How *else* can we define Information Flow in Neural Circuits?

Praveen Venkatesh*, Sanghamitra Dutta† and Pulkit Grover‡
Electrical & Computer Engineering, and the Center for the Neural Basis of Cognition, Carnegie Mellon University
*vpraveen@cmu.edu †sanghamd@andrew.cmu.edu ‡pulkit@cmu.edu

***Abstract*—Recently, we developed a systematic framework for defining and inferring flows of information about a specific message in neural circuits [2], [3]. We defined a computational model of a neural circuit consisting of computational nodes and transmissions being sent between these nodes over time. We then gave a formal definition of information flow pertaining to a specific message, which was capable of identifying *paths* along which information flowed in such a system. However, this definition also had some non-intuitive properties, such as the existence of "orphans"—nodes from which information flowed out, even though no information flowed in. In part, these non-intuitive properties arose because we restricted our attention to measures that were functions of transmissions at a single time instant, and measures that were *observational* rather than counterfactual. In this paper, we consider alternative definitions, including one that is a function of transmissions at multiple time instants, one that is counterfactual, and new observational definition. We show that a definition of information flow based on counterfactual causal influence (CCI) guarantees the existence of information paths while also having no orphans. We also prove that no observational definition of information flow that satisfies the information path property can match CCI in every instance. Furthermore, each of the definitions we examine (including CCI) is shown to have examples in which the information flow can take a non-intuitive path. Nevertheless, we believe our framework remains more amenable to drawing clear interpretations than classical tools used in neuroscience, such as Granger Causality.**

The full version of this paper is available online [1].

I. Introduction

There is a need to understand how information flows in various kinds of computational systems: particularly in fields such as neuroscience, where we wish to understand the inner workings of the brain [4]–[7], and in artificial neural networks, where we wish to analyze, prune, or assess the trustworthiness of AI systems [8]–[12]. Towards this, we recently proposed a computational model for such neural circuits, and defined a notion of information flow called M-information flow, pertaining to a specific message M in such a system [2], [3]. The primary goal of our previous work was to demonstrate that the intuitive mutual-information–based definition of flow does not satisfy very simple properties: information can "disappear" from the system and reappear at a later time instant, so we cannot always "track" how a message *flows* through the system. This necessitates a more involved definition, which uses conditioning in a particular way, to track the "information paths" along which the message flows.

However, M-information flow also has certain counter-intuitive features: for example, it allows for the existence of "orphans"—nodes from which M-information flows out, though no M-information flows in. This was partly because we chose to restrict ourselves to *observational* measures that are functions of transmissions at a *single* time instant. We did not examine counterfactual measures (which come from the field of causality [13]–[15] and cannot be estimated from passively observed data) and we only superficially examined how a definition based on multiple time instants can be employed.

The core contribution of the current work is an exploration of three alternative definitions of information flow. (i) A version of M-information flow with pruning, which is a function of transmissions at multiple time instants, and is a more detailed analysis of the same definition proposed in our previous work [2] (Section III); (ii) A counterfactual definition that closely matches our intuition in many cases, but cannot be estimated using passively observed data (Section IV); (iii) A modified M-information flow definition based on conditional mutual information, where we allow for functions to be applied to transmissions prior to conditioning—as stated, this is not computable in general, but might be more appealing in some settings (Section VI). We also prove an impossibility result: any observational measure that guarantees information paths will award information flow to edges that counterfactual causal influence would not (Section V).

We note that all three proposed definitions allow us to track information paths while also not having orphans (possibly after pruning). However, each definition we examine has its own shortcomings, giving rise to non-intuitive paths in at least some cases. Recognizing and understanding these shortcomings can help us determine which definition is better suited for a particular purpose. Despite no definition being ideal, this systematic framework lends itself much better to drawing clear interpretations than classical tools used in neuroscience, such as Granger Causality. We revisit this point in Section VII.

II. Background

We begin with a short recap of our computational system model and the definition of information flow about a message M discussed in [2]. We also restate two important properties of our M-information flow definition: firstly, that it guarantees the existence of "information paths" along which information about the message flows in the system; and secondly, that it suffers from "orphans". The definitions as well as the counterexample in this section are largely replicated from our previous work [2] with only minor modifications, in order to keep this paper self-contained.

A. The Computational System Model

Definition 1 (Time-unrolled graph): Let $\mathcal{G}^* = (\mathcal{V}^*, \mathcal{E}^*)$ be a fully-connected directed graph with N nodes, i.e., $\mathcal{V}^* = \{1, 2, \ldots, N\}$ and $\mathcal{E}^* = \mathcal{V}^* \times \mathcal{V}^*$. Also, let $\mathcal{T} = \{0, 1, \ldots, T\}$ be a set of time indices, where T is a positive integer representing the maximum time index. Then, a *time-unrolled graph* $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is constructed by indexing a fully-connected directed graph $\mathcal{G}^*$ using the time indices $\mathcal{T}$ as follows: (i) The nodes $\mathcal{V}$ consist of all nodes $\mathcal{V}^*$ in $\mathcal{G}^*$, subscripted by time indices $\mathcal{T}$, i.e., $\mathcal{V} = \{A_t : A \in \mathcal{V}^*, t \in \mathcal{T}\}$; (ii) The edges $\mathcal{E}$ connect nodes of *successive* times in $\mathcal{V}$, so they can be written in terms of the edges in $\mathcal{E}^*$ as $\mathcal{E} = \{(A_t, B_{t+1}) : (A, B) \in \mathcal{E}^*, t \in \mathcal{T}\}$. □

Remarks: (i) We denote the set of all nodes at time t by $\mathcal{V}_t$, and the set of all (outgoing) edges at time t by $\mathcal{E}_t$. So, for example, we will have $A_1 \in \mathcal{V}_1$ and $(A_1, B_2) \in \mathcal{E}_1$. (ii) The original fully-connected graph $\mathcal{G}^*$ has self-edges, so the time-unrolled graph will always have an edge (A_t, A_{t+1}) in $\mathcal{E}_t$ for every node $A_t \in \mathcal{V}_t$.

Definition 2 (Computational System): A *computational system* $\mathcal{C} = (\mathcal{G}, X, W, f)$ is a time-unrolled graph $\mathcal{G}$ that has *transmissions on its edges* which are constrained by *computations at its nodes*. The *input nodes* of the computational system compute a function of a *message*, M. We now elaborate upon these italicized terms:

2a) Transmissions on Edges

In a time-unrolled graph $\mathcal{G}$, let $X : \mathcal{E} \to \mathfrak{X}$ be a function that describes what random variable is being transmitted on a given edge, i.e., $X(E)$ is the random variable corresponding to the transmission on the edge E. Here, the range $\mathfrak{X}$ is the set of all random variables in some probability space.

For convenience, we define X applied to a *set of edges* as the set of random variables produced by applying X to each of those edges individually, i.e., for *any* subset $\mathcal{E}' \subseteq \mathcal{E}$,

$$X(\mathcal{E}') = \{X(E) : E \in \mathcal{E}'\}. \tag{1}$$

We extend the use of this notation to other functions of nodes and edges that we define, going forward.

2b) Computation at a Node

Let $A_t \in \mathcal{V}_t$ be a node in the time-unrolled graph $\mathcal{G}$, at some time $t \geq 1$ (recall that $t \in \{0, 1, \ldots, T\}$). Let $\mathcal{P}(A_t)$ be the set of edges entering A_t, and $\mathbb{Q}(A_t)$ be the set of edges leaving A_t. Further, let us suppose that A_t is able to intrinsically generate the random variable $W(A_t)$ at time t, where $W(A_t) \perp\!\!\!\perp W(\mathcal{V} \backslash \{A_t\}) \; \forall \; A_t \in \mathcal{V}$, $W(\mathcal{V}_t) \perp\!\!\!\perp \{M, X(\mathcal{E}_{t-1})\}$ and the symbol "$\perp\!\!\!\perp$" stands for independence between random variables.[1] Then, the *computation* performed by the node A_t (for $t \geq 1$) is a deterministic function f_{A_t} that satisfies

$$f_{A_t}\big(X(\mathcal{P}(A_t)), W(A_t)\big) = X(\mathbb{Q}(A_t)). \tag{2}$$

Here, $X(\mathcal{E}_{t-1})$, $W(\mathcal{V} \setminus \{A_t\})$, $W(\mathcal{V}_t)$, $X(\mathcal{P}(A_t))$ and $X(\mathbb{Q}(A_t))$ all make use of the notation described in (1). Note that the definition above does not apply when $t = 0$; this is a special case which is discussed below.

[1] Strictly speaking, we require that M is not an ancestor of any $W(V_t)$ in the structural causal model underlying the computational system, i.e., interventions on M will not affect $W(V_t)$, even in a counterfactual setting [14].

2c) The Message and the Input Nodes

The *message* is a random variable M, which is of interest to the observer, and for which we shall define information flow. We assume that the message enters the computational system at (and only at) time $t = 0$. We formally define the *input nodes* of the system as those nodes of $\mathcal{G}$, at time $t = 0$, whose transmissions statistically depend on the message M: $\mathcal{V}_{\text{ip}} \coloneqq \{A_0 \in \mathcal{V}_0 : I\big(M; X(\mathbb{Q}(A_0))\big) > 0\}$, where $\mathbb{Q}(A_0)$ represents the set of edges leaving the node A_0.

To remain consistent with Definition 2b, we define the computation performed by an input node $A_0 \in \mathcal{V}_{\text{ip}}$ as a function f_{A_0} that satisfies $f_{A_0}\big(M, W(A_0)\big) = X(\mathbb{Q}(A_0))$, and the computation performed by a non-input node at time $t = 0$, $A_0 \in \mathcal{V}_0 \backslash \mathcal{V}_{\text{ip}}$, as a function f_{A_0} that satisfies $f_{A_0}\big(W(A_0)\big) = X(\mathbb{Q}(A_0))$. As before, $W(A_0) \perp\!\!\!\perp W(\mathcal{V}_0 \backslash \{A_0\}) \; \forall \; A_0 \in \mathcal{V}_0$ and $W(\mathcal{V}_0) \perp\!\!\!\perp M$. □

B. Defining Information Flow

Definition 3 (M-information Flow): We say that an edge $E_t \in \mathcal{E}_t$ has M-information flow if

$$\exists \; \mathcal{E}_t' \subseteq \mathcal{E}_t \setminus \{E_t\} \quad \text{s.t.} \quad I\big(M; X(E_t) \,|\, X(\mathcal{E}_t')\big) > 0. \tag{3}$$

Analogously, a *collection* of edges at the same time instant, $\mathcal{R}_t \subseteq \mathcal{E}_t$, is said to have M-information flow if

$$\exists \; \mathcal{E}_t' \subseteq \mathcal{E}_t \setminus \mathcal{R}_t \quad \text{s.t.} \quad I\big(M; X(\mathcal{R}_t) \,|\, X(\mathcal{E}_t')\big) > 0. \tag{4}$$

That is, we say an edge E_t (at time t) has M-information flow if, *conditioned* on the transmissions of some subset $\mathcal{E}_t'$ also at time t, $X(E_t)$ has mutual information with M. The rationale behind this is explained after Counterexample 1. □

Note: Henceforth, "information flow about M" may refer to *any* measure of information flow, but M-information flow refers specifically to Definition 3.

C. The Information Path Property

Definition 4 (Path): In any computational system $\mathcal{C}$, suppose $\mathcal{A}$ and $\mathcal{B}$ are two disjoint sets of nodes in $\mathcal{V}$. Then, a *path* from $\mathcal{A}$ to $\mathcal{B}$ is any ordered set of nodes $\{V^{(0)}, V^{(1)}, \ldots, V^{(L)}\}$ that satisfies (i) $V^{(0)} \in \mathcal{A}$; (ii) $V^{(L)} \in \mathcal{B}$; and (iii) $(V^{(i-1)}, V^{(i)}) \in \mathcal{E}$ for every $1 \leq i \leq L$, where L is a positive integer indicating the length of the path. We refer to the set $\{(V^{(i-1)}, V^{(i)})\}_{i=1}^{L}$ as the *edges of the path*. □

Definition 5 (M-Information Path): An *M-information path* from $\mathcal{A}$ to $\mathcal{B}$ is a *path* from $\mathcal{A}$ to $\mathcal{B}$, every edge of which carries information flow about M. □

Property 1 (Existence of an Information Path): In any computational system $\mathcal{C}$, suppose that at some time $t_{\text{op}} \in \mathcal{T}$, there is an "output node" $V_{\text{op}} \in \mathcal{V}$ whose outgoing edges $\mathbb{Q}(V_{\text{op}})$ satisfy $I\big(M; X(\mathbb{Q}(V_{\text{op}}))\big) > 0$. Then, there must exist an M-information path from the input nodes $\mathcal{V}_{\text{ip}}$ to V_{op}.

Theorem 1: Definition 3 satisfies Property 1.

The proof of this theorem was one of the main contributions of our earlier work, and can be found in [2]. We have reiterated the theorem statement alone for completeness.

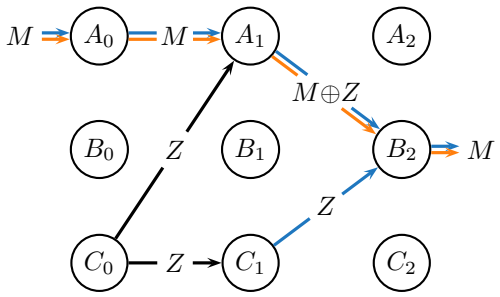

Fig. 1: The computational system for Counterexample 1, which also appeared in our previous work [2] (to avoid clutter, only edges relevant to the counterexample are depicted; all other edges are still present and their transmissions are assumed to be zero). Edges in blue have M-information flow (Definition 3) and those in orange are M-CCI'd (as described later in Section IV). Observe that the edges with $M \oplus Z$ as well as Z at time $t = 1$ have M-information flow as per Definition 3. This results in an orphan at C_1, since the only incoming edge of C_1 does not have M-information flow.

D. The No-Orphans Property

As pointed out in our earlier work [2], Definition 3 also has a very non-intuitive property: the existence of orphans.

Definition 6 (M-information Orphan): In a computational system $\mathscr{C}$, a node V_t is said to be an *M-information orphan* if its outgoing edges $\mathbb{Q}(V_t)$ have information flow about M, but its incoming edges $\mathscr{P}(V_t)$ do not. □

Property 2 (Absence of Orphans): M-information orphans must not exist in a computational system.

M-information flow (Definition 3) does *not* satisfy Property 2. This is illustrated by the following counterexample.

Counterexample 1: Consider the computational system depicted in Figure 1 (note that, in order to avoid unnecessary clutter, only edges with non-zero transmissions are shown in the figure). A_0 is the input node, which has the message $M \sim \text{Ber}(1/2)$ at time $t = 0$. The system's goal is to communicate M to the node B. It chooses the following strategy: at $t = 0$, A_0 transmits M to A_1. C_0 independently generates a different random number, $W(C_0) = Z \sim \text{Ber}(1/2)$, $Z \perp\!\!\!\perp M$, and sends this message to A_1, as well as C_1. A_1 then computes $M \oplus Z$ and passes the result to B_2, while C_1 sends Z to B_2. Here, the symbol "$\oplus$" stands for XOR, the exclusive-OR operator on two bits. B_2 is thus able to recover M by once again XOR-ing its inputs, $(M \oplus Z)$ and Z.

The edges shown in blue carry M-information flow: the edges transmitting M naturally carry M-information flow; even though $M \oplus Z$ and Z do *not* statistically depend on the message, they *conditionally* depend on the message given the other (recall Definition 3). That is, $I(M; M \oplus Z \,|\, Z) > 0$, and complementarily, $I(M; Z \,|\, M \oplus Z) > 0$. Hence, they *also* carry M-information flow.

Observe that the node C_1 is an M-information orphan, since the edge (C_1, B_2), transmitting Z, has M-information flow, but none of C_1's incoming edges have M-information flow. ■

Remark: Counterexample 1 essentially shows why Definition 3 is needed: a simpler definition that awards information flow to E_t if $I(M; X(E_t)) > 0$ would fail to identify the information path, because $M \oplus Z \perp\!\!\!\perp M$. The edge carrying $M \oplus Z$ thus plays the important role of maintaining the M-information path from A_0 to B_2 in this example.

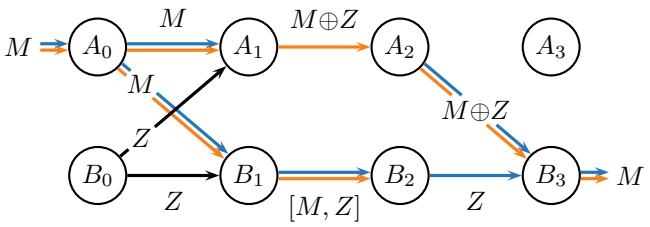

Fig. 2: The computational system corresponding to Counterexample 2, which demonstrates that pruning does not always remove edges with Z. Edges in blue have M-information flow per Definition 3 and those in orange are M-CCI'd (as described later in Section IV). Counterintuitively, in this example, the edge with $M \oplus Z$ does *not* carry M-information flow per Definition 3.

The existence of M-information flow on (C_1, B_2) (and hence the existence of M-information orphans) might seem rather counterintuitive in a way that M-information flow on $M \oplus Z$ does not. We likely feel this way because Z was never *computed* from M. In this sense, Z lacks some kind of "functional dependence" on M, which $M \oplus Z$ does not. This point is examined in greater detail from a causality perspective in Section IV. In the following section, we consider a simple pruning-based mechanism and determine whether this removes orphans and edges that transmit only Z.

III. M-Information Flow with Pruning

One way to avoid orphans might be to consider transmissions at more than one time instant when defining information flow: for instance, we could check for information flow at a previous time instant before assigning flow to a particular edge. The principled way to do this is to traverse paths backward from the output node to the input node, while systematically pruning all "stray" paths that lead to orphans. This process is described in the form of an Information Path Algorithm in [2, Section 5]. The algorithm relies on the fact that Definition 3 satisfies the information path property, so that a path leading backwards from the output node to the input nodes is always guaranteed to exist.

However, while this pruning mechanism removes orphans, it does not always remove edges like Z, which do not "functionally depend" on the message M. We next present a counterexample where an edge with $M \oplus Z$ is removed, instead of the edge with Z. It should be noted that this is a highly counterintuitive example, and is very unlikely to occur as such in practice. Nevertheless it shows that even with pruning, M-information flow is not completely devoid of shortcomings.

Counterexample 2 (Pruning does not remove Z-edges): Consider the computational system shown in Figure 2. Here, the message M is being communicated from A_0 to B_3 in the following manner: A_0 sends M to both A_1 and B_1, while B_0 generates $Z \sim \text{Ber}(1/2)$, $Z \perp\!\!\!\perp M$, and sends it to A_1 and B_1. The node A_1 then computes $M \oplus Z$ and passes it on to B_3 through A_2, while B_1 simply concatenates M and Z into a vector, $[M, Z]$ and sends it to B_2. B_2 then discards M, and passes on Z to B_3.

The result of this setup is that the edges shown in blue have M-information flow. In particular, the edge (A_1, A_2) carrying $M \oplus Z$, does *not* carry M-information flow: this is because $M \oplus Z$ does not depend on M by itself, and when conditioned on $[M, Z]$, naturally, M is treated as a constant and thus any mutual information with M goes to zero, i.e., $I(M; M \oplus Z \,|\, M, Z) = 0$. Thus, the only information path from the input node, A_0, to the output node, B_3, is the one that includes the edge (B_2, B_3), whose transmission is Z. In other words, if we were to prune edges that did not lead back to the input node A_0, we would end up removing (A_2, B_3), while (B_2, B_3), which carries Z, would remain intact. ■

The existence of such a counterexample makes the information path theorem proved in [2] all the more interesting and surprising. However, it also raises several questions: on the one hand, the existence of orphans seemed counterintuitive, because their outgoing transmissions seemed to "have nothing to do with the message M"; while on the other, Counterexample 2 shows that even the removal of orphans does not guarantee the removal of edges with such transmissions. This makes it all the more important to focus on such edges: how are we able to *intuitively* distinguish between transmissions that in some crude sense "functionally depend" on the message M (such as $M \oplus Z$), and those that do not (e.g. Z)? We argue that the answer to this question lies in the realm of causality, in a concept known as counterfactual causal influence.

IV. Counterfactual Causal Influence

Counterfactual causal influence [8]–[16] intuitively asks the question: for a particular *realization* of all random variables in the system, if M had instead taken a different value keeping everything else the same, how would the value of some other variable have changed? This turns out to be the key to formally understanding the intuitive notion of "functional dependence" discussed above. In this section, we show that a definition of information flow based on counterfactual causal influence satisfies the information path property while at the same time having no orphans.

Definition 7 (M-counterfactual causal influence): The transmission on some edge E_t can be written in terms of M and all past intrinsic random variables, $\underline{W}_t \coloneqq \cup_{\tau \le t} W(\mathcal{V}_\tau)$ as

$$X(E_t) = g(M, \underline{W}_t), \tag{5}$$

for some function g. Then, $X(E_t)$ (or equivalently, E_t) is said to be *counterfactually causally influenced by M (M-CCI'd)* if for some potential realization $\underline{w}_t$ of $\underline{W}_t$,

$$\exists\, m, m' \quad \text{s.t.} \quad g(m, \underline{w}_t) \neq g(m', \underline{w}_t). \tag{6}$$

M-CCI constitutes a definition of information flow in that it can be treated as an *indicator* of information flow about M on the edge E_t. The definition of M-CCI may also be applied in the same way to variables other than transmissions on edges. □

Theorem 2: M-CCI (Definition 7) satisfies Property 2, i.e., it does not give rise to M-information orphans. In other words, if at any node V_t, there exists an outgoing edge $E_t \in \mathbb{Q}(V_t)$ that is M-CCI'd, then there exists some incoming edge, $E'_{t-1} \in \mathcal{P}(V_t)$, which is also M-CCI'd.

Theorem 3: M-CCI (Definition 7) satisfies Property 1, i.e., it guarantees the existence of M-information paths. That is, if there is some "output node" $V_{\text{op}} \in \mathcal{V}$ that satisfies $I\big(M; X(\mathbb{Q}(V_{\text{op}}))\big) > 0$, then there exists a path from $\mathcal{V}_{\text{ip}}$ to V_{op} such that every edge of this path is M-CCI'd.

We defer the proofs to Appendix A, which appears in the full version of this document [1]. A brief combined proof outline for both theorems is provided below.

Proof outline for Theorems 2 and 3:

1) Link M-CCI for a single edge with that for a set of edges: if no edge in a set is individually M-CCI'd, then the set of all edges is not M-CCI'd. The converse is also true.
2) Show using Definition 2b that if the set of all incoming edges is not M-CCI'd, then the set of all outgoing edges is not M-CCI'd. Thus, no individual outgoing edge is M-CCI'd (by the converse in the previous point). With this, the contrapositive of Theorem 2 is proved.
3) Prove that if an edge is not M-CCI'd, then its transmission can have no mutual information with M.
4) Then, work backwards from the output node in Theorem 3 by recursively using Theorem 2 to show that an information path to the input nodes exists. This proves Theorem 3. ■

As shown by the orange edges in Figs. 1 and 2, M-CCI captures the *intuitively correct* edges in these examples, e.g., $M \oplus Z$ is considered to have information flow based on M-CCI, while Z is not. This raises the question of how close we can get to M-CCI with purely observational measures, which we address in the very next section. However, we should also note here that M-CCI is not without caveats: it is not observational (i.e., cannot be estimated from passively observed data) and it can produce information paths that could be considered spurious (as we will show in Example 3).

V. The Limitations of Observational Measures

In this section, we prove an impossibility result which shows that *no* observational measure that satisfies the information path property can be made to assign information flow only to edges that are M-CCI'd. First, we formally define what we mean by observational measures.

Definition 8 (Observational measures of information flow): A definition of information flow is said to be *observational* if it depends only on samples of $X(\mathcal{E})$ and M. In effect, the measure depends only on the joint distribution $p(X(\mathcal{E}), M)$, which we assume can be estimated from data. □

In contrast, interventional and counterfactual measures require knowledge outside of the joint distribution $p(X(\mathcal{E}), M)$: we must also know how the joint distribution *changes* when one or more variables are intervened upon, or held fixed to a constant value. We next state the impossibility result, deferring its proof to Appendix B, which can be found in the full version [1].

Theorem 4: Any observational definition of information flow on the edge E_t that satisfies the information path property will, in some instances, assign information flow to edges that are not M-CCI'd.

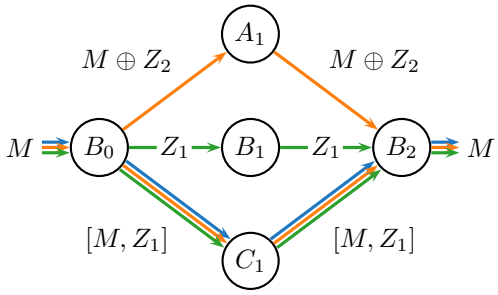

Fig. 3: The computational system from Example 3, showing the differences between Definitions 3, 7 and 9. Edges in blue, orange and green respectively have information flow as per Definitions 3, 7 and 9.

VI. One More Definition and an Example

Theorem 4 shows that observational measures are limited in that either they will not satisfy the information path property, or there will be instances where they award information flow to edges that are not M-CCI'd. However, we can ask if there are measures that satisfy the information path theorem, while at the same time provide more intuitive results upon pruning—e.g., measures that do not suffer from the counterintuitive problem discussed in Counterexample 2. In that spirit, we provide one more observational definition of information flow and show how it overcomes the problem discussed in Counterexample 2. Finally, we provide an example that brings out the differences between the three definitions presented here, and discuss their pros and cons.

Definition 9 (Modified M-information flow): We say that an edge E_t has modified M-information flow if there exists some subset of edges $\mathscr{E}'_t \subseteq \mathscr{E}_t \setminus \{E_t\}$, $\mathscr{E}'_t = \{E_t^{(i)}\}_{i=1}^k$ and some set of functions $\{h_i\}_{i=1}^k$ such that

$$I\big(M; X(E_t) \,\big|\, h_1(X(E_t^{(1)})), \ldots, h_k(X(E_t^{(k)}))\big) > 0. \quad (7)$$

In other words, an edge E_t has modified M-information flow, if there exist some other edges at time t, such that when conditioned on some *functions* of their individual transmissions, $X(E_t)$ has mutual information with M. □

Every edge that has M-information flow (Definition 3) also has modified M-information flow, since Definition 9 immediately reduces to Definition 3 if we restrict all h_i to be identity functions. However, the opposite is not true. Consider Fig. 2 for example: here, all blue edges, as well as the edge (A_1, A_2), will have modified M-information flow. This is because there exists a function of $[M, Z]$ (namely, $h([M, Z]) := Z$), such that when $M \oplus Z$ is conditioned on $h([M, Z])$, we get non-zero mutual information with M. Thus, we may be avoiding some of the more non-intuitive corner cases in which M-information flow does not supply the "intuitively correct" information path.

Modified M-information flow also suffers from many of the same drawbacks as M-information flow: it still has orphans (e.g., in Fig. 1, only blue edges have modified M-information flow, so C_1 will be an orphan). Furthermore, as stated, Definition 9 is not computable, as the range of the h_i can be arbitrarily large in dimension.

Example 3 (All definitions are imperfect): We use one last example to show that M-CCI and modified M-information flow are also not perfect, and to bring out their differences. Consider the computational system shown in Fig. 3. We take $M, Z_1, Z_2 \sim$ i.i.d. Ber$(1/2)$. Note that $M \oplus Z_2$ is M-CCI'd; however, since Z_2 no longer persists in the system, all information about M has been destroyed through the XOR with Z_2. In other words, M-CCI identifies an information path which can have no computational value whatsoever.

On the other hand, the edges with Z_1 have modified M-information flow, because $[M, Z_1]$ admits the function $M \oplus Z_1$. But since Z_1 does not interact with M (save possibly *within* the node B_2), it could be argued that these edges should not carry information flow about M either.

Example 3 also shows that there can be M-CCI'd edges that do not have (either original [2] or modified) M-information flow; edges *not* M-CCI'd but that *have* modified (and possibly original) M-information flow; and edges that have all three. ■

VII. Discussion and Conclusion

Choosing the right definition for a particular quantity is often a hard task, and might be problem- and context-dependent, as evidenced by the multitude of definitions for entropy [17], [18]. The choice of definition is also often dictated by the trade-offs that we are willing to live with. In the case of information flow, if we are in a setting where we can examine counterfactual effects (e.g., when simulating an artificial neural network), then M-CCI provides an intuitive definition, with the caveat that it may also identify some irrelevant edges. On the other hand, if we can only make observational measurements, then M-information flow with pruning goes a long way, save for some corner cases (such as Counterexample 2). We hope that these holes are also plugged when using modified M-information flow, especially in conjunction with a pruning algorithm that can remove orphans. Further work is needed to understand if there are other instances where modified M-information flow succeeds or fails in some important way.

Ultimately, it should be noted that this systematic framework for information flow, while not providing a single answer, still overcomes many of the fundamental challenges faced by classical techniques used for examining information flow. In the neuroscientific literature, Granger causality [5], [19]–[21] has long been used as a heuristic measure of information flow, despite several criticisms [22]–[29], including the well-known fact that it is not truly representative of causation [13]. Indeed, interpreting Granger causal influence as information flow may also be questionable, as we have shown in past work [2], [29]. Given the systematic approach we have taken in defining information flow here, a natural question that arises is what connection our definition has to true causation. Our results imply that an edge has information flow about M per any of our three definitions, if *some* intervention on M can change the marginal distribution of a transmission. Similarly, edges whose transmissions statistically depend (unconditionally) on the message have information flow according to all three definitions, meaning that they are also M-CCI'd.

REFERENCES

[1] Full version of this paper with appendices. [Online]. Available: https://praveenv253.github.io/assets/doc/papers/2020--isit--full-paper.pdf

[2] P. Venkatesh, S. Dutta, and P. Grover, "Information flow in computational systems," *arXiv:1902.02292 [cs.IT]*, 2019.

[3] ——, "How should we define information flow in neural circuits?" in *2019 IEEE International Symposium on Information Theory (ISIT)*, July 2019, pp. 176–180.

[4] J. Almeida, A. R. Fintzi, and B. Z. Mahon, "Tool manipulation knowledge is retrieved by way of the ventral visual object processing pathway," *Cortex*, vol. 49, no. 9, pp. 2334–2344, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0010945213001329

[5] A. Brovelli, M. Ding, A. Ledberg, Y. Chen, R. Nakamura, and S. L. Bressler, "Beta oscillations in a large-scale sensorimotor cortical network: Directional influences revealed by Granger causality," *Proceedings of the National Academy of Sciences*, vol. 101, no. 26, pp. 9849–9854, 2004. [Online]. Available: http://www.pnas.org/content/101/26/9849

[6] M. Bar, K. S. Kassam, A. S. Ghuman, J. Boshyan, A. M. Schmid, A. M. Dale, M. S. Hämäläinen, K. Marinkovic, D. L. Schacter, B. R. Rosen, and E. Halgren, "Top-down facilitation of visual recognition," *Proceedings of the National Academy of Sciences*, vol. 103, no. 2, pp. 449–454, 2006. [Online]. Available: https://www.pnas.org/content/103/2/449

[7] A. S. Greenberg, T. Verstynen, Y.-C. Chiu, S. Yantis, W. Schneider, and M. Behrmann, "Visuotopic cortical connectivity underlying attention revealed with white-matter tractography," *Journal of Neuroscience*, vol. 32, no. 8, pp. 2773–2782, 2012. [Online]. Available: http://www.jneurosci.org/content/32/8/2773

[8] S. Dutta, P. Venkatesh, P. Mardziel, A. Datta, and P. Grover, "An information-theoretic quantification of discrimination with exempt features," in *Proceedings of the AAAI Conference on Artificial Intelligence (To Appear)*, 2020. [Online]. Available: https://sites.google.com/site/sanghamitraweb/academic-articles

[9] A. Datta, S. Sen, and Y. Zick, "Algorithmic transparency via quantitative input influence: Theory and experiments with learning systems," in *IEEE Symposium on Security and Privacy*, 2016, pp. 598–617.

[10] P. W. Koh and P. Liang, "Understanding black-box predictions via influence functions," in *Proceedings of the 34th International Conference on Machine Learning*, vol. 70. JMLR.org, 2017, pp. 1885–1894.

[11] P. Adler, C. Falk, S. A. Friedler, T. Nix, G. Rybeck, C. Scheidegger, B. Smith, and S. Venkatasubramanian, "Auditing black-box models for indirect influence," *Knowledge and Information Systems*, vol. 54, no. 1, pp. 95–122, 2018.

[12] A. Henelius, K. Puolamäki, H. Boström, L. Asker, and P. Papapetrou, "A peek into the black box: exploring classifiers by randomization," *Data mining and knowledge discovery*, vol. 28, no. 5-6, pp. 1503–1529, 2014.

[13] J. Pearl, *Causality: Models, Reasoning and Inference*. Cambridge University Press, 2009.

[14] J. Peters, D. Janzing, and B. Schölkopf, *Elements of causal inference: Foundations and learning algorithms*. MIT press, 2017.

[15] M. J. Kusner, J. Loftus, C. Russell, and R. Silva, "Counterfactual fairness," in *Advances in Neural Information Processing Systems*, 2017, pp. 4066–4076.

[16] C. Russell, M. J. Kusner, J. Loftus, and R. Silva, "When worlds collide: integrating different counterfactual assumptions in fairness," in *Advances in Neural Information Processing Systems*, 2017, pp. 6414–6423.

[17] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, July 1948.

[18] A. Rényi, "On measures of entropy and information," in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics*. The Regents of the University of California, 1961.

[19] C. W. J. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica: Journal of the Econometric Society*, pp. 424–438, 1969.

[20] S. L. Bressler and A. K. Seth, "Wiener–Granger causality: A well established methodology," *NeuroImage*, vol. 58, no. 2, pp. 323–329, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1053811910002272

[21] M. Kamiński, M. Ding, W. A. Truccolo, and S. L. Bressler, "Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance," *Biological Cybernetics*, vol. 85, no. 2, pp. 145–157, Aug 2001. [Online]. Available: https://doi.org/10.1007/s004220000235

[22] O. David, I. Guillemain, S. Saillet, S. Reyt, C. Deransart, C. Segebarth, and A. Depaulis, "Identifying neural drivers with functional MRI: an electrophysiological validation," *PLoS biology*, vol. 6, no. 12, p. e315, 2008.

[23] A. Roebroeck, E. Formisano, and R. Goebel, "The identification of interacting networks in the brain using fMRI: model selection, causality and deconvolution," *Neuroimage*, vol. 58, no. 2, pp. 296–302, 2011.

[24] O. David, "fMRI connectivity, meaning and empiricism. comments on: Roebroeck et al. The identification of interacting networks in the brain using fMRI: model selection, causality and deconvolution." *Neuroimage*, vol. 58, no. 2, pp. 306–309, 2011.

[25] P. A. Stokes and P. L. Purdon, "A study of problems encountered in Granger causality analysis from a neuroscience perspective," *Proceedings of the National Academy of Sciences*, vol. 114, no. 34, pp. E7063–E7072, 2017. [Online]. Available: http://www.pnas.org/content/114/34/E7063

[26] J. Andersson, "Testing for Granger causality in the presence of measurement errors," *Economics Bulletin*, 2005.

[27] H. Nalatore, M. Ding, and G. Rangarajan, "Mitigating the effects of measurement noise on Granger causality," *Physical Review E*, vol. 75, no. 3, p. 031123, Mar 2007. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.75.031123

[28] M. Gong, K. Zhang, B. Schölkopf, D. Tao, and P. Geiger, "Discovering temporal causal relations from subsampled data," in *Proceedings of The 32nd International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, vol. 37. PMLR, Jul 2015, pp. 1898–1906. [Online]. Available: http://proceedings.mlr.press/v37/gongb15.html

[29] P. Venkatesh and P. Grover, "Is the direction of greater Granger causal influence the same as the direction of information flow?" in *53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Sept 2015, pp. 672–679.

APPENDIX A
PROOFS FROM SECTION IV

A. Proof of Theorem 2

We first prove a simple lemma, which connects M-CCI for a single edge and for a set of edges.

Lemma 5: For any set $\mathscr{E}'_t \subseteq \mathscr{E}_t$, if there exists some edge $E_t \in \mathscr{E}'_t$ which is M-CCI'd, then $X(\mathscr{E}'_t)$ is also M-CCI'd. The converse is also true.

Proof: We start by enumerating the edges in $\mathscr{E}'_t$. Suppose $|\mathscr{E}'_t| =: k$. Then, we can write $\mathscr{E}'_t = \{E_t^{(i)}\}_{i=1}^k$. Now, we note that the set $X(\mathscr{E}'_t)$ is simply the collection of all transmissions in $\mathscr{E}'_t$. Therefore, we can write

$$X(\mathscr{E}'_t) = \big\{X(E_t^{(i)}) : E_t^{(i)} \in \mathscr{E}'_t\big\} \tag{8}$$

$$= \big\{g_{X(E_t^{(i)})}(M, \underline{W}_t) : E_t^{(i)} \in \mathscr{E}'_t\big\} \tag{9}$$

$$=: h(M, \underline{W}_t), \tag{10}$$

where $g_{X(E_t^{(i)})}$ is as defined in Definition 7 and h is a function that can be written in terms of the $\{g_{X(E_t^{(i)})}\}$. Now, if any one $E_t^{(j)} \in \mathscr{E}'_t$ is M-CCI'd, then there will be some set of values m, m' and $\underline{w}_t$ such that $g_{X(E_t^{(j)})}(m, \underline{w}_t) \neq g_{X(E_t^{(j)})}(m', \underline{w}_t)$. Thus, $h(m, \underline{w}_t) \neq h(m', \underline{w}_t)$, and hence $\mathscr{E}'_t$ is M-CCI'd.

Conversely, if no edge $E_t \in \mathscr{E}'_t$ is M-CCI'd, we would have $g_{X(E_t^{(i)})}(m, \underline{w}_t) = g_{X(E_t^{(i)})}(m', \underline{w}_t)\ \forall\ m, m', \underline{w}_t$. Hence, it follows that $h(m, \underline{w}_t) = h(m', \underline{w}_t)\ \forall\ m, m', \underline{w}_t$. Thus $X(\mathscr{E}'_t)$ is not M-CCI'd. This proves the lemma. ∎

Remark: Lemma 5 might seem trivial, at least in the case of M-CCI, but it is actually a crucial step in the proof of the information path property. In particular, the equivalent of Lemma 5 does not hold for mutual information in the converse, i.e., it is *not* true that if $X(\mathscr{E}'_t)$ has non-zero mutual information with M, then some edge $E_t \in \mathscr{E}'_t$ also has non-zero mutual information with M. Two edges' transmissions may individually have no mutual information about M, while jointly having non-zero mutual information about M. The failure of this lemma is the reason that a definition of information flow based on mutual information (as mentioned in Counterexample 1) does not satisfy the information path property.

Proof of Theorem 2: For there to be no orphans, the following must hold: at any node V_t, if there exists an outgoing edge $E_t \in \mathbb{Q}(V_t)$ that is M-CCI'd, then there exists some incoming edge, $E'_{t-1} \in \mathscr{P}(V_t)$, which is also M-CCI'd.

First, note that if all incoming edges of V_t are *not* M-CCI'd, i.e. E_{t-1} is not M-CCI'd $\forall\ E_{t-1} \in \mathscr{P}(V_t)$, then the *set* of incoming edges $\mathscr{P}(V_t)$ is not M-CCI'd. This is a direct consequence of the converse of Lemma 5.

Next, recall from Definition 2b that $X(\mathbb{Q}(V_t)) = f_{V_t}\big(X(\mathscr{P}(V_t)), W(V_t)\big)$. We have already shown that $\mathscr{P}(V_t)$ is not M-CCI'd, and since M is not an ancestor of $W(V_t)$ in the structural causal model (SCM) corresponding to the computational system (see footnote 1), $W(V_t)$ is also not M-CCI'd. Thus, $\mathbb{Q}(V_t)$ is not M-CCI'd. Therefore, by Lemma 5, no individual outgoing edge, $E_t \in \mathbb{Q}(V_t)$, can be M-CCI'd.

Hence, by the contrapositive of the above statements, if there *is*, in fact, some outgoing edge of V_t, $E_t \in \mathbb{Q}(V_t)$, that is M-CCI'd, then there must also be an incoming edge, $E'_{t-1} \in \mathscr{P}(V_t)$, that is M-CCI'd. ∎

B. Proof of Theorem 3

Again, we first prove a simple lemma which links M-CCI with mutual information.

Lemma 6: If some variable $Y := h(M, \underline{W})$ is not M-CCI'd (where $\underline{W}$ does not have M as an ancestor in the SCM corresponding to the computational system), then $I(M; Y) = 0$.

Proof: Since Y is not M-CCI'd, we have that

$$h(m, \underline{w}) = h(m', \underline{w}) \quad \forall\ m, m', \underline{w}, \tag{11}$$

where $\underline{w}$ takes values in the set of possible realizations of the random variable $\underline{W}$. Thus, h is effectively independent of M, and we can write

$$Y = h(M, \underline{W}) =: h_0(\underline{W}). \tag{12}$$

Assuming all distributions are discrete, we can use summations to write:

$$p_{Y,M}(y, m) = \sum_{\underline{w}} p_{Y,M,\underline{W}}(y, m, \underline{w}) \tag{13}$$

$$= \sum_{\underline{w}} p_{Y|M,\underline{W}}(y \,|\, m, \underline{w}) p_{M,\underline{W}}(m, \underline{w}) \tag{14}$$

$$= \sum_{\underline{w}} \delta\big(y, h(m, \underline{w})\big) p_{M,\underline{W}}(m, \underline{w}) \tag{15}$$

$$= \sum_{\underline{w}} \delta\big(y, h_0(\underline{w})\big) p_M(m) p_{\underline{W}}(\underline{w}) \tag{16}$$

$$= p_M(m) \sum_{\underline{w}} \delta\big(y, h_0(\underline{w})\big) p_{\underline{W}}(\underline{w}) \tag{17}$$

$$=: p_M(m) c(y) \tag{18}$$

where in the above, δ is the Kronecker Delta function, which takes a value of 1 when its arguments are equal, and zero otherwise. In (15) we have made use of the fact that Y is a deterministic function of M and $\underline{W}$ to write $p_{Y|M,\underline{W}}$ as a δ-function, and in (16), we relied on the fact that $M \perp\!\!\!\perp \underline{W}$. Thus, we have shown that $p_{Y,M}$ can be factorized into functions purely in y and m. This implies that $Y \perp\!\!\!\perp M$, and hence $I(M; Y) = 0$. ∎

Proof of Theorem 3: Recall the theorem statement: if there is some "output node" $V_{\text{op}} \in \mathscr{V}$ that satisfies $I\big(M; X(\mathbb{Q}(V_{\text{op}}))\big) > 0$, then there exists a path from $\mathscr{V}_{\text{ip}}$ to V_{op} such that every edge of this path is M-CCI'd.

So, let us start by assuming that there is some V_{op} such that $I\big(M; X(\mathbb{Q}(V_{\text{op}}))\big) > 0$. Then, by the contrapositive of Lemma 6, we must have that $\mathbb{Q}(V_{\text{op}})$ is M-CCI'd. We can then repeatedly use Theorem 2 to find edges leading backwards in time to the input nodes. Applying Theorem 2 at time $t = t_{\text{op}}$, we find there must be some edge $E_{t-1} \in \mathscr{P}(V_{\text{op}})$ which is M-CCI'd. Following this edge backwards, suppose it originated from some node $V_{t-1} \in \mathscr{V}_{t-1}$. Once again, we can apply Theorem 2 at V_{t-1} to find another edge at time $t-2$ which is

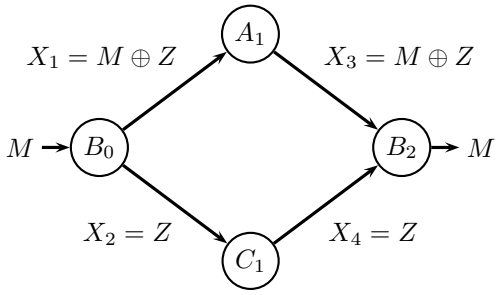

Fig. 4: The computational system used in the proof of Theorem 4. Only the edges on the upper path with $M \oplus Z$ are M-CCI'd, however, the joint distribution is symmetric with respect to Z and $M \oplus Z$. As a result, any observational measure that gives information flow to $M \oplus Z$ must also give information flow to Z.

M-CCI'd. In this manner, we can find a path leading all the way back to time $t = 0$, to some node V_0. Finally, we must argue that $V_0 \in \mathcal{V}_{\text{ip}}$ based on the fact that one of its outgoing edges, say E_0, is M-CCI'd.

At time $t = 0$, Definition 2c implies that the outgoing edges of each node in $\mathcal{V}_{\text{ip}}$ have mutual information with M, i.e., $X(\mathbb{Q}(U_0))$ depends on M for every $U_0 \in \mathcal{V}_{\text{ip}}$. By the contrapositive of Lemma 6, this implies that for every $U_0 \in \mathcal{V}_{\text{ip}}$, $\mathbb{Q}(U_0)$ is M-CCI'd. Then, by Lemma 5, we know that there must exist some particular edge in each $\mathbb{Q}(U_0)$ which is also M-CCI'd. So we have shown that each node in $\mathcal{V}_{\text{ip}}$ has at least one outgoing edge which is M-CCI'd. But we also need to show that these are the *only* edges that are M-CCI'd, and that we cannot trace an information path all the way back to some $V_0' \in \mathcal{V}_0 \setminus \mathcal{V}_{\text{ip}}$. To show this, we once again make use of Definition 2c, which states that for each $U_0' \in \mathcal{V}_0 \setminus \mathcal{V}_{\text{ip}}$, $X(\mathbb{Q}(U_0')) = f_{U_0'}(W(U_0))$. Thus, each $X(\mathbb{Q}(U_0'))$ is a deterministic function of $W(U_0')$, which in turn is not M-CCI'd. Thus, $\mathbb{Q}(U_0')$ cannot be M-CCI'd, and hence no individual edge $E_0' \in \mathbb{Q}(U_0')$ can be M-CCI'd for any $U_0' \in \mathcal{V}_0 \setminus \mathcal{V}_{\text{ip}}$. This proves that the information path we have traced backwards from V_{op} must lead to $\mathcal{V}_{\text{ip}}$.

Thus, there exists a path from $\mathcal{V}_{\text{ip}}$ to V_{op}, such that every edge of this path is M-CCI'd. ∎

Appendix B
Proofs from Section V

Proof of Theorem 4: Consider the computational system given in Fig. 4. Similar to the computational system in Counterexample 1, the node B_0 is trying to communicate M to B_2. However, this time, it generates Z itself, and sends $X_1 = M \oplus Z$ to A_1, while sending $X_2 = Z$ to C_1. A_1 and C_1 act merely as relay nodes, passing on $M \oplus Z$ and Z (which we label as X_3 and X_4 respectively) to B_2. Finally, B_2 computes M by XOR-ing its inputs.

The theorem statement asks us to consider any observational measure of information flow which satisfies the information path property. In the context of Fig. 4, the only possible information paths are (B_0, A_1, B_2) and (B_0, C_1, B_2). Therefore, any measure that satisfies the information path property will award information flow to at least one of the pairs (X_1, X_3) or (X_2, X_4).

Any observational definition of information flow would have to be a function of X_1, X_2, X_3, X_4 and M only (refer Definition 8). For convenience, denote $\underline{X} := [X_1, X_2, X_3, X_4] = [M \oplus Z,\ Z,\ M \oplus Z,\ Z]$. Consider the joint distribution $p_{M,\underline{X}}(m, \underline{x})$:

$$p_{M,\underline{X}}(m,\underline{x}) \tag{19}$$
$$= \sum_{z \in \{0,1\}} p(m, \underline{x}, z) \tag{20}$$
$$\overset{(a)}{=} \sum_{z \in \{0,1\}} p_M(m) p_Z(z) p_{\underline{X}\,|\,M,Z}(\underline{x}\,|\,m, z) \tag{21}$$
$$\overset{(b)}{=} \frac{1}{4} \sum_{z \in \{0,1\}} p_{\underline{X}\,|\,M,Z}(\underline{x}\,|\,m, z) \tag{22}$$
$$\overset{(c)}{=} \frac{1}{4} \sum_{z \in \{0,1\}} \delta(x_1, m \oplus z)\delta(x_2, z)\delta(x_3, m \oplus z)\delta(x_4, z), \tag{23}$$
$$= \frac{1}{4}\Big[\delta(x_1, m \oplus 0)\delta(x_2, 0)\delta(x_3, m \oplus 0)\delta(x_4, 0) + \delta(x_1, m \oplus 1)\delta(x_2, 1)\delta(x_3, m \oplus 1)\delta(x_4, 1)\Big], \tag{24}$$

where in (a), we made use of the fact that $M \perp\!\!\!\perp Z$; in (b), we relied on the the fact that M and Z are both $\text{Ber}(1/2)$ random variables; and in (c), δ represents the Kronecker Delta function, and we have used the fact that $\underline{X}$ is a deterministic function of M and Z. Note that when $m = 0$,

$$p_{M,\underline{X}}(0,\underline{x}) = \frac{1}{4}\Big[\delta(x_1, 0)\delta(x_2, 0)\delta(x_3, 0)\delta(x_4, 0) + \delta(x_1, 1)\delta(x_2, 1)\delta(x_3, 1)\delta(x_4, 1)\Big], \tag{25}$$

and when $m = 1$,

$$p_{M,\underline{X}}(1,\underline{x}) = \frac{1}{4}\Big[\delta(x_1, 1)\delta(x_2, 0)\delta(x_3, 1)\delta(x_4, 0) + \delta(x_1, 0)\delta(x_2, 1)\delta(x_3, 0)\delta(x_4, 1)\Big]. \tag{26}$$

In both cases, observe that $p_{M,\underline{X}}$ is symmetric in $\underline{X}$ in a very specific way: the ordered pair (x_1, x_3) may be swapped with the pair (x_2, x_4) to no effect (i.e., $M \oplus Z$ and Z are statistically symmetric with respect to M). In the limit of large samples, any observational measure will be some functional of $p_{M,\underline{X}}$. Thus, if X_1 and X_3 are awarded information flow, so too must X_2 and X_4, by basic symmetry. This means that if the information path property holds, then all edges in Fig. 4 will have information flow about M according to any observational definition. Thus, Fig. 4 describes an instance where any observational measure that satisfies the information path property awards information flow to edges that are not M-CCI'd. ∎