Regular Article

Predictors and predictands of linear response in spatially extended systems

Umberto Maria Tomasini[1,2,3,4,a] and Valerio Lucarini[3,4,b]

[1] Dipartimento di Fisica e Astronomia, Universitá di Padova, Padua, Italy
[2] Institute of Physics, EPFL, Lausanne, Switzerland
[3] Department of Mathematics and Statistics, University of Reading, Reading, UK
[4] Centre for the Mathematics of Planet Earth, University of Reading, Reading, UK

Abstract The goal of response theory, in each of its many statistical mechanical formulations, is to predict the perturbed response of a system from the knowledge of the unperturbed state and of the applied perturbation. A new recent angle on the problem focuses on providing a method to perform predictions of the change in one observable of the system using the change in a second observable as a surrogate for the actual forcing. Such a viewpoint tries to address the very relevant problem of causal links within complex system when only incomplete information is available. We present here a method for quantifying and ranking the predictive ability of observables and use it to investigate the response of a paradigmatic spatially extended system, the Lorenz '96 model. We perturb locally the system and we then study to what extent a given local observable can predict the behaviour of a separate local observable. We show that this approach can reveal insights on the way a signal propagates inside the system. We also show that the procedure becomes more efficient if one considers multiple acting forcings and, correspondingly, multiple observables as predictors of the observable of interest.

1 Introduction

Elements of response theory

Response theory is an area of statistical physics that provides general methods for predicting the changes in the statistical properties of an observable of interest Ψ from the knowledge of the applied perturbation and of the statistical properties of the unperturbed system. The expectation value of Ψ in the perturbed system is expressed as a perturbative series, where the zeroth-order term is the expectation value of Ψ in the unperturbed system. The higher order terms are expressed in terms of response functions that contain information about the higher order statistics of the unperturbed system and the applied forcing

A cornerstone in the development of response theory came with the work by Kubo [1,2], who considered the case of weakly perturbed systems near thermodynamic equilibrium. When a system is in this kind of steady state, it obeys detailed balance and features no net currents. In Kubo's theory, from the knowledge of the statistics of the unperturbed system, and in particular of the correlations describing its spontaneous fluctuations, it is possible to compute, in linear approximation, the response of the system to any (weak) perturbation. This is the key idea behind the fluctuation–dissipation theorem (FDT) [2], which establishes a link between the forced and free fluctuations in the perturbative regime [3]. A generalised version of the FDT valid for higher moments has been proposed by [4] in the spirit of Zubarev's generalization of Kubo's results [5]. We will discuss from now on the special and very relevant case of linear response theory (LRT) and of its applications.

The analysis of the response to perturbations of systems that are far from thermodynamics equilibrium requires a more general approach than what provided by Kubo's theory. Indeed, nonequilibrium systems are ubiquitous and their investigations has both great theoretical relevance and allows one to study many important real-life examples in material science, physics, ecology, fluid dynamics, climate science, biology, among others. Often, one says that after transients have died out, in absence of time-dependent forcings, a nonequilibrium system is in the so-called nonequilibrium steady state, where detailed balance is not obeyed, and, in many cases, dissipative processes are associated with a contraction of the phase space and with the production of entropy [6]. Studying the response of NESS systems to perturbations is of great relevance both for theory and applications. At this regard, a rigorous and crucial development in the context of deterministic dynamics was provided by Ruelle [7–10], who rigorously derived a LRT for smooth observable of Axiom

[a] e-mail: umberto.tomasini@epfl.ch
[b] e-mail: v.lucarini@reading.ac.uk (corresponding author)

A dynamical systems. Ruelle's results have been been recast and extended by methods of functional analysis [11,12]. Roughly speaking, Axiom A systems provide an extremely useful example of chaos characterized by the presence of a clear separation between the expanding and contracting directions in the tangent space, formalized by the concept of uniform hyperbolicity. Crucially, Axiom A systems possess a special ergodic invariant measure—the so-called Sinai–Ruelle–Bowen measure [7]—that makes them excellent candidates for describing non-equilibrium physical systems. While Axiom A systems are mathematically quite special, their practical relevance is clarified by the chaotic hypothesis proposed by Gallavotti and Cohen [13,14], which can be see an extension of the ergodic hypothesis to the non-equilibrium case.

The chaotic hypothesis is supportive of the fact that LRT should de facto work in a large class of chaotic dynamical systems. As opposed to the equilibrium case, general non-equilibrium dynamical systems obeying Axiom A dynamics possess an invariant measure that is singular with respect to Lebesgue and is supported on a strange attractor. Ruelle proved that the response operator is given by the sum of two contributions. One is related to the dynamics on the unstable and central manifolds and can be framed as a FDT result, while the second one is related to the dynamics occurring on the stable manifold, and cannot be framed as a FDT result because of the non-smoothness of the measure [10]. In other words, the natural fluctuations are not equivalent to the forced perturbations along the stable directions [15–17]. The direct computation of these two contributions is far from trivial [18–20], but the use of adjoint and shadowing methods has recently led to promising results in this direction [21–23].

LRT can also be rigorously established for stochastic dynamical systems [20,24–26]. Under rather general hypotheses, adding noise makes the invariant measure absolutely continuous with respect to Lebesgue, so that the FDT fully holds [27]. In this setting, the obtained formula is called the Kubo–Agarwal formula, which reduces to the Kubo formula for equilibrium systems. The addition of a noise term has to be justified by the nature of the considered problem. This stochastic perspective becomes relevant in many complex systems, where the focus is on coarse-grained dynamics, which is effectively stochastic as a result of the presence of microscopic degrees of freedom. Note that the coarse-grained dynamics is in general non-Markovian, with memory effect becoming negligible in the limit of infinite time-scale separation between the fast and slow variables [28–32].

The relationship between response theory in deterministic and stochastic systems has been thoroughly discussed in [33], where the authors also propose a very well-developed justification, not based on the chaotic hypothesis, for the broad pragmatic applicability of LRT in a vast class of deterministic chaotic systems.

Nowadays, LRT has an important role in the investigations of a vast range of systems, see, e.g., [20,34–40]. As an example, in the case of climate science, LRT makes it possible to perform climate change projections using climate models of different levels of complexity. This amounts to computing the time-dependent measure supported on pullback attractor of the climate [41] by constructing response operators for a suitably defined reference climatic state [39]. This viewpoint allows one to investigate ways to control the future pathways of climate change [42] and to make an assessment of potential and pitfalls of geoengineering strategies [43].

Recently, LRT has been extended in such a way that explicit formulas are given for describing how adding a forcing to a system changes its $n-$point correlations [44]. Another recent application of LRT has focused on detecting and characterising phase transitions in a network of coupled identical agents undergoing a stochastic evolution [45]. A recently published special issue showcases several emerging areas of applications for LRT [46].

Majda and Qi have recently shown the existence of a coherent thread connecting LRT, sensitivity analysis, model reduction techniques, uncertainty quantification, and control of high-dimensional chaotic and stochastic systems [47,48]. The theme of studying the response using incomplete information on the system is the main topic of the present contribution .

Predictors and predictands

A different angle on the problem of defining the response of a system to perturbations proposes a fairly general method that allows one to relate the response of different observables of a system undergoing a perturbation. This method could be useful in situations, where we need to interpret experimental data and we do not have necessarily perfect knowledge of the properties of the system—e.g., we do not know the specific features of the applied forcing, have access to only a limited number of degrees of freedom of the system, or ignore altogether the evolution equation of the system itself—but can measure instead multiple observables at the same time [49]. This is closely related to finding solutions to nonlinear rational least squares problems [50,51].

The goal is to understand to what extent we can use perturbed observables as surrogates of the perturbation to reconstruct the time behaviour of other observables. It turns out that, if we know the time behaviour of one observable Ψ_1, we can always reconstruct diagnostically, within linear approximation, the time behaviour of another observable Ψ_2 through a surrogate response function.

Instead, if the goal is to actually predict the future state of the observable Ψ_2, by means of a prognostic relation, it turns out that not all choices of predictor are equally successful to perform the prognosis of a given predictand, because the surrogate response function might be, as opposed to the standard response (Green) function, not causal, i.e., its support is not limited to non-negative times.

The analysis performed on the Lorenz '96 model [52–54] in [49] explicitly showed that the response of an observable Ψ_1 to a given perturbations could be predicted by the response of a certain observable Ψ_2 but could not be predicted by the response of a third one Ψ_3, because, in the latter case, the surrogate response function has a non-causal component.

This paper

Following the seminal contribution by Granger [55], more and more attention has been recently paid to finding rigorous ways for defining and detecting causal links within complex systems [56], and climate science has been a very successful fertile of application of such ideas [57–60].

In this paper, we address this class of problems using a different viewpoint. We want here to make a step forward compared to [49] with the goal to better quantify the skill of different observables in predicting the response of a given observable. Here, we propose a way to evaluate how much their corresponding surrogate response functions are not causal. This problem can emerge in a variety of situations, where we have more non-predictive surrogate response functions and we want to choose between them the one which provides the best prediction. For example, we could have a set of observables $\{\Psi_1, \ldots, \Psi_m\}$ and we could aim at finding out which one(s) can be used to predict most accurately the response of another observable Ψ_a. Any observable whose surrogate response function is predictive is equally good at this regard. If only one among the m observables features a predictive surrogate response function, the choice is obvious. The choice becomes less straightforward when all the observables $\{\Psi_1, \ldots, \Psi_m\}$ are not predictive. We would like to have a quantitative method that allows us to choose the most predictive one among them. Another setting where such a method could be helpful is when we want to understand whether an observable Ψ_a predicts better Ψ_b or viceversa, as a better predictive power can be linked to a causal link or a flow of information with a definite verse from an observable to the other.

Specifically, we investigate causal links in the context of a spatially extended system undergoing chaotic dynamics, namely, the paradigmatic Lorenz 96 (L96) model [52–54]. We investigate the response of the system to localised forcings, and we observe the system by looking at its local properties in different regions. We assess whether the response of local observables can be used to predict the response of other local observables of interest. One of the goals of this analysis is to see whether we can detect a clear indication of the way signals propagate inside the system. The recent contributions [61,62] pursue closely related scientific goals. We also explore the possibility suggested in [49] that the predictive skill of the surrogate approach proposed here could dramatically improve if one perturbs the system with multiple forcings and uses multiple observables as surrogate predictors.

The rest of the paper is structured as follows. In Sect. 2, after briefly reviewing the surrogate LRT [49], we present the predictability index (PI), which aims at quantifying how much a surrogate response function is non-predictive. We will remark that the presence of such an index provides the opportunity to build an hierarchy of observables in terms of their predictive power of other observables. In Sect. 3 we apply the surrogate LRT on the L96 model, considering local perturbations and local observables. We will also explore the impact of adding information gathered from global observables to predict the local forced response. In Sect. 4 we summarise the main findings of this study and present perspectives for future research. A set of Appendices provides supplementary material of possible interest for the reader. In Appendix A we provide an illustrative example to better explain the meaning of the PI. In Appendix B we provide evidence of the fact that in our experiments and data analysis we are in a linear regime of response. In Appendix C we clarify some asymptotic properties of the response of the Lorenz '96 model, while in Appendix D some specific properties of the surrogate response functions are discussed.

2 Surrogate response theory

Following Ruelle [8–10], we recapitulate very informally some basic elements of LRT by studying the effect of perturbing an Axiom A system of the form $\dot{x} = F(x)$, where $x \in \mathcal{M}$, a smooth compact manifold of dimension D. We introduce the following complex pattern of forcing, consisting of N independent perturbations:

$$F(x) \rightarrow F(x) + \sum_{l=1}^{N} e_l(t) G_l(x), \qquad (1)$$

where $G_1(x), \ldots, G_N(x)$ are $D-$dimensional smooth vector fields, while $e_1(t), \ldots e_N(t)$ are time patterns. In linear approximation, which is relevant in the case the applied forcings are small, it is possible to write the change in the expectation value of any smooth observable Ψ as follows:

$$\delta \langle \Psi \rangle(t) = \sum_{j=1}^{N} \int_{-\infty}^{\infty} \mathrm{d}\tau e_j(\tau) \Gamma_{\Psi, G_j}(t - \tau), \qquad (2)$$

where we have introduced the linear (Green) response functions Γ_{Ψ, G_j}, which mediate the effect of the time pattern of the perturbation e_j at time $\tau < t$ on the observable Ψ at time t. The response function can be seen as the expectation value in the unperturbed system of a highly nontrivial observable that depends on the applied forcing and on Ψ:

$$\Gamma_{\Psi, G_j}(t) = \Theta(t) \int \rho_0(\mathrm{d}y) G_j(y) \nabla_y (\Psi(y(t))), \qquad (3)$$

where Θ is the Heaviside distribution, which determines the causality of the response functions, ρ_0 is the invariant measure of the unperturbed system and $\Psi(y(t)) = S^t \Psi(y) = \exp(\mathcal{L}_0 t)\Psi(y)$ is the value of the observable Ψ at time t following the evolution in time according to the dynamics of the unperturbed system, with initial condition y. We have that S^t is the (unperturbed) Koopman operator and $\mathcal{L}_0 = F \cdot \nabla$ is its generator. By applying the Fourier transform to Eq. (2), we obtain the following identitys:

$$\delta\langle\Psi\rangle(\omega) = \sum_{j=1}^{N} e_j(\omega)\chi_{\Psi,G_j}(\omega), \qquad (4)$$

where $\chi_{\Psi,G_j}(\omega)$, $j = 1, \ldots, N$ are the so-called susceptibilities. Since the response functions $\Gamma_{\Psi,G_j}(t)$, $j = 1, \ldots, N$ are causal, under standard conditions of integrability the corresponding susceptibilities are analytic in the upper complex ω-plane [63,64].

2.1 Surrogate response functions

A different angle of the problem in LRT has been introduced in [49], where response relations between perturbed observables are built. These relations can be useful in a large variety of contexts, where the knowledge of the forcing is just partial, and we want to use perturbed observables to diagnose the state of other perturbed observables. We consider $N + 1$ independent observables $\Psi_1(x), \ldots, \Psi_N(x)$. In [49] it is shown that it is possible to express the linear change of the expectation value of an observable Ψ_{N+1} as a function of the ones of the other N observables:

$$\delta\langle\Psi_{N+1}\rangle(\omega) = \sum_{l=1}^{N} J_{N+1,l}(\omega)\delta\langle\Psi_l\rangle(\omega), \qquad (5)$$

where we take the surrogate of the N forcings using the other N observables through the surrogate susceptibilities $J_{N+1,l}$, $l = 1, \ldots, N$. In [49] explicit expressions for the surrogate susceptibilities $J_{N+1,l}(\omega)$ are derived:

$$\begin{pmatrix} J_{N+1,1}(\omega) \\ \vdots \\ J_{N+1,N}(\omega) \end{pmatrix} = \begin{pmatrix} \chi_{\Psi_1,G_1}(\omega) & \cdots & \chi_{\Psi_N,G_1}(\omega) \\ \vdots & \ddots & \vdots \\ \chi_{\Psi_1,G_N}(\omega) & \cdots & \chi_{\Psi_N,G_N}(\omega) \end{pmatrix}^{-1}$$
$$\times \begin{pmatrix} \chi_{\Psi_{N+1},G_1}(\omega) \\ \vdots \\ \chi_{\Psi_{N+1},G_N}(\omega) \end{pmatrix}. \qquad (6)$$

Once we have obtained the surrogate response functions, we plug them into Eq. (5), obtaining the following expression:

$$\delta\langle\Psi_{N+1}\rangle(t) = \sum_{l=1}^{N} \int_{-\infty}^{+\infty} d\tau H_{N+1,l}(t-\tau)\delta\langle\Psi_l\rangle(\tau). \qquad (7)$$

where $H_{N+1,l}(t)$ is the inverse Fourier transform of $J_{N+1,l}(\omega)$, $l = 1, \ldots, N$.

Note that the previous relation does not involve the time patterns of the acting forcings. This has important practical consequences. If we are able to derive the functions $H_{N+1,l}(t)$ or $J_{N+1,l}(\omega)$, $l = 1, \ldots, N$ from a probe experiment performed with a (broadband) time pattern of forcing, we can use Eq. (7) to reconstruct the time evolution of $\delta\langle\Psi_{N+1}\rangle(t)$ for any time pattern of the forcing using as input the time evolution of $\delta\langle\Psi_j\rangle(t)$, $l = 1, \ldots, N$. This will investigated in the next Sect. 3.2. This inverse problem can also be approached by considering monochromatic forcings and scanning across frequencies, in the spirit of the algorithms proposed in [50,51].

Equation (7) can be employed to diagnose, at the linear level in the perturbation, the time behaviour of Ψ_{N+1} by means of other N observables. On the other hand, it is not always possible to perform the prognosis of Ψ_{N+1} using the same observables, which can be useful in prediction problems. This requires that the response functions $H_{N+1,l}$ $\forall$ $l = 1, \ldots, N$ are causal, i.e., their support is in the non-negative domain.

Let us expand more on the case $N = 1$, to better clarify some issues associated with the surrogate LRT. Equations (6) and (7) become

$$J_{2,1}(\omega) \equiv \frac{\chi_{\Psi_1,G}(\omega)}{\chi_{\Psi_2,G}(\omega)},$$
$$\delta\langle\Psi_2\rangle(t) = \int_{-\infty}^{\infty} H_{2,1}(t-\tau)\delta\langle\Psi_1\rangle(\tau). \qquad (8)$$

The surrogate response function $H_{2,1}$ is predictive—i.e., it has support only on the non-negative domain—if and only if its Fourier transform $J_{2,1}$ has no poles in the upper complex ω−plane. Since the numerator $\chi_{\Psi_1,G}(\omega)$ is analytic in the upper complex ω plane, loss of predictability is realised only if the response function $\chi_{\Psi_2,G}(\omega)$ at the denominator has complex zeros in the upper complex ω plane, see discussion in [49].

Indeed, we expect that for a given forcing not all choices of predictors and predictands are equally successful in terms of predictive power. For instance, if there is a causal relation in a feedback or a flow of information linking observable Ψ_1 and Ψ_2, one expects the presence of an asymmetry in the mutual predictive power.

Lastly, we remark that the surrogate response function could have a singular component in 0, as it is noted in [49]. There is a close link between the short-time behaviour of $\Gamma_{\Psi_1,G}$ and the high-frequency behaviour of its Fourier transform:

$$\Gamma_{\Psi_1,G}(t) \approx \alpha_{\Psi_1,G}\Theta(t)t^\alpha \Leftrightarrow \chi_{\Psi_1,G}(\omega)$$
$$\approx (\alpha_{\Psi_1,G}\,\alpha!\,i^{\alpha+1})\frac{1}{\omega^{\alpha+1}} \qquad (9)$$

As a consequence, it is possible to obtain the asymptotic behaviour of the Fourier transform of the surrogate response function $J_{2,1}$ using the asymptotic behaviour

of the response functions. In particular, if for large values of ω we have that $\chi_{\Psi_1,G} \approx 1/\omega^{\alpha_1+1}$ and $\chi_{\Psi_2,G} \approx 1/\omega^{\alpha_2+1}$, we derive

$$J_{2,1}(\omega) \underset{\omega\to\infty}{\propto} \frac{1/\omega^{\alpha_1+1}}{1/\omega^{\alpha_2+1}}$$
$$\underset{\omega\to\infty}{\propto} \omega^{\alpha_2-\alpha_1}. \qquad (10)$$

If $\alpha_1 < \alpha_2$ $J_{2,1}(\omega)$ diverges for large values of ω, while it converges to a non-vanishing constant for $\alpha_1 = \alpha_2$. Hence, in these cases the surrogate response function $H_{2,1}(t)$ will have a singular component $S_{2,1}(t)$ in $t = 0$, because the Fourier transform of $(-i\omega)^j$ is $\delta^j(t)$, i.e., the jth derivative of the delta function $\delta(t)$. We can then write

$$H_{2,1}(t) = S_{2,1}(t) + K_{2,1}(t), \qquad (11)$$

where $K_{2,1}(t)$ $(S_{2,1}(t))$ is the non-singular (singular) component. On the contrary, if $\alpha_1 > \alpha_2$ the surrogate response function $H_{2,1}(t)$ has no singular component $S_{2,1}$.

Note that the exponent α describing the short time behaviour of the response function controls how long it takes for the observable to feel the effect of the forcing. The higher the exponent, the slower is the build-up of the effect of the forcing on the observable. Hence, it makes sense to use Ψ_1 to predict Ψ_2 only if $\alpha_1 \leq \alpha_2$, i.e., if Ψ_1 feels the applied perturbation before or approximately at the same time as Ψ_2. From now on we then exclude the case $\alpha_1 > \alpha_2$.

The part of the surrogate response function $H_{2,1}$ that is practically usable for predictions has support restricted to the non-negative domain, and can then be expressed as follows:

$$H_{2,1}^c(t) = \Theta(t)K_{2,1}(t) + S_{2,1}(t). \qquad (12)$$

In practice, since the outputs of numerical simulations have finite precision, $H_{2,1}^c(t)$ can be reconstructed from data as follows:

$$H_{2,1}^c(t) \equiv \lim_{\varepsilon\to 0^+} \left(\Theta(t+\varepsilon)H_{2,1}(t) \right), \qquad (13)$$

where the $\varepsilon-$ regularization has been introduced to include in the definition of $H_{2,1}^c$ possible singular components of $H_{2,1}(t)$ at $t = 0$.

2.2 Quantifying the ability to predict

The presence of the non-causal component in the surrogate response function $H_{2,1}$ hinders the prediction of Ψ_2 at time t using just the time behaviour of Ψ_1 up to time t. An interesting problem is to actually quantify the non-causal component of the surrogate response function. This quantification would allow to build an hierarchy of observables in terms of their predictive power of other observables.

From the discussion above, we then have that the surrogate response functions are of the form:

$$H_{2,1}(t) = S_{2,1}(t) + K_{2,1}(t)$$
$$= s_{2,1}\delta(t) + K_{2,1}(t), \qquad (14)$$

where the constant $s_{2,1} \in \mathbb{R}$ can also be zero. We propose to quantify the ability of the observable 1 to predict the observable 2 with the predictability index (PI), which is defined as follows:

$$R(H_{2,1}) \equiv \frac{\|K_{2,1}^{nc}(t)\|_1}{\|K_{2,1}^c(t)\|_1 + \|S_{2,1}\|_1}, \qquad (15)$$

where

$$K_{2,1}^c(t) = \Theta(t)K_{2,1}(t), \qquad K_{2,1}^{nc}(t) = \Theta(-t)K_{2,1}(t), \qquad (16)$$

hence $K_{2,1}^c$ is the causal part of the non-singular component of the surrogate response function, while $K_{2,1}^{nc}$ is its non-causal part. In addition, $\|\bullet\|_1$ is the standard L_1 norm. The PI depends on the system under investigation, the space pattern of the forcing $G(x)$ and on the observables Ψ_1 and Ψ_2.

The PI is non-negative and vanishes if and only if the surrogate response function is predictive, because its support includes only the non-negative domain. A large value for the PI indicates that the response of the observable 1 is a poor predictor of the response of the observable 2. Moreover, since this method revolves around the surrogate response function, it does not depend on the chosen time pattern. We will actually see the effectiveness of this indicator in the L96 model in Sect. 3. A few pedagogical examples can also be found in Appendix A.

We can generalize the indicator given in Eq. (15) to the case when we use Ψ_l, $l = 1, \ldots, N$ observables as predictors of the observable Ψ_{N+1}, as in Eq. 5. For each surrogate response function $H_{N+1,l}(t)$, with $l \in \{1, \ldots, N\}$, we define its singular part $S_{N+1,l}(t)$ and its non-singular part $K_{N+1,l}(t)$, which, in turn, can be split into the non-causal component $K_{N+1,l}^{nc}$ and the causal component $K_{N+1,l}^c$. We assume that all these surrogate response functions are of the type Eq. (14). We then define

$$R(\{H_{N+1,l}\}_{l=1,\ldots,N}) \equiv \frac{\sum_{l=1}^N \|K_{N+1,l}^{nc}(t)\|_1}{\sum_{l=1}^N \left(\|K_{N+1,l}^c(t)\|_1 + \|S_{N+1,l}(t)\|_1 \right)}. \qquad (17)$$

3 The Lorenz '96 model

3.1 Model formulation

The L96 model [52–54] is a paradigmatic model that provides a metaphor of some essential properties of the

Fig. 1 Plots of the response functions $\Gamma_{i,k}$ for $i = \{k-2, k-1, k, k+1, k+2\}$ and $\epsilon = 1$. We also portray $\Gamma_{\Psi_1,k}$, the response function for the mean momentum Ψ_1 (Eq. 26). see also Fig. 14
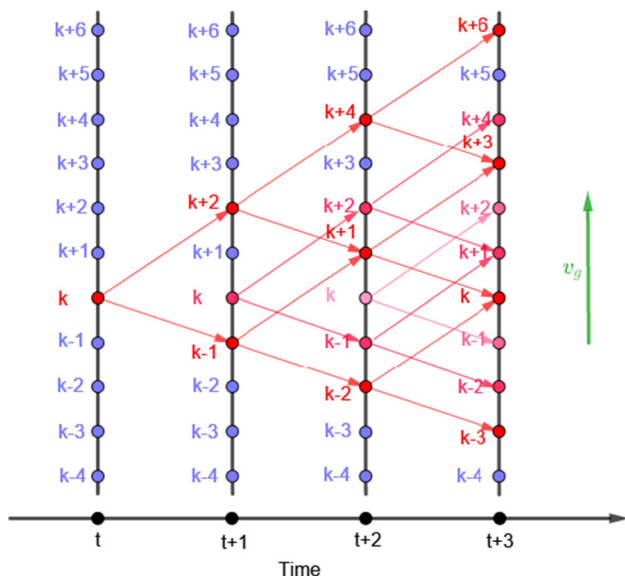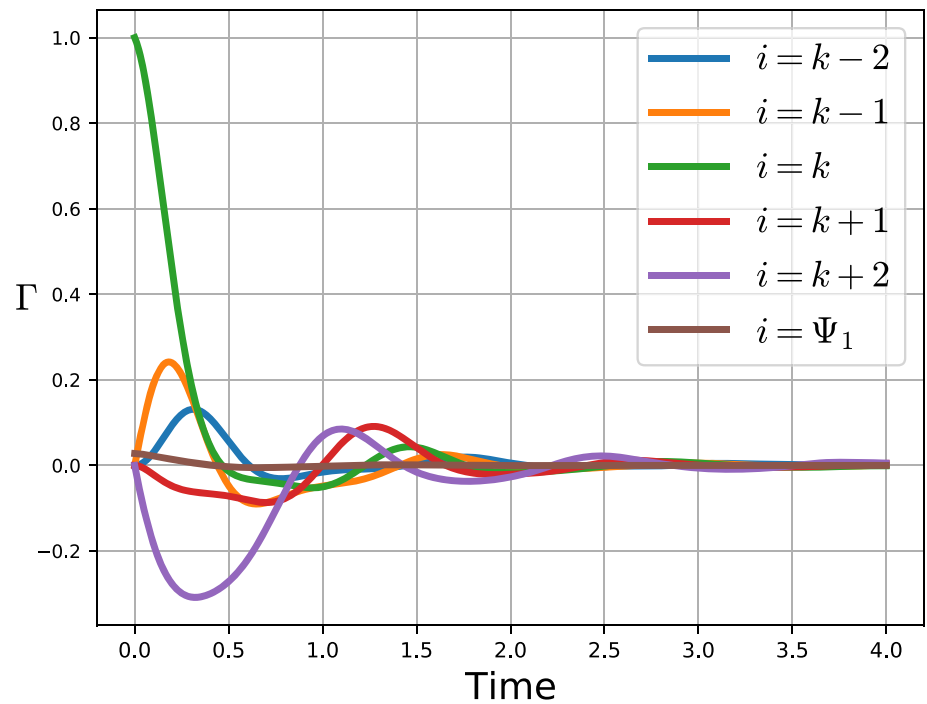

Fig. 2 Propagation of the perturbation for small times in the L96 model perturbed locally in x_k starting from a time t. The vertical lines are the dynamical variable x_k taken at different time instants, while the horizontal line below is the time axis. We have discretized the time in unit time steps for clarity purposes. At a given time, we have coloured the dynamical variables which feel directly the perturbation from the perturbed variables of the instant before with red. The colour loses its intensity as time goes by. We can notice that there are more red-coloured variables above the site x_k than below. This is consistent with the fact that the information propagates from the dynamical variables x_{k+j} with $j < 0$ to the dynamical variables x_{k+j} with $j > 0$, with a velocity given by the group velocity v_g, shown with a green arrow in the figure

dynamics of the atmosphere. It is defined on a lattice of N grid points and the evolution of the variables x_i, $i = 1, \ldots, N$ is given by the following system of ordinary differential equations:

$$\dot{x}_i = x_{i-1}(x_{i+1} - x_{i-2}) - \gamma x_i + F, \qquad (18)$$

where F is a constant controlling the external forcing, γ (usually set to a unitary value, as done also here) modulates the strength of the dissipation, and the quadratic term on the right hand side describes a non-standard advection process. The system obeys periodic boundary conditions, so that $x_{i-N} = x_i = x_{i+N}$ for all values of i. The qualitative properties of the dynamics of the L96 model changes dramatically accordingly to the values of F and N. In particular, it can be shown that, when $\gamma = 1$, the dynamics of the model is chaotic for $F \geq 5$ and the system becomes to a good approximation extensive as $N \geq 20$ [65]. One can show that travelling waves appear on top of the turbulent background in the chaotic regime, with phase velocity directed towards decreasing values of i [52]. Instead, the group velocity, which marks the direction of the propagation of information within the system, is directed in the opposite direction, towards increasing values i [66]. Detailed analyses of the properties of the L96 model can be found in [65, 67, 68], where the reader can find also an extensive bibliography. Recently, two extensions of the L96 model have been proposed, one able to accommodate for a complex interplay between dynamics and thermodynamics [66], and one featuring multiple competing attractors [69].
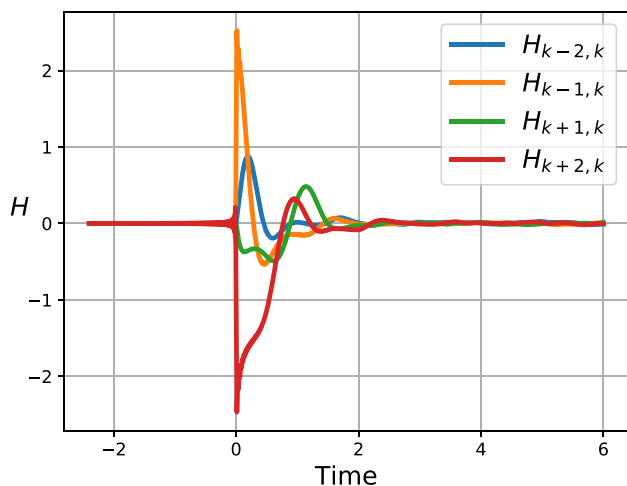
Fig. 3 Surrogate response functions $H_{i,k}$ for $i \in \{k-2, k-1, k, k+1, k+2\}$

3.2 Numerical simulations

3.2.1 Linear response functions

We will now apply the formalism of the surrogate LRT to the L96 dynamical system Eq. (18). We will set ourselves in the chaotic regime by choosing $N = 36$ and $F = 8$. In [49], the focus was on global perturbations, which impact directly all the variables x_i, $i = 1, \dots N$, and on global observables. We want now to take a different route, focusing on local perturbations and local observables, which had initially been proposed in [17]. In particular, we choose the following spatial pattern of applied forcing:

$$G_i(x) = \epsilon \delta_{ik}, \quad i = 1, \dots N \quad (19)$$

where δ_{ik} is the Kronecker delta, which has unitary value if the two indices are identical and vanishes otherwise, and ϵ is a real number which measures the magnitude of the perturbation. Equation (19) implies that we apply an extra forcing only at kth grid point. We will consider as observables of interest the dynamical variables x_j. With these choices of the perturbation and the observables, the problem we are addressing amounts to asking to what extent a perturbed variable at location i can predict the future state of another perturbed variable at location j after the system has been perturbed locally at location k. Given the discrete symmetry of the system, we expect that these properties will depend only on the relative position of i, j, and k, but, since the propagation of the information in this system is directional, we expect that they will not depend only on the distance between these locations.

Let $\Gamma_{i,k}$ be the response function of the perturbed variable x_i to the perturbation with spatial pattern Eq. (19), located in x_k. Along the lines of [49], we estimate $\Gamma_{i,k}$ by considering a probe with time pattern a Dirac's delta: $e(t) = \delta(t)$. We then run a long simulation with a random initial condition using adaptive

Runge–Kutta method of order 4 implemented through the Python function `solve_ivp`. We discard an initial transient of length T_{tr} and then we create an ensemble of M initial conditions $\{\bar{x}^{(k)}\}$, chosen each T_{gap} time units of the simulations. There is considerable flexibility in choice of T_{tr} and T_{gap}. T_{tr} depends on the chosen initial condition and should be much larger than the time it takes for the orbit to be come in close proximity of the attractor, while T_{gap} should be much longer than the inverse of the first Lyapunov exponent of the system to guarantee—at all practical levels—that the initial conditions are virtually independent. Our first estimate of the response function $\Gamma_{i,k}^+$ is obtained by averaging over such an ensemble of M initial conditions the quantities $\delta x_i^{(k)}(t)$:

$$\Gamma_{i,k}^+(t) = \frac{1}{\epsilon M} \sum_{k=1}^{M} \delta x_i^{(k)}(t), \quad (20)$$

where $\delta x_i^{(k)}(t)$ is the difference between the value of the variable x_i at time t in the perturbed and unperturbed run, both having the same initial condition $\bar{x}^{(k)}$. We then repeat the same procedure by switching the sign of the forcing: $\epsilon \rightarrow -\epsilon$, obtaining $\Gamma_{i,k}^-(t)$. Our estimate of the response function is then given by

$$\Gamma_{i,k}(t) = \frac{\Gamma_{i,k}^+(t) + \Gamma_{i,k}^-(t)}{2}, \quad (21)$$

The last step allows to remove the second-order correction to the linear response and considerably increases the precision of the estimate [70]. The linearity of the responses has been tested to hold very well up to $\epsilon = 1$, see Appendix B. We choose $\epsilon = 1$ for our numerical studies in order to have a good signal-to-noise ratio while being within the regime of linear response. We remark that the procedure above ensures that the response function is causal.

The response functions $\Gamma_{i,k}$ for $i = \{k-2, k-1, k, k+1, k+2\}$ obtained for $M = 2 \cdot 10^6$ and $\epsilon = 1$ are shown in Fig. 1. We focus on the dynamical variables nearby the perturbation site, because the intensity of the response decreases dramatically as we move further away, as discussed below.

3.2.2 Hierarchy in predictive power

Propagation of the perturbation signal

We now look at the leading order term of short-time behaviour of $\Gamma_{i,k}$ for positive times ($\Gamma_{i,k}$ vanishes for negative times):

$$\Gamma_{i,k}(t) \underset{t \rightarrow 0^+}{\approx}$$

$$\begin{cases} 1 + \mathcal{O}(t), & i = k \\ \langle C_i^{(1)} \rangle_0 + \mathcal{O}(t^2). & i \in \{k-1, k+2\} \\ \langle C_i^{(q)} \rangle_0 t^q + \mathcal{O}(t^{q+1}), & i \in \{k-q, k+2q-3, k+2q\}, q \geq 2, \end{cases}$$

Fig. 4 **a** Non-singular part of the surrogate response function $K_{k+2,k-1}$. **b** Surrogate response functions $H_{k-2,k-1}$ and $H_{k+1,k-1}$
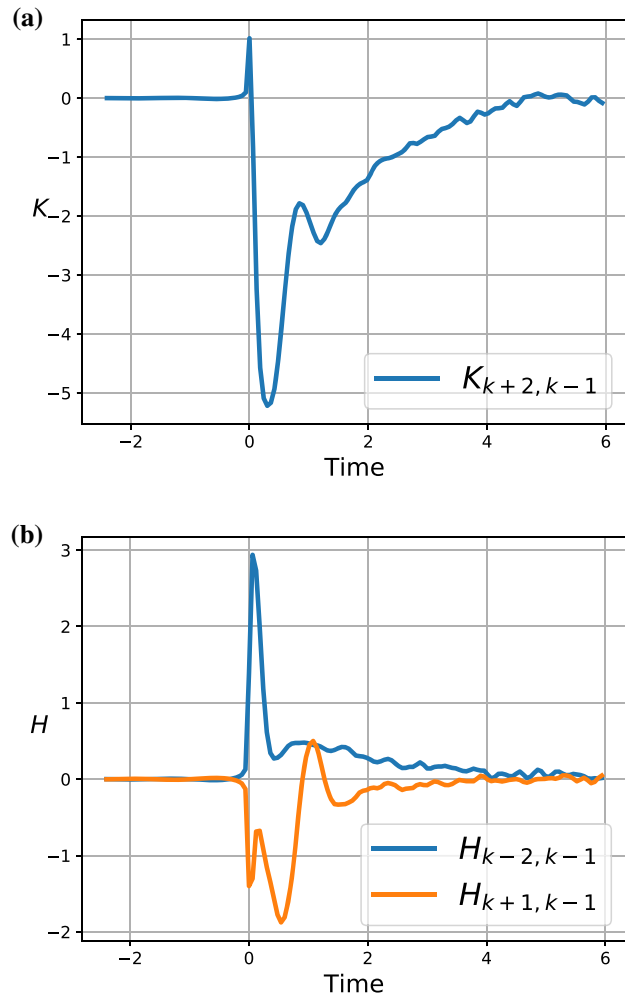
where the coefficient $\langle C_i^{(q)} \rangle_0$ is the expectation value of the function $C_i^{(q)}$ taken over the stationary measure ρ_0, see Appendix C. In particular, we notice that at $t = 0$ the response function $\Gamma_{i,k}$ is equal to 1 if $i = k$ and it is vanishing for $i \neq k$. This is intuitive: at $t = 0$ the perturbation is felt in all its intensity just at the grid point that has been directly perturbed. As t increases, the perturbation propagates also to the other grid points x_i, with a time scale determined by the leading order term given in Eq. (22). Note that the perturbation propagates more efficiently towards the right $(i > k)$ than towards left $(i < k)$, since for each dynamical variable x_{k-q} at the left of x_k with leading term t^q there are two dynamical variables x_{k+2q-3} and x_{k+2q} (for $q > 1$) at its right with the same leading term. We can have a clearer intuition of the actual propagation of the signal in space and in time by looking at the cartoon in Fig. 2. It is remarkable that the site $k + 3$ or $k + 5$ are affected by the forcing later than the sites $k + 4$ or $k + 6$, respectively, in agreement with the well-known fact that advection in the L96 system is non-standard. This further indicates that the advection taking place in the L96 system is a non-standard one.

(22)

The high-frequency asymptotic behaviour of the susceptibilities corresponding to the response functions given in Eq. (22) can be derived using Eq. (9).

Hierarchy of the dynamical variables

We want now to build a hierarchy of the dynamical variables in terms of their predictive power of the other dynamical variables. In particular, this hierarchy is closely related to how the perturbation signal propagates in the system: the sooner a dynamical variables feels the perturbation and the more predictive it will be. We construct the surrogate response function using Eq. (8) and by then applying the cutoff introduced in Eq. (12). We define as $H_{i,j}$ the surrogate response function that allows one to reconstruct the variable i using the variable j as surrogate forcing. We indicate with $K_{i,j}$ $(S_{i,j})$ its non-singular (singular) component, and with $H_{i,j}^c$ its causal component,

The variable with the highest predictive power is obviously x_k, since it responds immediately to the perturbation applied at the site k. Indeed, the non-causal component of the surrogate response functions shown in Fig. 3 is entirely negligible. The second most predictive variables are x_{k-1} and x_{k+2}. A few surrogate

Fig. 5 **a** Non-singular part of the surrogate response function $K_{k-1,k+2}$. **b** Surrogate response functions $H_{k-2,k+2}$ and $H_{k+1,k+2}$
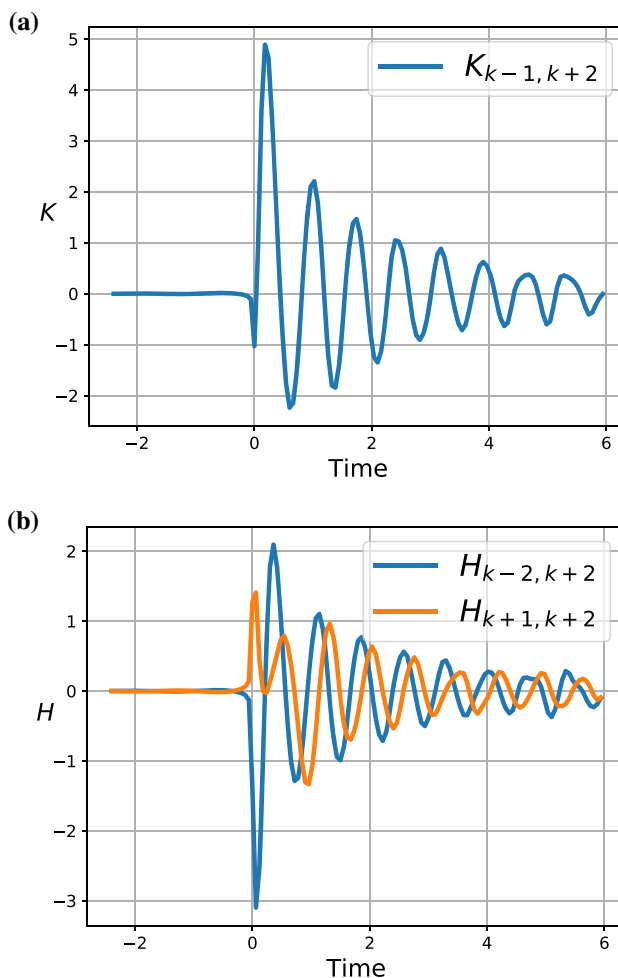
response functions employing them as predictors are shown in Figs. 4 and 5. We can observe that these surrogate response functions show a rather small non-causal component. Notice that just the non-singular components Eq. (11) are shown in the figures, for clarity purposes. The third most predictive variables between the ones we have considered are x_{k-2} and x_{k+1}. In Fig. 6, we show the non-singular parts of the related surrogate response functions $K_{k-2,k+1}$ and $K_{k+1,k-2}$. We can notice that their non-causal components are more relevant than the ones considered before, because the information retained by the predictors is degraded.

We can better quantify the importance of the non causal component of the surrogate response functions through the PI introduced in Eq. (15). The results are presented in Table 1 for the surrogate response functions between the dynamical variables x_{k-2}, x_{k-1}, x_{k+1} and x_{k+2}. Looking at the values provided by the PI, we observe that the weight of the non causal part is bigger when the predictors are x_{k-2} and x_{k+1}. This is in agreement with the hierarchy described above.

Another aspect we wish to analyse is whether one can define a preferential direction for performing the prediction. If we take two variables x_i and x_j, we can ask ourselves whether it is better to use x_i to predict x_j or the other way around. Of course, this issue makes sense

Table 1 PI for the surrogate response functions between the dynamical variables x_{k-2}, x_{k-1}, x_{k+1} and x_{k+2}

	x_{k-2}	x_{k-1}	x_{k+1}	x_{k+2}
x_{k-2}	.	.	0.083	.
x_{k-1}	0.013	.	0.019	0.0022
x_{k+1}	0.58	.	.	.
x_{k+2}	0.0068	0.0033	0.014	.

In the first column there are the predictors, while in the first row there are the predictands

in the case the two variables x_i and x_j have the same rank (otherwise we would just use the higher ranked variable as a predictor). This is the case of x_{k-1} and x_{k+2} and of x_{k-2} and x_{k+1}. Making use of Table 1, we see that in both cases the variable with lower index (more to the left) is the better predictor. This is consistent with the fact that the group velocity v_g of the travelling waves in L96, which controls the flow of information, is positive (from left to right): the variables that are situated upstream predict better than those situated downstream.

Now we test the actual predictive ability of the surrogate response function computed above. We perturb

Fig. 6 **a** Non-singular part of the surrogate response function $K_{k+1,k-2}$. **b** Non-singular part of the surrogate response function $K_{k-2,k+1}$

(a)

(b)
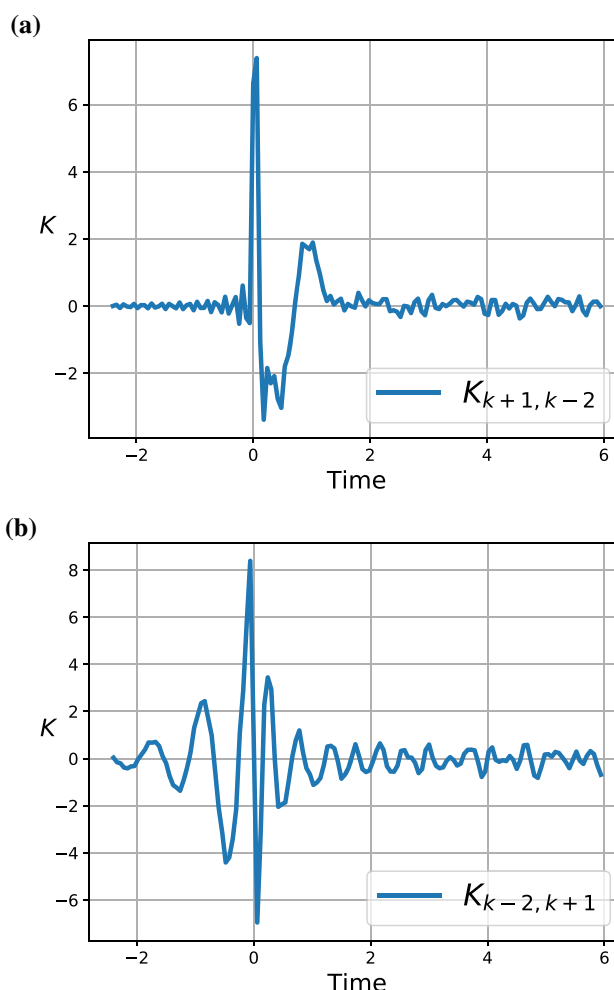
the L96 system with the vector field having spatial pattern $G(x) = \epsilon \delta_{i,k}$ as above and having the following time pattern:

$$e(t) = \Theta(t) - \Theta(t - \tau), \qquad (23)$$

with $\tau = 5$. This time pattern seems relevant because corresponds to the act of switching abruptly on and off the perturbation, and keeping it active for a time scale that is longer than the inverse of the first Lyapunov exponent of the system (about 0.6 time units). Therefore, it allows to appreciate both transient and longer term features of the response of the system. We then test the skill of the variable x_k in predicting x_j, with $j \in \{k-2, k-1, k+1, k+2\}$:

$$\delta\langle x_i\rangle(t) = \int_{-\infty}^{\infty} d\tau \, H_{ij}^c(t-\tau)\delta\langle x_j\rangle(\tau), \qquad (24)$$

where we use only the causal component of the surrogate response function. The predictions are shown in Fig. 7, where we can notice that the agreement between the actual response and the prediction made through the surrogate response functions, where x_k is the predictor is very good in all cases.

We now consider the second most predictive dynamical variables x_{k-1} and x_{k+2}. The predictions are shown in Fig. 8. We can notice that these predictors work quite well: despite not being directly perturbed by the forcing, they retain a lot of information. We also remark that some discrepancy between prediction and the actual response emerges in terms of mismatch of the oscillations taking place during the plateau of the forcing.

Finally, we take into account the predictions made using the variables x_{k-2} and x_{k+1}. As we can see in Fig. 9, the predictions are clearly less successful than in previous cases, even though they show a qualitative agreement with the actual responses. This is explained by the greater relevance of the non-causal components of the related surrogate response functions $H_{k-2,k+1}$ and $H_{k+1,k-2}$, as it can be seen in Fig. 6; see also Table 1. Remarkably, be comparing the two panels of Fig. 9, we can clearly see the asymmetry between the mutual predictive power of x_{k-2} (better predictor) and x_{k+1} (worse predictor) discussed above.

3.3 Making predictions with more observables

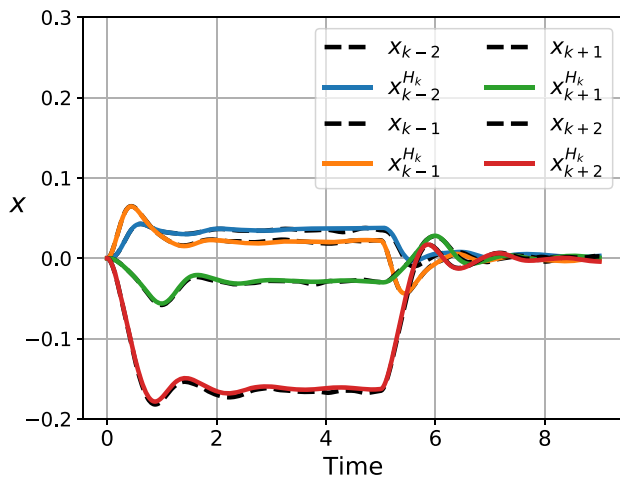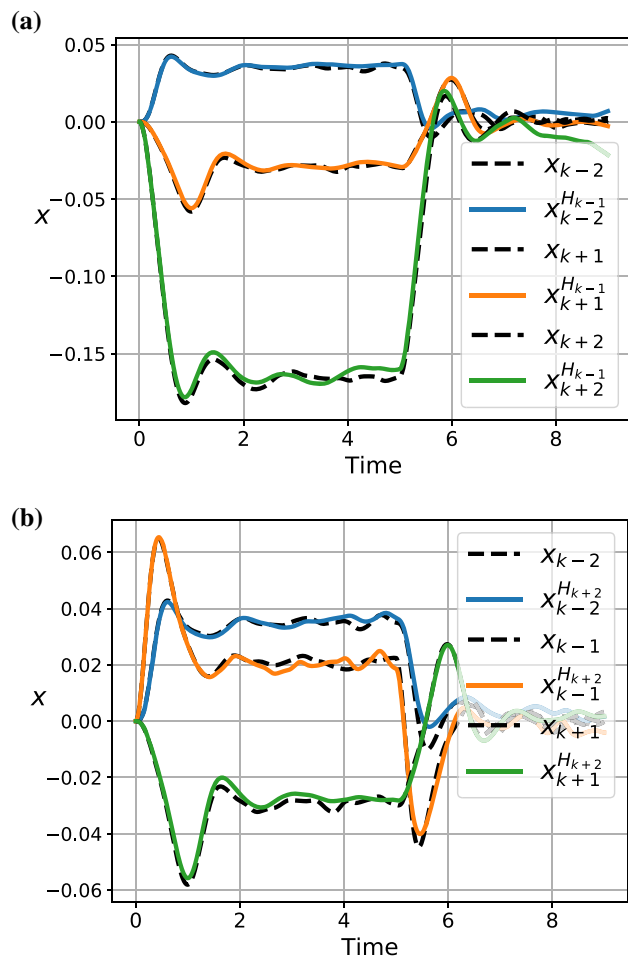We have shown above that some local variables cannot well predict other local variables, as a result of how

Fig. 7 Comparison between the response of the perturbed L96 system to perturbation with spatial pattern Eq. (19) and time pattern Eq. (23) with $\tau = 5$ and the predictions made using the surrogate response functions $H_{i,k}$ for $i \in \{k-2, k-1, k+1, k+2\}$

the perturbation signal propagates across the system. Following the discussion in Sect. 2, we test whether this can be improved by applying two independent forcings

and, correspondingly, using two predictors. The idea is that by adding a forcing and a predictor we are able to learn more about the properties of the system and of its response. We then add, on top of the perturbation described by Eq. (19) with time pattern given by Eq. (23), a second forcing that impacts the viscosity of the system acting at the kth grid point:

$$G_i^{(2)}(x) = -x_i\, \delta_{ik} \epsilon_2. \tag{25}$$

The forcing is applied using the temporal pattern

$$e^{(2)}(t) = \Theta(t) - \Theta(t - \tau_2),$$

where $\tau_2 = 3$.

The corresponding response functions $\Gamma_{i,k}^{(2)}$ for $i = \{k-2, k-1, k, k+1, k+2\}$ obtained for $M = 2 \cdot 10^6$ and $\epsilon = 0.1$ are shown in Fig. 10. We have tested the linearity of the response of the system for values of ϵ_2 ranging from 0.01 up to 2 and found that nonlinear corrections are negligible for $\epsilon_2 \leq 0.25$; see Appendix B. Note that here we need to consider smaller values of ϵ_2 compared to the case of the forcing $G^{(1)}$ because of the presence of the factor x_i ($|x_i|$ is typically larger than one in the unperturbed runs). We then perform

Fig. 8 Same as Fig. 7, but looking at predictions performed using x_{k-1} in **a** and with x_{k+2} in **b**

(a)

(b)

Fig. 9 Same as Fig. 7, but looking at predictions performed using x_{k-2} in **a** and with x_{k+1} in **b**

(a)

(b)
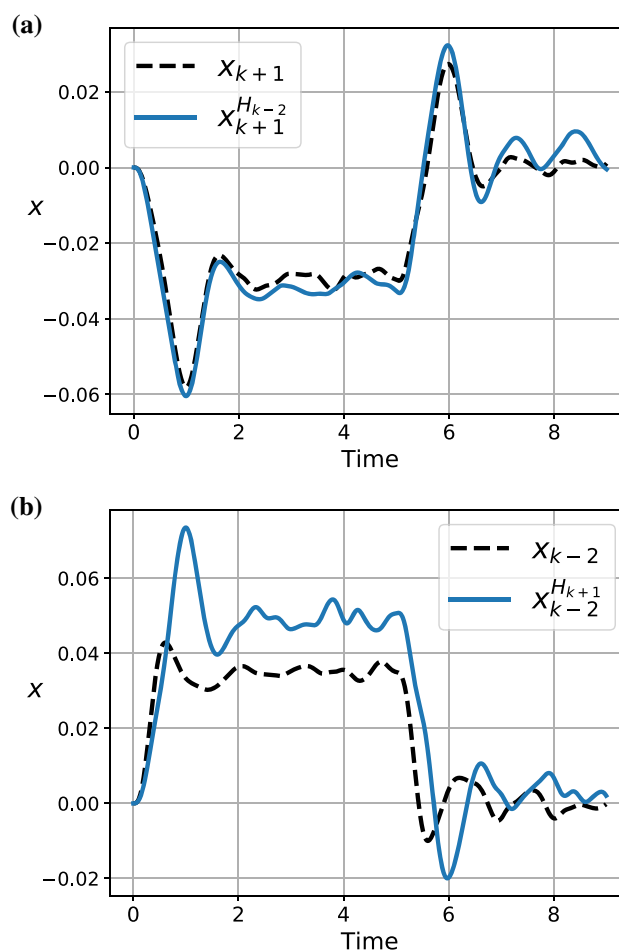
the data analysis for the case of the combined forcings $G^{(1)}$ and $G^{(2)}$ using $\epsilon_2 = 0.1$.

As shown in Eq. (5) for $N = 2$, it is then necessary to choose a second observable as predictor. We choose the following global observable:

$$\Psi_1(t) \equiv \frac{1}{N} \sum_{i=1}^{N} x_i. \qquad (26)$$

Note that Ψ_1 is usually interpreted as the mean momentum of the system [16]. The corresponding response function $\Gamma_{\Psi_1,k}(t)$ ($\Gamma^{(2)}_{\Psi_1,k}(t)$) to the forcing defined in Eq. (19) (Eq. 25) is portrayed in Fig. 1 (Fig. 10). By symmetry, $\Gamma_{\Psi_1,k}(t)$ and $\Gamma^{(2)}_{\Psi_1,k}(t)$ do not depend on k. This choice is motivated by the fact that we wish to simulate the situation, where a local observer, e.g., x_{k+1}, uses information on its own local state and on global properties of the system to predict the state of another local observer, e.g., x_{k-2}. We perform the prediction using the following formula:

$$\delta \langle x_{k-2} \rangle(t) = H^c_{k-2,\Psi_1}(t) * \delta \langle \Psi_1 \rangle(t)$$
$$+ H^c_{k-2,k+1}(t) * \delta \langle x_{k+1} \rangle(t). \qquad (27)$$

These surrogate response functions have in general different poles than the ones previously defined in the case

of just one forcing, see discussion in Appendix D. In Fig. 11, the surrogate response function H_{k-1,Ψ_1} and the non-singular component $K_{k-2,k+1}$ of the surrogate response function $H_{k-2,k+1}$ are shown. By comparing Figs. 6b and 11b, we note that the non-causal component of the response function associated to x_{k-2} is greatly reduced once a second forcing and a second observable are used. In addition, we also note that the amplitude of the surrogate response function associated to x_{k-2} is greatly reduced, implying that most of the information on x_{k-2} is drawn from Ψ_1 in the case analysed here. The improvement in our ability to predict x_{k-2} is summarised in Table 2. By comparing the dashed lines in Figs. 12 and 9, one notices that adding the second forcing given in Eq. (25) has little effect on the actual response of x_{k-2}; indeed, the contribution to the response is smaller by a factor $\mathcal{O}(10^{-2})$ (not shown) with respect to what coming from the forcing given in Eq. (19). But, instead, our ability to predict using surrogate response functions increases substantially when we add Ψ_1 as predictor, even if such observable has little information on the local properties of the system. This is due to the fact that adding a second perturbation to the system and a second observable as predictor regularises our problem. Indeed, a greater predictive skill is obtained even if we perform simulations with smaller values of ϵ_2 than what reported above (not shown).
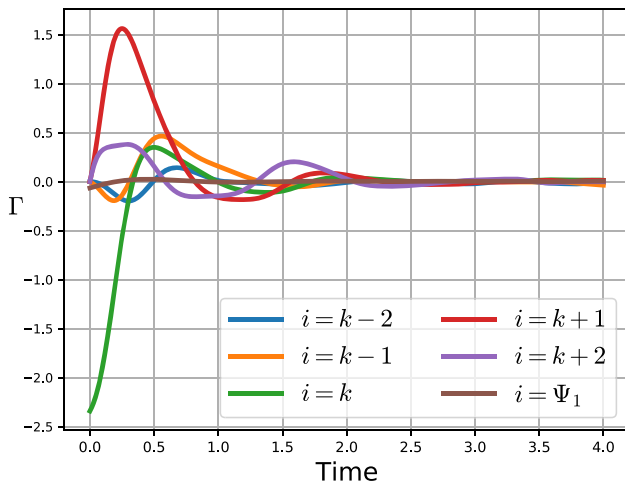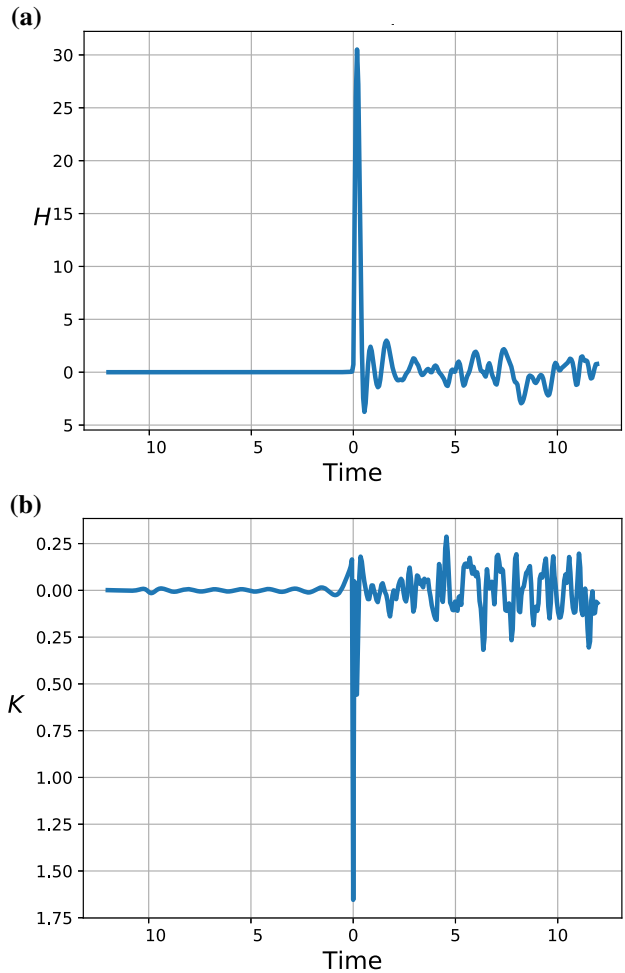
Fig. 10 Plots of the response functions $\Gamma_{i,k}^{(2)}$ for $i = \{k - 2, k - 1, k, k + 1, k + 2\}$ and $\epsilon = 0.1$. We also portray $\Gamma_{\Psi_1,k}^{(2)}$, the response function for the mean momentum Ψ_1 (Eq. 26). See also Fig. 15

4 Conclusions

In this paper we have explored the possibility of using, in the linear regime, an observable of a perturbed system as predictor of the response of another observable of the same system perturbed by a forcing. While specific conditions need to be met to be able to perform an actual prediction, it is always possible to reconstruct a posteriori the desired signal. The procedure requires gathering first some information on the relationship between the response of the two observables . Such a knowledge can be obtained by looking at the linear response of the two observables to the forcing undergoing a broadband temporal modulation (e.g., a kick in the form of a Dirac's delta). Then, the approach allows to use the response of one observable to reconstruct and, when possible, predict, the response of the second observable for any temporal pattern of modulation of the forcing. Hence, the first observable is used as a surrogate for the external forcing. The theory clarifies that the ability of two observable to predict each other is not the same, and allows for the treatment of N independent forcings to the system.

This approach can be very useful in the case we are facing an inverse problem, where we have limited information on the system and in particular on the acting

Fig. 11 **a** Surrogate response function H_{k-2,Ψ_1}. **b** Non-singular part of the surrogate response function $K_{k-2,k+1}$

(a)

(b)

Table 2 PI for the surrogate response function $H_{k-2,k+1}$ when one forcing $G^{(1)}$ is used (left) and PI for the surrogate response functions H_{k-2,Ψ_1} and H_{k-2,Ψ_1} when two forcings $G^{(1)}$ and $G^{(2)}$ are applied
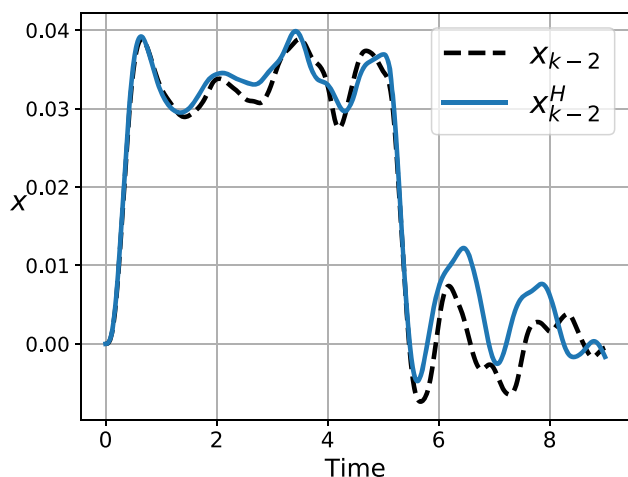
$G^{(1)}$	$G^{(1)}$ and $G^{(2)}$
0.58	0.0063

Fig. 12 Response of x_{k-2} to two acting forcings (dashed line) and prediction obtained using x_{k+1} and Ψ_1 as predictors

external forcing, while we can observe multiple features of the system at the same time and we want to be able to use internal feedbacks as surrogate forcings. This general viewpoint is closely linked to the vast class of problems associated with finding causal links in complex systems and it is of potential interest in many research area dealing with complexity, such as neurosciences and geosciences. Note that the very science of paleoclimatic reconstruction and the definition of proxy data implicitly uses some of the ideas presented here. Another area of applications of surrogate response theory is the analysis of the response of spatially extended system to perturbations, which has been the focus on this contribution.

We have focused on the interplay between applying localised forcing and observing the system at specific locations in the vicinity of where the forcing is applied using the L96 model as benchmark system. We have shown that, closely following the way signals propagate in the L96 model, one can establish a hierarchy of observables in terms of their ability to predict each other, where higher ranking observables are characterised by being impacted earlier by the applied perturbation. Such a hierarchy can be analytically motivated by looking at the asymptotic properties of suitably defined response functions. The prognostic ability of an observable with respect to a predictand can be quantified by evaluating the relative weight of the causal vs. of the non-causal components of the corresponding surrogate response function. Indeed, one can also verify that, one considers two observables that are

impacted by the applied forcing with the same time delay, it is in general true that the ability to predict each other is not symmetric.

The presence of such an asymmetry is associated with the group velocity on travelling waves in the system: variables with lower index predict better variables with higher index than vice versa. Finally, we have shown that implementing a more general form of surrogate response theory, where two independent forcings are applied and two observables are used as predictors, improves the quality of the prediction at local level even if the second observable used as predictor is a global one, because we are able to regularise the inverse problem addressed in this work compared to the simpler setting, where one forcing is applied and one observable is used as predictor. Our results are suggestive of the possibility of using the partial information gathered from the response of some observables in spatially extended system to reconstruct and predict the response for other observables when it is hard or impossible to know the exact form of the forcings.

Appendix A: Predictability Index: a simple example

We now consider a simple example to show the main idea behind the effectiveness of the PI introduced in Eq. (15). We consider the following idealised surrogate response function:

$$H(t) = \Theta(t)A\,e^{-at} + \Theta(-t)B\,e^{bt} + C\delta(t), \qquad (A1)$$

where the coefficients a and b are positive, while the constants A, B and C are real numbers. The non-causal and casual components of the non-singular part are respectively
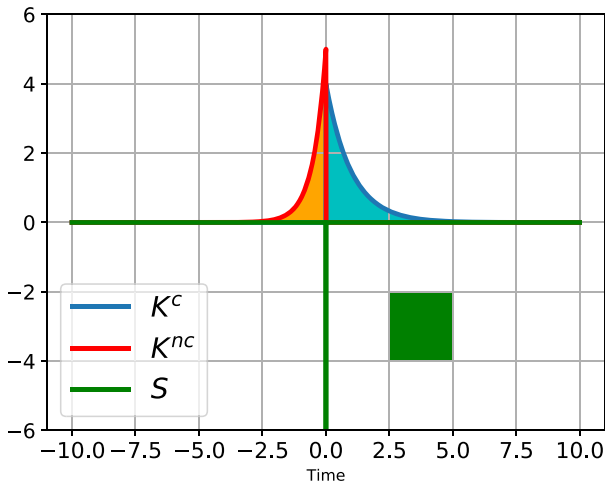
Fig. 13 Causal (K^c) and non-causal (K^{nc}) components of the non-singular part and singular part S of the surrogate response function given in Eq. (A1). The areas under the curves correspond to the various contributions appearing in Eq. (A2). The contribution coming from the singular term is represented as a rectangle. We have used $A = 4$, $B = 5$, $C = -5$, $a = 1$ and $b = 2$; in this case, $R = 5/18$

$K^{nc}(t) = \Theta(-t)B\,e^{bt}$ and $K^c(t) = \Theta(t)A\,e^{-at}$, while the singular part is $S(t) = C\delta(t)$. Note that the time scale related to the causal component is $1/a$, while the time scale related to the non-causal component is $1/b$. The smaller is the coefficient b, the larger is the time scale and the importance of the related component in the response function. The PI for the surrogate response function given in Eq. (A1) is

$$R(H) = \frac{\|K^{nc}(t)\|_1}{\|K^c(t)\|_1 + \|S(t)\|_1} \qquad (A2)$$

where $\|K^{nc}(t)\|_1 = |B|/b$, $\|K^c(t)\|_1 = |A|/a$, and $\|S(t)\|_1 = |C|$. Clearly, its value depends on the interplay of the time scales of the causal ($1/a$) and non-causal ($1/b$) components and of the magnitude of the three components A, B, and C; see Fig. 13 for a graphical representation of the contributions of the various terms.

Appendix B: Test of the linearity of the response in the Lorenz '96 model

When performing an analysis based on LRT, it is clearly essential to make sure that we are indeed studying the system in a regime where higher order terms in the perturbative expansion accounting for the full response are negligible.

Figure 14 portrays the estimates for various response functions $\Gamma_{i,j}(t)$ obtained using a range of values for the intensity ϵ of the perturbation applied according to Eq. (19). In all cases we have followed the procedure described in Sect. 3.2. The $\epsilon = 1$ curves are also reported in Fig. 1. Figure 15 provides the corresponding information on the range of linearity of the response of the system to the perturbation applied according to Eq. (25).

Appendix C: Asymptotic properties of the response in the Lorenz '96 model

Following [17], it is possible to obtain the behaviour of the response functions $\Gamma_{i,k}$ for $t \to 0^+$ in the L96 system as follows. Let us consider the response function $\Gamma_{i,k}$ in our case, with space pattern $G(x) = \epsilon\delta_{ik}$ and perturbed observable x_i:

$$\begin{aligned} \Gamma_{i,k} &= \Theta(t)\int \rho_0(\mathrm{d}x)\sum_l G\,\delta_{kl}\partial_l(x_i(t)) \\ &= \Theta(t)\int \rho_0(\mathrm{d}x)\,\partial_k(x_i(t)), \end{aligned} \qquad (C1)$$

where ∂_k stands for the derivation with respect to $x_k(0)$ and ρ_0 is the steady-state distribution over the initial condition $x(0)$ from which the trajectory $x(t)$ starts. We now perform a Taylor expansion of $x_i(t)$ at $t = 0^+$ and we exploit the evolution equation Eq. (18):

$$\begin{aligned} \partial_k(x_i(t)) \approx\ & \delta_{i,k} + t(C^{(1)}_{k-1}\delta_{i,k-1} + C^{(1)}_{k+1}\delta_{i,k+1} \\ & + C^{(1)}_{k+2}\delta_{i,k+2} - \delta_{ik}) + \mathcal{O}(t), \end{aligned} \qquad (C2)$$

where $C^{(a)}_p$ is the coefficient related to the dynamical variable p of the term of order a in the expansion above (we have that $C^{(0)}_k = \delta_{ik}$).

The coefficients for the linear terms ($k = 1$) give

$$\begin{aligned} C^{(1)}_{k-1} &= x_{k-2}(0) \\ C^{(1)}_{k+1} &= (x_{k+2}(0) - x_{k-1}(0)) \\ C^{(1)}_{k+2} &= -x_{k+1}(0). \end{aligned} \qquad (C3)$$

Considering that $\int \rho_0(dx)\,x_i = \int \rho_0(dx)\,x_j\ \forall i, j = 1, \ldots, N$ because of the discrete symmetry of the system in the unperturbed state, we have the linear term in the expansion of $\Gamma_{k+1,k}$ vanishes.

We can repeat the same derivation for all the grid points of the system, taking into account that the leading term of $\Gamma_{k+2q-1,k}$ is t^{q+1} instead of t^q. We obtain

$$\begin{aligned} \Gamma_{i,k}(t) \underset{t\to 0^+}{\approx}\ & \\ \times & \begin{cases} 1 + \mathcal{O}(t), & i = k \\ \langle C^{(1)}_i\rangle_0 t + \mathcal{O}(t^2). & i \in \{k-1, k+2\} \\ \langle C^{(q)}_i\rangle_0 t^q + \mathcal{O}(t^{q+1}), & i \in \{k-q, k+2q-3, k+2q\},\, q \geq 2, \end{cases} \end{aligned} \qquad (C4)$$

where the averages are over the stationary measure ρ_0.

Appendix D: Singular components of the surrogate response

1. Case of one forcing
We consider the perturbation with spatial pattern given by Eq. (19) and time pattern Eq. (23). The $\omega \to \infty$ asymptotic behaviour of $J_{k-1,k+2}(\omega)$ is given by

$$\lim_{\omega\to\infty} J_{k-1,k+2}(\omega) = \frac{\lim\limits_{\omega\to\infty}\chi_{k-1,k}(\omega)}{\lim\limits_{\omega\to\infty}\chi_{k+2,k}(\omega)} = \frac{\langle C^{(1)}_{k-1}\rangle_0}{\langle C^{(1)}_{k+2}\rangle_0} = -1 \qquad (D1)$$
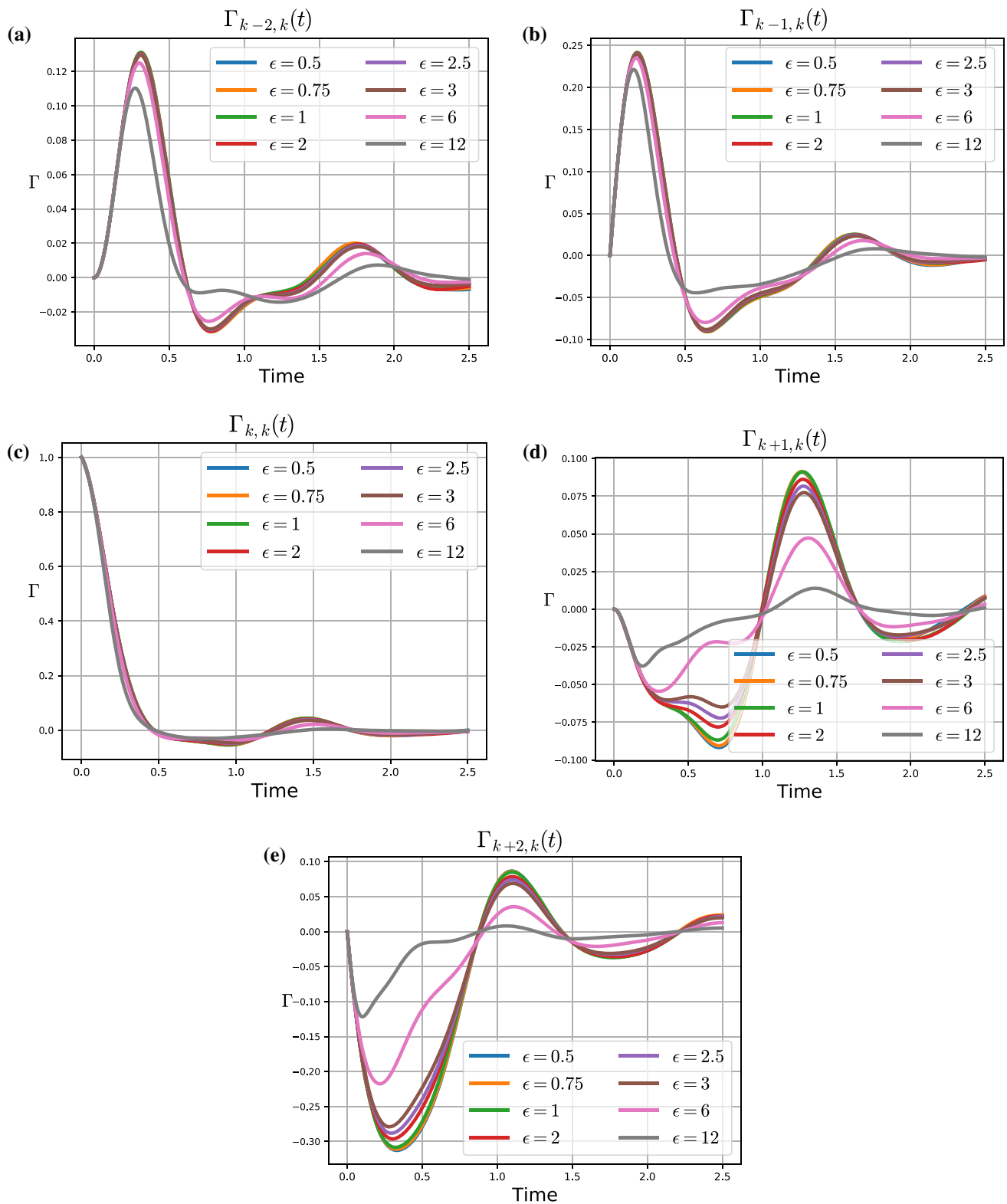
Fig. 14 Testing the linearity of the response of the system to the perturbation given in Eq. (19). Different estimates for the response functions have been obtained using different values of ϵ
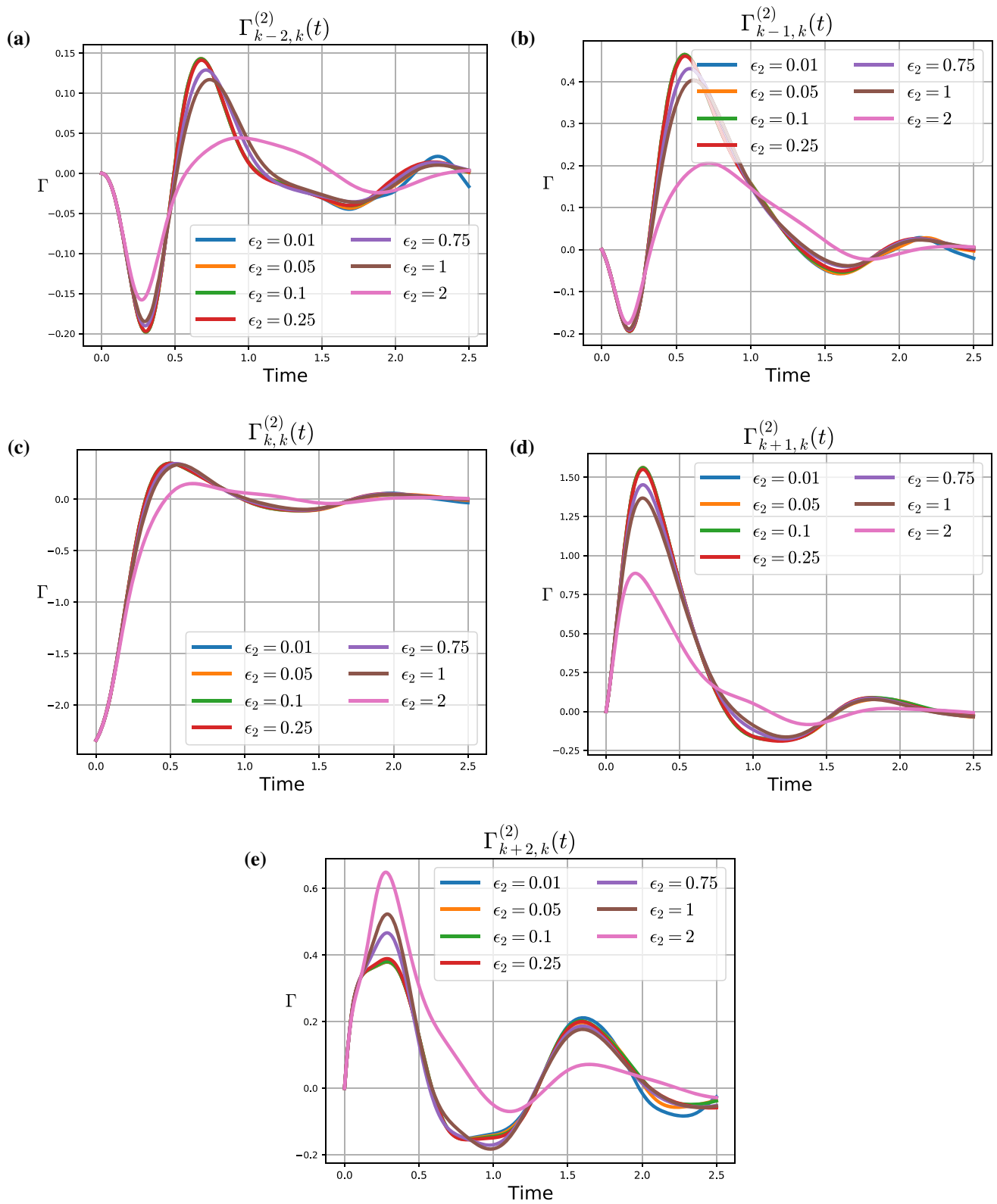
Fig. 15 Same as Fig. 14, but for the perturbation given in Eq. (25)

Table 3 Limit behaviour for $\omega \to \infty$ of the Fourier transforms of the response functions for the observable Ψ_1, x_{k-2} and x_{k+1}, when two forcings $G^{(1)}$ and $G^{(2)}$ are applied

	$\omega \to \infty$
$\chi_{\Psi_1, G^{(1)}}$	$(iG_1/(N\omega))$
$\chi_{k+1, G^{(1)}}$	$-2iG_1 \langle C^{(2)}_{k+1} \rangle_0 / \omega^3$
$\chi_{k-2, G^{(1)}}$	$-2iG_1 \langle C^{(2)}_{k-2} \rangle_0 / \omega^3$
$\chi_{\Psi_1, G^{(2)}}$	$-(iG_2/(N\omega)) \cdot \langle x \rangle_0$
$\chi_{k+1, G^{(2)}}$	$2iG_2 \langle D^{(2)}_{k+1} \rangle_0 / \omega^3$
$\chi_{k-2, G^{(2)}}$	$2iG_2 \langle D^{(2)}_{k-2} \rangle_0 / \omega^3$

where we have used Eqs. (C3) and (C4). As a consequence, the non-singular component of the surrogate response function $K_{k-1,k+2}$ is obtained by performing the inverse Fourier transform of the function $J_{k-1,k+2}(\omega) + 1$.

We then consider the surrogate response functions $H_{k-2,k+1}$ and $H_{k+1,k-2}$. We proceed as above. We look at the asymptotic behaviour of the corresponding Fourier transforms:

$$\lim_{\omega \to \infty} J_{k-2,k+1}(\omega) = \frac{\lim_{\omega \to \infty} \chi_{k-2,k}(\omega)}{\lim_{\omega \to \infty} \chi_{k+1,k}(\omega)} = \frac{\langle C^{(2)}_{k-2} \rangle_0}{\langle C^{(2)}_{k+1} \rangle_0}, \quad \text{(D2)}$$

where $\langle C^{(2)}_{k-2} \rangle_0$ and $\langle C^{(2)}_{k+1} \rangle_0$ can be computed directly from short-time behaviour of the response functions $\Gamma_{k-2,k}$ and $\Gamma_{k+1,k}$:

$$\langle C^{(2)}_{k-2} \rangle_0 = \langle x_{k-3} x_{k-2} \rangle_0$$
$$\langle C^{(2)}_{k+1} \rangle_0 = -\langle x_{k-3} x_{k-2} \rangle_0 - \langle x_{k-2} x_k \rangle_0. \quad \text{(D3)}$$

We derive the non-singular components $K_{k-2,k+1}$ by performing the inverse Fourier transform of $J_{k-2,k+1}(\omega) - \langle C^{(2)}_{k-2} \rangle_0 / \langle C^{(2)}_{k+1} \rangle_0$. Using the statistics of the unperturbed L96 model, we obtain $\langle C^{(2)}_{k-2} \rangle_0 / \langle C^{(2)}_{k+1} \rangle_0 \approx -0.90$.

2. Case of two forcings

We now want to investigate the possible presence of singular components in the surrogate response functions H_{k-2,Ψ_1} and $H_{k-2,k+1}$, in the setting with the L96 system Eq. (18) perturbed with the two forcings introduced in Sect. 3.3. The explicit expression of their Fourier transform is given by Eq. (6) and it is the following:

$$J_{k-2,\Psi_1}(\omega)$$
$$= \frac{\chi_{k+1,G^{(2)}}(\omega)\chi_{k-2,G^{(1)}}(\omega) - \chi_{k+1,G^{(1)}}(\omega)\chi_{k-2,G^{(2)}}(\omega)}{\chi_{k+1,G^{(2)}}(\omega)\chi_{\Psi_1,G^{(1)}}(\omega) - \chi_{\Psi_1,G^{(2)}}(\omega)\chi_{k+1,G^{(1)}}(\omega)}$$
$$J_{k-2,k+1}(\omega)$$
$$= \frac{-\chi_{\Psi_1,G^{(2)}}(\omega)\chi_{k-2,G^{(1)}}(\omega) + \chi_{\Psi_1,G^{(1)}}(\omega)\chi_{k-2,G^{(2)}}(\omega)}{\chi_{k+1,G^{(2)}}(\omega)\chi_{\Psi_1,G^{(1)}}(\omega) - \chi_{\Psi_1,G^{(2)}}(\omega)\chi_{k+1,G^{(1)}}(\omega)}$$
$$\text{(D4)}$$

To do that, we take the limit for $\omega \to \infty$ of the relations Eq. (D4) and we plug the limits obtained for the response functions which appear in these relations, listed in Table 3. As a consequence, the limit behaviour of Eq. (D4) is the following:

$$\lim_{\omega \to \infty} J_{k-2,\Psi_1}(\omega) = 0$$
$$\lim_{\omega \to \infty} J_{k-2,k+1}(\omega) = s_{k-2,k+1}, \quad \text{(D5)}$$

where

$$s_{k-2,k+1} = \frac{-\langle x \rangle_0 \langle C^{(2)}_{k-2} \rangle_0 + \langle D^{(2)}_{k-2} \rangle_0}{\langle D^{(2)}_{k+1} \rangle_0 - \langle x \rangle_0 \langle C^{(2)}_{k+1} \rangle_0}, \quad \text{(D6)}$$

where $\langle C^{(2)}_{k-2} \rangle_0$ and $\langle C^{(2)}_{k+1} \rangle_0$ are given by Eq. (D3). Instead, $\langle D^{(2)}_{k-2} \rangle_0$ and $\langle D^{(2)}_{k+1} \rangle_0$ can be derived by looking at the $t \to 0^+$ behaviour of $\Gamma_{k+1,G^{(2)}}$ and $\Gamma_{k-2,G^{(2)}}$, or, equivalently, at the $\omega \to \infty$ behaviour of their Fourier transforms:

$$\langle D^{(2)}_{k-2} \rangle_0 = \langle x_{k-3} x_{k-2} x_k \rangle_0$$
$$\langle D^{(2)}_{k+1} \rangle_0 = -3\langle x_{k-2} x_k \rangle_0 + 3\langle x_{k-1} x_k \rangle_0 + 2\langle x_k x_{k+1} x_{k+3} \rangle_0$$
$$\quad - 2\langle x_k^2 x_{k-1} \rangle_0 - 2\langle x_k^2 x_{k-2} \rangle_0. \quad \text{(D7)}$$

Looking at Eq. (D5), we deduce that H_{k+1,Ψ_1} has no singular component. Instead, $K_{k-2,k+1}$ can be obtained by performing the inverse Fourier transform of $J_{k-2,k+1}(\omega) - s_{k-2,k+1}$. Using the statistics of the unperturbed L96 model, we obtain $s_{k-2,k+1} \approx 0.071$.

References

1. R. Kubo, Statistical-mechanical theory of irreversible processes. I. general theory and simple applications to magnetic and conduction problems. J. Phys. Soc. Jpn. **12**, 570–86 (1957)
2. R. Kubo, The fluctuation–dissipation theorem. Rep. Prog. Phys. **29**, 255–284 (1966)
3. Kubo's theory initially received a vehement criticism by Van Kampen [71] see in [35] a clear counter-argument to Van Karmen's argument
4. V. Lucarini, M. Colangeli, Beyond the linear fluctuation–dissipation theorem: the role of causality. J. Stat. Mech. (2012)
5. D. Zubarev, P. Gray, P.J. Shepherd, *Nonequilibrium Statistical Thermodynamics. Studies in Soviet Science: Physical Sciences* (Springer, New York, 1974)
6. G. Gallavotti, *Nonequilibrium and Irreversibility* (Springer, Cham, 2014)
7. D. Ruelle, *Chaotic Evolution and Strange Attractors: the Statistical Analysis of Time Series for Deterministic Nonlinear Systems* (Cambridge University Press, Cambridge, 1989)
8. D. Ruelle, Differentiation of SRB states. Commun. Math. Phys. **187**, 227–418 (1997)
9. D. Ruelle, Nonequilibrium statistical mechanics near equilibrium: computing higher-order terms. Nonlinearity **11**, 5–18 (1998)
10. D. Ruelle, A review of linear response theory for general differentiable dynamical systems. Nonlinearity **22**, 855–70 (2009)
11. O. Butterley, C. Liverani, Smooth Anosov flows: correlation spectra and stability. J. Mod. Dyn. **1**(2), 301–322 (2007)
12. V. Baladi, Linear response despite critical points. Nonlinearity **21**(6), T81 (2008)
13. G. Gallavotti, Chaotic hypothesis: Onsager reciprocity and fluctuation–dissipation theorems. Stat. Phys. **84**, 899–925 (1996)
14. G. Gallavotti, E.G.D. Cohen, Dynamical ensembles in stationary states. J. Stat. Phys. **80**, 931–70 (1995)

15. V. Lucarini, Response theory for equilibrium and non-equilibrium statistical mechanics: causality and generalized Kramers–Kronig relations. Stat. Phys. **131**, 543–58 (2008)

16. V. Lucarini, Evidence of dispersion relations for the nonlinear response of the Lorenz 63 system. J. Stat. Phys. **134**, 381–400 (2009)

17. V. Lucarini, S. Sarno, A statistical mechanical approach for the computation of the climatic response to general forcings. Nonlinear Processes Geophys. **18**, 7–28 (2011)

18. R.V. Abramov, A.J. Majda, Blended response algorithms for linear fluctuation–dissipation for complex nonlinear dynamical systems. Nonlinearity **20**, 2793–821 (2007)

19. R.V. Abramov, A.J. Majda, New approximations and tests of linear fluctuation-response for chaotic nonlinear forced-dissipative dynamical systems. Nonlinear Sci. **18**, 303–41 (2008)

20. C. Maes, M. Baiesi, An update on the nonequilibrium linear response. New J. Phys. **15**, 013004 (2013)

21. Q. Wang, Forward and adjoint sensitivity computation of chaotic dynamical systems. J. Comput. Phys. **253**, 1–13 (2013)

22. N. Chandramoorthy, Q. Wang, A computable realization of Ruelle's formula for linearresponse of statistics in chaotic systems (2020)

23. A. Ni, Approximating Ruelle's linear response formula by shadowing methods (2020)

24. M. Hairer, A.J. Majda, A simple framework to justify linear response theory. Nonlinearity **23**(4), 909–922 (2010)

25. P. Hanggi, H. Thomas, Linear response and fluctuation theorems for nonstationary stochastic processes. Z. Phys. B **22**, 295–300 (1975)

26. U. Seifert, T. Speck, Fluctuation–dissipation theorem in nonequilibrium steady states. EPL (Europhys. Lett.) **89**(1), 10007 (2010)

27. L. Rondoni, M. Colangeli, F. Vulpiani, Fluctuation–dissipation relation for chaotic non-Hamiltonian systems. J. Stat. Mech. **2012**, L04002 (2012)

28. R. Zwanzig, Memory effects in irreversible thermodynamics. Phys. Rev. **124**, 983–992 (1961)

29. H. Mori, Transport, collective motion, and Brownian motion. Progr. Theor. Phys. **3**, 423–455 (1965)

30. J. Wouters, V. Lucarini, Multi-level dynamical systems: connecting the Ruelle response theory and the Mori–Zwanzig approach. J. Stat. Phys. **151**, 850–860 (2013)

31. H. Liu, M.D. Chekroun, S. Wang, *Approximation of Stochastic Invariant Manifolds: Stochastic Manifolds for Nonlinear SPDEs I* (Springer, Berlin, 2015)

32. H. Liu, M.D. Chekroun, S. Wang, *Approximation of Stochastic Invariant Manifolds: Stochastic Manifolds for Nonlinear SPDEs II* (Springer, Berlin, 2015)

33. C.L. Wormell, G.A. Gottwald, Linear response for macroscopic observables in high-dimensional systems. Chaos Interdiscip. J. Nonlinear Sci. **29**(11), 113127113127 (2019)

34. H.C. Öttinger, *Beyond Equilibrium Thermodynamics* (Wiley, Hoboken, 2005)

35. R.L. Marini, U.B. Marconi, A. Puglisi, A. Vulpiani, Fluctuation–dissipation: response theory in statistical physics. Phys. Rep. **461**, 111 (2008)

36. V. Lucarini, F. Ragone, F. Lunkeit, Predicting climate change using response theory: global averages and spatial patterns. J. Stat. Phys. **166**(3), 1036–1064 (2017)

37. B. Cessac, Linear response in neuronal networks: from neurons dynamics to collective response. Chaos **29**, 103105 (2019)

38. A. Sarracino, A. Vulpiani, On the fluctuation–dissipation relation in non-equilibrium and non-Hamiltonian systems. Chaos **29**, 083132 (2019)

39. M. Ghil, V. Lucarini, The physics of climate variability and climate change. Rev. Mod. Phys. **92**, 035002 (2020)

40. V. Lembo, V. Lucarini, F. Ragone, Beyond forcing scenarios: predicting climate change through response operators in a coupled general circulation model. Sci. Rep. **10**(1), 8668 (2020)

41. M.D. Chekroun, E. Simonnet, M. Ghil, Stochastic climate dynamics: random attractors and time-dependent invariant measures. Phys. D Nonlinear Phenom. **240**(21), 1685–1700 (2011)

42. M. Aengenheyster, Q.Y. Feng, F. van der Ploeg, H.A. Dijkstra, The point of no return for climate action: effects of climate uncertainty and risk tolerance. Earth Syst. Dyn. **9**(3), 1085–1095 (2018)

43. T. Bódai, V. Lucarini, F. Lunkeit, Can we use linear response theory to assess geoengineering strategies? Chaos Interdiscip. J. Nonlinear Sci. **30**(2), 023124 (2020)

44. V. Lucarini, J. Wouters, Response formulae forn-point correlations in statistical mechanical systems and application to a problem of coarse graining. J. Phys. A Math. Theor. **50**(35), 355003 (2017)

45. V. Lucarini, G.A. Pavliotis, N. Zagli, Response theory and phase transitions for the thermodynamic limit of interacting identical systems. Proc. R. Soc. A Math. Phys. Eng. Sci. **476**(2244), 20200688 (2020)

46. G.A. Gottwald, Introduction to focus issue: linear response theory: potentials and limits. Chaos Interdiscip. J. Nonlinear Sci. **30**(2), 020401 (2020)

47. A.J. Majda, D. Qi, Strategies for reduced-order models for predicting the statistical responses and uncertainty quantification in complex turbulent dynamical systems. SIAM Rev. **60**(3), 491–549 (2018)

48. A.J. Majda, D. Qi, Linear and nonlinear statistical response theories with prototype applications to sensitivity analysis and statistical control of complex turbulent dynamical systems. Chaos Interdiscip. J. Nonlinear Sci. **29**(10), 103131 (2019)

49. V. Lucarini, Revising and extending the linear response theory for statistical mechanical systems: evaluating observables as predictors and predictants. J. Stat. Phys. **173**, 1698–1721 (2018)

50. B. Gustavsen, A. Semlyen, Rational approximation of frequency domain responses by vector fitting. IEEE Trans. Power Deliv. **14**(3), 1052–1061 (1999)

51. M. Berljafa, S. Güttel, The RKFIT algorithm for nonlinear rational approximation. SIAM J. Sci. Comput. **39**(5), A2049–A2071 (2017)

52. E. Lorenz, *Predictability: a Problem Partly Solved. Proc Seminar on Predictability*, vol. 1 (Berkshire, ECMWF, Reading, 1996), pp. 1–18

53. E. Lorenz, K. Emanuel, Optimal sites for supplementary weather observations. J. Atmos. Sci. **55**, 399–414 (1998)

54. E. Lorenz, Designing chaotic models. J. Atmos. Sci. **62**, 1574–1587 (2004)

55. C.W.J. Granger, Investigating causal relations by econometric models and cross-spectral methods. Econometrica **37**(3), 424–438 (1969)

56. J. Pearl, *Causality: Models, Reasoning, and Inference* (Cambridge University Press, Cambridge, 2009)
57. J. Runge, V. Petoukhov, J. Kurths, Quantifying the strength and delay of climatic interactions: the ambiguities of cross correlation and a novel measure based on graphical models. J. Clim. **27**(2), 720–739 (2014)
58. A. Hannart, J. Pearl, F.E.L. Otto, P. Naveau, M. Ghil, Causal counterfactual theory for the attribution of weather and climate-related events. Bull. Am. Meteorol. Soc. **97**(1), 99–110 (2016)
59. E. Saggioro, T.G. Shepherd, Quantifying the timescale and strength of southern hemisphere intraseasonal stratosphere–troposphere coupling. Geophys. Res. Lett. **46**(22), 13479–13487 (2019)
60. J. Runge, S. Bathiany, E. Bollt, G. Camps-Valls, D. Coumou, E. Deyle, C. Glymour, M. Kretschmer, M.D. Mahecha, J. Muñoz-Marí, E.H. van Nes, J. Peters, R. Quax, M. Reichstein, M. Scheffer, B. Schölkopf, P. Spirtes, G. Sugihara, J. Sun, K. Zhang, J. Zscheischler, Inferring causation from time series in earth system sciences. Nat. Commun. **10**(1), 2553 (2019)
61. M. Baldovin, F. Cecconi, A. Vulpiani, Understanding causation via correlations and linear response theory. Phys. Rev. Res. 2(4), 043436 (2020). https://doi.org/10.1103/PhysRevResearch.2.043436
62. C. Sarra, M. Baldovin, A. Vulpiani, Response and flux of information in extended non-equilibrium dynamics (2021). arXiv:2105.11940
63. H.M. Nussenzweig, *Causality and Dispersion Relations* (Academic Press, New York, 1972)
64. K.-E. Peiponen, V. Lucarini, J.J. Saarinen, E.M. Vartiainen, *Kramers–Kronig Relations in Optical Materials Research* (Springer, Berlin, 2006)
65. G. Gallavotti, V. Lucarini, Equivalence of non-equilibrium ensembles and representation of friction in turbulent flows: the Lorenz 96 model. J. Stat. Phys. **156**, 1027–1065 (2014)
66. G. Vissio, V. Lucarini, Mechanics and thermodynamics of a new minimal model of the atmosphere. Eur. Phys. J. Plus **135**(10), 807 (2020)
67. D.L. Van Kekem, A.E. Sterk, Predictability of extreme waves in the Lorenz-96 model near intermittency and quasi-periodicity. Complexity (2017)
68. D.L. Van Kekem, A.E. Sterk, Travelling waves and their bifurcations in the Lorenz-96 model. Phys. D. **367**, 38–60 (2018)
69. M. Gelbrecht, V. Lucarini, N. Boers, J. Kurths, Analysis of a bistable climate toy model with physics-based machine learning methods. Eur. Phys. J. Spec. Top. (2021). https://doi.org/10.1140/epjs/s11734-021-00175-0
70. A. Gritsun, V. Lucarini, Fluctuations, response, and resonances in a simple atmospheric model. Phys. D Nonlinear Phenom. **349**, 62–76 (2017)
71. N. Van Kampen, The case against linear response theory. Phys. Nor. **5**, 279–284 (1971)