




**Higher order definition of causality by optimally conditioned transfer entropy**Jakub Kořenek <sup>1,2</sup> Pavel Sanda <sup>1</sup> and Jaroslav Hlinka <sup>1,3,\*</sup><sup>1</sup>*Institute of Computer Science of the Czech Academy of Sciences, Pod Vodárenskou věží 271/2, 182 07 Prague, Czech Republic*<sup>2</sup>*Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Břehová 7, 115 19 Prague, Czech Republic*<sup>3</sup>*National Institute of Mental Health, Topolová 748, 250 67 Klecany, Czech Republic*

(Received 5 September 2024; accepted 5 March 2025; published 23 April 2025)

The description of the dynamics of complex systems, in particular the capture of the interaction structure and causal relationships between elements of the system, is one of the central questions of interdisciplinary research. While the characterization of pairwise causal interactions is a relatively ripe field with established theoretical concepts and the current focus is on technical issues of their efficient estimation, it turns out that the standard concepts such as Granger causality or transfer entropy may not faithfully reflect possible synergies or interactions of higher orders, phenomena highly relevant for many real-world complex systems. In this paper, we propose a generalization and refinement of the information-theoretic approach to causal inference, enabling the description of truly multivariate, rather than multiple pairwise, causal interactions, and moving thus from causal networks to causal hypernetworks. In particular, while keeping the ability to control for mediating variables or common causes, in case of purely synergistic interactions such as the exclusive disjunction, it ascribes the causal role to the multivariate causal set but not to individual inputs, distinguishing it thus from the case of, e.g., two additive univariate causes. We demonstrate this concept by application to illustrative theoretical examples as well as a biophysically realistic simulation of biological neuronal dynamics recently reported to employ synergistic computations.

DOI: [10.1103/PhysRevE.111.L042302](https://doi.org/10.1103/PhysRevE.111.L042302)

**Introduction.** The study of complex networks is a rapidly developing field with applications across various scientific disciplines such as neuroscience, climate research, computer science, economics, energetics, or game theory [1]. The general approach views a given system as a network of interacting subsystems. A central challenge is to estimate the pattern of interactions from observed data. The formal definition and methods for estimating causal effects from one element to another have been thoroughly studied.

A common approach is the Granger causality [2]—a concept based on two principles: the cause happens before its effect, and the cause carries some additional information about the effect (not included in the "rest of the universe"). A nonlinear generalization of this concept—transfer entropy [3,4]—is based on the same principles. While Granger causality is typically cast in the framework based on prediction via linear vector autoregressive processes, transfer entropy is an information-theoretic measure aiming to capture time-directed information transfer of arbitrary functional form. Indeed, for Gaussian processes, the two concepts are equivalent [5].

In this work, we are proposing an extension to the framework of Granger causality (seen as reduction in variance) or transfer entropy (seen as reduction in surprise). To understand the nature and motivation of the extension, it is elucidating to start by conceptualizing Granger causality as a fundamental refinement of the notoriously naive concept of causality as correlation by adding the requirement of the candidate cause holding correlation (prediction/information) on top of that included in the target past and the rest of the universe. Note that analogously transfer entropy is a refinement of mutual information by conditioning on target past (and potentially other sources). This key refinement allows ascribing the causal effect more conservatively, in particular controlling for common sources or mediating variables, however, is not free of interpretation problems. To alleviate these, we suggest yet another refinement that would in particular avoid ascribing causal status to variables, that only contribute through multivariate interactions. We aim to do this while keeping the original safeguards (against false causal inferences) of Granger's approach.

Let us consider a pristine example of purely higher order interaction between source variables in causing the target variables: two candidate source variables  $X_1, X_2 \stackrel{iid}{\sim} \text{Be}(0.5)$ , i.e., independent, each with a uniform Bernoulli distribution with probability  $p = 0.5$ ; and a target variable  $Y$  defined as  $Y = \text{XOR}(X_1, X_2)$ , or in the context of time series  $Y(t+1) = \text{XOR}(X_1(t), X_2(t))$ . Note that logical XOR is zero if the inputs are equal; otherwise, it is one. In this system, pairwise mutual information  $I(X_1, Y)$  between  $X_1$  and  $Y$  is equal to zero, and

\*Contact author: [hlinka@cs.cas.cz](mailto:hlinka@cs.cas.cz)

so is  $I(X_2, Y)$ . In other words, neither of the candidate source variables carries (on their own) information about the target variable  $Y$ ; they play a role, but only together. Nevertheless transfer entropy ascribes causal status to both  $X_1$  and  $X_2$  (by conditioning each one on the other). One may, however, argue that such representation by a graph with the causal links (from  $X_1$  to  $Y$  and from  $X_2$  to  $Y$ ) is obfuscating something important, as it does not distinguish this system from, say,  $Y = X_1 + X_2$ , where each variable indeed carries information about  $Y$  on its own. We suggest a refinement in interpreting the relationship between  $X_1$  (or  $X_2$ ) and  $Y$ : they are, in fact, independent, and thus it may not be appropriate to discuss a direct causal effect in this context. At the same time, representing the system by an empty graph and denying causal status to both variables seems incorrect as well, as that would collide with the case of completely unpredictable  $Y$ . Thus, as only the multivariate information  $I(\{X_1, X_2\}, Y)$  is nonzero (in particular 1 bit), we suggest to state that there is no causal link, but there is a causal hyperlink from the set  $\{X_1, X_2\}$  to  $Y$ . Consequently, this perspective necessitates a shift from representing the system using directed graphs to employing directed causal hypergraphs. Note that directed hypergraph is a natural generalization of a directed graph, and corresponds to a pair  $(V, E)$ , where  $V$  is a set of elements called nodes and  $E$  is a set of pairs of subsets of  $V$ . Each of these pairs  $(D, C) \in E$  is called a hyperedge; the vertex subset  $D$  is known as its tail, and  $C$  as its head. Less formally one can speak of hypernetworks with hyperlinks between (sets of) nodes, in a causal context from sources to targets. However, note that interpreting the graph-theoretical properties of “networks” inferred from statistical dependencies is in many situations problematic and generally a challenging enterprise; see [6] as an example.

While the XOR function may seem artificial, higher-order interactions are commonly discussed in the context of modeling complex systems in physics [7], neuroscience [8,9], ecology [10], social sciences [11], and many more. Despite this challenge being increasingly recognized in the causal inference community [12–14], formal refinement of multivariate causality is not yet fully resolved.

*Transfer entropy and problem statement.* In the following, we assume that for each target variable  $Y$ , a set of candidate source variables  $\mathbf{X} = \{X_1, \dots, X_n\}$  is known *a priori*. In practice, the assumption is justified either by Granger causality-like reasoning—the candidate source variables are those temporally preceding the target variable—or by (additional) theoretical arguments. The transfer entropy (the positivity of which defines the presence of causal effect from  $X_i$  to  $Y$ ) is defined as the conditional mutual information between  $X_i$  and  $Y$  conditioned on all other potential sources:

$$\text{TE}_{X_i \rightarrow Y} = I(X_i, Y | \mathbf{X} \setminus \{X_i\}). \quad (1)$$

As already mentioned, we aim to generalize the transfer entropy concept to more faithfully represent causal structure in the presence of higher-order causal interactions. A natural generalization is to define existence of a causal effect of a set  $\mathbf{X}_I = \{X_{i_1}, \dots, X_{i_k}\}$  on target  $Y$  as

$$I(\{X_{i_1}, \dots, X_{i_k}\}, Y | \mathbf{X} \setminus \{X_{i_1}, \dots, X_{i_k}\}) > 0. \quad (2)$$

Let us revisit the TE behavior in the XOR example:

$$\begin{aligned} X_1, X_2 &\stackrel{iid}{\sim} \text{Be}(0.5) \\ Y &= \text{XOR}(X_1, X_2). \end{aligned} \quad (3)$$

What would such generalization of TE conclude? First, it infers a causal link from the set  $\{X_1, X_2\}$ :  $I(\{X_1, X_2\}, Y) = 1$  bit. However, it also infers causal link from  $X_1$  (or  $X_2$ ) as

$$I(X_1, Y | X_2) = I(\{X_1, X_2\}, Y) - I(X_2, Y) = 1 \text{ bit}. \quad (4)$$

As a side note, in practice, these individual links might remain undetected, as commonly used causal network inference algorithms progress iteratively from testing univariate predictors and may stop before discovering such conditional dependence—see [13] for discussion, and [14] for an example of an updated (yet, therefore, slower) algorithm more robust in this regard. Also note that the XOR example is straightforwardly generalizable to the  $k$ -variable higher order interaction by considering, e.g.,  $Y$  as a sum of variables  $X_1, X_2, \dots, X_{k-1} \stackrel{iid}{\sim} \text{Be}(0.5)$  modulo 2, i.e., the parity function. Again, even the knowledge of all but one of the source variables provides no information about  $Y$  without knowing the last one; it is the interaction of all  $k - 1$  sources that provides the target predictability.

Setting aside the implementation details, even such multivariate TE definition does not distinguish the distinctive synergistic causal effect in the XOR example, compared to, e.g., the case of summing two independent normally distributed variables. Two issues arise: first, that TE concludes that  $X_1$  has a “causal effect” on  $Y$  although  $X_1$  does not hold any information about  $Y$ , might be considered superfluous, as the information “belongs” to the pair (which multivariate TE correctly marks). Second, for the sum function, as the pair does not carry information on top of each variable, one might not want to mark the pair as causal per se, but only each variable.

*Causal effect.* The first issue suggests the suitability of including an extra condition (of nonzero mutual information between the source and the target) in the definition of a causal effect. Indeed, mutual information alone cannot be considered a measure of any causality since it could only mean that variables are affected by a common source, however, it makes good sense to consider it as an additional necessary condition of a (direct) causal effect. Such a necessary condition would solve the problem with the abovementioned example (3), but the situation can be much more complicated. Consider the following system:  $X_1, \dots, X_p, \mathcal{E}_{p+1}, \dots, \mathcal{E}_n \stackrel{iid}{\sim} \text{Be}(0.5)$ ,

$$\begin{aligned} X_k &= X_1 + \mathcal{E}_k, \quad k \in \{p+1, \dots, n\} \\ Y &= \sum_{k=2}^p \text{XOR}(X_1, X_k) + \sum_{k=p+1}^n X_k. \end{aligned} \quad (5)$$

The system is constructed (see Fig. 1) such that  $X_1$  does not directly affect  $Y$  by itself, but it affects  $Y$  indirectly through mediating variables  $X_{p+1}, \dots, X_n$ . Also,  $X_1$  affects  $Y$  through pairwise interactions: each pair  $\{X_1, X_i\}$  for  $i \in \{2, \dots, p\}$  affects  $Y$ . Note that both the mutual information  $I(X_1, Y)$  and the fully conditional mutual information (transfer entropy)  $I((X_1, Y) | \mathbf{X} \setminus \{X_1\})$  are greater than zero, thus

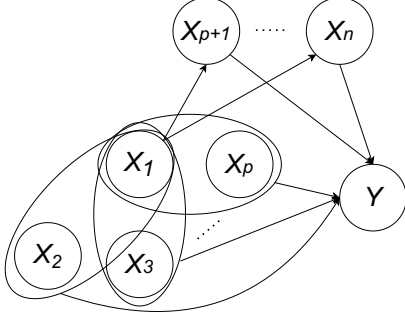


FIG. 1. Example system (5) illustrating that careful conditioning is necessary to uncover that  $X_1$  has only mediated and joint multivariate causal effects on  $Y$ . Linear terms are represented by oriented edges from a source element to the target element (the element on the left-hand side of the equation), while XOR terms are represented by oriented (hyper)edges from a pair of source elements to the target element.

suggesting a causal effect of  $X_1$  on  $Y$ . In fact, all of the conditional mutual informations  $I(X_1, Y|\mathbf{S})$  are positive, except  $I(X_1, Y|X_{p+1}, \dots, X_n)$ .

Indeed, if any of variables  $\{X_2, \dots, X_p\}$  is in the condition, the information is nonzero because of the effect of the predictor  $\{X_1, X_i\}$  [follows from chain rule, Eq. (4)]. On the other hand, if any of variables  $\{X_{p+1}, \dots, X_n\}$  is missing in the condition, information is nonzero because they are mediators of  $X_1$  and  $Y$  (information flows indirectly from  $X_1$  to  $Y$  through them). Thus, we suggest to define a causal effect as follows.

**Definition (unaccountable causal effect and optimally conditioned causal entropy).** Let  $\mathbf{X} = \{X_1, \dots, X_n\}$  be a set of candidate source variables and  $Y$  a target variable. There is a causal effect from set  $\mathbf{X}_I = \{X_{i_1}, \dots, X_{i_k}\}$  to  $Y$  if and only if

$$I(\{X_{i_1}, \dots, X_{i_k}\}, Y|\mathbf{S}) > 0 \quad (6)$$

for all  $\mathbf{S} \subseteq \mathbf{X} \setminus \{X_{i_1}, \dots, X_{i_k}\}$ .

The optimally conditioned transfer entropy is given as

$$\text{OCTE}_{\mathbf{X}_I \rightarrow Y} = \min_{\mathbf{S} \subseteq \mathbf{X} \setminus \mathbf{X}_I} I(\mathbf{X}_I, Y|\mathbf{S}). \quad (7)$$

In words, the causal effect is the minimum information a variable carries about a target, across conditioning it by all possible subsets of the "rest of the universe"—instead of considering only exactly the "rest of the universe" in the condition. Informally, a variable is only considered truly causal if it can not be conditioned out, whichever subset of other variables we choose.

Notably, the definition of the causal effect of a set of variables typically enforces the causal influence of an arbitrary set containing a variable with a causal effect. Consider, for example, a simple linear system where the variable  $Y$  is causally affected just by variable  $X_1$  plus some noise:  $Y = X_1 + \mathcal{E}_Y$ , and in the system, there are also other variables  $\{X_2, \dots, X_n\}$  independent on  $Y$  and  $X_1$ . Then, all sets containing  $X_1$  are also causally affecting  $Y$ . Note that similar behavior (supersets of causes defined as causes irrespective of further value in prediction) also holds for (a multivariate variant of) transfer entropy. In other words, although the definition of causal effect avoids attributing higher-order causal effects to individual variables, it still does not by itself distinguish multivariate

TABLE I. Example information-theoretical functionals for a triplet of variables coupled by the function  $\text{XOR}(X_1, X_2)$ , where  $X_1$  and  $X_2$  are independent random variables with distributions a)  $p_A X_1, X_2 \sim \text{Be}(0.5)$ , b)  $p_B X_1 \sim \text{Be}(0.5)$ ,  $X_2 \sim \text{Be}(0.8)$  c)  $p_C X_1 \sim \text{Be}(0.7)$ ,  $X_2 \sim \text{Be}(0.8)$ . Note that the presence of non-zero (conditional) information values depends not only on the coupling function, but substantially on the input variable distributions, even under independence of inputs.

	a)	b)	c)
$p(X_1 = 0, X_2 = 0, Y = 0)$	0.25	0.10	0.06
$p(X_1 = 0, X_2 = 1, Y = 1)$	0.25	0.40	0.24
$p(X_1 = 1, X_2 = 0, Y = 1)$	0.25	0.10	0.14
$p(X_1 = 1, X_2 = 1, Y = 0)$	0.25	0.40	0.56
$I(X_1, Y)$	0.00	0.28	0.24
$I(X_2, Y)$	0.00	0.00	0.08
$I(X_1, Y X_2)$	1.00	1.00	0.88
$I(X_2, Y X_1)$	1.00	0.72	0.72
$I(\{X_1, X_2\}, Y)$	1.00	1.00	0.96

causes that are trivial in the sense of inherited from its subsets from nontrivial higher-order multivariate causal effects.

We suggest one may add this distinction of a unique higher-order causal effect by interpreting differently the observation of multivariate causes in the presence/absence of the causal effects of the subsets. Further, to distinguish purely additive from synergistic multivariate causality, one may quantify the higher order causal information not induced trivially by the subsets; suitable approaches may include maximum entropy definition of higher order interactions [8,15], or partial information decomposition (PID) [16]. The relationship between OCTE and PID is inherently close but nontrivial. For three variables, PID can be constructed using OCTE: the unique information of an individual variable can be defined as its OCTE, while the redundancy and synergy elements are derived from the multivariate information. This construction satisfies the axioms of PID. However, a more natural approach, defining unique information and synergy directly as OCTE (with synergy corresponding to the OCTE of the set of source variables), fails to meet the PID axioms. A more sophisticated approach to PID, incorporating the strategy of powerset minimization as utilized in the analysis of how environmental factors influence transfer entropy between two variables, was recently proposed in [17].

Returning to the XOR example, note that the resulting causal graph of the system also depends on the distribution of the variables  $X_1$  and  $X_2$ . Consider three different distributions, see Table I. In all cases  $p_A, p_B, p_C$ ,  $X_1$  and  $X_2$  are independent variables with Bernoulli distribution.

We have already discussed the case A with uniformly distributed probabilities. In case B,  $X_1$  is uniformly distributed, but  $X_2 \sim \text{Be}(0.8)$ . While  $I(X_2, Y)$  remains zero,  $I(X_1, Y)$  is positive, as is  $I(X_1, Y|X_2)$ . Variable  $X_1$  thus has a causal effect on  $Y$  only due to the change in the distribution  $X_2$ . The further the variable  $X_2$  deviates from the uniform distribution, the more information  $X_1$  has about  $Y$  and the unique contribution of the pair  $\{X_1, X_2\}$  decreases. In case C, where both variables deviate from the uniform distribution, both  $X_1$  and  $X_2$  have a causal effect on  $Y$ , and they also have a unique contribution as a pair  $\{X_1, X_2\}$ .

It is worth mentioning that a similar challenge in defining causal effects, as seen in examples such as the XOR function, also arises in Pearl's approach to causality [18]. Pearl's total effect of a variable  $X$  on an outcome  $Y$  is defined as a difference in the probability distribution of  $Y$  when  $X$  is manipulated versus when  $X$  is observed naturally. Specifically, the total effect is determined by comparing the probabilities  $P(Y_x = y)$  and  $P(Y = y)$ . Here,  $P(Y_x = y)$  denotes the probability that  $Y$  takes the value  $y$  after an intervention on  $X$  (i.e., when  $X$  is set to a specific value), while  $P(Y = y)$  represents the probability that  $Y$  takes the value  $y$  under unperturbed, observational conditions (without any intervention on  $X$ ). If  $P(Y_x = y) \neq P(Y = y)$ , this inequality indicates that the intervention on  $X$  has caused a shift in the distribution of  $Y$ , suggesting the existence of a total causal effect.

Further, the direct causal effect of  $X$  on  $Y$ , holding other variables  $\mathbf{Z}$  constant, is defined as the difference  $P(Y_{x\mathbf{z}} = y) - P(Y_{x^*\mathbf{z}} = y)$ , where  $Y_{x\mathbf{z}}$  represents the outcome after both  $X$  and  $\mathbf{Z}$  are set to specific values ( $X = x$  and  $\mathbf{Z} = \mathbf{z}$ ), and  $Y_{x^*\mathbf{z}}$  represents the outcome after  $X$  is set to  $x^*$  (different from  $x$ ) while  $\mathbf{Z}$  are set equal to the previous case ( $\mathbf{Z} = \mathbf{z}$ ) [19]. This definition aims to isolate the direct effect of  $X$  on  $Y$ , removing the influence of other variables.

When applying the concept of total causal effect to the example  $Y = \text{XOR}(X, Z)$ , we observe that  $P(Y_x = y) = \text{Be}(0.5) = P(Y = y)$  for any value of  $x$ , interpreted as the absence of total causal effect of  $X$  on  $Y$ . However,  $Y_{x=0, z=0} = 0$  and  $Y_{x=1, z=0} = 1$ , therefore indicating a direct effect of  $X$  on  $Y$ .

The absence of total causal effect in the presence of a direct causal effect appears contradictory and difficult to interpret. However, we propose one can apply the same strategy as in the case of transfer entropy, in particular by redefining a variable as having a direct causal effect if and only if its perturbations affect the distribution irrespective of which subset of the other system variables is kept fixed.

**Dendritic computation of XOR.** Higher order (synergistic) dependencies occur in systems across disciplines, such as computer science or neuroscience. For instance, XOR based detectors are frequently used in algorithmic image edge detection [20,21]. For each pixel, the logical XOR mask is applied to a pair of adjacent pixels, and positive XOR value marks an edge in the picture. A similar processing principle can be found in some retinal ganglion cells, which respond to differences in intensity across the receptive field while ignoring spatially uniform input and make those neurons well posited to detect edges in visual images [22].

We further focus more closely on another example of higher order causal interactions in neuronal dynamics. Despite the long-standing assumption that computing logical XOR requires a circuit of connected neurons, recent work showed that calcium mediated dendritic action potentials (dCaAPs) in a single human cortical neuron can compute this function [23]. In essence, this computation is provided by anticoincidence behavior on the apical dendrite of the pyramidal neuron, where two simultaneous synaptic inputs surprisingly reduce dCaAP amplitude (in contrast to the traditional amplification of dendritic AP). We use the biophysical model of dCaAPs provided by [23], simulating its dynamics while dynamically modulating the synaptic inputs coming from two

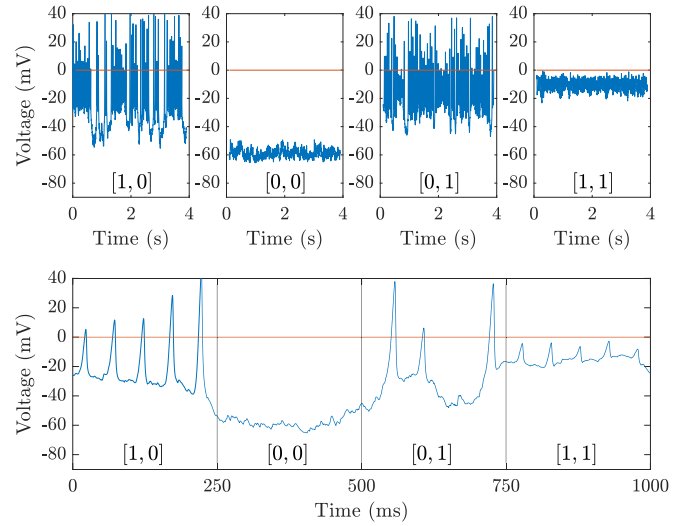


FIG. 2. Membrane voltage at the dendritic site of the dCaAP mechanism. Top: Dynamics of four logical configurations. In each panel, the system receives stable synaptic input for 4 s from two distinct synaptic pathways  $[X_1, X_2]$ , the associated logical values shown in labels at the bottom. Unequal input ( $[0,1]$  or  $[1,0]$ ) causes large dendritic spikes. Coincident activity in both pathways ( $[1,1]$ ) shows as depolarized compared to inactive pathways ( $[0,0]$ ) but will not make the neuron fire. Bottom. Altering input for each 250 ms window.

distinct sources, the activity of which corresponded to two parallel sequentially generated logical inputs, where low/high synaptic input represented logical value of 0/1, respectively. The resulting activity at the apical dendrite ("XOR gate") is shown in Fig. 2.

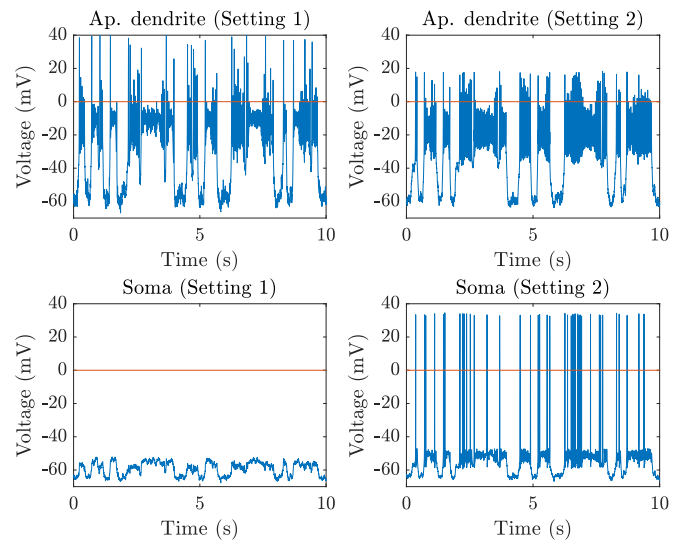


FIG. 3. Example of membrane voltage on the apical dendrite and soma of the neuron. Top: membrane potential at the dendritic site of dCaAP mechanism (as in Fig. 2). Bottom: membrane voltage at the soma of the neuron. Left: distant case. dCaAP mechanism is located 550  $\mu\text{m}$  from the soma (corresponds to [23], Fig. 3). Right: proximal case. dCaAP mechanism is located 287  $\mu\text{m}$  from the soma (corresponds to [23], Fig. S9). For model details, see the code deposition in modelDB [24].



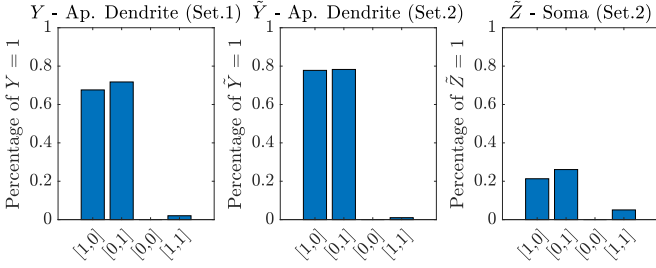


FIG. 4. Distribution of output variables  $Y$ ,  $\tilde{Y}$ , and  $\tilde{Z}$  depending on input configuration  $[X_1, X_2]$ . Left: distribution on the apical dendrite—distant case. Middle/right: distribution on the apical dendrite/soma—proximal case.

Following [23], we consider two settings: in setting 1, the location of the dCaAP mechanism on the dendrite is distant from the soma (Fig. 3, left), and in setting 2 the location is closer, thereby allowing some of the large dendritic spikes to trigger somatic spikes (Fig. 3, right).

The simulated 100 seconds were divided into 250 ms windows. Then we binarized to the apical dendrite voltages (Fig. 3, top) to obtain the target variable  $Y$  ( $\tilde{Y}$  for setting 2): we assigned a value of one to the  $Y$  if a spiking rate in the window is above 10 Hz, and zero otherwise (spike was defined as the membrane voltage exceeding 0 mV).

Figure 4 shows the measured distribution of the output variables. The distribution in all three cases follows XOR distribution, albeit for the somatic compartment, it is less accurate due to noisy transmission from the dendrite.

To determine causal relationships, the (conditional) mutual information between the input variables ( $X_1$ ,  $X_2$ , and  $\{X_1, X_2\}$ ) and the target variable  $Y$  was evaluated. As the finite sample estimate of mutual information is generally nonzero, even for independent variables, a statistical test is required. To evaluate the null hypothesis  $I(\mathbf{X}_I, Y|S) = 0$  at significance level  $\theta = 0.01$ , we use a permutation test [using  $N = 1000$  realizations of shuffled source variable(s)  $\mathbf{X}_I$  to generate the null distribution]. This breaks any dependency (pairwise or multivariate) between  $\mathbf{X}_I$  and  $Y$  while maintaining dependencies between  $Y$  and  $S$ .

For setting 1, we infer both  $I(X_1, Y) = 0$  and  $I(X_2, Y) = 0$ , while  $I(\{X_1, X_2\}, Y) = 0.46$ ; thus, we conclude by the definition of causality (6) there is no causal influence from  $X_1$  or  $X_2$ , but there is a causal (hyper)link from  $\{X_1, X_2\}$  to  $Y$  with a causal strength of  $OCTE_{\{X_1, X_2\} \rightarrow Y} = 0.46$ . For setting 2, we get the same causal structure between the synaptic input and activity on apical dendrite with a causal strength of  $OCTE_{\{X_1, X_2\} \rightarrow \tilde{Y}} = 0.56$ .

We can further evaluate the mediating role of the apical dendrite on the somatic spiking in setting 2. We add

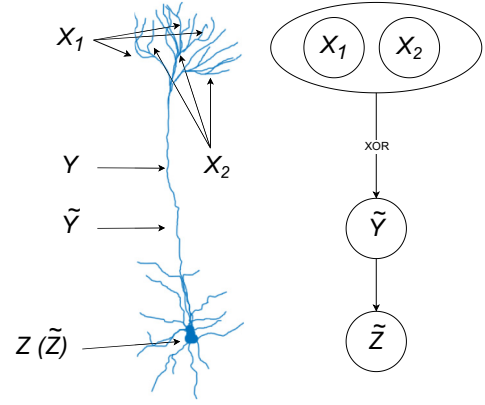


FIG. 5. Left: scheme of pyramidal neuron.  $X_1$  and  $X_2$  mark two distinct synaptic pathways (inputs),  $Y$  the output of the dCaAP mechanism on the apical dendrite ( $\tilde{Y}$  for setting 2),  $Z$  ( $\tilde{Z}$ ) the neuron soma. Right: corresponding causal diagram.

the variable  $\tilde{Z}$  representing somatic spiking to the monitored system (see right bottom corner of the Fig. 3). The set of potential sources for target variable  $\tilde{Z}$  are all subsets of  $\{X_1, X_2, Y\}$ . We obtain  $I(X_1, \tilde{Z}) = 0$ ,  $I(X_2, \tilde{Z}) = 0$ , thus neither  $X_1$  nor  $X_2$  are causal parents. Although  $I(\{X_1, X_2\}, \tilde{Z}) \neq 0$ , the predictor  $\{X_1, X_2\}$  is not a causal parent of  $\tilde{Z}$  because  $I(\{X_1, X_2\}, \tilde{Z}|\tilde{Y}) = 0$ . In this case, all information is mediated by  $\tilde{Y}$ , thus the link from  $\{X_1, X_2\}$  to  $\tilde{Z}$  is indirect. The only nontrivial causal parent of the  $\tilde{Z}$  variable is  $\tilde{Y}$  with causal strength  $OCTE_{\tilde{Y} \rightarrow \tilde{Z}} = 0.04$  (as discussed, all the supersets of  $\tilde{Y}$  also are causal parents of  $\tilde{Z}$ , however, as here they carry no extra information, being thus a purely formal, rather than unique, multivariate cause). The estimated causal scheme, a sketch of the neuron, is shown in Fig. 5.

**Conclusions.** We studied the problem of defining the causal effect in systems exhibiting higher-order causal patterns such as XOR-like interactions. We showed that the traditional approaches do not provide a sufficiently refined answer. We tackle this challenge by introducing a new, refined definition of causal structure and discuss its relation to the original definition. A tradeoff for the increased expressiveness of the causal hypernetwork definition is its exponential complexity due to formally requiring iteration over a powerset in the condition. This calls for iterative approximation algorithms already commonly utilized for transfer entropy estimation. The main theoretical challenge lies then in treating conveniently the cases of combined synergistic and direct, or even mediating, causal effect.

**Acknowledgments.** This work was supported by Grantová agentura České republiky, Project Nr. 21-17211S, by European Regional Development Fund, Project Nr. CZ.02.01.01/00/22\_008/0004643, and by Česká vysoká učen technická, Project Nr. SGS23/187/OHK4/3T/14.

- [1] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang, *Phys. Rep.* **424**, 175 (2006).
- [2] C. W. Granger, *Econometrica* **37**, 424 (1969).
- [3] T. Schreiber, *Phys. Rev. Lett.* **85**, 461 (2000).

- [4] K. Hlaváčková-Schindler, M. Paluš, M. Vejmelka, and J. Bhattacharya, *Phys. Rep.* **441**, 1 (2007).
- [5] L. Barnett, A. B. Barrett, and A. K. Seth, *Phys. Rev. Lett.* **103**, 238701 (2009).

- [6] J. Hlinka, D. Hartman, N. Jajcay, D. Tomeček, J. Tintěra, and M. Paluš, *Chaos* **27**, 035812 (2017).
- [7] F. Battiston, E. Amico, A. Barrat, G. Bianconi, G. Ferraz de Arruda, B. Franceschiello, I. Iacopini, S. Kéfi, V. Latora, Y. Moreno, M. M. Murray, T. P. Peixoto, F. Vaccarino, and G. Petri, *Nat. Phys.* **17**, 1093 (2021).
- [8] E. Schneidman, M. J. Berry, II, R. Segev, and W. Bialek, *Nature (London)* **440**, 1007 (2006).
- [9] S. Yu, H. Yang, H. Nakahara, G. S. Santos, D. Nikolić, and D. Plenz, *J. Neurosci.* **31**, 17514 (2011).
- [10] M. M. Mayfield and D. B. Stouffer, *Nat. Ecol. Evol.* **1**, 0062 (2017).
- [11] G. Cencetti, F. Battiston, B. Lepri, and M. Karsai, *Sci. Rep.* **11**, 7028 (2021).
- [12] M. A. Javidian, Z. Wang, L. Lu, and M. Valtorta, *Ann. Math. Artif. Intell.* **88**, 1003 (2020).
- [13] J. Kořenek and J. Hlinka, *Chaos* **30**, 013117 (2020).
- [14] A. Fotiadis, I. Vlachos, and D. Kugiumtzis, *Chaos* **34**, 033113 (2024).
- [15] E. Martin, J. Hlinka, and J. Davidsen, *Phys. Rev. E* **94**, 040301(R) (2016).
- [16] P. L. Williams and R. D. Beer, [arXiv:1004.2515](https://arxiv.org/abs/1004.2515).
- [17] S. Stramaglia, L. Faes, J. M. Cortes, and D. Marinazzo, *Phys. Rev. Res.* **6**, L032007 (2024).
- [18] J. Pearl, *Causality*, 2nd ed. (Cambridge University Press, Cambridge, UK, 2009).
- [19] J. Pearl, *Probabilistic and Causal Inference: The Works of Judea Pearl*, 1st ed. (Association for Computing Machinery, New York, NY, USA, 2022), pp. 373–392.
- [20] H. Al-Dmour and A. Al-Ani, *Expert Syst. Appl.* **46**, 293 (2016).
- [21] A.-V. Diaconu and I. Sima, in *Proceedings of the International Conference on Electronics, Computers and Artificial Intelligence-ECAI-2013* (IEEE, Piscataway, New Jersey, USA, 2013), pp. 1–6.
- [22] M. Meister and M. J. Berry, II, *Neuron* **22**, 435 (1999).
- [23] A. Gidon, T. A. Zolnik, P. Fidzinski, F. Bolduan, A. Papoutsis, P. Poirazi, M. Holtkamp, I. Vida, and M. E. Larkum, *Science* **367**, 83 (2020).
- [24] <https://modeldb.science/2016664>.