Decomposing causality into its synergistic, unique, and redundant components

Álvaro Martínez-Sánchez [1] ✉, Gonzalo Arranz [1] & Adrián Lozano-Durán [1,2]

Causality lies at the heart of scientific inquiry, serving as the fundamental basis for understanding interactions among variables in physical systems. Despite its central role, current methods for causal inference face significant challenges due to nonlinear dependencies, stochastic interactions, self-causation, collider effects, and influences from exogenous factors, among others. While existing methods can effectively address some of these challenges, no single approach has successfully integrated all these aspects. Here, we address these challenges with SURD: Synergistic-Unique-Redundant Decomposition of causality. SURD quantifies causality as the increments of redundant, unique, and synergistic information gained about future events from past observations. The formulation is non-intrusive and applicable to both computational and experimental investigations, even when samples are scarce. We benchmark SURD in scenarios that pose significant challenges for causal inference and demonstrate that it offers a more reliable quantification of causality compared to previous methods.

The quest for understanding causality is the cornerstone of scientific discovery[1]. It is through the exploration of cause-and-effect relationships that we are able to understand a given phenomenon and shape the course of events through deliberate actions[2]. This has accelerated the proliferation of methods for causal inference, as they hold the potential to drive progress across multiple scientific and engineering domains, such as climate research[3], neuroscience[4], economics[5], epidemiology[6], social sciences[7], and fluid dynamics[8,9], among others.

A central aspect of causality is the concept of physical influence[10]: manipulation of the cause manifests as changes in the effects[1,11–13]. For example, prolonged exposure to elevated air pollution levels has a causal connection to a higher incidence of chronic respiratory conditions[14]. The precise definition of causality remains elusive, yet it must be distinguished from the concepts of association and correlation. Association indicates a statistical relationship between two variables in which they have a tendency to co-occur more often than would be expected by random chance. Yet, association does not automatically imply causation[15]. Association may arise from shared causes, statistical coincidences, or the influence of confounding factors. An example of association can be observed in the increased rates of chronic respiratory diseases in regions undergoing significant

deforestation. Although it may seem that deforestation directly contributes to respiratory health issues, this might primarily be due to the confounding factor of air pollution. Correlation, on the other hand, refers to a particular type of association that measures the monotonic strength and direction of variables[15–18]. Correlation implies association but not causation; causation implies association but not correlation[15]. Discerning between causality, association, and correlation poses a significant challenge in the development of methods for causal discovery. Here, we introduce an approach for causal inference that facilitates the study of complex systems in a manner that surpasses simple correlational and associational analyses.

The first factor to consider is the nature of interaction among variables. Three building blocks serve as the foundations of causal interactions[1]: mediator, confounder, and collider effects. These interactions can intertwine and manifest concurrently, leading to more complex causal networks. Therefore, accurately capturing these interactions is key to faithfully characterizing more general causal patterns. Consider the three events denoted by A, B, and C:

- Mediator variables ($A \rightarrow B \rightarrow C$) emerge in the causal chain between variable A to variable C, with variable B acting as a bridge. In this scenario, B is often viewed as the mechanism or mediator

[1]Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA, USA. [2]Graduate Aerospace Laboratories, California Institute of Technology, Pasadena, CA, USA. ✉e-mail: alvaroms@mit.edu

responsible for transmitting the influence of A to C. Mediator variables help explain the underlying mechanisms by which an independent variable influences a dependent variable. A simple example is ↑ education level → ↑ job skills → ↑ income.

- Confounding variables ($A ← B → C$) act as the common cause for two variables: $B → A$ and $B → C$. They have the potential to create a statistical correlation between A and C, even if there is no direct causal link between them. Consequently, confounding variables can obscure or distort the genuine relationship between variables. Following the example above, air pollution → ↑ deforestation[19], and air pollution → ↑ respiratory health conditions[20].

- Collider variables ($A → B ← C$) represent the effect of multiple factors acting on the same variable: $A → B$ and $C → B$. This scenario is particularly relevant in nonlinear dynamical systems, where most variables are affected by multiple causes due to their coupling. A collider exhibits redundant causes when both A and C contribute to the same effect or outcome of B, creating overlapping or duplicative influences on the outcome. Consequently, redundant causes result in multiple pathways to the same effect. For instance, both hard work and high intelligence can independently contribute to the good grades of a student. Note that A and C may not necessarily be independent. A collider is synergistic if the combined effect of A and C on B surpasses their individual effects on B when considered separately. For example, two drugs, A and C, may be required in tandem to effectively treat a condition B, when each drug alone is ineffective.

The search for mathematical definitions of causality that accurately identify mediator, confounder, and collider effects remains an active area of research[21]. One of the most intuitive formulations of causality relies on the concept of interventions[1,10]. The approach offers a pathway for evaluating the causal effect that the process A exerts on another process B by setting A at a modified value $\tilde{A}$ and observing the consequences of the post-intervention in B. Despite its intuitiveness, interventional studies are not without limitations[3,22]. Causality with interventions is intrusive (i.e., it requires modification of the system) and costly (the experiments or simulations need to be repeated). When data are gathered from physical experiments, establishing causality through interventions may become highly challenging or impractical (e.g., we cannot use interventions to assess the causality in the stock market in 2008). Additionally, the notion of causality with interventions prompts questions about the type of intervention that should be introduced and whether this intervention could affect the outcome of the exercise as a consequence of forcing the system out of its natural state. Interventional studies can also pose ethical problems in fields such as neuroscience or climate science[23]. For example, they might involve manipulating neural functions in living organisms or altering natural environmental conditions, potentially leading to irreversible changes or damage.

The alternative approach to interventions involves discovering causal links through observations. Observational methods are predominantly data-driven and do not require alterations to the original system. In the recent years, the steady advancements in computational power coupled with the exponential growth of big data have significantly contributed to the widespread adoption of observational techniques. One of the pioneering approaches is rooted in the use of forecasting models. The concept was initially proposed by Wiener[24] and later quantified by Granger[25]. Granger causality (GC) measures the causality from the process B to A by evaluating how the inclusion of B in an autoregressive model reduces the forecast error for A. Originally developed for linear bivariate relationships, GC has since expanded to encompass nonlinear and multivariate scenarios[26–29], finding applications in diverse fields ranging from econometrics[30–32] to fluid dynamics[33], and biology[34,35].

Model-free approaches for causal discovery have also been proposed to overcome the limitations of GC. A leading method in this domain is convergent cross-mapping (CCM)[36] and its variants[37–41], which apply Takens' embedding theorem[42] to establish connections between variables and the attractor of the system. An alternative approach, known as continuity scaling[43], directly assesses causal relationships by examining the scaling laws governing the continuity of the system.

Information theory, the science of message communication[44], has also served as a framework for model-free causality quantification. The success of information theory relies on the notion of information as a fundamental property of physical systems, closely tied to the restrictions and possibilities of the laws of physics[45,46]. The grounds for causality as information are rooted in the intimate connection between information and the arrow of time. Time-asymmetries present in the system at a macroscopic level can be leveraged to measure the causality of events using information-theoretic metrics based on the Shannon entropy[44]. The initial applications of information theory for causality were formally established through the use of conditional entropies, employing what is known as directed information[47,48]. Among the most recognized contributions is transfer entropy (TE)[49], which measures the reduction in entropy about the future state of a variable by knowing the past states of another. Various improvements have been proposed to address the inherent limitations of TE. Among them, we can cite conditional transfer entropy (CTE)[50–53], which stands as the nonlinear, nonparametric extension of conditional GC[27]. Subsequent advancements of the method include multivariate formulations of CTE[45] and momentary information transfer[54], which extends TE by examining the transfer of information at each time step. Other information-theoretic methods, derived from dynamical system theory[55–58], quantify causality as the amount of information that flows from one process to another as dictated by the governing equations.

Another family of methods for causal inference relies on conducting conditional independence tests. This approach was popularized by the Peter-Clark algorithm (PC)[59], with subsequent extensions incorporating tests for momentary conditional independence (PCMCI)[23,60]. PCMCI aims to optimally identify a reduced conditioning set that includes the parents of the target variable[61]. This method has been shown to be effective in accurately detecting causal relationships while controlling for false positives[23]. Recently, new PCMCI variants have been developed for identifying contemporaneous links[62], latent confounders[63], and regime-dependent relationships[64].

The methods for causal inference discussed above have significantly advanced our understanding of cause-effect interactions in complex systems. Despite the progress, current approaches face limitations in the presence of nonlinear dependencies, stochastic interactions (i.e., noise), self-causation, mediator, confounder, and collider effects, to name a few. Moreover, they are not capable of classifying causal interactions as redundant, unique, and synergistic, which is crucial to identify the fundamental relationships within the system. Another gap in existing methodologies is their inability to quantify causality that remains unaccounted for due to unobserved variables. To address these shortcomings, we propose SURD: Synergistic-Unique-Redundant Decomposition of causality. SURD offers causal quantification in terms of redundant, unique, and synergistic contributions and provides a measure of the causality from hidden variables. The approach can be used to detect causal relationships in systems with multiple variables, dependencies at different time lags, and instantaneous links. We demonstrate the performance of SURD across a large collection of scenarios that have proven challenging for causal inference and compare the results to previous approaches.

Results

Theoretical background

Consider the collection of N time-dependent variables given by the vector $\boldsymbol{Q} = [Q_1(t), Q_2(t), ..., Q_N(t)]$. For example, Q_i may represent the

regional average of climatological variables (e.g., temperature, pressure,...) or the evolution of human heart rate. The components of $\boldsymbol{Q}$ are the observables and are treated as random variables. Our objective is to quantify the causality from the components of $\boldsymbol{Q}$ to the future of the target variable Q_j, denoted by $Q_j^+ = Q_j(t + \Delta T)$, where $\Delta T > 0$ is an arbitrary time increment. The vector $\boldsymbol{Q}$ can include variables at times less or equal to $t + \Delta T$, which allows us to identify both lagged and instantaneous dependencies.

SURD quantifies causality as the increase in information (ΔI) about Q_j^+ obtained from observing individual components or groups of components from $\boldsymbol{Q}$. The information in Q_j^+ is measured by the Shannon entropy[44], denoted by $H(Q_j^+)$, which represents the average number of bits required to unambiguously determine Q_j^+. It is also useful to interpret Shannon entropy as a measure of uncertainty. Processes that are highly uncertain (high entropy) are also the ones from which we gain the most information when their states are determined. Conversely, uncertainty is zero when the process is completely deterministic, indicating no information is gained when the outcome is revealed. Using the principle of forward-in-time propagation of information (i.e., information only flows toward the future)[45], $H(Q_j^+)$ can be decomposed as the sum of all causal contributions from the past and present:

$$H\left(Q_j^+\right) = \sum_{i \in \mathcal{C}} \Delta I_{\boldsymbol{i} \to j}^R + \sum_{i=1}^N \Delta I_{i \to j}^U + \sum_{\boldsymbol{i} \in \mathcal{C}} \Delta I_{\boldsymbol{i} \to j}^S + \Delta I_{\text{leak} \to j}, \quad (1)$$

where $\Delta I_{\boldsymbol{i} \to j}^R, \Delta I_{i \to j}^U$, and $\Delta I_{\boldsymbol{i} \to j}^S$ are the redundant, unique, and synergistic causalities, respectively, from the observed variables to Q_j^+, and $\Delta I_{\text{leak} \to j}$ is the causality from unobserved variables, referred to as the causality leak. Unique causalities are associated with individual components of $\boldsymbol{Q}$, whereas redundant and synergistic causalities arise from groups of variables from $\boldsymbol{Q}$. Consequently, the set $\mathcal{C}$ contains all combinations involving more than one variable. For instance, for $N = 2$, Equation (1) reduces to $H(Q_j^+) = \Delta I_{12 \to j}^R + \Delta I_{1 \to j}^U + \Delta I_{2 \to j}^U + \Delta I_{12 \to j}^S + \Delta I_{\text{leak} \to j}$. Figure 1 shows the diagram of the redundant, unique, and synergistic causalities for $N = 2$. The formal definitions of causality can be found in the Supplementary Materials. Here, we offer an interpretation of each term:

- Redundant causality from $\boldsymbol{Q_i} = [Q_{i_1}, Q_{i_2}, \ldots]$ to Q_j^+ (denoted by $\Delta I_{\boldsymbol{i} \to j}^R$) is the common causality shared among all the components of $\boldsymbol{Q_i}$, where $\boldsymbol{Q_i}$ is a subset of $\boldsymbol{Q}$. Redundant causality occurs when all the variables in $\boldsymbol{Q_i}$ contain the same amount of information

about Q_j^+. Therefore, any component of $\boldsymbol{Q_i}$ offers identical insight into the outcome of Q_j^+.
- Unique causality from Q_i to Q_j^+ (denoted by $\Delta I_{i \to j}^U$) is the causality from Q_i that cannot be obtained from any other individual variable $Q_k \neq Q_i$. This causality occurs when observing Q_i yields more information about some outcomes of Q_j^+ than observing any other isolated variable.
- Synergistic causality from $\boldsymbol{Q_i} = [Q_{i_1}, Q_{i_2}, \ldots]$ to Q_j^+ (denoted by $\Delta I_{\boldsymbol{i} \to j}^S$) is the causality arising from the joint effect of the variables in $\boldsymbol{Q_i}$. This causality occurs when more information about Q_j^+ is gained by observing a collection of variables simultaneously than by observing each variable individually.
- Causality leak represents the effect from unobserved variables that influence Q_j^+ but are not contained in $\boldsymbol{Q}$. This is the amount of information missing that would be required to unambiguously determine the future of Q_j after considering all observable variables collectively.

SURD exhibits several key properties that facilitate the precise identification of interactions by preventing the duplication of causality. This is illustrated in Fig. 1. First, the terms in Equation (1) are nonnegative and such that the sum of redundant, unique, and synergistic causalities equals the information shared between Q_j^+ and $\boldsymbol{Q}$, referred to as the mutual information $I(Q_j^+; \boldsymbol{Q})$[44,65,66]. SURD also satisfies that the mutual information between individual variables Q_i and Q_j^+, denoted as $I(Q_j^+; Q_i)$, is represented by the sum of unique and redundant causalities involving Q_i. This condition is consistent with the notion that causality from an individual variable to Q_j^+ is composed solely of unique and redundant causalities, while synergistic causalities emerge from the combined effects of two or more variables[67]. The information-theoretic formulation of SURD is also well-suited for capturing nonlinear dependencies, as well as deterministic and stochastic interactions, and self-causation.

The forward propagation of information from Equation (1) also lays the foundation for normalizing causality within SURD. Unique, redundant, and synergistic causalities to Q_j^+ are normalized by $I(Q_j^+; \boldsymbol{Q})$, such that their sum equals 1. Similarly, the causality leak is normalized by $H(Q_j^+)$, which bounds its values between 0 (indicating that all causalities to Q_j^+ are accounted for by $\boldsymbol{Q}$) and 1 (indicating that none of the causalities are accounted for by $\boldsymbol{Q}$). Figure 1 includes the results of SURD for three simple examples. Each case represents a system characterized
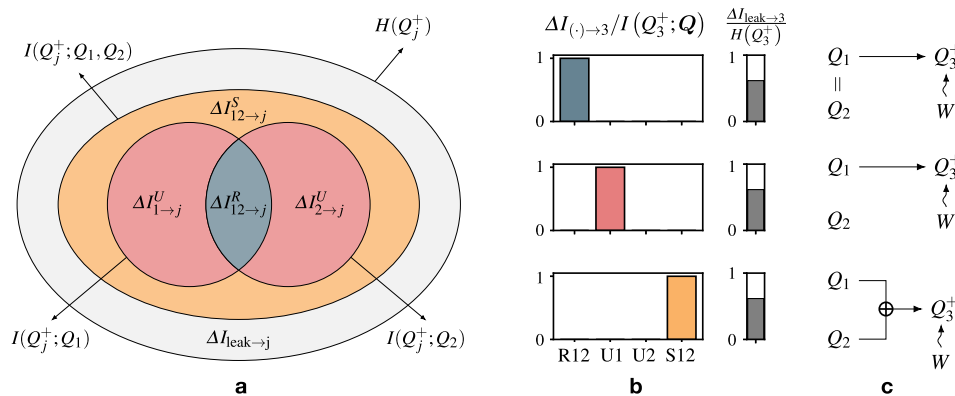

Fig. 1 | SURD: Synergistic-Unique-Redundant Decomposition of causality.
a Diagram of the decomposition of causal dependencies between a vector of observed variables $\boldsymbol{Q} = [Q_1, Q_2]$ (past) and a target variable Q_j^+ (future) into their synergistic (S), unique (U) and redundant (R) components (in yellow, red, and blue, respectively) and contributions to the total, $I(Q_j^+; Q_1, Q_2)$, and individual, $I(Q_j^+; Q_i)$, mutual information. The causality leak is represented in gray. A version of this diagram for three variables is shown in the Supplementary Materials.

b Redundant, unique, and synergistic causalities for the simple examples of (**c**) a duplicated input (top panel), an output equal to the first input (middle panel), and an exclusive-OR output (bottom panel). The notation used is such that $[R12, U1, U2, S12] \equiv [\Delta I_{12 \to 3}^R, \Delta I_{1 \to 3}^U, \Delta I_{2 \to 3}^U, \Delta I_{12 \to 3}^S]$. The target variable Q_3^+ is affected by external stochastic forcing W, which is independent of the observed variables in $\boldsymbol{Q}$. The effect of W is measured by the causality leak, represented by the gray bar.

Table 1 | Summary of the performance of different methods for causal inference

Case	CGC	CTE	CCM	PCMCI	SURD
Mediator variable	✗	✓	✗	✓	✓
Confounder variable	✓	✓	✓	✓[a]	✓
Synergistic collider variable	✗	✓[b]	✗	✓[b]	✓
Redundant collider variable	✗	✗	✗	✗	✓
Turbulent energy cascade	✗	✓[a]	✗	✓[a]	✓
Experimental turbulent boundary layer	✓	✓	✗	✗	✓
Lotka–Volterra prey-predator model[68,69]	✓	✓	✓	✗	✓
Three-interacting species system[40]	✗	✓[b]	✗	✗	✓
Moran effect model[70]	✓	✓	✓	✓	✓
One-way coupling nonlinear logistic difference system[36]	✗	✓	✓	✗	✓
Two-way coupling nonlinear logistic difference system[36]	✗	✓	✓	✗	✓
Stochastic system with linear time-lagged dependencies[77]	✓	✓	✗	✓[a]	✓
Stochastic system with non-linear time-lagged dependencies[35]	✗	✓	✗	✓	✓
Synchronization of two variables in logistic maps[78]	✗	✗	✗	✓[c]	✓
Synchronization of three variables in logistic maps[78]	✗	✗	✗	✗	✓
Uncoupled Rössler–Lorenz system[79,80]	✗	✓	✗	✓	✓
One-way coupled Rössler–Lorenz system[79,80]	✗	✓	✓	✓[a]	✓

The markers ✓ and ✗ denote consistent and inconsistent identification of causal links, respectively, according to the functional dependency of variables within the system. The methods considered are conditional Granger causality (CGC), conditional transfer entropy (CTE), convergent cross-mapping (CCM), Peter and Clark momentary conditional independence (PCMCI) based on conditional mutual information tests with a k-nearest neighbor estimator, and synergistic-unique-redundant decomposition of causality (SURD). A summary of the results for PCMCI with different independence tests is provided in the Supplementary Materials.
[a]The causality detected is consistent with the interactions in the system, although the causal strength of the links is weak.
[b]Causalities are detected but the method cannot discern whether they are synergistic or unique.
[c]The method cannot detect duplicated variables and redundant causalities.

Table 2 | Method for causal inference

Method	Multivariate	Nonlinear	Stochastic	Contemporaneous	Leak	Time-delay	Self-causation
CGC[26]	✓	✗	✓	✗	✗	✓	✓
CTE[53]	✓	✓	✓	✗	✗	✓	✓
CCM[36]	✗	✓	✗[a]	✓	✗	✗[b]	✗
PCMCI[23]	✓	✓	✓	✗[c]	✗	✓	✓
SURD	✓	✓	✓	✓	✓	✓	✓

List of methods for causal inference investigated, along with the methodological challenges each method is capable of addressing: multivariate relationships, nonlinear dependencies, stochastic (nondeterministic) processes, contemporaneous links, estimation of the causality leak, time-delayed dependencies, and self-causation.
[a]CCM aims to reconstruct the attractor manifold associated with two given variables, making it potentially effective for stochastic systems. However, the presence of increased dynamical noise can complicate the reconstruction process.
[b]An extension of the CCM method, extended CCM[39], introduces the concept of time-delayed causal interactions.
[c]A recent variant of the PCMCI method, PCMCI+[62], accounts for contemporaneous links.

exclusively by redundant, unique, or synergistic causality, respectively. These examples also allow us to introduce the notation used in the following figures: $[R12, U1, U2, S12] \equiv [\Delta I^R_{12\to3}, \Delta I^U_{1\to3}, \Delta I^U_{2\to3}, \Delta I^S_{12\to3}]$, where the index of the target variable Q_3 is omitted but it will be unambiguous from the context.

Validation

We validate SURD across multiple scenarios that pose significant challenges in causal inference. These include systems with mediator, confounder and colliders effects, Lotka–Volterra prey-predator model[68,69], three-interacting species system[40], the Moran effect model[70], turbulent energy cascade[71–73], experimental data for a turbulent boundary layer[74–76], deterministic and stochastic systems with time-lagged dependencies proposed by Sugihara et al.[36], Ding et al.[77], and Bueso et al.[35], logic gates, synchronization of logistic maps[78], and the coupled Rössler–Lorenz system[79–82]. A summary of the results is shown in Table 1, where the metric for success is based on whether the results are consistent with the functional dependencies of the system, rather than on the concrete value of the causal strength provided by each method.

The ability of SURD to identify causal relationships is compared with other methods for causal inference, which are also included in Table 1. The approaches considered are conditional Granger causality (CGC)[26], conditional transfer entropy (CTE)[53], convergent cross-mapping (CCM)[36,83], and Peter and Clark momentary conditional independence (PCMCI)[23]. Each method brings distinct capabilities to address specific challenges in the field of causal inference. To facilitate comparison, we have summarized the key properties associated with each approach in Table 2. This classification outlines the ability of each method to handle multivariate relationships, nonlinear dependencies, stochastic (nondeterministic) processes, contemporaneous links (i.e., those occurring at a smaller time-scale than the time resolution of the data), estimation of the causality leak (i.e., causality from unobserved variables), time-delayed dependencies, and self-causation. The table also highlights the ability of SURD to account for all the scenarios described above.

The findings in Table 1 show that, although there are methods capable of effectively tackling certain situations, SURD is consistently successful across all the cases considered. In particular, SURD offers a distinct advantage in the presence of redundant variables and
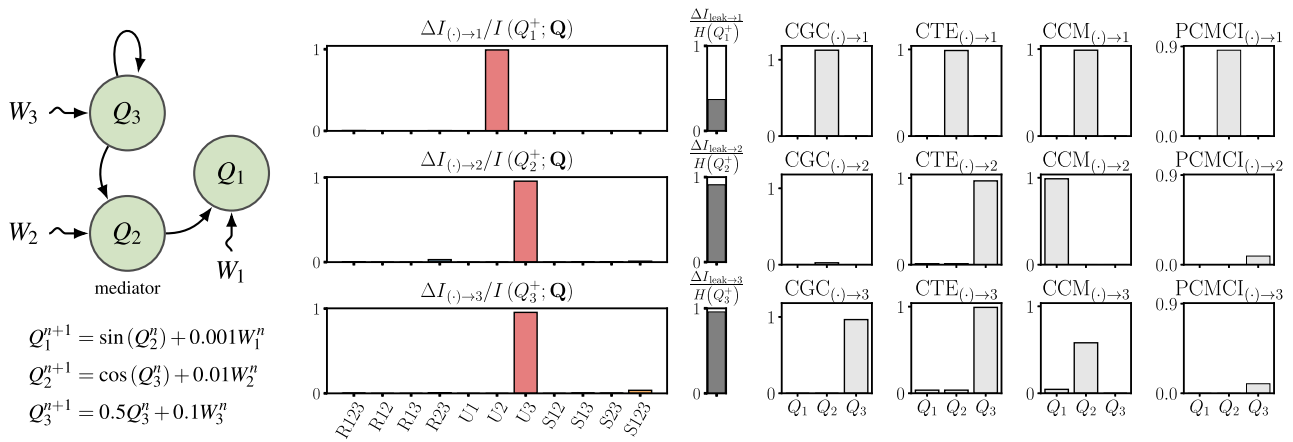
Fig. 2 | System with mediator variable. (Left panel) Schematic of the functional dependence among variables and system equations, where W_i represents unobserved, stochastic forcing on the variable Q_i. We use the notation $Q_i^n = Q_i(t_n)$, where n indicates the time step. (Center and right panels) Results from SURD with redundant (R), unique (U) and synergistic (S) causalities in blue, red and yellow, respectively. The notation employed is such that R123 denotes $\Delta I_{123 \to j}^R$ and so on. The gray bar is the causality leak. The results from CGC, CTE, PCMCI, and CCM are depicted on the right. In all methods but CCM, the value of the bar represent the strength of the causal link. In CCM, a causal link is detected only when the value converges to 1 as the length of the time series increases, but not otherwise. CGC and CTE use the same normalization as SURD. The values for SURD, CTE and CCM are upper bounded by 1. The values for PCMCI represent conditional mutual information and are unbounded. The equations to quantify causality by each of the methods are in the Supplementary Materials.

synchronization phenomena. In the following, we focus our discussion on the cases from Table 1 involving fundamental causal interactions: mediators, confounders, and colliders, as these are key to understand the success of SURD. We also discuss the application of SURD to unveil the causality in two turbulent flow scenarios. Readers are referred to the Supplementary Materials for a comprehensive discussion of the results presented in Table 1, as well as a detailed overview of the causal inference methods utilized and their implementation. Furthermore, the Supplementary Materials also contain a section on the application of SURD to select the most effective input variables for temporal forecasting, as well as a section that discusses the non-separability problem for nonlinear dynamical systems and demonstrates the robustness of SURD under such conditions.

Mediator variable. The first case investigated corresponds to the system $Q_3 \to Q_2 \to Q_1$, where Q_3 influences Q_1 through the mediator variable Q_2. Figure 2 displays a diagram illustrating the relationships among the variables, along with the results derived from SURD and other causal analysis methods. Note that while the exact value of the causal strength may be subject to debate, the predicted causal links by each method should maintain consistency with the arrows depicted in the diagram.

The causal contributions detected by SURD are the unique causalities $\Delta I_{3 \to 3}^U$, $\Delta I_{3 \to 2}^U$, and $\Delta I_{2 \to 1}^U$, which are consistent with the dependency of variables in the system. PCMCI and CTE yield similar results; however, CGC cannot identify the link $Q_3 \to Q_2$. CCM also fails to unambiguously capture the causal links, as the only causality converging to one as the length of the time series increases is $Q_1 \to Q_2$, which is inconsistent with the equations of the system.

SURD also offers an estimate of the causality leak, which exceeds 95% for Q_2 and Q_3. This is attributed to the influence of the stochastic forcing terms W_2 and W_3, respectively, which are assumed to be unknown. The causality leak for Q_1 is the lowest (below 50%), as the noise from W_1 is the smallest among the three W_i. Note that none of the methods offers any insight into the causality leak, and this is also the case for all the subsequent benchmarks.

Confounder variable. The second case (Fig. 3) corresponds to a system where Q_3 acts as a confounding variable for Q_1 and Q_2, i.e., $Q_1 \leftarrow Q_3 \to Q_2$. The presence of confounding effects is captured in SURD

by the synergistic causalities $\Delta I_{13 \to 1}^S$ and $\Delta I_{23 \to 2}^S$, while the self-induced causality from Q_3 is detected by $\Delta I_{3 \to 3}^U$. Other causalities manifest as $\Delta I_{1 \to 1}^U$ and $\Delta I_{3 \to 2}^U$, which are also consistent with the causal structure of the system. Regarding other methods, all of them correctly identified the confounding effects, although the link $Q_3 \to Q_1$ and the self-induced causalities identified by PCMCI are barely significant. This highlights another advantage of SURD: the relative importance of the causalities is easier to interpret, since the sum of their normalized values must always equal one. Also note that CCM cannot detect self-induced causalities by construction. The largest causality leak occurs for Q_3, which is consistent with the fact that the (unobserved) stochastic term acting on Q_3 is ten times larger than the (unobserved) stochastic terms acting on Q_1 and Q_2.

Collider with synergistic variables. Next, we consider the system $[Q_2, Q_3] \to Q_1$, where Q_2 and Q_3 act together to influence Q_1. In reality, Q_2 and Q_3 behave as a single random variable that drives Q_1. The results, presented in Fig. 4, demonstrate that SURD is able to detect the dominant synergistic effect of Q_2 and Q_3 on Q_1 through $\Delta I_{23 \to 1}^S$ along with the self-induced causalities $\Delta I_{2 \to 2}^U$ and $\Delta I_{3 \to 3}^U$. The smallest causality leak is associated with Q_1, as it is affected by the lowest stochastic forcing while being strongly influenced by the (observed) variables Q_2 and Q_3.

PCMCI and CTE also identify the self-influence of Q_2 and Q_3 and the effect of Q_2 and Q_3 on Q_1. However, neither of the two methods can label the interaction as synergistic and, hence, cannot show that both variables are required in combination, rather than individually, to affect Q_1. On the other hand, CGC and CCM are unsuccessful in identifying the interactions of Q_2 and Q_3 with Q_1.

Collider with redundant variables. The fourth case under examination explores the fundamental interaction $Q_2 \equiv Q_3 \to Q_1$, where Q_3 is identical to Q_2. In this scenario, Q_2 and Q_3 equally influence the future outcomes of Q_1. SURD identifies $\Delta I_{23 \to 2}^R = \Delta I_{23 \to 3}^R$ as the most significant causalities associated with Q_2 and Q_3, respectively, as shown in Fig. 5. Moreover, the identical causalities (and causality leaks) for both Q_2 and Q_3 suggest that they represent the same variable. SURD also identifies the influence of Q_2 and Q_3 on Q_1 mostly via $\Delta I_{23 \to 1}^R$ and $\Delta I_{12 \to 1}^S$. Given that Q_2 and Q_3 are identical, SURD assigns a nonzero value only to $\Delta I_{12 \to 1}^S$, but not to $\Delta I_{13 \to 1}^S$, to prevent the duplication of causality in compliance with Equation (1).
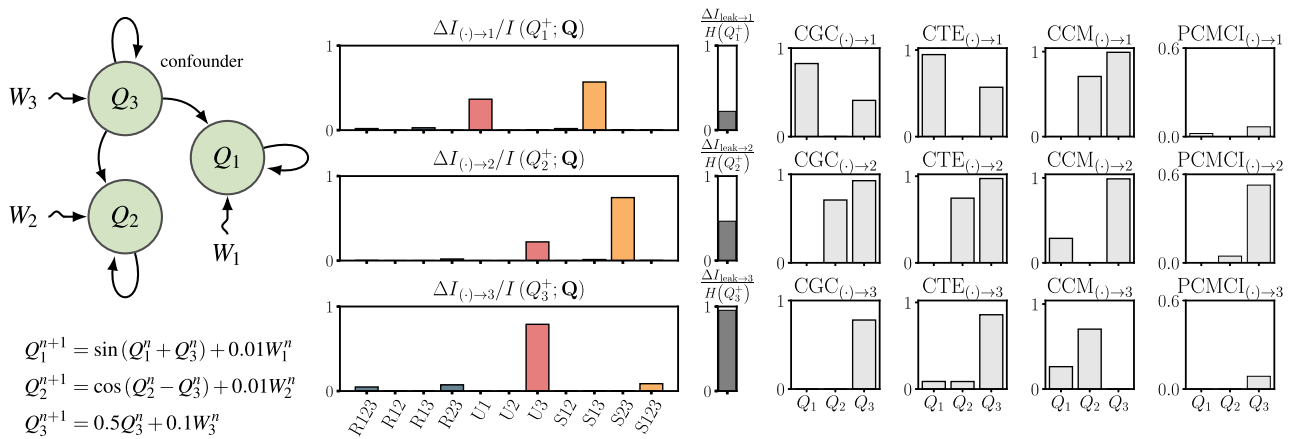
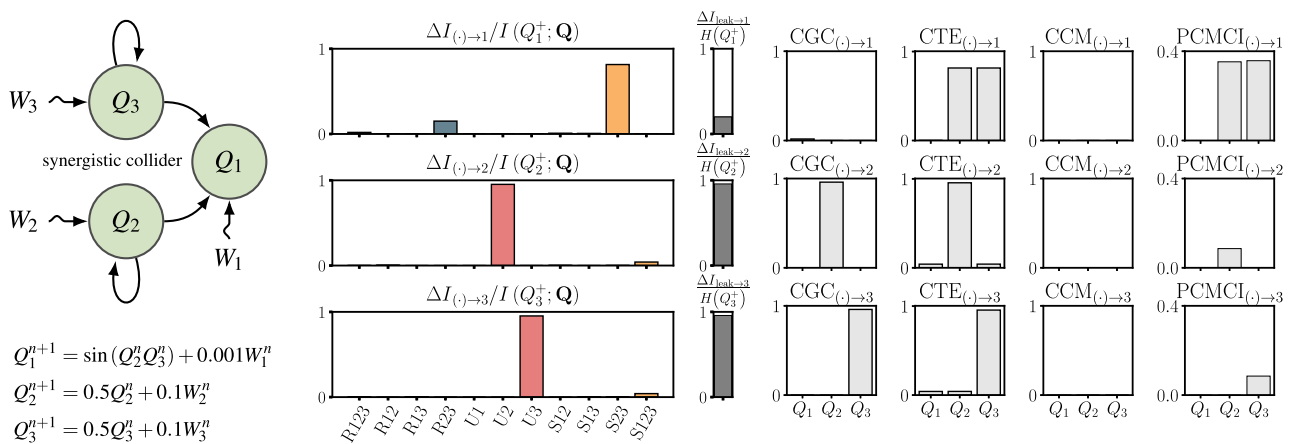**Fig. 3 | System with confounder variable.** Same as Fig. 2.

$$Q_1^{n+1} = \sin(Q_1^n + Q_3^n) + 0.01W_1^n$$
$$Q_2^{n+1} = \cos(Q_2^n - Q_3^n) + 0.01W_2^n$$
$$Q_3^{n+1} = 0.5Q_3^n + 0.1W_3^n$$

Fig. 4 | System with synergistic collider variables. Same as Fig. 2.

$$Q_1^{n+1} = \sin(Q_2^n Q_3^n) + 0.001W_1^n$$
$$Q_2^{n+1} = 0.5Q_2^n + 0.1W_2^n$$
$$Q_3^{n+1} = 0.5Q_3^n + 0.1W_3^n$$
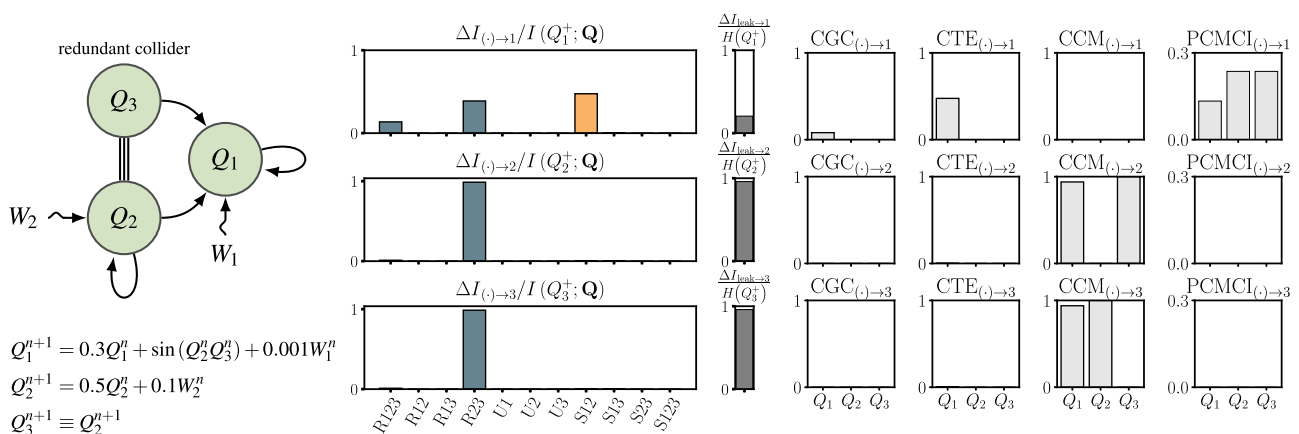
Fig. 5 | System with redundant collider variables. Same as Fig. 2. The symbol ≡ indicates that variables Q_2 and Q_3 are identical.

$$Q_1^{n+1} = 0.3Q_1^n + \sin(Q_2^n Q_3^n) + 0.001W_1^n$$
$$Q_2^{n+1} = 0.5Q_2^n + 0.1W_2^n$$
$$Q_3^{n+1} \equiv Q_2^{n+1}$$

For the rest of the methods, CGC and CTE are unable to identify any causal links between different variables. CCM identifies a bidirectional causal connection between Q_2 and Q_3, but cannot identify their influence to Q_1. Finally, PCMCI exhibits consistent results for Q_1; however, it does not offer a mechanism to identify the redundancy between Q_2 and Q_3.

Application to the energy cascade in turbulence

We apply SURD to investigate the causality of the energy cascade in turbulence, which serves as a primary example of a chaotic, multi-scale, high-dimensional system. The energy cascade is the transfer of kinetic energy from large to small scales in the flow (forward cascade), or vice versa (backward cascade), and has been the cornerstone of

most theories and models of turbulence since the 1940s[71–73,84–87]. However, understanding the dynamics of the energy transfer across scales remains an outstanding challenge. Given the ubiquity of turbulence, a deeper understanding of the energy transfer among flow scales could enable significant progress across various fields, ranging from plasma physics[88], combustion[89], climate[90], and astrophysics[91] to engineering applications in aero/hydrodynamics[92–95]. Despite the progress made in recent decades, the causal interactions of energy among scales in turbulent flows have received less attention. Here, we investigate the redundant, unique, and synergistic causalities involved in the energy transfer.

We use data from a high-fidelity simulation of isotropic turbulence in a triply periodic domain[87]. This case is the testbed used by the community to understand the fundamental physical processes of the energy cascade. The total number of degrees of freedom of the system is of the order of 10^9. The kinetic energy transfer was obtained by filtering the velocity field at four different length scales (denoted by Δ_i, for $i = 1, 2, 3, 4$) and calculating the energy flux across those scales. Causality is computed among the time signals of the volume-averaged energy transfer, denoted by Σ_i for $i = 1, 2, 3, 4$, where the index signifies energy transfer at scale Δ_i. The top panels of Fig. 6 depicts a visualization of the filtered velocity from the largest to the smallest flow scale, together with the time evolution of the energy transfer signals at that scale. The causal relationships among energy transfers identified by SURD are shown in the left panels of Fig. 6. The dominant contributions come from redundant and unique causalities, whereas synergistic causalities play a minor role. The unique causalities (depicted in red) vividly capture the forward energy cascade of causality from large to smaller scales, inferred from the non-zero terms $\Delta I^U_{1 \to 2}$, $\Delta I^U_{2 \to 3}$, and $\Delta I^U_{3 \to 4}$. Curiously, no unique causality is observed from smaller to larger scales, and any causality from the backward cascade arises solely through redundant relationships. In the context of SURD, this implies that no new information is conveyed from the smaller scales to the larger scales, which is consistent with recent views of the backward energy cascade in the literature[96,97]. From the modeling perspective, this justifies the success of subgrid-scale modeling in large-eddy simulation, as the information contained in the smaller scales is redundant and does not constitute a key ingredient in solving the closure model problem. The results obtained from SURD also provide support for classic hypotheses about the energy cascade from a new causal-effect perspective. Among them, we can cite Taylor's dissipation surrogate assumption[98] and the dissipation anomaly[99]. The former posits that the dissipation rate can be determined by large-scale dynamics, even if dissipation is formally a small-scale feature of the flow. SURD clearly supports this assumption due to the lack of unique and synergistic causality from small to large scales. The results from SURD are also consistent with the dissipation anomaly (i.e., the constant rate of energy dissipation despite decreasing viscosity), which is enabled by the forward directionality of the energy cascade process.

SURD also provides the causality leak ($\Delta I_{\text{leak} \to j}$) that measures the amount of causality unaccounted for due to unobserved variables. The largest causality leak occurs for Σ_2, where ~47% of the causality is carried by variables not included within [Σ_1, Σ_2, Σ_3, Σ_4]. This implies that there are other factors affecting Σ_2 that have not been accounted for and that explain the remaining 53% of the causality of the variable. Conversely, the energy transfer at the largest scale Σ_1 bears the smallest leak of 14%, which is due to the high value of the unique causality $\Delta I^U_{1 \to 1}$. The latter implies that the future of the largest scales is mostly determined by its own past.

Finally, the results from SURD are compared with the other methods. CGC and CCM do not support the hypothesis of a forward energy cascade, which disagrees with the consensus within the fluid dynamics community[85,87,100–104]. The formulation of CCM used in this study adheres to the original work by Sugihara et al.[36] However, more recent iterations of CCM, such as Extended CCM[39], which explicitly account for time delays, have demonstrated efficacy in accurately detecting causality in systems with strongly synchronized variables. Hence, these and other improved versions of CCM might be more suitable for analyzing the turbulent energy cascade, where smaller scales are enslaved to the larger ones. CTE and PCMCI are consistent with the forward propagation of energy, but the strength of the causal links detected is extremely weak. Beyond the failure of some methods to support the forward energy cascade hypothesis, different approaches also yield conflicting outcomes regarding the path followed by the energy across scales and the significance of the backward energy cascade. Additionally, none of the other previous methods offer quantification of missing causality due to unobserved variables, in contrast to the causality leak provided by SURD.

Application to experimental data from a turbulent boundary layer

The interaction of turbulent motions of different size within the thin fluid layers immediately adjacent to solid boundaries poses a significant challenge for both physical understanding and prediction. These layers are responsible for nearly 50% of the aerodynamic drag on modern airliners and play a crucial role in the first hundred meters of the atmosphere, influencing broader meteorological phenomena[94]. Here, we leverage SURD to investigate the interaction between flow velocity motions in the outer layer (far from the wall) and inner layer (close to the wall) of a turbulent boundary layer. Figure 7a illustrates the configuration used to examine the causal interactions between velocity motions. More specifically, the hypotheses under consideration are either (i) a dominant influence of motions far from the wall on those closer, indicating top-down causality (a.k.a. Townsend's outer-layer similarity hypothesis[105]), or (ii) the opposite scenario, where influences move from areas closer to the wall outward, suggesting bottom-up causality.

We use experimental data from a zero-pressure gradient turbulent boundary layer from the high Reynolds number wind tunnel at the University of Melbourne[74–76]. The friction Reynolds number is $Re_\tau = u_\tau \delta / \nu = 14{,}750$, based on the thickness of the boundary layer δ, the kinematic viscosity ν, and the average friction velocity at the wall u_τ. The time signals consists of the streamwise velocity at two wall-normal locations within the inner (I) and outer (O) layers, denoted by $u_I(t)$ and $u_O(t)$, respectively.

Figure 7b shows the redundant, unique, and synergistic causalities from SURD between the inner and outer layers. We use the subindices I and O to refer to causalities from/to $u_I(t)$ or $u_O(t)$, respectively. The primary observation is that the inner layer motions are predominantly influenced by the unique causality from the outer layer, $\Delta I^U_{O \to I}$. The redundant and synergistic causalities are lower, but they remain significant. Curiously, the unique causality $\Delta I^U_{I \to I}$ is zero, implying that, at the time scale considered, the inner layer motions are independent of their past history. For the outer-layer motions, most of the causality is self-induced $\Delta I^U_{O \to O}$ with no apparent influence from the inner layer. The results distinctly support the prevalence of top-down interactions: causality flows predominantly from the outer-layer large-scale motions to the inner-layer small-scale motions. The outcome is consistent with the modulation of near-wall scales by large-scale motions reported in previous investigations[106,107]. The lack of bottom-up causality from the inner to the outer layer also aligns with Townsend's outer-layer similarity hypothesis[105] and previous observations in the literature[108–113].

The causality leak, also shown in Fig. 7b, is 99% for both u_I and u_O. Such a high value implies that most of the causality determining the future of u_I and u_O is contained in other variables not considered in the analysis. This high value is unsurprising since most of the millions of degrees of freedom in the turbulent flow field have been neglected, and only two pointwise signals, u_I and u_O, are retained to evaluate the causality.
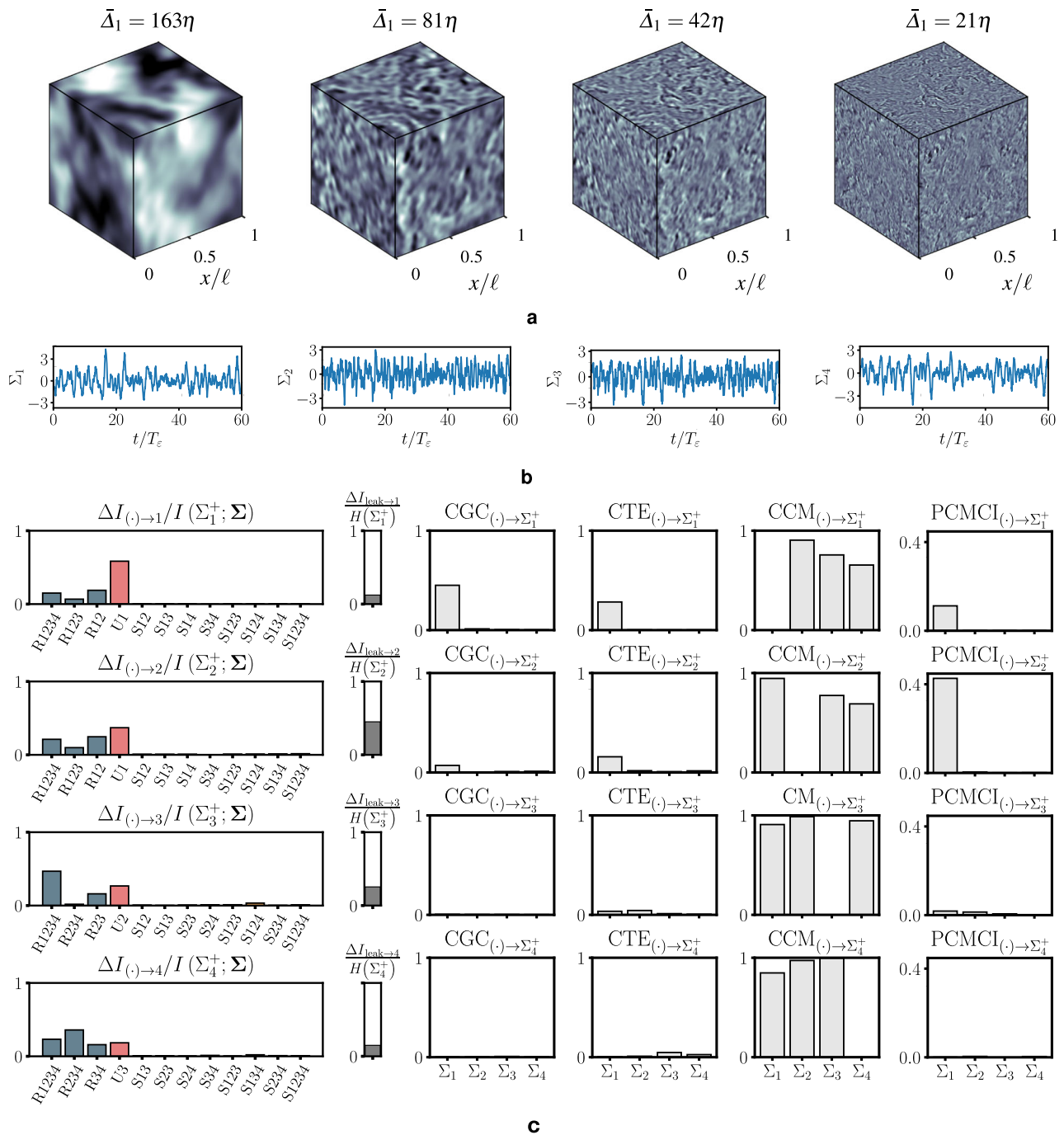
Fig. 6 | Causality in the turbulent energy cascade. a Visualization of the magnitude of the velocity field for four filter sizes Δ_i at the same instant. The parameter η is the Kolmogorov length-scale and represents the smallest scale in the flow. **b** An extract of the time history of $\Sigma_1, \Sigma_2, \Sigma_3,$ and Σ_4. The time is non-dimensionalized by T_ε, which is the integral time-scale of the flow. **c** Redundant (R), unique (U), and synergistic (S) causal contributions from SURD. Only the top 12 contributions, satisfying the condition $\Delta I_{(\cdot)\rightarrow j}/I(\Sigma_j^+; \boldsymbol{\Sigma}) \geq 10^{-3}$, are represented, where $\boldsymbol{\Sigma} = [\Sigma_1, \Sigma_2, \Sigma_3, \Sigma_4]$. The gray bar is the causality leak. The results of CGC, CTE, CCM, and PCMCI are depicted on the right.

Finally, the results from SURD are contrasted with other methods. In this case, CCM and PCMCI do not support the hypothesis of top-down interactions between velocity motions. The reason behind the failure of these methods is unclear, but it might be related to the high causality leak. CGC and CTE are consistent with the flow of causality from the outer-layer large-scale motions to the inner-layer small-scale motions. However, as already highlighted in previous cases, none of these methods offer a detailed decomposition into redundant, unique, and synergistic causality, nor they account for the effect of unobserved variables as quantified by the causality leak in SURD.

Discussion
The cases presented in this study show that the faithful quantification of causality remains elusive even in simple causal networks. The difficulties originate from a range of factors, including complexities introduced by mediator, confounding and collider effects; synergistic and duplicated variables; the influence of stochastic noise; and the presence of unobserved variables. SURD addresses these challenges by introducing several unique advances in the field of causal inference that go beyond the insights provided by other methods.
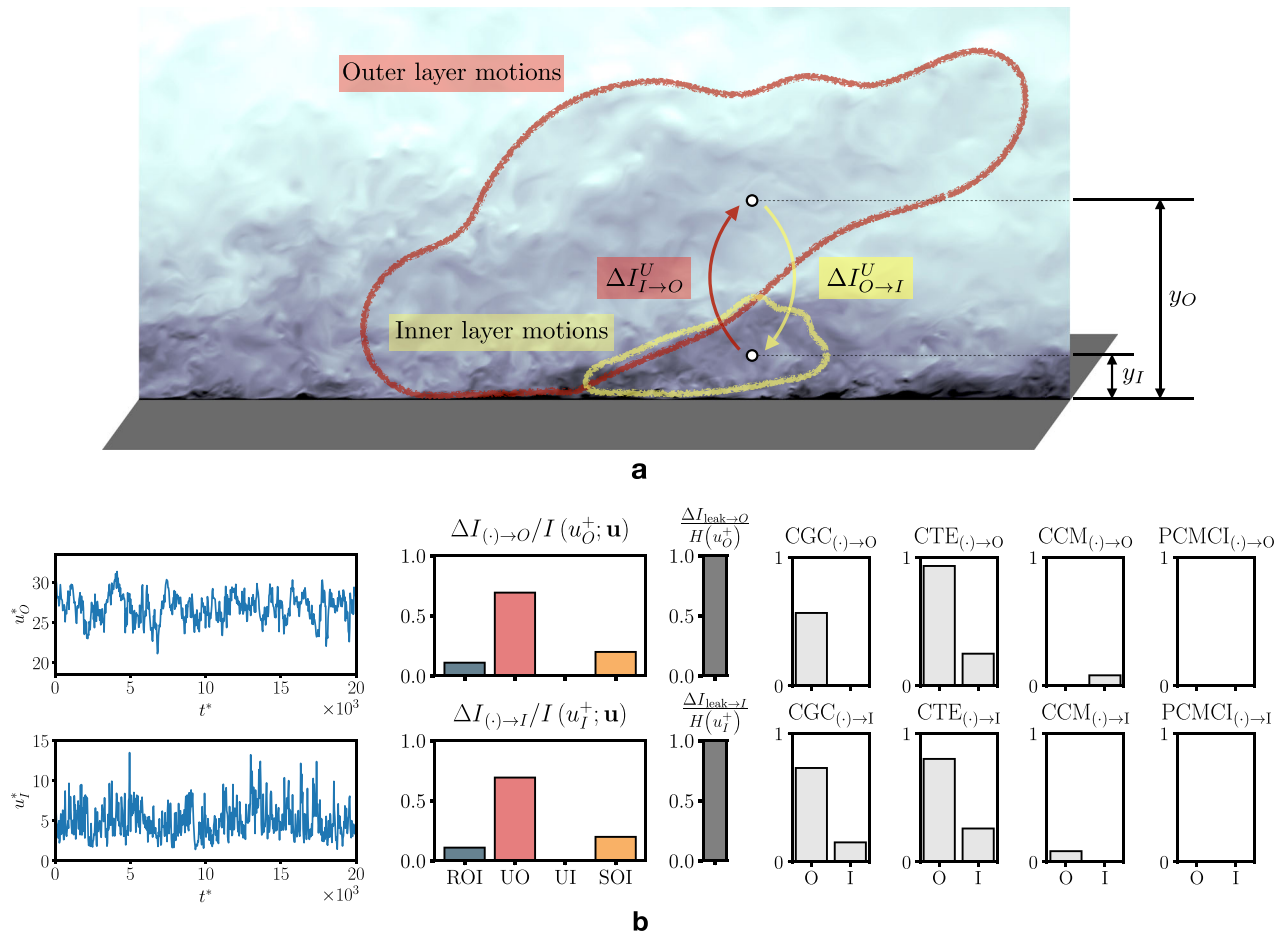
Fig. 7 | Causality between streamwise velocity motions in a turbulent boundary layer. a Schematic of outer-layer and inner-layer streamwise velocity motions in a turbulent boundary and their interactions via unique causality. The velocity signals $u_I(t)$ and $u_O(t)$ are experimentally measured at the wall-normal locations y_I and y_O, respectively, and are shown in the panel below. The superscript * denotes the inner scaling with friction velocity, u_τ, and kinematic viscosity, ν. **b** Redundant (R), unique (U), and synergistic (S) causalities among velocity signals in the inner (I) and outer (O) layer of a turbulent boundary layer. The gray bar is the causality leak. The results of CGC, CTE, CCM, and PCMCI are shown on the right. Details about data are provided in "Methods".

The first distinctive feature of SURD is its suitability for analyzing causal networks involving mediator, confounder, and collider effects—the building blocks of causal interactions between variables. The success of SURD in capturing these fundamental interactions stems from its ability to distinguish between redundant, unique, and synergistic causalities, which is lacking in previous methods. We have shown that the inability to disentangle redundant and synergistic causalities can obscure the relationships among variables within the system, leading to spurious causalities. In the case of PCMCI, incorporating a redundant variable into the set of conditioning variables may lead to the identification of erroneous links[23]. The challenge posed by redundant causality also extends to CTE[27,50–53], as it evaluates the causality between pairs of variables conditioned on the remaining set of observed variables. CGC encounters difficulties in the same situations as CTE, since the former is a linear, parametric version of the latter. Several attempts have been made in the literature to account for synergistic and redundant causality[45,114–119], for example, through the calculation of CTE in its multivariate form[45]. However, these methods may yield negative values of causality, which hinders the interpretability of the results. In contrast, SURD ensures the non-negativity of all the terms. This property significantly enhances the interpretability of SURD terms, allowing for a clear distinction between redundant, unique, and synergistic causality among variables.

SURD also introduces the concept of the causality leak, which quantifies the extent of causality that remains unaccounted for due to unobserved variables. The causality leak serves as a fundamental metric to evaluate the significance of the causal links identified. Low values of the causality leak imply that most of the causality is accounted for by the observed variables. Conversely, high values of the causality leak indicate that most of the causality can be attributed to hidden, exogenous variables. In such situations, SURD highlights the necessity of incorporating additional, currently overlooked variables into the analysis. The capability to detect and quantify missing causality is absent in other methods for causal inference.

The foundational principle of SURD from Equation (1), along with the non-negativity of causalities, also provides a natural normalization for causality that is both intuitive and easily interpretable. Unique, redundant, and synergistic causalities are normalized using the mutual information between all observed variables and the target, ensuring that their sum equals 1. This normalization measures the relative importance of each causality within the group of observed variables. Additionally, the causality leak is naturally normalized by the information content of the target variable, which bounds its value between 0 and 1. For example, in a system with three variables $[Q_1, Q_2, Q_3]$ where we are interested in the causes of Q_2, a normalized unique causality from $Q_1 \rightarrow Q_2$ equal to 90% implies that Q_2 and Q_3 play a minor role (i.e., 10%) in influencing the future of Q_2. However, the causality from $Q_1 \rightarrow Q_2$ would still be deemed insignificant if the causality leak of Q_2 is 99%, indicating that most of the causality to Q_2 resides in other variables not contained in the vector $[Q_1, Q_2, Q_3]$. Other methods, such as

CGC, do not offer bounded values nor a measure of causality leak. Outcomes from PCMCI (when based on correlations) and CCM can be normalized between 0 and 1; however, unlike SURD, the sum of causalities in PCMCI and CCM does not add up to 1 or to any conserved quantity. CTE shares many normalization properties with SURD and allows for the concept of causality leak; however, the possibility for negative values of causality in CTE complicates its interpretation[45].

Another essential aspect of SURD is its foundation in transitional probability distributions, which ensures its invariance to transformations such as shifting, rescaling, and other general invertible transformation of the variables. This is applicable to other methods such as CTE and PCMCI. SURD is also robust across scenarios with varying amounts of samples for causal inference, providing consistent causal links with fewer than a thousand samples. This capability is realized through the application of transport maps specifically designed for estimating high-dimensional conditional probability distributions[120]. Even in situations where the sample size is exceptionally small (e.g., of the order of hundreds) and the number of variables is large, SURD is still capable of identifying causal relationships up to a certain order of synergies [see "Methods"], while higher-order synergistic interactions can be considered part of the causality leak. Finally, methods such as CCM suffer from the presence of increased noise, which complicates the reconstruction of attractor manifolds and reduces the efficacy[121–123]. In contrast, our test cases have shown that SURD is reliable even in the presence of noise. In summary, SURD stands as an effective tool in the field of causal inference with the potential to drive progress across multiple scientific and engineering domains, such as climate research, neuroscience, economics, epidemiology, social sciences, and fluid dynamics, among others.

Methods

Assumptions for causal discovery

SURD is an observational non-intrusive method that operates within a probabilistic framework, where causal relationships emerge as a result of transitional probabilities between states. The method adheres to the principle of forward-in-time propagation of information, which states that causation cannot occur backward in time. This formulation is consistent with the identification of contemporaneous links, as these can be interpreted as causal influences acting on a time scale shorter than the measurement interval (e.g., inferring causality on an 8-h scale from daily measured data). Furthermore, the method incorporates the concept of causality leak, which acts as a mechanism of quality control by assessing the impact of unobserved variables. This measure alleviates the need for the assumption of causal sufficiency (i.e., all common causes of the variables must be accurately measured), as it offers a quantifiable measure of the extent to which information from unobserved variables remains unaccounted for. Additionally, the method is model-free, i.e., no prior knowledge about the system dynamics is required. This makes SURD appealing for applications involving deterministic or stochastic multivariate systems with linear and nonlinear dependencies. The method also assumes that the time signals are stationary, which ensures that their statistical properties do not vary over time. Finally, the method can identify cyclic causal relationships, provided that they adhere to the principle of forward-in-time propagation of information.

Synergistic-Unique-Redundant Decomposition

To perform the decomposition proposed in Equation (1), we rely on the concept of mutual information[44,65,66] between the target variable Q_j^+ and the vector of observed variables $\boldsymbol{Q}$. This quantity can be mathematically described as:

$$I\left(Q_j^+ ; \boldsymbol{Q}\right) = \sum_{q_j^+, \boldsymbol{q}} p\left(q_j^+, \boldsymbol{q}\right) \log_2\left(\frac{p\left(q_j^+, \boldsymbol{q}\right)}{p\left(q_j^+\right) p(\boldsymbol{q})}\right), \quad (2)$$

where q_j^+ and $\boldsymbol{q}$ represent all possible values or states of Q_j^+ and $\boldsymbol{Q}$, respectively. Mutual information measures how different the joint probability distribution $p(q_j^+, \boldsymbol{q})$ is from the hypothetical distribution $p(q_j^+)p(\boldsymbol{q})$, where q_j^+ and $\boldsymbol{q}$ are assumed to be independent. For instance, if Q_j^+ and $\boldsymbol{Q}$ are not independent, then $p(q_j^+, \boldsymbol{q})$ will differ significantly from $p(q_j^+)p(\boldsymbol{q})$. Hence, we assess causality by examining how the probability of Q_j^+ changes when accounting for $\boldsymbol{Q}$.

However, the source of causality might change depending on different states q_j^+ of the target variable Q_j^+. For example, Q_1 can only be causal to positive values of Q_j^+, whereas Q_2 can only be causal to negative values of Q_j^+. Therefore, this decomposition must be performed for all possible values of Q_j^+. To do that, we define the specific mutual information[124] from $\boldsymbol{Q}$ to a particular event $Q_j^+ = q_j^+$ as

$$\tilde{\imath}\left(q_j^+ ; \boldsymbol{Q}\right) = \sum_{\boldsymbol{q}} \frac{p\left(q_j^+, \boldsymbol{q}\right)}{p\left(q_j^+\right)} \log_2\left(\frac{p\left(q_j^+, \boldsymbol{q}\right)}{p\left(q_j^+\right) p(\boldsymbol{q})}\right) \geq 0. \quad (3)$$

Note that the specific mutual information is a function of the random variable $\boldsymbol{Q}$ (which encompasses all its states) but only a function of one particular state of the target variable (namely, q_j^+). Similarly to Equation (2), the specific mutual information quantifies the dissimilarity between $p(q_j^+, \boldsymbol{q})$ and $p(q_j^+)p(\boldsymbol{q})$ but in this case for the particular state $Q_j^+ = q_j^+$. The mutual information between Q_j^+ and $\boldsymbol{Q}$ is recovered by $I(Q_j^+ ; \boldsymbol{Q}) = \sum_{q_j^+} p(q_j^+)\tilde{\imath}(q_j^+ ; \boldsymbol{Q})$. For simplicity, we will use $\tilde{\imath}_i(q_j^+) = \tilde{\imath}(q_j^+ ; \boldsymbol{Q_i})$.

To perform the decomposition of the specific mutual information in its redundant $\Delta\tilde{\imath}_{\boldsymbol{i}}^R$, unique $\Delta\tilde{\imath}_i^U$, and synergistic $\Delta\tilde{\imath}_{\boldsymbol{i}}^S$ components, we quantify the increments in specific information $\Delta\tilde{\imath}$ about q_j^+ obtained by observing an individual or groups of components from $\boldsymbol{Q}$. For a given state q_j^+ of the target variable Q_j^+, the specific causalities $\tilde{\imath}$ are computed for all the possible combinations of past variables. These components are organized in ascending order, which allows to assign the redundant, unique, and synergistic causalities. Figure 8 shows an example of the decomposition of $\tilde{\imath}(q_j^+ ; \boldsymbol{Q})$ for a particular state of the target variable and for the three simple examples illustrated in Fig. 1. The quantities in Equation (1) are then obtained as the expectation of their corresponding values:

$$\Delta I_{\boldsymbol{i}\to j}^R = \sum_{q_j^+} p\left(q_j^+\right)\Delta\tilde{\imath}_{\boldsymbol{i}}^R\left(q_j^+\right), \quad \Delta I_{\boldsymbol{i}\to j}^U = \sum_{q_j^+} p\left(q_j^+\right)\Delta\tilde{\imath}_i^U\left(q_j^+\right),$$

$$\Delta I_{\boldsymbol{i}\to j}^S = \sum_{q_j^+} p\left(q_j^+\right)\Delta\tilde{\imath}_{\boldsymbol{i}}^S\left(q_j^+\right). \quad (4)$$

Comparison with other causality methods

We compared SURD with other established methods for causal inference in time series: CGC[26], CTE[27], CCM[36], and PCMCI[23]. The detailed calculation of the previous methods has been documented in the literature, and their corresponding source codes have been employed[23,77,83,125]. In this paper, an embedding dimension equivalent to the number of variables was used for the application of the CCM method, which was executed with a library size that ensured convergence of the prediction skill for all cases. Therefore, the CCM value should be close to 1 in order to consider a link significantly causal. We used the version of PCMCI based on conditional mutual information (CMI) independence test with the estimator k-nearest neighbor (k-NN)[60]. All cases were evaluated at a significance level of 1%. Furthermore, PCMCIA was estimated with αPC = 0.05 and CMI-kNN parameters kCMI = 0.1, $s$perm = 5, and B = 200 permutation surrogates. The same time lag was used for all the methods. The reader is referred to the Supplementary Materials for a more detailed discussion of the packages used, the validation with test cases provided in each of the
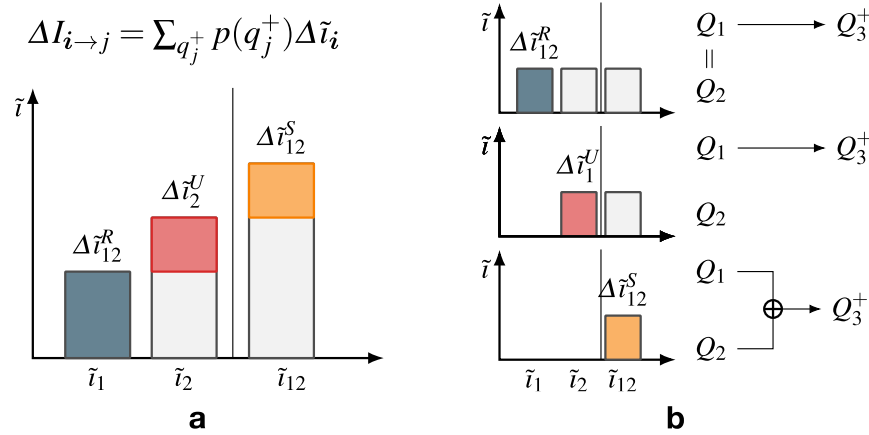
Fig. 8 | Schematic of the causal decomposition in SURD. a For a given state q_j^+ of the target variable Q_j^+, the specific causalities $\tilde{\imath}$ are computed for all the possible combinations of past variables. These components are organized in ascending order, which allows to assign the redundant (blue), unique (red), and synergistic (yellow) causalities. This process is performed for all possible values q_j^+ of Q_j^+. **b** Schematic of simple examples and associated specific mutual information for (top panel) duplicated input, (middle panel) output equal to first input, and (bottom panel) exclusive-OR output.

sources, and the effect of convergence for the CCM method. A summary of the results for PCMCI using different independence tests is also provided in the Supplementary Materials, where the optimal confidence interval α_{PC} during the initial condition selection phase (PC phase) was selected based on the Akaike Information criterion[126] from a default list of values, i.e., $\alpha_{PC} = [0.05, 0.1, 0.2, 0.3, 0.4, 0.5]$.

Validation data for mediator, confounder, and collider
The mediator, confounder, and collider systems considered comprise three variables $Q_1(t_n)$, $Q_2(t_n)$, and $Q_3(t_n)$ at discrete times $t_n = n$. The system is initially set to $Q_1(1) = Q_2(1) = Q_3(1) = 0$. A stochastic forcing, represented by $W_i(t_n)$, acts on $Q_i(t_n)$ and follows a Gaussian distribution with a mean of zero and a standard deviation of one. The computation of SURD is performed for a time lag of $\Delta T = 1$ using 100 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. CGC and CTE used the same samples as SURD, while CCM and PCMCI used 5×10^5 samples due to computational constraints. The latter methods were also evaluated using a smaller number of samples, which were one order of magnitude lower, and no significant differences were detected. An analysis of the impact of the number of samples and the sensitivity to partition refinement for SURD is provided in the Supplementary Materials.

Data for energy cascade in isotropic turbulence
The case chosen to study the energy cascade is forced isotropic turbulence in a triply periodic box. The data were obtained from a direct numerical simulation[127], which is publicly available in https://torroja.dmt.upm.es/turbdata/. In this simulation, the Navier–Stokes equations are numerically integrated by resolving the whole range of spatial and temporal scales of the flow. The conservation of momentum and mass equations for an incompressible flow are given by:

$$\frac{\partial u_i}{\partial t} + \frac{\partial u_i u_j}{\partial x_j} = -\frac{\partial \Pi}{\partial x_i} + \nu \frac{\partial^2 u_i}{\partial x_j \partial x_j} + f_i, \quad \frac{\partial u_i}{\partial x_i} = 0, \tag{5}$$

where repeated indices imply summation, $\boldsymbol{x} = [x_1, x_2, x_3]$ are the spatial coordinates, u_i for $i = 1, 2, 3$ are the velocities components, Π is the pressure, ν is the kinematic viscosity, and f_i is a linear forcing sustaining the turbulent flow[128]. The simulation was conducted with 1024^3 spatial Fourier modes, which is enough to accurately resolve all the relevant length-scales of the flow[127]. To quantify the transfer of kinetic energy

among eddies at different length scales over time, the i-th component of the instantaneous flow velocity in Equation (5), denoted as $u_i(\boldsymbol{x}, t)$, is decomposed into contributions from large and small scales according to $u_i(\boldsymbol{x}, t) = \bar{u}_i(\boldsymbol{x}, t) + u'_i(\boldsymbol{x}, t)$. The operator $(\bar{\cdot})$ signifies the low-pass Gaussian filter and is given by:

$$\bar{u}_i(\boldsymbol{x}, t) = \iiint_V \frac{\sqrt{\pi}}{\bar{\Delta}} \exp\left[-\pi^2 (\boldsymbol{x} - \boldsymbol{x}')^2 / \bar{\Delta}^2\right] u_i(\boldsymbol{x}', t) \mathrm{d}\boldsymbol{x}', \tag{6}$$

where $\bar{\Delta}$ is the filter width and V denotes integration over the whole flow domain. Examples of filtered velocity fields at four different filter widths are included in Fig. 6. The kinetic energy of the large-scale field evolves as

$$\left(\frac{\partial}{\partial t} + \bar{u}_j \frac{\partial}{\partial x_j}\right) \frac{1}{2} \bar{u}_i \bar{u}_i = -\frac{\partial}{\partial x_j}\left(\bar{u}_j \bar{\Pi} + \bar{u}_i \tau_{ij} - 2\nu \bar{u}_i \bar{S}_{ij}\right) + \tilde{\Sigma} - 2\nu \bar{S}_{ij} \bar{S}_{ij} + \bar{u}_i \bar{f}_i, \tag{7}$$

where $\tau_{ij} = (\overline{u_i u_j} - \bar{u}_i \bar{u}_j)$ is the subgrid-scale stress tensor, which represents the effect of the (filtered) small-scale eddies on the (resolved) large-scale eddies and $\bar{S}_{ij} = (\partial \bar{u}_i / \partial x_j + \partial \bar{u}_j / \partial x_i)/2$ denotes the filtered strain-rate tensor. The interscale energy transfer $\tilde{\Sigma}_i(\boldsymbol{x}, t; \bar{\Delta}_i)$ between the filtered and unfiltered scales is given by

$$\tilde{\Sigma}_i(\boldsymbol{x}, t; \bar{\Delta}_i) = \tau_{ij}(\boldsymbol{x}, t; \bar{\Delta}_i) \bar{S}_{ij}(\boldsymbol{x}, t; \bar{\Delta}_i), \tag{8}$$

which is the quantity of interest. The velocity field was low-passed filtered at four filter widths: $\bar{\Delta}_1 = 163\eta$, $\bar{\Delta}_2 = 81\eta$, $\bar{\Delta}_3 = 42\eta$, and $\bar{\Delta}_4 = 21\eta$. The filter widths are selected to represent four different flow scales and are located within the inertial range of the simulation: $L_\varepsilon > \bar{\Delta}_i > \eta$, for $i = 1, 2, 3$ and 4, where L_ε represents the size of the largest scales and η is the size of the smallest scales. Finally, we volume-averaged $\tilde{\Sigma}_i$ over the entire domain, denoted by Σ_i, which served as a marker for the dynamics of the energy cascade. The generated data are also resolved in time, with flow fields stored at intervals of $\Delta t = 0.0076T_\varepsilon$, where T_ε is the characteristic time of the largest flow scales. The simulation was intentionally run for an extended period to ensure the accurate computation of specific mutual information. The total simulated time after transient effects was equal to $165T_\varepsilon$. For a given target variable, Σ_j, the time delay ΔT_j used to evaluate causality was determined as the time required for maximum $\Delta I_{i \to j}^U$ with $j \neq i$, where Σ_j^+ is evaluated at $t + \Delta T_j$.

Data for turbulent boundary layer

The data used for analyzing inner/outer interactions in a turbulent boundary layer were obtained from a experimental campaign at the high Reynolds number wind tunnel at the University of Melbourne[74–76], which is publicly available in https://fluids.eng.unimelb.edu.au/. In this campaign, measurements were made at a streamwise distance of $x = 21.65\,\mathrm{m}$ from the trip at the test section inlet ($x = 0$), with a free-stream velocity of nominally 20 m/s. The boundary layer at this location has a thickness of $\delta = 0.361\,\mathrm{m}$, and the friction velocity is $u_\tau = 0.626\,\mathrm{m/s}$. Based on these values, the friction Reynolds number is $Re_\tau = u_\tau \delta / \nu = 14{,}750$. The data used in this study includes measurements of the streamwise velocity obtained using two synchronous hot-wire anemometry probes with an acquisition rate $\Delta t^* = 1.28$ at two different wall-normal locations: $y_I^* = 4.33$ (for the inner layer) and $y_O/\delta = 0.31$ (for the outer layer). The superscript * denotes the inner scaling with friction velocity, u_τ, and kinematic viscosity, ν. At each location, the acquisition time consists of three cycles of approximately $TU_\infty/\delta = 20{,}000$. Further details about the experimental setup can be found in ref. 74 and Marusic[76]. The time lag utilized to evaluate causality is $\Delta T^* \approx 756$, which corresponds to the time lag for maximum cross-induced unique causality.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data generated in this study as well as the analysis and simulation code have been deposited in a Zenodo database under identifier https://doi.org/10.5281/zenodo.13750918.

Code availability

The codes[129] developed for this work are available at: https://github.com/Computational-Turbulence-Group/SURD.

References

1. Pearl, J. *Causality: Models, Reasoning, and Inference* (Cambridge University Press, 2000).
2. Bunge, M. *Causality and Modern Science* (Dover Publications, 1979).
3. Runge, J., Gerhardus, A., Varando, G., Eyring, V. & Camps-Valls, G. Causal inference for time series. *Nat. Rev. Earth Environ.* **4**, 487–505 (2023).
4. Razi, A. & Friston, K. J. The connected brain: causality, models, and intrinsic dynamics. *IEEE Signal Process. Mag.* **33**, 14–35 (2016).
5. Chiou-Wei, S. Z., Chen, C. F. & Zhu, Z. Economic growth and energy consumption revisited—evidence from linear and non-linear Granger causality. *Energy Econ.* **30**, 3063–3076 (2008).
6. Rothman, K. J. & Greenland, S. Causation and causal inference in epidemiology. *Am. J. Public Health* **95**, S144–S150 (2005).
7. Hedström, P. & Ylikoski, P. Causal mechanisms in the social sciences. *Annu. Rev. Sociol.* **36**, 49–67 (2010).
8. Lozano-Durán, A., Bae, H. J. & Encinar, M. P. Causality of energy-containing eddies in wall turbulence. *J. Fluid Mech.* **882**, A2 (2020).
9. Martínez-Sánchez, Á. et al. Causality analysis of large-scale structures in the flow around a wall-mounted square cylinder. *J. Fluid Mech.* **967**, A1 (2023).
10. Eichler, M. Causal inference with multiple time series: principles and problems. *Philos. Trans. R. Soc. A* **371**, 20110613 (2013).
11. Barndorff-Nielsen, O. E. & Kluppelberg, C. *Complex stochastic systems* (Chapman and Hall, CRC, 2001).
12. Spirtes, P., Glymour, C. & Scheines, R. *Causation, Prediction, and Search* (The MIT Press, 2001).
13. Dawid, A. P. Influence diagrams for causal modelling and inference. *Int. Stat. Rev.* **70**, 161–189 (2002).
14. Kampa, M. & Castanas, E. Human health effects of air pollution. *Environ. Pollut.* **151**, 362–367 (2008).
15. Altman, N. & Krzywinski, M. Association, correlation and causation. *Nat. Methods* **12**, 899–900 (2015).
16. Pearson, K. & Galton, F. VII. Note on regression and inheritance in the case of two parents. *Proc. R. Soc. Lond.* **58**, 240–242 (1895).
17. Spearman, C. The proof and measurement of association between two things. *Am. J. Psychol.* **100**, 441–471 (1987).
18. Agresti, A. & Franklin, C. *Statistics: The Art and Science of Learning from Data* 5th edn (Pearson, 2024).
19. Horn, K. J. et al. Growth and survival relationships of 71 tree species with nitrogen and sulfur deposition across the conterminous U.S. *PLOS ONE* **13**, 1–19 (2018).
20. Duan, R.-R., Hao, K. & Yang, T. Air pollution and chronic obstructive pulmonary disease. *Chronic Dis. Transl. Med.* **6**, 260–269 (2020).
21. Camps-Valls, G. et al. Discovering causal relations and equations from data. *Phys. Rep.* **1044**, 1–68 (2023).
22. Eberhardt, F. & Scheines, R. Interventions and causal inference. *Philos. Sci.* **74**, 981–995 (2007).
23. Runge, J., Nowack, P., Kretschmer, M., Flaxman, S. & Sejdinovic, D. Detecting and quantifying causal associations in large nonlinear time series datasets. *Sci. Adv.* **5**, eaau4996 (2019).
24. Wiener, N. *The Theory of Prediction. Modern Mathematics for Engineers* Vol. 165 (Dover Publications, 1956).
25. Granger, C. W. J. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* **37**, 424–438 (1969).
26. Geweke, J. F. Measures of conditional linear dependence and feedback between time series. *J. Am. Stat. Assoc.* **79**, 907–915 (1984).
27. Barnett, L., Barrett, A. B. & Seth, A. K. Granger causality and transfer entropy are equivalent for Gaussian variables. *Phys. Rev. Lett.* **103**, 238701 (2009).
28. Barnett, L. & Seth, A. K. The MVGC multivariate Granger causality toolbox: a new approach to Granger-causal inference. *J. Neurosci. Methods* **223**, 50–68 (2014).
29. Barnett, L. & Seth, A. K. Granger causality for state-space models. *Phys. Rev. E* **91**, 040101 (2015).
30. Hiemstra, C. & Jones, J. D. Testing for linear and nonlinear Granger causality in the stock price-volume relation. *J. Financ.* **49**, 1639–1664 (1994).
31. Bell, D., Kay, J. & Malley, J. A non-parametric approach to non-linear causality testing. *Econ. Lett.* **51**, 7–18 (1996).
32. Abhyankar, A. Linear and nonlinear Granger causality: evidence from the UK stock index futures market. *J. Futur. Mark.* **18**, 519 (1998).
33. Tissot, G., Lozano-Durán, A., Jiménez, J., Cordier, L. & Noack, B. R. Granger causality in wall-bounded turbulence. *J. Phys. Conf. Ser* **506**, 012006 (2014).
34. Ancona, N., Marinazzo, D. & Stramaglia, S. Radial basis function approach to nonlinear Granger causality of time series. *Phys. Rev. E* **70**, 056221 (2004).
35. Bueso, D., Piles, M. & Camps-Valls, G. Explicit Granger causality in kernel Hilbert spaces. *Phys. Rev. E* **102**, 062201 (2020).
36. Sugihara, G. et al. Detecting causality in complex ecosystems. *Science* **338**, 496–500 (2012).
37. McCracken, J. M. & Weigel, R. S. Convergent cross-mapping and pairwise asymmetric inference. *Phys. Rev. E* **90**, 062903 (2014).
38. Clark, A. T. et al. Spatial convergent cross mapping to detect causal relationships from short time series. *Ecology* **96**, 1174–1181 (2015).
39. Ye, H., Deyle, E. R., Gilarranz, L. J. & Sugihara, G. Distinguishing time-delayed causal interactions using convergent cross mapping. *Sci. Rep.* **5**, 14750 (2015).
40. Leng, S. et al. Partial cross mapping eliminates indirect causal influences. *Nat. Commun.* **11**, 2632 (2020).

41. Brouwer, E. D., Arany, A., Simm, J., & Moreau, Y. Latent convergent cross mapping. in *International Conference on Learning Representations* (2021).
42. Takens, F. Detecting strange attractors in turbulence. In *Dynamical Systems and Turbulence, Warwick 1980* (eds Rand, D. & Young, L.-S.) 366–381 (Springer Berlin Heidelberg, 1981).
43. Ying, X. et al. Continuity scaling: a rigorous framework for detecting and quantifying causality accurately. *Research* **2022**, 9870149 (2022).
44. Shannon, C. E. A mathematical theory of communication. *Bell Labs Tech. J.* **27**, 379–423 (1948).
45. Lozano-Durán, A. & Arranz, G. Information-theoretic formulation of dynamical systems: causality, modeling, and control. *Phys. Rev. Res.* **4**, 023195 (2022).
46. Yuan, Y. & Lozano-Durán, A. Limits to extreme event forecasting in chaotic systems. *Phys. D* **467**, 134246 (2024).
47. Massey, J. Causality, feedback and directed information. In *Proc. 1990 Int. Symp. on Inform. Theory and its Applications*, 27–30 (1990).
48. Kramer, G. *Directed information for channels with feedback*. PhD Thesis, ETH Zürich (1998).
49. Schreiber, T. Measuring information transfer. *Phys. Rev. Lett.* **85**, 461 (2000).
50. Verdes, P. Assessing causality from multivariate time series. *Phys. Rev. E* **72**, 026222 (2005).
51. Lizier, J. T., Prokopenko, M. & Zomaya, A. Y. Local information transfer as a spatiotemporal filter for complex systems. *Phys. Rev. E* **77**, 026110 (2008).
52. Lizier, J. T., Prokopenko, M. & Zomaya, A. Y. Information modification and particle collisions in distributed computation. *Chaos* **20**, 037109 (2010).
53. Bossomaier, T., Barnett, L., Harré, M. & Lizier, J. T. *An Introduction to Transfer Entropy: Information Flow in Complex Systems* 1st edn (Springer International Publishing, 2016).
54. Pompe, B. & Runge, J. Momentary information transfer as a coupling measure of time series. *Phys. Rev. E* **83**, 051122 (2011).
55. Liang, X. S. & Kleeman, R. Information transfer between dynamical system components. *Phys. Rev. Lett.* **95**, 244101 (2006).
56. Liang, X. S. Information flow and causality as rigorous notions ab initio. *Phys. Rev. E* **94**, 052201 (2016).
57. Liang, X. S. Information flow within stochastic dynamical systems. *Phys. Rev. E* **78**, 031113 (2008).
58. Liang, X. S. The Liang-Kleeman information flow: theory and applications. *Entropy* **15**, 327–360 (2013).
59. Spirtes, P. & Glymour, C. An algorithm for fast recovery of sparse causal graphs. *Soc. Sci. Comput. Rev.* **9**, 62–72 (1991).
60. Runge, J. Conditional independence testing based on a nearest-neighbor estimator of conditional mutual information. In *Proc. Twenty-First International Conference on Artificial Intelligence and Statistics,* Vol. 84 (eds Storkey, A. & Perez-Cruz, F.) 938–947 (PMLR, 2018).
61. Runge, J. Modern causal inference approaches to investigate biodiversity-ecosystem functioning relationships. *Nat. Commun.* **14**, 1917 (2023).
62. Runge, J. Discovering contemporaneous and lagged causal relations in autocorrelated nonlinear time series datasets. In *Conference on Uncertainty in Artificial Intelligence* 1388–1397 (PMLR, 2020).
63. Gerhardus, A. & Runge, J. High-recall causal discovery for autocorrelated time series with latent confounders. In *Advances in Neural Information Processing Systems* Vol. 33 (eds Larochelle, H. et al.) 12615–12625 (Curran Associates, Inc., 2020).
64. Saggioro, E., de Wiljes, J., Kretschmer, M. & Runge, J. Reconstructing regime-dependent causal relationships from observational time series. *Chaos* **30**, 113115 (2020).
65. Kullback, S. & Leibler, R. A. On information and sufficiency. *Ann. Math. Stat.* **22**, 79–86 (1951).
66. Kreer, J. A question of terminology. *IEEE Trans. Inf. Theory* **3**, 208–208 (1957).
67. Ince, R. A. A. Measuring multivariate redundant information with pointwise common change in surprisal. *Entropy* **19**, 318 (2017).
68. Lotka, A. J. Elements of physical biology. *Nature* **116**, 461–461 (1925).
69. Volterra, V. Fluctuations in the abundance of a species considered mathematically. *Nature* **118**, 558–560 (1926).
70. Moran, P. A. The statistical analysis of the Canadian lynx cycle. *Aust. J. Zool.* **1**, 291–298 (1953).
71. Richardson, L. F. *Weather Prediction by Numerical Process* (Cambridge University Press, 1922).
72. Obukhov, A. M. On the distribution of energy in the spectrum of turbulent flow. *Izv. Akad. Nauk USSR, Ser. Geogr. Geofiz.* **5**, 453–466 (1941).
73. Kolmogorov, A. N. The local structure of turbulence in incompressible viscous fluid for very large Reynolds' numbers. *Dokl. Akad. Nauk SSSR* **30**, 301–305 (1941).
74. Baars, W. J., Talluru, K. M., Hutchins, N. & Marusic, I. Wavelet analysis of wall turbulence to study large-scale modulation of small scales. *Exp. Fluids* **56**, 188 (2015).
75. Baars, W. J., Hutchins, N. & Marusic, I. Reynolds number trend of hierarchies and scale interactions in turbulent boundary layers. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **375**, 20160077 (2017).
76. Marusic, I. Two-point high Reynolds number zero-pressure gradient turbulent boundary layer dataset. *University of Melbourne* https://figshare.unimelb.edu.au/articles/dataset/Two-point_high_Reynolds_number_zero pressure_gradient_turbulent_boundary_layer_dataset/12101088 (2020).
77. Ding, M., Chen, Y. & Bressler, S. Granger causality: basic theory and application to neuroscience. In *Handbook of Time Series Analysis: Recent Theoretical Developments and Applications*, (eds Schelter, B., Winterhalder, M. & Timmer, J.) 2437–2459 (Wiley-VCH, 2006).
78. May, R. M. Simple mathematical models with very complicated dynamics. *Nature* **261**, 459–467 (1976).
79. Lorenz, E. N. Deterministic nonperiodic flow. *J. Atmos. Sci.* **20**, 130–141 (1963).
80. Rössler, O. E. Continuous chaos. In *Synergetics: A Workshop Proceedings of the International Workshop on Synergetics at Schloss Elmau, Bavaria, May 2–7* 184–197. (Springer, 1977).
81. Quiroga, R. Q., Arnhold, J. & Grassberger, P. Learning driver-response relationships from synchronization patterns. *Phys. Rev. E* **61**, 5142–5148 (2000).
82. Krakovská, A. et al. Comparison of six methods for the detection of causality in a bivariate time series. *Phys. Rev. E* **97**, 042207 (2018).
83. Javier, P. J. E. causal-ccm: a Python implementation of Convergent Cross Mapping (2021).
84. Kolmogorov, A. N. A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number. *J. Fluid Mech.* **13**, 82–85 (1962).
85. Aoyama, T. et al. Statistics of energy transfer in high-resolution direct numerical simulation of turbulence in a periodic box. *J. Phys. Soc. Jpn.* **74**, 3202–3212 (2005).
86. Falkovich, G. Symmetries of the turbulent state. *J. Phys. A* **42**, 123001 (2009).
87. Cardesa, J. I., Vela-Martín, A. & Jiménez, J. The turbulent cascade in five dimensions. *Science* **357**, 782–784 (2017).
88. Yamada, T. et al. Anatomy of plasma turbulence. *Nat. Phys.* **4**, 721–725 (2008).
89. Veynante, D. & Vervisch, L. Turbulent combustion modeling. *Prog. Energy Combust. Sci.* **28**, 193–266 (2002).

90. Bodenschatz, E. Clouds resolved. *Science* **350**, 40–41 (2015).
91. Young, R. M. B. & Read, P. L. Forward and inverse kinetic energy cascades in Jupiter's turbulent weather layer. *Nat. Phys.* **13**, 1135–1140 (2017).
92. Sirovich, L. & Karlsson, S. Turbulent drag reduction by passive mechanisms. *Nature* **388**, 753–755 (1997).
93. Hof, B., De Lozar, A., Avila, M., Tu, X. & Schneider, T. M. Eliminating turbulence in spatially intermittent flows. *Science* **327**, 1491–1494 (2010).
94. Marusic, I., Mathis, R. & Hutchins, N. Predictive model for wall-bounded turbulent flow. *Science* **329**, 193–196 (2010).
95. Kühnen, J. et al. Destabilizing turbulence in pipe flow. *Nat. Phys.* **14**, 386–390 (2018).
96. Vela-Martín, A. & Jiménez, J. Entropy, irreversibility and cascades in the inertial range of isotropic turbulence. *J. Fluid Mech.* **915**, A36 (2021).
97. Vela-Martín, A. Subgrid-scale models of isotropic turbulence need not produce energy backscatter. *J. Fluid Mech.* **937**, A14 (2022).
98. Taylor, G. I. Statistical theory of turbulence. *Proc. R. Soc. Lond.* **151**, 444–454 (1935).
99. Frisch, U. *Turbulence: The Legacy of A. N. Kolmogorov* (Cambridge University Press, 1995).
100. Zhou, Y. Degrees of locality of energy transfer in the inertial range. *Phys. Fluids* **5**, 1092–1094 (1993).
101. Eyink, G. L. Locality of turbulent cascades. *Physica D* **207**, 91–116 (2005).
102. Mininni, P., Alexakis, A. & Pouquet, A. Large-scale flow effects, energy transfer, and self-similarity on turbulence. *Phys. Rev. E* **74**, 016303 (2006).
103. Aluie, H. & Eyink, G. L. Localness of energy cascade in hydrodynamic turbulence. II. Sharp spectral filter. *Phys. Fluids* **21**, 115108 (2009).
104. Domaradzki, J. A., Teaca, B. & Carati, D. Locality properties of the energy flux in turbulence. *Phys. Fluids* **21**, 025106 (2009).
105. Townsend, A. A. *The Structure of Turbulent Shear Flow* (Cambridge University Press, 1976).
106. Hutchins, N. & Marusic, I. Evidence of very long meandering features in the logarithmic region of turbulent boundary layers. *J. Fluid Mech.* **579**, 1–28 (2007).
107. Mathis, R., Hutchins, N. & Marusic, I. Large-scale amplitude modulation of the small-scale structures in turbulent boundary layers. *J. Fluid Mech.* **628**, 311–337 (2009).
108. Flack, K. A., Schultz, M. P. & Shapiro, T. A. Experimental support for Townsend's Reynolds number similarity hypothesis on rough walls. *Phys. Fluids* **17**, 035102 (2005).
109. Flores, O. & Jiménez, J. Effect of wall-boundary disturbances on turbulent channel flows. *J. Fluid Mech.* **566**, 357–376 (2006).
110. Busse, B. & Sandham, A. Parametric forcing approach to rough-wall turbulent channel flow. *J. Fluid Mech.* **712**, 169–202 (2012).
111. Mizuno, Y. & Jiménez, J. Wall turbulence without walls. *J. Fluid Mech.* **723**, 429–455 (2013).
112. Chung, D., Monty, J. P. & Ooi, A. An idealised assessment of Townsend's outer-layer similarity hypothesis for wall turbulence. *J. Fluid Mech.* **742**, R3 (2014).
113. Lozano-Durán, A. & Bae, H. J. Characteristic scales of Townsend's wall-attached eddies. *J. Fluid Mech.* **868**, 698–725 (2019).
114. Williams, P. L. & Beer, R. D. Nonnegative decomposition of multivariate information. *arXiv preprint arXiv:1004.2515* (2010).
115. Griffith, V. & Koch, C. Quantifying synergistic mutual information. in *Guided Self-Organization: Inception* 159–190 (Springer, 2014).
116. Griffith, V. & Ho, T. Quantifying redundant information in predicting a target random variable. *Entropy* **17**, 4644–4653 (2015).
117. Ince, R. A. Measuring multivariate redundant information with pointwise common change in surprisal. *Entropy* **19**, 318 (2017).
118. Gutknecht, A. J., Wibral, M. & Makkeh, A. Bits and pieces: understanding information decomposition from part-whole relationships and formal logic. *Proc. R. Soc. A* **477**, 20210110 (2021).
119. Kolchinsky, A. A novel approach to the partial information decomposition. *Entropy* **24**, 403 (2022).
120. Baptista, R., Marzouk, Y. & Zahm, O. On the representation and learning of monotone triangular transport maps. *Found. Comput. Math.* https://doi.org/10.1007/s10208-023-09630-x (2023).
121. Cobey, S. & Baskerville, E. B. Limits to causal inference with state-space reconstruction for infectious disease. *PLoS ONE* **11**, 1–22 (2016).
122. Mønster, D., Fusaroli, R., Tylén, K., Roepstorff, A. & Sherson, J. F. Causal inference from noisy time-series data—testing the convergent cross-mapping algorithm in the presence of noise and external influence. *Future Gener. Comput. Syst.* **73**, 52–62 (2017).
123. Runge, J. Causal network reconstruction from time series: From theoretical assumptions to practical estimation. *Chaos* **28**, 075310 (2018).
124. DeWeese, M. R. & Meister, M. How to measure the information gained from one symbol. *Netw. Comput. Neural Syst.* **10**, 325 (1999).
125. Lizier, J. T. JIDT: an information-theoretic toolkit for studying the dynamics of complex systems. *Front. Robot. AI.* **1**, 11 (2014).
126. Akaike, H. Akaike's Information Criterion. in *International Encyclopedia of Statistical Science* (ed Lovric, M.) (Springer, 2011).
127. Cardesa, J. I., Vela-Martín, A., Dong, S. & Jiménez, J. The temporal evolution of the energy flux across scales in homogeneous turbulence. *Phys. Fluids* **27**, 111702 (2015).
128. Rosales, C. & Meneveau, C. Linear forcing in numerical simulations of isotropic turbulence: physical space implementations and convergence properties. *Phys. Fluids* **17**, 095106 (2005).
129. Martínez-Sánchez, Á., Arranz, G. & Lozano-Durán, A. Decomposing causality in its synergistic, unique, and redundant components, SURD: Synergistic-Unique-Redundant Decomposition of causality. https://doi.org/10.5281/zenodo.13750918 (2024).

Acknowledgements

Author contributions

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

Competing interests

The authors declare no competing interests.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.