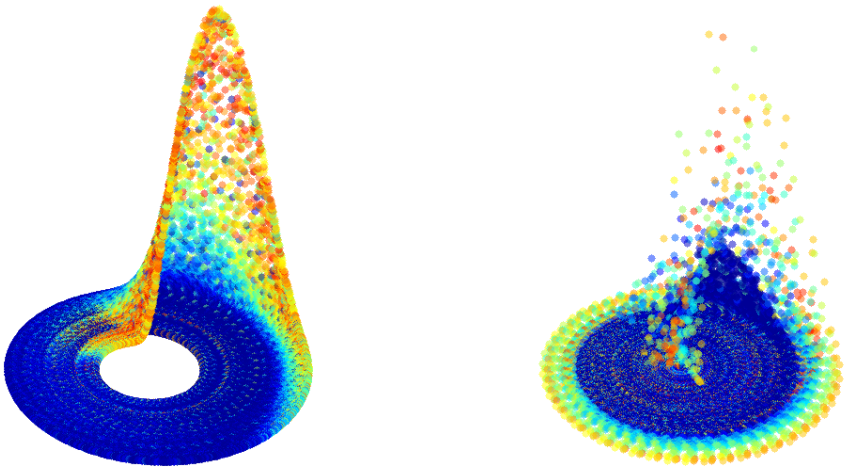
INFERRING CAUSAL INFLUENCES FROM EXPANSIVE DISTORTIONS BETWEEN STATE SPACE RECONSTRUCTIONS

ERIK LAMINSKI



Dissertation
in partial fulfilment of the degree
Doctor rerum naturalium (Dr. rer. nat.)

INFERRING CAUSAL INFLUENCES FROM EXPANSIVE DISTORTIONS BETWEEN STATE SPACE RECONSTRUCTIONS

submitted by
ERIK LAMINSKI , M. SC.

Universität Bremen
Fachbereich für Physik und Elektrotechnik
Institute für Theoretische Physik

Gutachter:
Prof. Dr. Pawelzik
Prof. Dr. Bornholdt

SUPERVISOR:
Prof. Dr. Pawelzik

LOCATION:
Bremen

DATE OF SUBMISSION:
June 2023

DATE OF COLLOQUIUM:
07th December 2023

Ohana means family.
Family means nobody gets left behind, or forgotten.

— Lilo & Stitch

ACKNOWLEDGMENTS

Es gibt viele Menschen bei denen ich mich bedanken will, seht mir nach, dass ich mich nicht in der Ausführlichkeit bei euch allen bedanken kann, die ihr verdient habt, aber dann würde dieser Anschnitt niemals enden.

Zunächst möchte ich mich bei meinem Betreuer Klaus Pawelzik für die vielen anregenden Diskussion bedanken, die in dieser Arbeit resultierten. Auch für das Vertrauen in mich in der Lehre bedanke ich mich herzlich.

Ich möchte mich auch bei den Mitgliedern meines Promotionskomitees, Stefan Bornholdt, Christopher Gies & Alberto Garcia Ortiz, für das Interesse an meiner Arbeit und die ermutigenden Worte in der Vorbereitung bedanken.

Vielen Dank Udo und David, nicht nur für die unglaublich tolle Vorbereitung auf das Kolloquium, sondern auch für Ideen, Ratschläge und offenen Ohren über die Jahre. Größte Dankbarkeit gebührt auch Maik: für die vielen Ideen, Ratschläge, Diskussion, türkischen Pizzen, Kaffees, (rauchfreie) Raucherpausen oder einfach nur quatschen. Ich glaube das Geräusch fallender Stifte wird mich noch eine Weile weiter heimsuchen.

Vielen Dank an Daniel, nicht nur für die Grundlage dieser Arbeit, aber auch für deine inspirierende Art. Vielen Dank auch an Alina und Federica für die wundervolle Zeit mit euch. Devo anche scusarmi per qualsiasi sigaretta aggiuntiva che ti ho indirettamente causato di fumare. Ich habe den Fasan wieder gesehen!

Ganz viel Dank an all die wundervollen Menschen, die ich über die Jahre im und ums Cognium kennenlernen durfte und für die vielen schönen Erinnerungen. Ich hoffe niemanden zu vergessen, vielen Dank: Axel, Nergis, Dmitriy, Ronja, Mahbod, Archili, Katharina, Foroogh, Enrique, Hendrik, Lisa, Mohammad, Lenny, Amy, Lynn, Monty. Ein extra Dankeschön geht noch an Agnes für all die großen und kleinen Dinge, die du für alle machst.

Vielen Dank an meine Freunde für Ablenkung, moralische Unterstützung und das ein oder andere Bierchen. Dankeschön Lasse, Julian, Lukas, Ronald, Nele, Kolja, Malte, Flo, Olli, Olga, Mammut, Timi, Jakob, Jante, Kevin, Sebi, Meentje, Haller.

Vielen Dank an meine Mannschaftskollegen und Freunde von Bremen 1860 Rugby und generell 1860. Ihr seid zu viele, als dass ich

alle aufzählen kann und will - außerdem weiß ich nicht wie man Beppini schreibt. Entschuldigt meine Abwesenheit in letzter Zeit auf dem Platz - ich gelobe Besserung. Vielen Dank für die gemeinsamen Freuden und Leiden, die eine willkommene Abwechslung zur Dissertation waren.

Schlussendlich gebührt mein Dank noch den allerwichtigsten Menschen: Ganz besonders möchte ich mich bei meiner Familie, meinen Eltern und meinem Bruder bedanken. Danke für Alles. Danke für eure Geduld. Danke für euer Verständnis. Danke für eure Ermutigung. Danke für euer Interesse. Danke dafür, dass ihr mich zu dem Menschen gemacht habt, der ich heute bin. Danke für eure endlose und niemals endende Unterstützung, ohne die diese Arbeit nicht möglich gewesen wäre.

ABSTRACT

State space reconstructions of nonlinear dynamical system contain within their metric and topological properties information about the causal influences between different observables. The expansive distortions among different observables not only reflect the directed coupling strengths, but also the dependency of effective influences on the systems temporally varying state. Estimation of expansions from pairs of time series is straightforward, either directly from intra-neighborhood relations or the mapping between reconstructions. Two approaches to compute expansive distortions are demonstrated using analytical and numerical analysis in a range of complex dynamical systems. The biggest challenge for the inference of causal influences is reached in synchronising systems or system perturbed by large amounts of noise. Remarkably, expansive distortions no only give insight into just the interaction scheme, but provide a time-dependent measure for these interaction. These new methods offer a potential tool to gain insight into interactions of (nonlinear) dynamical system for a wide range of disciplines.

PUBLICATIONS

Some ideas and figures have appeared previously in the publications listed below. Chapter 4 covers [1] - details on which parts and figures are reused, as well as my contributions to this publications are given at the beginning of said chapter. Chapter 5 deals with the entirety of [2] while providing additional details.

CONTENTS

LIST OF FIGURES

ACRONYMS

TC topological causality [1]
CPM cross projection method [2]
WGC Granger Causality [3]
CCM convergent cross mapping [4]

REOCCURRING NOTATIONS

$\mathbf{r}^x(t)$ state space reconstruction of time series $x(t)$, given by $\mathbf{r}^x(t) = [x(t), x(t-\tau), ..., x(t+(m-1)\tau]$

$L^i_j(k)$ size of a neighborhood: neighbor relations computed in i and neighborhood size measured in j

$d^i_j(k)$ logarithmic neighborhood size, with with neighbor relations computed in i and neighborhood size measured in j

1

MOTIVATION & INTRODUCTION

Nothing without cause, nothing without effect.

(Plutarch, 47-120 AD)

Why? - Trying to answer this simple question is a permanent companion in the human pursuit of knowledge and wisdom. Aspiration for insight and comprehension of our surrounding world is fueled by curiosity and the desire to answer this question.

Ancient natural philosophers, like Aristotle [5], did not limit themselves to the qualitative observation of nature, but strove to understand the underlying reasons. This interlacing of **causes** with their resulting **effects** enables inference and understanding of the underlying mechanisms. Understanding permits the selective control and manipulation of **causes** to achieve different **effects**, i.e. knowledge about **causes** for improved growth of crops provides more bountiful harvests.

But singling out the sole one fundamental **cause** of an event is difficult. Nevertheless, we possess an intuitive understanding of **causality** - a windmill does not cause the wind to blow, but is spun by the wind. Our intuition is not only based upon the correlated occurrences of wind and spinning windmills, but is supported by additional information. In fact, we are capable of stopping the windmill and observing whether the wind will continue blowing, supplying the needed auxiliary information to identify wind as a **cause** and spinning windmills as an **effect**.

Interventions are possible for (simple) man-made objects, like a windmill, but might be impossible, undesirable or too risky in other cases. We can not impact the motion of celestial objects. We do not want to change ecosystems. We can not risk to impact our climate systems in a possibly more harmful way. We want to identify possible **causes** and **effects**, but only by observing - without intervening.

Indeed, hypothetical interventions are the basis for Judea Pearls mathematical formulation of **causality** [6]. Here, counterfactual interventions uncover causal relations if a suitable structural model of a system exists. Judea Pearl covers this theory in [6] and also also provides a deeper dive into the more philosophical history of causality.

Nevertheless, there is a plethora of methods to infer *causal influences* from just observations. In general, multiple variables of a system are observed over time and the resulting time series are analyzed. Correlation measures [7–10] evaluated the simultaneous occurrence of events, but require additional information to infer *causal influences*.

Thus it is advantageous to employ directional methods. The most established is Wiener-Granger-causality [3, 11] inferring causal links from the prediction of future states. An alternative are for example methods originating from information theory, e.g. transfer entropy [12].

We developed two methods to infer *causal influences*, *topological causality* (TC) [1] and the *cross projection method* (CPM) [2]. The starting point for these methods is the work of Sugihara et al. [4, 13] that combines the prediction of future states with the reconstruction of the state space to infer *causal influences*. We were able to show that *causal influences* can be directly inferred from just the underlying properties of dynamical systems resulting from Takens theorem [14].

In essence, this is based on the following consideration: Let a system X unidirectional influence another system Y. Obviously, Y receives information about X. Furthermore, states of Y will contain information about states of a X, while states of X by assumption cannot have 'knowledge' about Y. For certain classes of dynamical systems Takens' Theorem [14–17] allows the reconstruction of the underlying state space. Due to incoming information from X subsystem Y can not only reconstruct its own dynamic, but more importantly the overall dynamic of (X, Y). If information is also transmitted from Y to X, X is also able to reconstruct the overall dynamic. Here, our key insight is that expansive local distortions of neighborhoods reflect the *causal influence* between system components X and Y [1].

The expansive distortions are already visible in a simple example: In Figure 1 the reconstruction of the state space of a system comprised of the subsystems Y (left) and X (right) are shown. On the one hand, (sub-)systems are unilateral coupled from X to Y (upper row) and on the other hand the (sub-)systems are uncoupled. In both cases, a generic reference point is chosen in Y and its neighbors are shown in blue in the state space of Y. Here, both neighborhoods are of somewhat identical size. Each point in Y is linked by its time index t to a corresponding point in X, these points are marked in blue in the state space of X (right column). There are clear differences in the distribution of these point for coupled and uncoupled systems. For an existing coupling the reconstruction using variables from Y uncovers the underlying manifold of the overall system (X, Y). Thus, the searched neighbors remain somewhat close in X. In contrast, for the case of no coupling the reconstruction of Y only contains its own dynamic - the distribution of neighbors in X is random.

This random distribution of neighbors equates to the maximum expansion and our proposed methods infer no *causal influence*. *Iff* there are interactions from X to Y the *causal influence* is larger than zero and

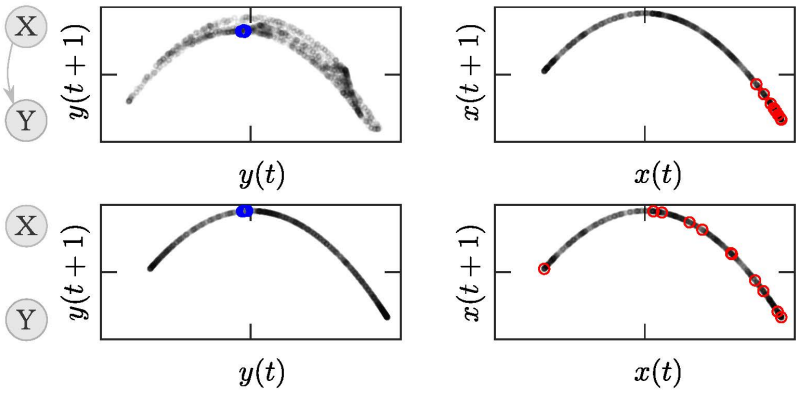


Figure 1: State spaces of an exemplary dynamical system, comprised of two subsystem X and Y. The reconstructed state space of Y is shown in the left column and for X in the right column. The upper row is a system with an unilateral interaction X to Y and the bottom row a system with no interactions. Blue dots signify the nearest neighbors of a reference point (left column) and their corresponding putative neighbors are shown in red (right column).

we propose that the local distortions of the neighborhood size are a suitable measure for *causal influence*.

Later, Section 3.6 will pick up upon this foundation of our work again, but first of all some fundamentals are needed that the first chapter of this work provides. The first half is dealing with a selection of required fundamental properties and terms from the field of dynamical systems. While the latter half elucidates the astonishing consequences of Takens' theorem and its extensions.

The following chapter gives an overview of methods inferring *causal influences* from time-series. Finally, it covers the aforementioned insight that expansive properties are a proxy for *causal influences* in more detail.

The third chapter deals with our first proposed method *topological causality* (TC) [1] published by Harnack, Laminski, Schünemann and Pawelzik. Initially, it is explained how to measure expansive properties directly from the mappings between state space reconstructions. Then a selection of successful applications is provided and the chapter ends with an example of a limit-cycle system for which TC fails to uncover the correct *causal influences* without additional modifications.

Ultimately, this lead to further improvements of *topological causality* and resulted in our second approach, the *cross projection method* (CPM), published by Laminski and Pawelzik [2]: Expansive properties can be more robustly estimated by not looking at expansions *among* different state space reconstructions but instead *within* each

reconstruction, respectively. The first part of this chapter elucidates this new method, while a second part deals with a selection of applications. Then, a set of suitable modifications and design choices are introduced. These modification are not only suitable for our own work, but can prove suitable for other related methods and were advantageous when dealing with synchronizing systems.

The last chapter of this work discloses some final thoughts, integrates the work in the existing literature, discusses future possibilities and highlights challenges.

2

DYNAMICAL SYSTEMS & TAKENS THEOREM

In the midst of chaos, there is also opportunity.

(Sun Tzu)

Before delving into methods for the detection of *causal influences*, some basics principles will be introduced in this chapter. The concepts for the detection of *causal influences* we developed - and relatable methods - are based on insights from the fields of *chaos theory* and *dynamical systems*. The central idea is the reconstruction of the state space using Takens' Theorem [14]. In the first part of this chapter we will focus on *dynamical systems*, while the second part focuses on Takens' Theorem.

The dynamic evolution of systems is imperative in the understanding and modelling of phenomena in all natural sciences. Applications range from pendulums in the Newtonian Mechanics, over predator-prey-relations in population dynamics [18] to the modeling of neuronal dynamics [19] in neuroscience. One of the first contributions is the work of french mathematician Henri Poincaré on the classical three body problem of celestial mechanics [20]. Broadly speaking, a dynamical system is a mathematical model describing the time evolution of a system. In general, a dynamical system can be depicted in the state space, where each point in time is represented by a point in state space. The evolution in time forms trajectories in state space, that are the basis for a range of characteristic properties of the dynamical system. For example, the analysis of stable trajectories identifies attracting structures in state space, referred to as *attractor*.

Often, the complete state of a dynamical system is not observable, but rather only a selection of variables is measurable. The fundamental insight of Takens is that neither all degrees of freedom, nor the underlying equations of the dynamical systems are necessary to construct the state space [14]. Indeed, a single observable and its delayed copies are sufficient to construct a homeomorphic copy of the attractor (of the dynamical system) in state space. Remarkably, this also works if the underlying dynamical system is unknown, although some characteristics like the dimension and time-delay have to be estimated.

Along the way this chapter is introducing a set of typical dynamical systems, e.g. the logistic map [21] and the Lorenz attractor [22]. Initially, these systems are illustrating basics concepts in this chapter, but we will encounter them in the following chapters again. Due to their simplicity and their relevance in ongoing research they provide an ideal testing ground for the detection of *causal influences*.

2.1 DYNAMICAL SYSTEMS

An (autonomous) dynamical system is linking something - typically the state ξ of a system - with an update rule for the evolution over time t. The update rule f for discrete time $t \in \mathbb{Z}$ is called a map and iterates the next states ξ_1 from an initial state $\xi(0)$.

$$\xi_1 = f(\xi(0)), \qquad \xi_t = f^t(\xi(0))$$

The state at time t is given by repeated application of the map t-times. For processes in the real physical world time is in general viewed as a continuous quantity - the time evolution is typical not discrete, but rather continuous in time, i.e. $t \in \mathbb{R}$. Here, the time evolution is given by the flow $\xi_i(t) = \Omega_t(\xi(0))$ or rather the rate of change of a state $\frac{d\xi(t)}{dt}$ - the velocity of the state. The typical formulation is a differential equation linking the current state $\xi(t)$ and its velocity:

Although, one might argue it is not.

$$\dot{\xi} := \frac{\xi(t)}{dt} = F(\xi(t)), \qquad \xi(t=0) = \xi_0$$

Combining this differential equation with an initial state $\xi(t=0) = \xi_0$ is called an *initial value problem*. The solution of this problem describes the time evolution of a single initial state ξ_0. The more general concept is the flow Ω_t of F, a mapping with identity relation, group law and Lipschitz continuity describing the global time evolution - the evolution of all possible initial states ξ_0.

One of the simplest examples of a *dynamical system* in (Newtonian) mechanics is the free undampened (mathematical) Pendulum Figure 2. The pendulum consist of a mass m attached to a mass-less string of length l and is under the effect of the the gravitational force $\vec{F} = m\vec{g}$. The deflection of the pendulum is given by the angle $\varphi(t)$, but the angle $\varphi(t)$ is not the complete state $\xi(t)$ of the pendulum. For example consider the position of the pendulum depicted in Figure 2, from the deflection alone it is not even possible to determine if the pendulum is moving counter- or clockwise. The pendulum has two degrees of freedom and thus the state is a tuple of two variables $\xi(t) = (\varphi(t), \omega(t))$, the angle $\varphi(t)$ and angular velocity $\omega(t)$. Consequently two ordinary differential equations are governing the time evolution of the (free friction less) pendulum:

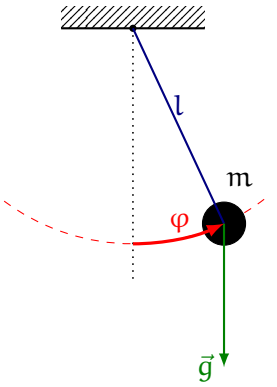


Figure 2: Schematic of the mathematical pendulum

$$\begin{aligned} \dot{\varphi}(t) &= \omega(t) \\ \dot{\omega}(t) &= -\frac{g}{l}\sin\varphi(t) \end{aligned} \tag{1}$$

These two equations and an initial condition fully describe the *state* of the pendulum for each point in time t, the first equation indicates the change in angle $\dot{\varphi}$, *angular velocity*, and the second the change in angular velocity $\dot{\omega}(t)$, *angular acceleration*. The solution of these differential equations enables one to derive all future states from one known state, e.g. $t = 0$. This state $(\varphi(t), \omega(t))$ also includes information about all past states and thus the whole previous time evolution of the dynamical system. This is especially evident if the space of all possible states is considered - the *state space*.

The mathematical pendulum for example can be simplified by small angle approximation or solved using elliptical integrals [23].

2.1.1 *State Space*

The aforementioned pendulum has a two dimensional state space spanned by the angle $\varphi(t)$ and the angular velocity $\omega(t)$ Figure 3, which represents all combinations of $\varphi(t)$ and $\omega(t)$. Figure 3 **(a)** shows the state space and one initial condition (red dot) of the pendulum, its time evolution is forming a circle-shaped (clockwise) trajectory in state space. After one complete oscillation the pendulum returns to its initial state (red dot) forming a closed ellipsoid in state space - states on it will remain evolving over time only along this path. However, this trajectory is not only unique for this initial condition, but rather all states on its path are valid starting states that share the same conserved total energy. A selection of typical trajectories in state

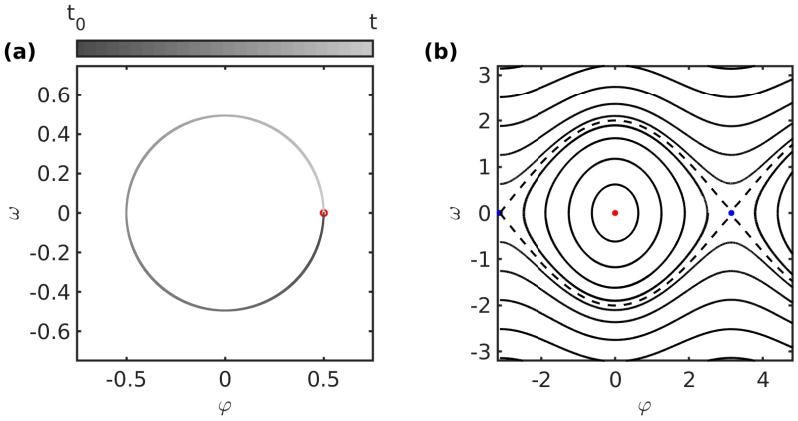


Figure 3: **(a)** State space (ω over φ) of the mathematical pendulum. A red circle signifies the initial state $(\varphi_0, \omega_0) = (0.5, 0)$ and the gray gradient shows the time evolution of the state, with darker gray corresponding to moments in time close to the initial state.
(b) Phase portrait of the mathematical pendulum. The dashed line marks the boundary between closed (oscillating) and open (circulating) trajectories. Colored dots mark stable (red) and unstable (blue) fixed points of the pendulum.

Both figures are generated by solving Equation 1 numerically (using a Runge-Kutta algorithm).

space (for the mathematical pendulum) is shown in Figure 3 **(b)**. This so called phase portrait enables to study the whole dynamic range

A finite dimensional state space, e.g. the manifold mappings or ordinary differential equations, is often referred to as phase space.

of a system, for the pendulum especially the behaviour for different total energy is captured. The trajectory from **(a)** and similar closed trajectories are delimited by dashed lines.

Confusingly circular trajectories in state space are not circular pendulum motions.

All closed trajectories in this area represent oscillations of the pendulum, i.e. the typical pendulum swings between maximum amplitudes. The trajectories outside of this area are circulating solution. Here, the energy is sufficient that the pendulum is rolling over its attachment point and circulates around it. Finally, two more trajectories are marked in red and blue colors - the two fixed points of the pendulum. Fixed points of a dynamical system are states invariant under time-evolution. The red point is the resting position of the pendulum and a stable fixed point, the blue point is the upside down position of the pendulum - an unstable fixed point.

2.1.2 *Discrete and Contionous Systems*

The aforementioned pendulum has a continuous analytic solution. Differential equations not necessarily have an analytic solution or it is more practical to evaluate them numerically and thus convert a continuous system into a discretized counterpart.

Consider for example the logistic function used by Pierre F. Verhulst to model population growth [24]:

More precisely $\frac{dN}{dt} = rN - \alpha N^2$, but we use the dimension-less case.

$$\frac{dx(t)}{dt} = Rx(t)(1 - x(t)) \tag{2}$$

The initial step to discretize a system is the choice of a appropriate discretization, instead of continuous time, appropriate time steps (for this system), e.g. of length $\Delta t = 1\,s$, are chosen. By choosing a discretization, the derivative $\frac{dx(t)}{dt}$ can be approximated by the (forward) finite difference $\frac{dx(t)}{dt}|_{t=t_0} \approx \frac{x(t_0+\Delta t)-x(t_0)}{\Delta t}$ and the continuous variables can be evaluated at the chosen time steps $x(t_0)$. Some arithmetic manipulations [25] and renaming yields the equation for the time-discrete logistic map [21]:

$$x(t+1) = rx(t)[1 - x(t)] \tag{3}$$

The shown equations appear analogous, but have distinct characteristic. For example, the logistic map (Equation 3) exhibits chaotic behavior in certain parameter ranges, while the continuous one dimensional differential equation (Equation 2) can not show chaotic behaviour [26]. The discretization of a system is not a trivial operation and the properties of the system can change, especially if a poor discretization was chosen. To avoid this, the later discussed systems of differential equations are integrated using a fourth order Runge-Kutta algorithm [27] and a small step size.

Small time steps means small in respect to the time scale of interest.

2.1.3 *Determinism and Chaos*

A property of dynamical system that was only discussed implicitly is *determinism*. The future states of a deterministic system are generated without randomness, intuitively this is true for the aforementioned pendulum (and the logistic map). Here, the pendulum position and velocity and its governing differential equation are sufficient to determine the future uniquely.

However, a deterministic system does not have to be fully predetermined, non-linearities of the dynamic and/or uncertain initial conditions can cause seemingly random erratic behaviour, i.e. diverging trajectories in state space. This sensitive dependence on initial condition is one of the formal criteria for a chaotic system [26]. Alongside determinism chaotic system feature aperiodic long-term dynamics.

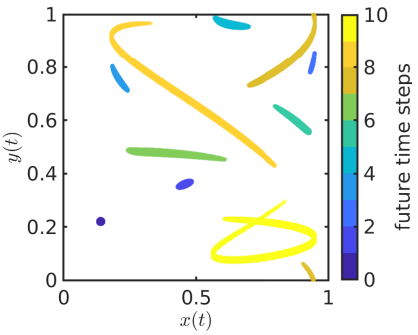


Figure 4: Sensitive dependence on initial conditions of the modified logistic map $x(t_1) = rx(t)[1-x(t)]$ & $y(t+1) = \mathrm{mod}_1[x(t)+y(t)]$. The color-bar denotes time steps after the initial state.

For example consider a two dimensional deterministic system, the first component being governed by the logistic map $x(t+1) = rx(t)[1-x(t)]$ and the second by $y(t+1) = \mathrm{mod}_1(x(t)+y(t))$. If we choose a state and the area around it in state space as a set of initial conditions (dark blue area in Figure 4) and evolve each initial condition over time, already a few time steps will be sufficient to spread the state over the state space. This sensitive dependence on initial conditions is typical for chaotic systems and can be quantified by the exponential divergence of trajectories - the Lyapunov exponent [28].

2.1.4 *Strange Attractor*

Studying the pendulum we already encountered an attractor and its counterpart the repellor - the fixed points marked in Figure 3. If a dampened pendulum is considered, it is even more straightforward what characterises an attractor. Here, the dissipative nature of the system will cause the it to strive towards the resting position (for large time scales) - a single point in state space attracting all trajectories.

Besides attracting fixed points there are also attracting orbits in state space. A simple example is a pendulum clock, where the driving force by the escapement mechanism counteracts the dissipative energy loss via friction and thus generates an attracting periodic orbit in state space.

Originated by David Ruelle and Floris Takens describing fluid flows.

The final type of attractor is the so called strange attractor , which will we will consider later extensively in form of the Lorenz and Rössler systems. A strange attractor has a fractal structure in state space, often times this fractal structure is based on the chaotic nature of the underlying dynamical system. Interestingly, trajectories in state space will converge to the attractor and become arbitrarily close to one another, but due to the chaotic nature two initial points will also diverge over time, but will not diverge from the attractor. This leads to the fascinating property of a globally stable, but locally unstable system.

There are also strange non chaotic attractors.

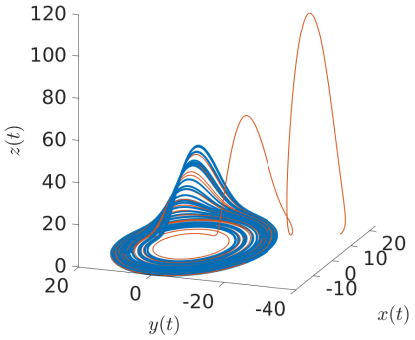


Figure 5: The attractor of the Rössler system (blue) and a trajectory (red) converging onto the attractor.

This section is merely supposed to be a brief insight; there is a plethora of textbooks covering this and related topics, e.g. [29]. As an example for an strange attractor consider the following system of three ordinary differential equations:

$$\begin{aligned} \dot{x} &= -(y+z) \\ \dot{y} &= x + ay \\ \dot{z} &= b + (x-c)z \end{aligned} \tag{4}$$

Equation 4 Known as the Rössler attractor [30] this system has be influential in the study chaos. In Figure 5 the Rössler attractor is shown in state space (blue). Furthermore we see the trajectory (red) of a initial conditions striving towards the attractor.

2.1.5 *Dimension*

Before we can talk about the center piece of this section, *Takens Theorem*, we have to briefly discuss *dimensions*. Not only is this vital for *Takens Theorem*, but we are able to infer information about the causal structure of a system from just the dimension of subsystems.

The previously shown Rössler system consists of three differential equations, intuitively we used a three dimensional euclidean space $(x(t), y(t), z(t))$ to visualize the system. But in fact, we have to distinguish between the space in which an object is embedded, and the

dimension of an object. The latter is an intrinsic property of an object and attributed to the degrees of freedom, i.e. the number of independent coordinates/parameters. Consider a (straight) line and a circle, both are one dimensional objects - one independent coordinate is sufficient to describe any position on the circle. But a two dimensional euclidean space is necessary to embed, i.e. visualize, the circle. Even worse, objects do not necessarily have an integer dimension, the Rössler system for example has neither a dimension of two nor a dimension of three.

Smooth continuous time dynamical systems need at least a dimension of two to exert chaotic behavior.

The classical example to derive the dimension of a fractal object is the British coastline [31]. Consider a circle and the coastline, both are measured with an increasingly more precise ruler. In the case of the circle the measurement will approach the correct circumference, but the length of the coastline will keep increasing with smaller and smaller gaps measured. Accounting for this, the coastline is covered in increasingly smaller square boxes. In doing so, one is able to derive the fractal dimension of the coastline. Because the number of boxes $N(\epsilon)$ necessary to cover a a one dimensional object increases linear with the size ϵ of said box, for a two dimensional object quadratic. In fact for a D-dimensional object N scales exponential with the dimension $N(1/\epsilon) \approx \epsilon^D$. The box counting dimension [32] is defined by:

Circles are more general for higher dimensions.

$$D_{box} = \lim_{\epsilon \to 0} \frac{\ln(N)}{\ln(1/\epsilon)} \tag{5}$$

where ϵ is the size of boxes and N the number of boxes.

Instead of an object, it is more interesting to consider the dimension of a (state) space filled by point clouds. Similar to the box counting dimension we can derive the dimension these points clouds. By covering the whole space with a multidimensional checkerboard of size ϵ. If we count the probability $P(\epsilon)$ of a checkerboard cell being populated while varying the size ϵ, we can define the information dimension [33]:

$$D_{inf} = \lim_{\epsilon \to 0^+} \sum_{i=1}^{N} \frac{P_i(\epsilon) \ln[P_i(\epsilon)]}{\ln(\epsilon)} = \lim_{\epsilon \to 0^+} \frac{< \ln[P_i(\epsilon)] >}{\ln(\epsilon)} \tag{6}$$

An added advantage of the information dimension is that it is derivable from basic properties of our measure for *causal influence* (Chapter 5).

2.2 STATE SPACE RECONSTRUCTION

In the previous section we covered the state space of well known systems, i.e. systems already modeled by mathematical equations or dynamics resulting from their known underlying equations. For (chaotic) dynamical system, observed in nature or studied in experiments, the full set of governing equations is often not available. But rather single or multiple variables are observed as a function of time. Already this seemingly incomplete information about a system can be sufficient to (re-)construct the state space of an attractor. In the following sections we will look at Takens' Theorem, its extensions by Sauer[17] and Stark [16, 34] and the foundation by Whitney [35]. Using these concepts the state space of a strange attractor can be constructed from a single measured variable.

2.2.1 *Embedding*

The basis for state space reconstruction is Whitney's embedding theorem [35], stating: 'Any smooth manifold $\mathcal{M}$ of dimension m can be embedded into $\mathbb{R}^{2m}$'. Meaning that the image of a m-dimensional manifold is unfolded in a $2m$ dimensional euclidean space - no points are mapped to the same point in euclidean space preserving its structure. Being more graphic one can consider the symbol '8', on this plane (two dimensional) paper there is a self intersection in the middle. An additional dimension is needed to completely unfold it. In the case of a m-dimensional system, there is a set of $2m$ independent observables that can be considered a map from the manifold of the system and thus Whitney's theorem implies that these observables are sufficient to reconstruct the state space.

2.2.2 *Takens' Theorem*

Floris Takens showed in 1981 [14] that only a single observable is necessary to reconstruct the state space. Instead of $2m+1$ signals $2m+1$ delayed copies $[x(t), x(t-\tau), x(t-2\tau), ..., x(t-(2m+1)\tau)]$ of a single observable are sufficient.

As a first intuition, consider a simple system like the pendulum that is governed by a set of equations of motion that only consists of position $\alpha(t)$ and its derivatives. A state in state space is given by the vector $[\alpha(t), \dot{\alpha}(t), \ddot{\alpha}(t), ...]$. The angular velocity $\dot{\alpha}(t)$ can be assessed via the difference quotient, similarly higher orders of derivatives can be assessed, thus the vector becomes

$$[\alpha(t), \frac{\alpha(t)-\alpha(t-\tau)}{\tau}, \frac{\alpha(t)-2\alpha(t-\tau)+\alpha(t-2\tau)}{\tau^2}, ...].$$

It is straightforward that for this simple system only delayed copies $\alpha(t)$ are necessary to construct the state space. In general, these copies do not have to be in the form of difference quotients, but rather $[\alpha(t), \alpha(t-\tau), \alpha(t-2\tau), ...]$ are sufficient.

As an example let us consider the three dimensional Lorenz system [22], governed by the following equations:

$$\begin{aligned} \dot{x} &= \sigma(y-x) \\ \dot{y} &= x(\rho-z)-y \\ \dot{z} &= xy-\beta z \end{aligned} \tag{7}$$

The system originates from meteorology and shows for a typical set of parameters (σ, ρ, β) a famously butterfly shaped attractor in state space, as shown Figure 6. This shape still remains visible if the state space is reconstructed using only one coordinate, e.g. $x(t)$ or $y(t)$ as shown in Figure 6. The trajectory in the space spanned by $x(t)$, $x(t+2\tau)$ and $x(t+2\tau)$ is shown in **(b)** and is homeomorphic to the original space meaning that all topological properties are preserved

In fact, due to symmetry $z(t)$ is an exception of the initial Takens' theorem.

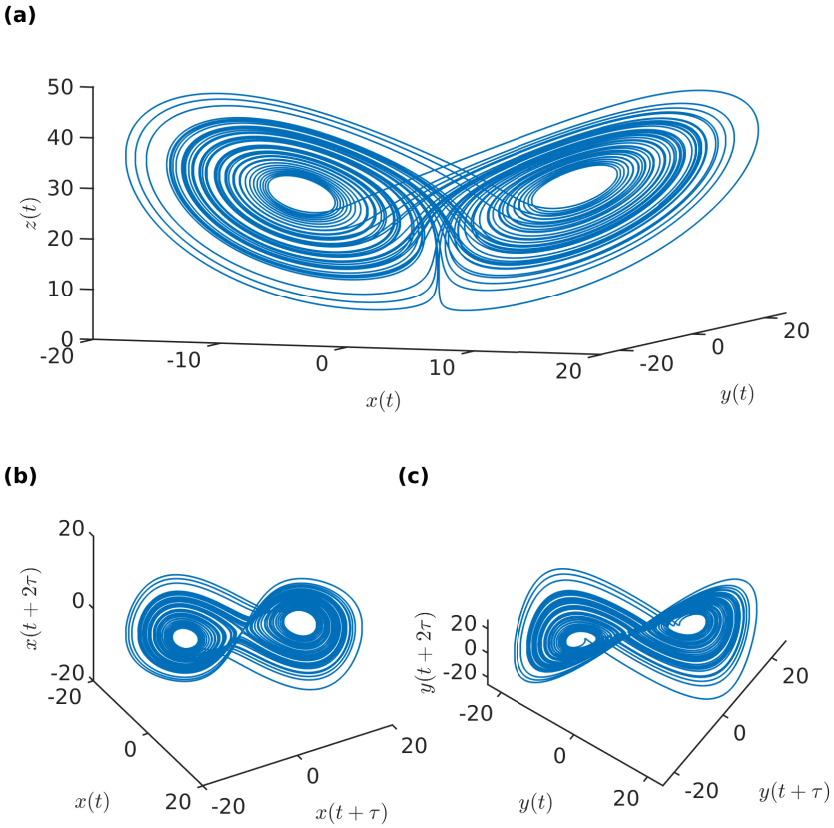


Figure 6: Lorenz attractor for $(\sigma, \rho, \beta) = (10, 28, 8/3)$ **(a)** and reconstructions of the Lorenz system from $x(t)$ **(b)** and $y(t)$ **(c)**.

The typical example for a homeomorphism are a donut and coffee cup.

in $[x(t), x(t+\tau), x(t+2\tau)]$ and the two shapes can be transformed into one another just by continuous stretching and bending.

An important consequence is the preservation of neighborhood relations, close neighbors in the original space (x, y, z) remain close neighbors in the reconstructed space $[x(t), x(t+\tau), x(t+2\tau)]$. This property is vital for defining causality measures in the following Chapter 3. Before continuing extensions of Takens theorem shall be considered briefly and the computation of parameters for the embedding are discussed. The theoretical requirements for embedding parameters are straightforward, but for practical purposes it is necessary to discuss the choice of embedding parameters. On a side note, it shall also be remarked that similar results were separately obtained by Packard et al. [15] and Aeyels [36].

2.2.3 *Mathematical Formulation*

The prior section introduced Takens theorem heuristically for the example of a Lorenz system. In fact, Takens' theorem is not a single theorem, but rather a collection of theorems for the embedding of continuous, discrete or discretization of continuous dynamical systems.

The time evolution for the continuous Lorenz system ($t \in \mathbb{R}$) in Figure 6 is described by the set of differential equations Equation 7, or a suitable discretization. Since the aforementioned Lorenz system is a continuous dynamical system with a time evolution governed by the flow $t \to \varphi(f_t(x)), x \in \mathcal{M}$, we will only show Takens Theorem for this type of system. The full set of theorems and their respective proofs are found in the original manuscript of Floris Takens [14]. In particular, the formulation for discrete system which will be our focus when dealing with simulated or experimental sets of data.

Typically, not the full set of variables is recorded by an observer outside of the system. An observable is a smooth function $\varphi : \mathcal{M} \to \mathbb{R}$, typically of one or a subset of system variables. In practice the observable $\varphi(t)$ can be contaminated by (external) measurement noise, e.g. $\varphi_t = f_t(x_t) + \eta_t$. The problem is this: if, for some dynamical system with time evolution $f_t()$ we know the functions $t \to \varphi_t(f_t(x)), x \in \mathcal{M}$, then how can we obtain information about the original dynamical system (and manifold) from this.

Theorem 1 (Takens Theorem [14]) *Let $\mathcal{M}$ be a compact manifold of dimension m. For pairs (f, φ), $f : \mathcal{M} \to \mathcal{M}$ a smooth diffeomorphism and $y : \mathcal{M} \to \mathbb{R}$ a smooth function, it is a generic property that the map $\Phi_{(f,\varphi)}(x) : \mathcal{M} \to \mathbb{R}^{2m+1}$, defined by*

$$\Phi_{(f,\varphi)}(x) = [\varphi(x), \varphi(f(x)), ..., \varphi(f^{2m}(x)] \tag{8}$$

is an embedding; by "smooth" we mean at least $\mathcal{C}^2$.

Note that the notation is changed in comparison to the original formulation.

Takens Theorem on its own is only valid for deterministic dynamical systems, more specifically delay coordinate maps. Covering a wider range of (dynamical) systems Sauer et al. generalised Takens' Theorem to fractal dimensions, enabling the state space recovery of fractal dimensional dynamical systems like the Lorenz-system [17]. Two further extensions validate Takens' Theorem for externally forced and stochastic systems [16, 34]. For the case of external stochasticity - measurement noise - filtering is useful to improve the state space reconstruction. There are also more complex filtering schemes like the so called Kalman-Takens algorithm [37] combining the recovery of unknown (underlying) variables and the reconstruction of the state space of said variables.

2.2.4 *Estimating the embedding parameters*

It is not by chance that the butterfly-shaped Lorenz attractor is clearly reflected in its reconstructions (Figure 6), the embedding delay τ was chosen carefully to fully unfold the attractor. For chaotic oscillators, like the Lorenz systems, a fourth of the average period is a good heuristic for an embedding delay. In general, a small embedding delay τ is preferable to prevent the attractor folding onto itself, theoretical assumptions what constitutes a good τ are discussed in [38].

In practice, a small τ still guaranteeing a linear independence of $x(t)$ and $x(t+\tau)$ is required. We have chosen the first minimum of the average mutual information [39] between $x(t)$ and $x(t+\tau)$ as a criterion for independence. Broadly speaking, mutual information measures how much information about $x(t)$ is uncovered by knowing $x(t+\tau)$, and vice versa. The first minima (of mutual information) depending on τ constitutes an embedding delay, for which $x(t)$ and $x(t+\tau)$ are sufficiently independent.

For the Lorenz system (Figure 6) τ was derived exemplary in Figure 11.

For unknown systems we have to also estimate an embedding dimension. For this purpose the *false nearest neighbor* algorithm of Kennel et al. [40] is suitable. A time series $x(t)$ is embedded for different increasing embedding dimensions n and for each embedding the nearest neighbors are searched. A mismatch in the neighbor relations while transitioning from a dimension n to a higher dimension $n+1$ indicates that the attractor is not fully unfolded in n. A suitable embed-

ding dimension m is a dimension for which said neighbor relations do not change.

Alternatively, there are also methods like the wavering product [41] estimating both, embedding delay τ and embedding dimension m. A more extensive and detailed discussion of suitable methods is provided in the overview paper by Bradley et al. [42].

There are also specialized methods to improve the reconstruction in certain cases. If multiple variables are observed, it can be advantageous to use mixed embeddings [43] from two or more variables. In the case of the aforementioned Lorenz system, a possible mixed embedding would be $(x(t), x(t+\tau), y(t+2\tau))$. Besides that, a non uniform choice of τ, i.e. $(x(t), x(t+\tau_1), y(t+2\tau_2))$, is improving the reconstruction if different timescales are present [44]. For the scope of this work, we restricted the reconstruction of the state space to an uniform embedding using a single observable. The embedding parameters (τ, m) are estimated using mutual information and a false nearest neighbor algorithm.

2.3 CONCLUSION

The purpose of this chapter was twofold. On the one hand, basic concepts from the field of dynamical systems were introduced. On the other hand, Takens theorem was introduced providing the basis for Chapter 5 and Chapter 4. The central idea is the possibility to uncover the topology, dynamics and neighbor relations of an underlying dynamical system from a single observable.

In the following pages, the observables are two scalar time-series $[x_(1), .., x(t), .., x(N)]$ and $[y(1), .., y(t), .., y(N)]$ of length N either from different (sub-)systems or from a joint dynamical system. Since the time evolution $f_t()$ of the time-series is implicitly included in the notation $x(t)$ and the measurement function is assumed to be $\varphi_t(x(t)) = x(t)$, the notation of Equation 8 is simplified. The embedding or state space reconstruction of an observed time-series $x(t)$ is given by:

$$\mathbf{r}^x(t) = [x(t), x(t+\tau), .., x(t+(m-1)\tau)] \tag{9}$$

In general, the latter sections consider two observables, $x(t)t$ and $y(t)$, and the relation between their respective reconstructions $\mathbf{r}^x(t)$ and $\mathbf{r}^y(t)$ that both share the same time-delay τ and dimension m, which are derived from the respective time-series.

3

CAUSALITY

The most merciful thing in the world, I think, is the inability of the human mind to correlate all its contents.

(H.P. Lovecraft)

The world surrounding us is dynamic and subject to change. In the previous chapter we contemplated a basic ideas of dynamical systems - the state. The state is comprised of all information necessary to progress towards the future. In general, we do not consider the overall state, but the state of subsystems and parts of systems; pendulums, balls or fluids, interacting with their surrounding environment.

A ball does not suddenly start moving due to some intrinsic state, but rather because of external influences, its likely someone kicked it - there is a **cause**. If a telephone is ringing, it is because somebody is calling. The motion of a ball and ringing of a phone are **effects** of their respective **cause** and succeed the **cause** in a predictable way. The temporal sequence always begins with a **cause** and ends with an **effect**. The **effect** can also become a **cause** - our ball in motion hits another ball and initiates its motion. In fact, one could argue chaining together **cause** and **effect** pairs describes a flow of time. This concept of **causality** has a long history ranging back to the beginnings of natural philosophy [5].

Annoyingly, Quantum Mechanics suggest this must not necessarily be the case.

The aim of this chapter is describing how these intuitive concepts and notions of **causality** can be transferred to the context of time series analysis. Already the introduction of this work differentiated in its choice of words between **causality** and *causal influences*. This demarcation is not only intentional, but necessary in the context of time-series. Here, it is not the ultimate goal to identify the single **cause**, but rather elucidate the structure of the underlying system. The inference of true **causality** always requires intervention into the dynamics of a system [6]. While *causal analysis* is only a step to learn or validate existing hypotheses about a system or process in question. The term *causal influence* is strongly linked to the employed method of *causal analysis* and its relation to structural properties of the investigate system, e.g. coupling weights.

The first part of this chapter illustrates that when dealing with time series it is reasonable to consider effective time-dependent *causal influences* to infer interactions. The second part covers a selection of state of the art methods to infer the interaction structure or *causal influences* from time-series. Finally, the last section returns to Figure 1 and the introduction elucidating how interactions between (sub-)systems lead to expansive distortions in the state space reconstructions and

thus preludes the introduction of our methods to infer *causal influences*.

3.1 FROM OBSERVED TIME SERIES TO CAUSAL INFLUENCES

The intuitive understanding of **cause** and **effect** pairs is transferable to an observed time-series (of a dynamical system) and the **cause** is identifiable in said time-series. But already small modifications of a dynamical system complicate the deciphering of **cause** and **effect** directly from the observed time series. Moreover, in the case of time-series not a single event, but rather the interaction between time-series is of interest.

As an example for an observed time series let us review a simple dynamical system, e.g. the pendulum (Equation 1), with a dampening force and perturbed by a short external pulse - a nudging of the pendulum. Instead of the whole system only the angle φ and the time series of the disturbance are observed and shown in Figure 7. For this simple case the temporal sequence is even visually identifiable, the (red) disturbance is succeeded by a dent in the time series of φ. The **cause** - δ-pulse - and the **effect** - the change in angle - are clearly distinguishable. A hypothetical measure of *causal influence*

Time-series were generated by solving the differential equations using a Runge-Kutta scheme

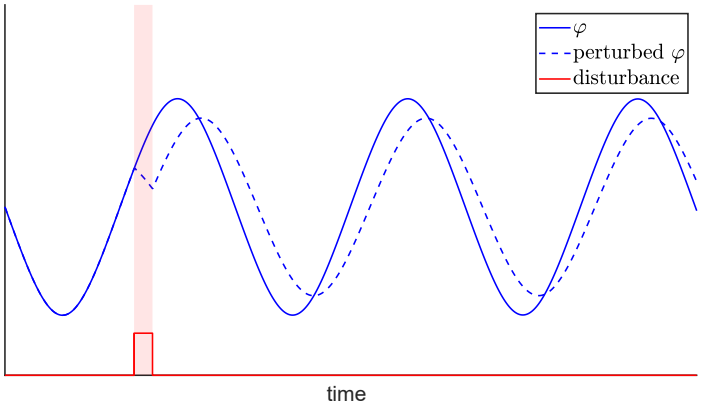


Figure 7: Time series (of the angle φ) of a harmonic oscillator shown in blue. The dashed line shows the angle of the harmonic oscillator disturbed by a delta-pulse (red). The light red area marks the area of causal influence.

C should only detect an interaction in the short period of the disturbance (light red). Otherwise there should be no significant *causal influence* $C = 0$.

Already for a slightly more complex case visual inspection of the time-series is ambiguous. Consider the same system perturbed by a periodic external force, i.e. $F_{external} = F\cos(\omega t)$, shown in red in Figure 8.

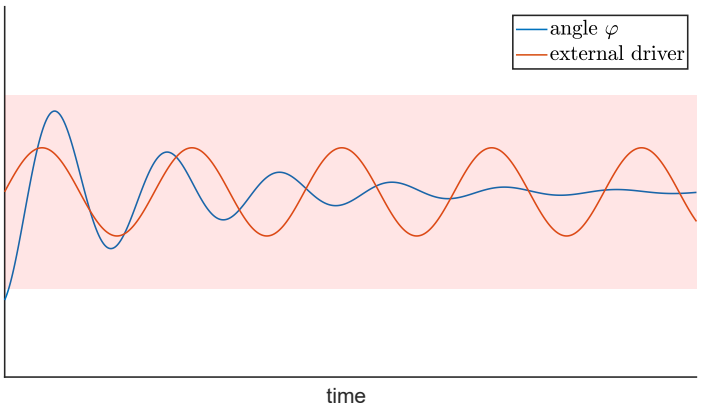


Figure 8: Time series (of the angle φ) of a harmonic oscillator with external driver shown in blue and the time series of the external force (red). The light red area marks a hypothetical area of causal influence from external driver to the angle φ.

Here, the external input forces the pendulum into another frequency. Clearly, the external driver is the **cause** for this change of motion. However, this is not evident from the observed time series alone - especially if only a short section is observed. Simple correlation based methods can quantify that the external forcing and the harmonic oscillator are synchronous, but cannot infer a direction of influence. In contrast, *topological causality* will infer a dominant influence originating from the external driver, at least until the oscillator is fully synchronous with the frequency of the external forcing.

Admittedly, other methods would do so as well.

In addition, the strength of the influence might vary according to the state of the system. This is even more evident for bilaterally interacting system, for example a predator-prey model:

$$\frac{dx}{dt} = x - \alpha xy \qquad \frac{dy}{dt} = -y + \beta xy \tag{10}$$

This Lotka-Volterra-model [18] describes two cross interacting species, where the prey population is inhibited by large predator populations and the predator populations is excited by large prey populations. The shifts in population size are shown in Figure 9.

In this system, individual parts influence each other depending on the system state - predator and prey populations swap between **cause** and **effect** over time. Thus, it is most suitable to describe the influences by a time-dependent index describing the causal relations. As an example the background is the *causal influence* inferred by CPM. This *causal influence* shifts with the state of the systems from time periods of dominant direction of *causal influence* from predator to prey to periods of 'equilibrium' and periods of (dominant) prey to predator influences.

CPM merely serves as a illustrations, thus any details are omitted (for now).

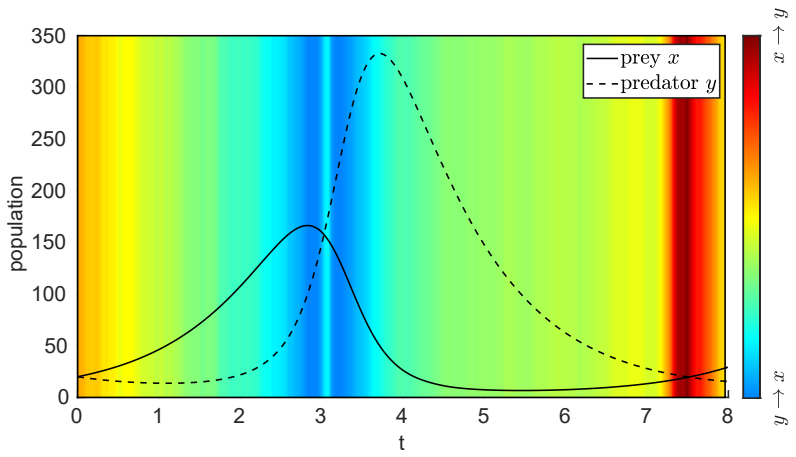


Figure 9: Predator (dashed) and prey (solid line) population of a Lotka-Volterra-model. A causal influence (derived by CPM) is signified by the background color. Red corresponds to strong influences from prey to predator and blue vice versa.

Systems, like the aforementioned predator-prey model, also pose two other common difficulties for a measure of *causal influences*. Apparent interaction can be generated by confounding factors, like environmental variables or shared interactions with other species. Besides that, interactions should remain detectable, if an intermittent (unobserved) part of a dynamical system facilitates the interaction(s).

Subsequently, dynamical systems with known structure are used as a testing ground. For these system the interactions are controlled by coupling weights $w_{i\to j}$, making the relation between these weights and the *causal influence* C relevant. In particular, changing the strength of a weight $w_{x\to y}$ - a stronger influence from X to Y - must be reflected in the *causal influence* $C_{x\to y}$. For small weights $C_{x\to y}$ this is a (linear) function of the weight $w_{x\to y}$.

If the interaction between identical subsystems is assessed, it is meaningful to consider *causal asymmetries* α, to uncover the dominant direction of influence :

The quantity shown in Figure 9 is α.

$$\alpha = \frac{C_{x\to y} - C_{y\to x}}{C_{x\to y} + C_{y\to x}}$$

For strong couplings and/or systems with highly similar dynamics systems parts can synchronize. For complete synchronization statements about *causal influences* are meaningless, since the individual dynamics no longer posses any degrees of freedom and none of the (sub-)systems has unique information. In this case, our methods will infer maximum *causal influence* in both directions. At least, up to complete synchronization the *causal asymmetry* is a reliable quantity.

Interventions in the form of noise or observation of transient dynamics can also prove helpful.

Besides that, there are other, more universal, limiting factors to account for when dealing with observed data. In general, the measure-

ment period and resolution is limited. Thus any method for inference of *causal influences* should need as little data as possible, while a significance criterion controls if sufficient data is available. The quality of data - measurement or intrinsic noise - should have preferably no or a predictable influence. In particular, no false positive *causal influences* should be induced by noise.

3.2 CORRELATION

Evaluating the existence and strength of interactions and influences in observed data traces back to the work of Sir Francis Galton [7]. Here, the correlation coefficient was designed as a purely statistical criterion quantifying the simultaneous occurrence of different features, e.g. different proportions of skeletons studied by Sir Francis Galton. Per construction, correlations are a symmetric measure and do not have to relate to *causal influences*.

3.2.1 *Pearson correlation coefficient*

Originally, the mathematical formulation of said correlation coefficient was published by Bravais [8], but in general, it is referred to as the Pearson correlation coefficient [9]:

$$\rho_{X,Y} = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} \tag{11}$$

The coefficient is composed of the covariance $\mathrm{cov}(X,Y)$ of the data sets X and Y, and their respective standard deviations σ_X and σ_Y. Its limiting cases are independent data sets X and Y equating to $\rho_{X,Y} = 0$ and $\rho_{X,Y} = 1$ for dependent data sets. Two data sets are dependent, if one is an affine transformation of the other.

The fundamental weakness of a linear coefficient are evident in (simple) toy examples, i.e. Anscombe's quartet [45] or the datasaurus dozen [46]. Here, diverse data sets yield the ambiguous result of matching correlation coefficients. Nonlinear measures, like Spearman correlation [10], remedy this limitation.

The data sets have the same simple properties (e.g. mean and variance), but distinct distributions.

3.2.2 *Cross correlation*

As aforementioned, we do not want to consider X and Y as an unsorted collection of data points, but rather time-resolved observables - the time-series $\{x(t)\}$ and $\{y(t)\}$. Here, it is sensible to incorporate temporal features of the data-set. In particular, influences can be evident after some time τ, which is reflected in a time-resolved variant of the Pearson correlation coefficient:

$$\rho_{x,y}(\tau) = \frac{\mathrm{cov}(x(t), y(t-\tau))}{\sigma_x \sigma_y} \tag{12}$$

In essence, $\rho_{x(t),y(t-\tau)}(\tau)$ is the same coefficient as Equation 11, but $x(t)$ and $y(t)$ are shifted against each other by τ. The chronological order of influences is most obvious by normalizing $\rho_{x(t),y(t-\tau)}(\tau)$ over τ.

3.3 WIENER-GRANGER CAUSALITY

A correlated occurrence of events must not and often does not correspond to a **cause** and **effect** pair - *correlation is not causation*. Moreover, a directionless measure, like correlation, is not suitable to determine **cause** and **effect** without additonal information.

Presumably, the most famous and widespread method for detection of *causal influences* is Granger Causality [3] (WGC) [3, 11]. Granger adapted the work of Wiener [47] formulating a (directional) prediction based approach that is based upon two fundamental assumptions:

(a) the **cause** occurs before the **effect**

(b) the **cause** contains unique information about the **effect**

In consequence, a **cause** can be used to predict an **effect**. The basic idea is simple, suppose we want to predict future values of a stochastic process X using all available information U and predict the future using all information, excluding some other stochastic process U/Y. If Y contains unique information about X, the prediction of X will be improved incorporating Y: Y Granger-causes X.

Typically, not all information U, but rather a subset of observed variables is available. Thus, in its most basic form Granger causality is limited to systems that can be modeled, e.g. in its simplest form by (linear) auto-regressive models. A simple bivariate linear autoregressive model is given by:

In general, the model can also include drift terms and noise with arbitrary variance.

$$\begin{pmatrix} x(t+1) \\ y(t+1) \end{pmatrix} = \begin{pmatrix} M_{xx} & M_{xy} \\ M_{yx} & M_{yy} \end{pmatrix} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} + \begin{pmatrix} \epsilon^x(t+1) \\ \epsilon^y(t+1) \end{pmatrix}$$

The future of both time-series is given by a linear combination of past values at time t and additional noise term $\epsilon^i(t) \in \mathcal{N}(0, \sigma^2)$.

In Figure 10 $x(t)$ and $y(t)$ are shown for $M = [0.88, 0.1; 0, 0.98]$ - a unidirectional influence from past values $y(t)$ to future values $x(t+1)$. In the shown time-series peaks and troughs of $y(t)$ are leading their counter-parts in $x(t)$. Thus, it is likely the future peaks in $x(t+1)$ can be better predicted using information from $y(t)$.

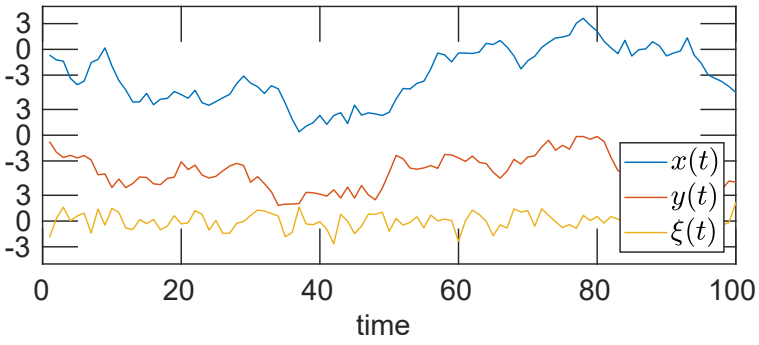


Figure 10: Time series (100 steps) generated by a bivariate autoregressive model using $M = [0.88, 0.1; 0, 0.98]$ and $\sigma_i = 1$. The blue line is the time-series $x(t)$ and the red line $y(t)$, the yellow line shows a random time-series $\xi(t)$ with $\mu_\xi = 0$ and $\sigma_\xi = \sigma_x$.

Quantifying this insight, the time-series $x(t)$ and $y(t)$ are used to fit a model, e.g. a linear vector autoregressive model:

$$\begin{aligned} \bar{x}(t+1) &= \sum_{i=0}^{p} A_{xx} x(t-i) + \sum_{i=0}^{q} A_{xy} y(t-i) + E_{x|y}(t) \\ \bar{y}(t+1) &= \sum_{i=0}^{p} A_{yx} x(t-i) + \sum_{i=0}^{q} A_{yy} y(t-i) + E_{y|x}(t) \end{aligned} \tag{13}$$

In this generalization of the generating model p past values up to $x(t)$ and q past values up to $y(t)$ are used to predict future values. Trivially, for the time-series shown above $p = q = 1$ is the best fit, but for arbitrary time-series a more general model is necessary.

Whether $y(t)$ and its past values are used in the prediction changes the residual or prediction error $E_x(t)$. If $E_x(t)$ is reduced by including $y(t)$, then Y Granger-causes X [48]. A simple way to introduce a measure for *causal influences* bound between 0 and 1 is using the variance of the residuals either incorporating $y(t)$ ($var(E_{x|y})$) or not incorporating $y(t)$ in the prediction, ($var(E_x)$):

A measure limited to $[0, 1]$ is not mandatory, but eases the comparison of different approaches.

$$GC_{y \to x} = \ln\left(\frac{var(E_x(t))}{var(E_{x|y}(t))}\right)$$

Here, no influence corresponds to 0 and maximum influence to 1.

In general, Wienger-Granger causality is not limited to linear vector autoregressive models. Different regression models can overcome limitations posed by the investigated time-series, e.g. in the case of non-linear [49] or non-stationary [50] problems. Due to its understandable and simple concept and its versatility Wiener-Granger causality is widely accepted, e.g. in neuroscience [51], biology [52], and economics [53].

In applications where *causal influences* in different frequency ranges are of interest, Wiener-Granger causality can also be transformed to Fourier space. The Fourier transformed vector autoregressive model is given by:

A Fourier transformation [54] is given by: $\hat{g}(f) = \int_{\infty}^{\infty} dx g(x) e^{-2\pi i x f}$.

$$\begin{pmatrix} A(f) & B(f) \\ C(f) & D(f) \end{pmatrix} \begin{pmatrix} x(f) \\ y(f) \end{pmatrix} = \begin{pmatrix} E_{x|y}(f) \\ E_{y|x}(f) \end{pmatrix}$$

For more details on how to infer a spectral Granger-causality from this transformed model refer to [55, 56].

3.4 MUTUAL INFORMATION AND TRANSFER ENTROPY

A more general variant of Wiener-Granger causality originates from information theory. The basic idea is to quantify the amount of shared and/or transferred information.

Naturally, the first step is quantifying the information within an observable, e.g. in the discrete variable X. The information entropy or so called Shannon entropy [57] is measuring exactly this:

Indeed, X can also be some time-series $x(t)$.

$$H(X) = \sum_x p(x) \log_2 \left(\frac{1}{p(x)}\right)$$

Notably, certain $(p(x) = 1)$ *and impossible* $(p(x) = 0)$ *outcomes do not contribute to the entropy.*

The probability of the occurrence x is $p(x)$, the second factor, $\log_2(\frac{1}{p(x)})$, is quantifying the amount of information of said event - rare events contributing more information. In principle, the base of the logarithm is arbitrary, but $\log_2()$ divides the entropy into units of *bits*. The toss of a fair coin contains one *bit* of information, the throw of a six sided die $\log_2(6) \approx 2.6$ *bits*. The amount of *bits* can also be thought of the smallest number of *yes/no-questions* to determine the outcome. For a coin the single question '*Heads?*' is sufficient, while for the case of the six-sided die the questions '*Larger 3?*' and '*Even?*' are not enough to determine the outcome.

Knowing that information is a measurable quantity, one can compare information of a variable X with another one Y or derive the amount of shared information. For the joint probability of the two variables $p_{XY}(x, y)$ and the marginals $p(x) = \sum_y p_{XY}(x, y)$ the mutual information [58] is then given by:

To avoid confusion with later notations the typical notation of $I(X;Y)$ *is replaced by* $MI(X;Y)$.

$$MI(X;Y) = \sum_{x,y} p_{XY}(x,y) \log \frac{p_{XY}(x,y)}{p_X(x) p_Y(y)}$$

The mutual information is the overlap or shared information of X and Y. In other words, if Y is known how much information is revealed about X or translated to *yes/no-questions*: how many less question have to be asked about X.

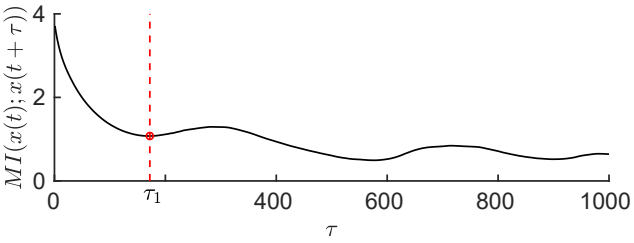


Figure 11: Mutual information $MI(x(t); x(t+\tau))$ between the time-series $x(t)$ and $x(t+\tau)$ of the Lorenz attractor (Figure 6) for τ between 0 and 1000. τ_1 (red line) is marking the first minima of $MI(x(t); x(t+\tau))$.

Importantly, mutual information is a symmetric quantity and thus suffers from the same drawbacks as correlations and is not suitable to infer *causal influences*. However, it is a useful measure to derive optimal parameters for state space reconstruction (Section 2.2.4). In Figure 11 the mutual information is shown for the times-series $x(t)$ of the Lorenz-attractor depicted in Figure 6. As a parameter for state space reconstruction the first (local) minimum at $\tau_1 = 172$ is a good choice, since $x(t)$ and $x(t+\tau)$ share little information - corresponding to an unfolded state space reconstruction.

A measure suitable to derive *causal influences* is the so called Transfer entropy [12] that shares the same basic thought as Wiener-Granger-causality and in the case of vector-auto-regressive models matches Wiener-Granger-causality. Since this short section only aims to illustrate the basic idea of methods from information theory, we will only illustrate the notion of transfer entropy using Figure 12. For a comprehensive and detailed view consider for example [58] or other textbooks on this topic.

Noteworthy, the interaction information can, in contrast to other quantities, also be negative.

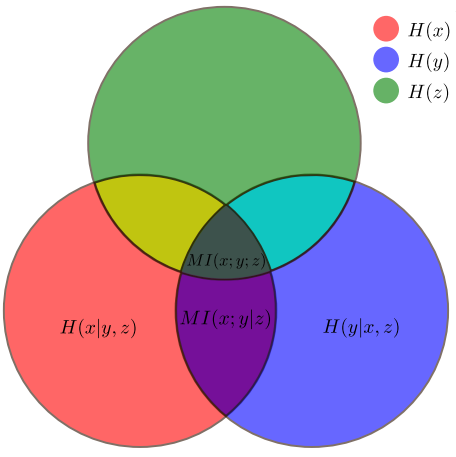


Figure 12: Venn diagram (of information) illustrating information theoretic quantities. Each circle (blue, red, green) is the entropy of the respective variable. A selection of relevant areas are labeled. Mutual information is the pairwise overlap of circles, the (dark green) overlap between all circles is the interaction information and the purple area the transfer entropy $MI(x; y|z)$.

In Figure 12 each circle represents the information of one variable, i.e. X, Y, Z. The intersection of two circles, i.e. the union of dark green and purple colored areas represents the amount of shared information (the mutual information, $MI(x;y)$). More interesting is just considering the purple area. The amount of shared information between Y and X that is not contained in Z.

This can be easily transferred to time-series and then gives an analog measure to Wiener-Granger causality. Suppose blue circle is the information within a time-series $[y(t), .., y(t-L)]$, and the information of another time-series is the union of the other two circles, future values $x(t)$ depicted by the red circle and past values $x(t-1), .., x(t-L)$ by the green. Naturally, future and past share some information, but there can be information in future $x(t)$ (red) that is contained in past values of $y(t), .., y(t-L)$ (blue), but not covered its own past (green). The resulting quantity is called transfer entropy or conditional mutual information and measures the amount of this information [12]:

$$T_{y\to x} = MI[(x(t); \{y(t-1), .., y(t-L)\} | \{x(t-1), .., x(t-L)\}]$$

This directly links to Wiener-Granger causality: More information about $x(t)$ in the past of $y(t), .., y(t-L)$ also means a better prediction of this future.

3.5 CAUSAL INFLUENCES FROM STATE SPACE RECONSTRUCTIONS

3.5.1 *Existing Methods - Convergent Cross Mapping*

The convergent cross map convergent cross mapping [4] (CCM) method proposed by Sugihara et al. [4, 13] shares the basic idea of WGC estimating *causal influences* from the prediction of future values. However, these predictions are not based on past values of the time series, but rather points in the vicinity on the underlying manifold are used for prediction.

Consider two observed time series $\{x(t)\}$ and $\{y(t)\}$ that causally influence the respective other. Due to Takens theorem [14] the underlying manifolds, $\mathbf{r}^x(t)$ and $\mathbf{r}^y(t)$, can be reconstructed. Furthermore due to the bilateral interaction, both reconstructions recover the whole overall state space (X, Y) and are homeomorphic copies of each other.

In particular, the nearest neighbors around a reference point remain close-by. Thus, CCM projects the nearest neighbors t_x around $\mathbf{r}^x(t_x)$ to the other space $\mathbf{r}^y(t_x)$. The new neighbors are weighted based on their distance to the reference point t in X and used to predict $y(t+1)$. The prediction error is then quantified via the Pearson cross correlation coefficient between predicted next time-step and actual next step $\rho(y(t+1), \hat{y}(t+1)|M_X)$. Varying the amount of available data reveals

the underlying *causal influences*, since the predictive power scales differently based on the interaction strength.

In Figure 13 one of the examples from Sugihara et al. [4] is reproduced. Two time-series from an experimental measurement of a predator-prey systems are shown in **(a)**. The time-series were analyzed using the same approach as [4], the resulting cross map skill from predator to prey (blue) and vice versa (red) is depicted in **(b)**. The higher cross mapping skill is inferred for the top-down direction originating from the predator. In combination with the reverse direction convergent cross mapping infers a bidirectional interaction with a stronger *causal influence* from prey to predator.

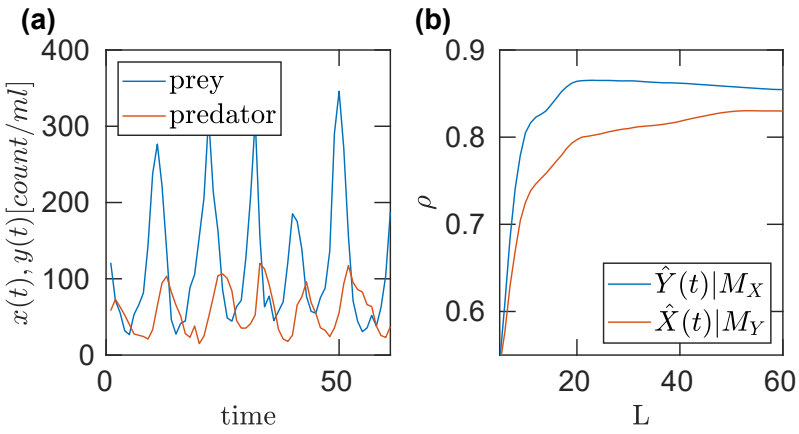


Figure 13: **(a)** Time series of a predator-prey system. The blue curve shows the prey population and red the predator. The population size is measured as count/mL. The analysis was done using the same parameters as in [4] - the time-series were embedded using $\tau = 1$ and $m = 3$, **(b)** Cross map skills measured by Pearson cross correlation coefficient ρ. The blue curves show ρ for the mapping prey to predator and red vice versa.

While Sugihara et al. showed in their original [4] and following work [59] the power of this new method, it could be shown in other works that CCM is flawed. Yielding wrong directions of dominant influences for synchronizing system [60–62], e.g. in the case of flue-humidity interactions.

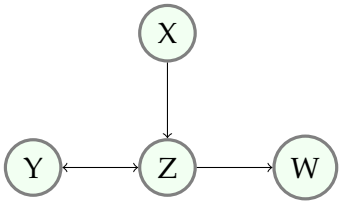
At heart, both our later proposed methods (*topological causality* and *cross projection method*), are motivated by CCM, but render the prediction of future states unnecessary. Instead the transferred information between system is inferred from metric properties of the state space reconstructions. In fact, already the metric and topological properties of these manifolds allow to infer causal relations [1, 63–65].

For more details on CCM, refer to [4] and its supplementary materials.

3.5.2 *Dimension of (sub)-systems*

The basis for CCM, and other methods alike, is provided by the fundamental properties of (coupled) dynamical systems resulting from Takens theorem [14]. In principle, these fundamental properties are sufficient to infer *causal influences* by comparing the dimensions of (sub-)systems.

Consider a system comprised of subsystems X, Y, Z, W as shown in Table 1 and their respective observed time-series. The reconstruction from these time-series differ, especially their dimension.



observable	reconstructed system
X	$\{X\}$
Y	$\{X, Y, Z\}$
Z	$\{X, Y, Z\}$
W	$\{X, Y, W, Z\}$

Table 1: **(left)** Interaction structure of a system comprised of 4 subsystems (X, Y, Z, W), arrows indicating the direction of interaction.
(right) Observable and the respective reconstructable (sub)-system.

Since X has only outgoing interactions no information about the other subsystem is available in the observable $x(t)$ resulting in a reconstructed system of only the subsystem $\{X\}$. Other subsystems further downstream (in the flow of information) also have information about the former subsystems i.e. an observable $z(t)$ of subsystem Z is able to reconstruct the joint space of $\{X, Y, Z\}$. Indeed, the joint space $\{X, Y, Z\}$ is also a reconstruction from the observable $y(t)$ due to the backwards connection from Z to Y. Only an observable from subsystem W yields the a reconstruction of the full overall system.

As illustrated in Table 1 the reconstruction using different observables contains information about the interactions between (sub-)systems. In fact, the dimension of the reconstructed space indicates the share of information about a subset of the overall system included in a single observable and can be used as binary criterion for interactions between subsystems [66]. In the example shown above the space reconstructed from W has the highest dimension and represents the whole space of the system $\{X, Y, W, Z\}$. A higher dimension corresponding to an observable that reconstructed a greater fraction of the whole system and thus being later in the interaction-chain, e.g. having more incoming connections from different subsystems.

Often, *causal influences* of systems with bidirectional interactions are especially interesting, for these it is not sufficient to only consider the

dimension. Here, not only the dimension of the reconstructed spaces, but additional properties of reconstructions have to be used.

3.6 EXPANSIVE DISTORTIONS OF NEIGHBORHOODS

Expansive properties as a proxy for *casual influences* were touched upon in the introduction and illustrated in Figure 1. The images shown are based on a set of (coupled) logistic maps:

$$\begin{aligned} x(t+1) &= x(t)[R_x(1-x(t)) - w_{y\to x}y(t)] \\ y(t+1) &= y(t)[R_y(1-y(t)) - w_{x\to y}x(t)] \end{aligned} \tag{14}$$

The introduction briefly elucidates expansive properties for the case of uncoupled and unilaterally coupled logistic maps. Here, we also show the case of bilaterally coupled maps in Figure 14.

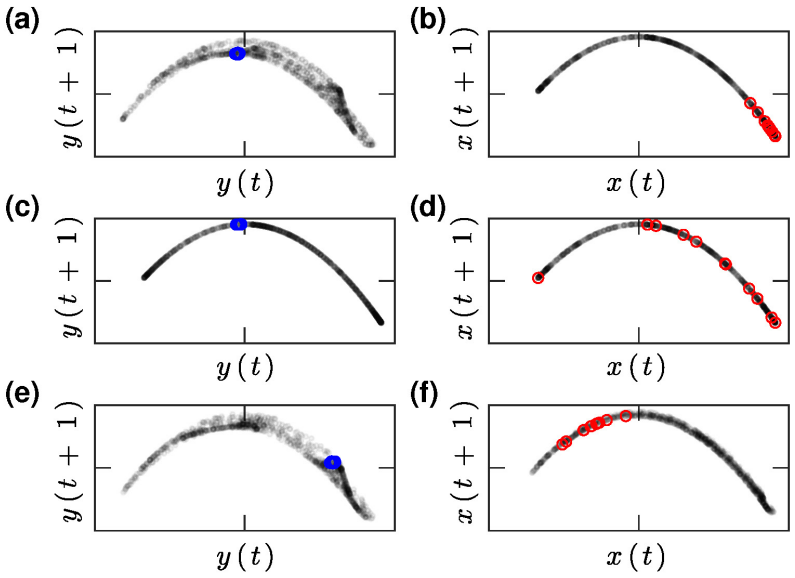


Figure 14: Reconstruction $\mathbf{r}^y(t)$ **(a)** and $\mathbf{r}^x(t)$ **(b)** of the state space of two noise free unilaterally coupled logistic maps with $w_{y\to x} = 0, w_{x\to y} = 0.3, R_x = R_y = 3.82$. 10^3 data points are shown in grey and the ten nearest neighbours in $\mathbf{r}^y(t)$ around a reference point are shown in blue in **(a)** and their images in X (red circles) in **(b)**. **(c)** & **(d)** show the equivalent results for uncoupled coupled logistic maps ($w_{y\to x} = 0, w_{x\to y} = 0$), while **(e)** & **(f)** show reconstructions for bilaterally coupled logistic maps ($w_{y\to x} = 0.1, w_{x\to y} = 0.2$). The embedding dimension is $m = 4$ and the time-delay $\tau = 1$ for all cases.

The picture above shows the state spaces $\mathbf{r}^y(t)$ and $\mathbf{r}^x(t)$ reconstructed from time series of the aforementioned logistic maps. The rows from top to bottom deal with unilateral coupled, **(a)** & **(b)**, uncoupled, **(c)** & **(d)**, and bidirectionally, **(e)** & **(f)**, coupled logistic maps.

Already, the shown reconstructions are insightful and clearly show the increasing dimension postulated in the previous section. With an existing connection from X to Y (middle row), the one dimensional line in the reconstructions space widens to a two dimensional object suggesting *causal influences* from X to Y.

The basis for the link between expansive properties and *causal influences* is the following simple consideration: Let a system X unidirectional influence another system Y. Obviously, Y receives information about X. Furthermore, states of Y will contain information about states of a X, while states of X by assumption cannot have 'knowledge' about Y. For certain classes of dynamical systems Takens' Theorem [14–17] allows the reconstruction of manifolds of the dynamic (as shown in Figure 14). Due to incoming information from X subsystem Y can not only reconstruct its own dynamic, but more importantly the overall dynamic of (X, Y). A key insight is that expansive local distortions reflect these *causal influence* between system components X and Y.

Consider a reference point t for the unilaterally coupled maps. The closest neighbors t_k of the reference point are marked in blue in **(d)**. Their respective time index t_k links them with a set of putative neighbors in the reconstruction in X **(c)**. If Y would have unlimited information about X, neighbors around a reference point in $\mathbf{r}^x(t)$ and $\mathbf{r}^y(t)$ were identical. In general, this is not the case, also the putative neighbors shown in **(c)** are more spread than a typical set of neighbors. A stronger coupling weight $w_{x\to y}$, results in more transferred information and more localized neighbors. The expansion of neighborhoods is thus inversely related to coupling weight and *causal influence*.

For vanishing weight (upper row) neighbors are spread of the whole reachable state space and are on par with random neighbors. If information is also transmitted from Y to X - the systems are bilateral interacting - X is also able to reconstruct the overall state space (bottom row). Both state reconstruction are homeomorphic to the overall dynamic - topologically they are equivalent and local neighborhoods are only distorted in relation to the respective other space.

The following chapters deal with our two methods, *topological causality* and *cross projection method*, quantifying these expansive properties to infer *causal influences*. *Topological causality* estimates the expansive properties directly from the local mappings between state space reconstructions. This allows to derive *causal influences* in a straightforward, analytically tractable manner.

Expansive properties can be more robustly estimated by looking not at expansion *between* different state space reconstructions but instead *within* each reconstruction, respectively. This is done by comparing the sizes of local neighbourhoods in one state space reconstruction

with both, the projections of neighbours from the other state space reconstruction and random neighbours. The resulting expansive distortions of neighborhoods in the manifold not only capture possible dimensional conflicts among state reconstructions [63], but also allows for a statistical criterion to control for false positive detection of *causal influences*. A fundamental advancement of CPM from TC is the notion that the neighborhood sizes are only compared in one space rendering this method invariant to differing scales of the state space reconstructions and less sensitive to fluctuations of densities.

Chicharro et al. [67] improved these distance based methods by looking at the rank of neighbors instead of their distance. However, our methods have the advantage of being time-resolved, resulting in state-dependent measure for *causal influences* that can also infer changes in the dominant direction of influence.

Or areas of strong interaction in state space.

Using inter neighborhood relations to estimate causality is not a novel concept of CPM (introduced in Chapter 5). There is a selection of other methods also using the expansive properties either directly or indirectly. Čenys et al. [68] termed the average distance in the corresponding space the *mean conditional dispersion* and noticed that for small ϵ there will be a dependence on the coupling strength. Various other distance based approaches, e.g. [69], were proposed that use a fixed number of neighbors instead of a fixed neighborhood size. These methods compare the distance to the nearest neighbors with the conditioned distance and/or the mean distance to all other points.

Our methods: topological causality [1] and cross projection method [2].

3.7 CONCLUSION

The notion of **causality** is strongly associated with intervention. When dealing only with observed time-series it is more suitable to consider *causal influences*. The inference of *causal influence* is used as a tool to elucidate hidden or not apparent interactions between systems or system parts. It is supposed to facilitate a better description of a system.

The simplest approach to investigate the interaction structure are correlations. However, undirected measures, like correlations, lack the ability to infer a direction of influence. State of the art methods (for inference of *causal influences*) are able to do this by assessing the prediction quality of future states. In contrast, to these methods, our proposed approach solely relies on expansive distortions of state space reconstruction and provides a time resolved measure for *causal influences*.

'This' = deriving an directed measure for causal influences

4

TOPOLOGICAL CAUSALITY

The true delight is in the finding out rather than in the knowing.

(Isaac Asimov)

The closing part of previous chapter already introduced the key insight and basis of topological causality [1] (TC). Local distortions in the mappings between state-space reconstructions of different observables reflect the time dependent efficacy of causal links among the underlying system components.

The causality measure, *topological causality*, derived from this insight is analytically accessible for simple systems and for more complicated ones can be estimated in a model free, data driven manner. For deterministic systems the expansive distortions of mappings among time-delay state space reconstructions from different observables not only reflect the directed coupling strengths, but also the dependency of effective influences on the system's temporally varying state.

This chapter addresses said concept of *topological causality*, introduced by Harnack, Laminski, Schünemann and Pawelzik in [1]. The first section elucidates how to to uncover *causal influences* from the expansive properties of state space reconstruction and especially how to measure the expansions. The second section covers some basic results as a proof of concept. Both of these sections are also covered in the publication [1] and its supplementary material in a similar manner. The last section showcases two examples, where *topological causality* fails to infer the correct underlying direction of *causal influences*.

4.0.1 *Contributions to Topological Causality*

The original publication was designed by Daniel Harnack and Klaus R. Pawelzik. Daniel Harnack wrote the manuscript for publication, did most of the calculations, analysed and performed simulations and prepared the figures. I contributed to development and testing of numerical procedures, provided a draft text for the section on numerical methods in the paper and designed Figure 17. Maik Schünemann provided the mathematics for invariance under transformations and the connection to information theory as well as draft for the corresponding section in the paper.

In particular, Figure 15 and Figure 16 and their encompassing text have strong correlations with [1]. The notation was overhauled to better match with the rest of this work. The section dealing with analytical results is partly based on supplementary materials of [1] and

on my master thesis [70], part of these results are also featured in the supplement of [1]. While the last section is discussed in the Bachelor thesis of Ronja Gronemeyer [71], to which I contributed in the supervision.

4.1 FROM EXPANSIVE DISTORTIONS TO TOPOLOGICAL CAUSALITY

Following the argumentation in Harnack et al. [1], consider a dynamical system Z composed of two subsystems, X and Y, governed by

$$\dot{\mathbf{x}}(t) = f(\mathbf{x}, w_{y\to x}\mu_x(\mathbf{y})) \qquad \dot{\mathbf{y}}(t) = g(\mathbf{y}, w_{x\to y}\mu_y(\mathbf{x})),$$

where $\mu_i(\mathbf{i})$ denote fixed scalar functions and $w_{i\to j}$ coupling constants. $\mathbf{x}(t)$ and $\mathbf{y}(t)$ are the multidimensional states of the individual systems components.

$i, j \in \{x, y\}$

The trajectories $(\mathbf{x}(t), \mathbf{y}(t))$ form an invariant manifold in the state space of the joint dynamical system. A manifold in a delay coordinate space is visited by $\mathbf{r}^i(t) = [i(t), i(t+\tau), (i(t+2\tau), \ldots, i(t+(m-1)\tau))]$ - the reconstruction (Equation 9) via Takens theorem.

Only a single variable of the multidimensional states $\mathbf{x}$ *and* $\mathbf{y}$ *is used for reconstruction.*

The two manifolds, $\mathbf{r}^x$ and $\mathbf{r}^y$, are topologically equivalent if homeomorphic mappings between the manifolds exist. If both $w_{x\to y} \neq 0$ and $w_{y\to x} \neq 0$, also homeomorphic mappings between the reconstructions exist, i.e. $M_{x\to y}$, from $\mathbf{r}^x$ to $\mathbf{r}^y$, and $M_{y\to x}$, from $\mathbf{r}^y$ to $\mathbf{r}^x$. To illustrate how these mappings between reconstructions relate to directed effective influence, consider the following thought experiment:

For a more explicit example consider the neighbor relations in Figure 14.

First, a system with unidirectional interaction is observed, i.e. $w_{x\to y} \neq 0$; $w_{y\to x} = 0$. By virtue of Takens' theorem, a unique mapping $M_{y\to x}$ from reconstruction $\mathbf{r}^y$ to $\mathbf{r}^x$ exists. However, the reverse direction $M_{x\to y}$ does not exist, since X has no information about Y. This is illustrated in Figure 15 **(a)** by a joint manifold $(\mathbf{r}^x, \mathbf{r}^y)$ lying folded over $\mathbf{r}^x$ but uniquely over $\mathbf{r}^y$: $\mathbf{r}^y(t)$ is not uniquely determined for all states $\mathbf{r}^x(t)$. Locally, $M_{x\to y}$ can be attributed a diverging expansion property: since close neighbors of a given point $\mathbf{r}^x(t)$ correspond to distant parts of the joint density $(\mathbf{r}^x, \mathbf{r}^y)$, i.e. local expansions extend to macroscopic scales. In contrast, close neighbors around a reference point $\mathbf{r}^y(t)$ remain in the vicinity of the corresponding point $\mathbf{r}^x(t)$, the mapping in this direction preserves the neighborhood relations. Note here that somewhat counter-intuitively the influence from x to y is reflected in the 'backward' mapping $M_{y\to x}$: the existence of $M_{y\to x}$ implies an interaction from X to Y.

$M^t_{i\to j}$ *is the local linearization of the mapping* $M_{i\to j}$.

Now increasing $w_{y\to x}$ while keeping $w_{x\to y} > w_{y\to x}$ leads to mutual but asymmetric interactions. Reconstructions $\mathbf{r}^x$ and $\mathbf{r}^y$ will now, both,

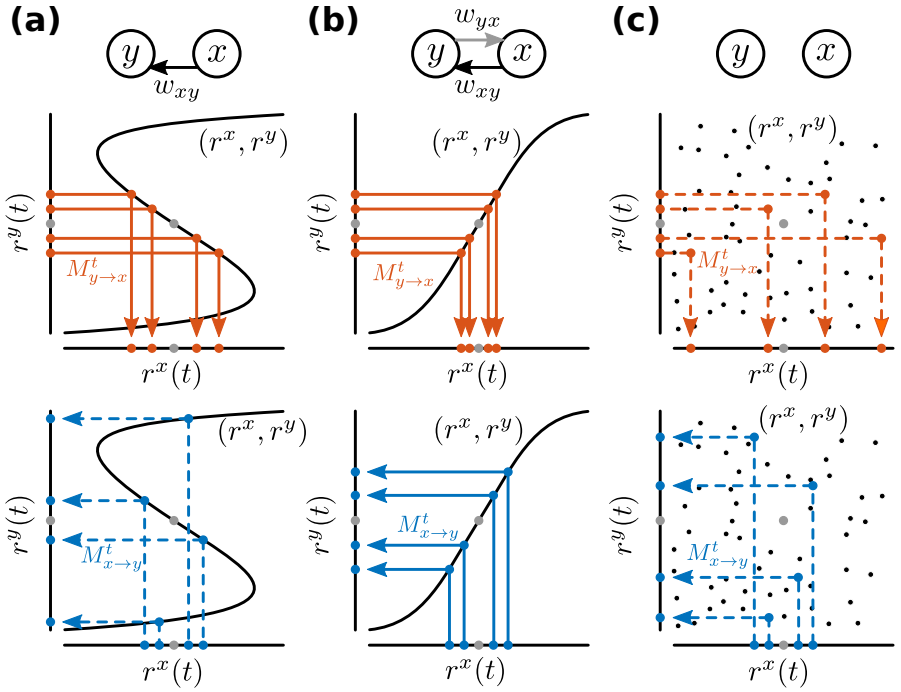


Figure 15: The relation of points $\mathbf{r}^x$ and $\mathbf{r}^y$ on multidimensional manifolds illustrated in 1-d. The joint manifold represented by $(\mathbf{r}^x, \mathbf{r}^y)$ can be interpreted as the function mediating the mappings $M_{i\to j}$ between both spaces, and local linearizations $M^t_{i\to j}$ of the mappings as the slope around a reference point. **(a)** When only $w_{x\to y} \neq 0$, a one-to-one mapping $M_{y\to x}$ from $\mathbf{r}^y$ to $\mathbf{r}^x$ exists, but not in the reverse direction: $\mathbf{r}^y(t)$ is not uniquely determined for all states $\mathbf{r}^x(t)$. Locally, $M^t_{x\to y}$ can be attributed a diverging expansion property: since close neighbors of a given point $\mathbf{r}^x(t)$ correspond to distant parts of the joint density $(\mathbf{r}^x, \mathbf{r}^y)$, i.e. local expansion extend to macroscopic scales. The dashed lines visualize the non-uniqueness. **(b)** Here, both couplings are non-zero, but $w_{x\to y} > w_{y\to x}$. Larger independence of x implies a stronger expansion by $M_{x\to y}$ than by $M_{y\to x}$ at most reference points, which is indicated by the higher slope of $(\mathbf{r}^x, \mathbf{r}^y)$ when seen from $\mathbf{r}^x$. **(c)** If no coupling exists, the expansion diverges in both directions.

reveal the same global system state. However, the weaker coupling from Y to X implies that a neighborhood of states around $\mathbf{r}^y(t)$ is larger than the corresponding neighborhood around $\mathbf{r}^x(t)$ for most reference points t: Both $\mathbf{r}^x$ and $\mathbf{r}^y$ are driven away from their state at time t by a combination of internal dynamics and the external influence from the other variable, but Y is more so due to the stronger coupling $w_{x\to y}$. This entails that $M^t_{x\to y}$ will be more expanding than the mapping $M^t_{y\to x}$, and the joint manifold $(\mathbf{r}^x, \mathbf{r}^y)$ lying uniquely over both reconstruction spaces, but more 'steeply' over $\mathbf{r}^x$ (Figure 15 **(b)**).

If $w_{y\to x}$ is now decreased again to approach 0, the expansion $e^t_{x\to y}$ will increase until it diverges at the point where $(\mathbf{r}^x, \mathbf{r}^y)$ folds in on it-

self as seen from $\mathbf{r}^x$ (Figure 15 **(a)**). This happens at $w_{y\to x} = 0$, where the map $M_{x\to y}$ looses uniqueness and corresponding points to neighbors in $\mathbf{r}^x$ lie scattered over the whole dynamical range of $\mathbf{r}^y$. Thus we equate infinite expansion to the non-existence of the corresponding mapping.

Consequently, when the couplings among X and Y vanish altogether, both component systems will behave independently and the density of the resulting joint manifold factorizes. When observed from reference states $\mathbf{r}^x(t)$ and $\mathbf{r}^y(t)$, the mappings can be considered infinitely expanding, since for most reference points close neighbors correspond to distant points in the respective other space (Figure 15 **(c)**).

In hindsight, metric causality would be a more suitable name.

Following these topological considerations, local expansions of the mappings between reconstruction manifolds of two observables can be utilized for a graded measures of the directed *causal influences* between component systems represented by these observables.

Assuming that the mappings between reconstruction are differentiable, $M^t_{i\to j}$ denotes the local linearization (Jacobian matrix) of $M_{i\to j}$ at the reference point t: Given that $\{t^i_1, ..., t^i_k\}$ are the time indices of the nearest neighbors in $\mathbf{r}^i$ to the reference point $\mathbf{r}^i(t)$, $M^t_{i\to j}$ is the linear approximation of the mapping that projects $\{\mathbf{r}^i\left(t^i_1\right), ..., \mathbf{r}^i\left(t^i_k\right)\}$ to $\{\mathbf{r}^j\left(t^i_1\right), ..., \mathbf{r}^j\left(t^i_k\right)\}$. In practice, we analyze the expansion $e^t_{i\to j}$ of $M^t_{i\to j}$, which is determined by the singular values $\sigma^t_k(M^t_{i\to j})$ of $M^t_{i\to j}$ larger than 1:

Note that neighbors are searched in i are projected to $\mathbf{r}^j$.

$$e^t_{i\to j} = \prod_k \max\left(1, \sigma^t_k(M^t_{i\to j})\right) . \tag{15}$$

The expansion $e^t_{i\to j}$ is inversely related to the strength of *causal influence* $j \to i$. Moreover, this expansion is not only depending on the coupling weights, but also on the state of the system, so that different regions in state space could be characterized by different directions of causal dominance. To measure such state dependent asymmetry of *causal influence*, we define an index $-1 \leqslant \alpha^t \leqslant 1$:

$$\alpha^t = \frac{\log(e^t_{x\to y}) - \log(e^t_{y\to x})}{\log(e^t_{x\to y}) + \log(e^t_{y\to x})} \tag{16}$$

This definition is motivated by the relation of the log expansions to loss of certainty in information theoretical terms (Supplementary Materials of [1]). Additionally, it is suitable to define a mean causal asymmetry $-1 \leqslant \alpha \leqslant 1$ index by averaging over the states visited during the dynamics:

$$\alpha = \frac{\langle \log(e^t_{x\to y}) - \log(e^t_{y\to x})\rangle_t}{\langle \log(e^t_{x\to y}) + \log(e^t_{y\to x})\rangle_t} \tag{17}$$

However, α and α^t do not differentiate between the qualitatively different situations of balanced strong and balanced weak influence. For this purpose and the case of unilateral interactions it is more feasible to quantify the strength of *causal influence* (TC) instead of an asymmetry index:

$$C_{x\to y} = \frac{1}{1+\langle \log\left(e^t_{y\to x}\right)\rangle_t}$$
$$C^t_{x\to y} = \frac{1}{1+\log\left(e^t_{y\to x}\right)} \tag{18}$$

$C \in [0,1]$ satisfies the following two fundamental intuitions about *causal influences*: TC from component system X to Y vanishes if no causal link exists ($w_{x\to y} = 0$), and for small couplings it is a monotonous function of the coupling weight $w_{x\to y}$. Note that also here the important distinction between TC and coupling weight holds: C depends on the coupling weights as well as on the current state of the system and the internal dynamics of each component.

For certain coupling schemes of logistic maps even a linear function of $w_{i\to j}$

4.2 EXEMPLARY RESULTS FOR TOPOLOGICAL CAUSALITY

Again, consider time series generated by coupled logistic maps (Equation 2) as an example. The state space of the system is shown in Figure 16 **(a)**, for each point its color signifies the respective time-resolved asymmetry index α^t.

The mixture of colors results from the projection to the $x(t) - y(t)$-plan, later Figure 20 shows an unfolded space (for unilateral interaction).

The asymmetry index α^t changes for different regions of the state space. However, for close states these changes are continuous - similar states share a similar causal asymmetry.

These fluctuation are also reflected in the representation of α^t over time (Figure 16 **(b)**). The changes of the dominant direction of *causal influences* corresponds to various dynamical regimes among the time courses of $\mathbf{r}^x(t)$ and $\mathbf{r}^y(t)$ and are rather volatile. This is due to the nature of logistic maps, where a future state is in a completely different region of state space as its predecessor. This also obvious to see in $x(t)$ and $y(t)$, shown in Figure 16 **(b)** & **(c)**. When the influence from X to Y is stronger than the reverse direction (blue region), i.e. $e^t_{x\to y} > e^t_{y\to x}$, the trajectory of $y(t)$ shows stronger fluctuations than the one of $x(t)$.

Lastly, the average causal asymmetry α is shown for different combinations of coupling weights in Figure 16 **(d)**. The fact that $\alpha \neq 0$ for $w_{x\to y} = w_{y\to x}$ reflects the difference between the dynamical equations for $x(t)$ and $y(t)$, and highlights again the expansion is not a mere proxy for the coupling weight, but actually measures the effective influence exerted along the causal link.

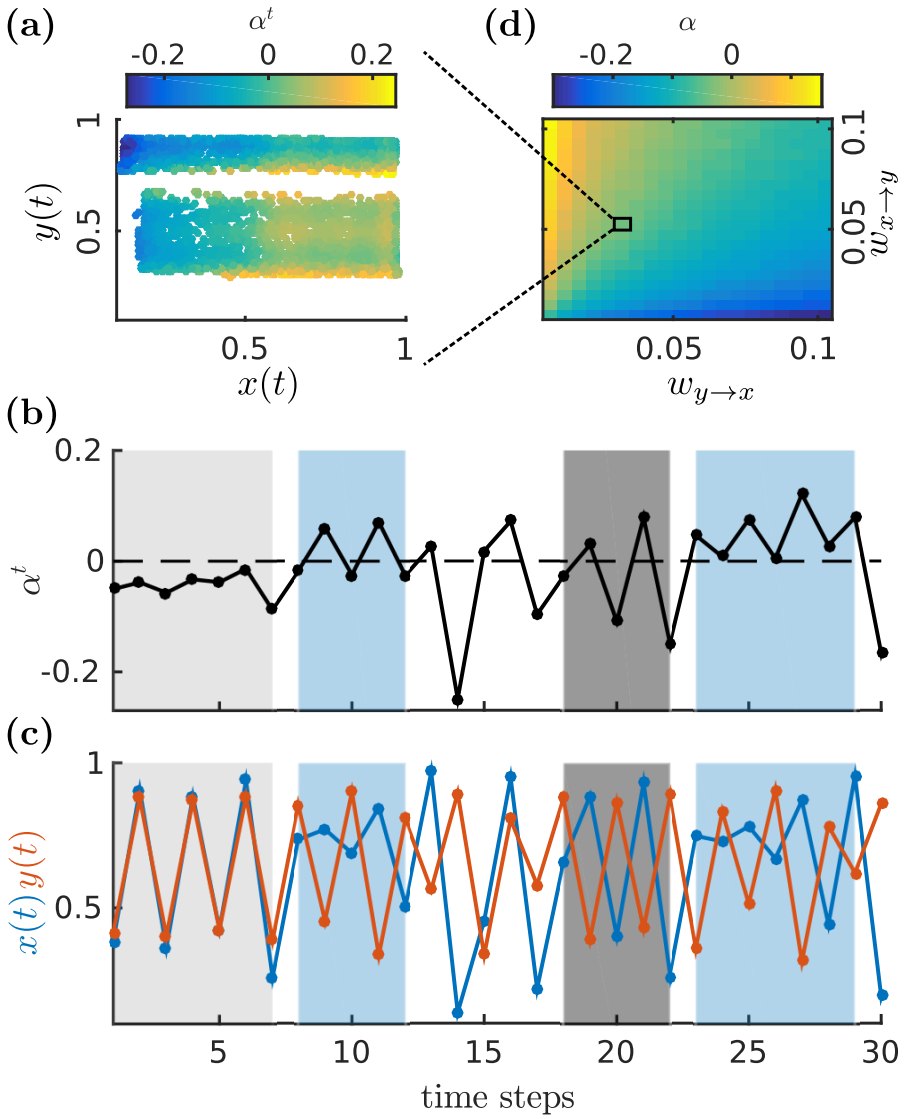


Figure 16: **(a)** The state-space dependent α_t in a system given by Equation 14 is shown for $w_{x\to y} = 0.05$, $w_{y\to x} = 0.02$. **(b)** A segment of α^t and the corresponding time courses **(c)** of $x(t)$ and $y(t)$ for the same coupling weights. Different regimes of dominant causal direction give rise to different dynamical motifs. If α^t is close to 0 for subsequent time points (light gray), x and y synchronize. If α^t varies strongly around 0 (dark gray), X and Y desynchronize. When the *causal influence* from one variable to the other is dominant, here from X to Y (blue), the trajectory of $y(t)$ shows higher amplitude excursions than the one of $x(t)$. **(d)** The mean asymmetry index α for the same system with varying coupling strengths.

As an example of a more complex case that is not analytically tractable, consider a system of coupled Rössler equations [30] described by:

$$\begin{aligned}
\dot{x}_i(t) &= -f_i y_i(t) - z_i(t) + w_{j\to i} z_j(t) \\
\dot{y}_i(t) &= f_i x_i(t) + 0.1 y_i(t) \\
\dot{z}_i(t) &= 0.1 + z_i(t)[x_i(t) - 14], \quad i = 1, 2
\end{aligned} \tag{19}$$

The frequency parameters $f_1 \approx 0.99$ and $f_2 \approx 0.85$ are incommensurable and coupling weights are set to $w_{1\to2} = 0.2$, $w_{2\to1} = 0.6$. The observables $y_i(t)$ are shown in Figure 17 **(d)**. Both variables are not directly part of the interaction between the individual oscillators. Thus, their respective time-series do not show any effects of the interaction.

The chosen interaction by the additive term $w_{j\to i} z_j(t)$ guarantees a strong state dependence of the interaction. Strong *causal influences* are only expected if the driving z_i component deviates from the x_i-y_i-plane. Exactly, in these peaks the derived measures for *causal influence* show the strongest asymmetry, **(a)** and **(b)**. The state, and more importantly time, dependence of the interaction is even more evident, if the time resolved *causal influence* $C^t_{2\to1}$ is shown with the variable $z_2(t)$ that drives the interaction in this direction **(c)**.

Causal influences can also be derived for other coupling schemes, but the clear state dependence is most prevalent using $z_i(t)$

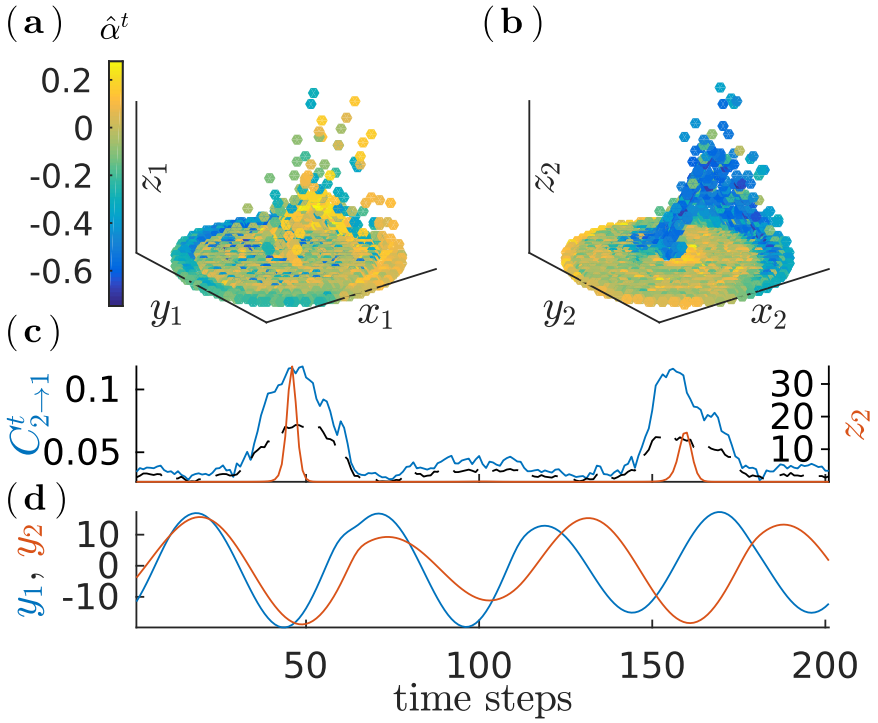


Figure 17: Two bidirectionally coupled Rössler systems Equation 19 with $w_{1\to2} = 0.2$, $w_{2\to1} = 0.6$. The time series of 10^5 data points were generated with a step size of $dt = pi/25$ and embedded with dimension $m = 13$ and $\tau = 12$. α^t and $C^t_{2\to1}$ are shifted to be aligned with the temporal mean $t + 1/2(m-1)\tau$ of the corresponding reconstructions $\mathbf{r}^i(t)$. **(a)** & **(b)** Local asymmetry α^t of 10^4 points shown on projections of the attractor to each system. **(c)** $C^t_{2\to1}$ (blue) for 200 consecutive time steps and the corresponding time series of z_2 (red). The dashed line marks chance level. **(d)** Time series y_1 and y_2 used to estimate $C^t_{2\to1}$.

4.2.1 *Analytical Results*

For some (simple) system, e.g. coupled logistic maps Equation 2, TC remains analytically tractable. Consider logistic maps, for simplicity the internal dynamics are substituted with $g_x(t)$ & $g_y(t)$ respectively:

$$x(t+1) = x(t)[R_x(1-x(t)) - w_{y\to x}y(t)] = g_x(t) - w_{y\to x}x(t)y(t)$$
$$y(t+1) = y(t)[R_y(1-y(t)) - w_{x\to y}y(t)] = g_y(t) - w_{x\to y}y(t)x(t)$$

For small perturbations, i.e. small neighborhoods, the mapping from $\mathbf{r}^x$ to $\mathbf{r}^y$ can be linearized around a reference point $\mathbf{r}^i(t) = (i(t), i(t+1))$, such that $(j(t), j(t+1))^T = M^t_{i\to j}(i(t), i(t+1))^T$ is

$$M^t_{i\to j} = \frac{1}{w_{ji}j(t)} \begin{pmatrix} g_i'(t)w_{ji}j(t) & -1 \\ (g_i'(t)w_{ji}i(t))(g_j'(t)w_{ij}j(t)) & -(g_j'(t)w_{ij}i(t)) \end{pmatrix}$$

with $g_i'(t) = R_i(1-2i(t))$ being the derivative of the internal dynamics of $i(t)$. In this example one singular value of $M^t_{i\to j}$ is greater 1 and attributed to the expansive direction of the (linearized) mapping. Thus, the expansion is

The arrows in the subscript are removed for compactness ($w_{ij} = w_{i\to j}$).

$$e^t_{j\to i} = \frac{1}{|w_{i\to j}|}\chi(w_{x\to y}, w_{y\to x}, x(t), y(t)).$$

The leading factor $1/w_{i\to j}$ is dominating the expansion for small weights and $\chi(w_{x\to y}, w_{y\to x}, x(t), y(t))$ is mostly constant. Only for large weights the expansion becomes a function of all weights and the system state.

Thus, comparing analytical and numerical results it is sensible to introduce an asymmetry index purely derived from (small) weights:

$$\alpha_w = \frac{w_{y\to x} - w_{x\to y}}{w_{y\to x} + w_{x\to y}} \tag{20}$$

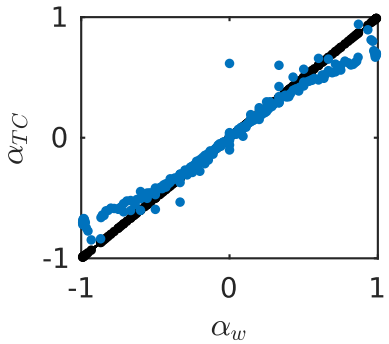


Figure 18: Causal asymmetry α_{TC} for two coupled logistic maps over the asymmetry of weights α_w. $N = 10^3$ data points were embedded with an embedding dimension $m = 4$ and time-delay $\tau = 1$ for different coupling weights.

For small coupling this is in good agreement with the analytic results, in Figure 18 the causal asymmetry α is shown over α_w for coupling weights forming an uniformly distributed α_w. As can be seen the

More Details are provided in the supplement of [1].

asymmetries are in good agreement apart from the the limiting cases $\alpha_w = \pm 1$. Here, one coupling weight is big, thus the state of the system also contributes strongly to the expansion and to α_{TC}.

Numerically, TC appears straightforward to compute by fitting the Jacobian of the mapping between local neighborhoods and then calculating the singular values. However, in practice we found two consecutive principle component analysis are needed to be numerically stable, for more details refer to the supplemental material of [1]. For synchronizing systems and/or system disturbed by strong noise this proved not reliable enough requiring an overhauled or improved variant of TC.

4.3 SYNCHRONIZATION IN COUPLED LIMIT-CYCLES

So far, all examples covered cases in which TC obtained reliable and somewhat predictable results, but in two cases TC fails. The first problem is noise leading to wrong neighbors and thus affecting the computation of mappings between reconstructions. We managed to (somewhat) overcome this in [1] by refining the numerical methods, in particular by including the aforementioned principle component analysis. Furthermore, there is a wide range of methods to reduce noise, rendering this problem, in general, manageable. The second problem is the inference of the wrong dominant direction of influence (in certain systems) and this is shown in the following example. In the end, this lead us to refine TC with the approach covered in Chapter 5.

In hindsight, synchronization is a problem for all methods inferring causal influences.

For a system prone to synchronization, consider a set of two coupled Fitzhugh-Nagumo neurons. This simplified neuronal model consists of a voltage-like and a recovery variable described by two coupled differential equations allowing to model the biological phenomena underlying spike-generation:

$$\begin{aligned} \dot{x_1} &= x_1 - \frac{1}{3}x_1^3 - x_2 + I + \Omega_{i\to j}, \qquad \Omega_{i\to j} = w_{i\to j}\frac{1}{1+e^{-3x_i}} \\ \dot{x_2} &= f(x_1 - 0.8x_2 + 0.7) \end{aligned} \tag{21}$$

On their own each neuron forms a limit cycle in the state space if the input current I is between $0.4 \leqslant I \leqslant 1.42$. Two neurons are linked via the coupling functions $\Omega_{i\to j}$. Already, for an unilateral coupling $w_{i\to j} = 0.75$ & $w_{j\to i} = 0$ they are synchronizing and form a shared limit cycle. In Figure 19 **(a)** the projection of this limit cycle on the x-y-plane is shown with the color indicating the causal asymmetry α^t.

Indeed, non-identical Fitzhugh-Nagumo neurons also synchronize.

The *causal influence* changes depending on the state of driving and driven subsystem. However, for a unilateral coupling scheme this

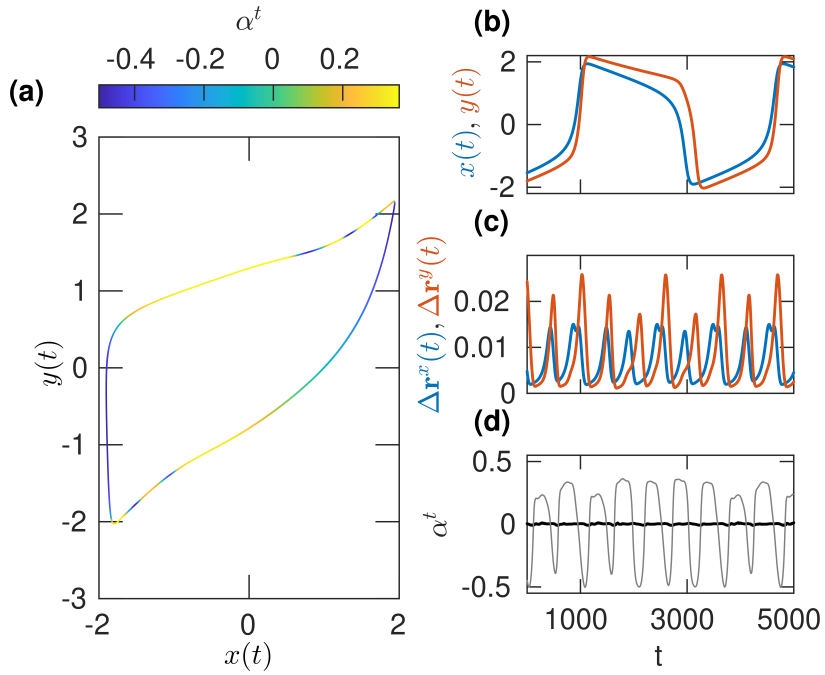


Figure 19: Two unilaterally coupled ($w_{i\to j} = 0.75$ & $w_{j\to i} = 0$) Fitzhugh-Nagumo-Neurons: The differential equations are solved using a step size of $dt = 0.1$ and $N = 10^5$ data points of the time series were embedded using an embedding dimension of $m = 5$ and time-delay of $\tau = 1000dt$ using $k = 100$ neighbors. **(a)** shows the projection of the state space as x-y-plane. The color coding reflects the causal asymmetry α_t. **(b)** A section of the time-series of the voltage-variable $x(t)$ (blue) and $y(t)$ (red), the single oscillation is equivalent to one cycle in the state space. **(c)** The (euclidean) distance of two consecutive time steps (in the reconstruction spaces $\mathbf{r}^x(t)$ and $\mathbf{r}^y(t)$) as a proxy for instantaneous velocity. **(d)** The causal asymmetry α_t (grey) and a renormalized asymmetry $\hat{\alpha}_t$ (black)

must not change the sign of α and thus predicting the wrong direction of *causal influence*.

Taking a closer look reveals changes in velocity along the the limit cycle. Since dt is fixed, a good proxy for the velocity is the distance between two (temporal) consecutive states, i.e. $\mathbf{r}^x(t)$ and $\mathbf{r}^x(t + dt)$, shown in Figure 19 **(c)**. Comparing these velocities with the *causal asymmetry* shows that the changes in α_t are caused by different velocities along the limit-cycle.

In fact, re-normalizing the expansions with the local velocities

$$\hat{e}_{i\to j} = e_{i\to j}\frac{v_{\mathbf{r}^x(t)}}{v_{\mathbf{r}^y(t)}}$$

yields an asymmetry $\hat{\alpha}^t$ invariant against these velocities changes (Figure 19 **(c)**) that is either 0 or positive. This is a better agreement

for a synchronizing unilateral system: Either the right direction of *causal influences* is detected or the systems are synchronized and thus indistinguishable yielding $\alpha_t = 0$.

More details, e.g. different coupling schemes (with matching results) were studied in the Bachelor thesis of Ronja Gronemeyer [71].

4.4 CONCLUSION

Expansive distortions in state space reconstructions reflect *causal influences*. The concept of *topological causality* computes these distortions from the local mappings between state space reconstructions resulting in a time-resolved measure for *causal influence*. In simple systems TC is fully analytically tractable and in systems with weak interactions TC has a linear relation to coupling weights.

In contrast to existing methods like [3, 4], TC infers these *causal influences* directly from the properties of the state space. Synchronous behaviour in a systems (like the neuronal model) can cause TC to infer wrong directions of dominant influence requiring a new approach.

5

CROSS PROJECTION METHOD

All that is gold does not glitter, Not all those who wander are lost..

(J.R.R. Tolkien)

The previous chapter (Chapter 4) introduced the basic idea of *Topological causality*: Local distortions in the mappings between state-space reconstructions based on different observables reflect the time dependent efficacy of causal links among the underlying system components.

However, as shown in the end of the previous chapter (Section 4.3) it is particularly difficult to infer the dominant direction of *causal influence* for synchronising systems. This resulted in a refined approach called cross projection method [2] (CPM) - a more stable advancement of TC that will be covered in this chapter.

The basic idea of this approach is straightforward and closely resembles the concept of TC, but uses the fact that already the neighborhood relations are sufficient to quantify expansive distortions. If (sub-)systems are fully connected, the corresponding points to a set of neighbors in X, will not be randomly distributed in Y. The extent of the distribution of this projected set of neighbors is directly linked to the expansive distortions derived by TC.

Using these inter neighborhood relations to estimate *causal influences* is not new. For interrelated subsystems spatial neighbors within a small range ϵ in one reconstruction will have corresponding points in the other reconstruction that are within a range of similar scale. Čenys et al. [68] derived the average distance of neighborhoods noticing that for small distances there will be a dependence on the coupling strength. Various other distance based approaches, as e.g. in [69] were proposed that use a fixed number of neighbors instead of a fixed neighborhood size. These methods compare the distance to the nearest neighbors with the conditioned distance and/or the mean distance to all other points. Chicharro et. al. [67] improved these distance based methods by quantifying the rank of neighbors instead of their distance. In doing so their method was independent from different subsystem scales and the detection of *causal influences* was improved. While these methods have their merits and drawbacks their relation to TC is at most indirect.

This chapter covers the concept of (CPM). Part of the results where already published by Laminski (myself) and Pawelzik in [2]. Additional sections are added to add more context or details that where not covered in original publication. The first section introduces this

alternative method, in particular how to infer expansive properties and thus *causal influences* from inter neighbor relations. The next addresses basic properties of CPM, dependence on coupling weights, noise and the amount of available data. Then some additional (and necessary) improvements of CPM are introduced overcoming some of the more challenging applications. These are introduced in the final section, ending with a selection of results from experimentally measured time-series.

5.1 EXPANSIVE PROPERTIES FROM INTER-NEIGHBORHOOD RELATIONS

Like Topological Causality this alternative approach relies on the expansion of manifolds of reconstructed observables by combining local and global properties of the relations among nearest neighbours of reference points in both reconstructions to estimate said expansions.

Again, consider time series from dynamical systems composed of two subsystems X and Y. For example, two unilaterally ($w_{y\to x} = 0$ and $w_{x\to y} = 0.3$) coupled logistic maps:

$$\begin{aligned} x(t+1) &= x(t)[R_x(1-x(t)) - w_{y\to x}y(t)] + \eta_x(t) \\ y(t+1) &= y(t)[R_y(1-y(t)) - w_{x\to y}x(t)] + \eta_y(t) \end{aligned} \tag{22}$$

with reflecting boundaries, system parameters R_i and subjected to additive Gaussian noise $\eta_i(t) \in \mathcal{N}(0,\sigma)$, $\langle \eta_i(t)\eta_j(t')\rangle = \delta_{i,j}\delta_{t,t'}\sigma^2$.

For now, we will consider the noise-free case. Here, the time-delay reconstruction of subsystem X can be projected to a single dimension and therefore the whole system can be visualized in three dimensions by showing $x(t)$ over the $(y(t), y(t+1))$-plane (Figure 20).

A 2 dimensional manifold in $\mathbf{r}^x(t) = (x(t), x(t+1))$

It shall be noted that the space X is actually curved, but the curvature is sufficiently small, monotonic and non self-intersecting to not influence local neighborhood relations. Thus, the visualisation in Figure 20 only using $x(t)$ to represent the space X is justified. Besides that, the space $(x(t), x(t+1))$ can parameterised using parabolic coordinates yielding the same neighborhood relations.

t_l^x are time-indices of neighbors in $\mathbf{r}^x$

Each point (shown in Figure 20) is identified by its unique time index $t \in \mathbb{N}$ as well as its location in both reconstructions. For a reference point t we determine the k nearest neighbours in both reconstructed spaces t_l^x and $t_l^y, l = 1, ..., k$. The (euclidean) distances from the reference point to these neighbours in their respective origin space are $L_x(t, t_l^x)$ and $L_y(t, t_l^y)$. Furthermore each neighbor also has a corresponding location - via the time-index - in the respective other reconstruction space, this distance between reference point and putative neighbors is denoted as $L_x(t, t_l^y)$ and $L_y(t, t_l^x)$.

Notation: the subscript of L_i is denoting the space in which distances are measured and the superscript in t_l^i the space in which neighbors have been searched.

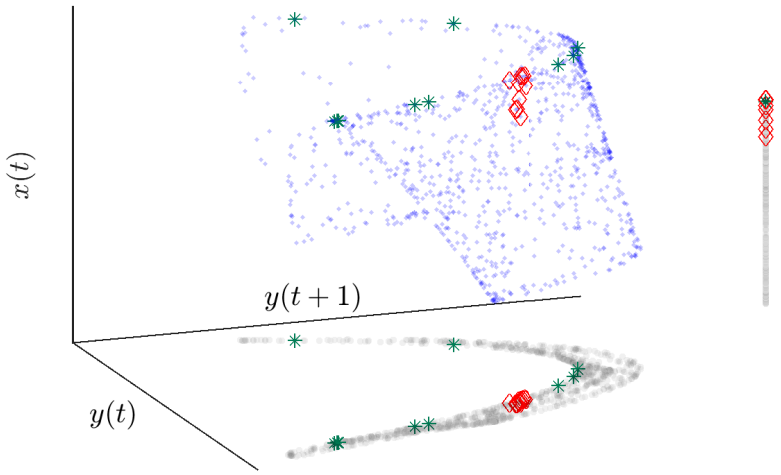


Figure 20: $x(t)$ over $y(t)$ and $y(t+1)$ for the noise free unilaterally coupled logistic maps with $w_{y\to x} = 0, w_{x\to y} = 0.3, R_x = R_y = 3.82$, together with the projection of the manifold on the $y(t+1)$-$y(t)$-plane and the $x(t)$-axis. 10^3 data points are shown in blue the 10 nearest neighbours of a reference point are shown in colour in both, the manifold and the projections (grey). The first 10 neighbours searched in Y (X) are shown in red (green).

Looking at the neighbors t_l^x of a given reference point (green asterisk in Figure 20), it is evident that they are close to each other in their origin space X, but dispersed over the whole space Y. This random distribution in Y is to be expected: If there is no interaction from Y to X, no information about Y can be transferred to X. In contrast, since $w_{x\to y} > 0$, information about X is contained in Y and thus the neighbors t_l^y (red diamonds in Figure 20), are not only localized around the reference point in Y, but also their images remain close together in X.

Coupling weights $w_{i\to j}$ lead to expansive distortions in the reverse direction $J \to I$

With increasing coupling $w_{x\to y}$ the neighbors in X become more localized resulting in the limiting case of perfect information preservation - the neighbours t_l^x and t_l^y become identical. In the other limiting case, $w_{y\to x} = 0$, no information about X is included in Y and the images of the neighbors are spread randomly over the whole space. Thus, by calculating a distance based on random neighbors, we obtain a chance-level as an upper bound for the distance of neighbors $L_i(t, t_l^{i*})$. More details for a suitable chance-level are provided in Section 5.3.2.

A more detailed view on the neighborhoods is depicted in Figure 21 by showing the distances of the first $k = 1, .., 20$ neighbors for some

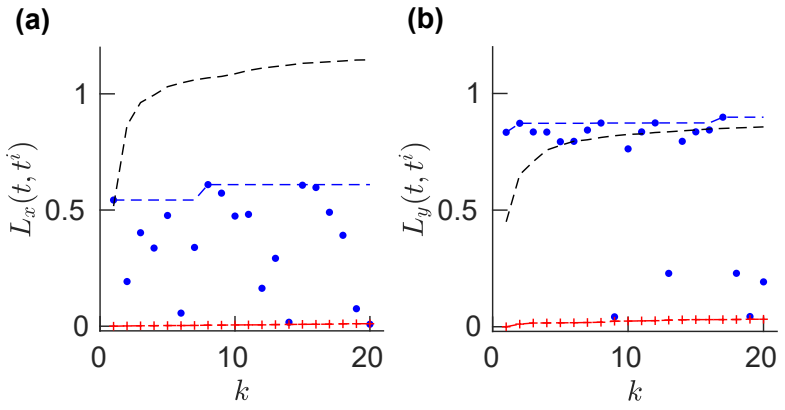


Figure 21: **(a)** Distances $L_x(t, t^x_k)$ from a reference point t to the k-th neighbor (red pluses) and the k-th putative neighbor $L_x(t, t^y_k)$ (blue dots). Blue and red dashed lines mark the respective size of the overall neighborhood, measured as the maximum distance of neighbors up to k. A dashed black line marks a chance level (average cumulative maximum distance of random neighbors) **(b)** Same as **(a)**, but for distances $L_y(t, t^i_k)$ measure in the other reconstruction $\mathbf{r}^y$.

arbitrary reference point. Red pluses signify the distances to the original neighbors of a reference point t, while the blue dots are distances from the reference point derived using the metric of X and the neighbor relations of Y **(a)** and vice versa **(b)** - so called putative neighbors. The dashed lines show a proxy for the size of the neighborhood of k neighbors. Here, the maximum distance up to the k-th neighbor is chosen to reflect the neighborhood size, other suitable options are discussed in Section 5.3.4.

The neighborhood size of the chance-level (black) and the original neighbors (red) limit (on average) the distances of the putative neighbors. If the neighbors t^y contain information about X, the putative neighbors in $L_x(t, t^y)$ in X will be between those limits. The more information is contained, the closer $L_x(t, t^y)$ will resemble the original neighborhood size. In contrast, the putative neighbors $L_y(t, t^x)$ are close to or above chance-level, since the neighborhood t^x has no information about Y and the putative neighbors resemble random points.

In general, not the causal relations of a single reference points is of interest, but rather global properties of the entire time series. Thus, the spread of neighbors in I searched in J is quantified as the mean logarithmic size of the neighborhood $d^j_i(k)$:

x^*, y^* *denoting random neighbors.*

$$d^j_i(k) = \left\langle \log(\max[(L_i(t, t^j_l)]_{l=1..k}) \right\rangle_{t \wedge E} \quad i, j = x, y, x^*, y^* \tag{23}$$

The random neighbourhoods are averaged over both ensembles E and the reference points, whereas for the non-random neighbours the average is only over the reference points.

Ensemble refers to E *different sets of random neighbors.*

These neighborhood sizes can be visualized in plots of the mean logarithmic neighbourhood sizes $d_i^j(k)$ as functions of $\kappa = \psi(k) - \log(N)$, where k is the number of neighbours, N the amount of data and ψ is the digamma-function (Figure 22).

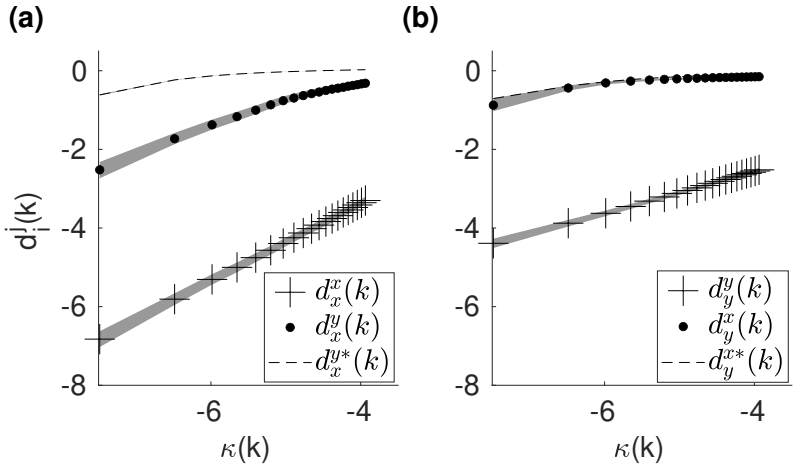


Figure 22: Logarithmic neighbourhood sizes $d_i^j(k)$ over $\kappa(k) = \psi(k) - \log(N)$ for noise free unilateral coupled logistic maps ($w_{x\to y} = 0, w_{y\to x} = 0.2$ and $R_x = R_y = 3.82$) from $N = 10^3$ data points. The observables were embedded with $m = 4$ and $\tau = 1$ and 100 ensembles were used for chance-level-estimation. **(a)** $d_x^x(k)$ shown as plus signs, $d_x^y(k)$ shown as dotted line and the respective chance-level $d_x^{y^*}(k)$ (dashed line). **(b)** $d_y^y(k)$ shown as plus signs, $d_y^x(k)$ shown as dotted line and the respective chance-level $d_y^{x^*}(k)$ (dashed line). Furthermore the five-fold standard error is shown for each $d_i^j(k)$ as a grey shade.

This particular choice of the abscissa allows for an unbiased estimate of the fractal (information) dimensions D_1 (Equation 6) of the subsystems by estimating the inverse slope of $D_1^x = [\Delta d_x^x(k)/\Delta\psi(k)]^{-1}$ (resp. D_1^y) [72]. In the present example the manifold reconstructed from observable X **(a)** has a smaller dimension than the attractor reconstructed from the influenced observable Y **(b)**. While dimensional conflicts provide a sufficient criterion for the direction of *causal influence* in unilaterally coupled deterministic systems [63], they are useless for mutually coupled systems. In this general case a different criterion for determining the dominant direction of *causal influence* is needed.

For this purpose not only $d_x^x(k)$, but also $d_x^y(k)$ and $d_x^{y*}(k)$ are taken into account. Here, the size of a random neighborhood $d_x^{y*}(k)$ is used as an upper bound. If information is transferred from X to Y, the image of neighbors in Y will (in X) be between this upper bound and

$d_x^x(k)$, Figure 22 **(a)**. In contrast, no information is conveyed from Y to X, thus $d_y^x(k)$ matches the random neighborhood $d_y^{x*}(k)$. With $d_x^x(k)$ providing a lower bound and $d_x^*(k)$ an upper bound for the size of the k-th neighbourhood, the size $d_y^x(k)$ can be used to define a measure for the relative amount of information preserved within the neighbourhood in y. With these insights, we use the ratio of the distance between $d_x^x(k)$ and $d_x^y(k)$ and the chance-level $d_x^{y^*}(k)$ to define the *causal influence* (of CPM) $\tilde{I}_{x\to y}(k)$:

$$\tilde{I}_{x\to y}(k) = \frac{d_x^{y*}(k) - d_x^y(k)}{d_x^{y^*}(k) - d_x^x(k)} \tag{24}$$

Note that $0 \leqslant \tilde{I}_{x\to y} \leqslant 1$ and therefore $\tilde{I}_{x\to y} \simeq 0$ means that x does not influence Y at all. If $\tilde{I}_{x\to y} \simeq 1$ Y 'knows everything' about X, which suggests that X has a strong influence on Y.

In analogy to TC (Equation 17), we introduce a measure for the asymmetry of *causal influences* α:

$$\alpha = \frac{I_{y\to x} - I_{x\to y}}{I_{y\to x} + I_{x\to y}}$$

To determine significance we use the standard error (SE) of the mean logarithmic neighbourhood sizes $SE(d_i^j)$, as shown e.g. in Figure 22 **(a)**. With increasing neighborhood size k the $d_i^j(k)$ will inevitable coincide limiting significance to small k. Therefore significance is determined for each neighborhood size $l = 1, .., k$ separately, a significant difference is given if there is no overlap between the 5-fold SE's (i.e. $d_x^{x*} - 5SE(d_x^{x*}) > d_x^y + 5SE(d_x^y)$).

In all further examples we use the average $I_{i\to j} = \left\langle \tilde{I}_{i\to j}(k) \right\rangle_k$ for obtaining a *causal influence* that is independent of a specific neighborhood. Here k is always small, e.g. 20, and we require that at least 75% of all contributions $k = 1, .., l$ to the mean are significant.

75% of 20 neighbors is not any specific statistical criterion, but proved reliable in practice.

5.2 FUNDAMENTAL PROPERTIES

As a described in Chapter 3, certain fundamental properties are desirable for a measure of *causal influences*. For starters, this section will look at the relation of coupling weights and *causal influences* and the influence of internal and external perturbations by noise. For simple examples (logistic maps), in which interactions are governed by coupling weights $w_{i\to j}$, a measure for *causal influence* must be function of said weights $w_{i\to j}$. Furthermore, the impact of data limitation is considered as well.

5.2.1 *Dependence on coupling weights*

Firstly, we investigate the dependence of our measure on the coupling weights. For this purpose, two logistic maps are coupled bilaterally and the *causal influence* is determined for different coupling weights. As a start, consider the mean logarithmic neighborhood sizes $d_i^j(k)$ shown in Figure 23 for two sets of coupling weights. The $d_i^j(k)$ for

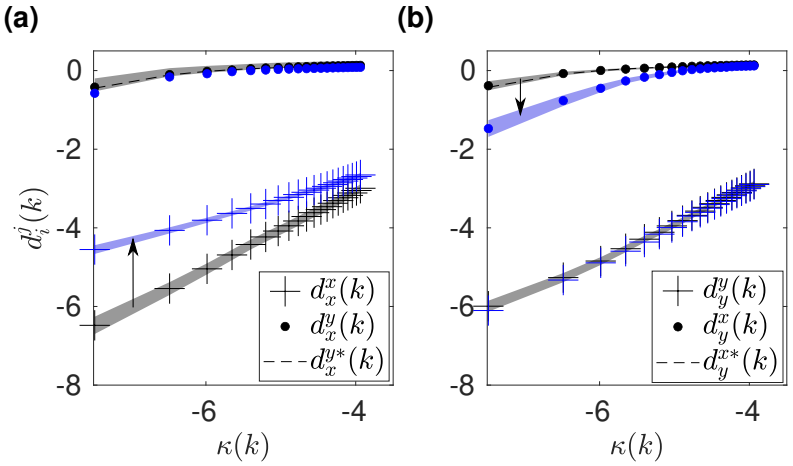


Figure 23: Logarithmic neighbourhood sizes $d_i^j(k)$ over κ for noise free unilateral coupled logistic maps ($w_{x\to y} = 0.00316, w_{y\to x} = 10^{-4}$ (black) and $w_{y\to x} = 10^{-1}$ (blue) and $R_x = R_y = 3.92$) from $N = 10^3$ data points. The observables were embedded with $m = 4$ and $\tau = 1$ and 10^2 ensembles were used for chance-level-estimation. **(a)** $d_x^x(k)$ shown as plus signs, $d_x^y(k)$ shown as dotted line and the respective chance-level $d_x^{y^*}(k)$ (dashed line). **(b)** $d_y^y(k)$ shown as plus signs, $d_y^x(k)$ shown as dotted line and the respective chance-level $d_y^{x^*}(k)$ (dashed line). Furthermore the five-fold standard error is shown for each $d_i^j(k)$ as a grey shade. For the sake of clarity, the standard error for the blue curves is only shown for the non overlapping curves ($d_x^y(k)$ and $d_y^y(k)$).

two small weights $w_{x\to y} = 3.16 \cdot 10^{-}3$ and $w_{y\to x} = 10^{-4}$ are shown in black. Notably, $d_x^y(k)$ and $d_y^x(k)$ are within the respective chance-levels - the coupling weights are too small to detect a meaningful *causal influence*. Furthermore, the slope of $d_x^x(k)$ and $d_y^y(k)$ are matching. Both state space reconstructions of the subsystems have roughly the same dimension and the reconstruction space only represents the respective subsystem. Increasing the weight $w_{x\to y} = 10^{-3}$ leads (primarily) to changes in $d_x^y(k)$ and $d_y^y(k)$ (shown in blue).

On one hand, the slope of $d_y^y(k)$ decreases in Figure 23 **(a)** equating to an increase of the dimension of the reconstruction $\mathbf{r}^y$. The increased weight $w_{x\to y}$ enables more information transfer from X to Y and now not only the subsystem Y, but rather to the overall system (X, Y) is reconstructed in $\mathbf{r}^x$. The *causal influence* $I_{x\to y}$ remains unchanged, since

*Remember, each causal influence is only derived from quantities in either **(a)** or **(b)**.*

the enumerator in Equation 24 still vanishes ($d_x^{x*}(k) - d_x^y(k) \approx 0$). On the other hand, neighbors in Y contain information about X leading to $d_x^y(k)$ deviating from the chance-level (blue dots in **(b)**). This increases the *causal influence* $I_{y\to x}$.

A wider range for coupling weights and the resulting *causal influences* is shown in Figure 24. Here, one weight, $w_{x\to y}$, is fixed to one of the three values (0, 0.00316, 0.032) and the other, $w_{y\to x}$, is changed over three orders of magnitudes. The resulting *causal influence* $I_{y\to x}$ (black symbols) is a monotonic function of its corresponding coupling weight $w_{y\to x}$ and is rather independent of the three counter directional couplings $w_{x\to y}$ (circles, triangles, squares). The reverse direction $I_{x\to y}$ (grey symbols) is largely unaffected by the weight $w_{y\to x}$ not only for a small range of $w_{y\to x}$, but over all three shown different magnitudes of coupling strength. However, for couplings smaller than 10^{-3} no significant *causal influence* is detectable.

black symbols = { square, triangle, circle }

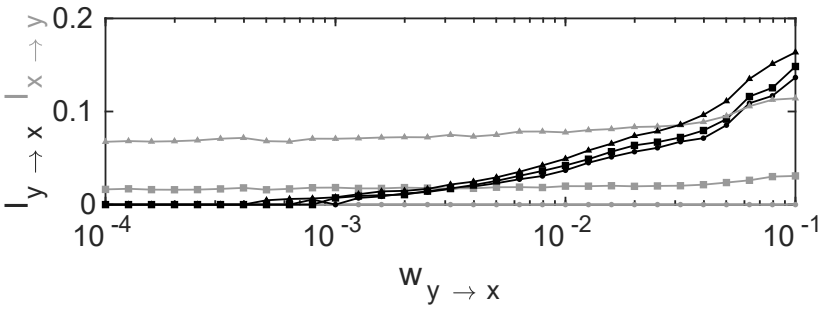


Figure 24: Causal Influence between bilaterally coupled logistic maps. $2 \cdot 10^4$ data points were generated with $R_x = R_y = 3.92$, the observed time series were embedded with $m = 4$ and $\tau = 1$ and we used the $k = 20$ nearest neighbors to determine the causal influence and $E = 10^3$ permutations for the chance-level. Black lines show $I_{y\to x}$ corresponding to the varied coupling $w_{y\to x}$, grey lines show the reverse direction with the coupling $w_{x\to y}$ fixed at 0 (circles), 0.00316 (squares) and 0.032 (triangles).

5.2.1.1 *Relation to Topological Causality*

For the example of logistic maps analytical results of TC were shown in Section 4.2.1. In particular, it was shown that the expansion estimated by TC is directly proportional to the coupling weights, $e_{i\to j} \propto \frac{1}{w_{j\to i}}$. For CPM we are not able to derive analytical results, but at least for coupled logistic maps the relation to the coupling weights can be derived indirectly.

Recapping briefly, TC derives the expansion from a fixed number of nearest neighbors k in I and the corresponding putative neighbors in J. The product of the singular values σ_l^t larger one of the mapping

Putative neighbors in J refers to a selection of points associated due to their neighborhood in I

$M_{i\to j}$ between these neighbors is the expansion, Equation 15:

$$e^t_{i\to j} = \prod_l \max(1, \sigma^t_l(M^t_{i\to j}))$$

In [1] it was shown that the logistic maps in question only have one singular value larger one - a single expansive direction. The expansion is thus equivalent to the largest singular value σ_{max} and the largest eigenvalue $\lambda_{max} = (\sigma^t_{max})^2$ respectively. Here, the expansion also characterizes the relative change of neighborhood sizes going from I to J. This is also reflected in the ratio of (logarithmic) neighborhood size, which is a good proxy for the eigenvalue λ^d_{max}:

$$e^d_{i\to j} = \lambda^d_{max} \approx \frac{exp(< d^j_i(k) >_t)}{exp(< d^i_i(k) >_t)}$$

Here, the enumerator $exp(< d^i_i(k) >_t)$ is the size of the original neighbors and the denominator $exp(< d^i_i(k) >_t)$ reflects the size of the corresponding putative neighbors. Depending on the choice of metric, this ratio either underestimates (e.g. euclidean distance) or matches the eigenvalue (Chebyshev-distance). Analogous to TC a causal asymmetry α^d can be derived, shown in Figure 25.

$$\alpha^d = \frac{e^d_{i\to j} - e^d_{j\to i}}{e^d_{i\to j} + e^d_{j\to i}}$$

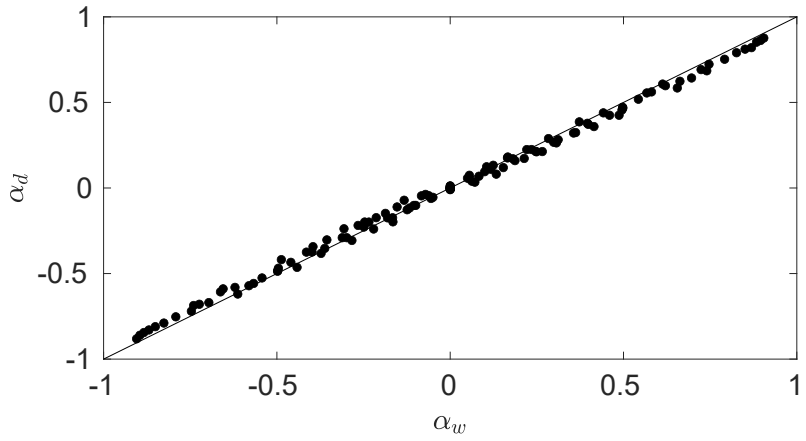


Figure 25: Causal asymmetry between bilaterally coupled logistic maps. 10^4 data points were generated with $R_x = R_y = 3.92$, the observed time series were embedded with $m = 4$ and $\tau = 1$ and we used the $k = 20$ nearest neighbors to determine the causal influence and $E = 10^2$ permutations for the chance-level. Different combinations of coupling weights $w_{x\to y}$ and $w_{y\to x}$ between 0.005 and 0.1 were chosen.

In comparison this yields a good agreement with the asymmetry of TC and the coupling weights.

Atleast for the simple case of coupled logistic maps.

5.2.2 *Sensitivity & Specificity*

The *causal influence* of small coupling weights is not always possible to detect. Taking a close look at small coupling weights in Figure 24, it is evident that small weights ($< 10^{-3}$) were not detectable. These small *causal influences* are only detectable if large(r) amounts of data are used. For illustration, we can look upon the $d_i^j(k)$ to understand how an increased amount data points facilitates the detection of these small *causal influences*. This time $d_i^j(k)$ are shown for fixed coupling weights $w_{x\to y}$ and $w_{y\to x}$, but the amount of data is increased from $N = 10^3$ (black curves) to $N = 8 \cdot 10^3$ (blue curves).

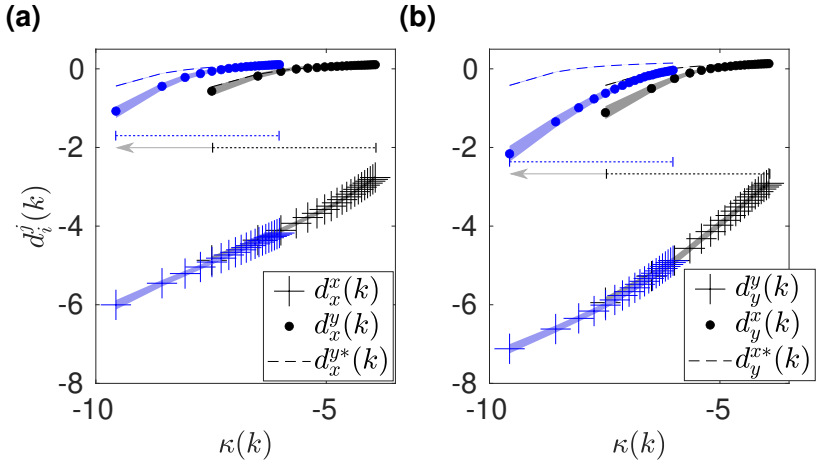


Figure 26: Logarithmic neighbourhood sizes $d_i^j(k)$ over κ for noise free unilateral coupled logistic maps ($w_{x\to y} = 0.005, w_{y\to x} = 0.05$ and $R_x = R_y = 3.92$) from $N = 10^3$ (black) data points and $N = 8 \cdot 10^3$ (blue). The observables were embedded with $m = 4$ and $\tau = 1$ and 10^2 ensembles were used for chance-level-estimation. **(a)** $d_x^x(k)$ shown as plus signs, $d_x^y(k)$ shown as dotted line and the respective chance-level $d_x^{y^*}(k)$ (dashed line). **(b)** $d_y^y(k)$ shown as plus signs, $d_y^x(k)$ shown as dotted line and the respective chance-level $d_y^{x^*}(k)$ (dashed line). Furthermore the five-fold standard error is shown for each $d_i^j(k)$ as a grey (or light blue) shade.

Assuming the density changes are somewhat uniformly.

Since $\kappa(k) = \psi(k) - \log(N)$ is a function of both the number of neighbors k and amount of data N, changing N leads to a shift along the x-axis. Besides that, more data changes the density of the state space resulting in a shrinkage of local neighborhoods. In other words, more data equates to using a smaller original neighborhood (d_x^x & d_y^y) as a basic quantity. In combination, this results in d_x^x and d_y^y shifting along their respective slopes. In contrast, the average size of random neighborhoods is not impacted by the increased density, thus the chance-level (blue dots) is just shifted horizontally. The size of the putative neighborhoods d_x^y & d_y^x shifts as well and more importantly due the more localized original neighbors the size of the putative neighborhoods decreases. This is especially evident for the small cou-

pling weight $w_{x\to y} = 0.005$, where $d_x^y(k)$ now significantly deviates from chance-level enabling the detection of a weight induced *causal influence* previously too small for detection.

A wider range of different amounts of data is shown in Figure 27. Due to symmetry only $I_{x\to y}$ needs to be shown in Figure 27. The

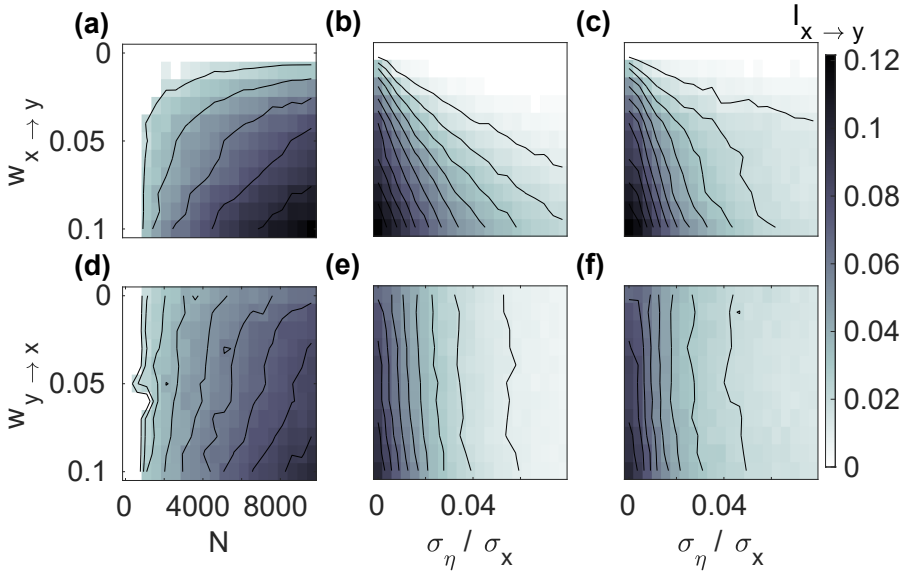


Figure 27: Bilaterally coupled logistic maps $R_x = R_y = 3.92$ embedded with $\tau = 1$ and $m = 4$, 10^3 ensembles were generated to estimate chance-level. The upper row, **(a)-(c)**, shows the causal influence $I_{x\to y}$ for $w_{y\to x} = 0.05$ and varying $w_{x\to y}$ between 0 and 0.1. The bottom row, **(d)-(f)**, shows the same causal influence, but for a fixed $w_{x\to y} = 0.05$ and varying $w_{y\to x}$ between 0 and 0.1. **(a)**,**(d)** Varying amount of data between 10^3 and 10^4. **(b)**,**(e)** Causal influence for additive internal noise. The amplitude of the noise is varied between 0 and 8% of the standard deviation of the unperturbed system. **(c)**,**(f)** Additive external noise is added to the observed time-series. The standard deviations of noise is varied between 0 and 8% of the amplitude of the unpolluted system. For all noise polluted results time-series of 10^4 data points were used and for better visualisation contour lines mark lines of equal causal influence.

upper row shows a varying weight $w_{x\to y}$ between 0 and 0.1 and $w_{y\to x} = 0.05$ is fixed and vice versa in the bottom row. For better visualisation isolines are used to mark equal *causal influences*. For this system at least 10^3 data points are needed to detect significant *causal influences* and the required amount of data points increases for declining coupling strength. Using the example of the *causal influence* $I_{x\to y}$, which depends on the varying coupling weight $w_{x\to y}$, it is evident that there is not enough data to significantly detect a *causal influence* for small $w_{x\to y}$. Note that regardless of the amount of data, no false positives are detected.

Alongside insufficient amounts of data intrinsic and external noise might obscure the underlying causal structure. To demonstrate noise robustness we injected intrinsic and external additive Gaussian noise for varying coupling weights in the logistic maps Figure 27 **(b),(c) & (e),(f)**. The internal noise is added according to Equation 22. The standard deviation of the noise is symmetric for both (sub-)systems $\sigma = \sigma_{\eta_x} = \sigma_{\eta_y}$ and varied between 0% and 8% of the standard deviation of the unperturbed systems. While noise lowers the estimated *causal influence* in both cases, in this example *causal influence* is still correctly determined.

However, the *causal influence* still depends on the coupling weight and amount of induced noise. Thus, it has to be noted that for the case of strong noise $\sigma_\eta/\sigma \to \infty$ and fixed weights the *causal influence* will converge to zero. In particular, small coupling weights will not be detected, although increased amounts of data will remedy this to some degree.

A closer look, how noise affected the *causal influences* via the $d_i^j(k)$ is shown in Figure 28 and Figure 29. Since logarithmic neighborhood sizes for both, intrinsic and external noise, closely resemble one another, only external noise, shown in Figure 28, is discussed and internal noise is only shown for the sake of completeness.

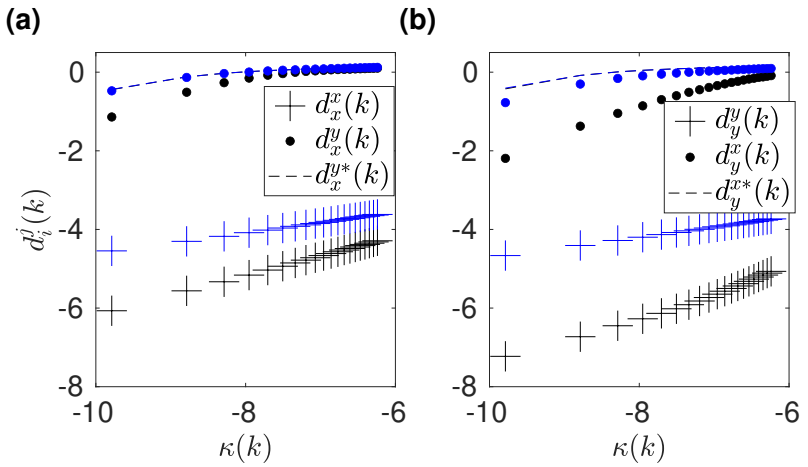


Figure 28: Logarithmic neighbourhood sizes $d_i^j(k)$ over κ for bilaterally coupled logistic maps ($w_{x\to y} = 0.005, w_{y\to x} = 0.05$ and $R_x = R_y = 3.92$) from $N = 10^3$ data points. The observables were embedded with $m = 4$ and $\tau = 1$ and 10^2 ensembles were used for chance-level-estimation. Black curves show the results for systems free of noise and blue curves show additive (uncorrelated) external noise of $\sigma = 0.0125$ on both maps.**(a)** $d_x^x(k)$ shown as plus signs, $d_x^y(k)$ shown as dotted line and the respective chance-level $d_x^{y^*}(k)$ (dashed line). **(b)** $d_y^y(k)$ shown as plus signs, $d_y^x(k)$ shown as dotted line and the respective chance-level $d_y^{x^*}(k)$ (dashed line).

The black symbols correspond to an unperturbed system, while the blue symbols are external noise of $\sigma^2 = 0.0125$. The random neighborhoods are largely invariant to noise (overlapping dashed lines). In contrast, the nearest neighbor distances, $d_x^x(k)$ and $d_y^y(k)$, increase, since noise blurs the shape of the attractor. In particular, in unexplored directions of state space. For example, the logistic maps no longer has its signature parabola shape in state space, but gains an additional thickness. However, the crucial effect of noise is on the quantities are $d_x^y(k)$ and $d_y^x(k)$ (dotted lines). These quantities are indirectly affected through $d_x^x(k)$ and $d_y^y(k)$, but moreover the injection of noise destroys neighborhood relations.

The state space is no longer parabola, but banana-shaped.

Here, two effects are influencing $d_x^y(k)$ (and $d_x^y(k)$). Briefly, only consider just $d_x^y(k)$. On one hand noise in Y shifts false nearest neighbors into the vicinity of a reference point in Y. These are indeed the closest points to a reference point in Y, but their adjacency is only due to to noise and thus their corresponding points in X have an arbitrary distance to the reference point. On the other hand even if the correct neighbors are classified in Y, their corresponding points in X are rearranged due to the noise in this part of the subsystem. In combination, both effects increase $d_x^y(k)$ and results in a decrease of *causal influence* in Figure 27. Furthermore, this suggest that a more reliable way to determine the nearest neighbors will improve the detection of *causal influences*. This will be discussed in section Section 5.3.6.

Already moving average filters can lead to better nearest neighbors in the presence of noise.

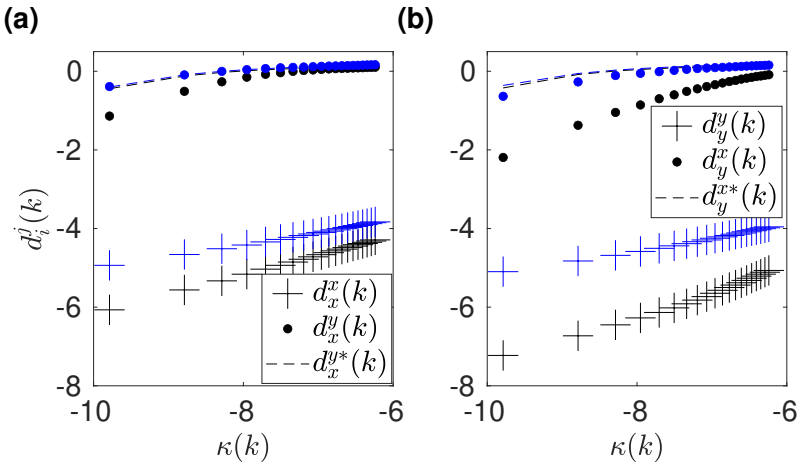


Figure 29: Logarithmic neighbourhood sizes $d_i^j(k)$ over κ for bilaterally coupled logistic maps ($w_{x\to y} = 0.005, w_{y\to x} = 0.05$ and $R_x = R_y = 3.92$) from $N = 10^3$ data points. The observables were embedded with $m = 4$ and $\tau = 1$ and 10^2 ensembles were used for chance-level-estimation. Black curves show the results for systems free of noise and blue curves show additive (uncorrelated) intrinsic noise of $\sigma = 0.0125$ on both maps. **(a)** $d_x^x(k)$ shown as plus signs, $d_x^y(k)$ shown as dotted line and the respective chance-level $d_x^{y^*}(k)$ (dashed line). **(b)** $d_y^y(k)$ shown as plus signs, $d_y^x(k)$ shown as dotted line and the respective chance-level $d_y^{x^*}(k)$ (dashed line).

5.2.3 *Application of CPM to a 10-species Lotka-Volterra system*

As an example for a more complex system dealing with both, limited amounts of data and intrinsic noise, we consider a ten species Lotka-Volterra system [73] modeled as a generalisation of a discrete two species Ricker model with added stochasticity:

$$\mathbf{N}_{t+1} = \mathbf{N}_t \exp(-\mathbf{r} + W\mathbf{N}_t + \eta_t), \eta_t \in \mathcal{N}(0, 0.1\mathbf{I})$$

The interaction matrix W is represented in Fig. Figure 30 **(a)** and the intrinsic growth rates are given by

$$\mathbf{r} = -[3, 2.1, 1.12, 2.8, 1.4, 2.1, 1.12, 3, 2.1, 1.12]^T \tag{25}$$

The interaction matrix inferred using CPM is shown in the Figure 30 **(b)**. CPM is correctly recovering the block-wise structure of the interaction matrix, but due to the limited amount of data (N = 300) the smallest interactions can not correctly by determined in all 25 simulations.

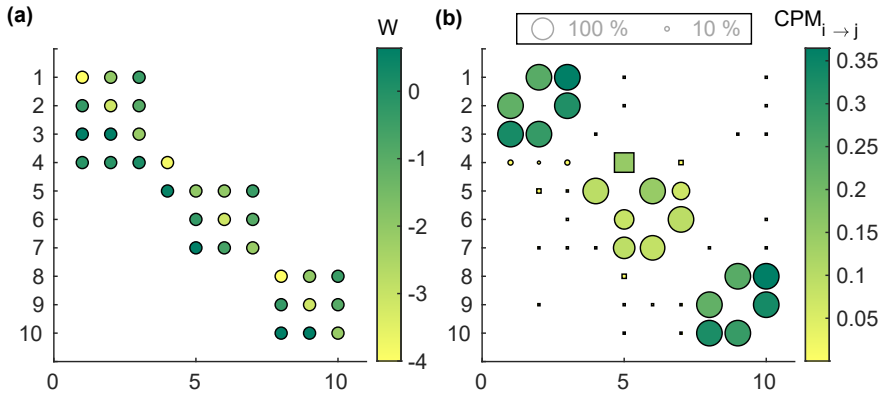


Figure 30: (a) Coupling matrix W of a 10-Species Ricker model. A darker color signifies a larger entry in W. (b) Interaction matrix W derived using CPM, circles represent true positive entries and squares false positives. The interaction matrix was derived from time series of length 300 and embedded using a delay of $\tau = 1$ and embedding dimension of $m = 8$ using the first $k = 20$ neighbors. The simulation was repeated 25 times, the radius of the circle is proportional to the proportion of detection. The color of the circle is the average value of CPM over all 25 simulations.

It shall be noted, the presentation, in particular the choice of radius instead of volume to represent the proportion of detection, is in analogy to [73].

5.2.4 *Time continuous systems - Lorenz attractor*

Up to now, the *causal influence* was contemplated in time discrete systems, but likewise determination of the direction of *causal influence* is viable in time continuous systems, e.g. two coupled Lorenz-systems. The Lorenz systems are coupled by their x-components and are given by the set of following differential equations:

$$\begin{aligned} \dot{x}_i &= -\mu_i(x_i - y_i) + \Omega_{j\to i} + \sigma\eta_i^x \\ \dot{y}_i &= \rho_i x_i - y_i - x_i z_i + \sigma\eta_i^y \\ \dot{z}_i &= -\theta_i z_i + x_i y_i + \sigma\eta_i^z \\ \text{with} \quad &< \eta_i^\lambda \eta_j^\gamma >= \delta_{ij}\delta_{\lambda\gamma}\delta(t - t'), \end{aligned} \tag{26}$$

with the typical set of parameters $\mu_i = 28$, $\rho_i = 8/3$ and $\theta_i = 10$. To account for the intrinsic noise the differential equations are solved by the (explicit) Euler scheme. The external additive Gaussian noise is superimposed onto the measured observable, in this case $x_i + \sigma\eta_i^{ext}$. The coupling function is chosen linearly by $\Omega_{j\to i} = w_{j\to i}x_j$. In this

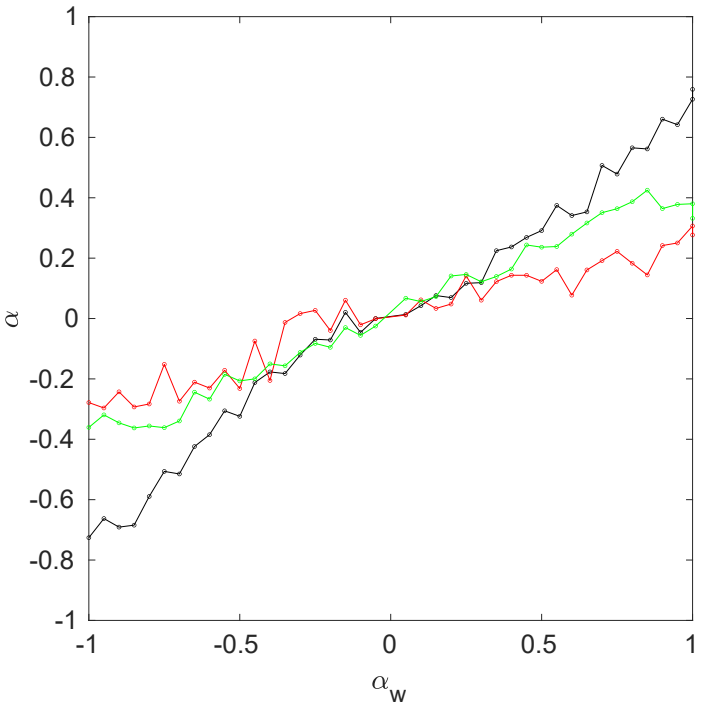


Figure 31: Asymmetry index between two Lorenz oscillators ($N = 10^4$ data points) with slightly different frequencies and coupled by their y-components ($\theta_{1/2} = 10, \rho_1 = 28.5, \rho_2 = 27.5, \theta_{1/2} = 8/3$). The time series were embedded with $m = 9$ and $\tau = 10$, using $k = 20$ neighbors and $E = 10^2$ ensembles were used to estimate chance-level. The noise free asymmetry is shown in black. Coloured circles represent internal noise with $\sigma = 2$ (green) and additive external noise with $\sigma = 2$ (red).

case the causal asymmetry has a (somewhat) linear functional relation with the asymmetry derived from weights, ($\alpha_w = \frac{w_{y\to x} - w_{x\to y}}{w_{y\to x} + w_{x\to y}}$). In Figure 31 the coupling weights $w_{i\to j}$ are chosen in a way that α_w is spread uniformly. Even for both, external (red) and intrinsic (green) noise, the direction of *causal influence* is determined correctly by α. However, in the presence of noise the *causal influence* is weakened and the asymmetry less pronounced, which is in agreement with the previously considered example of logistic maps.

However, since the change in *causal influence* depends on the coupling weight, strong asymmetric noise injection, e.g. strong noise in X and no noise in Y, could disguise the *causal influence* enough leading to a wrong classification of the dominant direction of *causal influence*. These spurious influences and possible remedies to prevent it even in the case of strong noise are discussed in a later section, Section 5.4.

5.3 DESIGN CHOICES

While the basic idea of CPM was already introduced in Section 5.1 some details and alternative design choices have not been discussed. This section will explain the choice of chance-level, discuss a time-resolved variant of CPM, elaborate on the estimation of neighborhood size, compare rank and distance-based CPM and finally go over improvements in the selection of nearest neighbors. The first section are more or less straightforward and typical approaches already established in literature, the latter section (selection of neighbors) could also disclose beneficial ideas for related methods.

5.3.1 *Temporal Neighbors*

In systems like the logistic maps (temporal) consecutive neighbors are far apart. But in continuous systems like the Rössler system (shown in Figure 32) time evolves along paths in state space. Thus, a neighborhood around a reference point also contains its temporal neighbors. In particular, for small integration step-sizes dt many neighbors are temporal neighbors. The neighborhood size is supposed to be invariant against changes in δt and temporal correlations with the reference point shall be avoided. Thus, the immediate temporal neighbors ($t \pm 2m\tau$) are excluded from the neighborhood search.

In fact, there is an argument to be made to only allow the closest neighbor per path, the neighborhood size then corresponds to the spread in the Poincaré-section [74]. In practice, we rarely found many close neighbors from the same path, not warranting the considerably larger computing cost. However, this might be a specific property of

The shown example is simplified projection.

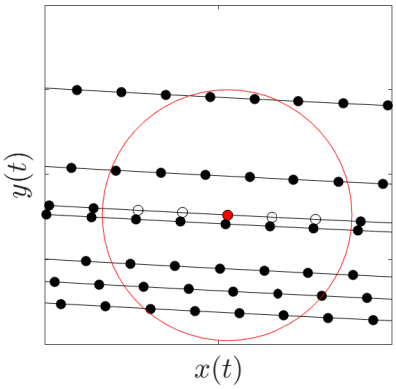


Figure 32: Section of the state space of a Rössler system. Lines connect consecutive points in time. Empty circles represent the temporal neighbors of the red reference point.

the investigated systems and could prove necessary in other scenarios.

5.3.2 *Choice of Chance-Level*

The straightforward and simple way to define a chance-level for the neighborhood size is to choose a set of k random neighbors t_l^{x*}. The extent of this random neighborhood is then the chance-level for the neighborhood size of k neighbors.

Instead an ensemble E of surrogate neighbors is generated by shifting the time indices of all actual neighbors of each reference point with a random time shift δt_E. The shift is random, but also excludes temporal neighbors $\delta t_E > \pm 2m\tau$. In essence, the time series are shifted against each other. When applying this random time shift to the time index of neighbors t_l^x in space X, a given reference point t becomes associated with a set of points $t_l^{x^*}$ and has distances $L_y(t, t_l^{x^*})$ to the images of these points in space Y.

In doing so causal relations of neighborhoods between different observables are removed while the temporal correlations within the set of surrogate neighbors are preserved. In the results presented in this work the chance-level is derived from temporally shifted random neighborhoods.

5.3.3 *Time-resolved CPM*

In Chapter 4 it was shown that TC ([1]) is able to provide a time-resolved measure for *causal influences*. In fact, CPM can be modified to retain a temporal resolution. For this purpose, the temporal average in Equation 23 is omitted from the logarithmic neighborhood size:

$$\bar{d}_i^j(k,t) = \left\langle \log(\max[(L_i(t,t_l^j)]_{l=1..k}) \right\rangle_E$$

In analogy to the original *causal influences* (Equation 24) and *asymmetries* time-resolved variants, $\bar{I}_{i\to j}(t)$ and $\bar{\alpha}(t)$, are derived for CPM. The chance-level $\bar{d}^*_i(k,t)$ can still be averaged over the ensembles, but the other $\bar{d}^j_i(k,t)$ are volatile and depend strongly on single neighbors. Therefore, in Figure 33 all $\bar{d}^j_i(k,t)$ smaller than the corresponding chance-level contribute to $\bar{I}_{i\to j}(t)$ and no criterion for significance is used.

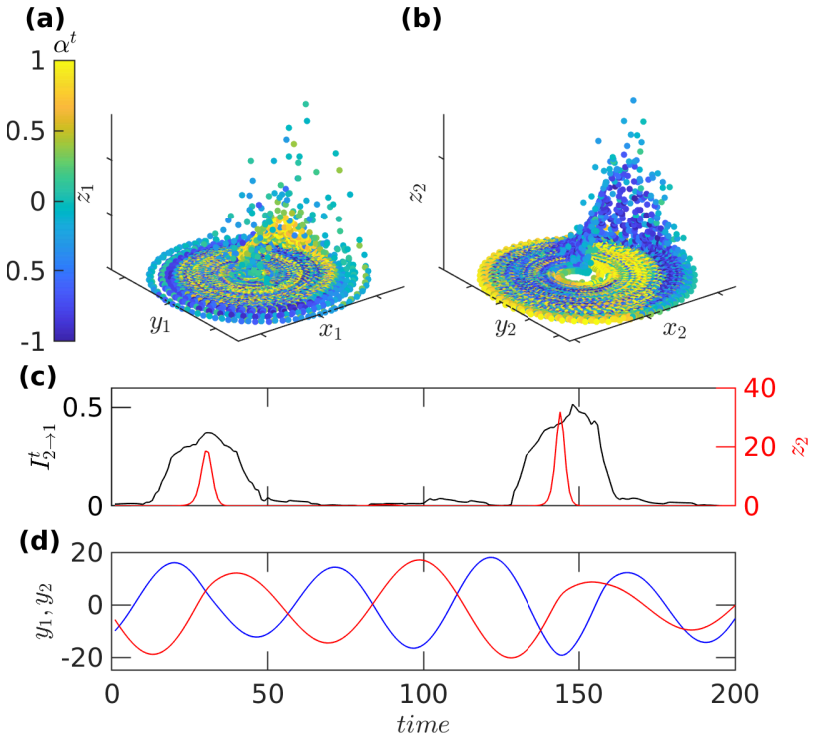


Figure 33: Two bidirectionally coupled Rössler systems with $w_{2\to1} = 0.2$ and $w_{1\to2} = 0.6$. The time series of 10^5 data points were embedded with dimension $m = 13$ and $\tau = 3$. α^t and $I^t_{2\to1}$ are shifted to be aligned with the temporal mean $t + 1/2(m-1)\tau$ of the corresponding reconstructions $\mathbf{r}^{x_i}(t)$. **(a)** & **(b)** Local asymmetry α^t of 10^4 points shown on projections of the attractor to each system. **(c)** $I^t_{2\to1}$ (black) for 200 consecutive time steps and the corresponding time series of z_2 (red). **(d)** Time series y_1 and y_2 used to estimate $I^t_{2\to1}$.

The example used earlier are two bidirectional coupled Rössler-systems, described by:

$$\begin{aligned}
\dot{x}_i(t) &= -f_i y_i(t) - z_i(t) + w_{j\to i} z_j(t) \\
\dot{y}_i(t) &= f_i x_i(t) + 0.1 y_i(t) \\
\dot{z}_i(t) &= 0.1 + z_i(t)[x_i(t) - 14]
\end{aligned}$$

The systems are coupled via additive input of the respective other z_j-component into x_i. $\{f_1, f_2\}$ are set to $\{0.99, 0.85\}$ and $y_1(t)$ and $y_2(t)$ are used as measurements from the individual systems.

The *causal measures* (of CPM) are shown in Figure 33 **(a)** & **(b)** ($\bar{\alpha}(t)$) and **(c)** ($\bar{I}_{x\to y}(t)$). By choosing the coupling function in the way shown above, a strong *causal influence* $i \to j$ is only expected if the driving z_i component deviates from 0. Both, $\alpha(t)$ and $\bar{I}_{i\to}(t)$, capture this temporal structure, which is not obvious from the time courses of the observed time series y_1, y_2 (Figure 33 **(d)**). This is especially evident in the direct match with $z_2(t)$ **(c)**. These results match the results of TC qualitatively and even outperform it in some-sense: There is no *causal influence* from 2 to 1 for $z_2 = 0$, $I^t_{2\to1}$ drops to approximately zero, while TC still inferred some influence for this case (Figure 17).

Side note: the title-page shows CPM for the unilateral coupled case.

5.3.4 *Neighborhood Size*

In Section 5.1 the neighborhood size was estimated as the maximum distance of k neighbors to a reference point. In other word, the neighborhood size is the hyper-sphere covering all k neighbors. It is evident that this choice is strongly dependent on the position of the most distant neighbors. Being less prone to single outliers should be achievable by using the mean or median distance of the neighbors as a measure of neighborhood size. In Figure 34 all three measures are

All three are still averaged over reference points.

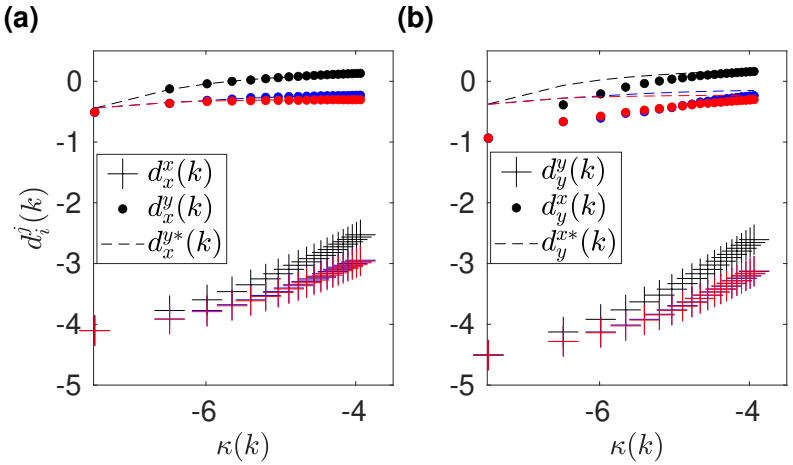


Figure 34: Logarithmic neighbourhood sizes $d^j_i(k)$ over κ for unilateral coupled logistic maps ($w_{x\to y} = 0.00316, w_{y\to x} = 0.1$ and $R_x = R_y = 3.92$) from $N = 10^3$ data points and $N = 8 \cdot 10^3$. The logistic maps are perturbed by additive intrinsic Gaussian noise $\mathcal{N}(0, 0.01 \cdot \mathbb{1})$. The observables were embedded with $m = 4$ and $\tau = 1$ and 10^2 ensembles were used for chance-level-estimation. **(a)** $d^x_x(k)$ shown as plus signs, $d^y_x(k)$ shown as dotted line and the respective chance-level $d^{y^*}_x(k)$ (dashed line). **(b)** $d^y_y(k)$ shown as plus signs, $d^x_y(k)$ shown as dotted line and the respective chance-level $d^{x^*}_y(k)$ (dashed line). Different colors represent different measures for neighborhood size, black lines are the maximum distance, blue lines the mean and red lines the median.

shown for a system of two coupled logistic maps perturbed by intrin-

sic noise. The mean logarithmic neighborhood size is derived from the maximum (back), mean (blue) and median (red) distance of k neighbors. All three are qualitatively equivalent and only vary quantitatively. In general, noise does not impact outliers stronger, but rather changes the whole neighborhood layout. Thus, other approaches, discussed in Section 5.3.6, are necessary.

The curves for mean and median are strongly overlapping and can be hard to see in Figure 34

5.3.5 *The Rank as a proxy for Neighborhood Size*

Another alternative to estimate the size of neighborhoods can be achieved by ranking neighbors. The basic idea: neighbors are ranked according to their distance to the reference point, i.e. the k-th neighbor searched in I has the rank $R_i(t, t^i_k) = k$. The corresponding point in J is linked by their shared time index t^i_k. Finding this time index in the ranking $R_j(t, t^j_k)$ yields the ranking of these putative neighbors $R_j(t, t^i_k)$. These ranks are comparable to the neighborhood size $L^i_j(k)$. Comparing with $d^i_j(k)$ it is sensible to look upon the mean logarithmic rank ${}^R d^i_j(k) =< \log(R_j(t, t^i_k)) >_t$. Figure 35 shows the comparison between maximum rank and maximum distance of neighbors. Again, both measures are only quantitatively different. Naturally, besides the maximum rank, mean or median rank can be considered as well.

The superscript d just denotes that these ${}^R d^i_j(k)$ are the analog to $d^i_j(k)$ but derived from ranks.

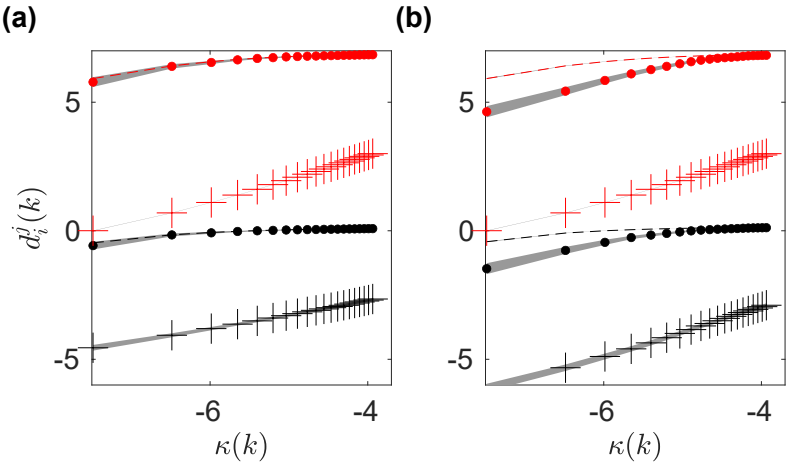


Figure 35: Logarithmic neighbourhood sizes $d^j_i(k)$ over κ for noise free unilateral coupled logistic maps ($w_{x\to y} = 0.00316, w_{y\to x} = 0.1$ and $R_x = R_y = 3.92$) from $N = 10^3$ data points. Red lines use the ranking of neighbors compute an neighborhood size. The observables were embedded with $m = 4$ and $\tau = 1$ and 10^2 ensembles were used for chance-level-estimation. **(a)** $d^x_x(k)$ shown as plus signs, $d^y_x(k)$ shown as dotted line and the respective chance-level $d^{y^*}_x(k)$ (dashed line). **(b)** $d^y_y(k)$ shown as plus signs, $d^x_y(k)$ shown as dotted line and the respective chance-level $d^{x^*}_y(k)$ (dashed line). Black lines correspond to the maximum distances, while red lines are using the rank of neighbors as a proxy for neighborhood size.

5.3.6 *Noise perturbed neighborhoods*

If a system is subject to noise, neighbor relations change. False neighbors are moved into neighborhoods and correct neighbors are removed. Section 5.2.2 shows that this affects the size of neighborhoods in both spaces.

To improve the neighborhood relations a measure to evaluate the quality of neighborhoods is needed. For this purpose neighbors are searched in the reconstruction of the unperturbed time series $x(t)$. The time-series is from the example system of a unilaterally coupled Lorenz→Lorenz system. Like in the previous section, neighbors can be ranked according to their distance to the reference point t, e.g. the k-th neighbor in X has rank $R^0_x(t, t^x_k) = k$. This is true for all reference points and their respective k-th neighbor, trivially the average rank of all k-th neighbors is $\left\langle R^0_x(t, t^x_k)\right\rangle = k$.

Next the time series $x(t)$ is perturbed with noise $\tilde{\mathbf{x}}(t) = \mathbf{x}(t) + \eta_x, \eta_x(t) \in \mathcal{N}(0, \sigma)$ and the neighborhood search is repeated. Each point in the perturbed and unperturbed time series remains identified by the same time-index t, but neighborhood relations changed due to noise. The l-th neighbor in $\tilde{X}$ is associated with the time index $\tilde{t}_l$. This time index is linked to a time index t_k in X and we can find the corresponding rank of the l-th neighbor in X, e.g. $R^\eta_x(t, \tilde{t}_l) = R^0_x(t, t_k = \tilde{t}_l)$. On average these ranks will be larger for the perturbed time series and increase with the standard deviation of the added noise. If a neighborhood in the perturbed time series is closely related to the unperturbed neighborhood, this will be reflected in a small increase in the average rank. Strongly distorted neighborhoods will be identifiable by large rank increases.

The superscript 0 denotes the rank of a noise free time-series.

Figure 36 (a) shows the average rank of the first 200 neighbors of perturbed time series generated by Equation 26 for a selection of noise standard deviations between 1.28 and 2 (full lines, with lighter colors signifying weaker noise). The ranks of the unperturbed (original) time series are shown as dashed line.

With increasing noise neighborhoods are stronger distorted and the rank R^η_x is increasing signifying a *worse* conformity with the unperturbed neighborhood relations.

To reduce the rank and thus improve the neighborhood relations in the presence of noise we here discuss three approaches. Firstly, a more general distance metric can be used. For this purpose we use the Minkowski distance as a generalization of euclidean and Manhattan

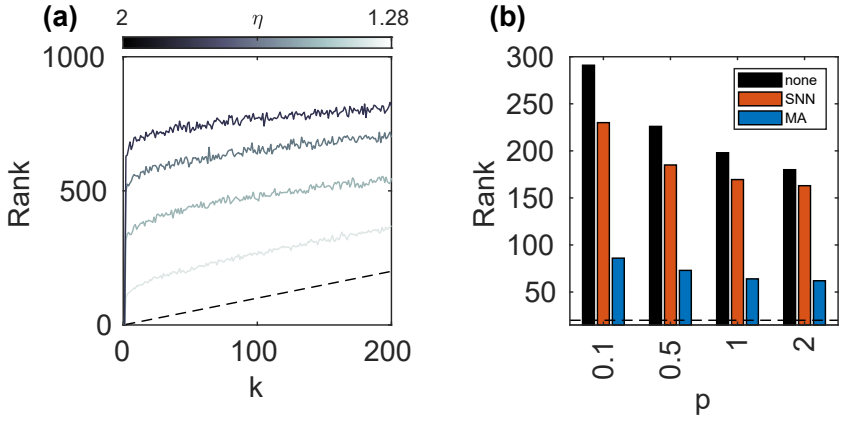


Figure 36: **(a)** Average rank of neighbors over the number of neighbors k for different noise levels shown in shades of grey. The dashed line marks the base line of the neighbor rank of a unperturbed time-series. **(b)** Average rank of the first $k = 20$ neighbors for 4 different Minkowski distances (0.1 to 2). To improve the neighborhood relations, moving average (blue), shared nearest neighbors (red) and no method (black). For each curve/bar a time series of $N = 2^{11}$ data points was generated by Equation 26. For the moving average filter with a window size of 8 time steps was used before searching neighbors. (blue bars) The shared nearest neighbors were chosen as a subset of the first 200 neighbors that also shared 40 neighbors with a reference point (red bars). The black bars show the rank for a perturbed time series.

distances. In doing so the distance between two points $A = (a_1, a_2, ..)$ and $B = (b_1, b_2, ..)$ is defined by:

$$D(A, B) = \left(\sum_i^n |a_i - b_i|^p \right)^{(1/p)}$$

An exponent of $p = 2$ being the Euclidean distance, p between 0.1 and 2 are shown on the abscissa in Figure 36 **(b)**. Secondly, a moving average filter is chosen to smooth the time series $\tilde{\mathbf{x}}(t)$ before even searching for neighbors. Lastly, the Jarvis-Patrick Clustering algorithm [75] is used to validate the shared nearest neighbors (SNN) of each reference point. As depicted in 36 **(b)** preprocessing the data by moving average filtering attains the strongest improvement. The moving average of a time-series $x(t)$ over p steps is given by:

$$MA_p(x(t)) = \frac{1}{p} \sum_{i=t-p+1}^{t} x(i)$$

The filter is also used to improve the noise perturbed unilateral coupled Lorenz→Lorenz system used in Figure 39 **(a)**. Here, it is evident that the improved neighborhoods also improve the detection of *causal*

influences and prevent the detection of misleading results in this case of strong unilateral coupling weights in combination with strong one-sided noise. This is particularly clear for large coupling weights were the causal differences ΔI mostly retain a positive sign.

5.4 SYNCHRONIZATION

In Section 4.3 a set of Fitzhugh-Nagumo neurons was introduced as a challenging problem for TC. Retaining some results a re-normalization of TC was introduced. However, the difficulties are only partly due to flaws of TC, but rather results of the chosen system. In particular, oscillations and synchronization affect the inference of *causal influences*, irrespective of the chosen measure.

If two (sub-)systems governed by the same underlying equations but slightly different parameter values are interacting, synchronous behavior is prevalent for a wide range of parameters. The Fitzhugh-Nagumo analyzed by TC (Section 4.3) are given by:

$$\dot{x_1} = x_1 - \frac{1}{3}x_1^3 - x_2 + I + \Omega_{i\to j}, \qquad \Omega_{i\to j} = w_{i\to j}\frac{1}{1+e^{-3i_1}}$$
$$\dot{x_2} = f_i(x_1 - a_i x_2 + b_i)$$

The parameter values a_i, b_i and f_i were chosen slightly different, but due to strong coupling weights $w_{i\to j}$ the previously considered example are two synchronous (sub-) systems. Here, statements about *causal influences* are meaningless, maximal influences with no *causal asymmetry* are to be expected. Thus, a criterion to determine whether a system is already fully synchronous is desirable.

In the case of the Fitzhugh-Nagumo neurons strong coupling weights enforce synchronous dynamics. Different oscillation periods of the subsystems can further aid or prohibit synchronization.

Oscillations periods are derived for uncoupled systems.

The state space of the affected system $\mathbf{r}^x(t)$ is shown for a small (unilateral) coupling weight $w_{y\to x} = 0.3$ and slower driving system Y in Figure 37 **(a)** and for a faster driving system in **(d)**. The limit-cycle in **(d)** is inflated, the affecting system introduced additional degrees of freedom leading to an increase of the dimension of the system. In contrast, for the faster driving system, in **(a)**, the reconstructions forms a joint limit cycle and the dimension collapses to the dimension of the affecting system - both subsystems are synchronous.

Whether the (sub-)system are synchronous is directly reflected in the *causal influence* $I_{i\to j}$. The *causal influence* for varying coupling weights between 0 and 0.5 and two different ratios of frequencies f_x/f_y is shown in **(b)** and **(e)**. The blue curves correspond to the *causal influence* $I_{y\to x}$ - in direction of the non zero coupling weight $w_{y\to x}$. The red curves show the reverse direction. Already for coupling weights

$w_{y\to x} \approx 0.15$ a slow driving system forces the other subsystem to synchronize, **(b)**. While for a faster driving system at least the dominant direction of *causal influences* is detectable for all shown weights **(e)**.

All frequencies were chosen to be in-commensurable

For a fixed coupling weight $w_{y\to x}$ certain frequency ratios can also lead to synchronization. In **(c)** the *causal influence* $I_{i\to j}$ is shown for different frequency ratios between $f_x/f_y \approx 1/3$ and $f_x/f_y \approx 3/1$ and a fixed unilateral coupling $w_{y\to x} = 0.3$. Again, only for a considerably faster driving system meaningful *causal influences* are inferable. Only for this case the two (sub-)systems are not synchronous. For vanishing coupling weights, **(f)** CPM detects false positive *causal influences*, if the frequency ratio is approximately 1. Here, the uncoupled systems just appear to be synchronous.

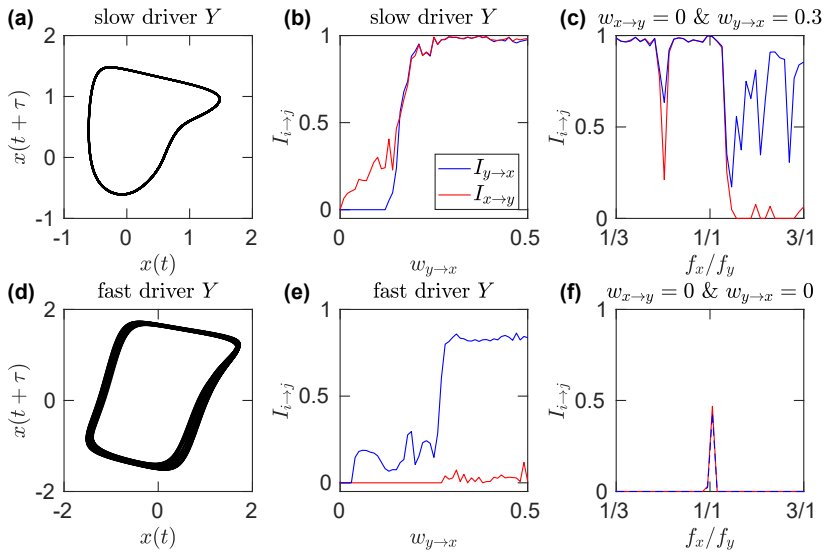


Figure 37: **(a)** & **(d)** show the state space reconstruction $\mathbf{r}^x$ of the voltage variable of two unilateral coupled Fitzhugh-Nagumo neurons ($w_{y\to x}$) with a time-delay $\tau = 110$, equaling roughly a quarter period. The shown reconstruction is the state space of the driven subsystem for slow driving system **(a)** and a fast driving system **(d)**. The causal influence between the two unilateral coupled Fitzhugh-Nagumo neurons is shown for varying coupling weight $w_{y\to x}$ between 0 and 0.5. 10^3 reference points were chosen from time series of 10^4 data points that were embedded using $\tau = 100$ and $m = 5$. **(c)** shows the causal influence for a slow driving systems, the blue curves correspond to $I_{y\to x}$ (in direction of the changing coupling weight) and the red curve the reverse direction $I_{x\to y}$. **(e)** shows the same results for a fast driving system. **(c)** shows $I_{i\to j}$ for different frequency ratios between $f_x/f_y = 1/3$ and $3/1$ for fixed unilateral coupling $w_{x\to y} = 0$ and $w_{y\to x} = 0.3$. **(f)** shows $I_{i\to j}$ for different frequency ratios between $f_x/f_y = 1/3$ and $3/1$ and not interaction.

Instead of inspecting the state space for changes of dimension, the fundamental properties of our method, in particular the logarithmic neighborhood size $d_i^i(k)$, is linked to the information dimension. Thus we will use this property as a measure for synchronization, while investigating the *causal influence* in time continuous systems with increasing degrees of synchronization. The information dimension is derived as the the inverse slope of $d_i^i(k)$ over $\psi(k)$, $D_1 \approx \left[\Delta d_i^i(k)/\Delta\psi(k)\right]^{-1}$. For unilateral coupling this is a sensitive measure, since (as aforementioned) for complete synchronization the dimension of the driven system will drop to the one of the driving system. As an example we use a Rössler-Lorenz-system as analysed in [62]:

$$
\begin{aligned}
\dot{x}_1 &= -6(y_1 + z_1) \\
\dot{y}_1 &= 6(x_1 + 0.2y_1) \\
\dot{z}_1 &= 6[0.2 + z_1(x_1 - 5.7)] \\
\dot{x}_2 &= -10(x_2 - y_2) \\
\dot{y}_2 &= 28x_2 - y_2 - x_2z_2 + wy_1^2 \\
\dot{z}_2 &= -8/3z_2 + x_2y_2
\end{aligned}
$$

Since two different dynamical systems are interacting, synchronization only occurs for strong coupling weights. In this system synchronization sets in at a critical value of $w \approx 2.5$, where the dimensions of the reconstructions approach each other (grey dashes in Figure 38 **(b)**). However, the dimensions remain different indicating that full synchronization is not achieved.

CPM is not only able to detect the correct direction of *causal influence* before, but also after this critical coupling is reached (Figure 38 **(a)**). We found that intrinsic noise can even improve the detectability of causal asymmetries (dotted lines in Figure 38). This result is in contrast to other approaches [62], where e.g. Convergent Cross Mapping fails to detect the correct causal direction for strong and weak couplings (38 **(d)**). Furthermore CPM is able to detect the critical coupling at which synchronization sets in. In the case of CPM at this critical coupling the *causal influence* is increased strongly while the dominant direction is still correctly determined, e.g. the asymmetry index does not change its sign even for large w. Besides CPM and CCM we also show results for method M from [67] (Figure 38 **(c)**). This methods performs similar to CPM, but have difficulties detecting the correct direction of *causal influence* for very strong coupling weights.

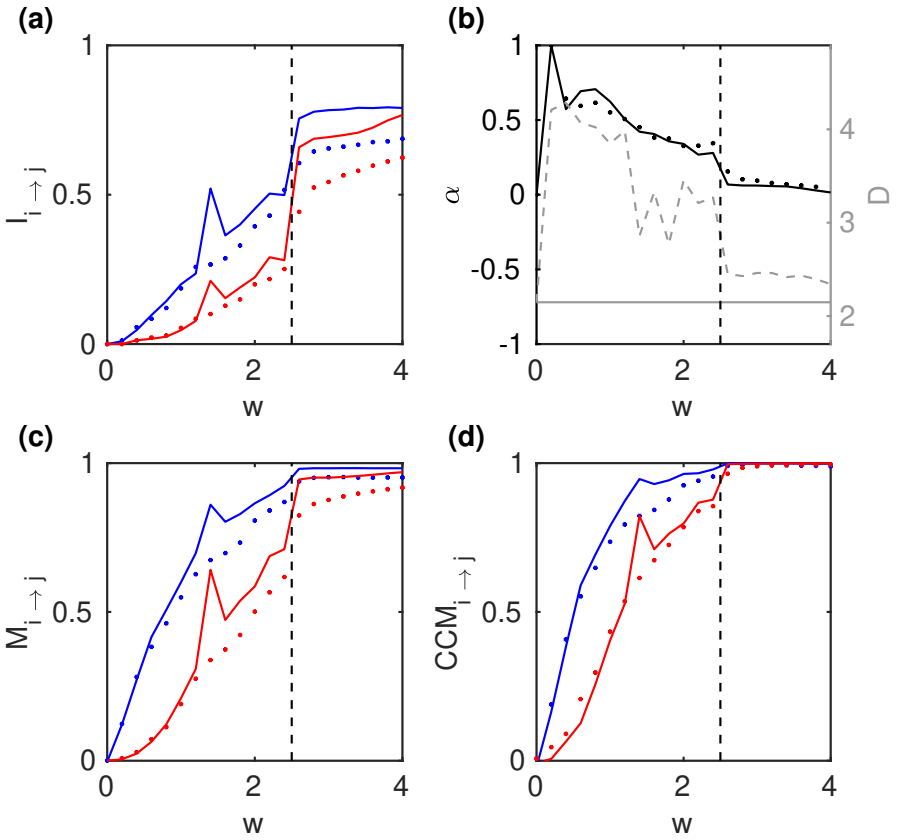


Figure 38: **(a)** Causal influence for the (unilaterally) coupled Lorenz → Rössler system derived from 10^3 reference points chosen from time series of length $3 \cdot 10^4$. The time series were embedded with $\tau = 2$ and $m = 7$ and $E = 10^3$ ensembles were used to derive a chance-level. Solid lines show the noise free results and dotted lines the noise perturbed results. The causal influence in direction of the coupling $I_{x\to y}$ is shown in blue, the reverse direction in red. **(b)** The asymmetry α for the noise free (solid lines) and perturbed system (dotted lines). Grey Lines show the Information Dimension estimated from the respective time series, solid lines being the Lorenz-system and dashed line the Rössler-system. **(c)** Causal Influence (M-Method) $I_{x\to y}$ (blue) and $I_{y\to x}$ (red) for the noise free (solid-lines) and for the systems perturbed by Gaussian white noise (asterisks). **(d)** Causal Influence (Convergent Cross Mapping) $I_{x\to y}$ (blue) and $I_{y\to x}$ (red) for the noise free (solid-lines) and for the systems perturbed by Gaussian white noise (asterisks).

5.4.1 *Asymmetric Noise*

A particular challenging case is a unilaterally coupled system influenced by asymmetric additive external noise. Due to the strong asymmetry in noise and coupling weights, it is rather difficult to infer the correct direction of *causal influence*. As an example for such a system Chicharro et al. [67] used two coupled Lorenz-oscillators (Equa-

tion 26) with the unilateral coupling $\Omega_{1\to2} = w_{1\to2} =: w$. This system synchronizes for strong coupling weights $w = w_{i\to j} \approx 9.8$. Since the asymmetry index α is volatile for small coupling weights, we consider only the enumerator $\Delta I = I_{y\to x} - I_{x\to y}$ instead of the full asymmetry index α. Analogous to Chicharro et al. [67] 100 ensembles were generated for each coupling weight between 0.05 and $0.05 \cdot 1.05^{121}$ and we show the estimated average causal asymmetry in Figure 39 **(a)**.

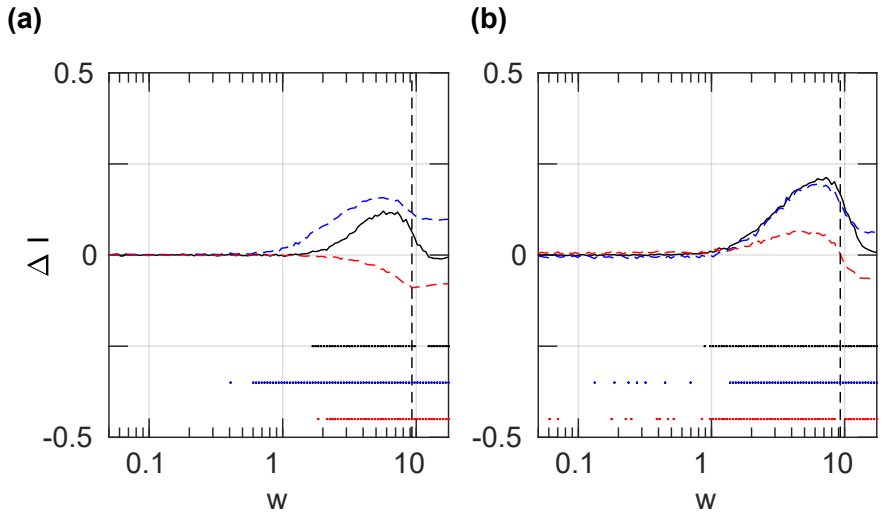


Figure 39: **(a)** ΔI derived for a system of two coupled Lorenz oscillators ($\theta_{1/2} = 10, \rho_1 = 39, \rho_2 = 35, \theta_{1/2} = 8/3$, $\Omega_{1\to2} = w_{1\to2}(x_1 - x_2)$, $\Omega_{2\to1} = 0$). The unilateral coupling weight $w_{x\to y}$ is increasing from 0 to 17. A fourth-order Runge-Kutta algorithm with step size 0.005 and down-sampled in intervals of 0.03 time units integrated the system. After removing the transient $N = 2^{11}$ data points are used for further analysis and are injected with additive noise either in x_1, x_2 or kept free of noise. We use Gaussian white noise with amplitude of $\sigma = 0.95\sigma_{x_i}$ and all time-series are embedded with $m = 8$ and $\tau = 4$. The solid black line show the noise free ΔI, dashed blue lines show perturbations in x_1 and dashed red lines in x_2. All results are averaged over 100 ensembles for each coupling weight and noise condition. **(b)** shows corresponding results using a moving average filter with a window of 8 time steps before further analysis. Dots mark ΔI with a significant difference from 0 derived by a Wilcoxon signed rank test ($p = 0.001$).

ΔI maintains a positive sign, reflecting the unilateral structure of the Lorenz $\to$ Lorenz-system. However, after synchronization (vertical dashed line) ΔI decreases reflecting the entrainment of the second Lorenz-Subsystem and the resulting similarity of both subsystems. To validate that ΔI is different from 0 a Wilcoxon signed-rank test is used, black dots below mark significant differences from 0. The colored lines show ΔI for asymmetric noise added to the driven Lorenz (blue) or the driving Lorenz (red). If the affecting subsystem (blue) is perturbed, the correct sign of ΔI is still detected. Furthermore for

large weights - synchronization - the unilateral system structure is better detected. In contrast, for noise in the affected subsystem ΔI has the wrong sign for a large range of weights and thus masks the underlying causal structure. As seen above in Figure 39 causality detection in synchronizing system under the influence of strong noise can lead to misleading results. This is not only true for CPM, but for a wide range of methods - in the case of unidirectional noise, methods either suffer if the affecting or the affected system part is perturbed by noise Figure 40. In general noise destroys the neighborhood relations and thus methods using such wrong neighbourhoods can yield misleading results. As discussed in the previous section to improve the detection of *causal influences* the original neighbourhood relations have to be better preserved or recovered in the presence of noise.

5.4.2 *Comparison*

Previously, Chicharro et al. [67] compared different methods for causality detection for the Lorenz→Lorenz system. Here, we also show the corresponding results for Figure 39 for three different methods. Topological Causality [1] in Figure 40, **(a)** & **(d)**, the M-method from Chicharro et. al., **(b)** & **(e)**, and Sugihara's CCM method [4] in **(c)** & **(f)**.

Black (solid) lines correspond to a noise-free system. Here, all methods maintain a positive sign, reflecting the unilateral structure of the Lorenz → Lorenz-system. If noise is injected in only one part of the system all methods struggle to different degrees. For noise in the affecting system (blue dashed lines) M retains a positive sign and only for strong coupling (after synchronisation) M and CCM change the sign indicating the detection of the wrong coupling structure. TC fails completely to detect the correct direction. For noise injected in the affected system TC and M detect the correct direction, i.e. a positive sign. CCM detects a wrong direction for weak coupling weights. In **(d)-(f)** the moving average filter was also used to improve the neighborhood. Similar as to CPM also M and CCM now better retain a positive sign and thus the correct causal structure, if the improved neighborhoods are used. In contrast, TC still detects the wrong direction of *causal influence*, albeit the results are improved.

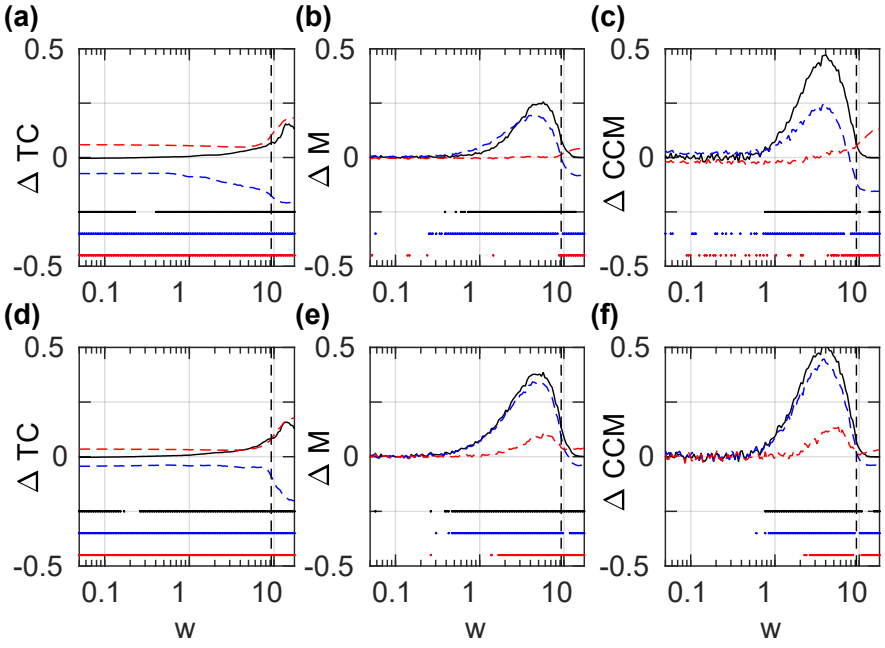


Figure 40: **(a)**ΔTC, **(b)** ΔM, **(c)** ΔCCM derived for a system of two coupled Lorenz oscillators ($\theta_{1/2} = 10, \rho_1 = 39, \rho_2 = 35, \theta_{1/2} = 8/3$, $\Omega_{1\to2} = w_{1\to2}(x_1 - x_2)$, $\Omega_{2\to1} = 0$). The unilateral coupling weight $w_{x\to y}$ is increasing from 0 to 17. A fourth-order Runge-Kutta algorithm with step size 0.005 and down-sampled in intervals of 0.03 time units integrated the system. After removing the transient $N = 2^{11}$ data points are used for further analysis and are injected with either noise in x_1, x_2 or kept free of noise. The injected noise is Gaussian white noise with amplitude of $\sigma = 0.95\sigma_{x_i}$ and all time-series are embedded with $m = 8$ and $\tau = 4$. The solid black line show the noise free ΔI, dashed blue lines show perturbations in x_1 and dashed red lines in x_2. All results are averaged over 100 Ensembles for each coupling weight and noise condition. **(d)**, **(e)**, **(f)** show corresponding results using a moving average filter with a window of 8 time steps before before further analysis. Dots mark a significant positive or negative ΔX derived by a Wilcoxon signed rank test ($p = 0.001$).

5.5 EXPERIMENTAL MEASUREMENTS

Finally, the viability of CPM is demonstrated for a selection of experimental results. The results presented here were already discussed in the context of other methods for causal inference, e.g. CCM. Thus the focus will be an reproducing previous results. The first two examples are (pairwise) Heart- & breath-rate data [12] and climate data studied in [76]. The third example, EEG-data, is already discussed in the supplement of TC [1].

5.5.1 *Heart- & Breath-rate*

The first set of experimental data is the bi-variate data Set B of the Santa Fe Time Series Competition [77] and features heart rate and breathing rate of a patient with sleep apnea. Schreiber [12] used Transfer Entropy to identify an asymmetry with the dominant direction of *causal influence* from heart- to breath-rate.

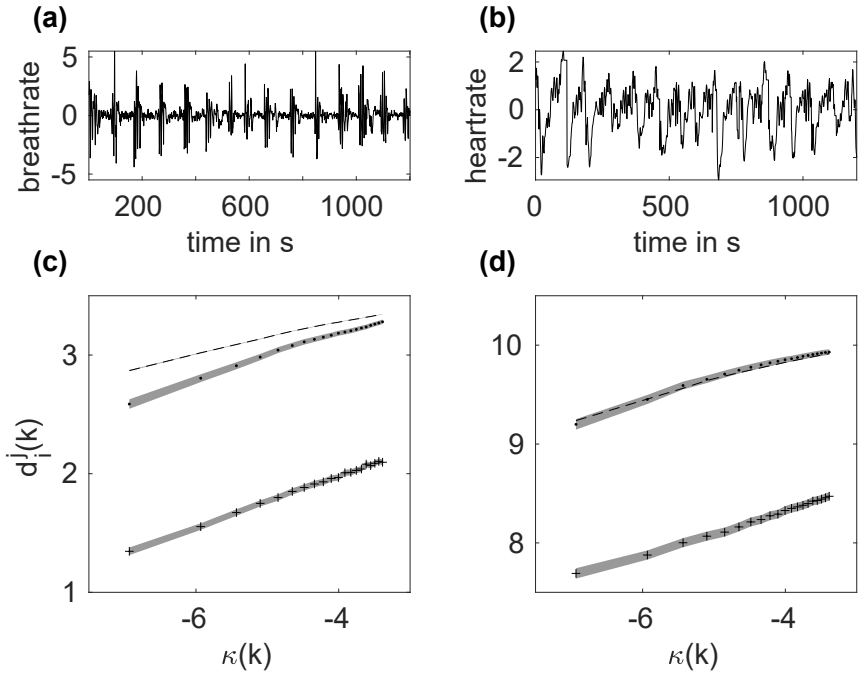


Figure 41: Heart-rate **(a)** and breath-rate **(b)** of a sleeping apnea patient. **(c)**, **(d)** show logarithmic neighborhood sizes for the heart and breathing rate. The data was sampled at 2Hz and embedded with $\tau = 6$ and $m = 5$, all $n = 601$ data points were chosen as reference points. **(c)** $d_H^H(k)$ shown as solid line, $d_H^B(k)$ shown as dotted line and the respective chance-level $d_H^{B^*}(k)$ (dashed line). Due to the small amount of data only the threefold standard deviation was used to determine the chance-level. **(d)** $d_B^B(k)$ shown as solid line, $d_B^H(k)$ shown as dotted line and the respective chance-level $d_B^{H^*}(k)$ (dashed line).

In agreement with Schreiber [12] the analysis is done on the same section of $n = 601$ data points that where sampled at a rate of 2Hz, shown in Figure 41 **(a)** & **(b)**. Schreiber concluded that the dominant direction of influence is from heart to breath signal and agreeing 'with the observation that the patient breathes in bursts which seem to occur whenever the heart rate crosses some threshold'. Using CPM this asymmetry is even more evident in the logarithmic neighborhood size $d_i^j(k)$ (Figure 41 **(c)**, **(d)**) and the resulting *causal influence*. Only a significant *causal influence* from heart-rate to breathrate $I_{H\to B} = 0.07$ is detected. This is also consistent with results us-

ing nonlinear Granger Causality [78] and also remains valid if the sampling rate is changed, e.g. doubled ($I_{H\to B} = 0.06$) or halved ($I_{H\to B} = 0.07$).

5.5.2 CO_2, CH_4 *and Temperature*

The second set of time series stems from climate research, for which van Nes used CCM to infer *causal influences* [76]. The goal of the publication was the validation of feedback effects between temperature T and climate gases, CO_2 and CH_4.

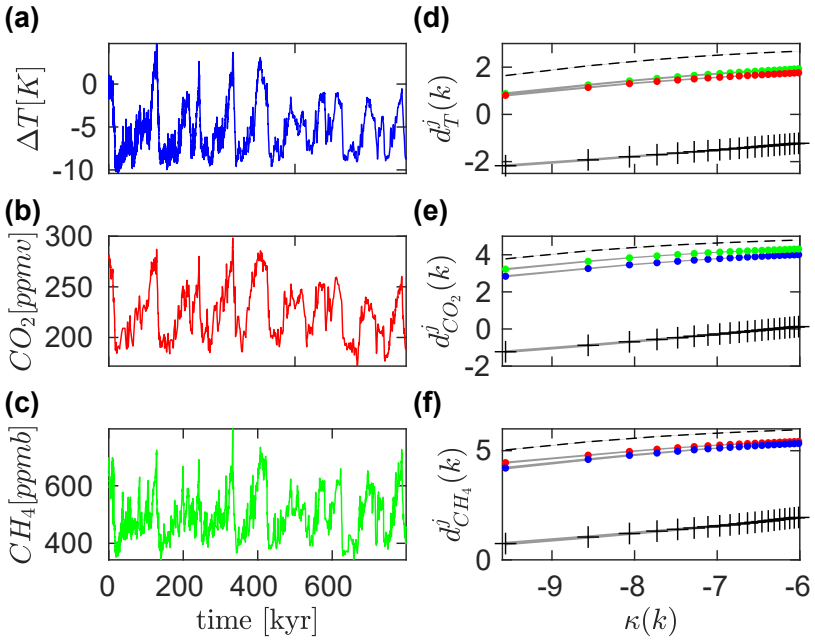


Figure 42: Time-series of T **(a)**, CO_2 **(b)** and CH_4 **(c)** interpolated at a sampling rate of 100 years [79]. **(d)** Logarithmic neighbourhood sizes $d^j_i(k)$ over κ. The interpolated time-series were embedded with $m = 5$ and $\tau = 100$kyr and 10^2 ensembles were used for chance-level-estimation. $d^T_T(k)$ shown as plus signs and the respective chance-level $d^*_T(k)$ as a dashed line. Colored dots show $d^i_T(k)$, the color is corresponding to T (blue), CO_2 (red) and CH_4 (green). **(e)** & **(f)** show analogue results for $C0_2$ and CH_4.

All three time-series are measurements obtained from Ice-cores. The original data were published in [80] (T), [81] (CO_2) and [82] (CH_4). Due to the nature of the recorded data additional preprocessing has to be performed. Measurement of the observables is conducted in different depth of Ice-cores: Downcore parts of the Ice-core correspond to times longer ago. However, the sampling intervals are non regular, more recent times are sampled in shorter intervals than past times. Furthermore the measurements are conducted on different Ice-cores with differing sampling steps.

Remedying this, all three time-series are interpolated with a sampling rate of 100 years [79], shown in Figure 42. Consistent results (with the following results) are also obtained for different sampling rates or binning the data into bins of equal size containing at least one data point. Furthermore, the temperature is not directly measured, but inferred from the deuterium content δD. For more details on the prepocessing refer to the aforementioned references [76, 79–82].

The different choice of chance-level only effects the first ≈ 5 neighbors slightly, qualitative the results would remain the same.

In contrast to previous results random neighbors were chosen for the chance-level. This allows the concurrent display of neighbors searched in the two other reconstructions, e.g. $d_T^{CO_2}(k)$ (red) and $d_T^{CH_4}(k)$ (green) are shown in **(d)**. The concurrent display allows to already infer which connection is stronger from inspecting the $d_i^j(k)$ in comparison to one another, e.g. CO_2 is stronger effecting T than CH_4. In this case the *causal influences* are somewhat equal, thus the difference in $d_i^j(k)$ is also not large.

Originally, this was done to save space.

The *causal influences* resulting from the shown (logarithmic) neighborhood size are shown in Table 2. In agreement with [76], both, influences from temperature to CO_2 and CH_4 and in the reverse direction are detected. Like CCM (in [76]), CPM is detecting a somewhat stronger outgoing influence from T compared to the feedback connection.

from \ to	T	CO2	CH4
T	1	0.2222	.1933
CO2	0.1715	1	0.1089
CH4	0.1560	0.1324	1

Table 2: Causal influences $I_{i\to j}$ between T, CO_2 and CH_4.

5.5.3 *Conclusion*

This chapter introduced CPM as measure for expansive distortions and thus *causal influences*. The measure can be directly computed from the neighbor relations resulting in a more resilient and straightforward advancement of the concept of *topological causality*.

The following pages discussed the impact of systems parameters like coupling weights, amount of available data and noise on CPM using a range of different dynamical systems as examples. A particular challenging set of systems was found in systems approaching synchronization. Here, CPM was still able to infer correct dominant directions of influences, at least up to synchronization. The point of synchronization, measured by information dimension, can be directly derived from the basic quantities of CPM.

Finally, *causal influences* are inferred in a selection of experimentally measured time-series. Here, the results of CPM agree with previously used approaches.

6

SYNTHESIS

6.1 SUMMARY OF RESEARCH FINDINGS

Interconnected (sub)-systems follow a shared overall dynamic. Here, observables from a single variable allow the time-delay reconstruction of the complete state space, using Takens theorem [14–17]. These reconstructions contain information about *causal influences* and the underlying structure of connections. In particular, expansive distortions in time-delay state space reconstructions systematically reflect effective, time-resolved influences among parts of non-separable deterministic system. These expansive distortions are reflected in the mapping between state space reconstructions and also in the inter neighborhood relations.

In Chapter 4, *topological causality* (TC) [1] is introduced as a well founded and analytically tractable method. TC measures the expansive properties from the mapping between state space reconstructions. The singular values of the mapping are inversely related to the *causal influence*. In doing so, not only an overall *global* value for inter system interaction, but rather the time-resolved *causal influence* can be computed. The estimation of local mappings can be (numerically) challenging, in particular, when only limited amounts of observations are available, when the system's strict determinism is violated (intrinsic noise), and when measurement noise contaminates the data. Even greater challenges pose synchronizing systems that can cause TC to reveal wrong directions of dominant influence.

A more refined approach to compute these expansive distortions in state space reconstructions, *Cross Projection method* (CPM) [2], is presented in Chapter 5. Here, the basic idea remains the same - expansive distortions are related to *causal influences*. But CPM derives these (the expansive properties) directly by projecting neighbor relations from the affected system(-part) to the respective other system(-part). Since only neighbor relations are transferred, the state space reconstruction from different observables are fully separated. This separation ensures invariance against different scales in the observables and thus no additional preprocessing, e.g. z-score or quantile transformation are required. The opportunity to directly compare projected neighborhoods and random neighborhoods grants a good criterion to prevent the detection of false positives. Furthermore, by not estimating the mapping between state space reconstructions the computational cost

is lowered and CPM can remain a local measure even in higher dimensions. By altering the computation of neighbor size, e.g. by adapting a different metric, CPM can be tailored to the requirements of the investigated data, increasing the reliability, in particular in cases of strong noise pollution. Furthermore, CPM retains important properties of TC, like the ability to infer time-resolved *causal influences*, while also giving further insights like the dimension of reconstructed system(-parts).

The mapping is, in general, a dimension × dimension matrix.

Or by looking at mean instead of maximum size of neighborhoods.

Besides the straightforward result of our two methods to infer *causal influence*, the application of said methods also provides additional insights: e.g. the effects of noise on methods of this kind, the handling of synchronizing systems and the potential to use noise as an aid when dealing with synchronization.

Measurement noise not only affects intermediate quantities, like neighborhood size, state-space mapping or prediction error, but already the location of neighbors in state space reconstructions. The composition of neighborhoods is changed by removing correct nearest neighbors and adding false nearest. Improving the neighborhood relations directly enhances the quality of *causal influences* and can help compensate for noise. Similar problems affect also other methods based on the same heuristics.

Intermediate in the sense, that these quantities are computed between neighbor search and causal influence.

Originally, TC was overhauled to achieve reliable results in systems of (simple) coupled oscillators, e.g. Fitzhugh-Nagumos-neurons. A wide range of coupling weights and/or frequencies force (these) oscillators to synchronize and not retain any own intrinsic degrees of freedom. In such cases *causal influences* are meaningless. We found the collapse of the dimension of the driven system onto the dimension of the driver is a good criterion to quantify at which point the systems are fully synchronized. Up to this transition CPM is still able to determine the correct direction of *causal influence*. Afterwards, meaningful *causal influences* can only be derived if the synchronization is broken up, e.g. by transient dynamics or noise. While small amounts of noise prove helpful, larger amounts distort the state space reconstructions and mask the underlying influences even for unilateral interactions.

Point in parameter space.

Despite the negatives, a great advantage of TC and CPM is the inference of even time-resolved *causal influences*. While the inference of the structure of interactions is insightful, a time resolved an thus time-dependent measure allows to also recognize opportune moments or system states to interact with and thus provides a more insightful view into unknown systems.

6.2 INTEGRATION IN PREVIOUS RESEARCH

The introduction already established the kind of system or observed data which have been of interest in this work: Time-series generated or observed from experiments are analyzed retroactive. Thus, methods based on intervention, like Judea Pearl's do-operator [6], are not suitable in this case. Furthermore, we are not looking to acquire correlations between observables that require additional information to establish directions of *causal influences*, but are rather seeking a directed measure for *causal influences* solely based on the observables.

Observed time series are sorted according to their time-index t and linked to each other by t. Comparing the prediction of the future using own predecessors and predecessors of other time-series yields a (first) measure for *causal influences* - Wiener-Granger causality [3]. Generalization of Wiener-Granger causality link this concept more directly to information theory [12].

In general, observed time-series are represented as one-dimensional vectors, but the underlying dynamical system is often multidimensional. The states of the underlying system can be reconstructed using Takens theorem [14]. It is clear, if the (reconstructed) state $\mathbf{r}^i(t)$ describes the system at time t, the future of similar states $\mathbf{r}^i(\tilde{t}+1)$ are related to the future of $\mathbf{r}^i(t)$. In particular, the prediction of these future states can be improved using these close neighbors, yielding a state-space method relatable to Wiener-Granger causality, proposed by Sugihara et al. [4]. Trivially, this works rather well in (pseudo-)oscillating systems, where close states to a reference point exist. Sugihara's *convergent cross mapping* and other methods demonstrate that *causal influences* can be inferred from state space reconstructions.

Similar meaning close-by in the state space reconstruction.

The fundamental difference of our methods [1, 2] and related approaches [64, 67] is the inference of *causal influences* from just properties of state space reconstructions.

Difference regarding WGC, CMM and related methods.

On of our insights is that precisely the expansive distortions in state space reconstructions reflect the *causal influences*. This idea is introduced as the the concept of *Topological causality* (TC) by Harnack, Laminski, Schünemann and Pawelzik [1]. In (simple) systems, a direct relation between TC and the basic properties of system interactions - coupling weights and system state - can be derived analytically. The time-resolved nature of TC allows a more detailed view into the interaction structure. In particular, changes in the interactions between systems due to different states are revealed by TC.

time-resolved, also implies state-resolved

Our second approach, *cross projection method* (CPM) [2] is strongly linked to TC, and even has a direct functional relation to TC. However, not quantifying expansive distortions from the mapping between state space reconstruction, but from neighbor relations has clear advan-

For coupled logistic maps.

tages over TC, i.e. higher robustness to noise or lower computational cost. Furthermore, by just transferring neighbor relations to the respective other space, state space reconstructions from different observables are fully separated. Thus, the comparison of neighborhood sizes and our resulting *causal influence* is invariant against observables of different scales.

Quantifying *causal influences* from neighbor relations/distances in state space reconstructions was already proposed by Chicharro et al. [64, 67]. In general, CPM and the methods of Chicharro et al. differ in the details of how to quantify neighborhood size and computation of the resulting measure for *causal influences*. However, CPM also retains additional properties. By considering different numbers of neighbors the logarithmic neighborhood sizes derived by CPM are directly associated with the Information dimension, allowing an additional criterion for the inference of interactions [66]. Last but not least, CPM retains the same time-resolved properties as TC. To be fair, the methods of Chicharro et al. can refined too in this regard.

When listing a selection of methods to infer *causal influences* it would be straightforward to pose the question: 'What is the method of choice?' In my perspective there is no 'all-in-one' approach. The best method for a specific application depend on the properties of the observable in question. For example, in systems with strong stochastic properties the reconstruction of the state space can be challenging. Thus, WGC might be preferable to our proposed approaches. In general, a reasonable approach is to use several different methods. This also agrees with ongoing research, e.g. in [83] the authors discuss key strength and limitations of several state of the art methods. Ambiguous results from different methods can offer deeper insights into limitations of different methods and how to overcome these limitations.

6.3 LIMITATIONS OF TC AND CPM

Roughly speaking, the limitations of the proposed method(s), TC and CPM, can be divided into requirements for the considered observables and conceptual limitations.

Some general requirements for the investigated observables are discussed in Chapter 5, focusing on quantities like length of time series and noise-level and their effect on CPM. Other non-discussed requirements are more specific properties for the observable, like stationarity. In general, we require the observables to have a valid state space reconstruction due to Takens theorem or its extension [14–17]. For the shown dynamical systems this is guaranteed, but for experimental results this does not have to be true.

Actually, the z-component of the Lorenz system is also an exception.

In particular, for limited available data weak interactions might prohibit information transfer from far apart system components to be included in the state space reconstruction of a distant observable. The amount of required data, naturally, also depends on the dimension of the underlying system and the time-delay τ used for reconstruction. Higher dimensional systems need more available data to sample the state space reconstruction (somewhat) densely and provide sufficient neighbors around a reference point. For example, a single period of a limit cycle is not sufficiently sampled, since all neighbors are unsuitable temporal neighbors.

far apart in the structure of the system

Avoiding temporal correlation, we exclude these temporal neighbors from analysis.

In general, simple limit cycles of oscillating systems can be challenging for the inference of *causal influences*. If only a few oscillations periods from observables with similar frequencies are available, these systems can appear to have strong *causal influences* or even seem synchronous - even if the parts of the system are not connected whatsoever. I would treat these systems with caution to not predict spurious *causal influence* despite of limited available data, but rather due to limited data. Other methods, like convergent cross mapping [4] seem to not be affected by limited available data. Although, convergent cross mapping has problems dealing with synchronization and was found to yield wrong directions of dominant influences with in synchronous systems [60–62].

More interesting, are fundamental properties of the investigated systems that deteriorate the inference of *causal influences*. Trivially, small interactions are hard to detect and/or might be non-significant due to the available amount of data. The other limit case, strong coupling weights and the resulting synchronization of (sub-)systems is the most challenging (for our methods), but also offers the opportunity for further insights. Before synchronization we are able to make meaningful statements on direction and strength of *causal influences*. Once, both system parts are synchronized this is no longer possible, not even for unilaterally coupled system. Whether synchronization is achieved depends on the investigated system and the parameters, e.g. coupling weights and oscillation frequency.

For example, in the considered coupled Fitzhugh-Nagumo neurons a faster driving system leads to earlier synchronization for smaller coupling weights. Even for no interaction the system parts can appear synchronized, especially if only short time-series of (sub-)systems with similar frequency are considered. We found the only way to deal with synchronized systems is to artificially break the synchronization. For example, by considering transient dynamics where subsystems are still not converged to the limit-cycle or the introduction of additional noise allowing for degrees of freedom than can be transmitted via the coupling weights.

Transient dynamics can be induced, e.g. by shot-noise.

6.4 THEORETICAL IMPLICATIONS

The fundamental insight of TC [1] is that expansive distortions between state space reconstructions are reflecting *causal influences*. These expansive properties can derived from mappings between state space reconstruction (TC) or neighbor relations within state space reconstructions (CPM) and are direct consequence of properties of the reconstruction.

In the past a related idea was used for determining the quality of mappings [41, 84], where the ratio of the inflations *within* the respective spaces was expected to be close to one for homeomorphy while for topology violations it systematically deviates from one. The concept of expansive distortions also directly links to the prediction quality derived by Sugihara's convergent cross mapping [4]. Naturally, close neighbors in state space reconstructions have similar states and provide a good base for prediction.

When system dynamics are too similar or even synchronous, we cannot infer causal relations, if the respective reconstructions have no unique degrees of freedom. This also agrees with results from different interacting dynamical system, e.g. a Roessler and a Lorenz system. Since these subsystems are so different in their respective structure even for strong coupling some own degrees of freedom are preserved. Thus, even for strong coupling weights CPM is able to infer at least the correct dominant direction of *causal influence*. The information also reflects this by collapsing (after synchronization) to the dimension of the driving system, but still retains different dimensions in both subsystems.

In identical (coupled) systems that synchronize with increasing coupling weights, we found the collapse of the information dimension to provide a suitable criterion quantifying the point at which systems parts are synchronous. Up to this point CPM (and selection of other methods [4, 67]) are able to infer the correct direction of influence. Asymmetrically injected noise to either driver or driven system lead to the detection of the wrong direction influence. Interestingly, the different methods are failing for different of the two noise scenarios. CPM and the methods proposed by Chicharro et al. [67] share the same basic quantities - local neighborhood sizes, and only differ in their further usage. We suspect detailed insights in the difference of these methods yields the possibility for a combined approach that copes with both, noise in the driver and the driven system.

6.5 DIRECT EXTENSIONS OF TC AND CPM

In Chapter 5 some additions and modifications for our proposed frameworks were implied covering more recent ideas. In particular, the choice of correct and non-temporally related neighbors is presented. These approaches already show (some) success, but further improvements can be made.

For example, the discussed dynamical systems show specific directions in state space that are affected by changes in coupling weights. Fitzhugh-Nagumo neurons show an expansion orthogonal to the temporal evolution of the system, while logistic maps are stretched (with increasing coupling weight) orthogonally in regards to the manifold of the uncoupled system. It seems, in both cases, a selection of neighbors along these directions will have a more direct reflection of coupling weights (and thus *causal influences*). Simple approaches to select this subset of neighbors can be achieved using principal component analysis or related methods select neighbors only in specific direction (in state space reconstructions).

Both, TC and CPM, rely on local expansions, reflected in small local neighborhoods. It seems different system scales are affected by noise to differing degrees. Thus, it might prove beneficial to investigate larger and/or different neighborhood sizes, if these still reflect *causal influences*. Varying the scale at which a system is viewed also ties in nicely with the concept of causal emergence [85] suggesting that cause and effect must not necessarily manifest at the same scale.

In fact, considering local scaling already enabled us to derive the information dimension directly from neighborhood sizes. Already, this dimension can gives insight in the interactions between system parts [66], this binary criterion for *causal influence* can be more directly incorporated in our framework.

All analysis done in this work is done as a pairwise causal analysis using the same parameters for both observed time-series. However, already the case of two oscillators with different frequencies suggests different time-delays τ will be beneficial, since the best reconstructions of the state space of (pseudo-)oscillating systems is achieved using one fourth of the period as time-delay. Furthermore, one of the perks of CPM is the complete separation of state space reconstructions - only information of neighborhoods is transferred. Thus, it is straightforward to use different schemes for state space reconstruction. For example, the combination of two or more observed time-series. This also indicates a possible generalization of CPM using already established ideas like *causal influences* conditioned on other vari-

ables or combinations of different observables for better state space reconstruction.

In my opinion, the most astonishing property of CPM (and TC) is the time- and state-resolved inference of *causal influences* which also offers additional applications. For example this could uncover time-dependence of coupling weights or the on/off-switching for interactions.

Finally, the ending sections of Chapter 5 discussed some examples of real measured data. This section was limited and just a reproduction of examples, already analysed by previous methods. More interesting would be examples where other methods fail and we prevail.

6.6 BROADER ISSUES TO BE COVERED IN FUTURE WORK

Just like the proposed frameworks, TC and CPM, other methods, especially methods based on state space reconstruction, suffer from the same drawbacks and offer similar improvements. Thus, our ideas to improve neighborhoods should also prove beneficial for CCM and other methods alike.

State space methods also have the drawback of operating in the time-domain. Methods like Wiener-Granger-Causality can also be transformed to the frequency-domain. This is especially advantageous for experimental results with many short observed time-series. Here, approaches like multi-tapering allow the combination of the short observables into a good representation of the frequency domain, while the state space methods have to rely on the short single representations of state-space.

Dealing with state space reconstructions a sufficiently dense sampling of the complete state space is needed. In particular, systems with high dimensional dynamics and/or few available data points represent the state space only incompletely. For example, CCM was shown to yield wrong results in synchronous systems [13, 61]. Also the shown results of short time-series of the 10-species Lotka-Volterra system (Figure 9) shows in some simulations the inference of false positive *causal influences*. A valid criterion that the density of the state space is sufficiently full will be helpful in this case and prevent false positives.

In general, the field is missing a meaningful criterion which method is to be chosen for different available time-series and their respective properties. In particular, noise, synchronization and/or oscillations are detrimental for a range of methods. Although, at least for a selection of methods [83] provides a guide on drawbacks, advantages and future challenges.

7

CONCLUSION

The simple **cause** and **effect** notion of causality is at odds with reciprocal interactions in dynamical system. Feedback loops entangle **cause** and **effect** resulting in observed time-series that are both, **cause** and **effect**. Measures to infer the underlying or hidden interactions from these observables quantify the effective *causal influence* [1–4, 12, 64].

The dynamic evolution of interconnected systems cannot be separated, the system behaves as a whole. Takens theorem [14] allows the reconstruction of the state space of this overall dynamic from a single measured observable. State space reconstruction from different observable allow to infer the *causal influence* from local expansive distortions.

Topological causality (TC) [1] computes these expansive distortions from mappings between state space reconstructions of different observables, resulting in a state-dependent effective measure for *causal influences* among parts of non-separable systems. Furthermore, TC is a model-free approach to infer *causal influences* that is also analytically tractable in simple examples.

The revised successor of TC is the *cross projection method* (CPM) [2]. CPM derives expansive distortions by projecting neighbor relations between state space reconstructions and comparing the size of these projections with domestic neighborhoods. Basing the comparison on the metric of a single state space reconstructions yields a scale invariant, robust approach to quantify expansive distortions.

Both, TC and CPM, have drawbacks, in particular, in systems prone to synchronization. Here, only CPM is able to infer correct *causal influences* up to synchronization. However, the lack of unique information in the individual state space reconstructions of a fully synchronous systems also corresponds to no expansive distortions. Systems lacking any unique degrees of freedom mark reasonable upper bound for the applicability for methods inferring *causal influences*.

Remarkably, TC and CPM allow to infer state- and time-resolved *causal influences* for a range of (deterministic) dynamical systems. Thus, not only revealing the underlying interaction structure, but also giving insight into system states that are opportune moments for external interventions.

BIBLIOGRAPHY

[1] Daniel Harnack, Erik Laminski, Maik Schünemann, and Klaus Richard Pawelzik. "Topological Causality in Dynamical Systems." In: *Phys. Rev. Lett.* 119 (9 2017), p. 098301. DOI: 10.1103/PhysRevLett.119.098301. URL: https://link.aps.org/doi/10.1103/PhysRevLett.119.098301.

[2] Erik Laminski and Klaus R. Pawelzik. "Reliable detection of causal asymmetries in dynamical systems." In: *Phys. Rev. E* 107 (1 Jan. 2023), p. 014214. DOI: 10.1103/PhysRevE.107.014214. URL: https://link.aps.org/doi/10.1103/PhysRevE.107.014214.

[3] C. W. J. Granger. "Investigating Causal Relations by Econometric Models and Cross-spectral Methods." In: *Econometrica* 37.3 (1969), pp. 424–438. ISSN: 00129682, 14680262. URL: http://www.jstor.org/stable/1912791.

[4] George Sugihara, Robert M May, H Ye, C Hsieh, E R Deyle, and M Fogarty. "Detecting Causality in Complex Ecosystems." In: *Science* 334 (2012), pp. 496–500.

[5] Aristotle. "Chapter 2." In: *Metaphysics* (350 BC).

[6] Judea Pearl. *Causality. Models, Reasoning, and Inference*. 2nd ed. Cambridge, UK: Cambridge University Press, 2009. ISBN: 978-0-521-89560-6. DOI: 10.1017/CBO9780511803161.

[7] Joseph Lee Rodgers and W. Alan Nicewander. "Thirteen Ways to Look at the Correlation Coefficient." In: *The American Statistician* 42.1 (1988), pp. 59–66. DOI: 10.1080/00031305.1988.10475524. eprint: https://doi.org/10.1080/00031305.1988.10475524. URL: https://doi.org/10.1080/00031305.1988.10475524.

[8] Auguste Bravais. *Analyse mathématique sur les probabilités des erreurs de situation d'un point*. Impr. Royale, 1844.

[9] Israel Cohen, Yiteng Huang, Jingdong Chen, Jacob Benesty, Jacob Benesty, Jingdong Chen, Yiteng Huang, and Israel Cohen. "Pearson correlation coefficient." In: *Noise reduction in speech processing* (2009), pp. 1–4.

[10] Charles Spearman. "The proof and measurement of association between two things." In: (1961).

[11] C.W.J. Granger. "Testing for causality: A personal viewpoint." In: *Journal of Economic Dynamics and Control* 2 (1980), pp. 329–352. ISSN: 0165-1889. DOI: https://doi.org/10.1016/0165-1889(80)90069-X. URL: https://www.sciencedirect.com/science/article/pii/016518898090069X.

[12] Thomas Schreiber. "Measuring Information Transfer." In: *Phys. Rev. Lett.* 85 (2 2000), pp. 461–464. DOI: 10.1103/PhysRevLett.85.461. URL: https://link.aps.org/doi/10.1103/PhysRevLett.85.461.

[13] Hao Ye and George Sugihara. "Information leverage in interconnected ecosystems: Overcoming the curse of dimensionality." In: *Science* 353.6302 (2016), pp. 922–925.

[14] Floris Takens. "Detecting strange attractors in turbulence." In: *Dynamical systems and turbulence, Warwick 1980*. Springer, 1981, pp. 366–381.

[15] Norman H Packard, James P Crutchfield, J Doyne Farmer, and Robert S Shaw. "Geometry from a time series." In: *Physical review letters* 45.9 (1980), p. 712.

[16] J. Stark, D.S. Broomhead, M.E. Davies, and J. Huke. "Takens embedding theorems for forced and stochastic systems." In: *Nonlinear Analysis: Theory, Methods & Applications* 30.8 (1997). Proceedings of the Second World Congress of Nonlinear Analysts, pp. 5303 –5314. ISSN: 0362-546X. DOI: https://doi.org/10.1016/S0362-546X(96)00149-6. URL: http://www.sciencedirect.com/science/article/pii/S0362546X96001496.

[17] Tim Sauer, James A Yorke, and Martin Casdagli. "Embedology." In: *Journal of statistical Physics* 65.3 (1991), pp. 579–616.

[18] Alfred J. Lotka. *Contribution to the Theory of Periodic Reactions*. Jan. 1909. DOI: 10.1021/j150111a004. URL: https://doi.org/10.1021/j150111a004.

[19] A. L. Hodgkin and A. F. Huxley. "A quantitative description of membrane current and its application to conduction and excitation in nerve." In: *The Journal of Physiology* 117.4 (1952), pp. 500–544. DOI: https://doi.org/10.1113/jphysiol.1952.sp004764. eprint: https://physoc.onlinelibrary.wiley.com/doi/pdf/10.1113/jphysiol.1952.sp004764. URL: https://physoc.onlinelibrary.wiley.com/doi/abs/10.1113/jphysiol.1952.sp004764.

[20] Henri Poincaré. *New methods of celestial mechanics*. Vol. 2. National Aeronautics and Space Administration, 1967.

[21] R May. "Simple mathematical models with very complicated dynamics. Nature, 261: 459-467." In: (1976).

[22] Edward N. Lorenz. "Deterministic Nonperiodic Flow." In: *Journal of Atmospheric Sciences* 20.2 (1963), pp. 130 –141. DOI: 10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2. URL: https://journals.ametsoc.org/view/journals/atsc/20/2/1520-0469_1963_020_0130_dnf_2_0_co_2.xml.

[23] Hermann von Helmholtz. *Vorlesungen über die Dynamik discreter Massenpunkte*. Vol. 1. Verlag von Johann Ambrosius Barth, 1898.

[24] J.G. Garnier and A. Quetelet. *Correspondance mathématique et physique*. Bd. 10. Impr. d'H. Vandekerckhove, 1838. URL: https://books.google.de/books?id=8GsEAAAAYAAJ.

[25] Marcos Lima, Fabio Claro, Wallace Ribeiro, Sergio Xavier, and Alejandro López-Castillo. "The Numerical Connection between Map and its Differential Equation: Logistic and Other Systems." In: *International Journal of Nonlinear Sciences and Numerical Simulation* 14 (Feb. 2013), pp. 77–85. DOI: 10.1515/ijnsns-2011-0032.

[26] Steven H Strogatz. *Nonlinear dynamics and chaos with student solutions manual: With applications to physics, biology, chemistry, and engineering*. CRC press, 2018.

[27] J.R. Dormand and P.J. Prince. "A family of embedded Runge-Kutta formulae." In: *Journal of Computational and Applied Mathematics* 6.1 (1980), pp. 19–26. ISSN: 0377-0427. DOI: https://doi.org/10.1016/0771-050X(80)90013-3. URL: https://www.sciencedirect.com/science/article/pii/0771050X80900133.

[28] Angelo Vulpiani, Fabio Cecconi, and Massimo Cencini. *Chaos: from simple models to complex systems*. Vol. 17. World Scientific, 2009.

[29] Robert L Devaney. *An introduction to chaotic dynamical systems*. CRC press, 2021.

[30] Otto E Rössler. "An equation for continuous chaos." In: *Physics Letters A* 57.5 (1976), pp. 397–398.

[31] Benoit Mandelbrot. "How long is the coast of Britain? Statistical self-similarity and fractional dimension." In: *science* 156.3775 (1967), pp. 636–638.

[32] Kenneth Falconer. *Fractal geometry: mathematical foundations and applications*. John Wiley & Sons, 2004.

[33] Alfréd Rényi. "On the dimension and entropy of probability distributions." In: *Acta Mathematica Academiae Scientiarum Hungarica* 10.1 (1959), pp. 193–215.

[34] Jaroslav Stark. "Delay embeddings for forced systems. I. Deterministic forcing." In: *Journal of Nonlinear Science* 9.3 (1999), pp. 255–332.

[35] Hassler Whitney. "The Self-Intersections of a Smooth n-Manifold in 2n-Space." In: *Annals of Mathematics* 45.2 (1944), pp. 220–246. ISSN: 0003486X. URL: http://www.jstor.org/stable/1969265 (visited on 08/17/2022).

[36] Dirk Aeyels. "On the number of samples necessary to achieve observability." In: *Systems & Control Letters* 1.2 (1981), pp. 92–94.

[37] Franz Hamilton, Tyrus Berry, and Timothy Sauer. "Kalman-Takens filtering in the presence of dynamical noise." In: *The European Physical Journal Special Topics* 226.15 (2017), pp. 3239–3250.

[38] Martin Casdagli, Stephen Eubank, J.Doyne Farmer, and John Gibson. "State space reconstruction in the presence of noise." In: *Physica D: Nonlinear Phenomena* 51.1 (1991), pp. 52–98. ISSN: 0167-2789. DOI: https://doi.org/10.1016/0167-2789(91)90222-U. URL: https://www.sciencedirect.com/science/article/pii/016727899190222U.

[39] Andrew M. Fraser and Harry L. Swinney. "Independent coordinates for strange attractors from mutual information." In: *Phys. Rev. A* 33 (2 1986), pp. 1134–1140. DOI: 10.1103/PhysRevA.33.1134. URL: https://link.aps.org/doi/10.1103/PhysRevA.33.1134.

[40] Matthew B. Kennel, Reggie Brown, and Henry D. I. Abarbanel. "Determining embedding dimension for phase-space reconstruction using a geometrical construction." In: *Phys. Rev. A* 45 (6 1992), pp. 3403–3411. DOI: 10.1103/PhysRevA.45.3403. URL: https://link.aps.org/doi/10.1103/PhysRevA.45.3403.

[41] Wolfgang Liebert, K Pawelzik, and HG Schuster. "Optimal embeddings of chaotic attractors from topological considerations." In: *EPL (Europhysics Letters)* 14.6 (1991), p. 521.

[42] Elizabeth Bradley and Holger Kantz. "Nonlinear time-series analysis revisited." In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 25.9 (2015), p. 097610. DOI: 10.1063/1.4917289. eprint: https://doi.org/10.1063/1.4917289. URL: https://doi.org/10.1063/1.4917289.

[43] Luca Faes, Giandomenico Nollo, and Alberto Porta. "Information-based detection of nonlinear Granger causality in multivariate processes via a nonuniform embedding technique." In: *Phys. Rev. E* 83 (5 2011), p. 051112. DOI: 10.1103/PhysRevE.83.051112. URL: https://link.aps.org/doi/10.1103/PhysRevE.83.051112.

[44] Min Han, Weijie Ren, Meiling Xu, and Tie Qiu. "Nonuniform State Space Reconstruction for Multivariate Chaotic Time Series." In: *IEEE Transactions on Cybernetics* 49.5 (2019), pp. 1885–1895. DOI: 10.1109/TCYB.2018.2816657.

[45] F. J. Anscombe. "Graphs in Statistical Analysis." In: *The American Statistician* 27.1 (1973), pp. 17–21. DOI: 10.1080/00031305.1973.10478966. eprint: https://www.tandfonline.com/doi/pdf/10.1080/00031305.1973.10478966. URL: https://www.tandfonline.com/doi/abs/10.1080/00031305.1973.10478966.

[46] Alberto Cairo. *The Functional Art: An introduction to information graphics and visualization*. New Riders, 2012.

[47] Norbert Wiener. "The theory of prediction." In: *Modern mathematics for engineers* (1956).

[48] A. Seth. "Granger causality." In: *Scholarpedia* 2.7 (2007). revision #127333, p. 1667. DOI: 10.4249/scholarpedia.1667.

[49] Yonghong Chen, Govindan Rangarajan, Jianfeng Feng, and Mingzhou Ding. "Analyzing multiple nonlinear time series with extended Granger causality." In: *Physics Letters A* 324.1 (2004), pp. 26 –35. ISSN: 0375-9601. DOI: http://dx.doi.org/10.1016/j.physleta.2004.02.032. URL: http://www.sciencedirect.com/science/article/pii/S0375960104002403.

[50] Wolfram Hesse, Eva Möller, Matthias Arnold, and Bärbel Schack. "The use of time-variant {EEG} Granger causality for inspecting directed interdependencies of neural assemblies." In: *Journal of Neuroscience Methods* 124.1 (2003), pp. 27 –44. ISSN: 0165-0270. DOI: http://dx.doi.org/10.1016/S0165-0270(02)00366-7. URL: http://www.sciencedirect.com/science/article/pii/S0165027002003667.

[51] Lionel Barnett and Anil K. Seth. "The {MVGC} multivariate Granger causality toolbox: A new approach to Granger-causal inference." In: *Journal of Neuroscience Methods* 223 (2014), pp. 50 –68. ISSN: 0165-0270. DOI: http://dx.doi.org/10.1016/j.jneumeth.2013.10.018. URL: http://www.sciencedirect.com/science/article/pii/S0165027013003701.

[52] Radhakrishnan Nagarajan and Meenakshi Upreti. "Granger Causality Analysis of Human Cell-Cycle Gene Expression Profiles." In: *Statistical Applications in Genetics and Molecular Biology* 9 (1 2010).

[53] Benjamin S. Cheng and Tin Wei Lai. "An investigation of co-integration and causality between energy consumption and economic activity in Taiwan." In: *Energy Economics* 19.4 (1997), pp. 435 –444. ISSN: 0140-9883. DOI: http://dx.doi.org/10.1016/S0140-9883(97)01023-2. URL: http://www.sciencedirect.com/science/article/pii/S0140988397010232.

[54] Ronald Newbold Bracewell and Ronald N Bracewell. *The Fourier transform and its applications*. Vol. 31999. McGraw-Hill New York, 1986.

[55] John Geweke. "Measurement of linear dependence and feedback between multiple time series." In: *Journal of the American statistical association* 77.378 (1982), pp. 304–313.

[56] Maciej Kamiński, Mingzhou Ding, Wilson A Truccolo, and Steven L Bressler. "Evaluating causal relations in neural systems: Granger causality, directed transfer function and statistical assessment of significance." In: *Biological cybernetics* 85 (2001), pp. 145–157.

[57] Claude E Shannon. "A mathematical theory of communication." In: *The Bell system technical journal* 27.3 (1948), pp. 379–423.

[58] Thomas M Cover and Joy A Thomas. "Elements of information theory 2nd edition." In: (2006).

[59] Anastasios A. Tsonis, Ethan R. Deyle, Hao Ye, and George Sugihara. "Convergent Cross Mapping: Theory and an Example." In: *Advances in Nonlinear Geosciences*. Ed. by Anastasios A. Tsonis. Cham: Springer International Publishing, 2018, pp. 587–600. ISBN: 978-3-319-58895-7. DOI: 10.1007/978-3-319-58895-7_27. URL: https://doi.org/10.1007/978-3-319-58895-7_27.

[60] Ethan R. Deyle, M. Cyrus Maher, Ryan D. Hernandez, Sanjay Basu, and George Sugihara. "Global environmental drivers of influenza." In: *Proceedings of the National Academy of Sciences* (2016). ISSN: 0027-8424. DOI: 10.1073/pnas.1607747113. eprint: https://www.pnas.org/content/early/2016/10/25/1607747113.full.pdf. URL: https://www.pnas.org/content/early/2016/10/25/1607747113.

[61] Edward B. Baskerville and Sarah Cobey. "Does influenza drive absolute humidity?" In: *Proceedings of the National Academy of Sciences* 114.12 (2017), E2270–E2271. ISSN: 0027-8424. DOI: 10.1073/pnas.1700369114. eprint: https://www.pnas.org/content/114/12/E2270.full.pdf. URL: https://www.pnas.org/content/114/12/E2270.

[62] Anna Krakovská and Filip Hanzely. "Testing for causality in reconstructed state spaces by an optimized mixed prediction method." In: *Phys. Rev. E* 94 (5 2016), p. 052203. DOI: 10.1103/PhysRevE.94.052203. URL: https://link.aps.org/doi/10.1103/PhysRevE.94.052203.

[63] Zsigmond Benkő, Ádám Zlatniczki, Dániel Fabó, András Sólyom, Loránd Erőss, András Telcs, and Zoltán Somogyvári. "Exact Inference of Causal Relations in Dynamical Systems." In: (2018). arXiv: 1808.10806 [q-bio.QM].

[64] Ralph G Andrzejak, Alexander Kraskov, Harald Stögbauer, Florian Mormann, and Thomas Kreuz. "Bivariate surrogate techniques: necessity, strengths, and caveats." In: *Physical review E* 68.6 (2003), p. 066202.

[65] R Quian Quiroga, A Kraskov, T Kreuz, and Peter Grassberger. "Performance of different synchronization measures in real data: a case study on electroencephalographic signals." In: *Physical Review E* 65.4 (2002), p. 041903.

[66] Bree. Cummins, Tomáš. Gedeon, and Kelly. Spendlove. "On the Efficacy of State Space Reconstruction Methods in Determining Causality." In: *SIAM Journal on Applied Dynamical Systems* 14.1 (2015), pp. 335–381. DOI: 10.1137/130946344. eprint: https://doi.org/10.1137/130946344. URL: https://doi.org/10.1137/130946344.

[67] Daniel Chicharro and Ralph G. Andrzejak. "Reliable detection of directional couplings using rank statistics." In: *Phys. Rev. E* 80 (2 2009), p. 026217. DOI: 10.1103/PhysRevE.80.026217. URL: https://link.aps.org/doi/10.1103/PhysRevE.80.026217.

[68] A Čenys, G Lasiene, and K Pyragas. "Estimation of interrelation between chaotic observables." In: *Physica D: Nonlinear Phenomena* 52.2-3 (1991), pp. 332–337.

[69] Jochen Arnhold, Peter Grassberger, Klaus Lehnertz, and Christian Erich Elger. "A robust method for detecting interdependences: application to intracranially recorded EEG." In: *Physica D: Nonlinear Phenomena* 134.4 (1999), pp. 419–430.

[70] Erik Laminski. "Master Thesis." In: (2017).

[71] Ronja Gronemeyer. "Bachelor Thesis." In: (2018).

[72] Peter Grassberger. "Finite sample corrections to entropy and dimension estimates." In: *Physics Letters A* 128.6 (1988), pp. 369 –373. ISSN: 0375-9601. DOI: https://doi.org/10.1016/0375-9601(88)90193-4. URL: http://www.sciencedirect.com/science/article/pii/0375960188901934.

[73] Frédéric Barraquand, Coralie Picoche, Matteo Detto, and Florian Hartig. "Inferring species interactions using Granger causality and convergent cross mapping." In: *Theoretical Ecology* 14.1 (2021), pp. 87–105.

[74] June Barrow-Green. *Poincaré and the three body problem*. 11. American Mathematical Soc., 1997.

[75] Raymond Austin Jarvis and Edward A Patrick. "Clustering using a similarity measure based on shared near neighbors." In: *IEEE Transactions on computers* 100.11 (1973), pp. 1025–1034.

[76] Egbert H Van Nes, Marten Scheffer, Victor Brovkin, Timothy M Lenton, Hao Ye, Ethan Deyle, and George Sugihara. "Causal feedbacks in climate change." In: *Nature Climate Change* 5.5 (2015), pp. 445–448.

[77] D Rigney, A Goldberger, W Ocasio, Y Ichimaru, G Moody, and R Mark. "Multi-channel physiological data: description and analysis." In: *Time Series Prediction: Forecasting the Future and Understanding the Past, A.S. Weigend and N.A. Gershenfeld (eds.), Reading, MA: Addison-Wesley* (1993), pp. 105–129.

[78] Daniele Marinazzo, Mario Pellicoro, and Sebastiano Stramaglia. "Kernel Method for Nonlinear Granger Causality." In: *Phys. Rev. Lett.* 100 (14 2008), p. 144103. DOI: 10.1103/PhysRevLett.100.144103. URL: https://link.aps.org/doi/10.1103/PhysRevLett.100.144103.

[79] F. Parrenin, V. Masson-Delmotte, P. Köhler, D. Raynaud, D. Paillard, J. Schwander, C. Barbante, A. Landais, A. Wegner, and J. Jouzel. "Synchronous Change of Atmospheric CO₂ and Antarctic Temperature During the Last Deglacial Warming." In: *Science* 339.6123 (2013), pp. 1060–1063. DOI: 10.1126/science.1226368. eprint: https://www.science.org/doi/pdf/10.1126/science.1226368. URL: https://www.science.org/doi/abs/10.1126/science.1226368.

[80] J Jouzel, V Masson-Delmotte, O Cattani, G Dreyfus, S Falourd, G Hoffmann, B Minster, J Nouet, JM Barnola, J Chappellaz, et al. "EPICA Dome C ice core 800KYr deuterium data and temperature estimates, IGBP PAGES/World Data Center for Paleoclimatology Data Contribution Series# 2007-091." In: *NOAA/NCDC Paleoclimatology Program, Boulder CO, USA* (2007).

[81] Dieter Lüthi, Martine Le Floch, Bernhard Bereiter, Thomas Blunier, Jean-Marc Barnola, Urs Siegenthaler, Dominique Raynaud, Jean Jouzel, Hubertus Fischer, Kenji Kawamura, et al. "High-resolution carbon dioxide concentration record 650,000–800,000 years before present." In: *nature* 453.7193 (2008), pp. 379–382.

[82] Laetitia Loulergue, Adrian Schilt, Renato Spahni, Valérie Masson-Delmotte, Thomas Blunier, Bénédicte Lemieux, Jean-Marc Barnola, Dominique Raynaud, Thomas F Stocker, and Jérôme Chappellaz. "Orbital and millennial-scale features of atmospheric CH4 over the past 800,000 years." In: *Nature* 453.7193 (2008), pp. 383–386.

[83] Jakob Runge, Sebastian Bathiany, Erik Bollt, Gustau Camps-Valls, Dim Coumou, Ethan Deyle, Clark Glymour, Marlene Kretschmer, Miguel D Mahecha, Jordi Muñoz-Marí, et al. "Inferring causation from time series in Earth system sciences." In: *Nature communications* 10.1 (2019), p. 2553.

[84] Hans-Ulrich Bauer and Klaus R Pawelzik. "Quantifying the neighborhood preservation of self-organizing feature maps." In: *IEEE Transactions on neural networks* 3.4 (1992), pp. 570–579.

[85] Erik P Hoel, Larissa Albantakis, and Giulio Tononi. "Quantifying causal emergence shows that macro can beat micro." In: *Proceedings of the National Academy of Sciences* 110.49 (2013), pp. 19790–19795.

DECLARATION

Hiermit versiche ich, dass ich

1. die Arbeit ohne unerlaubte fremde Hilfe angefertigt habe,
2. keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt habe,
3. die, den benutzen Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habt

Bremen,

Erik Laminski

COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". `classicthesis` is available for both LaTeX and LyX:

https://bitbucket.org/amiede/classicthesis/

Happy users of `classicthesis` usually send a real postcard to the author, a collection of postcards received so far is featured here:

http://postcards.miede.de/