

Towards Definition of Higher Order Causality in Complex Systems

Jakub Kořenek^{1,2}, Pavel Sanda¹, and Jaroslav Hlinka^{1,3*}

¹*Institute of Computer Science of the Czech Academy of Sciences,
Pod Vodárenskou věží 271/2, 182 07 Prague, Czech Republic*

²*Faculty of Nuclear Sciences and Physical Engineering,
Czech Technical University, Břehová 7, 115 19, Prague, Czech Republic and*

³*National Institute of Mental Health, Topolová 748, 250 67 Klecany, Czech Republic*

(Dated: September 16, 2024)

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 synergetic 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 synergetic computations.

I. 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]. It is useful to conceptualize 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*. This key refinement allows ascribing the causal effect more conservatively, in

particular controlling for common sources or mediating variables. In this work, we suggest yet another refinement that would avoid ascribing causal status to variables, that *only* contribute through multivariate interactions, while keeping the original safeguards of Granger's approach.

Another prominent approach to causality, developed by Judea Pearl [6], assumes that so-called interventions can be carried out, making it a more difficult concept to apply due to the need to be able to carry out the interventions and assess their effects (analytically or experimentally). On the other hand, unlike Granger causality or transfer entropy, Pearl's framework does not assume the temporal organization of the investigated data, allowing thus causal inference even without sampling time series with proper temporal resolution. Nevertheless, the approaches are to a large extent related and, in particular, describe the causal structure of the investigated system as a directed graph between the variables.

We suggest that it is useful to generalize this view to consider representing causal structure by hypergraphs rather than graphs/networks. Let us consider a pristine example of purely *higher order interaction* between source variables in causing the target variables: two candidate source variables $X_1, X_2 \stackrel{iid}{\sim} Be(0.5)$ and a target variable Y defined as $Y = XOR(X_1, X_2)$, or in the context of time series $Y(t+1) = XOR(X_1(t), X_2(t))$. 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 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. The XOR function provides a pristine ex-

* hlinka@cs.cas.cz

ample of purely *higher order interaction* between source variables in causing the target variables. Nevertheless, Granger would ascribe 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. The intuition suggests one should not exactly speak about the causal effect between X_1 (or X_2) and Y : they are even independent. At the same time, representing the system by an empty graph and denying causal status to both variables seems intuitively incorrect as well, as that would collide with the case of completely unpredictable Y . Thus, maybe most intuitively, as only the *multivariate* information $I(\{X_1, X_2\}, Y)$ is non-zero (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 . Overall, the system should not be represented by a directed graph but by a directed *hypergraph*.

While the XOR function may seem artificial, higher-order interactions are commonly discussed in the context of modelling complex systems in physics [7], neuroscience [8, 9], ecology [10], social sciences [11] and many more.

Despite this challenge is increasingly recognized in the causal inference community [12–14], formal refinement of causality methods such as transfer entropy in this direction is still missing – a gap we aim to fill.

II. TRANSFER ENTROPY & PROBLEM STATEMENT

In the following text, 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, such assumption is justified either by Granger causality-like reasoning: the candidate source variables are all those temporally preceding the target variable, or by (additional) theoretical arguments. The transfer entropy TE (the positivity of which defines the presence of causal effect from X_i to Y) is then defined as the conditional mutual information between X_i and Y conditioned on all other potential sources (i.e. $\mathbf{X} \setminus \{X_i\}$):

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

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 behaviour in the XOR example:

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

What would such generalization of TE (or Pearl’s direct causal effect inference) 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 (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 Y as a sum of variables $X_1, X_2, \dots, X_{k-1} \stackrel{iid}{\sim} Be(0.5)$ modulo 2. 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). On the other hand, in the case of the linear 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 the individual variables.

III. CAUSAL EFFECT

The first issue suggests the suitability of including an extra condition (of non-zero 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 above-mentioned example (3), but the situation can be much more complicated. Consider the following system:

$$\begin{aligned} X_1, \dots, X_p, \mathcal{E}_{p+1}, \dots, \mathcal{E}_n &\stackrel{iid}{\sim} Be(0.5), \\ X_k &= X_1 + \mathcal{E}_k, \quad k \in \{p+1, \dots, n\} \\ Y &= \sum_{k=2}^p XOR(X_1, X_k) + \sum_{k=p+1}^n X_k \end{aligned} \quad (5)$$

The system is constructed 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, suggesting thus 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 non-zero 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 non-zero 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 1. 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 $\{X_{i_1}, \dots, X_{i_k}\}$ 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:

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

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 behaviour (supersets of causes defined as causes irrespective of further value in prediction) also holds for (a multivariate variant of) transfer entropy or Pearl's causality. In other words, although the Definition 1 of causal effect avoids attributing higher-order causal effects to individual variables, it still does not by itself distinguish multivariate 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 [8, 15, 16] are the subject of further research.

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_1, p_2, p_3), X_1 and X_2 are independent variables with Bernoulli distribution.

X_1	X_2	$Y = XOR(X_1, X_2)$	p_1	p_2	p_3
0	0	0	0.25	0.10	0.06
0	1	1	0.25	0.40	0.24
1	0	1	0.25	0.10	0.14
1	1	0	0.25	0.40	0.56

Table I. Function $XOR(X_1, X_2)$, where X_1 and X_2 are independent random variables with distributions: *i*) $p_1 : X_1, X_2 \sim Be(0.5)$, *ii*) $p_2 : X_1 \sim Be(0.5), X_2 \sim Be(0.8)$ *iii*) $p_3 : X_1 \sim Be(0.7), X_2 \sim Be(0.8)$.

We have already discussed the first case with uniformly distributed probabilities. In the second case, X_1 is uniformly distributed, but $X_2 \sim Be(0.8)$. 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 the last case, 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 $\{X_1, X_2\}$.

IV. 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 [17, 18]. 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 [19].

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 [20]. In essence, this computation is provided by anti-coincidence behaviour 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 [20], simulating its dynamics while dynamically modulating the synaptic inputs coming from two 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. 1.

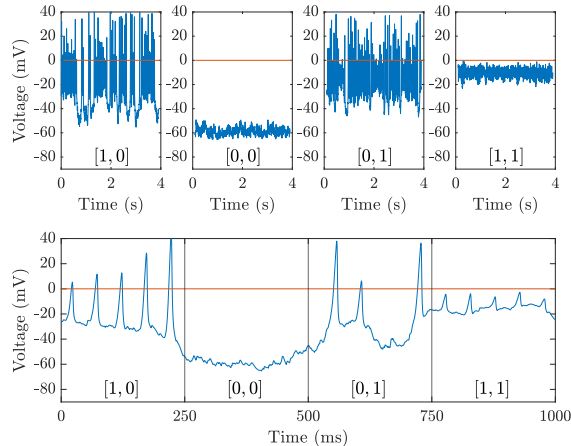


Figure 1. Membrane voltage at the dendritic site of the dCaAP mechanism. Top. Dynamics of 4 logical configurations. In each panel, the system receives stable synaptic input for 4 s from two distinct synaptic pathways $[X_1, X_2]$, the associated logical values shown in labels at the bottom. Unequal input ($[0,1]$ 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.

Following [20], we consider two settings: in setting 1, the location of the dCaAP mechanism on the dendrite is distant from the soma (Fig. 2, left), and in setting 2 the location is closer, thereby allowing some of the large dendritic spikes to trigger somatic spikes (Fig. 2, right).

The simulated 100 seconds were divided into 250 ms windows. Then we binarized to the apical dendrite voltages (Fig. 2, top) to obtain the target variable Y (\tilde{Y} for setting 2): we assigned a value of 1 to the Y if a spiking rate in the window is above 10 Hz and 0 otherwise (spike was defined as the membrane voltage exceeding 0 mV).

Fig. 3 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 $I(X, Y|Z) = 0$ at significance level $\theta = 0.01$, we use a permutation test (using $N = 1000$ realizations of shuffled source variable to generate the null distribution).

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 the setting 2, we get

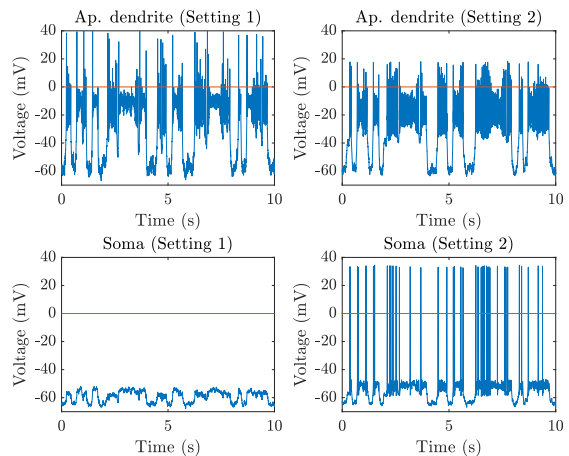


Figure 2. Example of membrane voltage on the apical dendrite and soma of the neuron. Top panels. Membrane potential at the dendritic site of dCaAP mechanism (as in Fig. 1). Bottom. Membrane voltage at the soma of the neuron. Left panels – distant case. dCaAP mechanism is located $550 \mu\text{m}$ from the soma (corresponds to [20], Fig. 3). Right – proximal case. dCaAP mechanism is located $287 \mu\text{m}$ from the soma (corresponds to [20], Fig. S9). For model details, see the code deposition in modelDB (<https://modeldb.science/2016664>).

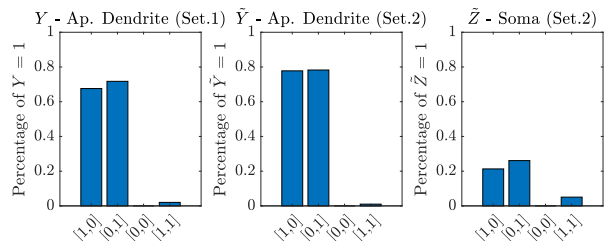


Figure 3. Distribution of output variables Y , \tilde{Y} and \tilde{Z} depending on input configuration $[X_1, X_2]$. Left. Distribution on the apical dendrite – distant case. Middle/Right. Distribution on the apical dendrite/soma – proximal case.

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 the variable \tilde{Z} representing somatic spiking to the monitored system (see right bottom corner of the Fig. 2). 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 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

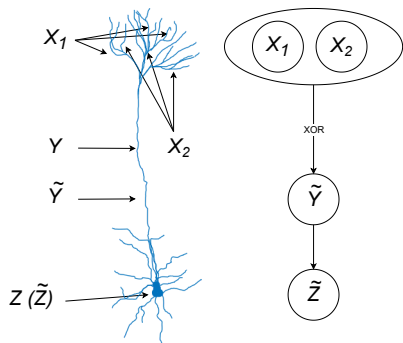


Figure 4. Left: Scheme of pyramidal neuron. X_1 and X_2 mark two distinct synaptic pathways (inputs), Y the output of the dCaAP mechanism on the apical dendrite (\tilde{Y} for setting 2), Z (\tilde{Z}) the neuron soma. Right: Corresponding causal diagram.

$OCTE_{\tilde{Y} \rightarrow \tilde{Z}} = 0.04$ (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 Figure 4.

V. 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 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 trade-off for the increased expressiveness of the 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 GACR 21-17211S, ERDF-Project Brain dynamics, No. CZ.02.01.01/00/22_008/0004643 and SGS23/187/OHK4/3T/14.

-
- [1] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang, Complex networks: Structure and dynamics, *Physics Reports* **424**, 175 (2006).
 - [2] C. W. Granger, Investigating causal relations by econometric model and cross spectral methods, *Econometrica* **37**, 424 (1969).
 - [3] T. Schreiber, Measuring information transfer, *Phys. Rev. Lett.* **85**, 461 (2000).
 - [4] K. Hlaváčková-Schindler, M. Paluš, M. Vejmelka, and J. Bhattacharya, Causality detection based on information-theoretic approaches in time series analysis, *Physics Reports* **441**, 1 (2007).
 - [5] L. Barnett, A. B. Barrett, and A. K. Seth, Granger Causality and Transfer Entropy Are Equivalent for Gaussian Variables, *Physical Review Letters* **103** (2009).
 - [6] J. Pearl, *Causality*, 2nd ed. (Cambridge University Press, Cambridge, UK, 2009).
 - [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, The physics of higher-order interactions in complex systems, *Nature Physics* **17**, 1093 (2021).
 - [8] E. Schneidman, M. J. n. Berry, R. Segev, and W. Bialek, Weak pairwise correlations imply strongly correlated network states in a neural population., *Nature* **440**, 1007 (2006).
 - [9] S. Yu, H. Yang, H. Nakahara, G. S. Santos, D. Nikolić, and D. Plenz, Higher-order interactions characterized in cortical activity, *Journal of Neuroscience* **31**, 17514 (2011).
 - [10] M. M. Mayfield and D. B. Stouffer, Higher-order interactions capture unexplained complexity in diverse communities, *Nature Ecology & Evolution* **1**, 0062 (2017).
 - [11] G. Cencetti, F. Battiston, B. Lepri, and M. Karsai, Temporal properties of higher-order interactions in social networks, *Scientific Reports* **11**, 7028 (2021).
 - [12] M. A. Javidian, Z. Wang, L. Lu, and M. Valtorta, On a hypergraph probabilistic graphical model, *Annals of Mathematics and Artificial Intelligence* **88**, 1003 (2020).
 - [13] J. Kořenek and J. Hlinka, Causal network discovery by iterative conditioning: Comparison of algorithms, *Chaos* **30**, 013117 (2020).
 - [14] A. Fotiadis, I. Vlachos, and D. Kugiumtzis, The causality measure of partial mutual information from mixed embedding (PMIME) revisited, *Chaos* **34**, 033113 (2024).
 - [15] P. L. Williams and R. D. Beer, Nonnegative decomposition of multivariate information, *ArXiv* **abs/1004.2515** (2010).
 - [16] E. Martin, J. Hlinka, and J. Davidsen, Pairwise network information and nonlinear correlations, *Physical Review E* **94** (2016).
 - [17] H. Al-Dmour and A. Al-Ani, A steganography embedding method based on edge identification and xor coding, *Expert systems with Applications* **46**, 293 (2016).
 - [18] A.-V. Diaconu and I. Sima, Simple, xor based, image edge detection, in *Proceedings of the International Conference on ELECTRONICS, COMPUTERS and ARTIFICIAL INTELLIGENCE-ECAI-2013* (IEEE, 2013) pp. 1–6.
 - [19] M. Meister and M. Berry, The neural code of the retina, *Neuron* **22**, 435 (1999).
 - [20] A. Gidon, T. A. Zolnik, P. Fidzinski, F. Bolduan, A. Pappoussi, P. Poirazi, M. Holtkamp, I. Vida, and M. E. Larkum, Dendritic action potentials and computation in human layer 2/3 cortical neurons, *Science* **367**, 83 (2020).