When Dependence Disappears: Faithfulness and Effective Independence in Nonlinear Dynamical Systems

Abstract: This article argues that nonlinear dynamical systems provide a structural mechanism for (near-)unfaithfulness that has received little attention in philosophical discussions of causal discovery. In such systems, observable statistical relations arise from invariant-measure averaging over state-dependent causal effects. Because local causal influences vary across the system's state space, their contributions may partially cancel, weakening statistical dependence between causally connected variables. The resulting "effective independence" can be empirically indistinguishable from genuine independence, posing challenges for dependence-based causal discovery methods. Unlike classical violations that rely on parameter fine-tuning, this form of (near-)unfaithfulness arises from structural features of system dynamics.

1. Introduction

The causal faithfulness condition plays a central role in causal discovery methods that infer causal structure from probabilistic data, particularly within the framework of causal Bayesian networks. Roughly, faithfulness requires that all and only those probabilistic independencies implied by the causal Markov condition appear in the observed distribution. When this condition holds, conditional independence relations in the data can be used to recover aspects of the underlying causal graph. Because of this role, the possibility that faithfulness might fail has long attracted attention in both philosophy and the sciences (Spirtes et al. 2000; Zhang and Spirtes 2008; Pearl 2009; Andersen 2013). A number of alternative assumptions weaker than faithfulness have therefore been proposed in order to address potential failures of the condition (Ramsey et al. 2012; Spirtes and Zhang 2014; Zhalama et al. 2017; Marx et al. 2021).[1]

One widely discussed mechanism for violating faithfulness is causal cancellation, where the effects of two or more causal paths offset one another. In such cases a variable may become probabilistically independent of one of its causes even though a causal connection remains present in the underlying graph. These scenarios are often regarded

---

[1] More recent work in causal discovery has increasingly recognized that methods based solely on conditional independence may encounter difficulties when applied to complex systems, especially those involving temporal structures, feedback, or latent variables. See Mooij et al.(2016), Peters et al.(2017), and Runge et al. (2019).

as rare because they require finely tuned parameter values: the set of parameter configurations that produce exact cancellation typically has Lebesgue measure zero (Spirtes et al. 2000). Debates about the plausibility of such violations have therefore largely focused on whether real systems are likely to realize these finely balanced parameterizations (Andersen 2013).

This article argues that nonlinear dynamical systems provide another largely overlooked structural mechanism for (near-)unfaithfulness that does not depend on finely tuned parameter values. In such systems the relationship between interacting variables often depends on the evolving state of the system itself. When the system possesses oscillatory or chaotic attractors,[2] observable probability distributions arise from invariant-measure averaging over hidden system states. Because local causal effects may vary across the attractor—sometimes even reversing sign—the resulting averaging process can substantially weaken or mask statistical dependence between causally connected variables. In contrast to the classical cancellation scenario, this phenomenon does not require delicate parameter fine-tuning. Instead, it arises from structural properties of the system dynamics, such as attractor geometry, regime switching, or oscillatory feedback. If this is correct, then unfaithfulness need not arise only from rare parameter cancellations but may instead emerge from generic features of nonlinear system dynamics.

It is true that what nonlinear dynamical systems most commonly produce is

---

[2] The attractor is the region of state space to which system trajectories converge.

approximate or effective independence[3] rather than full independence. Yet, in practice this distinction matters little for causal discovery, because empirical analyses typically rely on finite samples and thereby extremely weak statistical dependence may be empirically indistinguishable from genuine independence. Consequently, approximate independence can generate the same practical and epistemic difficulties for causal inference as exact faithfulness violations.

Note that the purpose of this article is not merely to point out that near-unfaithfulness can occur, which is already well documented in the causal discovery literature (Zhang and Spirtes 2008; Andersen 2013; Uhler et al. 2013; Spirtes and Zhang 2014). Rather, the claim defended here is that nonlinear dynamical systems provide a structural mechanism through which near-unfaithfulness can arise naturally. The article proceeds as follows. Section 2 reviews faithfulness and the classical cancellation mechanism. Section 3 argues that nonlinear dynamical systems can naturally generate approximate independence through invariant-measure averaging of state-dependent causal effects. Section 4 explains why such approximate independence creates the same practical and epistemic difficulties for causal discovery as full independence. Section 5 discusses two potential strategies for mitigating these difficulties.

2. The Causal Faithfulness Condition

[3] As will be discussed in Section 4, effective independence refers to statistical dependence that is so weak that, in finite samples, it becomes practically indistinguishable from independence.

Roughly, faithfulness says that *all and only* those (conditional and unconditional) probabilistic independencies entailed by the causal Markov condition hold (Spirtes et al. 2000). The violation of faithfulness occurs when, for example, there are probabilistic independencies that are not entailed by the causal Markov condition. To see this, consider Hesslow (1976)'s oft-cited example:

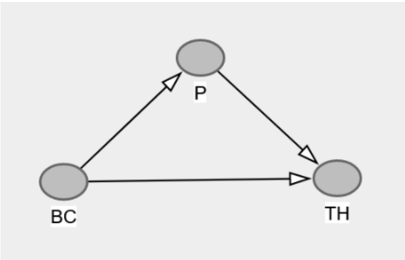

Figure 1. A causal graph with three variables.

In Figure 1, P refers to pregnancy, BC taking birth control pills, and TH the occurrence of thrombosis. BC increases the risk of TH, P also increases the risk of TH, and BC reduces the risk of P. Overall, BC reduces the risk of TH by reducing the risk of P along the causal path $BC \rightarrow P \rightarrow TH$. So, the causal effect of the path $BC \rightarrow TH$ sometimes might be cancelled out by the causal effect of the path $BC \rightarrow P \rightarrow TH$. This happens when the variables take on some exactly counterbalanced parameter values. Obviously, this violates faithfulness, because—thanks to cancelling out— BC may be probabilistically independent of TH but the causal Markov condition tells the opposite: they should be dependent. Thus, it is not a probabilistic independency entailed by the causal Markov condition.

It is disputed whether this type of unfaithfulness is rare. Spirtes et al. (2000) argue that unfaithfulness is rare because "The parameter values […] form a real space, and the set of points in this space that create vanishing partial correlations not implied by the Markov condition have Lebesgue measure 0" (41). Namely, it requires parameter configurations that form a set of Lebesgue measure zero. On the other hand, Andersen (2013) contends that Spirtes et al. (2000)'s argument relies on the unrealistic assumption that the parameter values are equally probable over their ranges; yet, as she argues, in real biological and ecological systems involving equilibrium-maintaining mechanisms, the parameter values are not equally probable but rather "disproportionately likely to be centered around the balanced values" (677). Notice that Andersen shares the view with Spirtes et al. (2000) that the supposed violations are essentially due to parameter fine-tuning.

However, there is another source of (near-)unfaithfulness philosophers have largely overlooked, which is related to nonlinear dynamical systems.

3. (Near-)Unfaithfulness in Nonlinear Dynamical Systems[4]

One key feature of nonlinear dynamical systems is state dependence, namely, the

---

[4] Notice that Zhang & Spirtes (2008), Uhler et al. (2013) and Spirtes & Zhang (2014) have already recognized that near-unfaithfulness can render causal discovery unreliable in practice, although they do not specifically address how near-unfaithfulness can arise in nonlinear dynamical systems.

relationship between interacting factors can vary as the state of the system changes (Deyle and Sugihara 2011; Sugihara et al. 2012; Chang et al. 2017). For instance, depending on the state of the system, one factor of the system can sometimes positively affect the second factor, but at some other times negatively affect it. Consequently, the two interacting factors may appear to be globally independent over some long period of time, even though they are in fact locally dependent at various states of the system.

One toy example helps illustrate the phenomenon. Suppose we have a simple predator-prey system, and let X represent predator abundance, Y prey growth rate and S ecological phase. Suppose further that there are only two phases of this ecological system, i.e., $S = +1$ during the recovery phase and $S = -1$ during the collapse phase. X and Y are causally linked because predator abundance affects prey growth, yet the causal effects change signs across these two phases. Namely, during the collapse phase predators suppress prey growth whereas during the recovery phase predator decline triggers prey recovery. If the system spends equal time in each phase, then X and Y may appear nearly independent in the aggregated data even though they are in fact causally linked.

Now let us show why this can occur a bit more formally. In nonlinear dynamical systems, (near-)unfaithfulness emerges naturally from invariant-measure[5] averaging

---

[5] In the study of dynamical systems, an *invariant measure* (often denoted as μ) is the mathematical description of the long-term distribution of a system's state on the attractor. E.g., imagine we have a stirred pot of water with a drop of ink. The water molecules are pushing the ink around in a complex,

over state-dependent causal effects. Consider a system of the form

$$Y_{t+1} = f(X_t, S_t)$$

where X_t and Y_{t+k} denote two interacting variables and S_t represents the (possibly high-dimensional) hidden system state evolving according to its own dynamics. In such systems the observable distribution takes the form[6]

$$P(Y_{t+1} \mid X_t) = \int P(Y_{t+1} \mid X_t, S_t = s)\, d\mu(s)$$

where μ is the invariant measure governing the long-run distribution of system states. The causal influence of X_t on Y_{t+1} is therefore typically state-dependent and can be characterized locally by the partial derivative

---

chaotic way. After a long time, the ink is perfectly distributed throughout the pot. If we look at any specific cubic centimeter of water, the probability of finding an ink particle there remains constant over time, even though the individual particles are still moving. The system is still dynamical (since the ink is moving), but the distribution of the ink is invariant (since it doesn't change).

[6] This expression should be understood heuristically: the observable statistical relationship between X_t and Y_{t+1} reflects an aggregation of local causal effects across the attractor rather than a single fixed structural coefficient.

$$\frac{\partial Y}{\partial X}(s) \text{ (evaluated at state } s).$$

Crucially, this local causal effect may vary substantially across the state space, not only in magnitude but also in sign. Since the system explores its attractor according to an invariant measure μ, the observable statistical relationship between X_t and Y_{t+1} reflects an average over the attractor. When the local causal effect changes sign across regions of the attractor that receive comparable invariant-measure weight, positive and negative contributions to the dependence between X_t and Y_{t+1} may cancel in the integral above. Consequently, the observed statistical dependence between X_t and Y_{t+1} may become weak and, in special cases, approach zero, despite the presence of causal influence throughout the system's dynamics (**Appendix** below provides a simple formal illustration showing how dynamical mixing can weaken observable statistical dependence between variables generated by a deterministic nonlinear system).

Importantly, this mechanism of cancellation does not rely on finely tuned parameter equalities of the sort typically discussed in the faithfulness literature. Instead, it arises naturally from structural features of nonlinear dynamical systems—specifically, from the geometry of the attractor and the distribution of trajectories across the system's state space. Although this mechanism typically produces only approximate rather than full independence, the resulting attenuation of dependence is sufficient to undermine causal discovery methods that rely on detecting statistical dependence, as will be discussed in the next section. In static linear systems, however, this mechanism generically does not arise. The reason is very simple: if we use a structural equation model such as $Y =$

$aX + \epsilon$ to capture these static linear systems, the causal influence of X on Y can be written as

$$\frac{\partial Y}{\partial X} = a$$

which is a constant over all the states of the system of interest. Hence, dependence disappears only if $a = 0$, meaning that causation is completely wiped out. This also shows that in static linear systems unfaithfulness requires parameter fine-tuning. Therefore, near-unfaithfulness in nonlinear dynamical systems can emerge as a generic consequence of the system's dynamics, rather than the result of accidental parameter tuning.

4. Full or Effective Independence?

One might rightly point out that full independence due to exact cancellation in nonlinear dynamical systems is relatively rare because exact cancellation is structurally extremely demanding and not many real systems exhibit that structure.[7] Instead, what we can

---

[7] Note that exact cancellation requires assumptions such as Haar-uniform noise on a compact group, perfect randomization mechanisms and very special probabilistic structures, and these assumptions are stronger than assumptions such as geometric symmetry of the attractor, balanced invariant-measure weights and sign-changing local causal effects that usually can be satisfied by ordinary nonlinear

legitimately have are at most systems that manifest near-exact cancellation that can only produce approximate independence. And this suffices to show that dependence may be weakened but not entirely masked. One might therefore conclude that faithfulness is rarely violated.

I am sympathetic to this diagnosis, and agree that exact cancellation might be relatively rare. Nevertheless, it remains true that near-exact cancellation creates exactly the same trouble for faithfulness as does exact cancellation. This is due to *effective independence* (Sugihara et al. 2012; Ye et al. 2015). Roughly, effective independence refers to the fact that two variables can be causally connected in the underlying dynamics but appear statistically independent (or nearly so) when analyzed using standard correlation or dependence-based methods applied to time-averaged observational data. So, even if causation does exist in the underlying dynamical system, observed statistical dependence can be too weak (or approach zero) such that it is undetectable in practice.[8] This is understandable given that in real data we never observe the exact distribution. Instead, we estimate it from a finite sample. So, if the difference

---

dynamical systems.

[8] Andersen (2013) also distinguishes genuine violations of faithfulness from apparent violations, and argues that apparent violations "generate precisely the same problems for inferring causal structure from probabilistic relationships in data as do genuine […] violations" due to practical reasons related to detectability (672).

$$| P(X_t, Y_{t+k}) - P(X_t)P(Y_{t+k}) |$$

is sufficiently small, then with realistic sample sizes we cannot statistically distinguish the true joint distribution from the product distribution. Thus, empirical tests will behave as if $X_t \perp Y_{t+k}$. This implies that in these scenarios exact-unfaithfulness (due to full independence) and near-unfaithfulness (due to effective independence) are practically and epistemically indistinguishable, even though they are conceptually and ontologically distinguishable.

Furthermore, unlike static linear systems where near-exact cancellation requires near-exact parameter fine-tuning, in nonlinear dynamical systems near-exact cancellation emerges naturally from attractor geometry, oscillations, feedback structures, etc. Therefore, this suggests that (near-)unfaithfulness may be relatively common in such systems.

5. Potential Solutions

5.1. Dynamical Bayesian networks

One way to resolve the trouble is to extend standard causal Bayesian networks in such a way that they can model systems evolving over time—dynamical Bayesian networks (DBNs) are one of those extensions (Dagum et al. 1992, 1995; Murphy 2002; Chang et al. 2023). Yet, even though DBNs can partially mitigate some of the challenges created by near-unfaithfulness in nonlinear dynamical systems, they do not fundamentally resolve the problem. Unlike static Bayesian networks, DBNs explicitly represent temporal causal structure by modeling probabilistic dependencies across time, typically through factorizations of the form $P(X_{t+1} \mid X_t)$. This temporal structure allows DBNs to represent lagged causal relationships such as $X_t \rightarrow Y_{t+1}$, which more closely mirrors the causal organization of many dynamical systems. By explicitly encoding time-lagged interactions, DBNs can reduce certain forms of dependence cancellation that arise when dynamical effects are collapsed into contemporaneous statistical associations. In particular, when causal influences unfold over multiple time steps, incorporating temporal lags can help separate distinct causal pathways that might otherwise be obscured in static models.

However, the effectiveness of DBNs remains limited in the presence of near-unfaithfulness generated by nonlinear state-dependent dynamics. In many nonlinear systems, the observable relationship between variables is obtained by averaging over hidden system states that evolve according to their own dynamics. For example, in a system of the form $Y_{t+1} = f(X_t, S_t)$, where S_t denotes an unobserved state variable, the observable conditional distribution $P(Y_{t+1} \mid X_t)$ results from integrating the conditional effects $P(Y_{t+1} \mid X_t, S_t)$ over the distribution of hidden states. Because the local causal effect of X_t on Y_{t+1} may vary substantially across different regions of the state space—sometimes even reversing sign—this averaging process can substantially attenuate observable dependence. Consequently, even when a genuine causal relationship exists, the resulting statistical dependence between X_t and Y_{t+1} may become arbitrarily weak, producing near-unfaithfulness. Since DBNs still rely on detecting conditional dependencies in the observed probability distribution, they remain vulnerable to this attenuation.

Moreover, nonlinear dynamical systems often involve high-dimensional latent states that are only partially observable through measured variables. When these hidden states are marginalized out, the resulting observable process can display weak or unstable dependence patterns that are difficult to recover through conditional independence tests. Although including additional time lags in a DBN can sometimes approximate hidden state information, this strategy cannot fully eliminate the loss of information induced by state-space projection. As a result, the independence structures exploited by DBN-based causal discovery algorithms may fail to reflect the true underlying causal interactions.

For these reasons, DBNs may alleviate certain temporal modeling deficiencies of static Bayesian networks but cannot fully overcome the structural fragility of faithfulness in nonlinear dynamical systems. When causal effects vary across system states and observable dependencies arise only through invariant-measure averaging, methods based solely on probabilistic dependence may systematically underestimate or fail to detect causal relationships. This limitation has motivated alternative approaches, such as state-space reconstruction methods, which attempt to recover causal structures from the geometry of dynamical trajectories rather than from global probabilistic dependencies alone. This is what we will explore in the next section.

5.2. Empirical dynamical modeling

Scientists during the last three decades have developed strategies tailored for detecting causal relationships in complex nonlinear dynamical systems with state dependence, strategies that no longer rely on probabilistic dependence. One such strategy is *empirical dynamical modeling* (EDM) (Sugihara and May 1990; Anderson et al. 2008; Sugihara et al. 2012; Ye et al. 2015). Roughly, EDM is built upon the idea derived from Takens (1981)'s mathematical theorem: one can reconstruct the behavior of a dynamical system solely from time series data, without hypothesizing any equations purported to underlie the behavior of the system.

In dynamical systems theory, the evolution of a system can be represented as trajectories on an attractor in state space. Takens' theorem (1981) shows that the geometry of such an attractor can often be reconstructed from time-series observations of a single variable using delay coordinates. Although the reconstructed attractor is only a "shadow" of the true system dynamics, it preserves essential structural features of the original dynamics. This insight provides the theoretical foundation for EDM, namely, that information encoded in the time series data of one variable can be recovered from information encoded in that of another variable. Inspired by this, novel methods have been proposed to identify causal relationships from time series data. For instance, Sugihara et al. (2012) have developed a method within the EDM framework called *convergent cross mapping* (CCM). The basic idea is that we simply examine "whether the time indices of nearby points on the Y manifold can be used to identify nearby points on the X manifold (Sugihara et al. 2012, 497)". If that is the case, then it is reasonable to infer that Y and X are causally connected.

More specifically, this method operates in this way. First, we create a few *delayed or lagged time series* for variable Y, so that we have a new set of time series, e.g., Y_t (the original unlagged time series), $Y_{t-\tau}$, $Y_{t-2\tau}$, $Y_{t-3\tau}$, etc. Second, we use these newly created time series as new coordinates to build a new manifold. Then, Takens' theorem rules that the new manifold is a "shadow" version of the original "real" manifold, which preserves the essential features of the "real" one; for example, if the "real" manifold encodes the information that X and Y are causally connected, the "shadow" manifold will also encode that information.

EDM illustrates a different strategy for causal discovery: instead of detecting causal structure through conditional dependencies, it attempts to recover causal relations from the geometric structure of the system's attractor. Consequently, it can recover causal relationships reliably in complex nonlinear dynamical systems even when there exist exact cancellations—so, unfaithfulness no longer poses a threat. However, it must be noted that, as a family of related methods still developing, EDM has its own limitations and there are situations where EDM might fall short of recovering the genuine underlying causal structures.[9]

---

[9] Critics point out that CCM can fail when synchronization occurs, shared periodic drivers exist, or noise or short data length prevents reliable attractor reconstruction (McCracken & Weigel 2014; Cobey &

6. Conclusion

This article has argued that violations of faithfulness may arise more naturally in nonlinear dynamical systems than is typically assumed. Unlike classical cases that depend on finely tuned parameter cancellation, nonlinear dynamics can generate near-unfaithfulness through invariant-measure averaging of state-dependent causal effects. Since local causal influences may vary across an attractor, their contributions can partially cancel in the observable distribution, weakening statistical dependence between causally connected variables. Although this process usually produces only approximate independence, such "effective independence" is often empirically indistinguishable from genuine independence in finite data. Consequently, nonlinear dynamical systems pose a systematic challenge for causal discovery methods that rely primarily on probabilistic independence and motivate the development of alternative approaches to causal inference.

Appendix

This appendix illustrates how dynamical mixing can weaken observable statistical dependence in nonlinear systems. Although the main text focuses on dependence

---

Baskerville 2016). For responses, see Sugihara et al. (2017).

between X_t and Y_{t+1}, the following result shows that mixing dynamics can drive statistical dependence arbitrarily close to independence at longer time lags.

Theorem: Approximate independence from mixing dynamics

Let $T: M \to M$ be a measurable map defining a deterministic dynamical system

$$Z_{t+1} = T(Z_t)$$

on a compact state space M with invariant probability measure μ.[10] Let $Z_0 \sim \mu$ and define the stationary process

$$Z_t = T^t(Z_0).$$

Assume the system is strongly mixing, meaning that for any measurable sets $A, B \subseteq M$,

$$\mu(A \cap T^{-k}(B)) \to \mu(A)\mu(B) \text{ as } k \to \infty.$$

Let the observable variables be

$$X_t = f(Z_t),\ Y_t = g(Z_t),$$

where $f, g: M \to \mathbb{R}$ are measurable functions. Then for measurable sets A, B,

$$P(X_t \in A,\, Y_{t+k} \in B) \to P(X_t \in A)P(Y_{t+k} \in B) \text{ as } k \to \infty.$$

(Note that although the system is deterministic, the stochastic process Z_t arises from sampling the initial state from the invariant measure. The mixing property therefore concerns the statistical dependence of events under this probability measure rather than

---

[10] Roughly, a compact state space denotes a bounded state space in which trajectories remain confined.

the functional dependence of individual trajectories.)

Proof: Since the system evolves according to

$$Z_{t+1} = T(Z_t),$$

the state at time t can be written as

$$Z_t = T^t(Z_0).$$

Thus,

$$Z_{t+k} = T^k(Z_t).$$

Let A, B be measurable sets and define

$$A_f = \{z \in M : f(z) \in A\},\ B_g = \{z \in M : g(z) \in B\}.$$

Then,

$$P(X_t \in A, Y_{t+k} \in B) = P(f(Z_t) \in A,\ g(Z_{t+k}) \in B).$$

Using $Z_{t+k} = T^k(Z_t)$, this becomes

$$P(X_t \in A, Y_{t+k} \in B) = P(Z_t \in A_f,\ T^k(Z_t) \in B_g).$$

Since $T^k(Z_t) \in B_g$ if and only if $Z_t \in T^{-k}(B_g)$, we obtain

$$P(Z_t \in A_f,\ T^k(Z_t) \in B_g) = P(Z_t \in A_f,\ Z_t \in T^{-k}(B_g)).$$

Hence,

$$P(X_t \in A, Y_{t+k} \in B) = P(Z_t \in A_f \cap T^{-k}(B_g)).$$

Since $Z_0 \sim \mu$ and μ is invariant under the dynamics, the process is stationary, so

$$P(Z_t \in A_f \cap T^{-k}(B_g)) = \mu(A_f \cap T^{-k}(B_g)).$$

By the strong mixing assumption,

$$\lim_{k\to\infty} \mu(A_f \cap T^{-k}(B_g)) = \mu(A_f)\mu(B_g).$$

By definition of the observables,

$$\mu(A_f) = P(X_t \in A),\ \mu(B_g) = P(Y_{t+k} \in B).$$

Substituting gives

$$\lim_{k\to\infty} P(X_t \in A, Y_{t+k} \in B) = P(X_t \in A)P(Y_{t+k} \in B).$$

Therefore,

$$X_t \perp Y_{t+k} \text{ as } k \to \infty.$$

Q.E.D.

References

Andersen H (2013) When to expect violations of causal faithfulness and why it matters. Philos Sci 80:672–683

Anderson CNK, Hsieh C, Sandin SA, et al (2008) Why fishing magnifies fluctuations in fish abundance. Nature 452:835–839. https://doi.org/10.1038/nature06851

Chang C-W, Ushio M, Hsieh C (2017) Empirical dynamic modeling for beginners. Ecol Res 32:785–796. https://doi.org/10.1007/s11284-017-1469-9

Chang J, Bai Y, Xue J, et al (2023) Dynamic Bayesian networks with application in environmental modeling and management: A review. Environ Model Softw 170:105835

Cobey S, Baskerville EB (2016) Limits to causal inference with state-space reconstruction for infectious disease. PloS One 11:e0169050

Dagum P, Galper A, Horvitz E (1992) Dynamic network models for forecasting. In: Uncertainty in artificial intelligence. Elsevier, pp 41–48

Dagum P, Galper A, Horvitz E, Seiver A (1995) Uncertain reasoning and forecasting. Int J Forecast 11:73–87

Deyle ER, Sugihara G (2011) Generalized theorems for nonlinear state space reconstruction. Plos One 6:e18295

Hesslow G (1976) Two notes on the probabilistic approach to causality. Philos Sci

43:290–292

Marx A, Gretton A, Mooij JM (2021) A weaker faithfulness assumption based on triple interactions. In: Uncertainty in Artificial Intelligence. PMLR, pp 451–460

McCracken JM, Weigel RS (2014) Convergent cross-mapping and pairwise asymmetric inference. Phys Rev E 90:062903. https://doi.org/10.1103/PhysRevE.90.062903

Mooij JM, Peters J, Janzing D, et al (2016) Distinguishing cause from effect using observational data: methods and benchmarks. J Mach Learn Res 17:1–102

Murphy KP (2002) Dynamic bayesian networks: representation, inference and learning. University of California, Berkeley

Pearl J (2009) causality: models, reasoning, and inference, 2nd edn. Cambridge University Press, Cambridge

Peters J, Janzing D, Scholkopf B (2017) Elements of causal inference: foundations and learning algorithms. MIT press

Ramsey J, Zhang J, Spirtes PL (2012) Adjacency-Faithfulness and Conservative Causal Inference

Runge J, Nowack P, Kretschmer M, et al (2019) Detecting and quantifying causal associations in large nonlinear time series datasets. Sci Adv 5:eaau4996. https://doi.org/10.1126/sciadv.aau4996

Spirtes P, Glymour CN, Scheines R (2000) Causation, prediction, and search. MIT press, Cambridge

Spirtes P, Zhang J (2014) A uniformly consistent estimator of causal effects under the k-triangle-faithfulness assumption. Stat Sci 662–678

Sugihara G, Deyle ER, Ye H (2017) Reply to Baskerville and Cobey: Misconceptions about causation with synchrony and seasonal drivers. Proc Natl Acad Sci 114:. https://doi.org/10.1073/pnas.1700998114

Sugihara G, May R, Ye H, et al (2012) Detecting causality in complex ecosystems. science 338:496–500

Sugihara G, May RM (1990) Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. Nature 344:734–741. https://doi.org/10.1038/344734a0

Takens F (1981) Detecting strange attractors in turbulence. In: Rand D, Young L-S (eds) Dynamical Systems and Turbulence, Warwick 1980. Springer, Berlin, Heidelberg, pp 366–381

Uhler C, Raskutti G, Bühlmann P, Yu B (2013) Geometry of the faithfulness assumption in causal inference. Ann Stat 436–463

Ye H, Beamish RJ, Glaser SM, et al (2015) Equation-free mechanistic ecosystem forecasting using empirical dynamic modeling. Proc Natl Acad Sci 112:E1569–E1576. https://doi.org/10.1073/pnas.1417063112

Zhalama, Zhang J, Mayer W (2017) Weakening faithfulness: some heuristic causal discovery algorithms. Int J Data Sci Anal 3:93–104. https://doi.org/10.1007/s41060-016-0033-y

Zhang J, Spirtes P (2008) Detection of Unfaithfulness and Robust Causal Inference. Minds Mach 18:239–271. https://doi.org/10.1007/s11023-008-9096-4