This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2021.3093301, IEEE Transactions on Automatic Control

Effects of Data Corruption on Network Identification using Directed Information

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 by applying control inputs. Here, causal structure identification via passive means is to be accomplished. With advancements in data processing technology, and 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 plagued by the effects of noise [11], asynchronous sensor clocks [12] and packet drops [13]. While considering the problem of identifying causal influences of a large network, it is fundamental to rigorously study such uncertainties and address detrimental effects of data corruption on network identification.

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 graph 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 directed interaction structure can be accurately recovered. However, 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. Recently, [19], [20], [21] defined and used *information transfer* to determine underlying causal interactions in dynamical systems. However, it is assumed that the data-streams are ideal with no distortions. [22] and [23] identify causal dependencies in network of LTI systems driven by unknown intrinsic noise inputs. However, in this article we consider nonlinear dependencies and study the problem of network reconstruction from corrupt data-streams.

The authors in [24], [25] use dynamical structure functions (DSF) for network recosntruction [26] 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 directed information-based network reconstruction from corrupt data-streams.

Despite its significance, little is known on the effects of measurement uncertainties on network identification. Assuming that the network structure is known, *errors-in-variables* framework for system identification with additive sensor noise is studied in [27], [28]. However, in this work we do not assume that the interaction structure is known. Recently in [29], 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 [30] focusing on networks with LTI interactions, authors provided characterization of the extent of spurious links that can appear due to data-corruption. However, the analysis is restricted to

The authors are with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455, USA.subra148@umn.edu, alampers@umn.edu, murtis@umn.edu

Work supported in part by NSF CMMI 1727096.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2021.3093301, IEEE Transactions on Automatic Control

LTI systems and determine the undirected structure of the networked system, and not to deduce the directions.

B. Our Contribution

In this article, we focus on identifying the Boolean structure of the network using non-invasive or passive means from corrupt data-streams. We consider networks with nonlinear and strictly causal dynamical interactions. Moreover, the endogenous noise exciting the system are not measured and hence, we assume a *blind* approach [31]. General analysis of network structure employing passive and and blind means with nonlinearities is challenging. We make an assumption that the endogenous noise affecting one node is independent of another and thus we deal with *target specified* network reconstruction.

We provide necessary and sufficient conditions that delineates the effects of data corruption on the directed network structure inferred using directed information. We present a tight characterization for the spurious links that arise due to corruption of data-streams by determining their location and orientation. Often, the knowledge of influence structure is required a priori to perform system identification in networked systems [15], [28], [32]. Thus, our results serve as a necessary first step in understanding what part of network reconstruction can be trusted to facilitate accurate system identification.

In [33], 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 a 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 [33] wherein only a proof sketch was provided. In addition, we establish convergence results for the estimator that we use to determine conditional directed information.

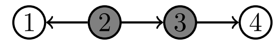
C. Paper Organization

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 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.

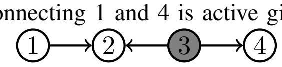
II. Preliminaries

A. Notations

Upper case letter Y denotes a random variable (r.v) while lower case letter y denotes a realization of r.v Y.
Caligraphic letter $\mathcal{Y}$ denotes the alphabet of r.v Y.
$y[\cdot]$ denotes a sequence and $y^{(t)}$ denotes the sequence $y[0], y[1], \dots y[t]$.
P_X represents the probability mass function (PMF) of a discrete random variable X or denotes the probability density function (PDF) of a continuous random variable X.



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



(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.

$\mathbb{E}[\cdot]$ denotes the expectation operator.
A *directed graph* G is denoted by 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 $i, j \in V$.
$i \to j$ indicates an edge or link from node i to node j in G.
$i - j$ denotes one of $i \to j$ or $j \to i$.

B. Graph Theory Definitions

This subsection gives a list of standard terminology from graphical models. It can be used as a reference for following sections. For further details, see [34].

Definition 1 (Children and Parents)**.** Given a directed graph $G = (V, A)$ and a node $j \in V$, the *children* of j are defined as $\mathcal{C}(j) := \{i | j \to i \in A\}$ and the *parents* of j as $\mathcal{P}(j) := \{i | i \to j \in A\}$.

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

Definition 3 (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 $v_i \to W_1 \to \cdots \to W_{k-2} \to v_j$ holds in G.

Definition 4 (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 5 (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 $v_i \to v_k \leftarrow v_j$ holds.

Definition 6 (Active Trail)**.** In a directed graph G, a trail $v_1 - v_2 - \cdots - v_n$ is *active* given a set of nodes Z if one of the following statements holds for $m \in \{2, \dots, n-1\}$ and every triple $v_{m-1} - v_m - v_{m+1}$ 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.

See Figure 1 for an illustration.

Definition 7 (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 8 (Directed Cycle)**.** A *directed cycle* from a node v_i to v_i in a directed graph, G, has the form $v_i \to W_1 \to \cdots \to W_k \to v_i$ for some set of nodes $\{W_n\}_{n=1}^{k}$ in G.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2021.3093301, IEEE Transactions on Automatic Control

Definition 9 (Directed Acyclic Graph)**.** A directed graph with no directed cycles is called a *directed acyclic graph* (DAG).

Definition 10 (Bayesian Network)**.** Suppose $G = (V, A)$ is a DAG whose N nodes represent random variables $a_1, \dots, a_N$. 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 11 (Faithful Bayesian network)**.** Suppose $G = (V, A)$ is a DAG whose N nodes represent random variables $a_1, \dots, a_N$. 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 $Y_i[\cdot]$ and a sequence $E_i[\cdot]$. We consider E_i such that P_{E_i} exists if $E_i[t]$ belongs to a continuous alphabet. The process E_i is considered to be target-specific, that is, E_i is 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 $\{Y_1, \dots, Y_N\}$ with a parent set $\mathcal{P}'(i)$ defined for $i = 1, \dots, N$. We consider strictly causal nonlinear dynamical relations. The generative model takes the form:

$$Y_i[t] = f_i\left(Y_i^{(t-1)}, \bigcup_{j \in \mathcal{P}'(i)} Y_j^{(t-1)}, E_i[t]\right), \tag{1}$$

where f_i's can be any nonlinear function such that P_{Y_i} is well defined if $Y_i[t]$ takes values in a continuous alphabet. f_i is a multivariate function that maps the past measurements of parent nodes of i, $\{Y_j^{(t-1)} : j \in \mathcal{P}'(i)\}$, previous measurements of the node i in $Y_i^{(t-1)}$, and the present realization of process noise, $E_i[t]$, to the present measurement of agent i, $Y_i[t]$.

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

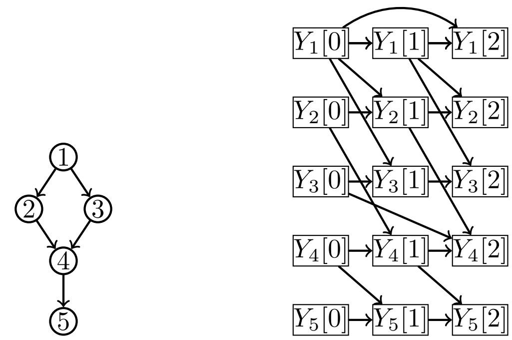
$$\begin{aligned}
Y_1[t] &= Y_1[t-1]Y_1[t-2] + E_1[t], \\
Y_2[t] &= Y_1[t-1] \cdot Y_2[t-1] + E_2[t], \\
Y_3[t] &= (Y_1[t-1] + Y_3[t-1]) \cdot E_3[t], \\
Y_4[t] &= Y_2[t-1]^2 + Y_3[t-2] + Y_4[t-1] + E_4[t], \\
Y_5[t] &= Y_5[t-1] \cdot Y_4[t-1] + E_5[t].
\end{aligned} \tag{2}$$

We remark that for any time instant t, the parent set $\mathcal{P}'(i)$ 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 $\mathcal{P}'(i)$ 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 describes the relationships between the stochastic processes in Y.



(a) Generative Graph G (b) DBN G' for 3 time slices

Fig. 2: This figure shows (a) generative graph, (b) its associated DBN for 3 time slices.

The generative graph associated with the example described in (2) is given by Fig. 2(a). When the time variable is unraveled we obtain the Dynamic Bayesian Network.

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 $S_{ij}[t] = \{t' : Y_j[t'] \in Y_j^{(t-1)}$ as an argument of f_i in expression of $Y_i[t]$ in (1)$\}$ for all $j \in \mathcal{P}'(i) \cup \{i\}$ and for all t. Consider the graph

$$G' = (V', A') \text{ where } V' = \left(\bigcup_{\substack{i \in V \\ t \in \{0,1,\dots T\}}} Y_i[t]\right) \text{ and}$$

$$A' = \bigcup_{\substack{i \in V \\ t \in \{0,1,\dots T\}}} \left(\bigcup_{j \in \mathcal{P}'(i) \cup \{i\}} \left(\bigcup_{k \in S_{ij}[t]} Y_j[k] \to Y_i[t]\right)\right)$$

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

$$P_{Y^{(T)}} = P_{Y_1[0]} \dots P_{Y_N[0]} \prod_{t=1}^{T} \prod_{i=1}^{N} P_{Y_i[t] \mid \mathcal{P}(Y_i[t])}, \tag{3}$$

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 $\{Y_i[t] : t = 0, 1, 2, \dots, T, \ i = 1, 2, \dots, N\}$ and is considered the *Dynamic Bayesian Network* for $\{Y_i : i = 1, 2, \dots, N\}$(see [34]). 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 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:

$$U_i[t] = g_i(Y_i^{(t)}, U_i^{(t-1)}, \zeta_i[t]), \tag{4}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2021.3093301, IEEE Transactions on Automatic Control

where g_i can be any multivariate function that maps the present and past values of uncorrupted data-streams in $Y_i^{(t)}$, the present value of an independent random process $\zeta_i[t]$, and past corrupt measurements in $U_i^{(t-1)}$ to the current corrupt measurement, $U_i[t]$, such that P_{U_i} exists if $U_i[t]$ takes values in a continuous alphabet. $\zeta_i[t]$ is such that P_{ζ_i} exists if $\zeta_i[t]$ belongs to a continuous alphabet, and is independent of E_i, Y_i for all $i \in 1, \cdots, N$, and $\zeta_j[t]$ for $i \neq j$. We highlight a few important perturbation models that are practically relevant.

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:

$$U_i[t] = Y_i[t + \zeta_i[t]]. \tag{5}$$

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:

$$U_i[t] = (L_i * Y_i)[t] + \zeta_i[t], \tag{6}$$

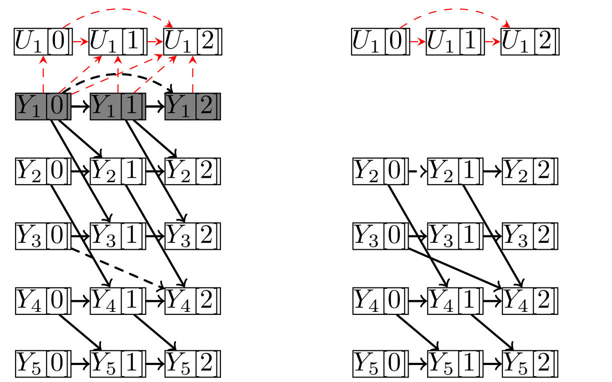
where L_i is a stable causal linear time invariant filter.

Packet Drops: Consider an IID Bernoulli process $\zeta_i[t]$ described by success probability, p_i. The measurement $U_i[t]$ corresponding to an ideal data-point $Y_i[t]$ packet reception at time t can be stochastically modeled as:

$$U_i[t] = \zeta_i[t] Y_i[t] + (1 - \zeta_i[t]) U_i[t-1]. \tag{7}$$

B. Perturbed Dynamic Bayesian Network

Here, we provide a discussion on how the DBN associated with the measured data-streams gets altered when the data-streams are subject to corruption. Note that the measured data-streams only includes the corrupted time-series for the nodes that are corrupted and the data-streams for those nodes that are not corrupted. The uncorrupted time-series for the corrupted nodes are not measured and are hence, not observed. When the time variable is unraveled we obtain the perturbed DBN (PDBN) that depicts the causal dependencies between the true data-streams for the network, and in addition shows the dependencies between the uncorrupted measurements and the corrupted values for the corrupted nodes, and between the corrupted measurements for each corrupted node. Thus, the perturbed DBN is the union of DBN when there is no data corruption and the causal dependencies for the time-series associated with the corrupted node. Figure 3(a) shows an example of a PDBN corresponding to the generative graph in Fig. 2(a) for three time slices. Here, node 1 data-streams are corrupt following a noisy filtering model described in (6).



(a) PDBN for 3 time slices (b) Measured Data-Stream

Fig. 3: Figure (a) shows Perturbed DBN G'_Z for 3 time slices when node 1 is corrupt. Node 1 ideal stream, Y_1, is shaded because it is not measured. Figure (b) only shows the causal relations between the measured data-streams without hidden Y_1. There are no direct causal connections between time-series U_1 and other time series nodes as there are no direct dynamic influences between them, and therefore, are not shown

Consider a generative graph $G = (V, A)$. 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 $G' = (V', A')$ 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 $Y_{\bar{Z}} = \{Y_j\}_{j \in \bar{Z}}$ where $\bar{Z} = V \setminus Z$. Due to corruption only U_Z and $Y_{\bar{Z}}$ are measured and observed. Denote the measured data-streams by $\mathcal{W} = U_Z \cup Y_{\bar{Z}}$. For all $j \in Z$, let $SU_j[t] = \{t' : U_j[t'] \in U_j^{(t-1)}$ is an argument of g_i in the expression of $U_j[t]$ in (4)$\}$ and let $SY_j[t] = \{t' : Y_j[t'] \in Y_j^{(t)}$ is an argument of g_i in the expression of $U_j[t]$ in (4)$\}$ for all t. Consider the graph $G'_Z = (V'_Z, A'_Z)$ where $V'_Z = V' \cup \left(\bigcup_{\substack{k \in Z \\ t \in \{0,1,\ldots T\}}} U_k[t] \right)$ and $A'_Z = A' \cup \left(\bigcup_{\substack{k \in Z \\ i \in SY_k[t]}} Y_k[i] \to U_k[t] \right) \cup \left(\bigcup_{\substack{k \in Z \\ i \in SU_k[t]}} U_k[i] \to U_k[t] \right)$ for all $t \in \{0, 1, 2, \ldots, T\}$. Note that the vertex set V'_Z consists of all measurements given by the set $\mathcal{W}$, and the uncorrupted versions Y_k of the corrupted versions U_k for $k \in Z$.

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

$$P_R = \left(\prod_{i \in V} P_{U_i[0]} \right) \cdot \left(\prod_{j \in Z} P_{Y_j[0]} \right) \cdot \left(\prod_{t=1}^{T} \prod_{i=1}^{N} P_{U_i[t] | \mathcal{P}(U_i[t])} \right) \cdot \left(\prod_{t=1}^{T} \prod_{j=1}^{N} P_{Y_j[t] | \mathcal{P}(Y_j[t])} \right), \tag{8}$$

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 associated with $U_Z \cup Y$.

Remark 1. Above, the discrete time sequences were considered to have finite horizon only to illustrate DBN and PDBN. The main results in this article characterizing the structure inference from corrupt data-streams holds for any horizon.

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 [18] defined and applied causally conditioned 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 [18] is:

Definition 12 (Causally Conditioned Directed Information)**.** The causally conditioned directed information (DI) from data-stream Y_j to Y_i is given by:

$$I(Y_j \to Y_i \parallel Y_{\bar{i}\bar{j}}) = \mathbb{E}\left[\log \frac{P_{Y_i \| Y_j, Y_{\bar{i}\bar{j}}}}{P_{Y_i \| Y_{\bar{i}\bar{j}}}}\right], \tag{9}$$

where $P_{Y_i\|Y_j,Y_{\bar{i}\bar{j}}} = \prod_{t=1}^{T} P_{Y_i[t]|Y_i^{(t-1)},Y_j^{(t-1)},Y_{\bar{i}\bar{j}}^{(t-1)}}$, $P_{Y_i\|Y_{\bar{i}\bar{j}}} = \prod_{t=1}^{T} P_{Y_i[t]|Y_i^{(t-1)},Y_{\bar{i}\bar{j}}^{(t-1)}}$ and $Y_{\bar{i}\bar{j}} = Y \setminus \{Y_i, Y_j\}$.

For the rest of the article, we drop the word 'causally' for convenience. Note that the conditional DI from Y_j to Y_i is positive if and only if the history of Y_j gives information about $Y_i[t]$ that could not have been obtained from Y_i's own history and the other signals from the network. So, if there is no directed edge from $j \to i$ in G, then we have $I(Y_j \to Y_i \parallel Y_{\bar{i}\bar{j}}) = 0$.

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 $I(Y_j \to Y_i \parallel Y_{\bar{i}\bar{j}}) > 0$.*

Remark 2. In [18], the authors assume positive distribution for the random processes in Y. The distribution is positive if $P_Y > 0$ for all joint sequences Y. This assumption avoids pathologies that arise in deterministic systems. For example, if $Y_1[t]$ are IID random variables, and $Y_2[t] = Y_1[t-1]$ and $Y_3[t] = Y_2[t-1]$, then $I(Y_2 \to Y_3 \parallel Y_1) = 0$, even though Y_3 depends on Y_2 The positivity assumption ensures that the computed expectations are non-negative and hence avoids false negatives for true edges in the generative graph and are therefore detected.

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

In this section we characterize the spurious edges that arise when using conditional DI to estimate network structure. In particular, we will show that under appropriate hypotheses, the estimated edges precisely correspond to edges in the *perturbed graph*, defined next.

Definition 13 (Perturbed Graph)**.** Let $G = (V, A)$ be a generative graph. Suppose $Z \subset V$ is the set of perturbed nodes with each perturbation model admitting a description provided by (4). The perturbed graph, $G_Z = (V, A_Z)$, is a directed graph where there is an edge $i \to j \in A_Z$ if and only if there is a trail, $trl_G : i = v_1 - v_2 - \cdots - v_{k-1} - v_k = j$ in G such that the following conditions hold:

P1) If $j \notin Z$, then $v_{k-1} \to j \in A$.
P2) For $m \in \{2, 3, \ldots, k-1\}$, if $v_{m-1} \to v_m \leftarrow v_{m+1}$, and $v_m \notin Z$, then $v_{m+1} \in Z$.
P3) If v_m is a node such that $v_{m-1} - v_m - v_{m+1}$ is a sub-path of the path $v_1 - \ldots - v_k$ and v_m is not a collider, then $v_m \in Z$.

Remark 3. Note that the existence of a trail that does not meet the 'if' conditions in P1), P2) and P3) guarantees that $i \to j \in A_Z$. For example, if $i \to j \in A$ then $i \to j \in A_Z$. Indeed, if $j \notin Z$ then $i \to j \in A_Z$ 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 $i \to j$. So, $i \to j \in A_Z$.

Definition 14 (Spurious Links)**.** Let $G = (V, A)$ be a generative graph, $Z \subset V$ be the set of perturbed nodes and $G_Z = (V, A_Z)$ be the perturbed graph. Spurious links are those links $i \to j \in A_Z$ that do not belong to A.

The conditions in P1-P3 specifies a path characterization based on the location of corrupt nodes. This defines the paths through which spurious probabilistic relations are introduced due to data corruption. These probabilistic relations are captured by trails in PDBN that become active due to data-corruption. The following theorem precisely gives a relationship between the active trails in PDBN and the directed edges in the perturbed graph. The proof is given in appendix I.

Theorem 2. *Consider a generative graph, $G = (V, A)$, consisting of N nodes. Let $Z = \{v_1, \ldots, v_n\} \subset V$ 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 $Y_{\bar{Z}} := \{Y_j\}_{j \in \bar{Z}}$ where $\bar{Z} = V \setminus Z$. Let the measured data-streams be $\mathcal{W} = U_Z \cup Y_{\bar{Z}} = \{W_1, W_2, \ldots, W_N\}$. Let the perturbed graph be $G_Z = (V, A_Z)$ and its associated PDBN be $G'_Z = (V'_Z, A'_Z)$. If $i \to j \notin A_Z$, then d-sep$(W_j[t], W_i^{(t-1)} \mid \{W_i^{(t-1)}, \mathcal{W}_{\bar{j}\bar{i}}^{(t-1)}\})$ holds in G'_Z for all $t > 0$.*

We will now show that if conditional directed information, $I(W_i \to W_j \parallel \mathcal{W}_{\bar{j}\bar{i}})$, are computed using corrupted data-streams, and were applied for causal structure inference, then we infer the perturbed graph that contains spurious links.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2021.3093301, IEEE Transactions on Automatic Control

(a) True Graph, G. (b) Perturbed Graph, G_Z.

Fig. 4: This figure illustrates the intuition behind spurious links in Example 1. Figure 4(b) shows the perturbed graph inferred. Spurious links are shown in red and the true edges are depicted in black.

Corollary 1. *Consider a generative graph, $G = (V, A)$, consisting of N nodes. Let $Z = \{v_1, \ldots, v_n\} \subset V$ 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 $Y_{\bar{Z}} := \{Y_j\}_{j \in \bar{Z}}$ where $\bar{Z} = V \setminus Z$. Let the measured data-streams be $\mathcal{W} = U_Z \cup Y_{\bar{Z}} = \{W_1, W_2, \ldots, W_N\}$. Let the perturbed graph be $G_Z = (V, A_Z)$. If $I(W_i \to W_j \parallel \mathcal{W}_{\overline{ji}}) > 0$, then $i \to j \in A_Z$.*

Proof. We will show that if $i \to j \notin A_Z$, then $I(W_i \to W_j \parallel \mathcal{W}_{\overline{ji}}) = 0$. Suppose, $i \to j \notin A_Z$. 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$(W_j[t], W_i^{(t-1)} \mid W_i^{(t-1)}, \mathcal{W}_{\overline{ji}}^{(t-1)})$ holds in G'_Z. In other words, this implies $P_{W_j[t]|W_j^{(t-1)}, W_i^{(t-1)}, \mathcal{W}_{\overline{ji}}^{(t-1)}} = P_{W_j[t]|W_j^{(t-1)}, \mathcal{W}_{\overline{ji}}^{(t-1)}}$ will hold true for all t and thus, $I(W_i \to W_j \parallel \mathcal{W}_{\overline{ji}}) = 0$. $\square$

The following example illustrates the intuition for the presence of spurious links in the perturbed graph.

Example 1. Consider a generative graph as shown in Figure 4 (a). Suppose node 3 is subject to packet drop corruption model in (7) 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 (7). The measured data-streams are $\{W_1 = Y_1, W_2 = Y_2, W_3 = U_3\}$. Since measurements of node 3 are corrupted, measurements of Y_1 and Y_2 can give useful information for predicting states at node 3 that would not be available in the noisily measured history of U_3. Thus, $I(W_1 \to W_3 \parallel W_2) > 0$ and $I(W_2 \to W_3 \parallel W_1) > 0$. The perturbed graph is shown in figure 4 (b).

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 (1), for all agents $i \in \{1, 2, \ldots, N\}$, and all $j \in \mathcal{P}'(i)$, 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 (4), there is a number $k_i \geq 1$ such that g_i always takes $Y_i[t - k_i]$ 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 (1).
B2) If a node $i \in Z$, then $Y_i[t]$ is an argument of g_i in (4).

Remark 4. The above assumption states that the dynamics in generative model (1), $Y_i[t]$ depends on at least one previous measurement value of its parent nodes. Similarly, for the perturbation model (4), the corrupt value $U_i[t]$ depends causally on uncorrupted measurement value. We consider strictly causal interactions in the generative model and causal interactions in the corruption model and are therefore realistic in many practical physical systems.

The following theorem asserts that if $i \to j \in A_Z$ then there exists a corresponding active trail in perturbed DBN. The proof is given in appendix II.

Theorem 3. *Consider a generative graph, $G = (V, A)$, consisting of N nodes. Let $Z = \{v_1, \ldots, v_n\} \subset V$ 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 $Y_{\bar{Z}} := \{Y_j\}_{j \in \bar{Z}}$ where $\bar{Z} = V \setminus Z$. Let the measured data-streams be $\mathcal{W} = U_Z \cup Y_{\bar{Z}} = \{W_1, W_2, \ldots, W_N\}$. 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 $\{W_j^{(t-1)}, \mathcal{W}_{\overline{ji}}^{(t-1)}\}$ in G'_Z, for some $t > 0$.*

Under the following assumption we can in fact show that $I(W_i \to W_j \parallel W_{\overline{ji}}) > 0$ 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 following Assumption 1, if $i \to j \in A_Z$, then $I(W_i \to W_j \parallel W_{\overline{ji}}) > 0$.*

Proof. By theorem 3, if $i \to j \in A_Z$, then there exists an trail in PDBN between $W_i^{(t-1)}$ and $W_j[t]$ that is active given $\{W_j^{(t-1)}, \mathcal{W}_{\overline{ji}}^{(t-1)}\}$ in G'_Z, for some $t > 0$. Under faithfulness assumption, this implies $P_{W_j[t]|W_j^{(t-1)}, W_i^{(t-1)}, \mathcal{W}_{\overline{ji}}^{(t-1)}} \neq P_{W_j[t]|W_j^{(t-1)}, \mathcal{W}_{\overline{ji}}^{(t-1)}}$. Thus, $I(W_i \to W_j \parallel \mathcal{W}_{\overline{ji}}) > 0$. $\square$

Remark 5. The faithfulness assumption is justified as the unfaithful probability distributions are restricted to a set of Lebesgue measure zero [34]. Here, system parameters for which the algebraic conditions for the conditional independence hold true with true dynamical dependencies must belong to the set of measure zero.

V. Estimation of Directed information

Given the time-series $\mathcal{W}$, the reconstruction of the perturbed graph is accomplished by (i) computing the conditional di-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2021.3093301, IEEE Transactions on Automatic Control

rected information, $I(W_i \to W_j \parallel W_{\overline{ji}})$ (ii) placing a link from node i to j if $I(W_i \to W_j \parallel W_{\overline{ji}}) > 0$. Thus, the algorithm requires computation of $I(W_i \to W_j \parallel W_{\overline{ji}})$ for all pairs of nodes (W_i, W_j) in $\mathcal{W}$. Toward computing the conditional directed information we refer to methods based on Context-Tree-Weighting (CTW) in [35], which provide estimates on conditional probability mass function (PMF) of a time-series admitting values in a finite alphabet. For the reminder of this section, we consider f_i in (1) and g_i in (4) to be such that $Y_i[t]$ and $U_i[t]$ belong to finite alphabet. Here, from time sequence $x^{(n)}$ (recall the notation of $x^{(n)}$), the PMF $Q(x[i] \mid x^{(i-1)})$ for all $i = 1, \ldots, n$ is computed where n is the length of the sequence. Q is also called as sequential probability assignment for a sequence $x^{(n)}$. Furthermore it is shown in [36] that Q computed is a *Universal Probability Assignment* as discussed next.

A. Universal Probability Assignment

The following definition characterizes the probability mass function Q in relation to the true mass function P in terms of the length of the time-series. It establishes that as the horizon of the time-series is extended, the sequential probability assignment estimate, Q, approaches the true PMF P.

Definition 15 (Universal Probability Assignment)**.** Let P be the true joint PMF of $x^{(n)}$. Then, a probability assignment Q is called as *universal* if the following holds:

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}\left[\log \frac{P(x^{(n)})}{Q(x^{(n)})}\right] = 0, \tag{10}$$

where estimated joint PMF for $x^{(n)}$ is given by $Q(x^{(n)}) = Q(x[0])Q(x[1] \mid x[0])Q(x[2] \mid x^{(1)}])\cdots Q(x[n] \mid x^{(n-1)})$. Similarly, $P(x^{(n)})$ can be factorized.

For the rest of the article Q is estimated by CTW algorithm which is a universal probability assignment as discussed in [36] for each time series. The only assumptions made are that the sequences belong to a finite alphabet and are stationary and ergodic Markov sequences of a bounded order D. That is, for a Markov sequence X, $P(x[t]|x^{(t-1)}) = P(x[t]|x_{t-l}^{t-1})$ where $l \leq D$. The CTW algorithm uses a weighted distribution to take into account of all possible D-bounded Markov sources and estimates the sequential probability, $Q(x[t]|x^{(t-1)})$ for every symbol $x[t]$ given the past observations. The computational complexity of CTW algorithm is linear in horizon length n, of the sequences considered.

B. Pairwise Estimation of Directed Information

Here, a pairwise estimator of directed information between a pair of random process proposed by [36] is described. 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:

$$I(X \to Y) = H(Y) - H(Y \parallel X) \tag{11}$$

where $H(Y) = \mathbb{E}[-\log P(Y)]$ and $H(Y \parallel X) = \mathbb{E}[-\log P(Y \parallel X)]$ denotes the entropy of Y and the causally conditioned entropy [37] respectively.

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

$$I_r(X \to Y) = \lim_{n\to\infty} \frac{1}{n} I(x^{(n)} \to y^{(n)}). \tag{12}$$

The directed information rate in (12) characterizes the directed information from X to Y in the limiting sense of the horizon being infinite. Let $H_r(Y) := \lim_{n\to\infty} \frac{1}{n} H(y^{(n)})$ and let $H_r(Y \parallel X) := \lim_{n\to\infty} \frac{1}{n} H(y^{(n)} \parallel x^{(n)})$. Thus, if $H_r(Y)$ and $H_r(Y \parallel X)$ converge, then I_r is convergent. That is,

$$I_r = H_r(Y) - H_r(Y \parallel X). \tag{13}$$

In [36], the following DIR estimator is defined:

$$\hat{I}(x^{(n)} \to y^{(n)}) = \frac{1}{n}\left\{\sum_{i=1}^{n} \sum_{y[i]\in\mathcal{Y}} Q(y[i] \mid x^{(i-1)}, y^{(i-1)}) \cdot \log \frac{1}{Q(y[i] \mid y^{(i-1)})}\right\} - \frac{1}{n}\left\{\sum_{i=1}^{n} \sum_{y[i]\in\mathcal{Y}} Q(y[i] \mid x^{(i-1)}, y^{(i-1)}) \cdot \log \frac{1}{Q(y[i] \mid x^{(i-1)}, y^{(i-1)})}\right\} \tag{14}$$

In [36], consistency results for estimating directed information (DI) between a pair of random processes from data was proposed. In this article we provide consistency results of the conditional directed information estimator by showing convergence in almost sure sense (denoted as P-a.s).

C. 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:

$$I(X \to Y \parallel Z) = H(Y \parallel Z) - H(Y \parallel X, Z). \tag{15}$$

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

$$I_r(X \to Y \parallel Z) = \lim_{n\to\infty} \frac{1}{n} I(x^{(n)} \to y^{(n)} \parallel z^{(n)}). \tag{16}$$

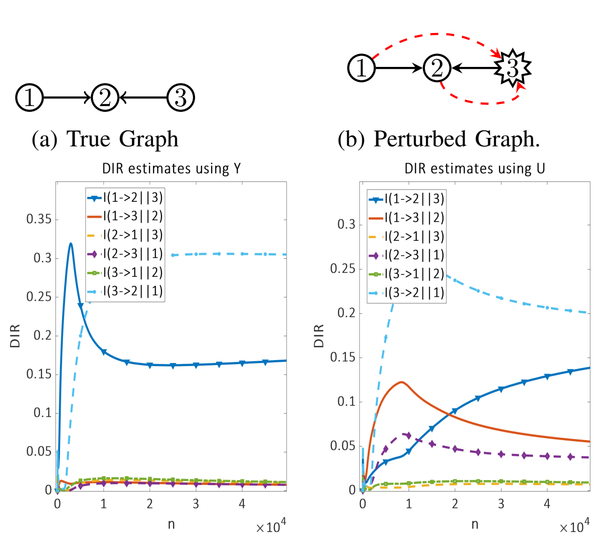
Let $H_r(Y \parallel X, Z) := \lim_{n\to\infty} \frac{1}{n} H(y^{(n)} \parallel x^{(n)}, z^{(n)})$. Thus, if $H_r(Y \parallel Z)$ and $H_r(Y \parallel X, Z)$ converge, then I_r is convergent. That is,

$$I_r = H_r(Y \parallel Z) - H_r(Y \parallel X, Z). \tag{17}$$

The conditional directed information estimator $\hat{I}(x^{(n)} \to y^{(n)} \parallel z^{(n)})$ is defined as under:

$$\hat{I}(x^{(n)} \to y^{(n)} \parallel z^{(n)}) = \frac{1}{n}\sum_{i=1}^{n} \sum_{y[i]\in\mathcal{Y}} Q(y[i] \mid x^{(i-1)}, y^{(i-1)}, z^{(i-1)}) \cdot \log \frac{1}{Q(y[i] \mid y^{(i-1)}, z^{(i-1)})} - \frac{1}{n}\sum_{i=1}^{n} \sum_{y[i]\in\mathcal{Y}} Q(y[i] \mid x^{(i-1)}, y^{(i-1)}, z^{(i-1)}) \cdot \log \frac{1}{Q(y[i] \mid x^{(i-1)}, y^{(i-1)}, z^{(i-1)})} \tag{18}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2021.3093301, IEEE Transactions on Automatic Control

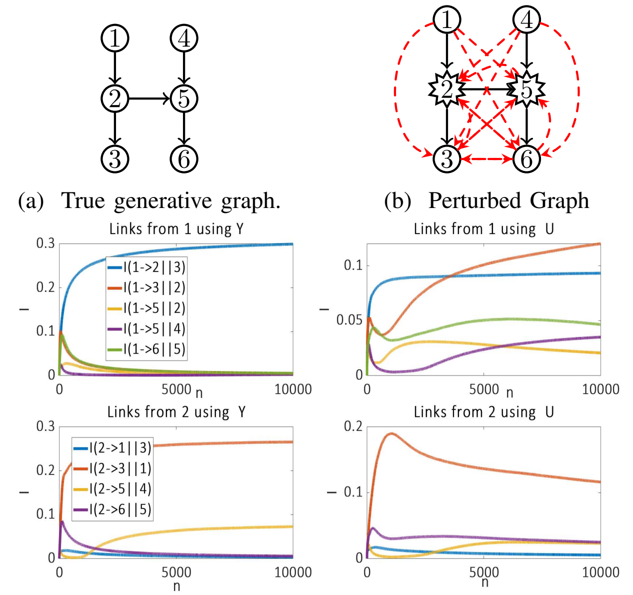


(a) True Graph

(b) Perturbed Graph.

(c) Comparison of directed information estimates between perfect measurements and corrupted data-streams. DIR I is shown along X-axis and the sample length n is along Y-axis.

Fig. 5: This figure shows how unreliable measurements at node 3 results in spuriously inferring a causal influence from $1 \to 3$ and $2 \to 3$. 5(b) shows the perturbed graph inferred. Spurious edges are shown in red while true edges are in black.



(a) True generative graph.

(b) Perturbed Graph

(c) Comparison of directed information rate (DIR) estimates for links from nodes 1 and 2, between ideal data-streams Y and uncertain measurements U. DIR I is shown along X-axis and the sample length n is along Y-axis.

Fig. 6: 6(a) shows true generative graph. 6(c) depicts DIR estimates to detect links from nodes 1 and 2 using ideal measurements Y and when there is corruption at nodes 2 and 5. 6(b) shows the perturbed graph inferred. The spurious links are shown in red and the true edges are shown in black. With cascaded perturbations, more spurious links are inferred.

The following theorem establishes the consistency result in estimating conditional DIR as defined in (18). The proof is given in appendix III.

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

$$\lim_{n\to\infty} \hat{I}(x^{(n)} \to y^{(n)} \parallel Z^{(n)}) = I_r(X \to Y \parallel Z) \qquad P\text{-a.s,} \tag{19}$$

For computing (18), first $Q(x[i], y[i], z[i] \mid x^{(i-1)}, y^{(i-1)}, z^{(i-1)})$ and $Q(y[i], z[i] \mid y^{(i-1)}, z^{(i-1)})$ are estimated using CTW for all realizations of tuples $(x[i], y[i], z[i])$ and $(y[i], z[i])$. The estimated probabilities are tabulated and the required marginalized conditional probabilities $Q(y[i] \mid x^{(i-1)}, y^{(i-1)}, z^{(i-1)})$ and $Q(y[i] \mid y^{(i-1)}, z^{(i-1)})$ in (18) are computed from this table for entropy estimation.

D. 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 until the horizon. We used the estimator described in (18) to compute DIR. For both the networks, the horizon length are in the order 10^4.

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

$$\begin{aligned} Y_1[t] &= E_1[t], \\ Y_2[t] &= (|Y_1[t-1] - Y_1[t-2]| \cdot Y_3[t-1]^2 \cdot E_2[t]) \bmod 3, \\ Y_3[t] &= E_3[t] \end{aligned}$$

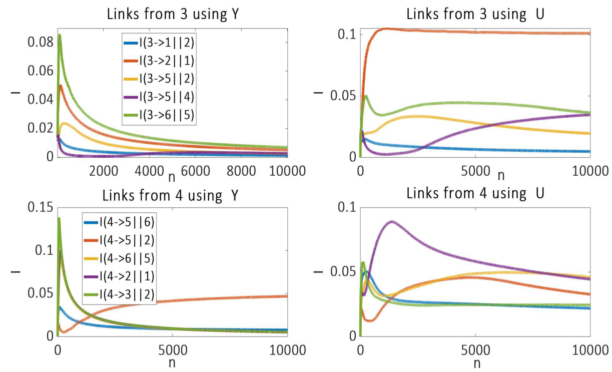
where $E_1[t] \sim \text{Categorical}(3, [0.15, \ 0.35, \ 0.5])$, $E_2[t] \sim \text{Categorical}(3, [0.35, \ 0.35, \ 0.3])$ and $E_3[t] \sim \text{Categorical}(3, [0.4, \ 0.2, \ 0.4])$. Each of $Y_1[t], Y_2[t]$ and $Y_3[t]$ has a finite alphabet $\{0, 1, 2\}$.

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

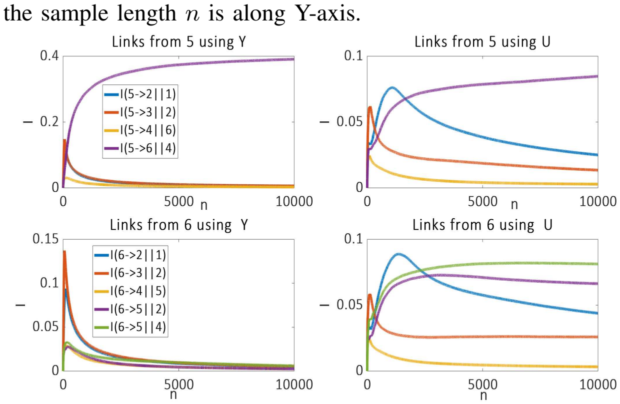
$$U_3[t] = \begin{cases} (Y_3[t] + U_3[t-1]) \bmod 3, & \text{with probability } 0.55 \\ (Y_3[t-1] + U_3[t-1]) \bmod 3, & \text{with prob. } 0.45. \end{cases}$$

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



(a) A comparison of DIR estimates to detect links from nodes 3 and 4 using ideal measurements and when there is corruption at nodes 2 and 5 is shown. DIR I is shown along X-axis and the sample length n is along Y-axis.



(b) A comparison of DIR estimates to detect links from nodes 5 and 6 using ideal measurements and when there is corruption at nodes 2 and 5 is shown. DIR I is shown along X-axis and the sample length n is along Y-axis.

Fig. 7: DI estimates to detect links from nodes 3,4,5 and 6. Notice the large number of non-zero DIR estimates computed from corrupt measurements corresponding to links from nodes 3 and 6 which had no children in the true generative graph that now has lot of children nodes in G_Z.

generative model are as follows:

$$\begin{aligned} Y_1[t] &= E_1[t], \\ Y_2[t] &= Y_1[t-1]||E_2[t], \\ Y_3[t] &= Y_2[t-1]||E_3[t], \\ Y_4[t] &= E_4[t], \\ Y_5[t] &= (Y_2[t-1]||Y_4[t-1])\&E_5[t], \\ Y_6[t] &= Y_5[t-1]||E_6[t] \end{aligned}$$

where $E_1[t] \sim$ Bernouilli(0.55), $E_2[t] \sim$ Bernouilli(0.5), $E_3[t] \sim$ Bernouilli(0.2), $E_4[t] \sim$ Bernouilli(0.4), $E_5[t] \sim$ and $E_6[t] \sim$ Bernouilli(0.3)and '$||$' is logical 'OR' operation while '&' is logical 'AND' operation. Each of $Y_1[t], Y_2[t], \dots, Y_6[t]$ has a finite alphabet $\{0, 1\}$. The perturbations considered here are time-origin uncertainties at nodes 2 and 5. The corruption models takes the form:

$$U_2[t] = \begin{cases} Y_2[t-2], & \text{with probability } 0.5, \\ Y_2[t], & \text{with probability } 0.5, \end{cases}$$

$$U_5[t] = \begin{cases} Y_5[t-2], & \text{with probability } 0.5, \\ Y_5[t], & \text{with probability } 0.5. \end{cases}$$

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

VI. Conclusion

We studied the problem of inferring directed graphs for a large class of networks that admit nonlinear and strictly causal interactions between several agents. We provided necessary and sufficient conditions that delineated the effects of data corruption on the directed network structure inferred using directed information. We presented a tight characterization for the spurious links that arise due to corruption of data-streams by determining their location and orientation. 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. Another interesting direction would be to consider network reconstruction for non-target specific nonlinear dynamical systems. Non-target specific network reconstruction for linear systems studied in [38] and [39] may yield useful insights in this direction. Moreover, it would be interesting to characterize effects of data corruption in other network reconstruction methods. Verifying identifiability conditions [14], [15] and quantifying error in the identified transfer function due to data corruption can be an interesting line of work.

Appendix I
Proof for Theorem 2

We will show that if $i \to j \notin A_Z$, then there is no trail between $W_i^{(t-1)}$ and $W_j[t]$ that is active given $\{W_j^{(t-1)}, \mathcal{W}_{\bar{j}\bar{i}}^{(t-1)}\}$ in the PDBN G'_Z, for all $t > 0$. However, trail in G'_Z between $W_i^{(t-1)}$ and $W_j[t]$ can contain future state node $\alpha_{b_m}[t_m]$ such that $t_m \geq t$ and $b_m \in V$. Note that such nodes are not observed in $\{W_j^{(t-1)}, \mathcal{W}_{\bar{j}\bar{i}}^{(t-1)}\}$ and can therefore make the trail in PDBN active. However, the following lemma proves that such nodes actually makes the trail inactive given $\{W_j^{(t-1)}, \mathcal{W}_{\bar{j}\bar{i}}^{(t-1)}\}$.

Lemma 1. *Consider a generative graph, $G = (V, A)$, consisting of N nodes. Let $Z = \{v_1, \dots, v_n\} \subset V$ 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 $Y_{\bar{Z}} := \{Y_j\}_{j \in \bar{Z}}$ where $\bar{Z} = V \setminus Z$. Let the measured*

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2021.3093301, IEEE Transactions on Automatic Control

data-streams be $\mathcal{W} = U_Z \cup Y_{\bar{Z}} = \{W_1, W_2, \ldots, W_N\}$. Let $G' = (V', A')$ be the dynamic Bayesian network (DBN) associated with G and $G'_Z = (V'_Z, A'_Z)$ be the perturbed DBN. If $i \to j \notin A$ 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 $t_m \geq t$ and $b_m \in V$, then for all $t > 0$, the trail is not active given $\{W_j^{(t-1)}, \mathcal{W}_{\bar{j}\bar{i}}^{(t-1)}\}$.

Proof. Consider any trail from a node in $W_i^{(t-1)}$ to $W_j[t]$ in G'_Z. Denote this by $trlG'_Z := W_i[t_1] = \alpha_{b_1}[t_1] - \alpha_{b_2}[t_2] - \cdots - \alpha_{b_{r-1}}[t_{r-1}] - \alpha_{b_r}[t_r] = W_j[t]$ where $0 \leq t_1 < t$. Here, b_k denotes the corresponding vertex in V for $k = \{1, 2, \ldots, r\}$. Also, $\alpha_{b_k}[t_k] = U_{b_k}[t_k]$ if $b_k \in Z$ or $\alpha_{b_k}[t_k] = Y_{b_k}[t_k]$ otherwise. For compact notation, set $\theta := \{W_j^{(t-1)}, \mathcal{W}_{\bar{j}\bar{i}}^{(t-1)}\}$.

The trail has length at least 3. As $i \to j \notin A$ and if $j \notin Z$, 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_i nor U_i. Thus, there is no direct link of the form $\alpha_i[t'] \to \alpha_j[t"]$ in G'_Z, for any $t', t"$. In particular, $W_i[t_1] \to W_j[t] \notin G'_Z$. Thus, there are at least 3 nodes in the trail, $trlG'_Z$.

Unobserved collider in trail. Without loss of generality, choose $t_m = \max\{t_1, \ldots, t_{r-1}\} \geq t$. Consider the sub-trail $subtrl' := \alpha_{b_{m-1}}[t_{m-1}] - \alpha_{b_m}[t_m] - \alpha_{b_{m+1}}[t_{m+1}]$ 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 $\alpha_{b_{m-1}}[t_{m-1}]$, $\alpha_{b_m}[t_m]$, and $\alpha_{b_{m+1}}[t_{m+1}]$ is a collider not in θ and therefore the trail $trlG'_Z$ cannot be active given θ.

Suppose $t_m > t_{m-1}$ and $t_m > t_{m+1}$. Then, $subtrl'$ is of the form, $\alpha_{b_{m-1}}[t_{m-1}] \to \alpha_{b_m}[t_m] \leftarrow \alpha_{b_{m+1}}[t_{m+1}]$. Note that, as $t_m \geq t$, it follows that neither $\alpha_{b_m}[t_m]$ nor any of its descendants can be in θ 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 $p \in V$, $Y_p[t_p]$ does not dynamically depend on any $Y_q[t_p]$ for $q \in \{p, \mathcal{P}'(p)\}$. By the perturbation model described by (4), for any $q \in Z$, $U_q[t_q]$ dynamically depends only on $\{U_q^{(t_q-1)}, Y_q^{(t_q)}\}$. 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 $t_m > t_{m-1}$, $subtrl'$ is in fact $\alpha_{b_{m-1}}[t_{m-1}] \to \alpha_{b_m}[t_m] = U_{b_m}[t_m] \leftarrow \alpha_{b_{m+1}}[t_{m+1}] = Y_{b_m}[t_m]$. Therefore, $\alpha_{b_m}[t_m]$ is a collider and as $t_m \geq t$, it is not observed in θ.

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 $\alpha_{b_{m-1}}[t_{m-1}] - \alpha_{b_m}[t_m] = Y_{b_m}[t_m] \to \alpha_{b_{m+1}}[t_{m+1}] = U_{b_m}[t_m] \leftarrow \alpha_{b_{m+2}}[t_{m+2}]$ in $trlG'_Z$. Therefore, $\alpha_{b_{m+1}}[t_{m+1}]$ is a collider not observed in θ. □

Proof of Theorem 2

For rest of the proof, denote $\theta := \{W_j^{(t-1)}, \mathcal{W}_{\bar{j}\bar{i}}^{(t-1)}\}$. Note that if $i \to j \notin A_Z$, then there is no directed edge from i to j in G, and every trail from i to j in G violates at least one of the conditions of Definition 13. We will consider these cases separately and show that no active trail exists in G'_Z in each case. Denote any trail connecting a node in $W_i^{(t-1)}$ and $W_j[t]$ in G'_Z, by $trlG'_Z := W_i[t_1] = \alpha_{b_1}[t_1] - \alpha_{b_2}[t_2] - \cdots - \alpha_{b_{r-1}}[t_{r-1}] - \alpha_{b_r}[t_r] = W_j[t]$ where $0 \leq t_1 < t$ and b_k denotes the corresponding vertex in V for $k = \{1, 2, \ldots, r\}$. Here, $\alpha_v[t_v] = U_v[t_v]$ if $v \in Z$ or $\alpha_v[t_v] = Y_v[t_v]$ otherwise. Using Lemma 1, if any t' in $\{t_2, \ldots, t_{r-1}\}$ is such that $t' \geq t$, then $trlG'_Z$ is not active. Now, consider $0 \leq t_1, t_2, t_3, \ldots, t_{r-1} < t$. We will first show that any such trail in G'_Z, $trlG'_Z$, can be mapped to a trail in G, $trlG := i = v_1 - v_2 - v_3 \ldots v_{k-1} - v_k = j$ as follows:

Initialize: $k = 1$ and $v_1 = b_1$.
for $l = 1 : r - 1$ **do**
  if $b_{l+1} \neq b_l$ in $\alpha_{b_l}[t_l] - \alpha_{b_{l+1}}[t_{l+1}]$ along $trlG'_Z$ **then**
    Set $v_{k+1} = b_{l+1}$.
    Add edge $v_k - v_{k+1}$ with the same direction as $\alpha_{b_l}[t_l] - \alpha_{b_{l+1}}[t_{l+1}]$.
    Set $s_k = t_l$ and $\tau_{k+1} = t_{l+1}$
    Set $k = k + 1$
  end if
end for

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

Now, let us reason out why such a construction is always feasible. To this, we claim that for any successive pair $\alpha_{b_l}[t_l] - \alpha_{b_{l+1}}[t_{l+1}]$, either $b_l = b_{l+1}$ or, $b_l \neq b_{l+1}$ and $b_l - b_{l+1} \in A$ with the same direction as in $\alpha_{b_l}[t_l] - \alpha_{b_{l+1}}[t_{l+1}]$. Assume $\alpha_{b_l}[t_l] \to \alpha_{b_{l+1}}[t_{l+1}]$. (The case of $\alpha_{b_l}[t_l] \leftarrow \alpha_{b_{l+1}}[t_{l+1}]$ is similar). Then, either $t_l = t_{l+1}$ or $t_l < t_{l+1}$. Consider, $t_l = t_{l+1}$. Then, the link must have the form $Y_{b_l}[t_l] \to U_{b_l}[t_l]$, as this is the only instantaneous influence defined in (1) or (4). Thus, $b_l = b_{l+1}$ in this case.

Suppose, $t_l < t_{l+1}$. Either, $b_{l+1} \in Z$ or $b_{l+1} \notin Z$. Consider $b_{l+1} \in Z$. By the perturbation model described by (4), $\alpha_{b_l}[t_l] \in \{Y_{b_{l+1}}^{(t_{l+1}-1)}, U_{b_{l+1}}^{(t_{l+1}-1)}\}$. Therefore, $b_l = b_{l+1}$. Suppose, $b_{l+1} \notin Z$. Then, $\alpha_{b_{l+1}}[t_{l+1}] = Y_{b_{l+1}}[t_{l+1}]$. By the generative model in (1), we either have dynamic dependence on self-history or history of other nodes. That is, $\alpha_{b_l}[t_l] \in \{Y_{b_l}^{(t_{l+1}-1)}, \bigcup_{q \in \mathcal{P}'(b_{l+1})} Y_q^{(t_{l+1}-1)}\}$. Then, $b_l = b_{l+1}$ when there is dependence on self-history. Otherwise, $b_l \in \mathcal{P}'(b_{l+1})$. Thus, $b_l \to b_{l+1} \in A$. Let us consider an example- from a trail of the form $U_1[t_1] \leftarrow Y_1[t_2] \leftarrow Y_2[t_3] \to Y_3[t_4] \to Y_3[t_5] \to U_3[t]$ in G'_Z, a trail $trlG$ in G can be constructed as $1 \leftarrow 2 \to 3$.

Additionally, we may assume that for $m = 2, \cdots, r-1$ we have that $\alpha_{b_m}[t_m] \neq W_i[t_m]$ in $trlG'_Z$. If $\alpha_{b_m}[t_m] = W_i[t_m]$ for some $m > 1$, then the sub-trail of $trlG'_Z$, $W_i[t_m] = \alpha_{b_m}[t_m] - \alpha_{b_{m+1}}[t_{m+1}] - \cdots - \alpha_r[t_r] = W_j[t]$ is a trail from $W_i[t_m] \in W_i^{(t-1)}$ to $W_j[t]$. This trail is of strictly shorter length than $trlG'_Z$. Thus, if the shorter trail cannot be active then the longer trail, $trlG'_Z$, cannot be active either. Also, by following the construction procedure described above, this condition implies that $v_l \neq i$ for $l = 2, 3, \cdots, k$ in $trlG$. Call this condition $loop_i$. To summarize, let $trlG := i = v_1 - v_2 - v_3 \ldots v_{k-1} - v_k = j$ be any trail connecting i and j in G constructed by following the above procedure from the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2021.3093301, IEEE Transactions on Automatic Control

trail $trlG'_Z$: $W_i[t_1] = \alpha_{b_1}[t_1] - \alpha_{b_2}[t_2] - \cdots - \alpha_{b_{r-1}}[t_{r-1}] - \alpha_{b_r}[t_r] = W_j[t]$. Since, $i \to j \notin A_Z$, this trail must violate any of the conditions P1), P2) and P3). We will now consider these cases separately and prove that there is no corresponding active trail in G'_Z.

If condition P1) is violated, then $trlG$ must have that $j \notin Z$ and $v_{k-1} \leftarrow j$. In this case, $W_j = Y_j$. Then, either $b_{r-1} = j$ or $b_{r-1} \neq j$. By construction of $trlG$, if $b_{r-1} \neq j$, then $b_{r-1} = v_{k-1}$. As $v_{k-1} \leftarrow j$, we must then have $\alpha_{b_{r-1}}[t_{r-1}] \leftarrow \alpha_{b_r}[t_r]$. However, this implies $t_r = t < t_{r-1}$ which violates the condition that $0 \leq t_1, t_2, t_3, \ldots, t_{r-1} < t$. Thus, $b_{r-1} = j$. That is, $\alpha_{b_{r-1}}[t_{r-1}] = Y_j[t_{r-1}]$. As $t_{r-1} < t$ and $j \notin Z$ we have $\alpha_{b_{r-1}}[t_{r-1}] = Y_j[t_{r-1}] \to \alpha_{b_r}[t] = Y_j[t]$ as a sub-trail of $trlG'_Z$. Clearly, $y_j[t_{r-1}]$ is not a collider. As $t_{r-1} < t$, we have $y_j[t_{r-1}] \in \theta$. Thus the trail cannot be active.

Recall the definitions of s_k and τ_{k+1} during construction of the trail in G. If condition P2) is violated, then a sub-path of $trlG$, $v_{m-1} \to v_m \leftarrow v_{m+1}$, must have a collider, v_m, such that $v_m \notin Z$ and $v_{m+1} \notin Z$ where $m = \{2, 3 \cdots, k-1\}$. If $v_{m+1} = j$ and $\tau_{m+1} = t$, P1) also fails, and the argument above shows that the trail in G'_Z is not active. If $v_{m+1} = j$ and $\tau_{m+1} < t$ then we have that $\alpha_{v_{m+1}}[\tau_{m+1}] = Y_{v_{m+1}}[\tau_{m+1}] \in \theta$ which is an observed node along the trail and is not a collider. Thus, the trail $trlG'_Z$ cannot be active. So, assume that $v_{m+1} \neq j$. By condition $loop_i$, $m+1 \neq i$. As $v_m \leftarrow v_{m+1} \in trlG$, by construction we must have $Y_{v_m}[s_m] = \alpha_{v_m}[s_m] \leftarrow \alpha_{v_{m+1}}[\tau_{m+1}] = Y_{v_{m+1}}[\tau_{m+1}]$ along $trlG'_Z$ with $\tau_{m+1} < s_m < t$. Note that since $v_{m+1} \notin Z$ and $\tau_{m+1} < t$, $\alpha_{v_{m+1}}[\tau_{m+1}] = Y_{v_{m+1}}[\tau_{m+1}]$ is an observed non-collider in θ. Thus, the trail cannot be active.

Finally consider the case that P3) is violated. Then along the trail, $trlG$, in G, there must be a sub-trail $v_{m-1} - v_m - v_{m+1}$ 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 $trlG$ 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 $trlG'_Z$ 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 $trlG'_Z$ which is a non-collider. Then, as $v_m \notin Z$, we must have that $\alpha_{v_m}[t_m] = W_{v_m}[t_m] = Y_{v_m}[t_m]$. Note that $v_m \neq i$ by condition $loop_i$. As $t_m < t$, $\alpha_{v_m}[t_m]$ is an intermediate non-collider node in θ and is thus observed. Hence, $trlG'_Z$ cannot be active. □

Appendix II
Proof for Theorem 3

Suppose $i \to j$ is in A_Z. Then there is a trail, trl_G, described by $i = v_1 - v_2 - \cdots - v_k = j$ in G satisfying conditions in Definition 13. 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 $trlG$. Otherwise, set $t_l = t_{l+1} + k_{v_l v_{l+1}}$ if $v_l \leftarrow v_{l+1}$ holds in $trlG$. Such a construction is feasible because by condition C 1), numbers $k_{v_{l+1}v_l}$ and $k_{v_l v_{l+1}}$ exists for all $l \in \{1, 2, \ldots, k-1\}$ and at all times. Thus, we have a trail $Y_i[t_1] - Y_{v_2}[t_2] - Y_{v_3}[t_3] - \ldots - Y_{v_{k-1}}[t_{k-1}] - Y_j[t_k]$. For all $m \in \{1, 2, \ldots, k\}$ if $v_m \in Z$, there exists a number $k_m > 0$ following conditions C 2). If B 2) also holds, then $k_m \geq 0$. Let $t > \max\{t_1, \ldots, t_{k-1}\}$, and for all $m \in \{1, 2, \ldots, k\}$ if $v_m \in Z$, let $t > t_m + k_m$ 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 C 2) $U_i[t_1 + k_i] \leftarrow Y_i[t_1]$ holds true. Choose t sufficiently large so that $t > t_1 + k_i$ also holds. As $j \in Z$, using C 2), t can be sufficiently large so that we have $Y_j[t_k] \to U_j[t]$ where $t = t_k + k_j$ and $k_j \geq 1$. If B 1) holds, then we can choose t sufficiently large such that at the end of the trail we take s 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 $W_i[t_1 + k_i] = U_i[t_1 + k_i] \leftarrow Y_i[t_1] - Y_{v_2}[t_2] - Y_{v_3}[t_3] - \cdots - Y_{v_{k-1}}[t_{k-1}] - Y_j[t_k] \to U_j[t] = W_j[t]$, or $W_i[t_1 + k_i] = U_i[t_1 + k_i] \leftarrow Y_i[t_1] - Y_{v_2}[t_2] - Y_{v_3}[t_3] - \cdots - Y_{v_{k-1}}[t_{k-1}] - Y_j[t_k] \to Y_j[t_k + k'_j] \to \cdots \to Y_j[t_k + sk'_j] \to U_j[t] = W_j[t]$ with $t > \max\{t_1 + k_i, t_1, \ldots, t_k, \ldots, t_k + sk'_j\}$, 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 C 2) $U_i[t_1 + k_i] \leftarrow Y_i[t_1]$ 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: $W_i[t_1 + k_i] = U_i[t_1 + k_i] \leftarrow Y_i[t_1] - Y_{v_2}[t_2] - Y_{v_3}[t_3] - \cdots - Y_{v_{k-1}}[t_{k-1}] - Y_j[t] = W_j[t]$ with $t > \max\{t_1 + k_i, t_1, \ldots, t_{k-1}\}$, 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 $W_i[t_1] = Y_i[t_1] - Y_{v_2}[t_2] - Y_{v_3}[t_3] - \cdots - Y_{v_{k-1}}[t_{k-1}] - Y_j[t_k] \to U_j[t] = W_j[t]$, or of form $W_i[t_1] = Y_i[t_1] - Y_{v_2}[t_2] - Y_{v_3}[t_3] - \cdots - Y_{v_{k-1}}[t_{k-1}] - Y_j[t_k] \to Y_j[t_k + k'_j] \to \cdots \to Y_j[t_k + sk'_j] \to U_j[t] = W_j[t]$ exists in the perturbed DBN G'_Z with $t > \max\{t_1, \ldots, t_k, \ldots, t_k + sk'_j\}$, and for all $m \in \{1, 2, \ldots, k\}$ if $v_m \in Z$, $t > t_m + k_m$.
D) Consider the case $i \notin Z$ and $j \notin Z$. Following arguments presented in Case (B) we conclude that the trail $W_i[t_1] = Y_i[t_1] - Y_{v_2}[t_2] - Y_{v_3}[t_3] - \cdots - Y_{v_{k-1}}[t_{k-1}] - Y_j[t] = W_j[t]$ exists in the perturbed DBN G'_Z with $t > \max\{t_1, \ldots, t_{k-1}\}$, and for all $m \in \{1, 2, \ldots, k\}$ if $v_m \in Z$, $t > t_m + k_m$.

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

Sub-trails with colliders: For all the trails in G'_Z constructed under various cases above consider a sub-trail of the form $Y_{v_{m-1}}[t_{m-1}] \to Y_{v_m}[t_m] \leftarrow Y_{v_{m+1}}[t_{m+1}]$. Clearly, v_m cannot be either i or j. If $v_m \notin Z$ then as $t_m < t$, we have $Y_{v_m}[t_m] \in W_{\bar{j}\bar{i}}^{(t-1)}$ and thus the sub-trail is active. If $v_m \in Z$ then the corrupted version of $Y_{v_m}[t_m]$ is $U_{v_m}[t_m + k_{v_m}] = W_{v_m}[t_m + k_{v_m}]$ and as $t_m + k_{v_m} < t$, we have $W_{v_m}[t_m + k_m] \in W_{\bar{j}\bar{i}}^{(t-1)}$. Thus the collider $Y_{v_m}[t_m]$ has a descendant $W_{v_m}[t_m + k_{v_m}] \in \theta$. Thus the sub-trail remains active. Thus no collider can deactivate the trails in G'_Z.

Sub-trails with with no colliders: Now consider any node

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2021.3093301, IEEE Transactions on Automatic Control

$Y_{v_m}[t_m]$ which is not a collider. Note that in the trails for the cases (A), (B), (C), and (D), Y_j and Y_i can only appear as an intermediate node only if they are corrupted. In such cases, neither $Y_i[t_1]$ nor $Y_j[t_k]$ belong to θ. Thus, if Y_j or Y_i are intermediate nodes, they cannot deactivate the trails given θ. Consider an intermediate node $v_m \notin \{i, j\}$. From Definition 13P 3), v_m is corrupted. Thus $Y_{v_m}[t_m] \neq W_{v_m}[t_m]$ and $Y_{v_m}[t_m]$ cannot deactivate the trail as $Y_{v_m}[t_m] \notin \theta$. □

Appendix III
Proof for Theorem 4

To prove the theorem, we require two results from [36]. 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,*

$$\lim_{n\to\infty} Q(x[n] \mid x^{(n-1)}) - P(x[n] \mid x^{(n-1)}) = 0 \text{ } P\text{-}as. \quad (20)$$

Next, we will later use the following proposition which is a rephrased result from [36].

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 $\hat{H}(y^{(n)} \parallel x^{(n)}) = \frac{1}{n}\sum_{i=1}^{n}\sum_{y[i]\in\mathcal{Y}} Q(y[i] \mid x^{(i-1)}, y^{(i-1)}) \cdot \log \frac{1}{Q(y[i]|x^{(i-1)},y^{(i-1)})}$. Then,*

$$\lim_{n\to\infty} \hat{H}(y^{(n)} \parallel x^{(n)}) - H_r(Y \parallel X) = 0 \qquad P\text{-}a.s, \quad (21)$$

Recall the expression for conditional DI estimator from (18):

$$\hat{I}(x^{(n)} \to y^{(n)} \parallel z^{(n)}) = \frac{1}{n}\sum_{i=1}^{n}\sum_{y[i]\in\mathcal{Y}} Q(y[i] \mid x^{(i-1)}, y^{(i-1)}, z^{(i-1)}) \cdot \log \frac{1}{Q(y[i] \mid y^{(i-1)}, z^{(i-1)})} - \frac{1}{n}\sum_{i=1}^{n}\sum_{y[i]\in\mathcal{Y}} Q(y[i] \mid x^{(i-1)}, y^{(i-1)}, z^{(i-1)}) \cdot \log \frac{1}{Q(y[i] \mid x^{(i-1)}, y^{(i-1)}, z^{(i-1)})} \quad (22)$$

We will show that the first term (call it T1) in equation (22) converges to $H_r(Y \parallel Z)$ and the second term (call it T2) in (22) converges to $H_r(Y \parallel X, Z)$.

Convergence of T2: Let $V = \{X, Z\}$. Thus, T2 can be written as $\hat{H}(y^{(n)} \parallel v^{(n)}) = \frac{1}{n}\sum_{i=1}^{n}\sum_{y[i]\in\mathcal{Y}} Q(y[i] \mid v^{(i-1)}, y^{(i-1)}) \cdot \log \frac{1}{Q(y[i]|v^{(i-1)},y^{(i-1)})}$. Using, proposition 1, we thus have that $\lim_{n\to\infty} \hat{H}(y^{(n)} \parallel v^{(n)}) \to H_r(Y \parallel V)$ almost surely.

Convergence of T1: Subtract $H_r(Y \parallel Z)$ from $T1$ and express $T1 - H_r(Y \parallel Z) = F_n + S_n$ where,

$$F_n = \frac{1}{n}\sum_{i=1}^{n}\sum_{y[i]\in\mathcal{Y}} P(y[i] \mid x^{(i-1)}, y^{(i-1)}, z^{(i-1)}) \cdot \log P(y[i] \mid y^{(i-1)}, z^{(i-1)}) - \frac{1}{n}\sum_{i=1}^{n}\sum_{y[i]\in\mathcal{Y}} Q(y[i] \mid x^{(i-1)}, y^{(i-1)}, z^{(i-1)}) \cdot \log Q(y[i] \mid y^{(i-1)}, z^{(i-1)}), \quad (23)$$

$$S_n = -\frac{1}{n}\sum_{i=1}^{n}\sum_{y[i]\in\mathcal{Y}} P(y[i] \mid x^{(i-1)}, y^{(i-1)}, z^{(i-1)}) \cdot \log P(y[i] \mid y^{(i-1)}, z^{(i-1)}) - H_r(Y \parallel Z) \quad (24)$$

By ergodicity, S_n converges to zero almost surely. We need to show that F_n converges to zero almost surely. Rewrite $F_n = \frac{1}{n}\sum_{i=1}^{n}\beta_i$ where,

$$\beta_i = \sum_{y[i]\in\mathcal{Y}} P(y[i] \mid x^{(i-1)}, y^{(i-1)}, z^{(i-1)}) \cdot \log P(y[i] \mid y^{(i-1)}, z^{(i-1)}) - \sum_{y[i]\in\mathcal{Y}} Q(y[i] \mid x^{(i-1)}, y^{(i-1)}, z^{(i-1)}) \cdot \log Q(y[i] \mid y^{(i-1)}, z^{(i-1)}) \quad (25)$$

By Lemma 2, the CTW probabilities $Q(y[i] \mid x^{(i-1)}, y^{(i-1)}, z^{(i-1)})$ converges to true probabilities $P(y[i] \mid x^{(i-1)}, y^{(i-1)}, z^{(i-1)})$ almost surely. Therefore,

$$\lim_{i\to\infty} \beta_i = 0 \qquad \text{P-a.s.} \quad (26)$$

Hence, by Cesaro mean [40] we have:

$$\lim_{n\to\infty} F_n = \lim_{n\to\infty} \frac{1}{n}\beta_i = 0 \quad \text{P-a.s.} \quad \square \quad (27)$$

References

[1] M. Kretschmer, D. Coumou, J. F. Donges, and J. Runge, "Using causal effect networks to analyze different arctic drivers of midlatitude winter circulation," *Journal of Climate*, vol. 29, no. 11, pp. 4069–4081, 2016.
[2] A. Sendrowski, K. Sadid, E. Meselhe, W. Wagner, D. Mohrig, and P. Passalacqua, "Transfer entropy as a tool for hydrodynamic model validation," *Entropy*, vol. 20, no. 1, p. 58, 2018.
[3] N. Omranian, J. M. Eloundou-Mbebi, B. Mueller-Roeber, and Z. Nikoloski, "Gene regulatory network inference using fused lasso on multiple data sets," *Scientific reports*, vol. 6, p. 20533, 2016.
[4] D. S. Bassett and O. Sporns, "Network neuroscience," *Nature neuroscience*, vol. 20, no. 3, p. 353, 2017.
[5] P. Fiedor, "Networks in financial markets based on the mutual information rate," *Phys. Rev. E*, vol. 89, p. 052801, May 2014.
[6] S. P. Borgatti, A. Mehra, D. J. Brass, and G. Labianca, "Network analysis in the social sciences," *science*, vol. 323, no. 5916, pp. 892–895, 2009.
[7] C. Zhu, V. C. Leung, L. Shu, and E. C.-H. Ngai, "Green internet of things for smart world," *IEEE Access*, vol. 3, pp. 2151–2162, 2015.
[8] S. Yang, U. Adeel, Y. Tahir, and J. A. McCann, "Practical opportunistic data collection in wireless sensor networks with mobile sinks," *IEEE Transactions on Mobile Computing*, vol. 16, no. 5, pp. 1420–1433, 2016.
[9] D. Deka, S. Backhaus, and M. Chertkov, "Structure learning in power distribution networks," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 3, pp. 1061–1074, Sept 2018.

[10] J. D. Finkle, J. J. Wu, and N. Bagheri, "Windowed granger causal inference strategy improves discovery of gene regulatory networks," *Proceedings of the National Academy of Sciences*, vol. 115, no. 9, 2018.

[11] M. S. Stankovic, S. S. Stankovic, and K. H. Johansson, "Distributed time synchronization for networks with random delays and measurement noise," *Automatica*, vol. 93, pp. 126 – 137, 2018.

[12] H.-H. Cho, C.-Y. Chen, T. K. Shih, and H.-C. Chao, "Survey on underwater delay/disruption tolerant wireless sensor network routing," *IET Wireless Sensor Systems*, vol. 4, no. 3, pp. 112–121, 2014.

[13] A. S. Leong, S. Dey, and D. E. Quevedo, "Sensor scheduling in variance based event triggered estimation with packet drops," *IEEE Transactions on Automatic Control*, vol. 62, no. 4, pp. 1880–1895, 2017.

[14] H. H. Weerts, P. M. V. den Hof, and A. G. Dankers, "Identifiability of linear dynamic networks," *Automatica*, vol. 89, pp. 247 – 258, 2018.

[15] J. M. Hendrickx, M. Gevers, and A. S. Bazanella, "Identifiability of dynamical networks with partial node measurements," *IEEE Transactions on Automatic Control*, 2018.

[16] D. Materassi and M. V. Salapaka, "Signal selection for estimation and identification in networks of dynamic systems: a graphical model approach," *arXiv preprint arXiv:1905.12132*, 2019.

[17] ——, "On the problem of reconstructing an unknown topology via locality properties of the wiener filter," *IEEE transactions on automatic control*, vol. 57, no. 7, pp. 1765–1777, 2012.

[18] C. J. Quinn, N. Kiyavash, and T. P. Coleman, "Directed Information Graphs," *IEEE Transactions on Information Theory*, vol. 61, no. 12, pp. 6887–6909, 2015.

[19] S. Sinha, P. Sharma, U. Vaidya, and V. Ajjarapu, "Identifying causal interaction in power system: Information-based approach," in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE, 2017.

[20] ——, "On information transfer based characterization of power system stability," *IEEE Transactions on Power Systems*, 2019.

[21] S. Sinha and U. Vaidya, "On data-driven computation of information transfer for causal inference in discrete-time dynamical systems," *Journal of Nonlinear Science*, 2020.

[22] D. Hayden, Y. Yuan, and J. Goncalves, "Network reconstruction from intrinsic noise: Non-minimum-phase systems," *IFAC Proceedings Volumes*, vol. 47, no. 3, 2014.

[23] D. Hayden, Y. Yuan, and J. Gonçalves, "Network reconstruction from intrinsic noise: Minimum-phase systems," in *2014 American Control Conference*. IEEE, 2014.

[24] Y. Yuan, G. B. Stan, S. Warnick, and J. Goncalves, "Robust dynamical network structure reconstruction," *Automatica*, vol. 47, no. 6, pp. 1230 – 1235, 2011, special Issue on Systems Biology.

[25] V. Chetty, D. Hayden, J. Goncalves, and S. Warnick, "Robust signal-structure reconstruction," in *52nd IEEE Conference on Decision and Control*, Dec 2013, pp. 3184–3189.

[26] J. Goncalves and S. Warnick, "Necessary and sufficient conditions for dynamical structure reconstruction of lti networks," *IEEE Transactions on Automatic Control*, vol. 53, no. 7, pp. 1670–1674, Aug 2008.

[27] A. Dankers, P. M. Van den Hof, X. Bombois, and P. S. Heuberger, "Errors-in-variables identification in dynamic networks—consistency results for an instrumental variable approach," *Automatica*, vol. 62, 2015.

[28] N. Everitt, G. Bottegal, and H. Hjalmarsson, "An empirical bayes approach to identification of modules in dynamic networks," *Automatica*, vol. 91, pp. 144 – 151, 2018.

[29] J. Runge, "Causal network reconstruction from time series: From theoretical assumptions to practical estimation," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 28, no. 7, p. 075310, 2018.

[30] V. R. Subramanian, A. Lamperski, and M. V. Salapaka, "Network topology identification from corrupt data streams," in *IEEE 56th Annual Conference on Decision and Control (CDC)*, 2017, pp. 1695–1700.

[31] F. Sepehr and D. Materassi, "Blind learning of tree network topologies in the presence of hidden nodes," *IEEE Transactions on Automatic Control*, 2019.

[32] K. R. Ramaswamy and P. M. J. Vandenhof, "A local direct method for module identification in dynamic networks with correlated noise," *IEEE Transactions on Automatic Control*, 2020.

[33] V. R. Subramanian, A. Lamperski, and M. V. Salapaka, "Inferring directed graphs for networks from corrupt data streams (in progress)," in *IEEE 57th Annual Conference on Decision and Control (CDC)*, 2018.

[34] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*. The MIT Press, 2009.

[35] F. M. Willems, Y. M. Shtarkov, and T. J. Tjalkens, "The context-tree weighting method: basic properties," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 653–664, 1995.

[36] J. Jiao, H. H. Permuter, L. Zhao, Y. Kim, and T. Weissman, "Universal estimation of directed information," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6220–6242, 2013.

[37] G. Kramer, *Directed information for channels with feedback*. Hartung-Gorre, 1998.

[38] J. Adebayo, T. Southwick, V. Chetty, E. Yeung, Y. Yuan, J. Goncalves, J. Grose, J. Prince, G.-B. Stan, and S. Warnick, "Dynamical structure function identifiability conditions enabling signal structure reconstruction," in *IEEE 51st IEEE Conference on Decision and Control (CDC)*. IEEE, 2012.

[39] V. Chetty, J. Eliason, and S. Warnick, "Passive reconstruction of non-target-specific discrete-time lti systems," in *American Control Conference (ACC)*. IEEE, 2016.

[40] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.

Venkat Ram Subramanian received the B.Tech degree in electrical engineering from SRM University, Chennai, India, in 2014, and the M.S. degree in electrical engineering from the University of Minnesota, Minneapolis, in 2016. Currently, he is working towards a Ph.D. degree at the University of Minnesota. His Ph.D. research is on learning dynamic relations in networks from corrupt data-streams. In addition to system identification and stochastic systems, his research interests also include grid modernization and optimal energy management in Distributed Energy Resources (DER).

Andrew Lamperski (S'05–M'11) received the B.S. degree in biomedical engineering and mathematics in 2004 from the Johns Hopkins University, Baltimore, MD, and the Ph.D. degree in control and dynamical systems in 2011 from the California Institute of Technology, Pasadena. He held postdoctoral positions in control and dynamical systems at the California Institute of Technology from 2011–2012 and in mechanical engineering at The Johns Hopkins University in 2012. From 2012–2014, did postdoctoral work in the Department of Engineering, University of Cambridge, on a scholarship from the Whitaker International Program. In 2014, he joined the Department of Electrical and Computer Engineering, University of Minnesota as an Assistant Professor. His research interests include optimal control, optimization, and identification, with applications to neuroscience and robotics.

Murti Salapaka (SM'01–F'19) Murti Salapaka received the bachelor's degree from the Indian Institute of Technology, Madras, India, in 1991, and the Master's and Ph.D. degrees from the University of California, Santa Barbara, CA, USA, in 1993 and 1997, respectively, all in mechanical engineering. He was with Electrical Engineering department, Iowa State University, from 1997 to 2007. He is currently the Vincentine Hermes-Luh Chair Professor with the Electrical and Computer Engineering Department, University of Minnesota, Minneapolis, MN, USA. Prof. Salapaka was the recipient of the NSF CAREER Award and the ISU—Young Engineering Faculty Research Award for the years 1998 and 2001, respectively. He is an IEEE Fellow.