1}, v_l} \) and \( k_{v_{l+1}, v_l} \) follows for all \( l \in \{1, 2, \ldots, k-1 \} \) and at all times. Thus, we have a trail \( y_1[t_1] - y_2[t_2] - y_3[t_3] - \cdots - y_{k_{l-1}}[t_{k-1}] - y_{k_l}[t_k] \). For all \( m \in Z \), \( t_{m} \in Z \), \( v_{m} = Z \), there exists a number \( k_m \geq 0 \). Let \( t > \max\{t_1, \ldots, t_{k-1} \} \), and for all \( m \in \{1, 2, \ldots, k \} \) if \( v_{m} = Z \), \( t > t_{m} + k_{m} \) also hold. Depending on whether \( i \) or \( j \) is a perturbed node, we have four cases on either end of the above trail.

A) Consider the case \( i, j \in Z \). As \( i \in Z \), using condition C2, \( u_i[t_1 + k_i] \leftarrow y_i[t_1] \) holds true. Choose \( t \) sufficiently large so that \( t > t_1 + k_i \) also holds. As \( j \in Z \), using C2, \( t \) can be sufficiently large so that we have \( y_j[t_k] \to u_j[t] \) where \( t = t_k + k_j \). As \( k_j \geq 1 \), if \( B_1 \) holds, then we can choose \( t \) sufficiently large such that at the end of the trail we take 5 steps from \( y_j[t_k] \) to \( u_j[t] \) such that the tail is of the form \( y_j[t_k] \to y_j[t_k + k_j] \to \cdots \to y_j[t_k + sk_j] \to u_j[t] \) with \( t = t_k + sk_j + k_j \). Thus, the constructed trail in \( G_Z \) is either \( u_i[t_1 + k_i] = u_i[t_1 + k_i] \leftarrow y_i[t_1] - y_2[t_2] - y_3[t_3] - \cdots - y_{k_{l-1}}[t_{k-1}] - y_{k_l}[t_k] \to u_j[t] = w_j[t] \), or \( u_i[t_1 + k_i] = u_i[t_1 + k_i] \leftarrow y_i[t_1] - y_2[t_2] - y_3[t_3] - \cdots - y_{k_{l-1}}[t_{k-1}] - y_{k_l}[t_k] \to y_j[t_k + sk_j] \to u_j[t] = w_j[t] \) with \( t > \max\{t_1 + k_1, t_2, \ldots, t_{k-1}, t_k + sk_j \} \), and for all \( m \in \{1, 2, \ldots, k \} \) if \( v_{m} \in Z \), \( t > t_{m} + k_{m} \).

B) Consider the case \( i \in Z \) but \( j \notin Z \). Choose \( t \) as \( t_k \). As \( i \in Z \), using condition C2, \( u_i[t_1 + k_i] \leftarrow y_i[t_1] \) holds true. Choose \( t \) sufficiently large so that \( t > t_1 + k_i \) also holds. Thus, we have constructed a trail in \( G_Z \) which is of the form: \( w_i[t_1 + k_i] = u_i[t_1 + k_i] \leftarrow y_i[t_1] - y_2[t_2] - y_3[t_3] - \cdots - y_{k_{l-1}}[t_{k-1}] - y_{k_l}[t_k] \to u_j[t] = w_j[t] \) with \( t > \max\{t_1 + k_1, t_2, \ldots, t_{k-1}, t_k \} \), and for all \( m \in \{1, 2, \ldots, k \} \) if \( v_{m} \in Z \), \( t > t_{m} + k_{m} \).

C) Consider the case \( i \notin Z \) but \( j \in Z \). Following arguments presented in case (A) we conclude that the constructed trail of form \( w_i[t_1] = y_i[t_1] - y_2[t_2] - y_3[t_3] - \cdots - y_{k_{l-1}}[t_{k-1}] - y_{k_l}[t_k] \to u_j[t] = w_j[t] \), or of form \( w_i[t_1] = y_i[t_1] - y_2[t_2] - y_3[t_3] - \cdots - y_{k_{l-1}}[t_{k-1}] - y_{k_l}[t_k] \to u_j[t] = w_j[t] \).
Recall the expression for the conditional DI estimator from \cite{19}:

$$
\hat{I}(X^{(n)} \rightarrow Y^{(n)} \parallel Z^{(n)}) = \frac{1}{n} \sum_{i=1}^{n} Q(y[i] | X^{(i-1)}, Y^{(i-1)}, Z^{(i-1)}) - \log \frac{1}{Q(y[i] | X^{(i-1)}, Y^{(i-1)}, Z^{(i-1)})}
$$

We will now argue that in each of the cases above, the constructed trail is active given \( \theta := \{ w_{j}^{(t-1)}, \bar{W}_{ji}^{(t-1)} \} \).

Sub-trails with colliders: For all the trails in \( G_{Z}^{2} \) constructed under various cases above consider a sub-trail of the form \( y_{vm-1}[t_{m-1}] \rightarrow y_{vm}[t_{m}] \leftarrow y_{vm+1}[t_{m+1}] \). Clearly, \( v_{m} \) cannot be either \( i \) or \( j \). If \( v_{m} \notin \{ i, j \} \) then as \( t_{m} < t \), we have \( y_{vm}[t_{m}] \in \bar{w}_{ij}^{(t-1)} \) and thus the sub-trail is active. If \( v_{m} \in \{ i, j \} \) then the corrupted version of \( y_{vm}[t_{m}] = w_{vm}[t_{m} + k_{vm}] \) and as \( t_{m} < t \), we have \( y_{vm}[t_{m}] \notin \bar{w}_{ij}^{(t-1)} \). Thus the collider \( y_{vm}[t_{m}] \) has a descendant \( w_{vm}[t_{m} + k_{vm}] \in \theta \). Thus the sub-trail remains active. Thus no collider can deactivate the trails in \( G_{Z}^{2} \).

Sub-trails with no colliders: Consider any node \( y_{vm}[t_{m}] \) which is not a collider. Note that in the trails for the cases (A), (B), (C), and (D), \( y_{i} \) and \( y_{j} \) can only appear as an intermediate node only if they are corrupted. In such cases, neither \( y_{i}[t_{i}] \) nor \( y_{j}[t_{j}] \) belong to \( \theta \). Thus, if \( y_{i} \) or \( y_{j} \) are intermediate nodes, they cannot deactivate the trails given \( \theta \). Consider an intermediate node \( v_{m} \notin \{ i, j \} \). From Definition \ref{def:5}, \( v_{m} \) is corrupted. Thus \( y_{vm}[t_{m}] \neq w_{vm}[t_{m}] \) and \( y_{vm}[t_{m}] \) cannot deactivate the trail as \( y_{vm}[t_{m}] \notin \theta \). \( \square \)

APPENDIX B

PROOF FOR THEOREM \[\ref{thm:3}\]

To prove the theorem, we require two results from \cite{23}. The following lemma shows that with sufficiently large data, the conditional probability assignment by CTW converges to the true probability assignment for a Markov process.

**Lemma 2.** Let \( Q \) be the probability assignment described by CTW. Let \( X \) be a stationary and finite alphabet Markov process with finite Markov order which is bounded by the prescribed tree depth of CTW algorithm. Let \( P \) be the true probability for \( x \). Then,

$$
\lim_{n \to \infty} Q(x[n] | x^{(n-1)}) - P(x[n] | x^{(n-1)}) = 0 \quad \text{P-a.s.}
$$

(21)

Next, we will later use the following proposition which is a rephrased result from \cite{23}.

**Proposition 1.** Let \( Q \) be the probability assignment in the CTW algorithm. Suppose, \( X, Y \) are jointly stationary irreducible aperiodic finite-alphabet Markov processes whose order is bounded by the prescribed tree depth of the CTW algorithm. Let \( \hat{H}(Y^{(n)} \parallel X^{(n)}) = \frac{1}{n} \sum_{i=1}^{n} \sum_{y[i]} Q(y[i] | X^{(i-1)}, Y^{(i-1)}) \cdot \log \frac{1}{Q(y[i]|X^{(i-1)},Y^{(i-1)})} \). Then,

$$
\lim_{n \to \infty} \hat{H}(Y^{(n)} \parallel X^{(n)}) - H_{r}(Y \parallel X) = 0 \quad \text{P-a.s.}
$$

(22)

By Lemma \ref{lem:2} the CTW probabilities \( Q(y[i] | X^{(i-1)}, Y^{(i-1)}, Z^{(i-1)}) \) converges to true probabilities \( P(y[i] | X^{(i-1)}, Y^{(i-1)}, Z^{(i-1)}) \) almost surely. Therefore,

$$
\lim_{i \to \infty} \beta_i = 0 \quad \text{P-a.s.}
$$

(27)

Hence, by Cesaro mean \cite{31} we have:

$$
\lim_{n \to \infty} F_n = \lim_{n \to \infty} \frac{1}{n} \beta_i = 0 \quad \text{P-a.s.} \quad \square
$$

(28)
