

Observational causality by states and interaction type for scientific discovery

Álvaro Martínez-Sánchez^{1,†} and Adrián Lozano-Durán^{1,2}

¹*Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

²*Graduate Aerospace Laboratories, California Institute of Technology, Pasadena, CA 91125, USA and*

[†]*Corresponding author: alvaroms@mit.edu*

Causality plays a central role in understanding interactions between variables in complex systems. These systems often exhibit state-dependent causal relationships, where both the strength and direction of causality vary with the value of the interacting variables. In this work, we introduce a state-aware causal inference method that quantifies causality in terms of information gain about future states. The effectiveness of the proposed approach stems from two key features: its ability to characterize causal influence as a function of system state, and its capacity to distinguish between redundant and synergistic interactions. The method is validated across a range of benchmark cases in which the direction and strength of causality evolve in a prescribed manner with the state of the system. We further demonstrate the applicability of our approach in two real scenarios: the interaction between motions across scales in a turbulent boundary layer, and the Walker circulation phenomenon in tropical Pacific climate dynamics. Our results show that, without accounting for state-dependent causality as well as redundant and synergistic effects, traditional approaches to causal inference may lead to incomplete or misleading conclusions.

Keywords: causal inference, complex systems, turbulence, climate science

INTRODUCTION

Causality is a fundamental concept in understanding the interactions between variables in complex systems. It provides insights into how one variable influences another and guides the implementation of meaningful changes within these systems [1]. In many real-world scenarios, the nature of these interactions is not uniform across all states of the system; they often vary depending on the value of the variables involved. For example, in climate science, the direction and magnitude of causality between atmospheric variables can shift dramatically under different conditions, such as the wind flow direction in the Pacific Region during El Niño or La Niña events [2]. Similarly, in neuroscience, the causality among brain signals often depends on the activation intensity of surrounding regions [3]. These examples highlight the importance of developing methods that not only detect causal relationships on average, but also capture how these interactions vary with the state of the system—and consequently, over time. By addressing this need, state-aware causal analysis can drive progress in a wide range of scientific and engineering fields, including climate science [4], neuroscience [5], economics [6], epidemiology [7], social sciences [8], and fluid dynamics [9, 10].

To date, most causal inference methods offer an estimate of the average causal strength across all system states, without providing insight into how causality varies at the level of individual states. One such method is Convergent Cross Mapping (CCM) [11] and its variants [12–15], which infer causality by analyzing connections in the reconstructed attractor of the underlying dynamical system using different combinations of observed variables. A similar limitation arises in the Peter–Clark Momentary Conditional Independence (PCMCI) framework [16] and its extensions [17–19], which focus on identifying a minimal conditioning set that includes the causal parents of the target variable. Traditional approaches, such as Granger causality [20], also suffer from this limitation, as they assess causality by testing whether the past values of a source variable improve the prediction of a target variable in the future, but only in an average sense over all system states. Feature-ranking methods, such as the distributed information bottleneck [21], have gained popularity for isolating the most informative variables in systems to enhance predictive accuracy. However, these approaches do not explicitly address causality, let alone how causal influence may vary with the state of the system.

Among the methods that offer a decomposition of causality by states, one of the most intuitive formulations relies on the concept of interventions [22, 23]. Interventions actively modify one variable to observe its effect on another. A family of approaches within this framework is based on the concept of effective information [24–26], which quantifies the strength of causal interactions by measuring the change in the state of the system when a variable is replaced with maximally uncertain noise. These methods provide a state-dependent decomposition of causality by selectively intervening under specific conditions (e.g., intervening on a variable A only when $A = 1$ to measure the causal effect from that state). However, when data are gathered from physical experiments, establishing causality through interventions may be highly challenging, impractical [27], or even unethical in fields such as neuroscience or climate science [16]. Furthermore, the notion of causality based on interventions raises important questions about the type of intervention that should be introduced, and whether such interventions might affect the outcome by forcing the system out of its natural attractor.

Information theory [28] has also become a foundational framework for quantifying causality. The motivation for employing information-theoretic approaches stems from the recognition of information as a fundamental property of physical systems [29, 30]. Early applications of information-theoretic causality involved the use of conditional entropies to define directed information [31, 32]. One of the most recognized developments in this area is the introduction of transfer entropy (TE) [33], which quantifies how knowledge of the past states of one variable reduces uncertainty about the future state of another. Subsequent efforts to refine TE led to the development of conditional transfer entropy (CTE) [34–38], aimed at addressing multivariate analyses. Other information-theoretic approaches, inspired by dynamical systems theory [39–42], quantify causality as the amount of information flowing from one process to another, as dictated by the governing equations of the system. Similar to previous methods, these information-based approaches provide only an average measure of causal strength and do not offer insight into which specific states contribute most to the observed causality.

Several extensions of TE have been developed to account for state-dependent information transfer [43], although many are formulated primarily to quantify information dynamics rather than to perform causal inference. One of these extensions is local TE [35, 37, 44], which evaluates the information transfer between specific states of the (past) source and (future) target variables by analyzing the expectation term in the original TE formulation. This method preserves several desirable properties, including the ability to recover the standard TE when averaged over all states. Local TE can take negative values, reflecting a combination of contributions from both informative states (positive values) and misinformative states (negative values). Specific TE was later introduced [45], refining the interpretation of state-dependent information transfer by projecting local TE onto the states of the target variable. This projection yields a non-negative decomposition, enhancing interpretability; however, it considers only the past states of the source variables. Other approaches, such as ensemble TE [46–50] or time-varying Liang–Kleeman information transfer [51, 52], explore information transfer across varying time windows to detect shifts in causal structure over time. More recently, a local version of Granger causality [53]—viewed as the linear counterpart to local TE—has been proposed. This formulation offers a computationally efficient alternative for analyzing state-dependent causal interactions in linear systems with limited data.

While the methods above have advanced our understanding of state-dependent interactions, they are unable to distinguish between synergistic, unique, and redundant causal influences. It has been shown that the inability to disentangle these components can obscure the true causal structure of a system, often leading to spurious or misleading conclusions [27]. For example, consider the effect of two genes, A and B , on the expression of eye color, C . These interactions can manifest in several ways: *synergistic*, where both genes must be active together to produce a specific eye color—such as green—that neither gene can induce on its own; *unique*, where gene A contributes to blue eye color via a pathway independent of gene B ; and *redundant*, where either gene alone is sufficient to produce brown eyes due to overlapping biological pathways. Without a method capable of isolating these different types of causal contributions, such complexities remain hidden, and critical insights into causal relationships are missed. Recently, the synergistic-unique-redundant decomposition of causality (SURD) [27] was introduced, an information-theoretic method for causal inference that explicitly accounts for the distinct contributions of synergistic, unique, and redundant causal influences. Other information-theoretic approaches have also been recently developed to disentangle synergistic and redundant effects in transfer entropy [54, 55]. However, neither SURD nor these earlier methods were designed to capture state-dependent variations in causality within the systems. As a result, these methods cannot distinguish between scenarios in which a specific value of gene $A = 'a'$ exerts a unique causal influence on a particular state ' c ' of the condition C , and those in which its effect is redundant or synergistic with a specific value ' b ' of gene B .

In this work, we present a causal inference method that simultaneously accounts for synergistic, unique, and redundant causal influences, while also decomposing these contributions at the level of individual system states. We show that this approach outperforms existing causal inference methods, including the original SURD formulation, by providing a more detailed, state-aware characterization of causal interactions. We validate the method across a range of benchmark problems in which causal pathways depend on the system state in a controlled and known manner. We also demonstrate the applicability of our approach in two real-world scenarios—fluid dynamics and climate science—highlighting its ability to uncover state-specific causal relationships.

RESULTS

State-dependent and interaction-specific causality

Consider the collection of N time-evolving source variables given by the vector $\mathbf{Q} = [Q_1(t), Q_2(t), \dots, Q_N(t)]$. For example, Q_i may represent the activity level of a specific brain region over time, a binary variable indicating the presence or absence of an event (e.g., whether it rained on a particular day), or the daily average of a stock market index (e.g., S&P 500). The components of \mathbf{Q} are observables and are treated as random variables. A particular

realization of \mathbf{Q} is denoted by \mathbf{q} , representing a specific state (i.e., value). Following the examples above, q_i may correspond to a particular level of neural activity, one of the two states (0 or 1) of the binary variable indicating whether it rained or not, or a specific value of the stock market index (e.g., the closing value on a given day). These states are defined by the user according to the specific requirements of the problem. Our objective is to quantify the causal influence of the states of the source, \mathbf{q} , on the future state of a target variable, denoted as $q_j^+ = q_j(t + \Delta T)$, where $\Delta T > 0$ represents the future time interval over which causal influence is assessed. The variable Q_j is often included in \mathbf{Q} , but this is not required in general.

Our approach is formulated in three steps. First, following the principle of *forward-in-time information propagation*—i.e., information flows only toward the future [30]—we quantify causality among variables in terms of increments of information. Second, we separate these increments into causal and non-causal components. Third, we further decompose the causal increments into distinct types of interactions: synergistic, unique, and redundant contributions.

For step one, we adopt the definition of causality proposed in Ref. [27]. In this framework, causality is quantified as the increase in information about each future state q_j^+ gained by observing individual or groups of past states \mathbf{q} . The information content in Q_j^+ is measured by Shannon information (or entropy) [28], denoted as $H(Q_j^+)$, which represents the average number of bits needed to determine the value of Q_j^+ unambiguously. Shannon information can also be interpreted as a measure of uncertainty: highly uncertain processes (high entropy) yield greater information gain when their outcomes are revealed, whereas completely deterministic processes (zero entropy) provide no new information upon observation.

For the second step, we decompose the information in $H(Q_j^+)$ into a sum of information increments contributed by each past state of the observable vector \mathbf{Q} :

$$H(Q_j^+) = \sum_{q_j^+ \in Q_j^+} \sum_{\mathbf{q} \in \mathbf{Q}} [\Delta C(q_j^+; \mathbf{q}) + \Delta N(q_j^+; \mathbf{q})] + \Delta I_{\text{leak} \rightarrow j}, \quad (1)$$

where the terms $\Delta C(q_j^+; \mathbf{q}) > 0$ and $\Delta N(q_j^+; \mathbf{q}) \leq 0$ represent the causal and non-causal contributions, respectively, for a given source and target state. The term $\Delta I_{\text{leak} \rightarrow j}$ is referred to as the causality leak [30]. The mathematical expression for each of the terms in Eq. (1) is given in Methods and its derivation is detailed in Supplementary Materials. Here, we provide an interpretation of each term:

- The causal contribution $\Delta C(q_j^+; \mathbf{q}) > 0$ is the *positive* increment of information that a source state \mathbf{q} provides about a future target state q_j^+ , resulting in a decrease in uncertainty about q_j^+ .
- The non-causal contribution $\Delta N(q_j^+; \mathbf{q}) \leq 0$ is the *negative* increment of information obtained by observing \mathbf{q} , which leads to increased uncertainty about the future state q_j^+ .
- The causality leak $\Delta I_{\text{leak} \rightarrow j} \geq 0$ is the effect of *unobserved* variables that influence Q_j^+ but are not contained in \mathbf{Q} , i.e., the information of Q_j^+ that remains unexplained after having observed \mathbf{Q} .

To illustrate the distinction between $\Delta C(q_j^+; \mathbf{q})$ and $\Delta N(q_j^+; \mathbf{q})$, consider the task of forecasting whether it will rain tomorrow, denoted by Q_{rain}^+ , where $q_{\text{rain}}^+ = 1$ indicates rain and $q_{\text{rain}}^+ = 0$ indicates no rain. Suppose we use an atmospheric pressure sensor as the observable, Q_p , which can take on two values: low ($q_p = 0$) or high ($q_p = 1$). Within our causality framework, if observing low pressure ($q_p = 0$) provides a positive increment of information (i.e., reduces uncertainty) about the state $q_{\text{rain}}^+ = 1$, then $q_p = 0$ is considered causal and contributes to $\Delta C(q_{\text{rain}}^+; q_p)$. Conversely, if observing $q_p = 0$ yields a negative increment of information (i.e., increases uncertainty) about $q_{\text{rain}}^+ = 0$, then $q_p = 0$ is regarded as non-causal and contributes to $\Delta N(q_{\text{rain}}^+; q_p)$.

The final step is to decompose $\Delta C(q_j^+; \mathbf{q})$ and $\Delta N(q_j^+; \mathbf{q})$ into distinct types of interactions—namely, redundant, unique, and synergistic components [56]:

$$\Delta C(q_j^+; \mathbf{q}) = \sum_{\mathbf{i} \in \mathcal{P}} \Delta C_{\mathbf{i} \rightarrow j}^R + \sum_{i=1}^N \Delta C_{i \rightarrow j}^U + \sum_{\mathbf{i} \in \mathcal{P}} \Delta C_{\mathbf{i} \rightarrow j}^S, \quad (2)$$

where $\Delta C_{\mathbf{i} \rightarrow j}^R$, $\Delta C_{i \rightarrow j}^U$, and $\Delta C_{\mathbf{i} \rightarrow j}^S$ are the redundant, unique, and synergistic causalities, respectively, from the source states \mathbf{q} to target state q_j^+ . Unique causalities are associated with individual components of \mathbf{q} , whereas redundant and synergistic causalities arise from groups of components from \mathbf{q} . Consequently, the set \mathcal{P} contains all combinations involving more than one variable. For instance, for $N = 2$, Eq. (2) reduces to $\Delta C(q_j^+; \mathbf{q}) = \Delta C_{12 \rightarrow j}^R + \Delta C_{1 \rightarrow j}^U + \Delta C_{2 \rightarrow j}^U + \Delta C_{12 \rightarrow j}^S$. Similar to the causal and non-causal contributions, we provide an interpretation of redundant, unique and synergistic causalities, with the mathematical details available in the Methods section and the Supplementary Materials.

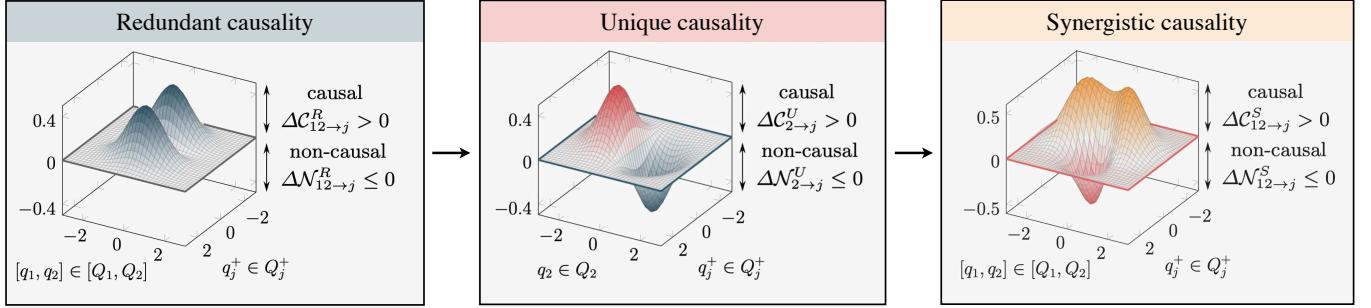


FIG. 1. Diagram of the decomposition of causal dependencies between a vector of observed states $\mathbf{q} = [q_1, q_2]$ and a future target state q_j^+ into synergistic (S), unique (U), and redundant (R) components—shown in orange, red, and blue, respectively. Note that redundant and synergistic interactions depend on all three states, $[q_1, q_2, q_j^+]$, but the source $[q_1, q_2]$ is shown along a single axis for schematic simplicity. The causal contributions are assigned following the order: redundant → unique → synergistic. As a result, causal states are color-coded according to the type of causality they contribute to, whereas non-causal states are shown using the color of the preceding causality level. For example, in the unique causality panel, causal states are shown in red, while non-causal states are shown in blue. This approach prevents the double counting of causal influences.

- Redundant causality from $\mathbf{q}_i = [q_{i_1}, q_{i_2}, \dots]$ to q_j^+ (denoted by $\Delta C_{i \rightarrow j}^R$) refers to the causal influence shared among all individual states in \mathbf{q}_i , where \mathbf{q}_i is a subset of \mathbf{q} . Redundant causality occurs when observing any of the states q_{i_1}, q_{i_2} , etc., provides identical amount of information about the outcome q_j^+ . Consider again the example of rain forecasting: we might use two pressure sensors placed in close proximity, both offering redundant causal information—i.e., each is equally useful for predicting whether it will rain or not.
- Unique causality from q_i to q_j^+ (denoted by $\Delta C_{i \rightarrow j}^U$) is the causal influence from q_i that cannot be obtained from any other individual variable different from Q_i . This occurs when observing q_i yields more information about q_j^+ than observing all the states of any other individual variable. In the previous example, this will occur when one pressure sensor provides more information about the likelihood of rain than the other.
- Synergistic causality from $\mathbf{q}_i = [q_{i_1}, q_{i_2}, \dots]$ to q_j^+ (denoted by $\Delta C_{i \rightarrow j}^S$) refers to the causal influence that arises from the joint effect of the states \mathbf{q}_i . Synergistic causality occurs when more information about q_j^+ is gained by observing the collection of states $[q_{i_1}, q_{i_2}, \dots]$ simultaneously than by observing all the states of subsets of \mathbf{q}_i individually. In the rain forecasting scenario, a synergy would take place when the information gained by measuring with two pressure sensors simultaneously is greater than the information provided by each sensor individually.

An analogous decomposition (redundant, unique, and synergistic) applies to the non-causal term $\Delta N(q_j^+; \mathbf{q})$. However, in this case, the information gain is negative—acquiring this knowledge actually increases uncertainty about future outcomes. Here, we focus on the causal contribution $\Delta C(q_j^+; \mathbf{q})$. The non-causal components also offer valuable insight, with further discussion and examples provided in the Supplementary Materials.

By definition, causal contributions are assigned following the order: redundant → unique → synergistic. Redundant causalities are identified first. The next increments of information that are not redundant are attributed to unique causality. Finally, the remaining causal influence is classified as synergistic. This implies that when contributions are labeled as causal (information gain) or non-causal (information loss), the gain/loss is referred to the previously identified redundant, unique, or synergistic causality. For example, in a system with one redundant, one unique, and one synergistic contribution: redundant causality reflects a gain relative to the information in the target q_j^+ ; unique causality reflects a gain relative to the redundant component; and synergistic causality reflects a gain relative to the unique one. This process is illustrated in the diagrams in Fig. 1, where the causal and non-causal contributions between the source states q_1 and q_2 and the target state q_j^+ are decomposed into its redundant, unique, and synergistic components. This approach ensures that causal contributions are not double-counted if they have already been accounted for in the previous type of interaction.

Our approach also enables the representation of causality as a function of time. The instantaneous value of causality at each time step is obtained by identifying the current state of the system and extracting the corresponding state-dependent value of $\Delta C_{i \rightarrow j}^R$, $\Delta C_{i \rightarrow j}^U$, and $\Delta C_{i \rightarrow j}^S$. As a result, it becomes possible to visualize the dynamic evolution of causal interactions, capturing temporal reversals and highlighting periods dominated by different types of causal contributions.

Finally, we define the cumulative causality over all possible states, after accounting for the uncertainty introduced

$$q_2(n+1) = \begin{cases} \sin[q_1(n)] + 0.1w_2(n) & \text{if } q_1(n) > 0 \\ q_2(n) + 0.1w_2(n) & \text{otherwise} \end{cases}$$

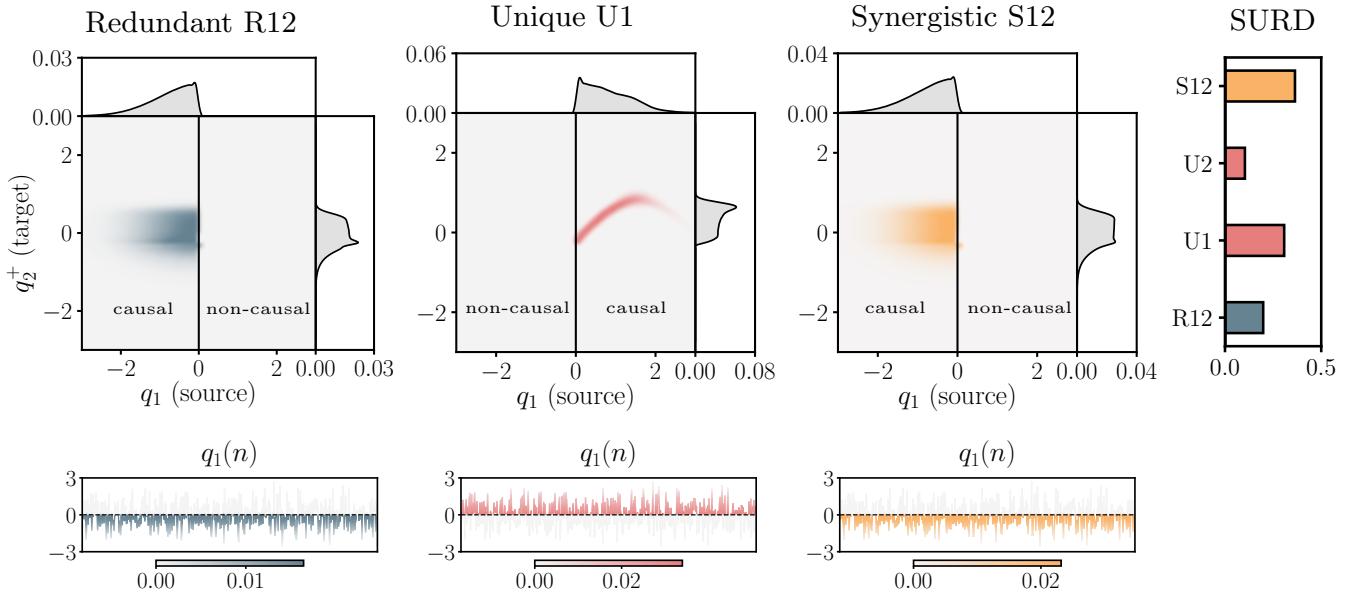


FIG. 2. State-dependent decomposition of causality for the source-dependent case. From left to right, the panels show: state-dependent redundant ($\Delta C_{12 \rightarrow 2}^R$), unique ($\Delta C_{12 \rightarrow 2}^U$), and synergistic ($\Delta C_{12 \rightarrow 2}^S$) causal contributions in blue, red, and orange, respectively; and the SURD causalities to the target variable Q_2^+ . The notation follows: U1 denotes the unique SURD causality from Q_1 to Q_2^+ , i.e., $\Delta I_{1 \rightarrow 2}^U$, with analogous definitions for other components (i.e., R and S). All SURD and state-dependent causal contributions are normalized by the mutual information $I(Q_2^+; Q_1, Q_2)$. The causal maps for redundant and synergistic components show averages over all states of q_2 . In each panel, the bottom row displays the temporal evolution of $q_1(n)$, color-coded according to the corresponding instantaneous state-dependent causal contribution.

by non-causal contributions, as $\Delta I_{i \rightarrow j}^\alpha = \sum_{q_j^+ \in Q_j^+} \sum_{q \in Q} [\Delta C_{i \rightarrow j}^\alpha + \Delta N_{i \rightarrow j}^\alpha]$, for $\alpha \in [R, U, S]$. These cumulative increments of information directly yield the causalities from SURD proposed in Ref. [27]. This is a key feature of our method, as SURD causality has been shown to outperform other definitions of causality across a wide range of benchmark cases.

Validation

We demonstrate our causal inference approach using validation cases representative of two common types of causal relationships: source-dependent and target-dependent causality. Each system consists of two source variables, Q_1 and Q_2 , and a target variable, Q_2^+ , designed such that the direction of causality varies with the intensity of either the source or the target. These benchmark cases are constructed so that the direction of causality for each state is known *a priori*, allowing us to assess the accuracy and reliability of the method. Importantly, the cases are also designed to exhibit qualitatively similar SURD causalities—i.e., comparable cumulative increments $\Delta I_{i \rightarrow 2}^\alpha$ —while the underlying causal flow reveals an entirely opposite structure as a function of the system state, a distinction that can only be revealed by a causality-by-state decomposition. Additional validation cases are presented in the Supplementary Materials.

Source-dependent causality

We consider the two-variable system, $[Q_1, Q_2]$, in which the direction of causality toward the future state Q_2^+ depends on the value of the source variable Q_1 . Specifically, $Q_1 \rightarrow Q_2^+$ when $q_1 > 0$, and $Q_2 \rightarrow Q_2^+$ when $q_1 \leq 0$. The equation governing the system is provided in Fig. 2. The variable Q_1 is sampled from a normal distribution with zero mean and unit variance. The variable Q_2 is influenced by an external noise term, W_2 , drawn from a normal distribution with mean -2 and unit standard deviation. This noise term is not included in the vector of observables

$$q_2(n+1) = \begin{cases} q_1(n) \sin[q_1(n)] + 0.1w_2(n) & \text{if } q_1(n) \sin[q_1(n)] + 0.1w_2(n) > 0 \\ q_2(n) + 0.1w_2(n) & \text{otherwise} \end{cases}$$

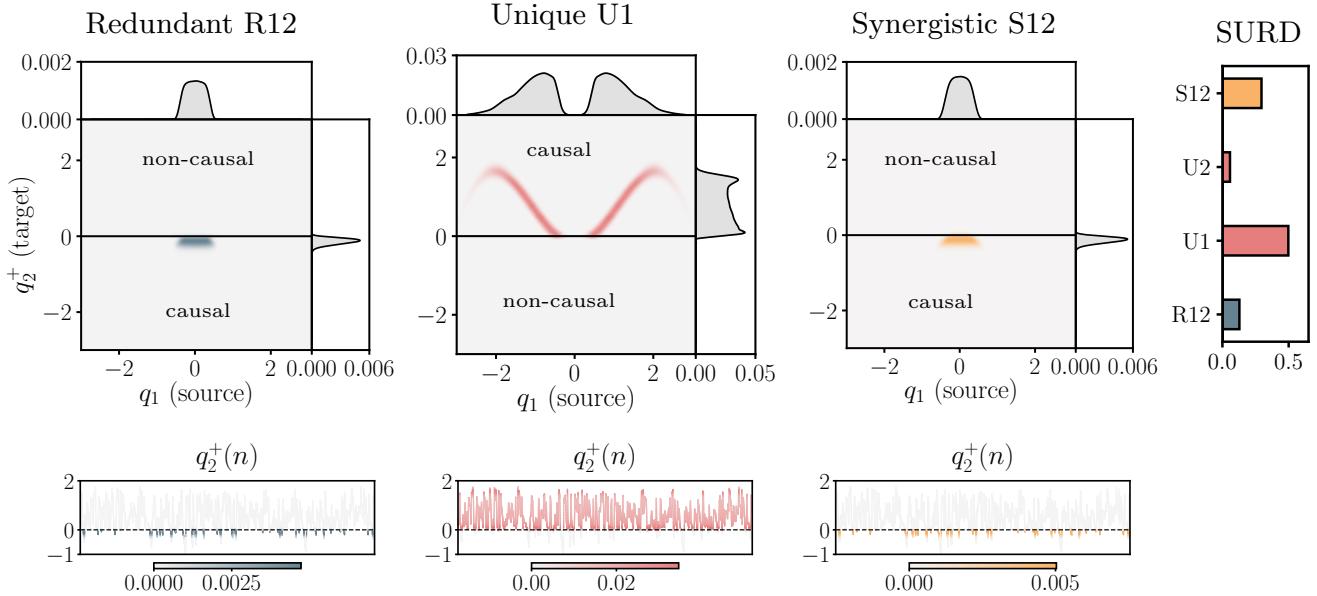


FIG. 3. State-dependent decomposition of causality for the target-dependent case. From left to right, the panels show: state-dependent redundant ($\Delta C_{12 \rightarrow 2}^R$), unique ($\Delta C_{1 \rightarrow 2}^U$), and synergistic ($\Delta C_{12 \rightarrow 2}^S$) causal contributions in blue, red, and orange, respectively; and the SURD causalities to the target variable Q_2^+ . The notation follows: U1 denotes the unique SURD causality from Q_1 to Q_2^+ , i.e., $\Delta I_{1 \rightarrow 2}^U$, with analogous definitions for other components (i.e., R and S). All SURD and state-dependent causal contributions are normalized by the mutual information $I(Q_2^+; Q_1, Q_2)$. The causal maps for redundant and synergistic causalities are shown for the q_2 that maximizes the total sum of individual contributions. In each panel, the bottom row displays the temporal evolution of $q_2^+(n)$, color-coded according to the corresponding instantaneous state-dependent causal contribution.

and serves to mimic unobserved perturbations commonly encountered in real-world systems.

The state-dependent redundant, unique, and synergistic causalities for the target variable Q_2^+ are shown in Fig. 2. The figure also includes the time evolution of the causalities, which is obtained by mapping the state-aware causality to the corresponding time instances.

For states with $q_1 > 0$, the results clearly show a nonzero contribution from unique causality ($\Delta C_{1 \rightarrow 2}^U$). In this regime, Q_1 acts as the sole driver of Q_2^+ , providing unique information that is not contained in Q_2 . For states with $q_1 < 0$, both redundant ($\Delta C_{12 \rightarrow 2}^R$) and synergistic ($\Delta C_{12 \rightarrow 2}^S$) causalities coexist. Redundancy arises because, in the region $q_1 < 0$, the information contained in q_1 is already provided (i.e., rendered redundant) by q_2 , which fully determines the outcome. This leads to duplicated information about Q_2^+ from both Q_1 and Q_2 . Synergistic causality emerges because two pieces of information are simultaneously needed to predict Q_2^+ , i.e., the value of Q_1 , to assess the condition $q_1 < 0$, and the value of Q_2 , to determine the outcome. Thus, Q_1 and Q_2 jointly provide information about Q_2^+ that neither can supply independently.

Fig. 2 shows the cumulative causal contributions to Q_2^+ based on the SURD causalities ($\Delta I_{i \rightarrow 2}^\alpha$). These results, which are consistent with the discussion above, will serve as a reference for the next benchmark case, where we demonstrate that the underlying state-dependent causal structures follow opposite trends despite exhibiting similar SURD causalities.

Target-dependent causality

The second benchmark case involves a system in which the direction of causality depends on the state of the target variable. As shown in Fig. 3, the system again consists of two variables, $[Q_1, Q_2]$, such that $f(Q_1) \rightarrow Q_2^+$ when $f(q_1) < 0$, and $Q_2 \rightarrow Q_2^+$ when $f(q_1) \geq 0$. Here, $f(\cdot)$ denotes a nonlinear function that maps the past of the source variable to the future of the target. The variable Q_1 is sampled from a normal distribution with zero mean and unit variance, while Q_2 is influenced by an external noise term, W_2 , drawn from a normal distribution with unit mean and

unit standard deviation.

The state-dependent redundant, unique, and synergistic causalities for the target variable Q_2^+ are shown in Fig. 3, along with the corresponding time evolution of the causalities. The analysis clearly shows that the unique causality, $\Delta C_{1 \rightarrow 2}^U$, is primarily determined by the state of the target variable rather than the source. In particular, positive target states ($q_2^+ > 0$) exhibit the strongest unique influence from Q_1 , consistent with the fact that Q_1 acts as the sole driver of Q_2^+ in this regime. In contrast, negative values of q_2^+ give rise to both synergistic ($\Delta C_{12 \rightarrow 2}^S$) and redundant ($\Delta C_{12 \rightarrow 2}^R$) causalities, as both Q_1 and Q_2 are required to predict Q_2^+ when $f(q_1) \leq 0$. This behavior mirrors the reasoning observed in the source-dependent benchmark case.

It is worth noting that the cumulative SURD causality shown in Fig. 3 resembles that of the source-dependent case discussed earlier, despite the fundamentally different interactions between the system variables. This illustrates a clear case where a state-aware decomposition of causality is essential to uncover causal pathways that would be overlooked by a state-averaged inference approach.

Application to flow interactions in wall turbulence

We apply our state-dependent causal inference method to identify the most causal flow states within a turbulent boundary layer, i.e., the chaotic fluid motion within a thin region adjacent to solid surfaces. Turbulent boundary layers are responsible for nearly 50% of the aerodynamic drag on modern airliners and play a critical role in the lower atmosphere, particularly within the first hundred meters, where they influence broader meteorological phenomena [57]. Understanding and modeling the interactions among turbulent motions of different scales within these layers remains a major challenge due to strong nonlinearities and high dimensionality of the system. Prior studies have documented the impact of large-scale motions in the outer layer on smaller-scale motions near the wall [27, 57–65], supporting the notion of top-down causality (a.k.a. Townsend’s outer-layer similarity hypothesis [66]). In this work, we focus on identifying the specific flow states that contribute to the causal influence from the outer layer flow (far from the wall) to the inner layer flow (near the wall).

We utilize data from a high-fidelity numerical simulation of turbulent flow over a flat plate with zero mean-pressure gradient, which provides a realistic representation of turbulent boundary layers encountered in engineering applications [67]. Fig. 4 shows an instantaneous visualization of the streamwise velocity. The flow conditions are characterized by the friction Reynolds number, $Re_\tau = u_\tau \delta / \nu$, which ranges from around 300 (inflow) to 700 (outflow), where δ is the boundary-layer thickness, ν is the kinematic viscosity, and u_τ is the friction velocity at the wall. The time signals analyzed here correspond to the streamwise velocity measured at two wall-normal locations, representing the inner and outer layers of the boundary layer. These velocity signals, denoted as $u_I(t)$ and $u_O(t)$, are located at wall-normal distances of $y_I = 4\nu/u_\tau$ and $y_O = 0.3\delta$, respectively. Fig. 4 presents a sample of these velocity signals at $Re_\tau = 500$.

The objective is to identify the states that contribute most significantly to the causal influence of the outer-layer velocity $u_O(t)$ on the future inner-layer velocity $u_I^+(t) = u_I(t + \Delta T)$. The time lag ΔT for the causal evaluation is selected to maximize the unique causality from $u_O(t)$ to $u_I^+(t)$, resulting in an optimal delay of approximately $\Delta T \approx 60\nu/u_\tau^2$. The causal influence in the reverse direction, from $u_I(t)$ to $u_O(t + \Delta T)$, is discussed in the Methods section. We adopt the notation $(\cdot)'$ to indicate standardization of the signals by their respective mean and standard deviation. Thus, negative values of u_O' represent below-average deviations of u_O , while positive values correspond to above-average deviations, with the same convention applied to u_I' .

We begin by analyzing the SURD causalities from $[u_I, u_O]$ to u_I^+ , as shown in Fig. 4. The redundant, unique, and synergistic components are denoted by $\Delta I_{OI \rightarrow I}^R$, $\Delta I_{OI \rightarrow I}^U$, and $\Delta I_{OI \rightarrow I}^S$, respectively, where the subscripts indicate contributions from the outer-layer (O) and inner-layer (I) velocities. The results demonstrate that the unique causality from u_O to u_I^+ dominates, followed by the synergistic and redundant components. In contrast, the unique causality from u_I to u_I^+ is negligible at the current time lag ΔT , since this influence occurs on timescales shorter than ΔT [68].

We now examine the state-dependent decomposition of causality, presented in Fig. 4, which also illustrates the temporal evolution of the individual causal components. The dominant form of causality is the unique component, which exceeds the redundant and synergistic components by more than an order of magnitude. The small redundant and synergistic causalities are concentrated in regions where $u_I^{+'} < 0$. This indicates that, during low-speed motions in the inner layer (i.e., velocities below the mean flow), part of the causal influence arises from overlapping information simultaneously conveyed by the outer- and inner-layer histories, while another portion is jointly determined by concurrent fluctuations in both layers.

The unique causal contribution is predominantly distributed within the first ($u_O' > 0, u_I^{+'} > 0$) and third ($u_O' < 0, u_I^{+'} < 0$) quadrants of the $u_O'-u_I^{+'}$ space. These quadrants represent states where outer- and inner-layer motions concurrently exceed or fall below the mean velocity. Moreover, the region of causal states $\Delta C_{O \rightarrow I}^U$ in the

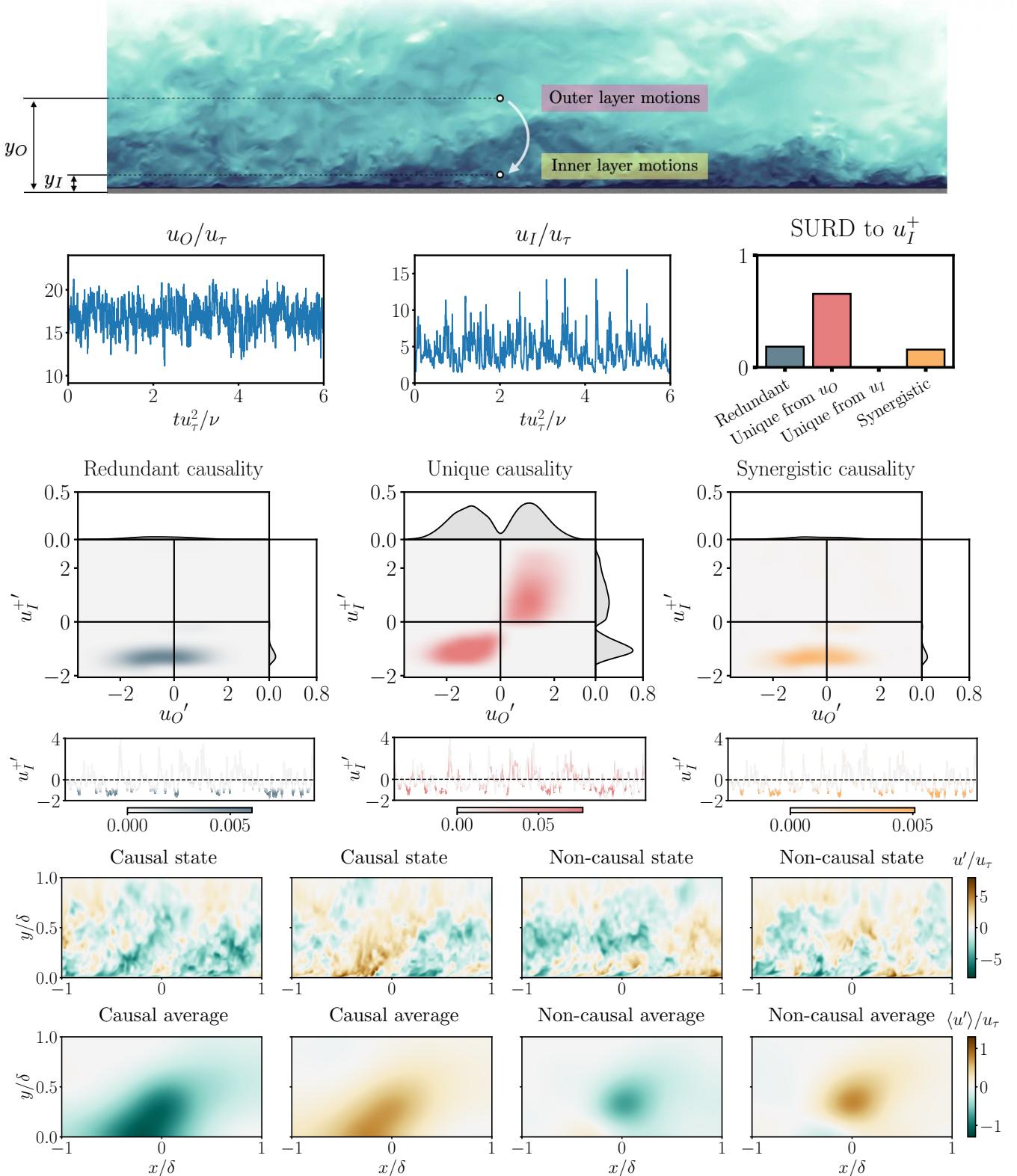


FIG. 4. Causality between inner and outer flow motions in a turbulent boundary layer. (a) Instantaneous visualization of the streamwise velocity field. (b) Temporal evolution of the streamwise velocity at two fixed wall-normal locations: the outer layer ($y_O = 0.3\delta$) and the inner layer ($y_I = 4\nu/u_\tau$). (c) Redundant ($\Delta I_{OI \rightarrow I}^R$), unique ($\Delta I_{(.) \rightarrow I}^U$), and synergistic ($\Delta I_{OI \rightarrow I}^S$) SURD causalities to the future inner-layer velocity $u_I^+ = u_I(t + \Delta T)$. The time lag ΔT is chosen to maximize cross-induced unique causality. (d) State-dependent redundant ($\Delta C_{OI \rightarrow I}^R$), unique ($\Delta C_{OI \rightarrow I}^U$), and synergistic ($\Delta C_{OI \rightarrow I}^S$) causalities as functions of the instantaneous streamwise velocities u_O (outer layer) and u_I^+ (inner layer). Note that for some of the panels, the color scale has been intentionally saturated to enhance visual contrast; however, the causalities are of small magnitude, as evidenced by the horizontal and vertical projections. (e) Time series of u_I^+ color-coded by the dominant state-dependent causal component at each time step. (f) Instantaneous and (g) conditionally averaged visualizations of flow fields corresponding to the causal and non-causal states, respectively.

first quadrant is wider than the third, consistent with previous studies [61] showing that large positive fluctuations in the outer layer exert a stronger influence on small-scale inner-layer structures compared to negative fluctuations. This observation supports the well-known modulation effect, whereby high-speed outer-layer events leave a stronger causal imprint on the inner-layer dynamics [61]. In contrast, the second and fourth quadrants do not contain causal contributions. Overall, these findings indicate that high-speed streaks propagating towards the wall act as dominant drivers of unique causality. Conversely, low-speed streaks are predominantly associated with redundant or synergistic causal interactions.

We conclude this section by examining the flow structures associated with causal and non-causal states. To this end, we analyze instantaneous realizations of the streamwise velocity field, $u(x, z, t)$, within a spatial region of size 2δ , conditioned on states of $u_O(t)$ and $u_I^+(t)$ corresponding to the four quadrants of the unique causality map $\Delta C_{O \rightarrow I}^U$ discussed previously. Fig. 4 shows example snapshots of $u(x, z, t)$ representative of each quadrant. To highlight characteristic flow patterns for each quadrant, we compute ensemble-averaged velocity fields, $\langle u(x, z, t) \rangle$, over all instances belonging to the respective quadrant. For instance, the flow field in the first quadrant is averaged over all time instants for which $u'_O > 0$ and $u'^+_I > 0$, and similarly for the other quadrants. The results (bottom of Fig. 4) reveal that causal states correspond to coherent velocity structures attached to the wall, extending beyond $y = 0.3\delta$, the wall-normal location of u_O . These findings indicate that causal interactions are associated with coherent flow structures—such as high- or low-speed streaks—that span across the inner and outer layers. Furthermore, the shape of these structures is in agreement with other averaged flow structures found in wall-bounded turbulence [69]. In contrast, the non-causal states (second and fourth quadrant) do not exhibit such structural coherence, with velocity fields appearing fragmented and detached from the wall. This fragmentation shows a breakdown of spatial coherence and diminished influence from the outer to the inner layer.

Application to Walker circulation

We apply our state-dependent causal inference method to analyze the Walker circulation in the tropical Pacific—a key atmospheric process with a well-established physical mechanism [72], and a subject of previous causal investigations [4, 16]. However, a state-aware analysis that accounts for redundant and unique causalities has not yet been performed. The Walker circulation is a large-scale atmospheric circulation pattern in the tropical Pacific, driven by the temperature gradient between the warm waters of the Western Pacific and the cooler Eastern Pacific. Under normal conditions, strong trade winds transport warm surface waters westward, leading to moist air rising over the Western Pacific due to enhanced convection. This air ascends, moves eastward aloft, cools, and descends over the Eastern Pacific. The circulation is completed as the cooler, drier air flows westward near the surface, forming a closed loop across the tropical Pacific. Fig. 5 illustrates this circulation pattern, which governs critical atmospheric and oceanic interactions in the Pacific basin and plays a central role in driving global weather patterns and climate variability.

We use data consisting of regional averages of surface pressure anomalies in the West Pacific (WPAC) and surface air temperature anomalies in the Central Pacific (CPAC) and East Pacific (EPAC) regions for the period 1948–2012, comprising 780 months [16]. The regions used for these averages are illustrated as shaded boxes in Fig. 5. Here, we focus on understanding what particular states of each variable are responsible for the Walker circulation phenomenon from the East to West Pacific close to the sea surface. To achieve this, we analyze the most relevant total redundant and cross-induced unique causalities to CPAC and WPAC involving at most two variables, which are shown in Fig. 5. The number of samples in the dataset was insufficient to reliably estimate synergistic causalities, which are therefore not included in the analysis.

The decomposition for these causalities is shown in Fig. 5, as a function of the values of the source and target variables. First, we focus on the redundant and cross-induced unique causalities to $WPAC^+$. In this case, unique causalities dominate and redundant causalities are small. The small redundant causality shows causal states for negative values of CPAC and $WPAC^+$, whereas the higher causal states for the unique causality occur when both CPAC and $WPAC^+$ are either above or below their respective means. Notably, states above the mean exhibit stronger causal influence compared to those below [73]. For the rest of the cases, i.e., where CPAC and $WPAC^+$ have opposite signs, the state-dependent unique causality is zero, which indicates that causality is already explained by the redundant components. These results are consistent with the observation that positive CPAC temperature anomalies (associated with El Niño conditions) have a more pronounced impact on Western Pacific surface pressure than negative anomalies (associated with La Niña conditions) [73, 74].

For $CPAC^+$, the most significant contributions are the redundant $[CPAC, EPAC] \rightarrow CPAC^+$ and unique $EPAC \rightarrow CPAC^+$ causalities. We find that the values of EPAC and $CPAC^+$ contributing most significantly to $EPAC \rightarrow CPAC^+$ come from the extreme values of EPAC and $CPAC^+$, far from the mean. In contrast, EPAC values near the mean are identified as non-causal, as their causality to $CPAC^+$ is already explained by the redundant component $[CPAC,$

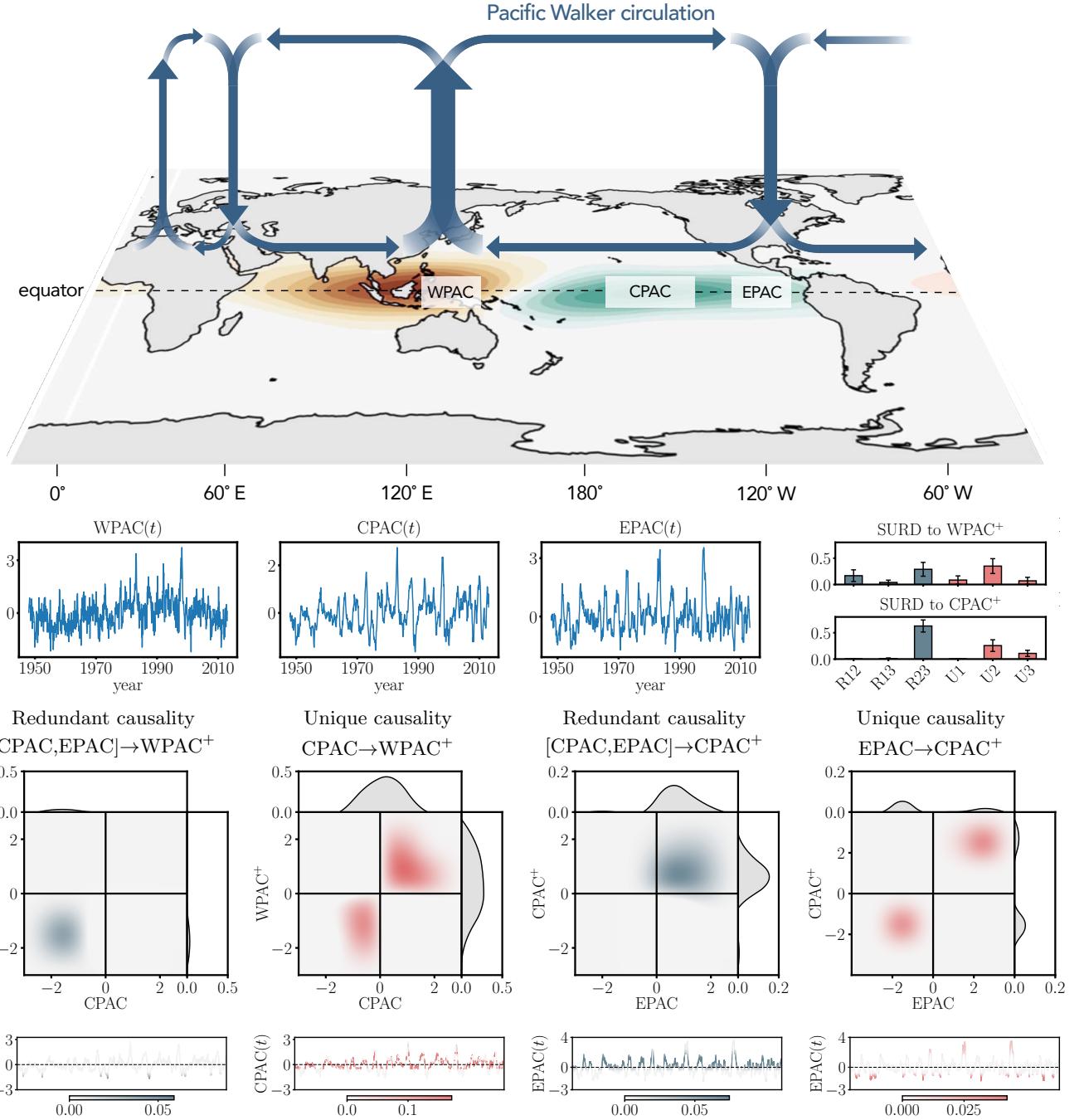


FIG. 5. Causality in the Pacific Walker circulation. Top: Schematic representation of the Pacific Walker circulation mechanism, showing the regions where regional averages were computed for the period 1948–2012 ($T = 780$ months): surface pressure anomalies in the West Pacific (WPAC) and surface air temperature anomalies in the Central Pacific (CPAC) and East Pacific (EPAC). The shaded boxes indicate the specific regions of interest. The schematic diagram was adapted from Ref. [70] and the data was taken from Watanabe et al. [71], which illustrates how local sea surface temperature (SST) warming affects the strength of the Pacific Walker circulation, with positive (orange) values indicating strengthening and negative (green) values indicating weakening. Middle top: Time evolution of WPAC, CPAC, and EPAC, along with redundant (R) and unique (U) SURD causalities to $\text{WPAC}^+ = \text{WPAC}(t + \Delta T)$ and $\text{CPAC}^+ = \text{CPAC}(t + \Delta T)$. The variables WPAC, CPAC, and EPAC are denoted as 1, 2, and 3, respectively. For example, the label $R13 \rightarrow \text{CPAC}^+$ indicates the redundant causality from WPAC and EPAC to CPAC^+ . The causal analysis is performed with a time lag of $\Delta T = 4$ months, which corresponds to the maximum cross-induced causality. Error bars on the SURD causalities represent the standard deviation across 50 bootstrap resamplings. Statistical significance was assessed via 1000 random permutations, yielding a p -value of zero for both redundant and unique components. Middle bottom: Decomposition of state-dependent causalities: redundant $[\text{CPAC}, \text{EPAC}] \rightarrow \text{WPAC}^+$, unique $\text{CPAC} \rightarrow \text{WPAC}^+$, redundant $[\text{CPAC}, \text{EPAC}] \rightarrow \text{CPAC}^+$, and unique $\text{EPAC} \rightarrow \text{CPAC}^+$ as a function of the source and target variable states. Note that in the left-most panel, the color scale has been intentionally saturated to enhance visual contrast; however, the actual causalities are of small magnitude, as indicated by the horizontal and vertical projections. Bottom: Temporal evolution of $\text{CPAC}(t)$ and $\text{EPAC}(t)$ color-coded according to the state-dependent causal contribution at that instant.

$\text{EPAC}] \rightarrow \text{CPAC}^+$, with causal states concentrated around positive values of EPAC and CPAC^+ . This result supports previous observations [75], where it was reported that positive temperature anomalies in the Central Pacific can originate from positive temperature anomalies in the Eastern Pacific [76].

Overall, our findings indicate that causal states for sea surface temperatures in the Central Pacific (CPAC) and Eastern Pacific (EPAC), as well as surface pressure in the Western Pacific (WPAC), predominantly occur when all three variables are either simultaneously above ($\text{EPAC} > 0$, $\text{CPAC} > 0$, $\text{WPAC} > 0$) or below ($\text{EPAC} < 0$, $\text{CPAC} < 0$, $\text{WPAC} < 0$) their respective means. In contrast, non-causal states emerge when there is a mismatch in trends between CPAC and WPAC, while EPAC remains near its mean. Specifically, non-causal interactions are observed when CPAC is below the mean and WPAC is above ($\text{CPAC} < 0$, $\text{WPAC} > 0$), or vice versa ($\text{CPAC} > 0$, $\text{WPAC} < 0$).

DISCUSSION

We have introduced an information-theoretic causal inference method that decomposes causality according to both the system state and the interaction type—namely, synergistic, unique, and redundant. This dual decomposition allows for the identification of causal system configurations, distinguishing them from non-causal states. The approach further enables visualization of causality over time, facilitating the detection of causal pathways. A key feature of the method is that summing the state-dependent causal and non-causal contributions recovers the SURD causality [27], which has previously been shown to outperform existing methods for causal inference across diverse scenarios.

Two benchmark cases were designed to represent common causal pathways: source-dependent and target-dependent causality. In both systems, the direction and strength of causal influence vary with the state of the interacting variables. Although the resulting SURD causalities for the two systems appear similar, their underlying causal flows are fundamentally different—a distinction clearly captured by our state-dependent causal inference method (Fig. 2). We have also compared our method to a state-dependent variant of conditional transfer entropy (CTE), detailed in the Methods section. Although CTE enables a state-wise evaluation of causality, it fails to detect source-dependent and target-dependent causal structures. This limitation arises because CTE aggregates all causal contributions into a single measure, without distinguishing between synergistic, unique, and redundant components. In general, we anticipate that existing causal inference methods—such as Granger causality, PCMCI, or convergent cross mapping—will likewise be unable to differentiate between source- and target-dependent causal pathways, since they rely on average causal strength and do not disentangle redundant and synergistic interactions.

We have applied our method to two real-world systems that had not previously been examined through the lens of state-level causality: scale interactions in a turbulent boundary layer and the Walker circulation in climate dynamics. In the turbulent boundary layer, our analysis revealed causal flow structures from the outer to the inner layer, demonstrating that high-speed streaks modulate near-wall motions in a state-specific manner (Fig. 4). In the Walker circulation case, we found that causal interactions between sea surface temperatures and surface pressure anomalies in the Eastern, Central, and Western Pacific emerge only when these variables deviate coherently from their mean values—highlighting the asymmetry between El Niño and La Niña regimes (Fig. 5). These detailed causal relationships are not captured by traditional causal inference methods that rely on state-aggregated measures.

Finally, it is informative to compare our method with other approaches proposed in the literature for analyzing the impact of source states on a target variable. Among these methods are local TE [35], specific TE [45], local GC [53], and the time-varying Liang–Kleeman (TvLK) information flow [51]. It is important to note that neither local TE nor specific TE are explicitly designed for causal inference—as discussed by their respective authors—but rather aim to quantify local information dynamics. Technical details of each approach are provided in the Supplementary Materials, and results for the two benchmark systems shown in Fig. 2 are discussed in Methods. Here, we offer a brief summary. Local TE decomposes information transfer across both source and target states, yielding informative (positive) and misinformative (negative) contributions, while local GC can be viewed as a variation of local TE tailored for systems with linear relationships. Specific TE provides a decomposition based solely on source states, which has the benefit of yielding only positive contributions. Finally, TvLK relies on an estimation of the original Liang–Kleeman information flow formulation, employing a square-root Kalman filter to estimate covariance matrices over time [51]. Although these approaches have proven useful for analyzing dynamical systems, they are not well suited to the problems considered here—namely, source-dependent and target-dependent systems—as their outcomes diverge from the intuitive causal pathways expected in such cases (see *Methods*). These limitations may be due to assumptions of linearity, difficulties in selecting an appropriate temporal scale, and the inability to distinguish redundant, unique, and synergistic contributions.

Overall, by providing a state-dependent, interaction-aware decomposition of causality, our method contributes a new perspective to the study of causal inference in complex systems. It offers a means to disentangle pathways for causal relationships and to identify specific conditions under which variables interact synergistically or redundantly. We hope that this approach can complement existing methods and help deepen our understanding of multiscale,

nonlinear systems in fields such as climate science, fluid mechanics, neuroscience, and beyond.

MATERIALS AND METHODS

Assumptions for causal discovery

The causal inference method proposed in this study follows an *observational* (i.e., non-interventional) approach within a probabilistic framework, inferring causal relationships from the transition probabilities between system states under the assumption that explicit interventions are not necessary. Specifically, we employ the notion of *causality* proposed in Ref. [27], in which causal influence is quantified through the informational gain associated with observing individual or groups of variables. The state-dependent causalities identified by our framework are consistent with the total causal contributions obtained from the SURD formulation [27].

We assume the *causal Markov condition*, which states that all relevant probabilistic information about a variable is contained in its direct causes (i.e., parents), rather than in indirect dependencies. This condition ensures that the inferred causal structure accurately reflects the true generative mechanism of the system. In addition, the method assumes *faithfulness*, i.e., observed statistical dependencies (or independencies) arise from the underlying causal structure itself, rather than from pathological parameter configurations or coincidences.

We further impose the principle of *forward-in-time information propagation* [30], which prohibits backward-in-time causation. To account for latent influences, our formulation incorporates the concept of a *causality leak* [27, 30], which quantifies the portion of causal influence that remains unexplained by the observed variables. This allows us to relax the standard assumption of *causal sufficiency*, which traditionally requires all common causes to be explicitly measured.

Finally, the method is *model-free*, requiring no prior knowledge of the governing equations or dynamics of the system. This generality makes the approach applicable to a wide variety of multivariate systems, whether deterministic or stochastic, and irrespective of whether the dependencies are linear or nonlinear. However, the analysis assumes that statistical properties of the input time series remain invariant over time. This assumption enables the estimation of slowly varying non-stationary dynamics, provided there is sufficient statistical support within those periods.

Decomposition of redundant, unique, and synergistic causalities by states

To perform the decomposition proposed in Eq. (2), we start with the SURD approach, which is based on a decomposition of mutual information [28, 77, 78]. The mutual information between a target variable Q_j^+ and a vector of source variables \mathbf{Q} is given by:

$$I(Q_j^+; \mathbf{Q}) = \sum_{q_j^+ \in Q_j^+} \sum_{\mathbf{q} \in \mathbf{Q}} p(q_j^+, \mathbf{q}) \log_2 \left(\frac{p(q_j^+, \mathbf{q})}{p(q_j^+)p(\mathbf{q})} \right) \geq 0, \quad (3)$$

where q_j^+ and \mathbf{q} denote the states of the target Q_j^+ and source \mathbf{Q} variables, respectively. Mutual information measures how different the joint probability distribution $p(q_j^+, \mathbf{q})$ is from the hypothetical distribution $p(q_j^+)p(\mathbf{q})$, where q_j^+ and \mathbf{q} are assumed to be independent. For instance, if Q_j^+ and \mathbf{Q} are not independent, then $p(q_j^+, \mathbf{q})$ will differ significantly from $p(q_j^+)p(\mathbf{q})$. The portion of information in Q_j^+ that remains unexplained by the source variables \mathbf{Q} is referred to as the causality leak. This can be quantified in closed form as the conditional Shannon information $H(Q_j^+ | \mathbf{Q})$ [28]:

$$H(Q_j^+ | \mathbf{Q}) = \sum_{q_j^+ \in Q_j^+} \sum_{\mathbf{q} \in \mathbf{Q}} -p(q_j^+, \mathbf{q}) \log_2 [p(q_j^+ | \mathbf{q})] \geq 0. \quad (4)$$

We decompose the mutual information into its redundant, unique, and synergistic components for each of the states of the target variable q_j^+ using the *specific* mutual information $\tilde{i}(Q_j^+ = q_j^+; \mathbf{Q})$, i.e., the mutual information between a fixed target state q_j^+ and all the source states in \mathbf{Q} [27, 79]:

$$\tilde{i}(q_j^+; \mathbf{Q}) = \sum_{\mathbf{q} \in \mathbf{Q}} \frac{p(q_j^+, \mathbf{q})}{p(q_j^+)} \log_2 \left(\frac{p(q_j^+, \mathbf{q})}{p(q_j^+)p(\mathbf{q})} \right) \geq 0. \quad (5)$$

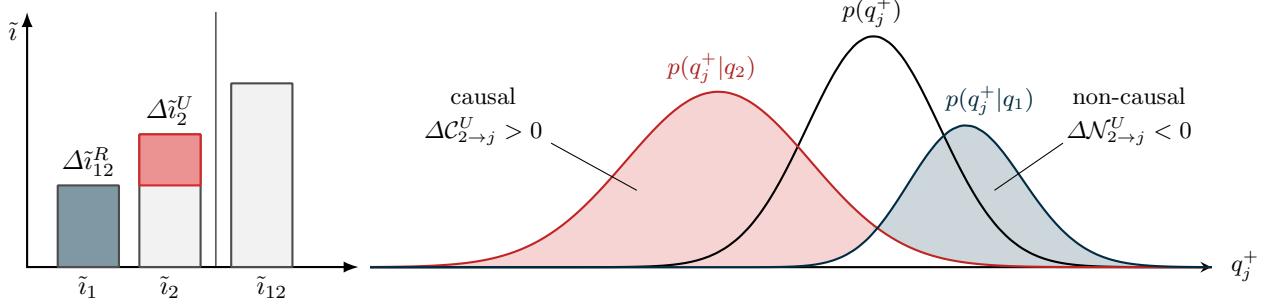


FIG. 6. Left: Ranking of specific mutual information between a target state q_j^+ and all combinations of two source variables $\mathbf{Q} = [Q_1, Q_2]$, namely \tilde{i}_1 , \tilde{i}_2 , and \tilde{i}_{12} . The redundant increment $\Delta\tilde{i}_{12}^R$ and the unique increment $\Delta\tilde{i}_2^U$ are computed based on this ranking. Right: State-dependent causality $\Delta C_{2 \rightarrow j}^U$ and non-causality $\Delta N_{2 \rightarrow j}^U$ for a system with target variable Q_j^+ and source variables Q_1 and Q_2 , as defined by Equation 6. In this example, it is assumed that the specific mutual information satisfies $\tilde{i}_2 > \tilde{i}_1$ for all $q_j^+ \in Q_j^+$. Positive values of $\Delta C_{2 \rightarrow j}^U$ correspond to states where knowledge of q_2 increases the likelihood of the target state q_j^+ relative to q_1 , i.e., $p(q_j^+ | q_2) > p(q_j^+ | q_1)$, indicating unique causality. Conversely, negative or zero values (denoted by $\Delta N_{2 \rightarrow j}^U$) identify non-causal states where q_2 provides less information about q_j^+ than q_1 . i.e., $p(q_j^+ | q_2) < p(q_j^+ | q_1)$.

For a given state q_j^+ of the target variable Q_j^+ , the specific causalities \tilde{i} are computed for all the possible combinations of source variables. These components are organized in ascending order. The increments of information between each \tilde{i} , denoted by $\Delta\tilde{i}$, define the redundant $\Delta\tilde{i}_{i \rightarrow j}^R$, unique $\Delta\tilde{i}_{i \rightarrow j}^U$, and synergistic $\Delta\tilde{i}_{i \rightarrow j}^S$ causalities. Fig. 6(left) shows an example for two source variables $\mathbf{Q} = [Q_1, Q_2]$ and a given state of the target variable q_j^+ . The example illustrates the specific mutual information \tilde{i}_1 , \tilde{i}_2 , \tilde{i}_{12} and the corresponding redundant $\Delta\tilde{i}_{12}^R$ and unique $\Delta\tilde{i}_2^U$ increments.

The redundant, unique, and synergistic information increments are further decomposed into causal and non-causal contributions for each state pair $q_j^+ - \mathbf{q}$. These causal redundant, unique, and synergistic contributions are denoted as $\Delta C_{i \rightarrow j}^R$, $\Delta C_{i \rightarrow j}^U$, and $\Delta C_{i \rightarrow j}^S$, respectively. The general mathematical expressions for $\Delta C_{i \rightarrow j}^\alpha$ with $\alpha \in \{R, U, S\}$ are detailed in the Supplementary Materials. Here, we provide the equations particularized for a system with two source variables $\mathbf{Q} = [Q_1, Q_2]$ and a target variable Q_j^+ , where $\tilde{i}_1 < \tilde{i}_2 < \tilde{i}_{12}$ is satisfied for all possible states $q_j^+ \in Q_j^+$:

$$\begin{aligned} \Delta C_{12 \rightarrow j}^R &= p(q_j^+, q_1, q_2) \max \left(0, \log_2 \frac{p(q_j^+ | q_1)}{p(q_j^+)} \right), \\ \Delta C_{2 \rightarrow j}^U &= \sum_{q_1 \in Q_1} p(q_j^+, q_1, q_2) \max \left(0, \log_2 \frac{p(q_j^+ | q_2)}{p(q_j^+ | q_1)} \right), \\ \Delta C_{12 \rightarrow j}^S &= p(q_j^+, q_1, q_2) \max \left(0, \log_2 \frac{p(q_j^+ | q_1, q_2)}{p(q_j^+ | q_2)} \right). \end{aligned} \quad (6)$$

Similar expressions can be derived for the non-causal components $\Delta N_{i \rightarrow j}^\alpha$, which are detailed in the Supplementary Materials.

Fig. 6 illustrates the interpretation of the unique causal and non-causal states, denoted by $\Delta C_{2 \rightarrow j}^U$ and $\Delta N_{2 \rightarrow j}^U$, respectively. In this context, causal states (in red) indicate that q_2 provides more information about the q_j^+ than q_1 , i.e., $\log_2 p(q_j^+ | q_2) > \log_2 p(q_j^+ | q_1)$. Conversely, non-causal states (in blue) indicate that q_2 offers less or the same information about q_j^+ than q_1 , i.e., $\log_2 p(q_j^+ | q_2) < \log_2 p(q_j^+ | q_1)$.

The SURD causalities, $\Delta I_{i \rightarrow j}^\alpha$ where $\alpha \in [R, U, S]$, are recovered by summing the causal and non-causal contributions across all possible states:

$$\Delta I_{i \rightarrow j}^\alpha = \sum_{q_j^+ \in Q_j^+} \sum_{\mathbf{q} \in \mathbf{Q}} [\Delta C_{i \rightarrow j}^\alpha(q_j^+; \mathbf{q}) + \Delta N_{i \rightarrow j}^\alpha(q_j^+; \mathbf{q})]. \quad (7)$$

Note that in SURD, the average causal effect of one variable on another is obtained by accounting for both causal and non-causal contributions. There are two main reasons for this construction. First, including both causal and non-causal contributions ensures consistency with the mutual information between variables, as illustrated in Fig. 1. This also guarantees consistency with the forward-in-time propagation of information condition from Eq. (1). The second reason is the need to define an average causal measure that discounts instances where knowledge of the source

variable actually increases uncertainty about the target. In practice, if only positive (causal) contributions were considered, one might mistakenly deem a variable Q_1 causal to another variable Q^+ based on a small subset of states q_1 where the influence is strong. However, it could occur that, in the vast majority of instances, knowledge of q_1 does not increase, and may even decrease, certainty about Q^+ . Including non-causal (negative) contributions ensures that these misleading instances are properly accounted for, leading to an average measure that more faithfully reflects the overall causal relationship between variables across all conditions. This is analogous to the distinction between informative and misinformative components in pointwise mutual information [80]. As a result, SURD avoids overestimating causality based on rare or atypical states and provides a more robust and interpretable measure of causal influence.

Finally, it is worth noting that other approaches for decomposing mutual information into redundant, unique, and synergistic components have been proposed in the literature, such as Partial Information Decomposition (PID) [81] and its variants [30, 82–86]. While these approaches provide valuable insights into the structure of mutual information, it was discussed in Ref. [27] that they either fail to satisfy all the properties required by our method or result in an unmanageable number of terms. Hence, our formulation does not follow PID, and the meanings of redundancy, uniqueness, and synergy adopted in this work differ from those in PID and related approaches. Instead, our framework focuses on quantifying the *increments of information* gained about the states of the target variable from different combinations of source variable states. Although this definition departs from earlier notions of redundancy, synergy, and uniqueness [81], it enables us, for example, to group synergistic components by specific order or to account for them through the concept of *causality leak*, thanks to the additivity of causal components. In our view, this simplifies the interpretation of scenarios involving many variables, where the number of decomposed terms can become prohibitively large, or situations where the number of observations is very low. For instance, PID results in a number of terms that grows according to the Dedekind numbers; in the case of nine variables, this decomposition yields over 10^{23} terms, whereas our method produces 512 causal terms (along with 512 non-causal counterparts)—which, although still large, remains computationally feasible. Despite our different notion of redundancy, uniqueness, and synergy, by focusing on incremental information, our method ensures that causality is not double-counted and that the minimal set of variables containing information about the target states can be systematically identified.

State-dependent conditional transfer entropy

We use a state-dependent version of conditional transfer entropy (CTE) to quantify directed information transfer between a target variable Q_j^+ and a vector of source variables \mathbf{Q} . The CTE from a source variable Q_i to the target Q_j^+ is defined as:

$$\overline{\text{CTE}}_{i \rightarrow j} = H(Q_j^+ | \mathbf{Q}_j) - H(Q_j^+ | \mathbf{Q}), \quad (8)$$

where \mathbf{Q}_j denotes the vector \mathbf{Q} excluding the component Q_i , and $H(\cdot | \cdot)$ is the conditional Shannon entropy [28]. This expression measures the additional information that Q_i provides about Q_j^+ beyond what is already available from the remaining variables in \mathbf{Q} . By expanding the conditional Shannon entropies in the definition above, we obtain a state-level representation of CTE:

$$\text{CTE}_{i \rightarrow j} = \sum_{q_j \in Q_j} p(q_j^+, q_i, \mathbf{q}_j) \log_2 \left(\frac{p(q_j^+ | q_i, \mathbf{q}_j)}{p(q_j^+ | \mathbf{q}_j)} \right). \quad (9)$$

For a system consisting of two source variables Q_1 and Q_2 , and a target variable Q_j^+ with joint probability distribution $p(q_j^+, q_1, q_2)$, the state-dependent conditional transfer entropy from the source variable Q_2 to the target Q_j^+ can be expressed as:

$$\text{CTE}_{2 \rightarrow j} = \sum_{q_1 \in Q_1} p(q_j^+, q_1, q_2) \log_2 \left(\frac{p(q_j^+ | q_1, q_2)}{p(q_j^+ | q_1)} \right), \quad (10)$$

which compares the log of the conditional probabilities $p(q_j^+ | q_1, q_2)$ and $p(q_j^+ | q_1)$, weighted by the joint probability $p(q_j^+, q_1, q_2)$. In contrast to the specific unique causality defined in Eq. (6), we observe that here the comparison is made between $p(q_j^+ | q_1, q_2)$ and $p(q_j^+ | q_1)$, while in the unique causality the comparison is between $p(q_j^+ | q_2)$ and $p(q_j^+ | q_1)$.

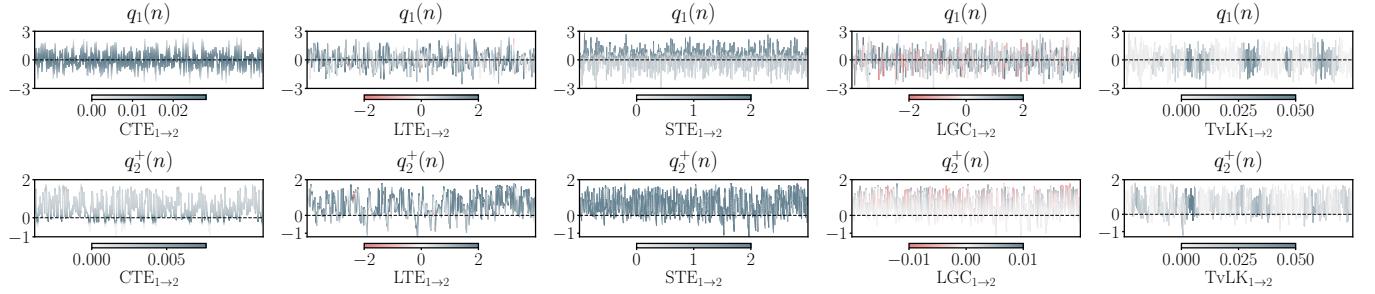


FIG. 7. Comparison of time-varying causal inference methods, including state-dependent conditional transfer entropy (CTE), local transfer entropy (LTE) [87], specific transfer entropy (STE) [45], local Granger causality (LGC) [53], and time-varying Liang–Kleeman (TvLK) information flow [51]. Results are shown for the two benchmark cases from Fig. 2: (top) System with source-dependent causality and (bottom) system with target-dependent causality.

Comparison with other methods

We compare the results for the two benchmark cases shown in Fig. 2 using five different methods that estimate time-varying measures of influence between variables: a state-dependent variant of conditional transfer entropy (CTE), local transfer entropy (LTE) [87], specific transfer entropy (STE) [45], local Granger causality (LGC) [53], and time-varying Liang–Kleeman (TvLK) information flow [51]. LTE and STE are designed to quantify interactions among variables in dynamical systems without necessarily asserting an underlying causal structure, whereas LGC and TvLK are explicitly formulated as causal inference methods.

For LTE and STE, we employ the implementation by Darmon and Rapp [45], where LTE estimates are obtained using a k -nearest neighbors (k -NN) estimator with $k = 5$ neighbors [88]. Following the authors' recommendations, the STE estimates are derived from LTE using $k_{\text{reg}} = \lfloor \sqrt{T} \rfloor$, where T denotes the total number of samples. This choice controls the degree of smoothing and balances the bias-variance trade-off in the estimation of STE. For LGC, we implement the formulation proposed in Ref. [53] and validate our implementation by reproducing the toy examples presented in that study. For TvLK, we adopt the implementation described in Ref. [51], applying a uniformly weighted moving average (UWMA) filter with a window length of 10 points, and a sliding estimation window of 100 points. We verified that varying the estimation window length between 30 and 150 points did not affect the conclusions. Further details on each method and their implementation are provided in the Supplementary Materials.

We applied the five methods described above to the source- and target-dependent causality cases considered in the validation section. The results are visualized in Fig. 7, which shows the time evolution of q_1 (for source-dependent causality) and q_2^+ (for target-dependent causality). The signals are colored according to the value of CTE, LTE, STE, LGC, or TvLK. The expectation is that, for source-dependent causality, only instances where $q_1 > 0$ should exhibit strong causal influence from q_1 to q_2^+ , while for target-dependent causality, strong causal influence from q_1 to q_2^+ should be detected when $q_2^+ > 0$. However, the results reveal that none of the methods consistently exhibit the expected trends.

Among the methods evaluated, STE performs best in the source-dependent case, where larger values are observed for $q_1 > 0$, consistent with the dependence $q_1 \rightarrow q_2^+$ when $q_1 > 0$. However, STE remains nonzero even for $q_1 < 0$, where q_2 is also required to predict q_2^+ , indicating that STE does not fully resolve the dependence structure of the system. In the case of target-dependent causality, STE projects the results from LTE onto the future states of the target variable. As a consequence, the results for STE in the target-dependent case are not consistent with the functional dependencies inherent to the system.

LTE, LGC, and TvLK do not offer clear evidence for either source-dependent or target-dependent interaction. This outcome is not surprising for LGC and TvLK, as these methods are constrained by the assumptions of linearity and Gaussianity. In the present cases, the dependencies between variables are nonlinear, complicating the detection of the *a priori* causal dependencies. TvLK estimates time-varying causal relationships by applying a square-root Kalman filter to track the evolution of covariance matrices based on the original Liang–Kleeman information flow formulation. While this approach can be effective when the underlying probability distributions evolve smoothly over time, it may lack the sensitivity required to detect abrupt, short-time-scale changes in causality, such as those presented in the benchmark cases shown in Fig. 2.

Finally, in CTE, all redundant, unique, and synergistic causal contributions are aggregated into a single measure, obscuring the interpretation of the interactions between variables. For example, CTE fails to reveal in the source-dependent case that when $q_1 > 0$, the variable Q_1 uniquely determines the future of Q_2 , whereas when $q_1 < 0$, both

TABLE I. Comparison of SURD and the proposed method. A checkmark (\checkmark) indicates the feature is supported; a cross mark (\times) indicates it is not.

Capability	SURD	Proposed method
Based on forward-in-time propagation of information	\checkmark	\checkmark
Interaction-specific decomposition (redundant, unique, synergistic)	\checkmark	\checkmark
State-dependent causality	\times	\checkmark
Causal vs. non-causal state distinction	\times	\checkmark
Temporal evolution of causality	\times	\checkmark

Q_1 and Q_2 are jointly required to predict Q_2^+ . As a result, the detailed structure of variable interactions is lost when all forms of causality are conflated into a single quantity.

In closing this comparison, it is important to emphasize that the results produced by the methods discussed above are not incorrect, nor do we suggest otherwise. Each of these approaches offers valuable insights into interactions and information dynamics in complex systems, within the scope of their respective assumptions and intended applications, as demonstrated in prior studies. The key conclusion here is that these methods are not specifically designed to quantify the types of interaction and causality targeted in the present work. In contrast, our framework is explicitly tailored to address these challenges, particularly in systems characterized by source- or target-dependent causal structures.

We conclude this section by comparing the current method with our previous approach, SURD. The key similarities and differences are summarized in Table I. Both SURD and the proposed method are grounded in the information-theoretic principle of forward-in-time information propagation. Similarly, both frameworks decompose causality according to the type of interaction—namely, redundant, unique, and synergistic contributions. However, SURD computes these causal contributions as averages over all system states, whereas the proposed method performs the decomposition at the level of individual states, both for the source and target variables. As a consequence, SURD may yield similar averaged causal attributions for systems with fundamentally different underlying state-dependent dynamics. This limitation is illustrated in Figs. 2 and 3, where source-dependent and target-dependent causal effects are discernible only through the proposed method. In addition, the current approach distinguishes between causal and non-causal components, with the latter corresponding to states whose knowledge increases uncertainty about the target. SURD does not offer this distinction. Another limitation of SURD is its inability to capture the temporal evolution of causality. In contrast, the proposed method allows for tracking such evolution, provided sufficient statistical support exists across the observed states. Overall, the current method enables a more detailed and state-resolved characterization of causal structure in complex systems, going beyond what is possible with existing causal inference frameworks.

Limitations

We discuss here some of the limitations of our method. First, the proposed causal framework is observational in nature—i.e., it infers causal relationships from statistical dependencies in time-resolved data without requiring interventions. While this allows for broad applicability to real-world systems where interventions are infeasible or unethical, it also introduces limitations: observational causality may not coincide with interventional or counterfactual definitions of causality, and can be confounded by hidden variables or latent dynamics.

Another key limitation is that the method is data-intensive, as it requires the estimation of probability distributions, which becomes increasingly challenging as the dimensionality of the dataset grows. This challenge primarily affects the quantification of synergistic causalities, which rely on higher-dimensional joint distributions compared to redundant and unique contributions. To mitigate this, our method permits restricting the order of synergistic interactions (e.g., limiting to pairwise synergistic causalities in systems with many source variables and limited data). The unestimated higher-order contributions can then be interpreted as a causality leak, thanks to the additive structure of our causal decomposition. Furthermore, recent advances in distribution estimation techniques, such as transport maps [89] and flow matching [90], can enhance the applicability of our approach to high-dimensional systems.

Another important consideration is that our method requires defining a partition of the state space of the system. In this work, we employed a uniform partition based on the values of the signals. However, in other contexts, phase-space partitioning may be tailored to the specific research question—for example, by distinguishing between regimes with and without extreme events. We believe that recent advances in quantized deep autoencoders, such as vector-quantized variational autoencoders [91] and finite scalar quantization methods [92], offer promising tools for constructing more efficient and data-adaptive state-space partitions.

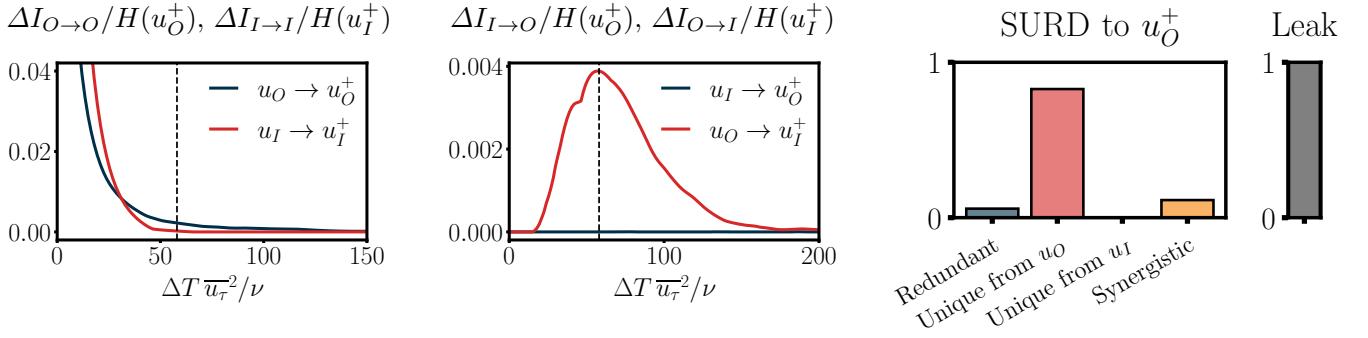


FIG. 8. Left: Self-induced ($\Delta I_{O \rightarrow O}^U / H(u_O^+)$, $\Delta I_{I \rightarrow I}^U / H(u_I^+)$) and cross-induced ($\Delta I_{I \rightarrow O}^U / H(u_O^+)$, $\Delta I_{O \rightarrow I}^U / H(u_I^+)$) SURD causalities as a function of the time lag. Time is normalized using inner units based on the streamwise-averaged friction velocity \bar{u}_τ and viscosity ν and causalities with the Shannon information of the corresponding target variable, i.e. $H(u)$. The dashed vertical line denotes the time lag selected for the causal analysis. Right: redundant ($\Delta I_{O \rightarrow I}^R$), unique ($\Delta I_{(\cdot) \rightarrow I}^U$), and synergistic ($\Delta I_{O \rightarrow I}^S$) causalities from SURD to the outer-layer motions, $u_O^+ = u_O(t + \Delta T)$.

Data for benchmark cases

The source-dependent and target-dependent benchmark cases comprise a system of two variables Q_1 and Q_2 at discrete times $t_n = n$. The system is initially set to $q_1(1) = q_2(1) = 0$. A stochastic forcing, represented by W_i , acts on Q_i and follows a Gaussian distribution. The computation of causalities is performed for a time lag of $\Delta T = 1$ using 75 uniform bins per variable. The integration of the system is carried out over 10^8 time steps, with the first 10,000 steps excluded from the analysis to avoid transient effects. In the analysis presented in Fig. 2, the causality leak for $q_2^+ = q_2(n+1)$ is 53% in the source-dependent case and 54% in the target-dependent case. This indicates that only 47% and 46% of the causality to Q_2^+ can be accounted for using Q_1 and Q_2 alone. The remaining unexplained part is due to the influence of W_2 , which also affects Q_2^+ but is not included in the analysis. An additional study with a system of three variables Q_1 , Q_2 , and Q_3 with similar functional relationships is reported in the Supplementary Materials, where we also discuss the structure of the state-dependent non-causal components.

Data for turbulent boundary layer

The data used for analyzing inner/outer interactions in a turbulent boundary layer were obtained for a dataset of a direct numerical simulation of the Navier–Stokes equations [93], where all temporal and spatial scales were resolved. The dataset covers the range of friction Reynolds number from $Re_\tau \approx 292$ to 729, with 10,000 flowfields stored spanning 26 eddy-turnover times (after transients) based on $\bar{\delta}/\bar{u}_\tau$, where the symbol $\bar{(\cdot)}$ denotes the average along the streamwise direction x . Further details about the numerical setup can be found in Towne et al. [67]. The data used in this study includes the temporal evolution of the streamwise velocity at two different wall-normal locations: $y_I = 4\nu/u_\tau$ (for the inner layer) and $y_O = 0.3\delta$ (for the outer layer). The analysis was conducted at every streamwise location within the boundary layer, leveraging the self-similarity of the flow, which results in 2×10^7 samples for causal analysis. The time lag utilized to evaluate causality was $\Delta T = 60\nu/u_\tau^2$, which corresponds to the time lag for maximum cross-induced unique causality. At this time lag, the causality leak to the inner-layer velocity u_I^+ is 99%. This high value is expected, given that the vast majority of the degrees of freedom of the system (associated with the full turbulent flow field) are not accounted for [27].

The temporal evolution of the self-induced causalities ($\Delta I_{O \rightarrow O}^U$ and $\Delta I_{I \rightarrow I}^U$) and cross-induced causalities ($\Delta I_{I \rightarrow O}^U$ and $\Delta I_{O \rightarrow I}^U$) is presented in Fig. 8. The causalities directed toward the outer-layer velocity u_O^+ at the same time lag are also shown in Fig. 8, revealing no significant contribution from the inner-layer velocity u_I . Finally, the probability distributions were discretized using 50 uniform bins per variable. Additional tests conducted with half and double the number of bins produced no significant differences in the results.

Data for climate science

The climate time series data are regional averages from the reanalysis for the period 1948–2012 with 780 months. WPAC denotes monthly surface pressure anomalies in the West Pacific, and CPAC and EPAC surface air temperature

anomalies in the Central and East Pacific. Anomalies are taken with respect to the whole period, which is publicly available in <https://psl.noaa.gov/data/gridded/data.ncep.reanalysis.html>. The time lag utilized to evaluate causality is $\Delta T = 4$ months, which maximized the unique causality across different variables and is within the range of lags analyzed in Runge *et al.* [16]. At this time lag, the causality leak to WPAC($t + \Delta T$) and CPAC($t + \Delta T$) is 79% and 65%, respectively, which quantifies the effect from variables not included in the analysis. The joint distribution of the variables is estimated using a k -NN density estimator with $k = 7$ neighbors. Alternative values of k within the range [4,8] were tested, showing no significant impact on the results. To quantify uncertainty in the results, we performed 50 bootstrap resamplings on the available data. Additionally, to assess statistical significance, we applied 1000 random permutations to compute the p -values of the redundant and unique causalities reported in Fig. 5, both of which yielded p -values of zero. In contrast, when including synergistic causalities, the p -values increased significantly, leading us to conclude that the latter causalities lack sufficient statistical support with the available data. More details about the statistical significance of the results for this case are shown in the Supplementary Materials, where we include the results for reduced sample sizes.

CODE AVAILABILITY

The codes developed for this work are available at: <https://github.com/ALD-Lab/SURD-states>.

ACKNOWLEDGEMENTS

The authors would like to thank Jakob Runge for providing the data for the climate science application and Gonzalo Arranz for his contributions to this work. This work was supported by the National Science Foundation under Grant No. 2140775 and MISTI Global Seed Funds and UPM. Á. M.-S. received the support of a fellowship from the "la Caixa" Foundation (ID 100010434). The fellowship code is LCF/BQ/EU22/11930094. The authors acknowledge the MIT SuperCloud and Lincoln Laboratory Supercomputing Center for providing HPC resources that have contributed to the research results reported within this paper.

- [1] M. Bunge, *Causality and Modern Science* (Dover Publications, New York, 1979).
- [2] X. Wu, Y. M. Okumura, and P. N. DiNezio, What controls the duration of el niño and la niña events?, *Journal of Climate* **32**, 5941 (2019).
- [3] D. Cordes, V. M. Haughton, K. Arfanakis, G. J. Wendt, P. A. Turski, C. H. Moritz, M. A. Quigley, and M. E. Meyerand, Mapping functionally related regions of brain with functional connectivity mr imaging, *American journal of neuroradiology* **21**, 1636 (2000).
- [4] J. Runge, A. Gerhardus, G. Varando, V. Eyring, and G. Camps-Valls, Causal inference for time series, *Nat. Rev. Earth Environ.* **4**, 487 (2023).
- [5] A. Razi and K. J. Friston, The connected brain: causality, models, and intrinsic dynamics, *IEEE Signal Process. Mag.* **33**, 14 (2016).
- [6] S. Z. Chiou-Wei, C. F. Chen, and Z. Zhu, Economic growth and energy consumption revisited — evidence from linear and nonlinear granger causality, *Energy Econ.* **30**, 3063 (2008).
- [7] K. J. Rothman and S. Greenland, Causation and causal inference in epidemiology, *Am. J. Public Health* **95**, S144 (2005).
- [8] P. Hedström and P. Ylikoski, Causal mechanisms in the social sciences, *Annu. Rev. Sociol.* **36**, 49 (2010).
- [9] A. Lozano-Durán, H. J. Bae, and M. P. Encinar, Causality of energy-containing eddies in wall turbulence, *J. Fluid Mech.* **882**, A2 (2020).
- [10] Á. Martínez-Sánchez, E. López, S. Le Clainche, A. Lozano-Durán, A. Srivastava, and R. Vinuesa, Causality analysis of large-scale structures in the flow around a wall-mounted square cylinder, *J. Fluid Mech.* **967**, A1 (2023).
- [11] G. Sugihara, R. May, H. Ye, C. hao Hsieh, E. Deyle, M. Fogarty, and S. Munch, Detecting causality in complex ecosystems, *Science* **338**, 496 (2012).
- [12] H. Ye, E. R. Deyle, L. J. Gilarranz, and G. Sugihara, Distinguishing time-delayed causal interactions using convergent cross mapping, *Sci. Rep.* **5**, 14750 (2015).
- [13] A. T. Clark, H. Ye, F. Isbell, E. R. Deyle, J. Cowles, G. D. Tilman, and G. Sugihara, Spatial convergent cross mapping to detect causal relationships from short time series, *Ecology* **96**, 1174 (2015).
- [14] E. D. Brouwer, A. Arany, J. Simm, and Y. Moreau, Latent Convergent Cross Mapping, in *International Conference on Learning Representations* (2021).
- [15] P. J. E. Javier, causal-ccm: a Python implementation of Convergent Cross Mapping (2021).

- [16] J. Runge, P. Nowack, M. Kretschmer, S. Flaxman, and D. Sejdinovic, Detecting and quantifying causal associations in large nonlinear time series datasets, *Sci. Adv.* **5**, eaau4996 (2019).
- [17] J. Runge, Discovering contemporaneous and lagged causal relations in autocorrelated nonlinear time series datasets, in *Conference on Uncertainty in Artificial Intelligence* (PMLR, 2020) pp. 1388–1397.
- [18] A. Gerhardus and J. Runge, High-recall causal discovery for autocorrelated time series with latent confounders, in *Advances in Neural Information Processing Systems*, Vol. 33, edited by H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin (Curran Associates, Inc., 2020) pp. 12615–12625.
- [19] E. Saggioro, J. de Wiljes, M. Kretschmer, and J. Runge, Reconstructing regime-dependent causal relationships from observational time series, *Chaos* **30**, 113115 (2020).
- [20] C. W. J. Granger, Investigating causal relations by econometric models and cross-spectral methods, *Econometrica* **37**, 424 (1969).
- [21] K. A. Murphy and D. S. Bassett, Information decomposition in complex systems via machine learning, *Proceedings of the National Academy of Sciences (PNAS)* **121**, e2312988121 (2024).
- [22] J. Pearl, *Causality: Models, Reasoning, and Inference* (Cambridge University Press, New York, 2000).
- [23] M. Eichler, Causal inference with multiple time series: principles and problems, *Philos. Trans. R. Soc. A* **371**, 20110613 (2013).
- [24] G. Tononi and O. Sporns, Measuring information integration, *BMC Neuroscience* **4**, 31 (2003).
- [25] E. P. Hoel, L. Albantakis, and G. Tononi, Quantifying causal emergence shows that macro can beat micro, *Proceedings of the National Academy of Sciences* **110**, 19790 (2013).
- [26] E. P. Hoel, L. Albantakis, W. Marshall, and G. Tononi, Can the macro beat the micro? integrated information across spatiotemporal scales, *Neuroscience of Consciousness* **2016**, niw012 (2016).
- [27] Á. Martínez-Sánchez, G. Arranz, and A. Lozano-Durán, Decomposing causality into its synergistic, unique, and redundant components, *Nature Communications* **15**, 9296 (2024).
- [28] C. E. Shannon, A mathematical theory of communication, *The Bell System Technical Journal* **27**, 379 (1948).
- [29] R. Landauer, The physical nature of information, *Physics Letters A* **217**, 188 (1996).
- [30] A. Lozano-Durán and G. Arranz, Information-theoretic formulation of dynamical systems: Causality, modeling, and control, *Phys. Rev. Res.* **4**, 023195 (2022).
- [31] J. Massey, Causality, feedback and directed information, in *Proc. 1990 Int. Symp. on Infom. Theory and its Applications* (1990) pp. 27–30.
- [32] G. Kramer, *Directed information for channels with feedback*, PhD Thesis, ETH Zürich, Zürich (1998).
- [33] T. Schreiber, Measuring information transfer, *Phys. Rev. Lett.* **85**, 461 (2000).
- [34] P. Verdes, Assessing causality from multivariate time series, *Phys. Rev. E* **72**, 026222 (2005).
- [35] J. T. Lizier, M. Prokopenko, and A. Y. Zomaya, Local information transfer as a spatiotemporal filter for complex systems, *Phys. Rev. E* **77**, 026110 (2008).
- [36] L. Barnett, A. B. Barrett, and A. K. Seth, Granger causality and transfer entropy are equivalent for gaussian variables, *Phys. Rev. Lett.* **103**, 238701 (2009).
- [37] J. T. Lizier, M. Prokopenko, and A. Y. Zomaya, Information modification and particle collisions in distributed computation, *Chaos* **20** (2010).
- [38] T. Bossomaier, L. Barnett, M. Harré, and J. T. Lizier, *An Introduction to Transfer Entropy: Information Flow in Complex Systems*, 1st ed. (Springer International Publishing, Cham, 2016).
- [39] X. S. Liang and R. Kleeman, Information transfer between dynamical system components, *Phys. Rev. Lett.* **95**, 244101 (2006).
- [40] X. S. Liang, Information flow and causality as rigorous notions ab initio, *Phys. Rev. E* **94**, 052201 (2016).
- [41] X. S. Liang, Information flow within stochastic dynamical systems, *Phys. Rev. E* **78**, 031113 (2008).
- [42] X. S. Liang, The Liang-Kleeman information flow: Theory and applications, *Entropy* **15**, 327 (2013).
- [43] W. Wang, A. Lozano-Durán, R. Helmig, and X. Chu, Spatial and spectral characteristics of information flux between turbulent boundary layers and porous media, *J. Fluid Mech.* **949**, A16 (2022).
- [44] J. T. Lizier, JIDT: An Information-Theoretic Toolkit for Studying the Dynamics of Complex Systems, *Front. Robot. AI* **1**, 10.3389/frobt.2014.00011 (2014).
- [45] D. Darmon and P. E. Rapp, Specific transfer entropy and other state-dependent transfer entropies for continuous-state input-output systems, *Physical Review E* **96**, 022121 (2017).
- [46] M. Martini, T. A. Kranz, T. Wagner, and K. Lehnertz, Inferring directional interactions from transient signals with symbolic transfer entropy, *Physical Review E* **83**, 011919 (2011).
- [47] M. Wibral, N. Pampu, V. Priesemann, F. Siebenhühner, H. Seiwert, M. Lindner, J. T. Lizier, and R. Vicente, Measuring information-transfer delays, *PLoS ONE* **8**, e55809 (2013).
- [48] P. Wollstadt, M. Martinez-Zarzuela, R. Vicente, F. J. Diaz-Pernas, and M. Wibral, Efficient transfer entropy analysis of non-stationary neural time series, *PLoS ONE* **9**, e102833 (2014).
- [49] G. Gómez-Herrero, W. Wu, K. Rutanen, M. Soriano, G. Pipa, and R. Vicente, Assessing coupling dynamics from an ensemble of time series, *Entropy* **17**, 1958 (2015).
- [50] J. Zhu, M. Chen, J. Lu, K. Zhao, E. Cui, Z. Zhang, and H. Wan, A fast and efficient ensemble transfer entropy and applications in neural signals, *Entropy* **24**, 1118 (2022).
- [51] D. F. T. Hagan, G. Wang, X. S. Liang, and H. A. J. Dolman, A time-varying causality formalism based on the liang–kleeman information flow for analyzing directed interactions in nonstationary climate systems, *Journal of Climate* **32**, 7521 (2019).

- [52] F. Zhou, D. F. T. Hagan, G. Wang, X. S. Liang, S. Li, Y. Shao, E. Yeboah, and X. Wei, Estimating time-dependent structures in a multivariate causality for land–atmosphere interactions, *Journal of Climate* **37**, 1853 (2024).
- [53] S. Stramaglia, T. Scagliarini, Y. Antonacci, and L. Faes, Local granger causality, *Physical Review E* **103**, L020102 (2021).
- [54] S. Stramaglia, L. Faes, J. M. Cortes, and D. Marinazzo, Disentangling high-order effects in the transfer entropy, *Phys. Rev. Res.* **6**, L032007 (2024).
- [55] J. Kořenek, P. Sanda, and J. Hlinka, Higher order definition of causality by optimally conditioned transfer entropy, *Phys. Rev. E* **111**, L042302 (2025).
- [56] Á. Martínez-Sánchez, G. Arranz, and A. Lozano-Durán, Decomposing causality in its synergistic, unique, and redundant components (2024), SURD: Synergistic-Unique-Redundant Decomposition of causality. <https://doi.org/10.5281/zenodo.13750918>.
- [57] I. Marusic, R. Mathis, and N. Hutchins, Predictive model for wall-bounded turbulent flow, *Science* **329**, 193 (2010).
- [58] K. A. Flack, M. P. Schultz, and T. A. Shapiro, Experimental support for townsend’s reynolds number similarity hypothesis on rough walls, *Phys. Fluids* **17**, 035102 (2005).
- [59] O. Flores and J. Jiménez, Effect of wall-boundary disturbances on turbulent channel flows, *J. Fluid Mech.* **566**, 357 (2006).
- [60] N. Hutchins and I. Marusic, Evidence of very long meandering features in the logarithmic region of turbulent boundary layers, *J. Fluid Mech.* **579**, 1 (2007).
- [61] R. Mathis, N. Hutchins, and I. Marusic, Large-scale amplitude modulation of the small-scale structures in turbulent boundary layers, *J. Fluid Mech.* **628**, 311 (2009).
- [62] B. Busse and A. Sandham, Parametric forcing approach to rough-wall turbulent channel flow, *J. Fluid Mech.* **712**, 169 (2012).
- [63] Y. Mizuno and J. Jiménez, Wall turbulence without walls, *J. Fluid Mech.* **723**, 429 (2013).
- [64] D. Chung, J. P. Monty, and A. Ooi, An idealised assessment of Townsend’s outer-layer similarity hypothesis for wall turbulence, *J. Fluid Mech.* **742**, R3 (2014).
- [65] A. Lozano-Durán and H. J. Bae, Characteristic scales of Townsend’s wall-attached eddies, *J. Fluid Mech.* **868**, 698 (2019).
- [66] A. A. Townsend, *The structure of turbulent shear flow* (Cambridge University Press, Cambridge, 1976).
- [67] A. Towne, S. T. M. Dawson, G. A. Brès, A. Lozano-Durán, T. Saxton-Fox, A. Parthasarathy, A. R. Jones, H. Biler, C.-A. Yeh, H. D. Patel, and K. Taira, A database for reduced-complexity modeling of fluid flows, *AIAA Journal* **61**, 2867 (2023).
- [68] J. Jiménez and A. Pinelli, The autonomous cycle of near-wall turbulence, *Journal of Fluid Mechanics* **389**, 335–359 (1999).
- [69] J. Jiménez, Cascades in wall-bounded turbulence, *Annual Review of Fluid Mechanics* **44**, 27 (2012).
- [70] T. DiLiberto, M. L’Heureux, E. Becker, N. Johnson, and R. Lindsey, The walker circulation: Enso’s atmospheric buddy (2014), published on NOAA Climate.gov.
- [71] M. Watanabe, T. Iwakiri, Y. Dong, and S. M. Kang, Two competing drivers of the recent walker circulation trend, *Geophysical Research Letters* **50**, e2023GL105332 (2023).
- [72] J. Bjerknes, Atmospheric teleconnections from the equatorial pacific, *Monthly Weather Review* **97**, 163 (1969).
- [73] W. Cai, M. J. McPhaden, A. M. Grimm, R. R. Rodrigues, A. S. Taschetto, R. D. Garreaud, B. Dewitte, G. Poveda, Y.-G. Ham, A. Santoso, *et al.*, Climate impacts of the el niño–southern oscillation on south america, *Nature Reviews Earth & Environment* **1**, 215 (2020).
- [74] C. Wunsch, Geophysical interplays: El Niño, La Niña, and the Southern Oscillation, *Science* **248**, 904 (1990).
- [75] X. Pan, T. Li, and J. Yu, Change of el niño onset location around 1970, *npj Climate and Atmospheric Science* **7**, 163 (2024).
- [76] A. Timmermann, S.-I. An, J.-S. Kug, F.-F. Jin, W. Cai, A. Capotondi, K. M. Cobb, M. Lengaigne, M. J. McPhaden, M. F. Stuecker, *et al.*, El niño–southern oscillation complexity, *Nature* **559**, 535 (2018).
- [77] S. Kullback and R. A. Leibler, On information and sufficiency, *Ann. Math. Stat.* **22**, 79 (1951).
- [78] J. Kreer, A question of terminology, *IEEE Trans. Inf. Theory* **3**, 208 (1957).
- [79] M. R. DeWeese and M. Meister, How to measure the information gained from one symbol, *Netw. Comput. Neural Syst.* **10**, 325 (1999).
- [80] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley-Interscience, USA, 2006).
- [81] P. L. Williams and R. D. Beer, Nonnegative decomposition of multivariate information, *arXiv preprint arXiv:1004.2515* (2010).
- [82] V. Griffith and C. Koch, Quantifying synergistic mutual information, in *Guided Self-Organization: Inception* (Springer, Berlin, Heidelberg, 2014) pp. 159–190.
- [83] V. Griffith and T. Ho, Quantifying redundant information in predicting a target random variable, *Entropy* **17**, 4644 (2015).
- [84] R. A. Ince, Measuring multivariate redundant information with pointwise common change in surprisal, *Entropy* **19**, 318 (2017).
- [85] A. J. Gutknecht, M. Wibral, and A. Makkeh, Bits and pieces: understanding information decomposition from part-whole relationships and formal logic, *Proc. R. Soc. A* **477**, 20210110 (2021).
- [86] A. Kolchinsky, A novel approach to the partial information decomposition, *Entropy* **24**, 403 (2022).
- [87] J. T. Lizier, Measuring the dynamics of information processing on a local scale in time and space, in *Local Measures of Information Storage and Transfer in Complex Systems* (Springer, 2014) pp. 1–21.
- [88] A. Kraskov, H. Stögbauer, and P. Grassberger, Estimating mutual information, *Phys. Rev. E* **69**, 066138 (2004).
- [89] R. Baptista, Y. Marzouk, and O. Zahm, On the representation and learning of monotone triangular transport maps, *Found. Comput. Math.* (2023).
- [90] T. Fjelde, E. Mathieu, and V. Dutordoir, An introduction to flow matching (2024).

- [91] A. Van Den Oord, O. Vinyals, *et al.*, Neural discrete representation learning, *Advances in neural information processing systems* **30** (2017).
- [92] F. Mentzer, D. Minnen, E. Agustsson, and M. Tschannen, Finite scalar quantization: VQ-VAE made simple, in *The Twelfth International Conference on Learning Representations* (2024).
- [93] A. S. Towne and A. Lozano-Durán, Turbulent boundary layer direct numerical simulations [Data set] (2022), University of Michigan - Deep Blue Data.

AUTHOR CONTRIBUTIONS STATEMENT

Á. M.-S.: Methodology, Software, Validation, Investigation, Data Curation, Writing – Original Draft, Writing – Review & Editing, Visualization. A. L.-D.: Ideation, Methodology, Writing – Review & Editing, Supervision, Resources, Funding acquisition.

COMPETING INTERESTS

The authors declare no competing interests.

Supplementary Material

Observational causality by states

Álvaro Martínez-Sánchez¹ and Adrián Lozano-Durán^{1,2}

¹*Department of Aeronautics and Astronautics, Massachusetts Institute of Technology
Cambridge, MA 02139, USA*

²*Graduate Aerospace Laboratories, California Institute of Technology, Pasadena, CA 91125, USA*

Contents

S1 Method formulation	2
S1.1 Fundamentals of information theory	2
S1.2 SURD: Synergistic-Unique-Redundant Decomposition of causality	3
S1.3 State-dependent decomposition of causality	6
S1.3.1 State-dependent redundant causalities	6
S1.3.2 State-dependent unique causalities	8
S1.3.3 State-dependent synergistic causalities	9
S1.4 Alternative state-dependent decomposition of causality	10
S1.4.1 Alternative definition I: Causality as positive and negative information increments	11
S1.4.2 Alternative definition II: Causality as absolute value of information increments	11
S2 Verification and validation	13
S2.1 Multivariate correlated Gaussian variables	13
S2.2 Source-dependent unique causality	13
S2.3 Target-dependent unique causality	15
S2.4 Source-dependent synergistic causality	16
S2.5 Coupled Rössler–Lorenz system	16
S3 Analysis of non-causal contributions	18
S3.1 Source-dependent and target-dependent systems	19
S3.2 Turbulent boundary layer and Walker circulation	20
S4 Other methods for quantifying state-dependent influence among variables	22
S4.1 State-dependent conditional transfer entropy	22
S4.2 Local transfer entropy	22
S4.3 Specific transfer entropy	23
S4.4 Local Granger causality	24
S4.5 Time-varying Liang–Kleeman information flow	25
S4.6 Comparison of methods	26
S5 Climate science application	29
S5.1 Effect of the number of samples	29

S1 Method formulation

S1.1 Fundamentals of information theory

Consider N quantities of interest at time t represented by the vector of observable variables $\mathbf{Q} = [Q_1(t), Q_2(t), \dots, Q_N(t)]$. We treat \mathbf{Q} as a random variable and consider a finite partition of the observable phase space $D = \{D_1, D_2, \dots, D_{N_D}\}$, where N_D is the number of cells, such that $D = \bigcup_{i=1}^{N_D} D_i$ and $D_i \cap D_j = \emptyset$ for all $i \neq j$ (i.e., disjoint cover of the observable phase space D). We use upper case Q to denote the random variable itself; and lower case q to denote a particular state contained in one D_i (also referred to as a value or an event) of Q . The probability of finding the system at state D_i at time t is $p(\mathbf{Q}(t) \in D_i)$, that in general depends on the partition D . For simplicity, we refer to the latter probability as $p(q)$.

The information contained in the variable \mathbf{Q} is given by [28]:

$$H(\mathbf{Q}) = \sum_q -p(q) \log_2[p(q)] \geq 0, \quad (\text{S1})$$

where the summation is over all the states of \mathbf{Q} . The quantity H is the Shannon information or entropy [28]. The units of H are set by the base chosen, in this case ‘bits’ for base 2. For example, consider a fair coin with $Q \in \{\text{heads, tails}\}$ such that $p(\text{heads}) = p(\text{tails}) = 0.5$. The information of the system ‘tossing a fair coin n times’ is $H = -\sum 0.5^n \log_2(0.5^n) = n$ bits, where the summation is carried out across all possible outcomes (namely, 2^n). If the coin is completely biased towards heads, $p(\text{heads}) = 1$, then $H = 0$ bits (taking $0 \log 0 = 0$), i.e., no information is gained as the outcome was already known before tossing the coin. The Shannon information can also be interpreted in terms of uncertainty: $H(\mathbf{Q})$ is the average number of bits required to unambiguously determine \mathbf{Q} . H is maximum when all the possible outcomes are equiprobable (indicating a high level of uncertainty in the state of the system) and zero when the process is completely deterministic (indicating no uncertainty in the outcome).

The Shannon information of \mathbf{Q} conditioned on another variable \mathbf{Q}' is defined as:

$$H(\mathbf{Q}|\mathbf{Q}') = \sum_{q,q'} -p(q, q') \log_2[p(q|q')]. \quad (\text{S2})$$

where $p(q|q') = p(q, q')/p(q')$ with $p(q') \neq 0$ is the conditional probability distribution, and $p(q') = \sum_q p(q, q')$ is the marginal probability distribution of \mathbf{Q}' . It is useful to interpret $H(\mathbf{Q}|\mathbf{Q}')$ as the uncertainty in the variable \mathbf{Q} after conducting the ‘measurement’ of \mathbf{Q}' . If \mathbf{Q} and \mathbf{Q}' are independent random variables, then $H(\mathbf{Q}|\mathbf{Q}') = H(\mathbf{Q})$, i.e., knowing \mathbf{Q}' does not reduce the uncertainty in \mathbf{Q} . Conversely, $H(\mathbf{Q}|\mathbf{Q}') = 0$ if knowing \mathbf{Q}' implies that \mathbf{Q} is completely determined. Finally, the mutual information between the random variables \mathbf{Q} and \mathbf{Q}' is

$$I(\mathbf{Q}; \mathbf{Q}') = H(\mathbf{Q}) - H(\mathbf{Q}|\mathbf{Q}') = H(\mathbf{Q}') - H(\mathbf{Q}'|\mathbf{Q}), \quad (\text{S3})$$

which is a symmetric measure $I(\mathbf{Q}; \mathbf{Q}') = I(\mathbf{Q}'; \mathbf{Q})$ representing the information shared among the variables \mathbf{Q} and \mathbf{Q}' . Figure S1 depicts the relationship between the Shannon information, conditional Shannon information, and mutual information.

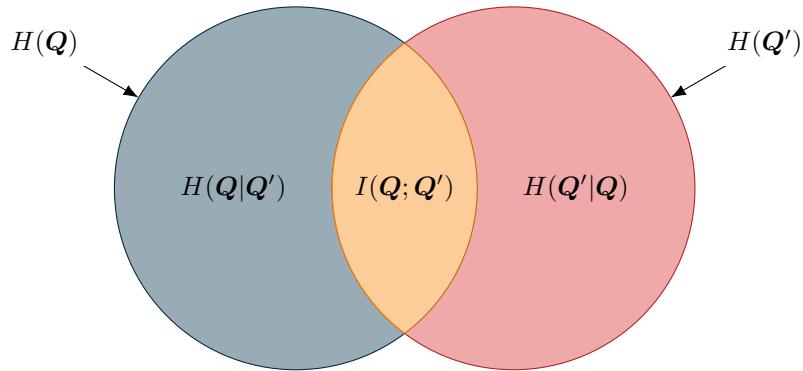


Figure S1: Venn diagram of the Shannon information, conditional Shannon information and mutual information between two random variables \mathbf{Q} and \mathbf{Q}' .

S1.2 SURD: Synergistic-Unique-Redundant Decomposition of causality

We provide a brief summary of the SURD formulation. For full derivations and discussion, the reader is referred to Ref. [27]. SURD quantifies causality as the increase in information (ΔI) about Q_j^+ obtained from observing individual components or groups of components from \mathbf{Q} . The information in Q_j^+ is measured by the Shannon entropy. Using the principle of forward-in-time propagation of information (i.e., information only flows toward the future) [30], $H(Q_j^+)$ can be decomposed as the sum of all causal contributions from the past and present:

$$H(Q_j^+) = \sum_{\mathbf{i} \in \mathcal{P}} \Delta I_{\mathbf{i} \rightarrow j}^R + \sum_{i=1}^N \Delta I_{i \rightarrow j}^U + \sum_{\mathbf{i} \in \mathcal{P}} \Delta I_{\mathbf{i} \rightarrow j}^S + \Delta I_{\text{leak} \rightarrow j}, \quad (\text{S4})$$

where $\Delta I_{\mathbf{i} \rightarrow j}^R$, $\Delta I_{i \rightarrow j}^U$, and $\Delta I_{\mathbf{i} \rightarrow j}^S$ are the redundant, unique, and synergistic causalities, respectively, from the observed variables to Q_j^+ , and $\Delta I_{\text{leak} \rightarrow j}$ is the causality from unobserved variables, referred to as the causality leak. Unique causalities are associated with individual components of \mathbf{Q} , whereas redundant and synergistic causalities arise from groups of variables from \mathbf{Q} . Hence, the set \mathcal{P} contains all combinations involving more than one variable. For instance, Eq. (1equation.0.1) can be expanded for $N = 3$ as

$$H(Q_j^+) \equiv \Delta I_{1 \rightarrow j}^U + \Delta I_{2 \rightarrow j}^U + \Delta I_{3 \rightarrow j}^U + \Delta I_{12 \rightarrow j}^R + \Delta I_{13 \rightarrow j}^R + \Delta I_{23 \rightarrow j}^R \quad (\text{S5a})$$

$$+ \Delta I_{12 \rightarrow j}^S + \Delta I_{13 \rightarrow j}^S + \Delta I_{23 \rightarrow j}^S + \Delta I_{123 \rightarrow j}^R + \Delta I_{123 \rightarrow j}^S + \Delta I_{\text{leak} \rightarrow j}. \quad (\text{S5b})$$

The SURD causalities are constructed using the specific mutual information [79] from $\mathbf{Q}_i = [Q_{i_1}, Q_{i_2}, \dots]$ to a particular event $Q_j^+ = q_j^+$, which is defined as

$$\tilde{i}(Q_j^+ = q_j^+; \mathbf{Q}_i) = \sum_{\mathbf{q}_i} p(\mathbf{q}_i | q_j^+) \log_2 \left(\frac{p(q_j^+ | \mathbf{q}_i)}{p(q_j^+)} \right) \geq 0. \quad (\text{S6})$$

Note that the specific mutual information is a function of the random variable \mathbf{Q}_i (which encompasses all its states) but only a function of one particular state of the target variable (namely, q_j^+). For the sake of simplicity, we will use the notation $\tilde{i}_i(q_j^+) = \tilde{i}(Q_j^+ = q_j^+; \mathbf{Q}_i)$. Similarly to Eq. (S3), the specific mutual information quantifies the dissimilarity between $p(q_j^+)$ and $p(q_j^+ | \mathbf{q})$ but in this case for the particular state $Q_j^+ = q_j^+$. The mutual information between Q_j^+ and \mathbf{Q}_i is recovered by $I(Q_j^+; \mathbf{Q}_i) = \sum_{q_j^+} p(q_j^+) \tilde{i}_i(q_j^+)$.

We now introduce the mathematical expressions for redundant, unique, and synergistic causalities (Fig. S2). These definitions are guided by the following intuition:

- Redundant causality from $\mathbf{Q}_i = [Q_{i_1}, Q_{i_2}, \dots]$ to Q_j^+ is the common causality shared among all the components of \mathbf{Q}_i , where \mathbf{Q}_i is a subset of \mathbf{Q} with two or more components.
- Unique causality from Q_i to Q_j^+ is the causality from Q_i that cannot be obtained from any other individual variable Q_k with $k \neq i$.
- Synergistic causality from $\mathbf{Q}_i = [Q_{i_1}, Q_{i_2}, \dots]$ to Q_j^+ is the causality arising from the joint effect of the variables in \mathbf{Q}_i .
- Redundant and unique causalities must depend only on probability distributions based on Q_i and Q_j^+ , that is, $p(q_i, q_j^+)$. On the other hand, synergistic causality must depend on the joint probability distribution of \mathbf{Q}_i and Q_j^+ , i.e., $p(\mathbf{q}_i, q_j^+)$.

For a given value $Q_j^+ = q_j^+$, the specific redundant, unique, and synergistic causalities are calculated as follows:

1. The specific mutual information are computed for all possible combinations of variables in \mathbf{Q} . This includes specific mutual information of order one ($\tilde{i}_1, \tilde{i}_2, \dots$), order two ($\tilde{i}_{12}, \tilde{i}_{13}, \dots$), order three ($\tilde{i}_{123}, \tilde{i}_{124}, \dots$), and so forth. One example is shown in Fig. S2(a).
2. The tuples containing the specific mutual information of order M , denoted by $\tilde{\mathcal{G}}^M$, are constructed for $M = 1, \dots, N$. The components of each $\tilde{\mathcal{G}}^M$ are organized in ascending order as shown in Fig. S2(b).

3. The specific redundant causality is the increment in information gained about q_j^+ that is common to all the components of $\mathbf{Q}_{\mathbf{j}_k}$ (blue contributions in Fig. S2c):

$$\Delta \tilde{i}_{\mathbf{j}_k}^R = \begin{cases} \tilde{i}_{i_k} - \tilde{i}_{i_{k-1}}, & \text{for } \tilde{i}_{i_k}, \tilde{i}_{i_{k-1}} \in \tilde{\mathcal{G}}^1 \text{ and } k \neq n_1 \\ 0, & \text{otherwise,} \end{cases} \quad (\text{S7})$$

where we take $\tilde{i}_{i_0} = 0$, $\mathbf{j}_k = [j_{k1}, j_{k2}, \dots]$ is the vector of indices satisfying $\tilde{i}_{j_{kl}} \geq \tilde{i}_{i_k}$ for $\tilde{i}_{j_{kl}}, \tilde{i}_{i_k} \in \tilde{\mathcal{G}}^1$, and n_1 is the number of elements in $\tilde{\mathcal{G}}^1$. For instance, in Fig. S2c, the specific redundant causality $\Delta \tilde{i}_{12}^R$ is calculated using \tilde{i}_2 and \tilde{i}_3 . In this case, the indices i_k and i_{k-1} are 1 and 3, respectively, and $\mathbf{j}_k = [1, 2]$ since $\tilde{i}_1 \geq \tilde{i}_{i_k}$ and $\tilde{i}_2 \geq \tilde{i}_{i_k}$.

4. The specific unique causality is the increment in information gained by Q_{i_k} about q_j^+ that cannot be obtained by any other individual variable (red contribution in Fig. S2c):

$$\Delta \tilde{i}_{i_k}^U = \begin{cases} \tilde{i}_{i_k} - \tilde{i}_{i_{k-1}}, & \text{for } i_k = n_1, \tilde{i}_{i_k}, \tilde{i}_{i_{k-1}} \in \tilde{\mathcal{G}}^1 \\ 0, & \text{otherwise.} \end{cases} \quad (\text{S8})$$

5. The specific synergistic causality is the increment in information gained by the combined effect of all the variables in \mathbf{Q}_{i_k} that cannot be gained by other combination of variables $\mathbf{Q}_{\mathbf{j}_k}$ (yellow contributions in Fig. S2c) such that $\tilde{i}_{j_k} \leq \tilde{i}_{i_k}$ for $\tilde{i}_{i_k} \in \tilde{\mathcal{G}}^M$ and $\tilde{i}_{j_k} \in \{\tilde{\mathcal{G}}^1, \dots, \tilde{\mathcal{G}}^M\}$ with $M > 1$ (dotted line in Fig. S2c):

$$\Delta \tilde{i}_{i_k}^S = \begin{cases} \tilde{i}_{i_k} - \tilde{i}_{i_{k-1}}, & \text{for } \tilde{i}_{i_{k-1}} \geq \max\{\tilde{\mathcal{G}}^{M-1}\}, \text{ and } \tilde{i}_{i_k}, \tilde{i}_{i_{k-1}} \in \tilde{\mathcal{G}}^M \\ \tilde{i}_{i_k} - \max\{\tilde{\mathcal{G}}^{M-1}\}, & \text{for } \tilde{i}_{i_k} > \max\{\tilde{\mathcal{G}}^{M-1}\} > \tilde{i}_{i_{k-1}}, \text{ and } \tilde{i}_{i_k}, \tilde{i}_{i_{k-1}} \in \tilde{\mathcal{G}}^M \\ 0, & \text{otherwise.} \end{cases} \quad (\text{S9})$$

6. The specific redundant, unique and synergistic causalities that do not appear in the steps above are set to zero.
 7. The steps (1) to (6) are repeated for all the states of Q_j^+ (Figure S2d).
 8. Redundant, unique, and synergistic causalities are obtained as the expectation of their corresponding specific values with respect to Q_j^+ ,

$$\Delta I_{\mathbf{i} \rightarrow j}^R = \sum_{q_j^+} p(q_j^+) \Delta \tilde{i}_{\mathbf{i}}^R(q_j^+), \quad (\text{S10a})$$

$$\Delta I_{\mathbf{i} \rightarrow j}^U = \sum_{q_j^+} p(q_j^+) \Delta \tilde{i}_{\mathbf{i}}^U(q_j^+), \quad (\text{S10b})$$

$$\Delta I_{\mathbf{i} \rightarrow j}^S = \sum_{q_j^+} p(q_j^+) \Delta \tilde{i}_{\mathbf{i}}^S(q_j^+). \quad (\text{S10c})$$

9. Finally, the average order of the specific causalities with respect to Q_j^+ is defined as

$$N_{\mathbf{i} \rightarrow j}^\alpha = \sum_{q_j^+} p(q_j^+) n_{\mathbf{i} \rightarrow j}^\alpha(q_j^+), \quad (\text{S11})$$

where α denotes R, U, or S, $n_{\mathbf{i} \rightarrow j}^\alpha(q_j^+)$ is the order of appearance of $\Delta \tilde{i}_{\mathbf{i}}^\alpha(q_j^+)$ from left to right as in the example shown in Figure S2. The values of $N_{\mathbf{i} \rightarrow j}^\alpha$ are used to plot $\Delta I_{\mathbf{i} \rightarrow j}^\alpha$ following the expected order of appearance of $\Delta \tilde{i}_{\mathbf{i} \rightarrow j}^\alpha$. All the causalities from SURD presented in this work are plotted in order from left to right, following $N_{\mathbf{i} \rightarrow j}^\alpha$.

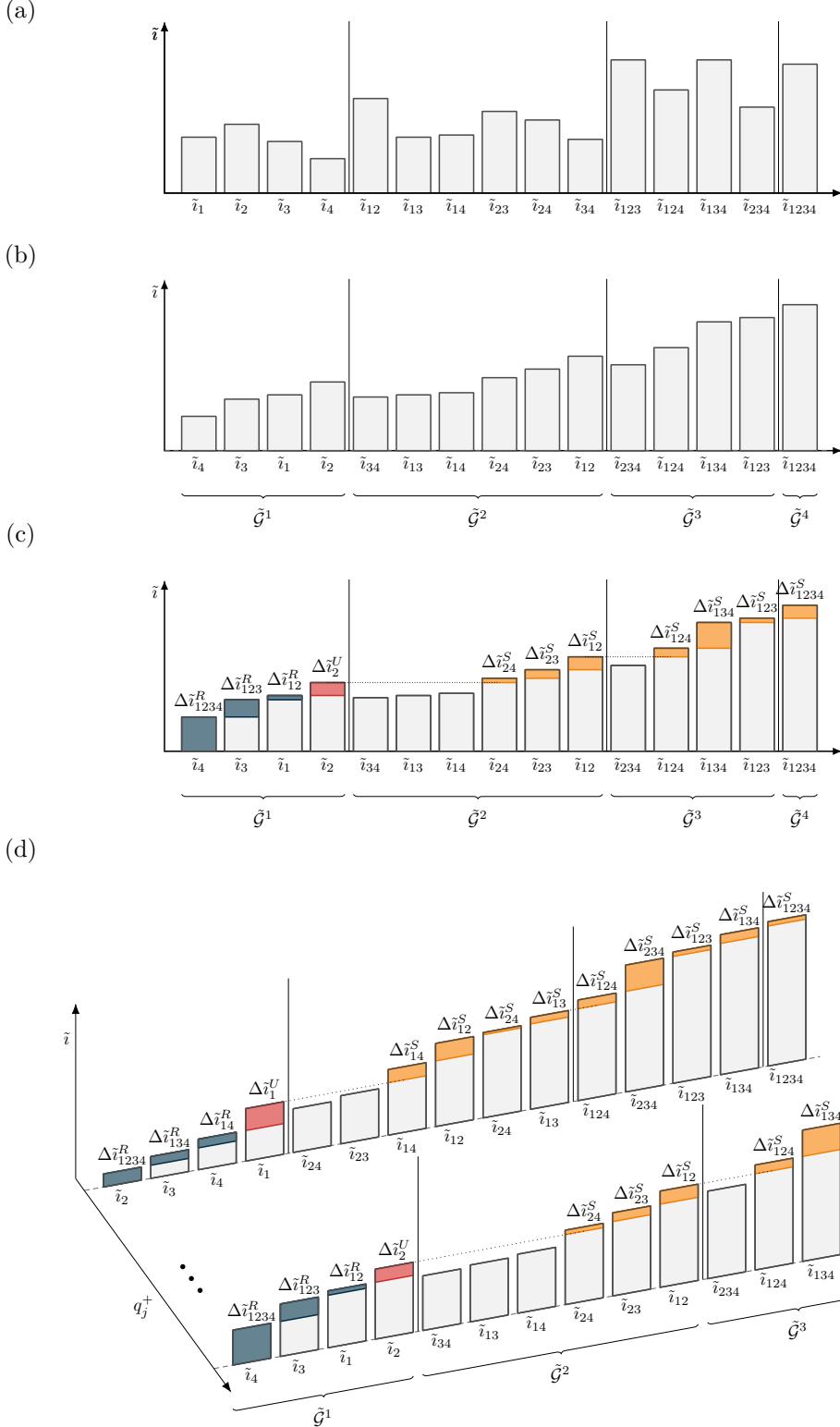


Figure S2: Schematic of the steps involved in the calculation of specific causalities. For a given state $Q_j^+ = q_j^+$, the panels illustrate: (a) all possible specific mutual information values for a collection of four variables; (b) tuples of specific mutual information with the components organized in ascending order; (c) the increments corresponding to specific redundant (blue), unique (red), and synergistic (yellow) causalities; and (d) illustration of specific causalities for different states of Q_j^+ .

S1.3 State-dependent decomposition of causality

The derivation of our state-dependent causality builds on the SURD formulation introduced above. The key distinction is that, rather than quantifying the aggregate contribution of the entire sets Q_j^+ and \mathbf{Q} , we evaluate causality at the level of individual state pairs (q_j^+, \mathbf{q}) . Specifically, we aim to measure the influence of each source state \mathbf{q}_i on each future state of the target variable, denoted as $q_j^+ = q_j(t + \Delta T)$, where $\Delta T > 0$ is an arbitrary time increment. To this end, we assess the information gain about q_j^+ obtained by observing single or groups of states \mathbf{q} from the observable \mathbf{Q} . The principle of forward-in-time information propagation can be recast as:

$$H(Q_j^+) = \sum_{q_j^+ \in Q_j^+} \sum_{\mathbf{q} \in \mathbf{Q}} [\Delta\mathcal{C}(q_j^+; \mathbf{q}) + \Delta\mathcal{N}(q_j^+; \mathbf{q})] + \Delta I_{\text{leak} \rightarrow j}, \quad (\text{S12})$$

where $\Delta\mathcal{C}(q_j^+; \mathbf{q}) > 0$ and $\Delta\mathcal{N}(q_j^+; \mathbf{q}) \leq 0$ denote the causal and non-causal contributions of the source state \mathbf{q} to the target state q_j^+ , respectively, and $\Delta I_{\text{leak} \rightarrow j}$ again accounts for the causality leak. The terms $\Delta\mathcal{C}(q_j^+; \mathbf{q})$ and $\Delta\mathcal{N}(q_j^+; \mathbf{q})$ can be further decomposed as:

$$\Delta\mathcal{C}(q_j^+; \mathbf{q}) = \sum_{i \in \mathcal{P}} \Delta\mathcal{C}_{i \rightarrow j}^R + \sum_{i=1}^N \Delta\mathcal{C}_{i \rightarrow j}^U + \sum_{i \in \mathcal{P}} \Delta\mathcal{C}_{i \rightarrow j}^S, \quad (\text{S13})$$

where $\Delta\mathcal{C}_{i \rightarrow j}^R$, $\Delta\mathcal{C}_{i \rightarrow j}^U$, and $\Delta\mathcal{C}_{i \rightarrow j}^S$ are the redundant, unique, and synergistic causalities, respectively, specific to each of the states of the observed variables to the states of the target variable q_j^+ . As in SURD causality, unique causalities are associated with individual components of \mathbf{Q} , whereas redundant and synergistic causalities arise from groups of variables from \mathbf{Q} . The set \mathcal{P} contains all combinations involving more than one variable. For instance, for $N = 2$, Eq. (2equation.0.2) reduces to $\Delta\mathcal{C}(q_j^+; \mathbf{q}) = \Delta\mathcal{C}_{12 \rightarrow j}^R + \Delta\mathcal{C}_{1 \rightarrow j}^U + \Delta\mathcal{C}_{2 \rightarrow j}^U + \Delta\mathcal{C}_{12 \rightarrow j}^S$. The same expression can be derived for the non-causal component $\Delta\mathcal{N}(q_j^+; \mathbf{q})$:

$$\Delta\mathcal{N}(q_j^+; \mathbf{q}) = \sum_{i \in \mathcal{P}} \Delta\mathcal{N}_{i \rightarrow j}^R + \sum_{i=1}^N \Delta\mathcal{N}_{i \rightarrow j}^U + \sum_{i \in \mathcal{P}} \Delta\mathcal{N}_{i \rightarrow j}^S. \quad (\text{S14})$$

Next, we detail how each component of $\Delta\mathcal{C}(q_j^+; \mathbf{q})$ and $\Delta\mathcal{N}(q_j^+; \mathbf{q})$ can be further decomposed according to the individual states of the target and source variables. We then demonstrate that this state-dependent formulation is equivalent to the total causality measures originally defined by SURD.

S1.3.1 State-dependent redundant causalities

In our decomposition, causal and non-causal contributions are accounted for in the sequence: redundant \rightarrow unique \rightarrow synergistic. We begin by deriving the state-dependent redundant terms. Denote by $\Delta\mathcal{C}_{i \rightarrow j}^R$ and $\Delta\mathcal{N}_{i \rightarrow j}^R$ the positive (causal) and negative (non-causal) contributions, respectively, from the joint state (\mathbf{q}_i, q_j^+) to $\Delta I_{i \rightarrow j}^R$, defined as

$$\Delta I_{i \rightarrow j}^R = \sum_{q_j^+ \in Q_j^+} \sum_{\mathbf{q}_i \in \mathbf{Q}} [\Delta\mathcal{C}_{i \rightarrow j}^R + \Delta\mathcal{N}_{i \rightarrow j}^R], \quad (\text{S15})$$

where $i = [i_1, i_2, \dots]$ indexes the source variables involved in the redundant causality, and $\mathbf{q}_i = [q_{i_1}, q_{i_2}, \dots]$ denotes their particular states. The expressions for $\Delta\mathcal{C}_{i \rightarrow j}^R$ and $\Delta\mathcal{N}_{i \rightarrow j}^R$ follow from Eqs. S6, S7, and S10:

$$\Delta I_{i \rightarrow j}^R = \sum_{q_j^+} \sum_{q_{i_k}} \sum_{q_{i_{k-1}}} p(q_j^+, q_{i_k}, q_{i_{k-1}}) \log_2 \left(\frac{p(q_j^+ | q_{i_k})}{p(q_j^+ | q_{i_{k-1}})} \right), \quad (\text{S16})$$

for $\tilde{i}_{i_k}, \tilde{i}_{i_{k-1}} \in \mathcal{G}_1$ with $k \neq n_1$, and zero otherwise. The index vector $i = [i_1, i_2, \dots]$ comprises those n satisfying $\tilde{i}_n \geq \tilde{i}_{i_k}$ (ordered so that \tilde{i}_{i_k} increases). For example, if $i = [1, 2]$ and $\tilde{i}_4 < \tilde{i}_3 < \tilde{i}_1 < \tilde{i}_2$ (see Fig. S2b), then $i_k = 1$

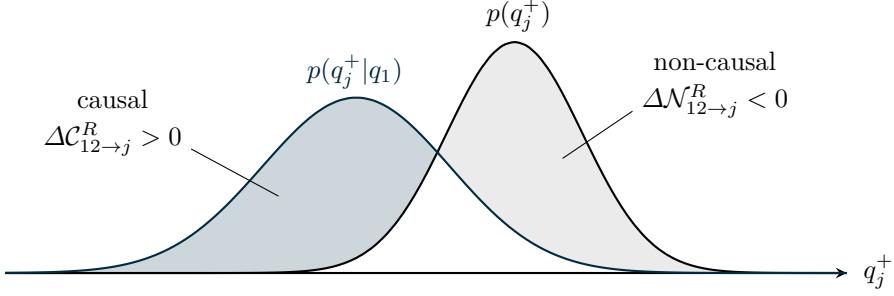


Figure S3: State-dependent redundant causality $\Delta\mathcal{C}_{12\rightarrow j}^R$ for a system with target variable Q_j^+ and source variables Q_1 and Q_2 . In this example, the specific mutual information $\tilde{\imath}(Q_j^+ = q_j^+; Q_2)$ is greater than $\tilde{\imath}(Q_j^+ = q_j^+; Q_1)$ for all states $q_j^+ \in Q_j^+$. Causal states ($\Delta\mathcal{C}_{12\rightarrow j}^R > 0$) indicate causal states where knowledge of q_1 increases the likelihood of the reference state q_j^+ , i.e., $p(q_j^+ | q_1) > p(q_j^+)$. Conversely, non-causal states ($\Delta\mathcal{N}_{12\rightarrow j}^R < 0$) corresponds to negative values where knowledge of q_1 decreases the likelihood of q_j^+ .

and $i_{k-1} = 3$. Therefore, the positive redundant causal component is given by:

$$\Delta\mathcal{C}_{\mathbf{i}\rightarrow j}^R = \begin{cases} \sum_{q_j^+} p(q_j^+, \mathbf{q}) \max \left(0, \log_2 \left(\frac{p(q_j^+ | q_{i_k})}{p(q_j^+ | q_{i_{k-1}})} \right) \right), & \text{for } \tilde{i}_{i_k}, \tilde{i}_{i_{k-1}} \in \tilde{\mathcal{G}}^1 \text{ and } k \neq n_1, \\ 0, & \text{otherwise,} \end{cases} \quad (\text{S17})$$

where \mathbf{i} lists all indices not in \mathbf{j} .

For example, in a system with two variables Q_1 and Q_2 targeting Q_j^+ with joint distribution $p(q_j^+, q_1, q_2)$, Eq. (S17) simplifies to

$$\Delta\mathcal{C}_{12\rightarrow j}^R = \begin{cases} p(q_j^+, q_1, q_2) \max \left(0, \log_2 \left(\frac{p(q_j^+ | q_1)}{p(q_j^+)} \right) \right), & \text{for } \tilde{i}_2 \geq \tilde{i}_1, \\ 0, & \text{otherwise.} \end{cases} \quad (\text{S18})$$

Figure S3 provides a schematic of the state-dependent redundant causality $\Delta\mathcal{C}_{12\rightarrow j}^R$. Positive values of $\Delta\mathcal{C}_{12\rightarrow j}^R$ indicate states where knowledge of q_1 increases the likelihood of q_j^+ , i.e., $p(q_j^+ | q_1) > p(q_j^+)$.

An analogous definition holds for the non-causal, state-dependent redundant causality, capturing those source states that increase uncertainty about q_j^+ relative to the previous level. It is given by:

$$\Delta\mathcal{N}_{\mathbf{i}\rightarrow j}^R = \begin{cases} \sum_{q_j^+} p(q_j^+, \mathbf{q}) \min \left(0, \log_2 \left(\frac{p(q_j^+ | q_{i_k})}{p(q_j^+ | q_{i_{k-1}})} \right) \right), & \text{for } \tilde{i}_{i_k}, \tilde{i}_{i_{k-1}} \in \tilde{\mathcal{G}}^1 \text{ and } k \neq n_1, \\ 0, & \text{otherwise.} \end{cases} \quad (\text{S19})$$

In the example from Fig. S3, non-causal values indicate states where knowledge of q_1 decreases the likelihood of q_j^+ , i.e., $p(q_j^+ | q_1) < p(q_j^+)$.

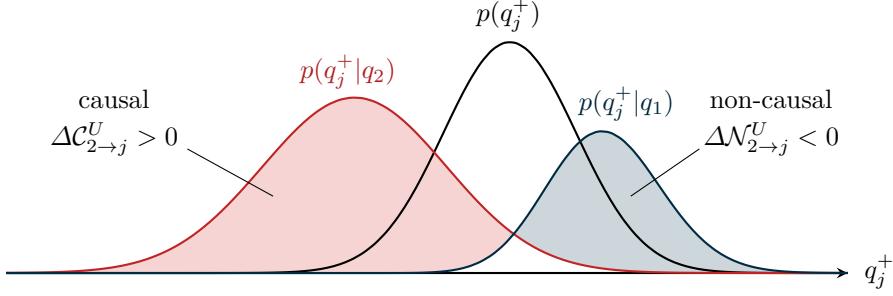


Figure S4: State-dependent unique causality $\Delta C_{2 \rightarrow j}^U$ for a system with target variable Q_j^+ and source variables Q_1 and Q_2 , as given by Eq. (S23). In this example, the specific mutual information $\tilde{i}(Q_j^+ = q_j^+; Q_2)$ exceeds $\tilde{i}(Q_j^+ = q_j^+; Q_1)$ for every state $q_j^+ \in Q_j^+$. Causal states ($\Delta C_{2 \rightarrow j}^U > 0$) indicate states where knowing q_2 increases the likelihood of q_j^+ relative to knowing q_1 , i.e., $p(q_j^+ | q_2) > p(q_j^+ | q_1)$. Conversely, non-causal states ($\Delta N_{2 \rightarrow j}^U < 0$) correspond to states where knowledge of q_2 decreases that likelihood, $p(q_j^+ | q_2) < p(q_j^+ | q_1)$.

S1.3.2 State-dependent unique causalities

We denote by $\Delta C_{i \rightarrow j}^U$ and $\Delta N_{i \rightarrow j}^U$ the positive (causal) and negative (non-causal) state-dependent contributions from (q_i, q_j^+) to the unique causality increment:

$$\Delta I_{i \rightarrow j}^U = \sum_{q_j^+ \in Q_j^+} \sum_{q_i \in Q_i} [\Delta C_{i \rightarrow j}^U + \Delta N_{i \rightarrow j}^U]. \quad (\text{S20})$$

Comparing Eq. (S20) with the full expression for $\Delta I_{i \rightarrow j}^U$ (Eqs. S6, S8, and S10) leads to

$$\Delta I_{i \rightarrow j}^U = \sum_{q_j^+} \sum_{q_i} \sum_{q_k} p(q_j^+, q_i, q_k) \log_2 \left(\frac{p(q_j^+ | q_i)}{p(q_j^+ | q_k)} \right), \quad (\text{S21})$$

which holds for $i = n_1$ and $\tilde{i}_i, \tilde{i}_k \in \tilde{\mathcal{G}}^1$, and is zero otherwise. This represents the expected log-ratio of the conditional probabilities $p(q_j^+ | q_i)$ and $p(q_j^+ | q_k)$. From this, we obtain the causal component:

$$\Delta C_{i \rightarrow j}^U(q_i, q_j^+) = \begin{cases} \sum_{q_k} p(q_j^+, q_i, q_k) \max\left(0, \log_2\left(\frac{p(q_j^+ | q_i)}{p(q_j^+ | q_k)}\right)\right), & \text{if } \tilde{i}_i \geq \tilde{i}_{k-1}, \\ 0, & \text{otherwise,} \end{cases} \quad (\text{S22})$$

where \tilde{i}_k is the second-largest element of $\tilde{\mathcal{G}}^1$.

For instance, in a two-variable system (Q_1, Q_2) targeting Q_j^+ with joint distribution $p(q_j^+, q_1, q_2)$, Eq. (S22) for the unique causality from Q_2 becomes

$$\Delta C_{2 \rightarrow j}^U = \begin{cases} \sum_{q_1} p(q_j^+, q_1, q_2) \max\left(0, \log_2\left(\frac{p(q_j^+ | q_2)}{p(q_j^+ | q_1)}\right)\right), & \text{if } \tilde{i}_2 \geq \tilde{i}_1, \\ 0, & \text{otherwise.} \end{cases} \quad (\text{S23})$$

The non-causal component contains those source states that increase uncertainty about q_j^+ compared to the largest redundant contribution, and is always non-positive:

$$\Delta N_{i \rightarrow j}^U(q_i, q_j^+) = \begin{cases} \sum_{q_k} p(q_j^+, q_i, q_k) \min\left(0, \log_2\left(\frac{p(q_j^+ | q_i)}{p(q_j^+ | q_k)}\right)\right), & \text{if } \tilde{i}_i \geq \tilde{i}_{k-1}, \\ 0, & \text{otherwise.} \end{cases} \quad (\text{S24})$$

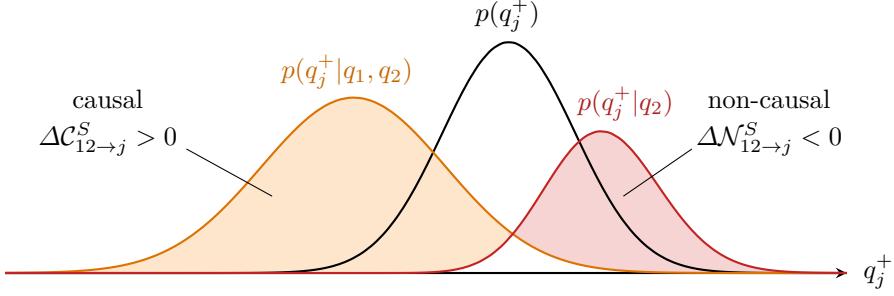


Figure S5: State-dependent synergistic causality $\Delta C_{12 \rightarrow j}^S$ for a system with target variable Q_j^+ and source variables Q_1 and Q_2 . In this example, the specific mutual information $\tilde{i}(Q_j^+ = q_j^+; Q_1, Q_2)$ is greater than $\tilde{i}(Q_j^+ = q_j^+; Q_2)$ for all states $q_j^+ \in Q_j^+$. Causal values ($\Delta C_{12 \rightarrow j}^S > 0$) indicate states where knowledge of both q_1 and q_2 increases the likelihood of the reference state q_j^+ compared to knowing only q_2 , i.e., $p(q_j^+ | q_1, q_2) > p(q_j^+ | q_2)$. Conversely, non-causal values ($\Delta N_{12 \rightarrow j}^S < 0$) implies that knowledge of q_1 and q_2 reduces the likelihood of q_j^+ relative to knowing only q_2 , i.e., $p(q_j^+ | q_1, q_2) < p(q_j^+ | q_2)$.

S1.3.3 State-dependent synergistic causalities

The derivation for the state-dependent synergistic causality parallels those for the redundant and unique terms. We introduce the positive (causal) and negative (non-causal) contributions $\Delta C_{\mathbf{i} \rightarrow j}^S$ and $\Delta N_{\mathbf{i} \rightarrow j}^S$ from the joint state (\mathbf{q}_i, q_j^+) to the synergistic increment $\Delta I_{\mathbf{i} \rightarrow j}^S$:

$$\Delta I_{\mathbf{i} \rightarrow j}^S = \sum_{q_j^+ \in Q_j^+} \sum_{\mathbf{q}_i \in \mathbf{Q}_i} [\Delta C_{\mathbf{i} \rightarrow j}^S + \Delta N_{\mathbf{i} \rightarrow j}^S], \quad (\text{S25})$$

where $\mathbf{i} = [i_1, i_2, \dots]$ indexes the source variables in the synergy and $\mathbf{q}_i = [q_{i_1}, q_{i_2}, \dots]$ their particular states. Hence, the dimensionality of $\Delta C_{\mathbf{i} \rightarrow j}^S$ grows with $|\mathbf{i}|$. For instance, if $\mathbf{q}_i = [q_1, q_3, q_4]$, then $\Delta C_{134 \rightarrow j}^S(q_1, q_3, q_4, q_j^+)$ depends on three source states in addition to q_j^+ . Comparing Eq. (S25) with the full form of $\Delta I_{\mathbf{i} \rightarrow j}^S$ (using Eqs. S6, S9, and S10) yields

$$\Delta I_{\mathbf{i} \rightarrow j}^S = \sum_{q_j^+} \sum_{\mathbf{q}_i} \sum_{\mathbf{q}_k} p(q_j^+, \mathbf{q}_i, \mathbf{q}_k) \log_2 \left(\frac{p(q_j^+ | \mathbf{q}_i)}{p(q_j^+ | \mathbf{q}_k)} \right), \quad (\text{S26})$$

where the index vector \mathbf{k} selects the variables with the next-highest specific mutual information after those in \mathbf{i} . By definition of synergistic causality (§S1.2), \mathbf{k} may either continue in the same order as \mathbf{i} or revert to the maximal entry of the previous order. This leads to the following decomposition of the causal component:

$$\Delta C_{\mathbf{i} \rightarrow j}^S = \begin{cases} \sum_{\mathbf{q}_k} p(q_j^+, \mathbf{q}_i, \mathbf{q}_k) \max \left(0, \log_2 \left(\frac{p(q_j^+ | \mathbf{q}_i)}{p(q_j^+ | \mathbf{q}_k)} \right) \right), & \tilde{i}_{\mathbf{i}} \geq \tilde{i}_{\mathbf{k}} \geq \max \{\tilde{\mathcal{G}}^{M-1}\}, \\ \sum_{\mathbf{q}_m} p(q_j^+, \mathbf{q}_i, \mathbf{q}_m) \max \left(0, \log_2 \left(\frac{p(q_j^+ | \mathbf{q}_i)}{p(q_j^+ | \mathbf{q}_m)} \right) \right), & \tilde{i}_{\mathbf{i}} \geq \tilde{i}_{\mathbf{m}} = \max \{\tilde{\mathcal{G}}^{M-1}\} > \tilde{i}_{\mathbf{k}}, \\ 0, & \text{otherwise,} \end{cases} \quad (\text{S27})$$

where \mathbf{m} indexes the element with the maximum specific mutual information in $\tilde{\mathcal{G}}^{M-1}$.

In the two-variable example (Q_1, Q_2) targeting Q_j^+ with joint distribution $p(q_j^+, q_1, q_2)$, this reduces to

$$\Delta C_{12 \rightarrow j}^S = \begin{cases} p(q_j^+, q_1, q_2) \max \left(0, \log_2 \left(\frac{p(q_j^+ | q_1, q_2)}{p(q_j^+ | q_2)} \right) \right), & \tilde{i}_{12} \geq \tilde{i}_2, \\ 0, & \text{otherwise.} \end{cases} \quad (\text{S28})$$

Figure S5 illustrates $\Delta C_{12 \rightarrow j}^S$ for a system consisting of two variables Q_1 and Q_2 and a target variable Q_j^+ with

a joint probability distribution $p(q_j^+, q_1, q_2)$. Positive values mark states where joint knowledge of q_1 and q_2 adds information beyond the unique effect of q_2 , i.e. $p(q_j^+ | q_1, q_2) > p(q_j^+ | q_2)$. Zero values indicate no synergy relative to the unique contribution from q_2 .

Finally, the non-causal component follows analogously:

$$\Delta\mathcal{N}_{i \rightarrow j}^S = \begin{cases} \sum_{\mathbf{q}_k} p(q_j^+, \mathbf{q}_i, \mathbf{q}_k) \min \left(0, \log_2 \left(\frac{p(q_j^+ | \mathbf{q}_i)}{p(q_j^+ | \mathbf{q}_k)} \right) \right), & \tilde{i}_i \geq \tilde{i}_k \geq \max\{\tilde{\mathcal{G}}^{M-1}\}, \\ \sum_{\mathbf{q}_m} p(q_j^+, \mathbf{q}_i, \mathbf{q}_m) \min \left(0, \log_2 \left(\frac{p(q_j^+ | \mathbf{q}_i)}{p(q_j^+ | \mathbf{q}_m)} \right) \right), & \tilde{i}_i \geq \tilde{i}_m = \max\{\tilde{\mathcal{G}}^{M-1}\} > \tilde{i}_k, \\ 0, & \text{otherwise.} \end{cases} \quad (\text{S29})$$

S1.4 Alternative state-dependent decomposition of causality

The definition of causality adopted here treats positive increments of information as causal, while negative increments are deemed non-causal. This choice is guided by the intuition that such a distinction aligns better with how we interpret causality in the context of understanding interactions among variables. Of course, this is only one possible definition and alternative formulations may also be considered plausible. Here, we present two such alternatives and discuss the reasons why we did not adopt them in our framework.

First, we illustrate the inevitability of obtaining negative terms when decomposing mutual information based on individual state pairs (q_j^+, \mathbf{q}) . Consider the decomposition

$$I(Q_j^+; \mathbf{Q}) = \sum_{q_j^+} \sum_{\mathbf{q}} [\Delta\mathcal{C}(q_j^+, \mathbf{q}) + \Delta\mathcal{N}(q_j^+, \mathbf{q})],$$

which can also be written explicitly as:

$$I(Q_j^+; \mathbf{Q}) = \sum_{q_j^+ \in Q_j^+} \sum_{\mathbf{q} \in \mathbf{Q}} \Delta\mathcal{I} = \sum_{q_j^+ \in Q_j^+} \sum_{\mathbf{q} \in \mathbf{Q}} p(q_j^+, \mathbf{q}) \log_2 \frac{p(q_j^+, \mathbf{q})}{p(q_j^+)p(\mathbf{q})}. \quad (\text{S30})$$

The term inside the sum is zero *if and only if* q_j^+ and \mathbf{q} are statistically independent, and it can be positive or negative depending on whether the joint probability $p(q_j^+, \mathbf{q})$ is greater or smaller than the product of the marginals $p(q_j^+)p(\mathbf{q})$.

As a simple example, consider the task of forecasting whether it will rain tomorrow, denoted by Q_{rain}^+ , where $q_{\text{rain}}^+ = 1$ indicates rain and $q_{\text{rain}}^+ = 0$ indicates no rain. Suppose we use atmospheric pressure, Q_p , as an observable, which can take only two values: low ($q_p = 0$) or high ($q_p = 1$). After collecting observations, we can construct a joint probability table for all combinations of these variables. For example, consider the observations in Table S1.

pressure	rain	$p(\text{pressure, rain})$	$\Delta\mathcal{I}(\text{pressure; rain})$
0	0	0.10	-0.10
0	1	0.70	0.16
1	0	0.15	0.24
1	1	0.05	-0.08

Table S1: Joint probability distribution of pressure and rain, along with their contributions to the mutual information $\Delta\mathcal{I}$. The marginal probabilities are: $p(\text{pressure} = 0) = 0.80$, $p(\text{pressure} = 1) = 0.20$, $p(\text{rain} = 0) = 0.25$, and $p(\text{rain} = 1) = 0.75$. The total mutual information is given by the sum of all values in $\Delta\mathcal{I}$, i.e., $I(\text{pressure; rain}) = 0.22$.

Given a table of probabilities, we can evaluate the expression in Eq. (S30) for all combinations of states. In the example above, observing that the pressure is low (i.e., $q_p = 0$) yields a negative increment of information for $q_{\text{rain}}^+ = 0$ and a positive increment for $q_{\text{rain}}^+ = 1$. Intuitively, this implies that a particular state \mathbf{q} of the source variables may be non-causal with respect to one target state $q_{j_1}^+$, while simultaneously being causal with respect to a different target state $q_{j_2}^+$. This behavior arises because, if for some pair $(\mathbf{q}, q_{j_1}^+)$ we have $p(\mathbf{q}, q_{j_1}^+) < p(\mathbf{q})p(q_{j_1}^+)$, then normalization of probabilities requires that this be offset by another pair $(\mathbf{q}, q_{j_2}^+)$ such that $p(\mathbf{q}, q_{j_2}^+) > p(\mathbf{q})p(q_{j_2}^+)$.

As a consequence, if there exists at least one pair of states (q_j^+, \mathbf{q}) such that the contribution to the mutual information $\Delta\mathcal{I}(q_j^+, \mathbf{q})$ is negative, then any decomposition of mutual information into terms that depend solely on

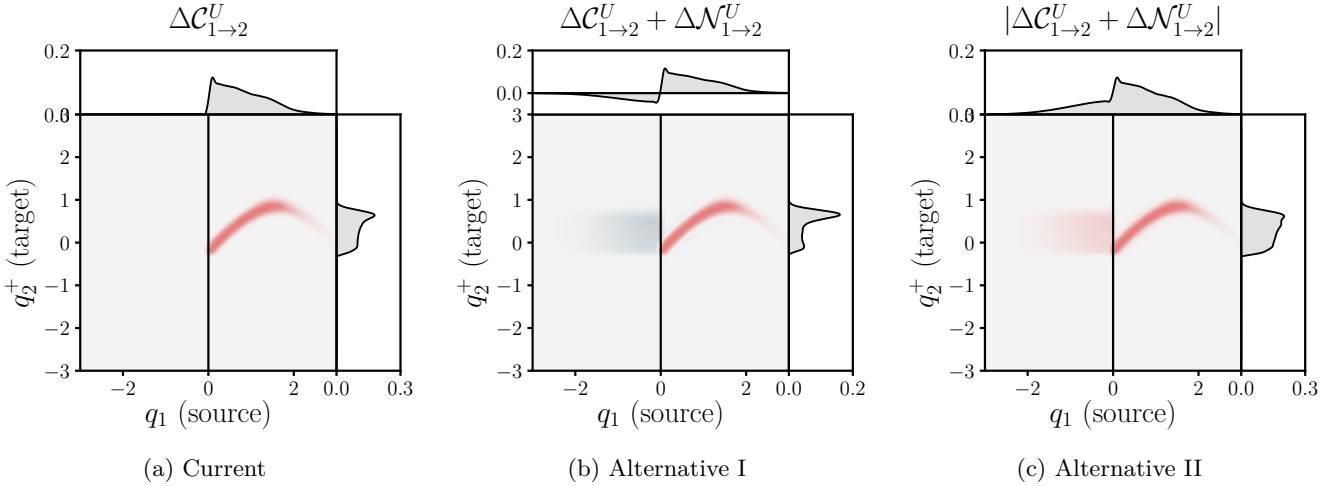


Figure S6: Comparison of different state-dependent causal decomposition methods: (a) Proposed method, which separates causal and non-causal contributions; (b) Alternative Method I, which combines causal and non-causal terms into a single quantity; and (c) Alternative Method II, which uses the absolute value of the combined contributions, discarding the sign information.

specific pairs (q_j^+, \mathbf{q}) must necessarily include at least one negative component. In particular, the sum $\Delta\mathcal{C}(q_j^+, \mathbf{q}) + \Delta\mathcal{N}(q_j^+, \mathbf{q})$ must be negative for some state pair (q_j^+, \mathbf{q}) . If this sum can take both positive and negative values, then each of the state-dependent causal components $\Delta\mathcal{C}_{i \rightarrow j}^\alpha$ and $\Delta\mathcal{N}_{i \rightarrow j}^\alpha$, with $\alpha \in \{R, U, S\}$, must also include negative values for at least one such pair.

S1.4.1 Alternative definition I: Causality as positive and negative information increments

The first alternative definition considered treats the sum of both positive and negative contributions as the effective measure of causality at the state level, namely $\Delta\mathcal{C}_{i \rightarrow j}^\alpha + \Delta\mathcal{N}_{i \rightarrow j}^\alpha$. The main advantage of this formulation is that the total SURD causalities can be directly recovered by summing this combined quantity over all states. In contrast, in our proposed approach, the total SURD causalities are obtained only when the causal and non-causal components are explicitly combined. While the alternative definition simplifies aggregation, it may reduce interpretability at the state level, since the resulting maps include both reinforcing and opposing contributions. Nevertheless, this formulation still preserves the key qualitative distinction between causal states—where $\Delta\mathcal{C}_{i \rightarrow j}^\alpha + \Delta\mathcal{N}_{i \rightarrow j}^\alpha > 0$ —and non-causal states—where $\Delta\mathcal{C}_{i \rightarrow j}^\alpha + \Delta\mathcal{N}_{i \rightarrow j}^\alpha < 0$ —consistent with our proposed method.

To illustrate the distinction between our proposed method and the alternative formulation discussed above, Fig. S6 presents a comparison based on the source-dependent benchmark case introduced in Fig. 2. State-dependent decomposition of causality for the source-dependent case. From left to right, the panels show: state-dependent redundant ($\Delta\mathcal{C}_{12 \rightarrow 2}^R$), unique ($\Delta\mathcal{C}_{12 \rightarrow 2}^U$), and synergistic ($\Delta\mathcal{C}_{12 \rightarrow 2}^S$) causal contributions in blue, red, and orange, respectively; and the SURD causalities to the target variable Q_2^+ . The notation follows: U1 denotes the unique SURD causality from Q_1 to Q_2^+ , i.e., $\Delta I_{1 \rightarrow 2}^U$, with analogous definitions for other components (i.e., R and S). All SURD and state-dependent causal contributions are normalized by the mutual information $I(Q_2^+; Q_1, Q_2)$. The causal maps for redundant and synergistic components show averages over all states of q_2 . In each panel, the bottom row displays the temporal evolution of $q_1(n)$, color-coded according to the corresponding instantaneous state-dependent causal contribution figure 2. In this example, the interaction between variables is defined *a priori* as $q_1 \rightarrow q_2^+$ if $q_1 > 0$, and $q_2 \rightarrow q_2^+$ if $q_1 \leq 0$. We focus on the unique causality from the source variable q_1 to the target variable q_2^+ , quantified by the term $\Delta\mathcal{C}_{1 \rightarrow 2}^U$. Under our proposed method, only states with $q_1 > 0$ are identified as uniquely causal, in agreement with the system functional dependence. In contrast, the alternative formulation, which uses the combined quantity $\Delta\mathcal{C}_{1 \rightarrow 2}^U + \Delta\mathcal{N}_{1 \rightarrow 2}^U$, also identifies states with $q_1 < 0$ as contributing states to $\Delta I_{1 \rightarrow 2}^U$. We argue that the latter is less useful for interpreting the functional structure of the system.

S1.4.2 Alternative definition II: Causality as absolute value of information increments

A second alternative approach defines state-dependent causality as the absolute value of the combined positive and negative increments: $|\Delta\mathcal{C}_{i \rightarrow j}^\alpha + \Delta\mathcal{N}_{i \rightarrow j}^\alpha|$. However, this formulation presents significant drawbacks. First, the total

causality defined by SURD is no longer preserved, as summing absolute values distorts the balance between positive and negative contributions. More critically, taking the absolute value eliminates the distinction between $\Delta\mathcal{C}_{i \rightarrow j}^\alpha$ and $\Delta\mathcal{N}_{i \rightarrow j}^\alpha$, effectively interpreting all state pairs as causal regardless of whether they reduce or increase uncertainty about the target. This conflation undermines interpretability: states in $\Delta\mathcal{C}_{i \rightarrow j}^\alpha$ provide useful, predictive information that reduces uncertainty about the future, whereas states in $\Delta\mathcal{N}_{i \rightarrow j}^\alpha$ increase uncertainty and may reflect noise, misleading associations, or indirect effects. While the latter may carry some informational relevance, their role is fundamentally different from genuinely causal states. As a result, treating both types equivalently may obscure the intuitive notion of causality as uncertainty reduction.

The results of this alternative formulation, applied to the source-dependent benchmark case, are shown in Fig. 2. State-dependent decomposition of causality for the source-dependent case. From left to right, the panels show: state-dependent redundant ($\Delta\mathcal{C}_{12 \rightarrow 2}^R$), unique ($\Delta\mathcal{C}_{1 \rightarrow 2}^U$), and synergistic ($\Delta\mathcal{C}_{12 \rightarrow 2}^S$) causal contributions in blue, red, and orange, respectively; and the SURD causalities to the target variable Q_2^+ . The notation follows: U1 denotes the unique SURD causality from Q_1 to Q_2^+ , i.e., $\Delta I_{1 \rightarrow 2}^U$, with analogous definitions for other components (i.e., R and S). All SURD and state-dependent causal contributions are normalized by the mutual information $I(Q_2^+; Q_1, Q_2)$. The causal maps for redundant and synergistic components show averages over all states of q_2 . In each panel, the bottom row displays the temporal evolution of $q_1(n)$, color-coded according to the corresponding instantaneous state-dependent causal contribution figure 2. We observe that this approach assigns equal importance to both $q_1 > 0$ and $q_1 < 0$, which can be misleading. While q_1 indeed plays a role in the regime where $q_1 < 0$ —as its sign determines which variable governs the behavior of q_2^+ —the information from q_2 is still necessary to fully determine the future state q_2^+ . Consequently, in this regime, q_2^+ cannot be uniquely determined by q_1 alone.

In conclusion, our proposed definition of state-dependent causality offers clearer interpretability. Additionally, the sum of all causal and non-causal components recovers the full mutual information between the source variables \mathbf{Q} and the target Q_j^+ , without requiring a redefinition of the global redundant, unique, and synergistic components in the SURD framework.

S2 Verification and validation

S2.1 Multivariate correlated Gaussian variables

We verify the implementation of the state-dependent causalities using a Gaussian system, for which an analytical solution can be obtained. We consider a system with two source variables, Q_1 and Q_2 , and a target variable, Q_j^+ , following a joint Gaussian distribution with arbitrary correlations between the variables. The joint probability density function $p(q_j^+, q_1, q_2)$ takes the form:

$$p(q_j^+, q_1, q_2) = \frac{1}{\sqrt{(2\pi)^3 \det \Sigma}} \exp \left(-\frac{1}{2} (\mathbf{z} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{z} - \boldsymbol{\mu}) \right), \quad (\text{S31})$$

where \mathbf{z} represents the vector of variables q_j^+ , q_1 , and q_2 ; $\boldsymbol{\mu} = [\mu_+, \mu_1, \mu_2]^\top$ is the mean vector of these variables, and Σ is the covariance matrix which quantifies the degree of dependence among the variables. Mathematically, the matrix is represented as:

$$\Sigma = \begin{bmatrix} \sigma_+^2 & \rho_1 \sigma_+ \sigma_1 & \rho_2 \sigma_+ \sigma_2 \\ \rho_1 \sigma_+ \sigma_1 & \sigma_1^2 & \rho_{12} \sigma_1 \sigma_2 \\ \rho_2 \sigma_+ \sigma_2 & \rho_{12} \sigma_1 \sigma_2 & \sigma_2^2 \end{bmatrix}, \quad (\text{S32})$$

where σ_+ , σ_1 , and σ_2 denote the standard deviations of q_j^+ , q_1 , and q_2 , respectively, and ρ_1 , ρ_2 , and ρ_{12} are the correlation coefficients that measure the linear dependencies between the pairs of these variables. In particular, ρ_1 denotes the correlation between Q_1 and Q_j^+ , ρ_2 the correlation between Q_2 and Q_j^+ , and ρ_{12} the correlation between Q_1 and Q_2 .

From Eq. (S31), we can directly obtain an analytical equation for $\Delta C_{i \rightarrow j}^\alpha$. For example, for the case in which $\rho_1 = \rho_{12} = 0$ and $\rho_2 \neq 0$, the analytical expression for $\Delta C_{2 \rightarrow j}$ is given by Eq. (S22):

$$\Delta C_{2 \rightarrow j}^U = \max \left(0, \frac{1}{\pi^{3/2} \sqrt{2(1 - \rho_2^2)}} \exp \left[-\frac{1}{1 - \rho_2^2} \left(\frac{q_2^2}{2} + \frac{q_j^{+2}}{2} + (1 - \rho_2) \frac{q_1^2}{2} - q_2 q_j^+ \rho_2 \right) \right] \log \left(\frac{\exp \left(-\frac{\rho_2(-2q_2 q_j^+ + q_2^2 \rho_2 + q_j^{+2} \rho_2)}{2(1 - \rho_2^2)} \right)}{\sqrt{1 - \rho_2^2}} \right)^{1/2} \right), \quad (\text{S33})$$

where we assume that the condition $\tilde{i}_2 \geq \tilde{i}_1$ is satisfied for all states of Q_j^+ . A similar expression can be obtained for the non-causal component $\Delta N_{2 \rightarrow j}$ using the function $\min(\cdot)$.

Figure S7 presents the causal and non-causal contributions to the unique causality, given by $\Delta C_{2 \rightarrow j}^U + \Delta N_{2 \rightarrow j}^U$, for a Gaussian system with three different combinations of correlation parameters: ρ_1 (correlation between Q_1 and Q_j^+) and ρ_2 (correlation between Q_2 and Q_j^+). These analytical results are compared against the numerical estimates obtained from our algorithmic implementation, which is used throughout this study. This comparison provides a verification of the numerical results for both the causal and non-causal components with respect to the exact analytical solutions.

S2.2 Source-dependent unique causality

We investigate additional benchmark cases to further assess the performance of our state-dependent causal inference method. To this end, we extend the benchmark systems analyzed in the validation section of the main text by introducing an additional target variable, Q_3^+ , whose dynamics mirror those of Q_2^+ but with an opposite conditioning structure. The variables Q_2^+ and Q_3^+ are constructed to be statistically independent of one another. This setup is designed to evaluate the ability of our method to distinguish state-dependent confounding effects, where the same source variable, Q_1 , influences different targets in distinct regions of its state space. Specifically, the interactions are defined such that $q_1 \rightarrow q_2^+$ for $q_1 > 0$, and $q_1 \rightarrow q_3^+$ for $q_1 < 0$. The dynamics of Q_2^+ are identical to those presented in Fig. 2State-dependent decomposition of causality for the source-dependent case. From left to right, the panels show: state-dependent redundant ($\Delta C_{12 \rightarrow 2}^R$), unique ($\Delta C_{1 \rightarrow 2}^U$), and synergistic ($\Delta C_{12 \rightarrow 2}^S$) causal contributions in blue, red, and orange, respectively; and the SURD causalities to the target variable Q_2^+ . The notation follows: U1 denotes the unique SURD causality from Q_1 to Q_2^+ , i.e., $\Delta I_{1 \rightarrow 2}^U$, with analogous definitions for other components (i.e., R and S). All SURD and state-dependent causal contributions are normalized by the mutual information $I(Q_2^+; Q_1, Q_2)$. The causal maps for redundant and synergistic components show averages over all states of q_2 . In each panel, the bottom row displays the temporal evolution of $q_1(n)$, color-coded according to the corresponding instantaneous state-dependent causal contribution figure.2, while the behavior of Q_3^+ is shown in Fig. S8. The source variable Q_1 is sampled from a standard normal distribution with zero mean and unit variance. The variables Q_2 and Q_3 are driven

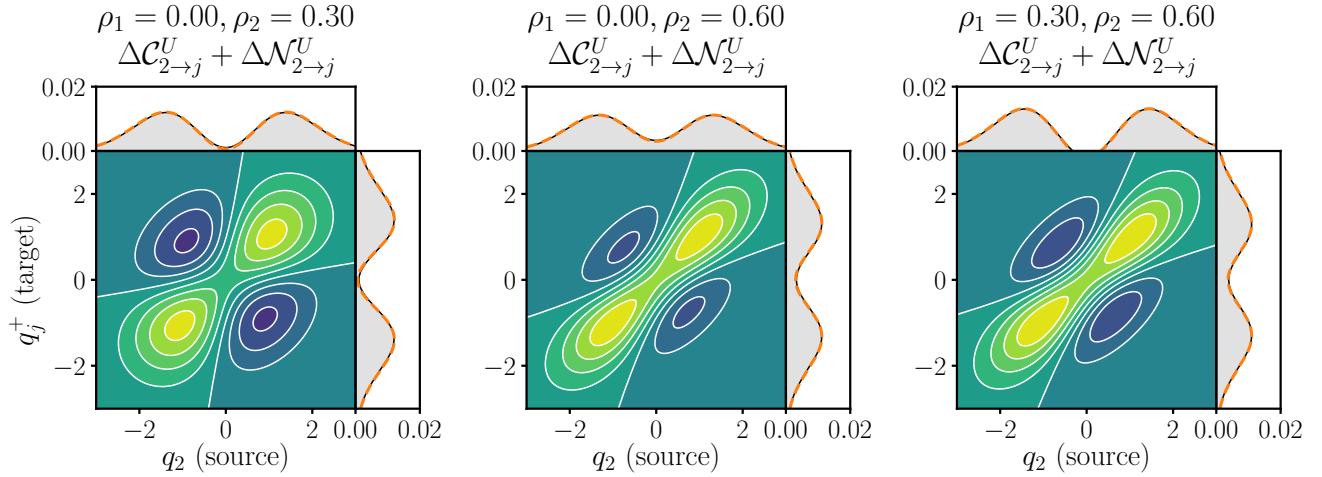


Figure S7: Causal $\Delta\mathcal{C}_{2 \rightarrow j}^U$ and non-causal $\Delta\mathcal{N}_{2 \rightarrow j}^U$ contributions to the unique causality $\Delta I_{2 \rightarrow j}^U$ for a system with two source variables Q_1 and Q_2 and a target variable Q_j^+ following a Gaussian distribution with three different combinations of correlation between Q_1 and Q_j^+ , i.e. ρ_1 , and between Q_2 and Q_j^+ , i.e. ρ_2 . The analytical results are shown with filled contours and grey surfaces, whereas the numerical results from SURD are shown with white line contours and orange dashed lines for all cases.

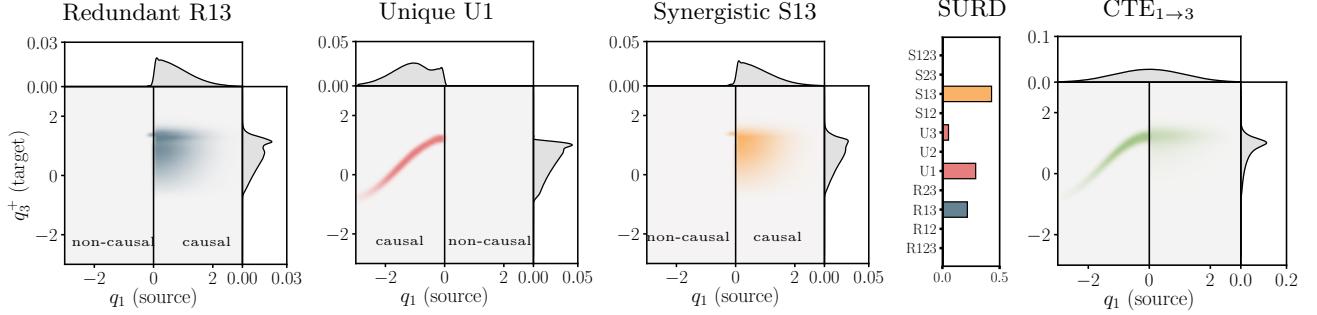
by independent stochastic forcings W_2 and W_3 , respectively, each modeled as Gaussian random variables with means of -2 and 2 , and unit standard deviation.

The state-dependent redundant, unique, and synergistic causalities for the target variable Q_2^+ are identical to those shown in Fig. 2. State-dependent decomposition of causality for the source-dependent case. From left to right, the panels show: state-dependent redundant ($\Delta\mathcal{C}_{12 \rightarrow 2}^R$), unique ($\Delta\mathcal{C}_{1 \rightarrow 2}^U$), and synergistic ($\Delta\mathcal{C}_{12 \rightarrow 2}^S$) causal contributions in blue, red, and orange, respectively; and the SURD causalities to the target variable Q_2^+ . The notation follows: U1 denotes the unique SURD causality from Q_1 to Q_2^+ , i.e., $\Delta I_{1 \rightarrow 2}^U$, with analogous definitions for other components (i.e., R and S). All SURD and state-dependent causal contributions are normalized by the mutual information $I(Q_2^+; Q_1, Q_2)$. The causal maps for redundant and synergistic components show averages over all states of q_2 . In each panel, the bottom row displays the temporal evolution of $q_1(n)$, color-coded according to the corresponding instantaneous state-dependent causal contribution figure 2, as the dynamics governing Q_2^+ remain unchanged and the newly introduced variable Q_3^+ does not influence Q_2^+ . Figure S8 shows the corresponding results for the target variable Q_3^+ . First, for the redundant causality $\Delta\mathcal{C}_{13 \rightarrow 3}^R$, we observe that positive states $q_1 > 0$ are causal, whereas negative states $q_1 < 0$ are non-causal. This indicates that causality is shared between Q_1 and Q_3 about Q_3^+ when $q_1 < 0$, as both Q_1 and Q_3 contribute information necessary to determine the future state of Q_3^+ . For the unique causality $\Delta\mathcal{C}_{1 \rightarrow 3}^U$, the results clearly show that negative q_1 states ($q_1 < 0$) are causal ($\Delta\mathcal{C}_{1 \rightarrow 3}^U > 0$). This implies that Q_1 uniquely drives Q_3^+ in this regime without requiring information from Q_3 . Finally, for the synergistic causality $\Delta\mathcal{C}_{13 \rightarrow 3}^S$, causal states appear for $q_1 > 0$, where both Q_1 and Q_3 are jointly required to predict Q_3^+ . This synergistic behavior implies that additional predictive information about Q_3^+ only emerges when considering the combination of Q_1 and Q_3 together, beyond what each variable can contribute individually through their unique causalities.

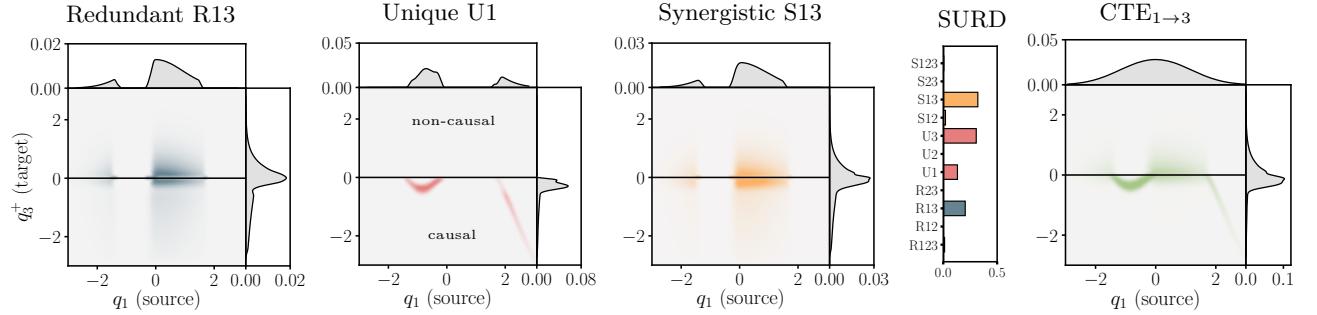
Figure S8 also illustrates the SURD causalities. The dominant causal influences originate from combinations of the source variables Q_1 and Q_3 , as reflected in the redundant ($\Delta I_{13 \rightarrow 3}^R$), unique ($\Delta I_{1 \rightarrow 3}^U$ and $\Delta I_{3 \rightarrow 3}^U$), and synergistic ($\Delta I_{13 \rightarrow 3}^S$) contributions. These results are consistent with the functional dependencies of the system, where Q_1 and Q_3 govern the future evolution of Q_3^+ depending on the specific state of Q_1 . The analysis further shows that the role of Q_2 in determining Q_3^+ is negligible, indicating that the method correctly identifies the relevant causal variables. However, as in the benchmark shown in the main text, SURD causalities do not reveal the specific system states in which each type of causality becomes active.

Additionally, we compare the results with those obtained using state-dependent CTE, shown on the right-hand side of Fig. S8. As in the case of Q_2^+ , the CTE framework aggregates all causal contributions into a single measure, which obscures the distinctions between redundant, unique, and synergistic interactions. In particular, CTE fails to reveal that when $q_1 < 0$, Q_1 uniquely determines the future of Q_3^+ , whereas when $q_1 > 0$, information from both Q_1 and Q_3 is required.

$$q_3(n+1) = \begin{cases} \cos[q_1(n)] + 0.1w_3(n) & \text{if } q_1(n) < 0 \\ 0.9q_3(n) + 0.1w_3(n) & \text{otherwise} \end{cases}$$



$$q_3(n+1) = \begin{cases} q_1(n) \cos[q_1(n)] + 0.1w_3(n) & \text{if } q_1(n) \cos[q_1(n)] + 0.1w_3(n) < 0 \\ q_3(n) + 0.1w_3(n) & \text{otherwise} \end{cases}$$



$$q_3(n+1) = \begin{cases} \sin[q_1(n)q_2(n)] + 0.1w_3(n) & \text{if } q_1(n)q_2(n) > 0 \\ w_3(n) & \text{otherwise} \end{cases}$$

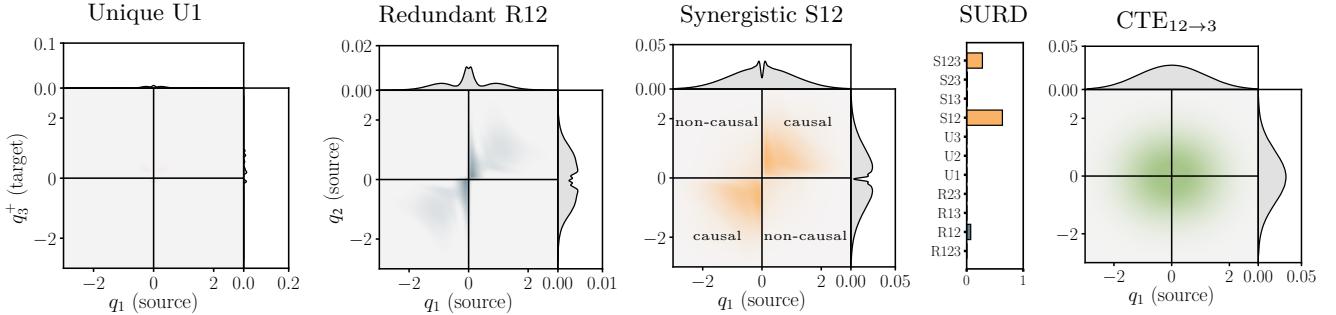


Figure S8: State-wise decomposition of causal states for (top) unique source-dependent, (middle) target-dependent cases, and (bottom) synergistic source-dependent cases. From left to right, panels show: state-dependent redundant (R), unique (U), and synergistic (S) causal contributions in blue, red, and yellow, respectively; the total causalities to the target variable Q_3^+ from SURD; and the state-dependent conditional transfer entropy (CTE). The notation used here is such that U1 represents the unique causality from Q_1 to Q_3^+ , i.e., $\Delta I_{1 \rightarrow 3}^U$, and similarly for other terms. The causal maps for redundant and synergistic causalities are average over all the states q_3 .

S2.3 Target-dependent unique causality

The second case illustrates the decomposition of causality in a target-dependent system, extending the validation example from the main text by introducing an additional variable, Q_3^+ , which also exhibits target-dependent dynamics. The objective is to evaluate whether the method can detect a different type of state-dependent behavior and effectively handle confounding effects, given that the source variable Q_1 influences both Q_2^+ and Q_3^+ . In this system, the relationship is defined such that $f(q_1) \rightarrow q_3^+$ when $f(q_1) < 0$, and $q_3 \rightarrow q_3^+$ when $f(q_1) \geq 0$, where $f(\cdot)$ is

a nonlinear function mapping the past of the source variable to the future state of the target. The dynamics of Q_2^+ are identical to those presented in the main text, ensuring that any observed differences can be attributed solely to the introduction of Q_3^+ and its interaction with Q_1 . The variable Q_1 is drawn from a standard normal distribution (mean zero, unit variance), while Q_3 is influenced by stochastic forcing W_3 , modeled as a Gaussian random variable with mean 1 and unit standard deviation. The dynamics of the target variable Q_3^+ are illustrated in Fig. S8.

First, it was assessed that the results for Q_2^+ remain unchanged after introducing the additional variable Q_3 , as the relevant causalities are identical to those reported in the main text. This demonstrates that the proposed method is robust to confounding effects, particularly in scenarios where a single variable simultaneously influences multiple targets.

The results for Q_3^+ are presented in Fig. S8. Similarly to the findings for Q_2^+ , we observe that the total causality maps obtained from SURD for Q_3^+ —both in the source-dependent and target-dependent regimes—appear to be remarkably similar, despite the underlying differences in the system interactions. This highlights the limitation of aggregate causal measures and underscores the value of a state-resolved decomposition, which provides a more faithful representation of the underlying causal structure. Figure S8 illustrates the state-dependent decomposition of redundant ($\Delta C_{13 \rightarrow 3}^R$), unique ($\Delta C_{1 \rightarrow 3}^U$), and synergistic ($\Delta C_{13 \rightarrow 3}^S$) causalities as functions of the source state q_1 and target state q_3^+ . The analysis clearly shows that the unique causality $\Delta C_{1 \rightarrow 3}^U$ varies primarily with the target state, rather than the source. In particular, negative values of the target variable ($q_3^+ < 0$) correspond to the strongest unique influence from Q_1 , reflecting the regime where Q_1 alone drives Q_3^+ . In contrast, both the synergistic $\Delta C_{13 \rightarrow 3}^S$ and redundant $\Delta C_{13 \rightarrow 3}^R$ causalities exhibit nonzero contributions across the full range of q_3^+ , indicating that Q_1 and Q_3 jointly provide information about Q_3^+ in both positive and negative target states.

Figure S8 also compares the state-dependent causalities with the results obtained from state-dependent CTE for Q_3^+ . As in the source-dependent case, state-dependent CTE is unable to disentangle redundant, unique, and synergistic causal interactions. For instance, it does not capture the unique causality from Q_1 to Q_3^+ when $q_3^+ < 0$, nor does it reveal that the states for which causality becomes redundant or synergistic.

S2.4 Source-dependent synergistic causality

This benchmark case involves a system in which the direction of causality varies synergistically depending on the joint intensity of two source variables. Specifically, we consider a system with three variables, where $[q_1, q_2] \rightarrow q_3^+$ when $q_1 q_2 > 0$, and $w_3 \rightarrow q_3^+$ otherwise. The source variables Q_1 and Q_2 are independent and sampled from a standard normal distribution (mean zero, unit variance). The target variable Q_3 is also influenced by a stochastic forcing term W_3 , which is independent, normally distributed with zero mean and unit variance, and not included in the set of observed variables.

The results for the state-dependent causalities are shown in Fig. S8, alongside the corresponding SURD causalities. The key insight from this analysis arises from the synergistic term $\Delta C_{12 \rightarrow 3}^S$, which clearly highlights strong synergistic effects in regions where $q_1 q_2 > 0$ —i.e., in the first and third quadrants of the q_1 – q_2 space. In contrast, no causality is detected in the regime where $q_1 q_2 \leq 0$, as the dynamics of Q_3^+ are entirely driven by the unobserved stochastic forcing W_3 in this region. The regime $q_1 q_2 > 0$ also exhibits a small amount of redundant causality, as captured by the SURD decomposition, though its magnitude remains relatively low. Moreover, the unique causalities from Q_1 and Q_2 are nearly zero throughout, reflecting the fact that neither source variable alone is sufficient to drive Q_3^+ —causal influence emerges only when both are considered jointly.

Finally, Fig. S8 also compares the state-dependent causalities with those obtained from CTE. Specifically, the figure presents the results for the combination $CTE_{12 \rightarrow 3}$, where no discernible difference is observed between the cases $q_1 q_2 > 0$ and $q_1 q_2 < 0$. Hence, CTE is unable to capture state-dependent relationships involving synergistic effects.

S2.5 Coupled Rössler–Lorenz system

We study a coupled version of the Lorenz system and the Rössler system. The former was developed by Lorenz as a simplified model of a viscous fluid flow. Rössler proposed a simpler version of the Lorenz's equations in order to

facilitate the study of its chaotic properties. The governing equations are:

$$\frac{dq_1}{dt} = -6[q_2 + q_3], \quad (\text{S34a})$$

$$\frac{dq_2}{dt} = 6[q_1 + 0.2q_2], \quad (\text{S34b})$$

$$\frac{dq_3}{dt} = 6[0.2 + q_3[q_1 - 5.7]], \quad (\text{S34c})$$

$$\frac{dq_4}{dt} = 10[q_5 - q_4], \quad (\text{S34d})$$

$$\frac{dq_5}{dt} = q_4[28 - q_6] - q_5 + c(q_2)q_2^2, \quad (\text{S34e})$$

$$\frac{dq_6}{dt} = q_4q_5 - \frac{8}{3}q_6, \quad (\text{S34f})$$

where $[Q_1, Q_2, Q_3]$ correspond to the Rössler system and $[Q_4, Q_5, Q_6]$ to the Lorenz system. The coupling between the systems is unidirectional from the Rössler system to the Lorenz system from $Q_2 \rightarrow Q_5$ via the parameter $c(q_2)$ depending on the intensity of q_2 . The aim is to infer the states of the Rössler system $[Q_1, Q_2, Q_3]$ that are the most causal to Q_5 from the Lorenz system. Therefore, the observable source variables are $\mathbf{Q} = [Q_1, Q_2, Q_3]$. The system was integrated for $10^6 t_{\text{ref}}$ where t_{ref} is the time for which $I(Q_1^+; Q_1)/I(Q_1; Q_1) = 0.5$. The time-lag selected for causal inference is the one that maximizes the cross-induced unique causality for Q_5 and 25 bins per variable were used to partition the observed phase space.

This section explores two coupling scenarios between the Rössler and Lorenz systems. In the first scenario, the coupling from q_2 (Rössler system) to q_5 (Lorenz system) is activated only when the states of q_2 are negative; in the second, the coupling is active only for positive q_2 values. When $q_2 > 0$, the Rössler attractor tends to occupy regions where q_3 reaches large values, whereas for $q_2 < 0$, q_3 remains negative and relatively stable. This state-dependent coupling results in distinct dynamic behaviors in the Lorenz system, giving rise to two different attractors depending on the coupling regime. The corresponding attractors are shown in Fig. S9(b) for the case where coupling is active for $q_2 < 0$, and in Fig. S9(c) for the case where coupling is active for $q_2 > 0$.

The results from the SURD decomposition for both coupling scenarios are presented in Figs. S9(d) and S9(e), respectively. In both cases, the most prominent unique causality originates from variable q_2 , quantified by $\Delta I_{2 \rightarrow 5}^U$. This indicates that q_2 has a direct and distinctive influence on the target variable q_5 . In addition, a strong synergistic causality from the pair (q_1, q_2) to q_5 , denoted by $\Delta I_{12 \rightarrow 5}^S$, is observed. This suggests that q_1 also contributes to the evolution of q_5 , despite not being directly coupled in the Lorenz equations. This influence arises due to the interaction between q_1 and q_2 within the original Rössler system, which becomes manifest over the integration time scale used in the analysis—ultimately causing all variables to appear in the right-hand side of the system equations.

As in the other cases discussed above, the Lorenz attractor exhibits visibly different structures under the two coupling scenarios. However, the total causal influence measured by SURD remains largely similar across both cases. This highlights once again that SURD alone may not fully capture the nuanced, state-dependent dynamics of the system. In such instances, decomposing causality into state-specific components offers additional insights. This decomposition is illustrated in Figs. S9(f) and S9(g), where it becomes clear that negative q_2 states are the most influential when the coupling is active only for $q_2 < 0$, while the opposite holds when the coupling is restricted to $q_2 > 0$. Moreover, the decomposition reveals that positive q_2 states still contribute relevant information about q_5 in the $q_2 < 0$ coupling regime, provided that $c \neq 0$, within the time scale considered for causal inference. This effect arises from the tendency of the system to transition into the $q_2 < 0$ regime shortly after experiencing large positive values of q_2 , thus allowing past positive states to carry predictive information about the future. A similar phenomenon occurs in the $q_2 > 0$ coupling scenario, where large negative q_2 states can retain residual influence on q_5^+ over the time-scale considered.

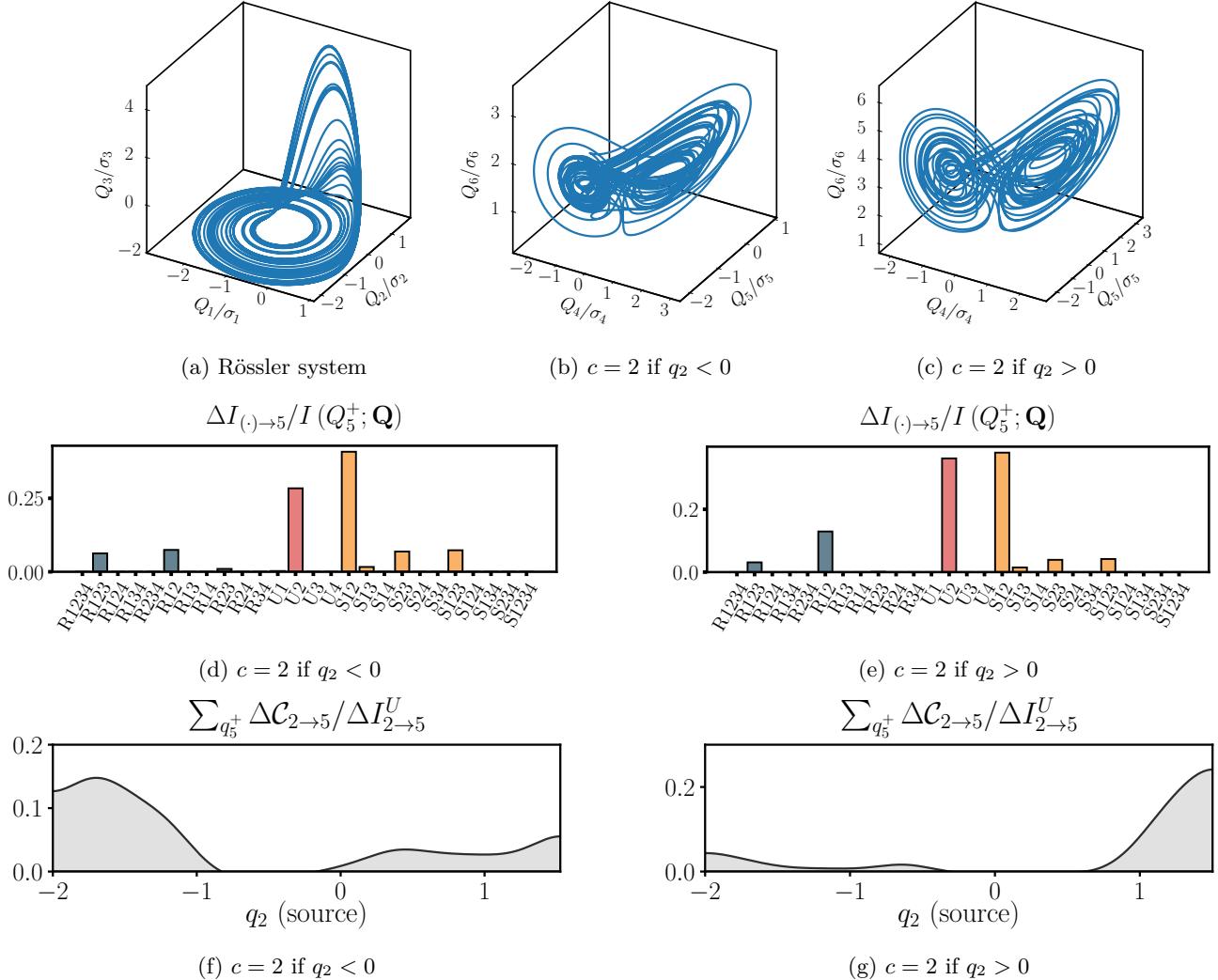


Figure S9: Coupled Rössler–Lorenz system. The upper panels show excerpts of the trajectories pertaining to (a) the Rössler system $[Q_1, Q_2, Q_3]$ and the Lorenz system $[Q_4, Q_5, Q_6]$ with (b) $c = 2$ if $q_2 < 0$ and (c) $c = 2$ if $q_2 > 0$. Redundant (R), unique (U), and synergistic (S) causalities from SURD among $[Q_1, Q_2, Q_3]$ for (d) $c = 2$ if $q_2 < 0$ and (e) $c = 2$ if $q_2 > 0$. The causality leak for each variable is also shown in the right-hand side bar. Panels (f) and (g) show the state-dependent decomposition of the unique causality $\Delta C_{2 \rightarrow 5}^U$ for both couplings.

S3 Analysis of non-causal contributions

We examine the non-causal contributions corresponding to the cases analyzed in the main text. These contributions arise in instances where knowledge of the source states leads to a loss of information—i.e., an increase in uncertainty—about the target variable, relative to the previous level of SURD causality. An alternative interpretation is that non-causal contributions indicate causal influence that has already been accounted for at a lower level in the SURD hierarchy. For example, consider a system with one redundant, one unique, and one synergistic S-causality component:

- Non-causal redundant contributions reflect a loss of information relative to the baseline information already contained in the target q_j^+ .
- Non-causal unique contributions represent a loss of information relative to what is already explained by the redundant component.
- Non-causal synergistic contributions represent a loss relative to the unique component.

In all cases, these non-causal terms are negative by construction. Importantly, the average over both causal and non-causal contributions recovers the total SURD causality at each level.

$$q_2(n+1) = \begin{cases} \sin[q_1(n)] + 0.1w_2(n) & \text{if } q_1(n) > 0 \\ 0.9q_2(n) + 0.1w_2(n) & \text{otherwise} \end{cases}$$

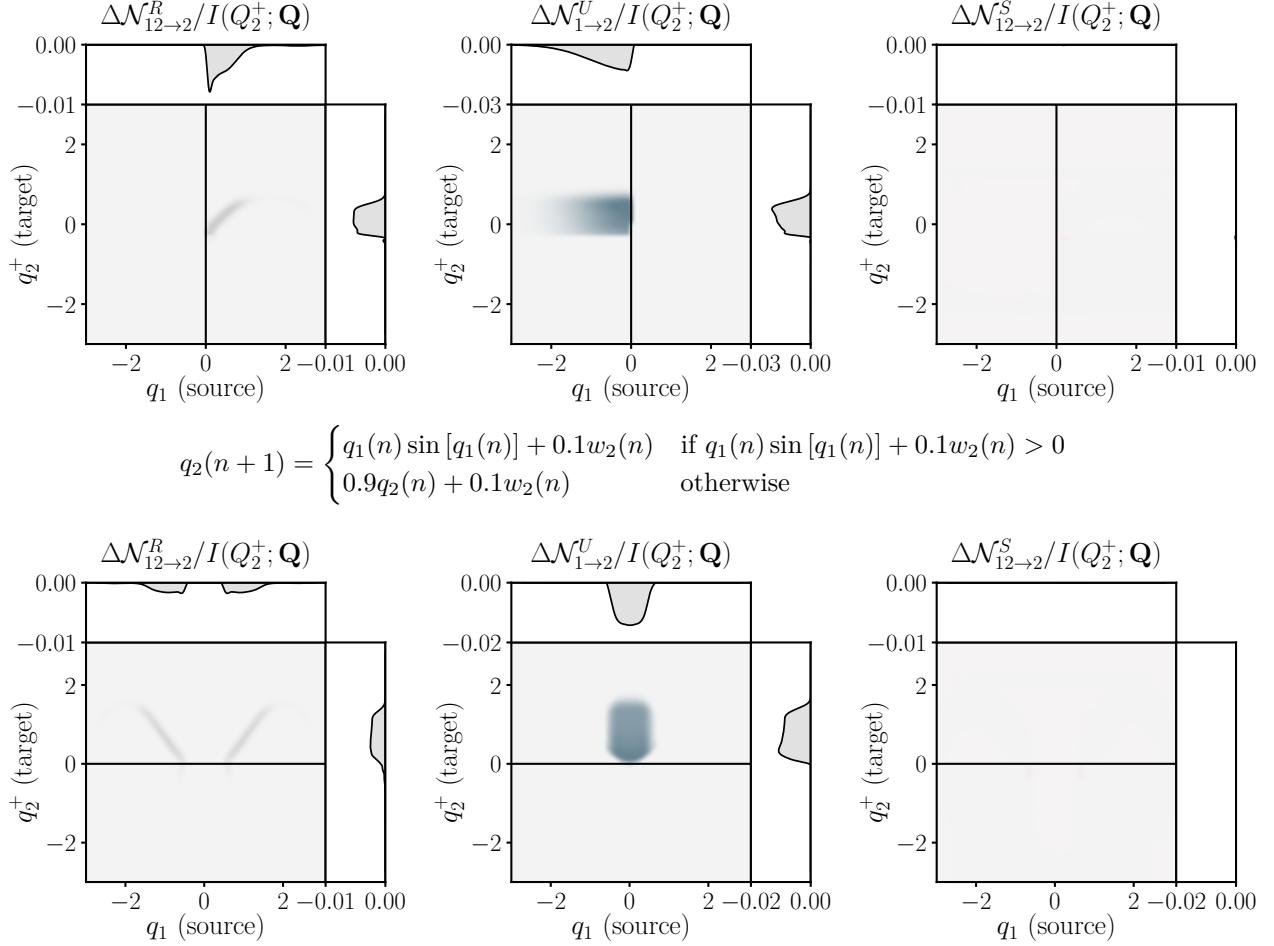


Figure S10: Non-causal redundant (R), unique (U), and synergistic (S) states for (top) source-dependent and (bottom) target-dependent benchmark cases. The maps for redundant and synergistic components illustrate (top) the sum over all states q_2 and (bottom) the specific state q_2 that maximizes the total sum of individual contributions.

S3.1 Source-dependent and target-dependent systems

First, we examine the non-causal components for the two benchmark systems (source-dependent and target-dependent) discussed in the validation section of the main text. Figure S10 illustrates the decomposition of non-causal states into their redundant, unique, and synergistic components for the target variable Q_2^+ in both cases.

Starting with the source-dependent system, recall that the causal states exhibited a strong unique contribution for $q_1 > 0$, where Q_1 uniquely drives Q_2^+ , as well as redundant and synergistic causalities for $q_1 < 0$, where Q_1 and Q_2 jointly influence Q_2^+ . For the non-causal contributions, the redundant component $\Delta\mathcal{N}_{12\rightarrow 2}^R$ is primarily located in the region $q_1 > 0$, where no additional information is gained beyond what is already contained in the target variable Q_2^+ . This is consistent with the fact that, in this regime, no redundancy arises between the q_1 and q_2 states. For the unique component $\Delta\mathcal{N}_{1\rightarrow 2}^U$, the non-causal states are concentrated around $q_1 < 0$, indicating that the causal influence of Q_1 in this region is already captured by the total redundant contribution $\Delta I_{12\rightarrow 2}^R$. Finally, the non-causal synergistic component $\Delta\mathcal{N}_{12\rightarrow 2}^S$ is negligible across the state space.

In the target-dependent system, causal states were previously identified in the region $q_2^+ > 0$ for unique causality, and in the region $q_2^+ < 0$ for redundant and synergistic causalities. Figure S10 shows the corresponding non-causal contributions. The redundant non-causal component displays the opposite trend of its causal counterpart: non-causal states are now concentrated in the region $q_2^+ > 0$, where no additional information is gained beyond what is already contained in the target variable $H(Q_2^+)$. This behavior is consistent with the functional dependencies of the system,

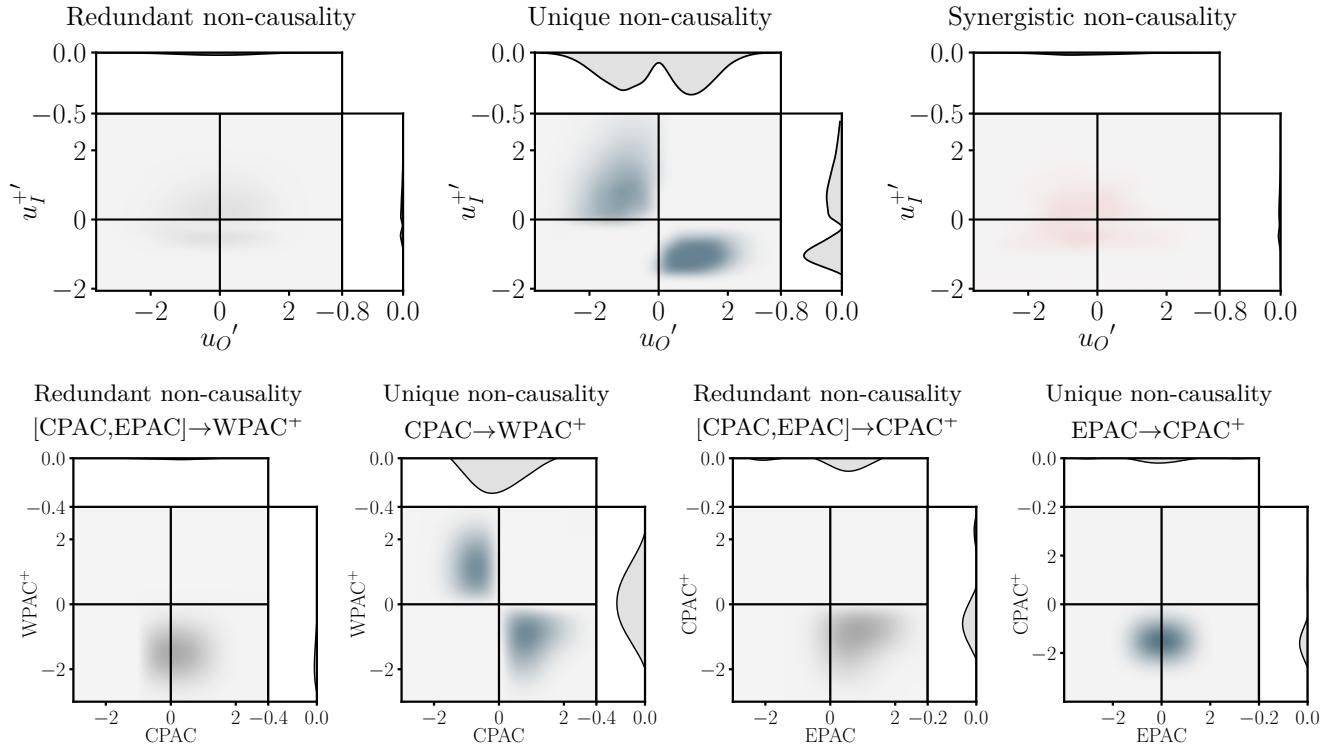


Figure S11: Non-causal redundant (R), unique (U), and synergistic (S) states for the two applications discussed in the main text: (top) inner/outer interactions in a turbulent boundary layer and (bottom) Walker circulation in the tropical Pacific. Note that for some of the panels, the color scale has been intentionally saturated to enhance visual contrast; however, the causalities are of small magnitude, as evidenced by the horizontal and vertical projections.

where no overlap of information between q_1 and q_2 was imposed for $q_2^+ > 0$. Additionally, non-causal contributions to the unique component appear in the same region ($q_2^+ > 0$), particularly around $q_1 = 0$, where no causal influence had been detected in the earlier analysis. This further confirms that these states contribute little to the predictability of Q_2^+ from Q_1 alone. Finally, the non-causal synergistic component $\Delta\mathcal{N}_{12 \rightarrow 2}^S$ remains negligible across the entire state space.

S3.2 Turbulent boundary layer and Walker circulation

We analyze the decomposition of non-causal contributions to the causalities in the turbulent boundary layer and Walker circulation applications, as shown in Fig. S11.

For the turbulent boundary layer, we observe that the redundant and synergistic non-causal components are essentially negligible. The dominant contribution arises from the unique non-causal component, which is primarily concentrated in the second ($u'_O > 0, u_I^+ < 0$) and fourth ($u'_O < 0, u_I^+ > 0$) quadrants of the $u'_O - u_I^+$ space. These quadrants correspond to states in which the outer- and inner-layer motions deviate from the mean velocity in opposite directions—indicating an asynchronous relationship between the two layers. Notably, the non-causal states in the fourth quadrant contribute more strongly to $\Delta\mathcal{N}_{OI \rightarrow I}^U$ than those in the second quadrant. This asymmetry aligns with the flow structures shown in Fig. 4Causality between inner and outer flow motions in a turbulent boundary layer. (a) Instantaneous visualization of the streamwise velocity field. (b) Temporal evolution of the streamwise velocity at two fixed wall-normal locations: the outer layer ($y_O = 0.3\delta$) and the inner layer ($y_I = 4\nu/u_\tau$). (c) Redundant ($\Delta I_{OI \rightarrow I}^R$), unique ($\Delta I_{(.) \rightarrow I}^U$), and synergistic ($\Delta I_{OI \rightarrow I}^S$) SURD causalities to the future inner-layer velocity $u_I^+ = u_I(t + \Delta T)$. The time lag ΔT is chosen to maximize cross-induced unique causality. (d) State-dependent redundant ($\Delta C_{OI \rightarrow I}^R$), unique ($\Delta C_{OI \rightarrow I}^U$), and synergistic ($\Delta C_{OI \rightarrow I}^S$) causalities as functions of the instantaneous streamwise velocities u_O (outer layer) and u_I^+ (inner layer). Note that for some of the panels, the color scale has been intentionally saturated to enhance visual contrast; however, the causalities are of small magnitude, as evidenced by the horizontal and vertical projections. (e) Time series of u_I^+ color-coded by the dominant state-dependent causal component at each time step. (f) Instantaneous and (g) conditionally averaged visualizations of flow fields corresponding to the causal

and non-causal states, respectively figure.4, where non-causal states are associated with a lack of structural coherence across the boundary layer. In these cases, no coherent flow structure spans both regions; instead, the velocity fields linked to second- and fourth-quadrant states appear fragmented and detached from the wall. This fragmentation reflects a breakdown of spatial coherence and a corresponding reduction in causal influence from the outer to the inner layer, consistent with the presence of non-causal states identified in Fig. S11.

For the Walker circulation cases, the decomposition of non-causal contributions is also shown in Fig. S11. For CPAC to WPAC⁺, the redundant non-causal component is negligible, while the unique non-causal contributions dominate the overall structure. The unique non-causal contribution from CPAC to WPAC⁺ is primarily concentrated in the second ($\text{CPAC} > 0$, $\text{WPAC}^+ < 0$) and fourth ($\text{CPAC} < 0$, $\text{WPAC}^+ > 0$) quadrants of the CPAC-WPAC⁺ state space. These quadrants correspond to states where the CPAC and WPAC⁺ anomalies differ in sign—that is, one variable deviates above its mean while the other falls below it. This misalignment suggests that, in these states, the causal influence from CPAC to WPAC⁺ is already captured by redundant information.

For EPAC to CPAC⁺, both redundant and unique non-causal contribution are relevant. We observe dominant non-causal contributions to the redundant interaction from [CPAC, EPAC] to CPAC⁺ for positive EPAC anomalies ($\text{EPAC} > 0$) and negative CPAC⁺ anomalies ($\text{CPAC}^+ < 0$). The unique non-causal contribution from EPAC to CPAC⁺ is primarily associated with states where EPAC fluctuates around its mean value and $\text{CPAC}^+ < 0$. These results highlight the assymetry between El Niño and La Niña events in the Walker circulation in the Pacific Region.

S4 Other methods for quantifying state-dependent influence among variables

We review four additional methods that provide state-dependent and/or time-varying measures of influence among variables: local transfer entropy (LTE) [87], specific transfer entropy (STE) [45], local Granger causality (LGC) [53], and time-varying Liang–Kleeman (TvLK) information flow [51]. LTE and STE are designed to quantify interactions among variables in dynamical systems without assuming an underlying causal structure, whereas LGC and TvLK are explicitly formulated as causal inference methods. For each approach, we provide a brief description and, where applicable, compare its performance against benchmark cases from the original studies to demonstrate the correct implementation and use of the method.

S4.1 State-dependent conditional transfer entropy

Transfer entropy (TE) is a non-parametric measure of directional information transfer between two time series [33]. The TE between two time signals quantifies how much past values of one variable help predict the present value of another. Here, we describe its multivariate extension, Conditional Transfer Entropy (CTE) [34, 30], which generalizes this concept by quantifying the information that a subset of variables \mathbf{Q}_i provides about a target Q_j^+ , conditioned on the remaining variables $\mathbf{Q}_{\bar{i}}$:

$$\overline{\text{CTE}}_{i \rightarrow j} = H(Q_j^+ | \mathbf{Q}_i) - H(Q_j^+ | \mathbf{Q}), \quad (\text{S35})$$

where $\mathbf{Q} = [\mathbf{Q}_i, \mathbf{Q}_{\bar{i}}]$. The expression above can be decomposed into state-dependent contributions following the same approach proposed in this study:

$$\text{CTE}_{i \rightarrow j} = \sum_{\mathbf{q}_{\bar{i}}} p(q_j^+, \mathbf{q}) \log_2 \frac{p(q_j^+ | \mathbf{q})}{p(q_j^+ | \mathbf{q}_{\bar{i}})}, \quad (\text{S36})$$

which allows one to identify the states of the source and target variables that contribute to the information transfer from \mathbf{Q}_i to Q_j^+ .

For example, in a system with the variables Q_1 , Q_2 , and target Q_j^+ , CTE simplifies to:

$$\text{CTE}_{2 \rightarrow j} = \sum_{q_1} p(q_j^+, q_1, q_2) \log_2 \left(\frac{p(q_j^+ | q_1, q_2)}{p(q_j^+ | q_1)} \right), \quad (\text{S37})$$

which denotes the ratio between the conditional probabilities $p(q_j^+ | q_1, q_2)$ and $p(q_j^+ | q_1)$ weighted by the joint probability $p(q_j^+, q_1, q_2)$. Comparing this result with the specific unique causality defined in Eq. (6equation.0.6), we observe that the key difference lies in the conditional probabilities: our method uses $p(q_j^+ | q_2)$ and $p(q_j^+ | q_1)$, rather than $p(q_j^+ | q_1, q_2)$ and $p(q_j^+ | q_1)$. This distinction allows our approach to more effectively isolate the influence of individual variable states by conditioning out the effects of other variables.

S4.2 Local transfer entropy

The localized definition of transfer entropy proposed by Lizier *et al.* [35] follows a similar derivation to the state-dependent CTE discussed above but focuses on the pointwise (state-specific) contributions to the expected information transfer. Accordingly, the local transfer entropy (LTE) from Q_i to Q_j^+ is defined as:

$$\text{LTE}_{i \rightarrow j} = \log_2 \frac{p(q_j^+ | \mathbf{q})}{p(q_j^+ | \mathbf{q}_{\bar{i}})}. \quad (\text{S38})$$

The approach exhibits several desirable properties: it is identically zero if and only if the total transfer entropy is zero, and its average over all source and target states recovers the total transfer entropy. Individual values can be either positive or negative—although the overall average is non-negative—with negative values interpreted as misleading or misinformative contributions.

We employ nearest-neighbor estimators for local transfer entropy based on the approach outlined in Refs. [44, 45]. Specifically, we adopt the Kraskov-Stögbauer-Grassberger (KSG) estimator for mutual information [88] in the context of TE computation. For a fixed nearest-neighbor parameter k and a single past time lag, the estimator of the local transfer entropy is defined as:

$$\hat{\text{LTE}}_{i \rightarrow j} = \psi(k) + \psi(N_{Q_j}(t; \rho_{t,k}) + 1) - \psi(N_{Q_j^+}(t; \rho_{t,k}) + 1) - \psi(N_{Q_i \times Q_j}(t; \rho_{t,k}) + 1), \quad (\text{S39})$$

where $\psi(\cdot)$ denotes the digamma function, and $\rho_{t,k}$ represents the distance to the k th-nearest neighbor of the state vector (Q_i, Q_j, Q_j^+) under the infinity norm. The function $N_S(t; \rho_{t,k})$ counts the number of sample points within a radius $\rho_{t,k}$ in the space \mathcal{S} . The TE can be recovered by averaging the LTE over all available samples. The parameter k controls the trade-off between bias and variance in the estimator $\tilde{\text{LTE}}_{i \rightarrow j}$. While the KSG-based estimator is asymptotically unbiased for any fixed k , finite sample sizes introduce a nonzero bias that decreases with smaller values of k . However, smaller k values also lead to increased estimator variance. Following standard practice, we fix $k = 4$ in all our analyses. To project the results onto a selected variable, we adopt the approach proposed in Ref. [45] for Q_j^+ . In this setting, the projection is carried out with respect to a selected variable, enabling the estimation of $\tilde{\text{LTE}}_{i \rightarrow j}$ as a function of both q_j^+ and q_i when projecting onto q_j .

S4.3 Specific transfer entropy

Darmon and Rapp [45] introduced specific transfer entropy (STE), an alternative to LTE that retains many of its desirable properties while introducing others, such as the non-negativity of individual terms and a clearer interpretation in terms of the state-dependent predictive impact of source states on target states. The STE formulation is derived by applying the law of iterated expectations to TE:

$$\text{TE}_{i \rightarrow j} = \sum_{q_j^+} \sum_{\mathbf{q}} p(q_j^+ | \mathbf{q}) \log_2 \frac{p(q_j^+ | \mathbf{q})}{p(q_j^+ | \mathbf{q}_t)} = \sum_{\mathbf{q}} p(\mathbf{q}) \sum_{q_j^+} p(q_j^+ | \mathbf{q}) \log_2 \frac{p(q_j^+ | \mathbf{q})}{p(q_j^+ | \mathbf{q}_t)}, \quad (\text{S40})$$

Darmon and Rapp [45] define STE using the inner conditional expectation from the decomposition of total transfer entropy. This quantity corresponds to the Kullback–Leibler divergence (D_{KL}) between $p(q_j^+ | \mathbf{q})$ and $p(q_j^+ | \mathbf{q}_t)$, and is therefore guaranteed to be non-negative:

$$\text{STE}_{i \rightarrow j} = \sum_{q_j^+} p(q_j^+ | \mathbf{q}) \log_2 \frac{p(q_j^+ | \mathbf{q})}{p(q_j^+ | \mathbf{q}_t)} = D_{KL} [p(q_j^+ | \mathbf{q}) || p(q_j^+ | \mathbf{q}_t)]. \quad (\text{S41})$$

An estimator of STE can be obtained from the LTE by performing a regression of the LTE values onto the past input-output states. To achieve this, we follow the approach in Ref. [45] and use a k_{reg} -nearest-neighbor regression as a nonparametric smoother. Given a nearest-neighbor parameter k_{reg} , the estimator of the STE is given by:

$$\tilde{\text{STE}}_{i \rightarrow j} = \sum_{t=p+1}^T W_{t, k_{\text{reg}}} \cdot \tilde{\text{LTE}}_{i \rightarrow j}(q_j^+, q_i, q_j), \quad (\text{S42})$$

where the weights $W_{t, k_{\text{reg}}}$ are defined as $1/k_{\text{reg}}$ if the pair $q_i - q_j$ is among the k_{reg} nearest neighbors, and zero otherwise. In this formulation, k_{reg} controls the degree of smoothing and balances the bias-variance trade-off in the estimation. Following standard practice, we set $k_{\text{reg}} = \lfloor \sqrt{T} \rfloor$, where T is the total number of samples. Notably, unlike kernel density-based estimators of the specific entropy rate, the k th-nearest-neighbor-based estimator can yield negative STE values.

We use the implementation and examples from Darmon & Rapp [45] to evaluate the performance of LTE and STE on the benchmark cases discussed in this study. To verify the implementation, we first consider the smooth threshold autorregressive model with an exogenous driver (STARX). The system consists of a nonlinear autorregressive model that incorporates an exogenous input (Q_1) whose value induces a smooth thresholding between two linear autorregressive models for the output (Q_2). Mathematically, the model can be expressed as:

$$q_1(n+1) = 0.8q_1(n) + w_1(n) \quad (\text{S43})$$

$$q_2(n+1) = \phi[q_1(n)] [0.5q_2(n) + 2w_2(n)] + \{1 - \phi[q_1(n)]\} [-0.5q_2(n) + w_2(n)], \quad (\text{S44})$$

where W_1 and W_2 are mutually independent, identically distributed standard normal random variables. The function $\phi[q_1(n)]$ controls the switching between two autoregressive models and corresponds to a sigmoidal function obtained by rescaling the cumulative distribution function. Figure S12 displays a segment of the time evolution of Q_1 and Q_2 , where gray boxes highlight intervals during which Q_1 remains negative for more than 10 consecutive time steps. We observe that when $q_1 < 0$, the output Q_2 exhibits low-amplitude oscillations, whereas for $q_1 > 0$, the oscillations in Q_2 are amplified. Additionally, Fig. S12 shows the results for LTE and STE computed for this system. As expected, LTE values fluctuate between positive and negative, while STE remains strictly non-negative. Peaks in STE align with the shaded regions (i.e., intervals where $q_1 < 0$). When examining STE as a function of both q_1 and q_2 states, we

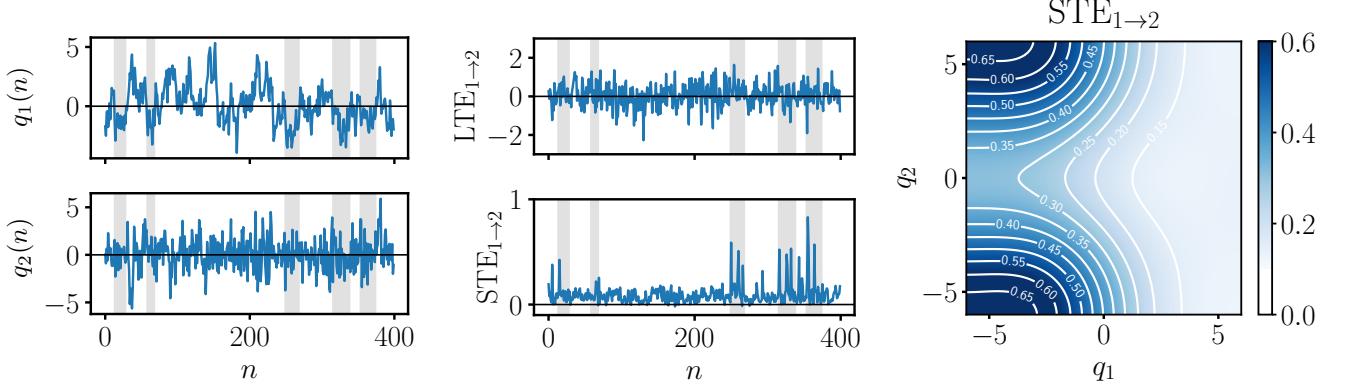


Figure S12: Verification example for local and specific transfer entropy from Darmon & Rapp [45]. The left panels show the temporal evolution of $q_1(n)$ and $q_2(n)$. The middle panels show the temporal evolution of local transfer entropy (LTE) as given by Eq. (S38) and specific transfer entropy (STE) as given by Eq. (S40). The gray boxes show the regions in which $q_1(n) < 0$ for at least 10 consecutive steps. The right panel depicts STE as a function of the values of source variables q_1 and q_2 .

find that the most significant influence from q_1 to q_2 occurs when $q_1 < 0$ and $|q_2| > 2.5$. These results are consistent with the results reported by Darmon and Rapp [45], which verifies our implementation.

S4.4 Local Granger causality

Local Granger causality (LGC), proposed by Stramaglia *et al.* [53], aims at providing a linear counterpart of LTE. In this case, the method is obtained from a vector autoregressive (VAR) model. Consider a system with n zero-mean variables $\mathbf{Q} = [Q_1, Q_2, \dots, Q_n]$, which is stationary and ergodic with time-invariant variances and covariances. The distribution of \mathbf{Q} is assumed to be a multivariate Gaussian in the stationary regime. Given these conditions, the LGC from a vector of source variables \mathbf{Q}_i to a target variable Q_j^+ is defined as:

$$\text{LGC}_{i \rightarrow j}(q_j^+, \mathbf{q}_i, \mathbf{q}_{\neq i}) = \text{GC}_{i \rightarrow j} + \mathcal{L}(q_j^+, \mathbf{q}_i, \mathbf{q}_{\neq i}), \quad (\text{S45})$$

where $\mathbf{q}_{\neq i}$ represents the states of the vector of variables in \mathbf{Q} that are not \mathbf{Q}_i . In the previous equation, the first term is independent of the states q_j^+ and \mathbf{q}_i , and corresponds to the classical Granger causality. In contrast, the term $\mathcal{L}(q_j^+, \mathbf{q}_i)$ captures the instantaneous, state-specific fluctuations around this baseline value [53]. Specifically, these terms can be mathematically expressed as:

$$\text{GC}_{i \rightarrow j} = \log \frac{|\Sigma_{ji}| |\Sigma_{j\neq i}|}{|\Sigma_{j\neq i}| |\Sigma_{\neq i}|}, \quad (\text{S46})$$

where $|\cdot|$ stands for the determinant of a matrix, and

$$\mathcal{L}(q_j^+, \mathbf{q}_i, \mathbf{q}_{\neq i}) = \mathbf{Z}_{\neq i}^\top \Sigma_{\neq i}^{-1} \mathbf{Z}_{\neq i} + \mathbf{Z}_{j\neq i}^\top \Sigma_{j\neq i}^{-1} \mathbf{Z}_{j\neq i} - \mathbf{Z}_{j\neq i}^\top \Sigma_{j\neq i}^{-1} \mathbf{Z}_{\neq i} - \mathbf{q}_{\neq i}^\top \Sigma_{\neq i}^{-1} \mathbf{q}_{\neq i}, \quad (\text{S47})$$

where $\mathbf{Z}_{\neq i} = [\mathbf{q}_i^\top \mathbf{q}_{\neq i}^\top]^\top$, $\mathbf{Z}_{j\neq i} = [q_j^+ \mathbf{q}_{\neq i}^\top]^\top$, and $\mathbf{Z}_{j\neq i} = [q_j^+ \mathbf{q}_{\neq i}^\top \mathbf{q}_i^\top]^\top$ are the observations of the present and past states of the variables, and $\Sigma_{\neq i} = \mathbb{E}[\mathbf{Z}_{\neq i} \mathbf{Z}_{\neq i}^\top]$, $\Sigma_{j\neq i} = \mathbb{E}[\mathbf{Z}_{j\neq i} \mathbf{Z}_{j\neq i}^\top]$, $\Sigma_{j\neq i} = \mathbb{E}[\mathbf{Z}_{j\neq i} \mathbf{Z}_{j\neq i}^\top]$, and $\Sigma_{\neq i} = \mathbb{E}[\mathbf{q}_{\neq i} \mathbf{q}_{\neq i}^\top]$ are the covariance matrices. Importantly, $\text{LGC}_{i \rightarrow j}$ retains the property that its time-average recovers the standard Granger causality, since the term $\mathcal{L}(q_j^+, \mathbf{q}_i, \mathbf{q}_{\neq i})$ has a vanishing expected value. However, individual local values can be either positive or negative. According to Stramaglia *et al.* [53], negative LGC values indicate instances where the knowledge of the source variables misleads the prediction of the target.

To estimate LGC in practice, we first fit a VAR model globally to the data, selecting the model order via the Akaike information criterion. We then compute the covariance matrices required for the evaluation of Eq. (S45). To verify the implementation of the method, we consider the following toy model with two variables $\mathbf{Q} = [Q_1, Q_2]$

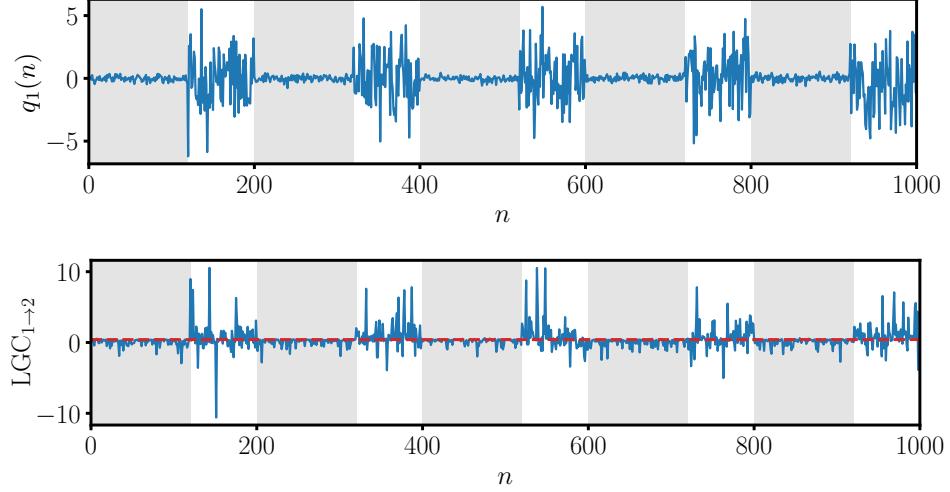


Figure S13: Time evolution of (top) $q_1(n)$ and (bottom) local Granger causality (LGC) from Q_1 to Q_2^+ for the toy model described in Eq. (S48). The dashed line in the bottom panel denotes the total Granger causality GC_{1to2} . Light gray boxes denote periods where $\sigma_1 = 0.2$, while white intervals correspond to $\sigma_1 = 2$.

reported by Stramaglia *et al.* [53]:

$$q_1(n) = w_1(n), \quad (S48)$$

$$q_2(n) = 0.2 q_2(n-1) + 0.4 q_1(n-1) + w_2(n), \quad (S49)$$

where $w_1(n)$ and $w_2(n)$ are independent white noise terms with standard deviation of σ_1 and σ_2 , respectively. In particular, σ_1 alternates between 0.2 and 2 at the switching times $n \in \{120, 200, 320, 400, 520, 600, 720, 800, 920\}$, while $\sigma_2 = 0.8$ remains constant throughout. An excerpt of the evolution of Q_1 is shown in Fig. S13. This system establishes a unidirectional causal relationship from $q_1(n)$ to $q_2^+ = q_2(n+1)$, which is modulated by the strength of the noise acting on Q_1 .

The LGC analysis for this system is presented in Fig. S13. The results exhibit pronounced fluctuations that span both positive and negative values, particularly during intervals where σ_1 is large. Notably, large negative values of local Granger causality arise when the noise term $w_2(n)$ drives the system in opposition to the causal influence of $q_1(n-1)$. According to the authors, such instances render the driver's contribution misinformative with respect to the target. Conversely, large positive fluctuations occur when the noise and the source variable act in the same direction, reinforcing the causal relationship. As discussed by Stramaglia *et al.* [53], these fluctuations reflect not merely variations in the noise amplitude, but rather the interplay between stochastic forcing and the underlying driver dynamics. The results reported here are consistent with those previously described by Stramaglia *et al.* [53], supporting the correct implementation of LGC.

S4.5 Time-varying Liang–Kleeman information flow

The last causal inference method considered in this study that supports time-resolved analysis is based on the concept of Liang–Kleeman (LK) information flow [40]. Consider a n -dimensional stochastic system governed by:

$$d\mathbf{q} = \mathbf{F}(\mathbf{q}, t) dt + \mathbf{B}(\mathbf{q}, t) d\mathbf{w}, \quad (S50)$$

where $\mathbf{F} = [F_1, F_2, \dots]$ is the drift vector representing the deterministic dynamics, \mathbf{B} is a matrix describing the stochastic perturbations, and \mathbf{w} is a vector of independent Wiener processes. Within this framework, the rate of information flow from a source variable Q_2 to a target variable Q_1 quantifies the amount of entropy contributed by Q_2 to the marginal entropy evolution of Q_1 . Mathematically, the information flow from Q_2 to Q_1 is given by:

$$T_{2 \rightarrow 1} = -\mathbb{E} \left[\frac{1}{p_1} \frac{\partial(F_1 p_1)}{\partial q_1} \right] + \frac{1}{2} \mathbb{E} \left[\frac{1}{p_1} \frac{\partial^2 [(b_{11}^2 + b_{12}^2)p_1]}{\partial q_1^2} \right], \quad (S51)$$

where p_1 is the marginal probability density function of Q_1 , F_1 is the drift component associated with Q_1 and \mathbb{E} is the expectation operator. The information flow is asymmetric; $T_{2 \rightarrow 1} \neq T_{1 \rightarrow 2}$, and it guarantees that if the evolution of Q_1 is independent of Q_2 , then $T_{2 \rightarrow 1} = 0$. In the special case where the system is linear, i.e., $\mathbf{F}(\mathbf{q}) = \mathbf{f} + A\mathbf{q}$ and \mathbf{B} is constant, and the joint distribution of $[Q_1, Q_2]$ follows a bivariate Gaussian distribution, Eq. (S51) simplifies to:

$$T_{2 \rightarrow 1} = \frac{\sigma_{12}}{\sigma_{11}} a_{12}, \quad (\text{S52})$$

where σ_{11} and σ_{12} are elements of the covariance matrix of $[Q_1, Q_2]$, and a_{12} is the coefficient representing the direct effect of Q_2 on Q_1 in the drift term. In practice, we do not know the parameters of the system, but an estimate can be obtained from time series:

$$T_{2 \rightarrow 1} = \frac{\Sigma_{12}\Sigma_{11}\Sigma_{2,d1} - \Sigma_{12}^2\Sigma_{1,d1}}{\Sigma_{11}^2\Sigma_{22} - \Sigma_{11}\Sigma_{12}^2}, \quad (\text{S53})$$

where Σ_{ij} denotes the sample covariance between Q_i and Q_j , and $\Sigma_{i,dj}$ is the sample covariance between Q_i and the time derivative of Q_j estimated via finite differences.

The time-varying extension of the LK formalism, known as TvLK [51], was developed by dynamically estimating covariances at each time step using a Kalman filter. Specifically, let $\mathbf{P}_{ij}(t)$ denote the covariance between variables Q_i and Q_j at time t , estimated via a square-root Kalman filter. The time-varying causality from Q_2 to Q_1 is then given by:

$$T_{2 \rightarrow 1}(t) = \frac{\mathbf{P}_{12}\mathbf{P}_{11}\mathbf{P}_{2,d1} - \mathbf{P}_{12}^2\mathbf{P}_{1,d1}}{\mathbf{P}_{11}^2\mathbf{P}_{22} - \mathbf{P}_{11}\mathbf{P}_{12}^2}, \quad (\text{S54})$$

Here, $\mathbf{P}_{i,dj}(t)$ represents the covariance between Q_i and the derivative of Q_j at time t .

We adopt the implementation described in Ref. [51] to estimate the TvLK information flow, applying a uniformly weighted moving average (UWMA) filter with a window length of 10 points and a sliding estimation window of 100 points. We verified that varying the estimation window length between 30 and 150 points had no significant impact on the results or the conclusions drawn from the analysis.

S4.6 Comparison of methods

A summary of the time-varying results obtained from existing methods for the benchmark cases was already provided in the main text. Here, we offer an additional comparison limited to those cases where the methods allow for a representation of causal influence on a source-target state map, similar to that used in our approach. These methods are CTE, LTE, and STE. For LGC and TvLK, such a map could in principle be constructed by averaging the time-resolved estimates over state bins. However, since these methods do not explicitly incorporate the same notion of state-dependent causality, we exclude them from this comparison. As discussed in the main text, we hypothesize that LGC would exhibit behavior similar to LTE in linear systems, while the estimates from TvLK information flow may lack the sensitivity required to capture the abrupt, short-time-scale causal transitions present in our benchmark cases.

We present the results for the source- and target-dependent benchmark cases analyzed in Fig. 2. State-dependent decomposition of causality for the source-dependent case. From left to right, the panels show: state-dependent redundant ($\Delta\mathcal{C}_{12 \rightarrow 2}^R$), unique ($\Delta\mathcal{C}_{1 \rightarrow 2}^U$), and synergistic ($\Delta\mathcal{C}_{12 \rightarrow 2}^S$) causal contributions in blue, red, and orange, respectively; and the SURD causalities to the target variable Q_2^+ . The notation follows: U1 denotes the unique SURD causality from Q_1 to Q_2^+ , i.e., $\Delta I_{1 \rightarrow 2}^U$, with analogous definitions for other components (i.e., R and S). All SURD and state-dependent causal contributions are normalized by the mutual information $I(Q_2^+; Q_1, Q_2)$. The causal maps for redundant and synergistic components show averages over all states of q_2 . In each panel, the bottom row displays the temporal evolution of $q_1(n)$, color-coded according to the corresponding instantaneous state-dependent causal contribution figure 2 and 3. State-dependent decomposition of causality for the target-dependent case. From left to right, the panels show: state-dependent redundant ($\Delta\mathcal{C}_{12 \rightarrow 2}^R$), unique ($\Delta\mathcal{C}_{1 \rightarrow 2}^U$), and synergistic ($\Delta\mathcal{C}_{12 \rightarrow 2}^S$) causal contributions in blue, red, and orange, respectively; and the SURD causalities to the target variable Q_2^+ . The notation follows: U1 denotes the unique SURD causality from Q_1 to Q_2^+ , i.e., $\Delta I_{1 \rightarrow 2}^U$, with analogous definitions for other components (i.e., R and S). All SURD and state-dependent causal contributions are normalized by the mutual information $I(Q_2^+; Q_1, Q_2)$. The causal maps for redundant and synergistic causalities are shown for the q_2 that maximizes the total sum of individual contributions. In each panel, the bottom row displays the temporal evolution of $q_2^+(n)$, color-coded according to the corresponding instantaneous state-dependent causal contribution figure 3 using CTE, LTE, and STE methods. We also include in the comparison the results for the additional target variable Q_3^+ , whose results using our proposed method were discussed in §S2. Figure S14 shows the results for the source-dependent case. We observe that CTE successfully detects the sinusoidal relationship between q_1 and q_2^+ for $q_1 > 0$, while also

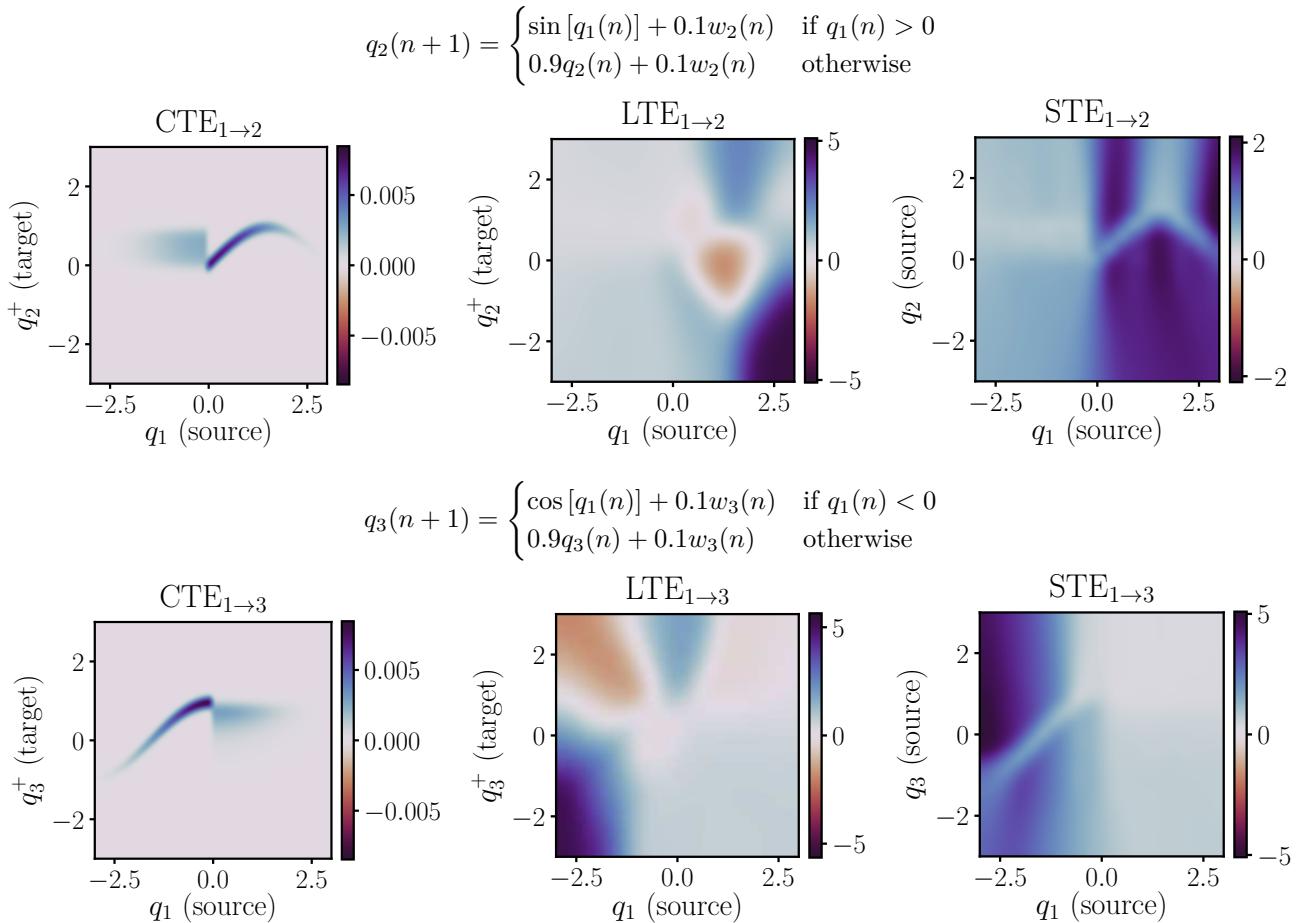


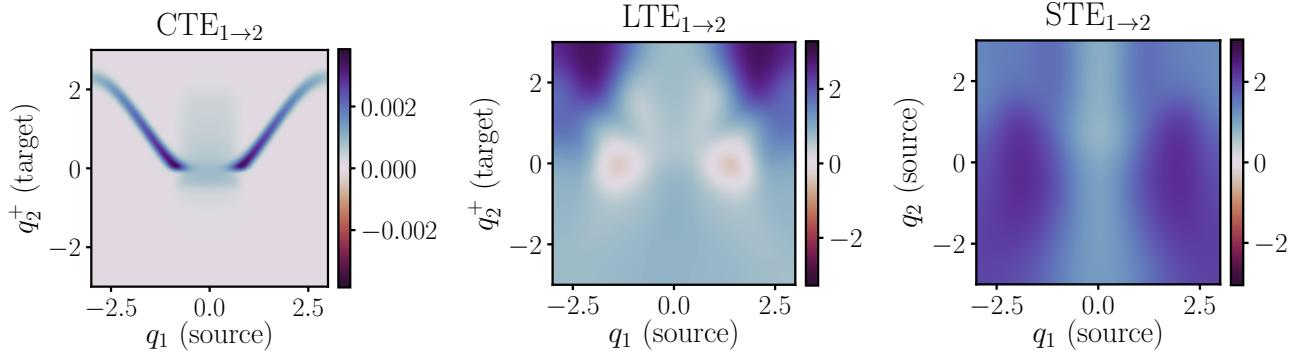
Figure S14: Results for the source-dependent system using conditional transfer entropy (CTE), local transfer entropy (LTE), and specific transfer entropy (STE).

identifying that $q_1 < 0$ provides some information about q_2^+ . However, CTE cannot clearly distinguish between the different types of causality in each case. In particular, CTE cannot detect that q_1 is the sole driver of q_2^+ for $q_1 > 0$ (unique causality) and that q_1 is required together with q_2 to drive q_2^+ for $q_1 < 0$ (synergistic causality). Our method accurately differentiates these dependencies, distinguishing between the unique dependency for $q_1 > 0$ and the synergistic dependency for $q_1 < 0$.

The maps for LTE and STE differ substantially from that of CTE. This discrepancy arises because LTE and STE are not weighted by the probability of each state, making them more sensitive to low-probability regions. As a result, their maps can be heavily influenced by rare states. Despite this difference, the highest values of $\text{LTE}_{1\rightarrow 2}$ and $\text{STE}_{1\rightarrow 2}$ occur for $q_1 > 0$, which is consistent with the relationship $q_1 \rightarrow q_2^+$ for $q_1 < 0$. Furthermore, both LTE and STE exhibit non-zero values across a range of states, reflecting the fact that q_1 is also required to drive the future state q_2^+ when $q_1 < 0$. A similar conclusion holds for q_3^+ across all three methods: the most significant contributions are observed in the region where $q_1 < 0$, again aligning with the underlying functional dependencies.

Next, we analyze the results for the target-dependent system. The results from CTE, LTE and STE are shown in Fig. S15. In this case, CTE successfully detects the non-linear relationship that occurs for $q_2^+ > 0$. However, it cannot fully differentiate between all types of causality, which leads to the identification of some information transfer even for $q_2^+ < 0$. According to the results from our method, this information transfer corresponds to redundant and synergistic causalities, while the unique causality from q_1 to q_2^+ occurs exclusively for $q_2^+ > 0$. This highlights the importance of disentangling causal relationships into their synergistic, redundant, and unique components at the state level. On the other hand, the most significant values for LTE are observed for $q_2^+ > 0$, which is consistent with the causal relationship from q_1 to q_2^+ occurring in this region. However, STE fails to capture dependencies based on target states, as it projects the results from LTE onto q_2^+ . A similar conclusion can be drawn for q_3^+ across all methods, although the most relevant values are now observed for $q_3^+ < 0$. Once again, STE is unable to capture target-dependent relationships.

$$q_2(n+1) = \begin{cases} q_1(n) \sin [q_1(n)] + 0.1w_2(n) & \text{if } q_1(n) \sin [q_1(n)] + 0.1w_2(n) > 0 \\ q_2(n) + 0.1w_2(n) & \text{otherwise} \end{cases}$$



$$q_3(n+1) = \begin{cases} q_1(n) \cos [q_1(n)] + 0.1w_3(n) & \text{if } q_1(n) \cos [q_1(n)] + 0.1w_3(n) < 0 \\ q_3(n) + 0.1w_3(n) & \text{otherwise} \end{cases}$$

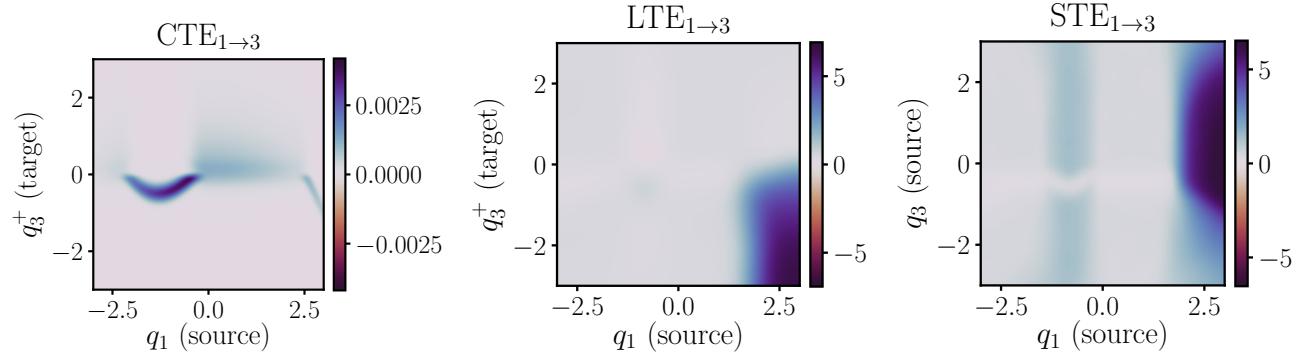


Figure S15: Results for the target-dependent system using conditional transfer entropy (CTE), local transfer entropy (LTE), and specific transfer entropy (STE).

The main conclusion from these results is that both LTE and STE differ significantly from the state-dependent CTE and the method proposed in this study. In particular, the maps generated by LTE and STE are heavily influenced by rare events with very low probabilities. However, as noted in the main text, this does not imply that the results from LTE and STE are incorrect. Instead, they offer a distinct local perspective on the total transfer entropy, emphasizing contributions from individual state realizations rather than probability-weighted averages.

S5 Climate science application

S5.1 Effect of the number of samples

In this section, we examine the impact of sample size on the climate science application. The climate time series consist of regional averages from reanalysis data spanning 1948–2012, totaling 780 monthly observations. This is the same dataset used in Ref. [16], and includes monthly surface pressure anomalies in the West Pacific (WPAC), and surface air temperature anomalies in the Central Pacific (CPAC) and East Pacific (EPAC).

To evaluate causality, we use a time lag of $\Delta T = 4$ months, which was found to maximize cross-induced unique causality between the variables. Figure S16 shows the evolution of state-dependent causal contributions across three different sample sizes ($N_{\text{samples}} = 390, 520$, and 780). In all cases, the joint distribution is estimated using a k -NN density estimator with $k = 7$ neighbors.

Due to sample size constraints, this analysis focuses exclusively on the redundant and unique causal contributions, as the estimation of synergistic causality is considerably more sensitive to data sparsity. To assess statistical significance, we conducted 1000 random permutations, yielding a p -value of zero for the combined redundant and

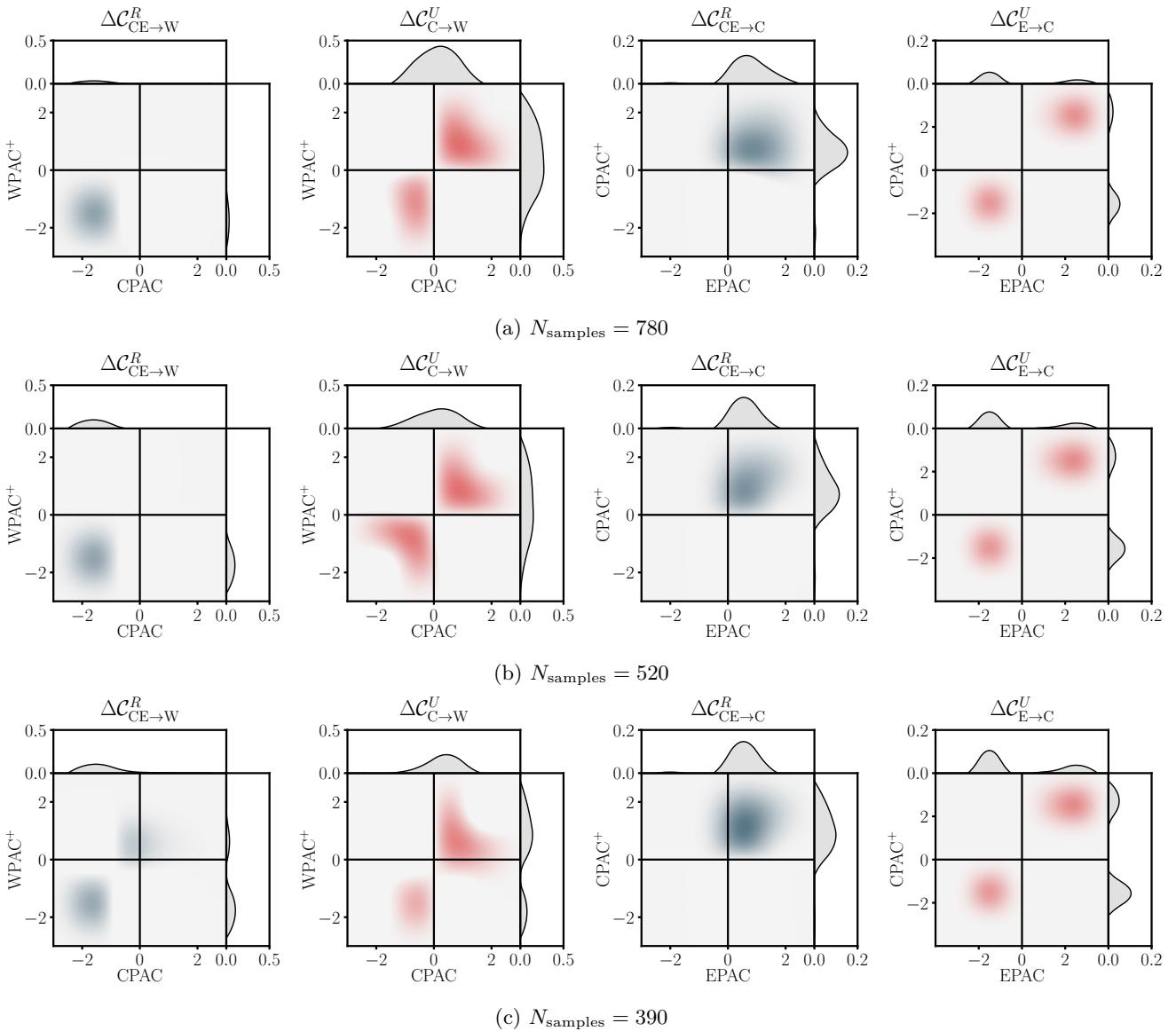


Figure S16: Effect of the number of samples on the state-dependent causalities $\Delta C_{CE \rightarrow W}^R$, $\Delta C_{C \rightarrow W}^U$, $\Delta C_{CE \rightarrow C}^R$, $\Delta C_{E \rightarrow C}^U$ for three different sample sizes: (a) $N_{\text{samples}} = 780$, (b) 520, and (c) 390. The state-dependent causalities are normalized with the mutual information.

unique causalities. In contrast, including the synergistic component in the test substantially increases the p -value, which highlights its reduced reliability under limited data and justifies its exclusion from the analysis.

The first key observation from these results is that the overall structure of the state-dependent causality maps remains consistent as sample size decreases. This robustness suggests that the main conclusions drawn from the full dataset are preserved, even under moderate data limitations. The only noticeable deviations appear at the lowest sample size, particularly for causal influences to WPAC⁺. Specifically, the redundant causality $\Delta C_{CE \rightarrow W}^R$ exhibits a new, localized region of positive causal values near the climatological mean of both CPAC and WPAC⁺. For the unique causality $\Delta N_{C \rightarrow W}^U$, we observe a small reduction of causal states in the region corresponding to $WPAC^+ < 0$ and $CPAC < 0$. This might be an indication that under reduced sample sizes, El Niño-like states (positive CPAC anomalies) remain more prominently causal, while La Niña-like states lose some significance. Notably, the state-dependent causalities to CPAC⁺ remain largely unchanged across all sample sizes.