CrossMark

Response Operators for Markov Processes in a Finite State Space: Radius of Convergence and Link to the Response Theory for Axiom A Systems

Valerio Lucarini[1,2]

Abstract Using straightforward linear algebra we derive response operators describing the impact of small perturbations to finite state Markov processes. The results can be used for studying empirically constructed—e.g. from observations or through coarse graining of model simulations—finite state approximation of statistical mechanical systems. Recent results concerning the convergence of the statistical properties of finite state Markov approximation of the full asymptotic dynamics on the SRB measure in the limit of finer and finer partitions of the phase space are suggestive of some degree of robustness of the obtained results in the case of Axiom A system. Our findings give closed formulas for the linear and nonlinear response theory at all orders of perturbation and provide matrix expressions that can be directly implemented in any coding language, plus providing bounds on the radius of convergence of the perturbative theory. In particular, we relate the convergence of the response theory to the rate of mixing of the unperturbed system. One can use the formulas derived for finite state Markov processes to recover previous findings obtained on the response of continuous time Axiom A dynamical systems to perturbations, by considering the generator of time evolution for the measure and for the observables. A very basic, low-tech, and computationally cheap analysis of the response of the Lorenz '63 model to perturbations provides rather encouraging results regarding the possibility of using the approximate representation given by finite state Markov processes to compute the system's response.

Keywords Markov process · Response theory · Radius of convergence · Perron–Frobenius operator · Ulam conjecture · Lorenz system

✉ Valerio Lucarini
    valerio.lucarini@uni-hamburg.de

1   CEN, Institute of Meteorology, University of Hamburg, Hamburg, Germany

2   Department of Mathematics and Statistics, University of Reading, Reading, UK

1 Introduction

1.1 A Brief Summary of Response Theory

The development of methods for computing the response of a complex system to small perturbations affecting its dynamics is the subject of very active investigation in many fields of science and of technology. Statistical mechanics provides tools for approaching such a problem through so-called response theories, which allow for evaluating the change in the properties of a system through suitably defined operators that factor in the statistical properties of the unperturbed system and the specific nature of the perturbation one wants to study.

One can see a response theory as a virtual experimental setting where one has at hand a given system, various measurement instruments, and a knob controlling the value of a parameter, and knows how to relate the position of the knob with the reading of the instruments. In other terms, response theories provide the basis for understanding the outcome of experiments, and, not by chance, physical sciences have been at the forefront of the theoretical investigation in this direction. The monumental contribution by [1] provided the basis and the explicit formulas needed for studying the impact of very general perturbations to statistical mechanical systems at equilibrium, as described by the canonical ensemble. The Kubo formulas are extremely useful for studying a large class of problems in e.g. transport, optics, and acoustics. A cornerstone of Kubo's theory is the fluctuation–dissipation relation, which enables connecting—within linear approximation—the free fluctuations of a system to its response to perturbations. This property is closely related to the celebrated diffusion law for the brownian motion and has been recently extend to a fully nonlinear case [2]. Despite its obvious relevance, Kubo's approach has been criticized for several reasons:

- it is not physically consistent in treating the transition from equilibrium to non-equilibrium dynamics, because it studies the impact on equilibrium systems of perturbations that drive them near (but out of) equilibrium, but does not clarify how a new stationary state is reached and maintained; additionally, it is not suited for studying the response to perturbations of non-equilibrium systems;
- it lacks mathematical rigour, as it is not clear which are the systems for which the response formulas apply, and why it should apply at all.

In [3–5] it was clarified that it is possible to establish a rigorous response theory for Axiom A [6] continuous or discrete time dynamical systems. One obtains that the invariant SRB measure is smooth with respect to the parameter ϵ that controls the strength of the perturbation changing the dynamics of the system from $\dot{\mathbf{x}} = F(\mathbf{x})$ to $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}) + \epsilon \mathbf{X}(\mathbf{x})$, in the case of continuous time evolution, and from $\mathbf{x}_{k+1} = \mathbf{F}(\mathbf{x}_k)$ to $\mathbf{x}_{k+1} = \mathbf{F}(\mathbf{x}_k) + \epsilon \mathbf{X}(\mathbf{x}_k)$, in the discrete case. We continue our discussion taking into consideration the continuous case.

We can introduce the unperturbed evolution operator $S_0^t = \exp(t\mathbf{F}\cdot)$, which moves forward in time any function of phase space $O(\mathbf{x})$ by an interval t according to the unperturbed dynamics, so that $O(\mathbf{x}(t)) = S_0^t O(\mathbf{x}(0))$, and its perturbed counterpart $S_\epsilon^t = \exp(t(\mathbf{F}+\epsilon\mathbf{X})\cdot)$, which instead describes the evolution in the perturbed system.

We define $\rho_0(d\mathbf{x})$ and $\rho_\epsilon(d\mathbf{x})$ the invariant measures of the unperturbed and perturbed states, respectively. In particular, one obtains that the expectation value of sufficiently smooth observables $O(\mathbf{x})$ in the perturbed state can be expressed in the form:

$$[O]_\epsilon = [O]_0 + \sum_{n=1}^{\infty} \epsilon^n \delta[O]_n, \tag{1}$$

where $[Q]_\epsilon = \int \nu_\epsilon(\mathrm{d}\mathbf{x}) Q(\mathbf{x})$ and $[Q]_0 = \int \nu_0(\mathrm{d}\mathbf{x}) Q(\mathbf{x})$, while the various terms of the perturbative expansion can be written as:

$$\delta[O]_n = \int \nu_0(\mathrm{d}\mathbf{x}) \int_0^\infty \mathrm{d}t_1 \cdots \int_0^\infty \mathrm{d}t_n \Lambda S_0^{t_1} \cdots S_0^{t_{n-1}} \Lambda S_0^{t_n} O(\mathbf{x}), \tag{2}$$

where $\Lambda(\bullet) = \mathbf{X} \cdot \nabla(\bullet)$. In particular, the linear term can be written as:

$$\delta[O]_1 = \int \nu_0(\mathrm{d}\mathbf{x}) \int_0^\infty \mathrm{d}t_1 \Lambda S_0^{t_1} O(\mathbf{x}), \tag{3}$$

All terms $\delta[O]_j$ can be written as an expectation value on the unperturbed measure of a new observable expressed as a functional of the background vector field $\mathbf{F}$, of the perturbative vector field $\mathbf{X}$, and of the observable O. The somewhat surprising conclusion we draw is that the invariant measure of the system, despite being supported on a strange geometrical set, is differentiable with respect to ϵ. Among the many merits of the Ruelle response theory, one can mention that a) it clarifies the mathematical framework needed for developing a response theory, whose main ingredient, roughly speaking, is the robustness deriving from having a uniformly hyperbolic dynamics on the attractor supporting an SRB measure; and b) it works seamlessly, in principle, in equilibrium and non equilibrium statistical mechanical systems, reducing to Kubo's formulas when considering the first scenario, *if one assumes that statistical mechanical systems are Axiom A*. Non-trivial implications of the nonequilibrium/equilibrium dichotomy regarding the validity of the fluctuation-dissipation relations are discussed in [2,5,7], while the a physical interpretation of the first and second order terms occurring in Ruelle's response formalism is provided in [8].

Of course, at this stage one needs to bridge the gap between mathematical formalism and physical meaningfulness, One manages to bring Ruelle's formalism into the realm of applicability by adopting the *chaotic hypothesis* [9,10], which basically says that a high-dimensional chaotic physical system can be treated at all practical purposes as if it were Axiom A if we focus on macroscopic observables. The chaotic hypothesis is the generalisation of the ergodic hypothesis, and provides a firm background for translating the mathematical properties of Axiom A systems into physically meaningful statements. Clearly, the chaotic hypothesis applies far from regimes of metastability and far from critical transitions, where entirely different phenomena appear. The chaotic hypothesis might also be practically problematic in the case one treats multiscale systems featuring many near-zero Lyapunov exponents; see discussion in [11].

Taking the point of view of the chaotic hypothesis, one has that, after transients have died out, nonequilibrium systems reach a nonequilibrium steady state (NESS) where the phase space is on the average contracting (with the rate of contraction corresponding, broadly speaking, to the entropy production of the system [12]), so that one can associate to the hyperbolic strange attractor supporting the invariant measure a Hausdorff dimension that is lower that the dimensionality of the phase space and, in general, not integer [6,13].

The last piece of the puzzle one needs to lay in order to sort out the above-mentioned criticisms to Kubo's theory relies on the physical interpretation of how a perturbed equilibrium system reaches a steady state. A convincing point of view on this relies on emphasizing the role of thermostats, which are large physical systems interacting with the system of interest in such a way to extract the excess of heat generated as result of the energy input due to the perturbation. Thermostats are also responsible for making it possible the set-up of stationarity in the case of forced and dissipative non equilibrium systems. An extensive treatment of the role of thermostats in equilibrium and nonequilibrium systems in the context of the chaotic hypothesis is given in [14]. We will not elaborate further on this aspect here.

1.2 Transfer Operator Approach

One can point out that the formulas above describe the impact of and expressed in terms of expectation values of a generic observable O, whereas one might like to derive directly results for the impacts of the perturbations on the invariant measure.

In [3–5] one constructs the response of the system to perturbations by following the changes in the individual trajectories and summing over the possible initial configurations distributed according to the unperturbed invariant measure. A different point of view on response theory focuses on studying the properties of the unperturbed and perturbed transfer operators and of their generators (see [15] for an introduction on these mathematical objects), through the construction of an appropriate framework of suitable (Banach) functional spaces where their actions are well defined, able to carefully treat the fundamental differences between the (smooth) unstable and (singular) stable manifolds of the Axiom A systems [16–19].

The evolution of the measure driven by the system $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$ up to time $t \geq 0$ starting from an initial condition at time $t = 0$ is described by the Perron–Frobenius transfer $\mathcal{L}^t$ (see, e.g., [15]), so that $\rho(\mathbf{x}, t) = \mathcal{L}^t \rho(\mathbf{x}, 0)$. We have that the family of $\{\mathcal{L}^t\}_{t \geq 0}$ forms a one-parameter semigroup, such that $\mathcal{L}^{t+s} = \mathcal{L}^t \mathcal{L}^s$ and $\mathcal{L}^0 = \mathbf{1}$. The Perron–Frobenius operator $\mathcal{L}^t$ is the adjoint of the evolution operator $S^t = \left(\mathcal{L}^t \right)^\top$, so that $\langle S^t O, \rho \rangle = \langle O, \mathcal{L}^t \rho \rangle$, where $\langle f, g \rangle$ is the action (computation of the expectation value) of the linear functional g (the probability measure) on the test function f (the observable). We have that $\mathcal{L}^t \nu_0 = \nu_0 \; \forall t \geq 0$, meaning that the invariant measure is an eigenvector corresponding to unitary eigenvalue of the Perron–Frobenius operator.

Assuming strong continuity and boundedness of the semigroup given by $\{\mathcal{L}^t\}_{t \geq 0}$, we can introduce the unperturbed Liouvillian operator L, which is the generator of the unperturbed Perron–Frobenius operator $\mathcal{L}^t = \exp(tL)$, and write the Liouville evolution equation for $\rho(\mathbf{x}, t)$ as follows [20]:

$$\partial_t \rho = -\nabla \cdot (\rho \mathbf{F}) = L\rho \tag{4}$$

One immediately obtains that $L\nu_0 = 0$. In general, the spectrum of L is complex and in a strip of finite width including and below the imaginary axis consists only of isolated eigenvalues of finite multiplicity corresponding to the Ruelle–Pollicott resonances, while below such a strip one finds the essential spectrum, which is responsible for the continuum of the power spectra of integrable observables. Furthermore, the presence of a unique SRB measure comes from the presence of a simple vanishing eigenvalue, while mixing properties result from the absence of any other eigenvalue along the imaginary axis. The relevance of these properties for constructing a response theory are discussed in great detail in [18, 19]. In [21] it is argued, using mathematical considerations and examples of geophysical relevance, that the presence of Ruelle–Pollicott resonances having real part close to zero may lead to the presence of rough parameter dependence, as the smoothness of the response if lost. Additionally, in [22], it is shown, along similar lines, that the crisis of a very high-dimensional chaotic attractor near a critical transition—namely, of a climate model in the vicinity of the tipping point responsible for the transition between warm and snowball climate [23–26]—can be detected and anticipated by looking at spectrum of the transfer operator.

We then have that the presence of the ϵ perturbation to the dynamics changes the Liouville equation as follows:

$$\partial_t \rho = -\nabla \cdot (\rho \mathbf{F}) - \epsilon \nabla \cdot (\rho \mathbf{X}) = L_\epsilon \rho, \tag{5}$$

so that we can introduce the perturbed Perron–Frobenius operator $\mathcal{L}_\epsilon^t = \exp(tL_\epsilon)$, which pushes forward in time the measure according to the perturbed dynamics: $\rho(\mathbf{x}, t) = \mathcal{L}_\epsilon^t \rho(\mathbf{x}, 0)$. Clearly, $\langle S_\epsilon^t O, \rho \rangle = \langle O, \mathcal{L}_\epsilon^t \rho \rangle$. Additionally, we have that $\mathcal{L}_\epsilon^t \nu_\epsilon = \rho_\epsilon \; \forall t \geq 0$ and $L_\epsilon \nu_\epsilon = 0$. While this approach is in some sense mathematically more problematic, because it is based on studying a partial differential equation instead of a finite dimensional dynamical system, it seems to provide a more comprehensive set of tools for studying the response of a system and relating it to its unperturbed fluctuations, see, e.g., [16], where Ruelle's formulas are obtained along these lines. See also a comprehensive review given in [18], where the applicability of the response theory beyond the case of Axiom A systems is discussed in detail..

One needs to emphasise that the transfer operator approach is more natural in all the cases when our interest focuses on studying the properties of the response of an ensemble of trajectories (initialised according to the unperturbed invariant measure) rather than on individual orbits of a system.

Note that in some applications there is not an obvious separation between the two approaches. Let's take the problem of constructing climate projections through the use of (extremely complex) numerical climate models, which is one of the core activities summarized in the IPCC reports [27]. Indeed, modelling centers are actively pursuing the preparation of multiple runs starting from an ensemble of initial conditions for a given scenario of forcing in order to estimate more accurately the uncertainties in the projections. Nonetheless, we will not experience an ensemble of realizations of the climatic evolution, but just one.

1.3 Computing the Response

The analysis of high-dimensional complex system in terms of direct numerical simulation and of time series analysis suffers from the (almost) ubiquitous curse of dimensionality, which makes it hard to represent correctly the details of the dynamics because computational complexity explodes with the number of degrees of freedom. The construction of efficient and accurate algorithms for studying the response of a complex system to perturbations faces serious difficulties. Let's focus now on the linear case. Some previous studies have emphasised the need for treating separately the contributions to the response coming from short and long-time delayed contributions in Eq. 3, and have underlined the need for reducing the complexity of the invariant measure by adding in the background state some stochastic forcing, able to smooth out the singularity of the SRB measure [28,29].

A promising way to deal with the actual computation of the scalar product in Eq. 3 is to use as time-dependent basis the covariant Lyapunov vectors [30,31], which automatically separate the contributions to the response coming from the unstable, neutral, and stable directions. This clarifies that the convergence of the formula given in Eq. 3 comes from the two distinct facts that (a) perturbations along the stable directions naturally decay, and (b) perturbations along the unstable directions grow in size, but are dominated by the loss of correlation due to mixing.

Recently, algorithms based upon adjoint methods have shown a good degree of accuracy and seem promising, even if scaling them up to high-dimensional systems has not been attempted yet [32,33]. A different approach to the problem has been proposed in [7,34–36], where, instead of trying to computing ab initio and directly the response given in Eq. 3, the authors construct it a posteriori, probing the system with some test forcings and using the formal properties of the theory to be able to predict the response for new patterns of forcings. One can say that by studying the differential response to similar yet differently modulated perturbations, it is possible to derive the overall response properties of the system.

1.4 This Paper

Any numerical representation of a continuum system builds upon the need of discretizing the phase space and, in the case of time-continuous system, of time.

In this case, we partition the phase space of the system in say N states $\phi_1, \ldots, \phi_N$. In many cases, the states are constructed by discretizing the phase space in a grid of boxes, which provide a (Galerkin) basis of orthogonal functions. We then construct an initial ensemble as defined by the occupancy $u_0^1, \ldots, u_0^N$ of each of the ϕ_i's, $i = 1, \ldots, N$, so that

$$u_0^i = \int d\mathbf{x} \rho(\mathbf{x}, 0) \mathbf{1}(\phi_i),$$

where $\mathbf{1}(A)$ is the characteristic function in the set A, and we want to approximate the evolution of such occupancies change with time, considering discrete time steps Δt, so that, to a good approximation the occupancy at time $k\Delta t$ is

$$u^i(k\Delta t) \sim \int d\mathbf{x} \mathcal{L}^{k\Delta t} \rho(\mathbf{x}, 0) \mathbf{1}(\phi_i).$$

Moreover, in such a discrete representation, we have that the value of an observable O in the state ϕ_i is given by its average

$$O(k\Delta t) = \frac{\int d\mathbf{x} \rho(\mathbf{x}, k\Delta t) \mathbf{1}(\phi_i) O(\mathbf{x})}{\int d\mathbf{x} \rho(\mathbf{x}, k\Delta t) \mathbf{1}(\phi_i)}. \tag{6}$$

Let's emphasize that when analyzing virtually any sort of complex system, almost invariably one proposes a natural spatial and temporal cut-off, so that one on not in fact interested in really being able to compute the response of any possible observable defined at any possible spatial and temporal resolution, whereas meso- or macroscopic properties are relevant. Going again to the useful example of climate science, it is commonly regarded as a good and useful question to learn about the change in the surface temperature in response to climate forcing on a spatial scale corresponding to say a continent or a fraction thereof, and on a temporal scale of say one year. Nobody would find useful nor intelligent to study the surface temperature response over extremely small temporal and spatial scales.

Empirically, using long numerical integrations and defining the set of finite states ϕ_i, $i = 1, \ldots, N$, we can construct the stochastic matrix $\mathcal{M}_{i,j}$ describing the probability of performing a transition from state ϕ_i to state ϕ_j in a period of time Δt. The same operation can in principle be performed using experimental and observational data. A fundamental issue at the core of such procedure is whether for some dynamical systems in the limit of finer and finer partitions covering the phase space (actually, the attractor of the system) with $N \to \infty$ one reconstructs the actual invariant measure of the original system. See in [37] a comprehensive discussion of such an issue, the so-called *Ulam conjecture*, and in [38] some extremely promising applications of finite state Markov processes for studying severely reduced representations of complex systems.

Following the idea that the performing the discretization of the phase space amounts to adding a stochastic perturbation of the original dynamical systems, with intensity going to zero with the scale of the actual partitions, and exploiting the fact that the SRB measure can be constructed as zero-noise limit (with measure that is absolutely continuous with respect to Lebesgue) of the physical measure, in [39,40] it has been proposed that the Ulam conjecture applies in the case of Axiom A systems, which are endowed with an SRB measure. The convergence in the case of Anosov diffeormorphism has indeed been proved provided one adds some noise of asymptotically vanishing intensity (through stronger than the noise induced

by the partition itself) to the underlying dynamics [41]. Somehow this is not so surprising because by adding noise one introduces a cutoff below which partitions do indeed work. At any practical level, these results suggest that in the case of Axiom A system constructing finite state Markov processes using Ulam partitions can do a pretty good job in simulating the true dynamics, if one consider reasonably well-behaved, smooth observables as test functions. Nonetheless, one has to note that different choices for the partitions can lead to very different rates of convergence [37]. See also the discussion and the numerical examples presented in [42].

Apart from the Ulam method, one can follow a mathematically more elegant but practically much harder way to construct finer and finer partitions. As well known, Axiom A systems possess Markov partitions, i.e. well-defined, metric independent, finite resolution representations of the phase space that refine themselves with the dynamics [6,14]. Such Markov partitions can be used to construct in the limit the actual SRB measure of the system, and, additionally, following [43], they provide a natural way to build finite Markov chains whose properties converge in the limit to those of the Perron–Frobenius operator of the system.

Having a response formulas in the finite case has direct relevance for finite Markov chains and for interpreting the results of reduced models. Another good reason to construct a response theory in a finite state space has to do with the fact that the response operators for Axiom A systems introduced by Ruelle can be written as expectation value of certain observables on the unperturbed SRB measure. Therefore, given what said above, one can hope to have convergence of the finite state reconstructed response operators to the corresponding *true* response operator in the limit of infinitely fine partitions of the dynamics. Actually, providing explicit formulas for the response operator for a finite state partition of a system the response operator and taking the limit for (suitably defined) finer and finer partitions could be interpreted as a rigorous way for constructing the actual response on the asymptotic SRB measure. One needs to note—see discussion in Sects. 2.1 and 3—that special attention has to be paid when studying the convergence of such operators.

In what follows, we present the derivation of the response formulas at all orders of perturbations (as well as the full nonlinear versions) for finite state spaces of arbitrary size N. All expressions are given in terms of the transitions matrix of the unperturbed system, to its corrections to the perturbation, and of the parameter controlling the strength of the perturbation. The interest we see in the calculations we present below is mostly three-fold:

- our results are obtained using basic linear algebra operations in finite dimensional spaces, which can used to interpret more complex operators acting on infinite dimensional spaces. It is also possible to use the finite dimensional expressions to derive, e.g., the the actual response operators for continuous time Axiom A dynamical systems;
- we are able to derive an explicit expression for the a lower bound to for the radius of convergence of the perturbative theory, and relate it with the mixing properties of the unperturbed system. We also find a (very tentative) expression for such a lower bound in the case of continuous time case Axiom A dynamical systems;
- our formulas can be translated into one-line commands in now widely available software tools like R, Octave, or MATLAB®. This might greatly facilitate the actual implementation of response operators. In particular, we can say that our results provide a direct translation of the response theory into a readily implementable algorithms.

The paper is organised as follows. In Sect. 2, we introduce some notation and provide basic properties of ergodic finite state Markov chains, which can be taken as mathematical model on its own or as finite precision representation of ergodic (in this case, Axiom A) systems. We also show how it is possible to find an exact expression for the impact of a perturbation

on the invariant measure of the Markov process and we study the radius of convergence of the perturbative expansion. In Sect. 3 we rephrase our results in terms of observables, by constructing straightforward adjoint operators in finite dimensions. In Sect. 4 we show how our findings agree with the response theory for continuous time systems when we suitably translate the matrix operations into operators. In Sect. 5 we present a simple yet instructive investigation of the response of the Lorenz '63 system [44] to perturbations using Ulam-like partitions and the formalism developed here. In Sect. 6 we recapitulate and discuss our results.

2 Response Operators for Finite-State Markov Processes

Let's consider an ergodic Markov process with a finite number of states defined by the N-component vector $\mathbf{u}$. We consider the infinite Markov chain generated as $\mathbf{u_0}, \mathcal{M}\mathbf{u_0}, \ldots \mathcal{M}^n\mathbf{u_0}$, $\ldots$ where $\mathbf{u}_0$ is the initial ensemble of states, and $\mathcal{M}_{i,j} \in \mathbb{R}^{N \times N}$ is the stochastic transition matrix determining the probability of reaching the state i at step n if at step $n-1$ we are in the state j. The process is taken to be stationary, so that $\mathcal{M}$ does not change with n. We remind that $\mathcal{M}$ is such that $\sum_{i=1}^{N} \mathcal{M}_{i,j} = 1$ and $\mathcal{M}_{i,j} \geq 0 \; \forall i, j = 1, \ldots, N$.

The invariant measure is obtained by solving the eigenvalue problem

$$\mathcal{M}\mathbf{u} = \lambda\mathbf{u}, \tag{7}$$

and selecting the unique solution with eigenvalue $\lambda = 1$. The corresponding (column) eigenvector $\mathbf{u_1}$[1] is the invariant measure of the system. We also remind that

$$\lim_{n \to \infty} \mathcal{M}^n \mathbf{z} = \alpha_1 \mathbf{u_1}, \quad \forall \mathbf{z}. \tag{8}$$

where $\{\lambda_j, \mathbf{u_j}\} \; j = 1, \ldots, N$ are the pairs of eigenvalues and eigenvectors of $\mathcal{M}$, where $\lambda_1 = 1$, $|\lambda_j| < 1$ if $j > 1$, and $\mathbf{z}$ can be expressed as $\mathbf{z} = \sum_{j=1}^{N} \alpha_j \mathbf{u_j}$.

Our goal is to find a formula for expressing the change in the invariant measure resulting from perturbing the transition matrix $\mathcal{M} \to \mathcal{M} + \epsilon m$.

We note that in order to preserve the Markov property of the system, m obeys the following constraint: $\sum_{j=1}^{N} m_{i,j} = 0$, so that $\sum_{j=1}^{N} \left(\mathcal{M}_{i,j} + \epsilon m_{i,j} \right) = 0$. Moreover, an additional constraint on ϵm comes from the fact that all elements of $\mathcal{M} + \epsilon m$ have to be positive. We define

$$\epsilon_+ = \max_{\epsilon} |\forall i, j \in \{1, \ldots, N\}, \mathcal{M}_{i,j} + \epsilon m_{i,j} \geq 0, \tag{9}$$

and

$$\epsilon_- = \min_{\epsilon} |\forall i, j \in \{1, \ldots, N\}, \mathcal{M}_{i,j} + \epsilon m_{i,j} \geq 0; \tag{10}$$

clearly, $\epsilon_- \leq 0 \leq \epsilon_+$, and the perturbed matrix is a stochastic matrix $\forall \epsilon \in [\epsilon_-, \epsilon_+]$. In order to have some *room* for studying the impacts of perturbations, we require that $\epsilon_+ - \epsilon_- > 0$. Such conditions show that, for a given $\mathcal{M}$, it makes sense to consider only a specific class of perturbation matrices m. Let's provide an example of an ill-chosen m: if $\mathcal{M}$ has two zero entries $\mathcal{M}_{i_1,j_1} = \mathcal{M}_{i_2,j_2} = 0$ and $m_{i_1,j_1} m_{i_2,j_2} < 0$, then we have $\epsilon_- = 0 = \epsilon_+$.

The new invariant measure is the unique solution to the eigenvalue problem:

$$(\mathcal{M} + \epsilon m)\mathbf{u} = \lambda\mathbf{u}, \tag{11}$$

---

[1] Most commonly Markov chains are constructed using row vectors; we use column vectors because we find it easier to perform formal matrix manipulations and because we are closer to the formulation most commonly implemented in scientific software.

with unitary eigenvalue. We define $\mathbf{v_1}$ as the invariant measure of the perturbed system. Our goal is to express it as a function of $\mathcal{M}$, m, ϵ and $\mathbf{u}$. This amounts to constructing a response theory. We first present the results of the explicit calculation, and then discuss issues of well-posedness of the problem and convergence of the procedure in Sect. 2.1. Let's express $\mathbf{v_1} = \mathbf{u_1} + \sum_{n=1}^{\infty} \epsilon^n \mathbf{w^n}$, so that we obtain:

$$(\mathcal{M} + \epsilon m)\left(\mathbf{u_1} + \sum_{n=1}^{\infty} \epsilon^n \mathbf{w^n}\right) = \mathbf{u_1} + \sum_{n=1}^{\infty} \epsilon^n \mathbf{w^n}, \tag{12}$$

Note that the first eigenvalue is *not* changed by the perturbation $\mathcal{M} \to \mathcal{M} + \epsilon m$, because also $\mathcal{M} + \epsilon m$ is a stochastic matrix. Using the definition of $\mathbf{u_1}$ we obtain a system of concatenated equations

$$(1 - \mathcal{M})\mathbf{w^1} = m\mathbf{u_1} \tag{13}$$

$$(1 - \mathcal{M})\mathbf{w^n} = m\mathbf{w^{n-1}}, \quad \forall n \in \mathbf{N}, n > 1. \tag{14}$$

We obtain

$$\mathbf{w^1} = \Psi_1 \mathbf{u_1} = (1 - \mathcal{M})^{-1} m \mathbf{u_1} \tag{15}$$

$$\mathbf{w^n} = \Psi_1 \mathbf{w^{n-1}}. \tag{16}$$

Given the recursive structure, we immediately derive the general formula:

$$\mathbf{w^n} = \Psi_n \mathbf{u_1} = \Psi_1^n \mathbf{u_1} = \prod_{j=1}^{n} \left((1 - \mathcal{M})^{-1} m\right) \mathbf{u_1}. \tag{17}$$

where $\Psi_n = \Psi_1^n$. Concluding, we have that:

$$\mathbf{v_1} = \mathbf{u_1} + \sum_{n=1}^{\infty} \epsilon^n \mathbf{w_n} = \mathbf{u_1} + \sum_{n=1}^{\infty} \epsilon^n \Psi_1^n \mathbf{u_1} = \mathbf{u_1} + \sum_{n=1}^{\infty} \epsilon^n \prod_{j=1}^{n} \left((1 - \mathcal{M})^{-1} m\right) \mathbf{u_1} \tag{18}$$

which provides the formula we have been looking for. We note that the term responsible for the nth order of perturbation to the measure can be expressed as

$$\lim_{\epsilon \to 0} \frac{1}{n!} \frac{d^n}{d\epsilon^n} \mathbf{v_1}. \tag{19}$$

Using the matrix identity $(1 - \mathcal{N})^{-1} = \sum_{k=0}^{\infty} \mathcal{N}^k$ with $\mathcal{N} = \epsilon \Psi_1 = \epsilon (1 - \mathcal{M})^{-1} m$, we can also formally express the previous result as:

$$\mathbf{v_1} = (1 - \epsilon \Psi_1)^{-1} \mathbf{u_1} = (1 - \epsilon(1 - \mathcal{M})^{-1} m)^{-1} \mathbf{u_1}. \tag{20}$$

Using again the matrix identity $(1 - \mathcal{M})^{-1} = \sum_{k=0}^{\infty} \mathcal{M}^k$, the previous expression can be rewritten as:

$$\mathbf{v_1} = (1 - \epsilon \Psi_1)^{-1} \mathbf{u_1} = \left(1 - \epsilon \sum_{k=0}^{\infty} \mathcal{M}^k m\right)^{-1} \mathbf{u_1}. \tag{21}$$

or

$$\mathbf{v_1} = \mathbf{u_1} + \sum_{n=1}^{\infty} \epsilon^n \prod_{j=1}^{n} \left(\sum_{k=0}^{\infty} \mathcal{M}^k m\right) \mathbf{u_1} \tag{22}$$

2.1 Well-Posedness and Convergence

In the previous equations, we have used somewhat carelessly the expression $(1 - \mathcal{M})^{-1}$. Unfortunately, the matrix $1 - \mathcal{M}$ is *not* invertible, because all of its columns sum up to zero, or, alternatively, because we know that 1 is an eigenvalue of $\mathcal{M}$. Nonetheless, the expression makes sense if we apply it to a vector belonging to span$\{\mathbf{u_2}, \ldots, \mathbf{u_n}\}$. We now want to prove that:

Lemma 1 *If $\mathcal{M}$ is a Markov transition matrix $\mathbb{R}^N \to \mathbb{R}^N$ with eigenvectors $(\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_N})$, and corresponding eigenvalues $(\lambda_1 = 1, \lambda_2, \ldots, \lambda_N)$, $1 > |\lambda_2| \geq \ldots |\lambda_N|$, and m is a matrix matrix $\mathbb{R}^N \to \mathbb{R}^n$ such that $\sum_{i=1}^{n} m_{i,j} = 0$, then $m\mathbf{z} \in$ span$\{\mathbf{u_2}, \ldots, \mathbf{u_n}\} \forall \mathbf{z} \in \mathbb{R}^n$.*

Proof Let's consider the vector $\mathbf{y} = m\mathbf{z}$. Its ith component can be written as $y_i = \sum_{j=1}^{N} m_{i,j} z_j$. Since $\sum_{i=1}^{N} m_{i,j} = 0$, we have that $\sum_{i=1}^{N} z_i = \sum_{i=1}^{N} \sum_{j=1}^{N} m_{i,j} z_j = 0$.

Let's now consider the kth eigenvector $\mathbf{u}_k$ of $\mathcal{M}$. We have $\sum_{j=1}^{N} \mathcal{M}_{i,j} u_{k;j} = \lambda_k u_{k;i}$. Since $\sum_{i=1}^{N} \mathcal{M}_{i,j} = 1$, taking the sum over the i components of the previous expression, we obtain: $\sum_{i=1}^{N} \sum_{j=1}^{N} \mathcal{M}_{i,j} u_{k;j} = \sum_{j=1}^{N} u_{k;j} = \lambda_k \sum_{j=1}^{N} u_{k;j}$. Therefore, either $\lambda_k = 1$, or $\sum_{j=1}^{N} u_{k;j} = 0$. We have that if $k > 1$, $\sum_{j=1}^{N} u_{k;j} = 0$.

We conclude that $\mathbf{y} = m\mathbf{z} \in$ span$\{\mathbf{u_2}, \ldots, \mathbf{u_N}\} \forall \mathbf{z} \in \mathbb{R}^N$. $\qquad \square$

Remark One needs note that finite numerical precision might cause troubles, so that one should be careful in eliminating any component along $\mathbf{u_1}$ at each before applying $\sum_{j=1}^{\infty} \mathcal{M}^j$. Note that we *must* use $\sum_{j=1}^{\infty} \mathcal{M}^j$ expression for $(1 - \mathcal{M})^{-1}$ in any code, because otherwise any software would give us automatically a `NaN` as error message.

Remark We wish to underline another method for avoiding the `NaN` problem mentioned above. Following [45], we introduce the fundamental matrix of the Markov chain as $\mathcal{Z} = (1 - \mathcal{M} + \mathcal{M}^{\infty})^{-1}$, where $\mathcal{M}^{\infty}$ is the limit matrix whose columns are all equal to $\mathbf{u}$. One can show that $\mathcal{Z}$ exists as the operation of inverse is well defined given the spectral properties of $\mathcal{M} - \mathcal{M}^{\infty}$ [39]. One can show that $\mathcal{M}^{\infty} m\mathbf{z} = 0 \forall \mathbf{z} \in \mathbb{R}^N$. Therfore, in all the previous Eqs. 16-22 we can substitute $(1 - \mathcal{M})^{-1} m = \sum_{j=0}^{\infty} \mathcal{M}^j m = \mathcal{Z}m = \sum_{j=0}^{\infty} (\mathcal{M} - \mathcal{M}^{\infty})^j m$.

Let's consider the problem of convergence of the expression in Eq. 18. We want to make sure that the L^1 norm of $\sum_{n=1}^{\infty} \epsilon^n \mathbf{w_n}$ does not diverge, and use this to find a bound for the value of ϵ. A simple way to approach this problem is to study the ratio of the L^1 norm of two consecutive terms in the previous series. Using Eqs. 15-16, we have:

$$\frac{\epsilon^n ||w_n||_1}{\epsilon^{n-1} ||w_{n-1}||_1} = \epsilon \frac{||(1 - \mathcal{M})^{-1} m w_{n-1}||_1}{||w_{n-1}||_1} \leq \epsilon \frac{||(1 - \mathcal{M})^{-1}||_1^* ||m w_{n-1}||_1}{||w_{n-1}||_1} \tag{23}$$

$$\leq \epsilon ||(1 - \mathcal{M})^{-1}||_1^* ||m||_1 \tag{24}$$

$$\leq \epsilon (1 - ||\mathcal{M}||_1^*)^{-1} ||m||_1 \tag{25}$$

where we use the submultiplicative property of the norm and we introduce a modified definition of the L^1 norm taking into account that the vector $m\mathbf{v} \in$ span$\{\mathbf{u_2}, \ldots, \mathbf{u_N}\} \forall \mathbf{v} \in \mathbb{R}^N$:

$$||\mathcal{Q}||_1^* = \sup_{v \in \text{span}\{\mathbf{u_2}, \ldots, \mathbf{u_N}\}, ||\mathbf{v}||_1 = 1} \frac{||\mathcal{Q}\mathbf{v}||_1}{||\mathbf{v}||_1}.$$

Using expression 25, we have that the perturbative expression converges if

$$|\epsilon| (1 - ||\mathcal{M}||_1^*)^{-1} ||m||_1 < 1 \to |\epsilon| < \epsilon_{max} = \frac{1 - ||\mathcal{M}||_1^*}{||m||_1}, \tag{26}$$

The previous expression provides an explicit bound for our calculations. We note that ϵ_{max} is finite because of the restriction imposed in the definition of the norm $|| \bullet ||^*$. Such a bound ensures also the invertibility of $(1 - \epsilon \Psi_1)^{-1}$. From the previous result, we find the following bound for the first order correction to the invariant measure:

$$||\epsilon w_1||_1 \leq \frac{\epsilon ||m||_1}{1 - ||\mathcal{M}||_1^*},$$

so that $||m||_1/(1 - ||\mathcal{M}||^*)$ can be though as a bound to the first order sensitivity of the measure to perturbations.

Using expression 24, we can derive a more generous bound for ϵ:

$$|\epsilon| < \epsilon_{max}^* = \frac{1}{||m||_1 ||(1 - \mathcal{M})^{-1}||_1^*} \geq \epsilon_{max}. \tag{27}$$

while $||m||_1 ||(1 - \mathcal{M})^{-1}||_1^*$ provides an additional (stricter) bound to the first order sensitivity. Note that in all the previous expressions we can substitute $||(1 - \mathcal{M})^{-1}||_1^*$ with $||\mathcal{Z}||_1$.

At this point, we wish to refer to previous results (see, e.g., [46]) providing bounds for the L^1 norm of the difference between the perturbed and unperturbed invariant measure:

$$||v_1 - u_1||_1 \leq \frac{\epsilon ||m||_1}{1 - \tau_{\mathcal{M}}(1)} \tag{28}$$

where $\tau_{\mathcal{M}}(1)$ is the so-called *ergodicity coefficient* [47] defined as:

$$\tau_{\mathcal{M}}(1) = \frac{1}{2} \sup_{i,j} ||\mathcal{M}(\mathbf{e_i} - \mathbf{e_j})||_1$$

with $\mathbf{e_i}$ indicating the unit vector having 1 at the ith entry and zero elsewhere. We remind that $\tau_1(\mathcal{M})$ is larger than any subdominant eigenvalue of $\mathcal{M}$, and $1/(1 - \tau_{\mathcal{M}}(1))$ can be taken as a definition of conditioning number of $\mathcal{M}$ [48]. Clearly if $\tau_{\mathcal{M}}(1)$ is close to 1, the bound given in Eq. 28 diverges. Note that $1/(1 - \tau_{\mathcal{M}}(1))$ is the bound to non-perturbative sensitivity mirroring the bound to the perturbative, linearized sensitivity given previously as $1/(1 - ||\mathcal{M}||_1^*)$. See also additional results presented in [49].

The sensitivity of the unperturbed measure to perturbations given in Eq. 28 can also be cast in terms $\rho_{\mathcal{M}}$, the smallest possible value for the constant controlling the rate of convergence of iterates $\mathcal{M}\mathbf{e_i}$, $\mathcal{M}^2\mathbf{e_i}$, ..., $\mathcal{M}^n\mathbf{e_i}$ to $\mathbf{u_1}$, so that $\forall n \in \mathbb{N}_+, \forall i \in 1, \ldots N$ we have that $||\mathcal{M}^n\mathbf{e_i} - \mathbf{u_1}||_1 \leq C\rho_{\mathcal{M}}^n$, $C \geq 1$ [46,48]. The sensitivity diverges as $\rho_{\mathcal{M}}$ approaches 1, i.e. when the unperturbed matrix has slow properties of convergence.

While the quantities $||\mathcal{M}||_1^*$, $\tau_{\mathcal{M}}(1)$, and $\rho_{\mathcal{M}}$ are indeed different, they all point to the fact that if the mixing rate of the unperturbed matrix $\mathcal{M}$ is slow—so that such quantities are close to 1 (so that $||(1 - \mathcal{M})^{-1}||_1^*$ and $||\mathcal{Z}||_1$ are very large)—then the sensitivity of the measure to perturbations is high. See in [21] a discussion of the link between slow mixing of a system and the presence of rough parameter dependence in its response to perturbations, with some examples of applications in a geophysical context.

Bringing together the results presented in Eqs. 9, 10 and in Eq. 27, we conclude that Eqs. 18–22 provide the exact expression for the invariant measure of the stochastic matrix $\mathcal{M} + \epsilon m$ $\forall \epsilon \in \{[-\epsilon_{max}^*, \epsilon_{max}^*] \cap [\epsilon_-, \epsilon_+]\}$.

3 Response Theory for Observables

Let's now look at the problem in terms of impact of the perturbation m on the expectation value of observables. Observables live in the dual space of the densities, and, given our convention, they are row vectors. They are approximated as having a constant value within each cell of the chosen partition of the phase space. The expectation value of the observable π with respect to a measure $\mathbf{w}$ can be written as $\langle \pi, \mathbf{w} \rangle$, where $\langle \bullet, \bullet \rangle$ denotes the scalar product. By definition, we have that $\langle \pi, A\mathbf{w} \rangle = \langle A^\top \pi, \mathbf{w} \rangle$, where $A^\top$ indicates the transpose (and adjoint, because we are studying real functions) of A.

Let's look at the change in the expectation value of the observable π as a result of $\mathcal{M} \to \mathcal{M} + \epsilon m$. We can write:

$$\langle \pi, \mathbf{v_1} \rangle = [\pi]_\epsilon = [\pi]_0 + \sum_{n=1}^{\infty} \epsilon^n \delta[\pi]_n \tag{29}$$

$$= \langle \pi, \mathbf{u_1} \rangle + \sum_{n=1}^{\infty} \epsilon^n \langle \Psi_n^\top \pi, \mathbf{u_1} \rangle \tag{30}$$

$$= \langle \pi, \mathbf{u_1} \rangle + \sum_{n=1}^{\infty} \epsilon^n \langle \left(\Psi_1^\top \right)^n \pi, \mathbf{u_1} \rangle, \tag{31}$$

where $[\pi]_0 = \langle \pi, \mathbf{u_1} \rangle$ is the expectation value of π in the unperturbed system, $[\pi]_\epsilon = \langle \pi, \mathbf{v_1} \rangle$ is the expectation value of π in the perturbed system, $\delta[\pi]_n$ is the nth order perturbation, which can be expressed as

$$\delta[\pi]_n = \lim_{\epsilon \to 0} \frac{1}{n!} \frac{\mathrm{d}^n}{\mathrm{d}\epsilon^n} \langle \pi, \mathbf{v_1} \rangle. \tag{32}$$

Moreover, $\Psi_n^\top$ is the nth order adjoint response operator, acting on the observables, which can be written as:

$$\Psi_n^\top = (\Psi_1^\top)^n = \prod_{j=1}^{n} m^\top \left(\sum_{k=1}^{\infty} \mathcal{M}^k \right)^\top. \tag{33}$$

We can also wrote Eq. 31 as:

$$\langle \pi, \mathbf{v_1} \rangle = \langle (1 - \epsilon \Psi_1^\top)^{-1} \pi, \mathbf{u_1} \rangle \tag{34}$$

$$= \left\langle \left(1 - \epsilon m^\top \left(\sum_{k=1}^{\infty} \mathcal{M}^k \right)^\top \right)^{-1} \pi, \mathbf{u_1} \right\rangle \tag{35}$$

$$= \langle (1 - \epsilon m^\top (1 - \mathcal{M}^\top)^{-1})^{-1} \pi, \mathbf{u_1} \rangle. \tag{36}$$

where the last two expressions provide the nonperturbative formulas.

Remark Equations 22 and 31 provide at all orders the response formulas for the discrete Markov process studied here. If we are constructing empirically the discrete phase space, we expect that different choices of the partitions, corresponding to different approximate representations of the full dynamics, will deliver different results in terms of response. Hence, our results can be model dependent, which is reasonable, as we are starting from a subjective choice on the way we approximate the phase space. In fact, one can empirically test the robustness of the obtained results against a set of given criteria by comparing whether the perturbations to a certain set of *relevant* observables weakly depend on the specific partition

used. We present a very preliminary (and encouraging) numerical study performed on the Lorenz '63 model [44] later in Sect. 5.

Moreover, as discussed in Sect. 1.4, if we construct finer and finer partitions of for studying the response of systems whose unperturbed dynamics features an SRB invariant measure (most notably in the case of Axiom A systems), and indeed if we follow the self-refining Markov partitions of the dynamics, our results should converge to the exact response theory built upon the true SRB measure.

One needs to note that Eq. 27 gives an estimate of the largest possible value of ϵ for *a given partition*, but we are are not sure whether the minimum over all the finer and finer partitions of ϵ_{max}^* is positive—this corresponds to imposing the uniform—in N—bound on the norm of $||(1 - \mathcal{M})^{-1}||_1^*$ or $||\mathcal{Z}||_1$.

In [39] it is shown that L^1 convergence of the finite state measure constructed using the Ulam method to the actual SRB measure is realized when $||\mathcal{Z}||_1$ grows asymptotically not faster than log N, where N is the number of states. The requirement we seem to have here for applying response theory here is unavoidably stricter because computing the response entails considering the expectation value of not necessarily well behaved observables, constructed through nontrivial operations of differentiation of the actual observables of which we want to study the sensitivity to perturbations, see Eq. 2 and [3–5]. This essential difficulty is exactly what motivates the point of view discussed in [18,50], where a delicate analysis of the relationship between tangent space of the unperturbed dynamics, the perturbation flow, and of the observable allow to set up a robust framework for the response theory.

Similarly, in our case, making the response theory work at practical level means having/choosing m and $\mathbf{u}$ in such a way that $||(1 - \mathcal{M})^{-1}||_1^*$ or $||\mathcal{Z}||_1$ grossly overestimates in terms of norm the effect of applying $(1 - \mathcal{M})^{-1}$ or equivalently $\mathcal{Z}$ in, e.g., Eq. 22. Additionally, a suitable choice of the observable π can help avoiding potential singularities in Eq. 36. In other terms, response theory can work much more easily once we get rid of or cure pathological cases.

4 Towards Continuous Time Dynamical Systems

We want to rephrase the previous results in the context of continuous time dynamical systems and derive some formulas previously presented in the literature concerning Axiom A systems. We coonsider a time continuous dynamical system of the form $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$ and study its response to the perturbation $\mathbf{F}(\mathbf{x}) \to \mathbf{F}(\mathbf{x}) + \epsilon \mathbf{X}(\mathbf{x})$. Correspondingly, as a result of the perturbation, the original invariant measure $\nu_0(d\mathbf{x})$ is changed into $\nu_\epsilon(d\mathbf{x})$. The Liouville equation describing the evolution of a given initial density of states $\rho(\mathbf{x})$ for the unperturbed system can be written as

$$\partial_t \rho(\mathbf{x}, t) = -\nabla \cdot (\mathbf{F}(\mathbf{x})\rho(\mathbf{x}, t)); \tag{37}$$

considering two instants of time separated by a small time interval dt, we have:

$$\rho(\mathbf{x}, t + dt) = \rho(\mathbf{x}, t) - dt\nabla \cdot (\mathbf{F}(\mathbf{x})\rho(\mathbf{x}, t)) = \mathcal{M}\rho(\mathbf{x}, t)$$
$$\mathcal{M} = 1 + dt\mathcal{F} \quad \mathcal{F} = -\nabla \cdot (\mathbf{F}(\mathbf{x})\bullet). \tag{38}$$

We understand that $\mathcal{M}$ is in this context the unperturbed Perron–Frobenius operator $\mathcal{L}_\epsilon^{dt}$ pushing forward the measure ρ from t to $t + dt$. When looking at the perturbed flow we have:

$$\rho(\mathbf{x}, t + dt) = \rho(\mathbf{x}, t) - dt\nabla \cdot (\mathbf{F}(\mathbf{x})\rho(\mathbf{x}, t)) - dt\epsilon\nabla \cdot (\mathbf{X}(\mathbf{x})\rho(\mathbf{x}, t))$$
$$= (\mathcal{M} + \epsilon m)\rho(\mathbf{x}, t), . \tag{39}$$

where

$$m = \mathrm{d}t\,\mathcal{X} \quad \mathcal{X} = -\nabla \cdot (\mathbf{X}(\mathbf{x})\bullet) \tag{40}$$

In this case, starting from Eq. 23, and considering that no normalization is applied to the perturbation operator, it is possible to propose a definition of ϵ^*_{max} for the continuous time dynamics taking inspiration from Eq. 27:

$$\epsilon^*_{max} = \frac{1}{||\mathcal{X}||_{\mathcal{B}}||\mathcal{F}^{-1}||^*_{\mathcal{B}}}, \tag{41}$$

such that the perturbative expansion converges if $\epsilon \leq \epsilon^*_{max}$, where $|| \bullet ||_{\mathcal{B}}$ describes the norm of the operator in the appropriate Banach space $\mathcal{B}$ it belongs to, while $|| \bullet ||^*_{\mathcal{B}}$ is such that the computation of the norm excludes the SRB measure. Note that ϵ^*_{max} is finite if both $||\mathcal{X}||_{\mathcal{B}}||$ and $||\mathcal{F}^{-1}||^*_{\mathcal{B}}$ are finite. This expression is admittedly tentative. As mentioned before, the problem of selecting appropriate functional spaces for constructing the response theory for Axiom A systems along the lines of studying the perturbations to the transfer operator requires a careful construction of suitable Banach spaces and of the related metrics [16,18,19] and is beyond the scope of this paper.[2]

4.1 Linear Response

We now want to derive the Ruelle response formulas for computing the linear correction to the invariant measure resulting from the perturbation. We write

$$\nu_\epsilon(\mathrm{d}\mathbf{x}) = \nu_0(\mathrm{d}\mathbf{x}) + \sum_{n=1}^{\infty} \epsilon^n \nu_n(\mathrm{d}\mathbf{x}), \tag{42}$$

where n indicates the order of perturbation. Let's first go back to the first order term in Eq. 15:

$$\epsilon \mathbf{w^1} = \epsilon \Psi_1 \mathbf{u_1} = \left(\sum_{k=1}^{\infty} \mathcal{M}^k \right) m \mathbf{u_1}. \tag{43}$$

Each term of the form $\mathcal{M}^k$ pushes forward up to time $t_k = k \times \mathrm{d}t$ what is positioned to its right. Summing over k in, in fact, amounts to looking forward in time. If we insert the definition of m given above, we get the integrating factor $\mathrm{d}t$, so that we obtain the following expression:

$$\nu_1(\mathrm{d}\mathbf{x}) = - \int_0^{\infty} \mathrm{d}t\, \nabla \cdot (\mathbf{X}(\mathbf{x}(t))\nu_0(\mathrm{d}\mathbf{x})), \tag{44}$$

where the evolution takes place according to the unperturbed system, and we have used the invariance of $\nu(\mathrm{d}\mathbf{x})$ with respect to such an evolution law.

By going into the dual space of the observables, we have that the change in the value of an observable $O(\mathbf{x})$ from time t to time $t + \mathrm{d}t$ in the unperturbed system can be written as:

$$\frac{\mathrm{d}}{\mathrm{d}t} O(\mathbf{x}(t)) = \mathbf{F}(\mathbf{x}) \cdot \nabla O(\mathbf{x}(t)), \tag{45}$$

so that

$$O(\mathbf{x}(t + \mathrm{d}t)) = O(\mathbf{x}(t)) + \mathrm{d}t\mathbf{F}(\mathbf{x}) \cdot \nabla O(\mathbf{x}(t)) = \mathcal{M}^\top O(\mathbf{x}(t)). \tag{46}$$

---

[2] Following [51], one might tentatively consider the norms of the operator acting between the Banach spaces $\mathcal{B}_{2,q}$ and $\mathcal{B}_{1,q+1}$.

where the operator $\mathcal{M}^\top = 1 + dt\,\mathcal{F}^\top = \mathbf{1} + dt\,\mathbf{F}(\mathbf{x}) \cdot \nabla(\bullet)$. Along the same lines, one derives that the perturbation operator $m^\top$ acting on the observable can be written as $m^\top = dt\,\mathcal{X}^\top = dt\,\mathbf{X}(\mathbf{x}) \cdot \nabla(\bullet)$. Furthermore, we introduce the following expansion for the expectation value of $O(\mathbf{x})$:

$$[O]_\epsilon = [O]_0 + \sum_{n=1}^\infty \epsilon^n \delta[O]_n, \tag{47}$$

where $[O]_\epsilon$ is the expectation value in the perturbed system, $[O]_0$ is the unperturbed expectation value, and the corrections are included in the summation.

Applying this expression to the first order term in Eq. 31–33:

$$\delta[\pi]_1 = \epsilon\langle \Psi_1^\top \pi, \mathbf{u}_1\rangle = \epsilon\left\langle \left(m^\top \left(\sum_{k=1}^\infty \mathcal{M}^k\right)\right)^\top \pi, \mathbf{u}_1\right\rangle. \tag{48}$$

we get:

$$\delta[O]_1 = \epsilon \int v_0(d\mathbf{x}) \int_0^\infty dt\,\mathbf{X}(\mathbf{x}) \cdot \nabla O(\mathbf{x}(t)) = \epsilon \int v_0(d\mathbf{x}) \int_0^\infty dt\,\Lambda S_0^t O(\mathbf{x}) \tag{49}$$

which is exactly the original version of Ruelle's linear response formula given in Eq. 3.

One needs to note that what in Ruelle's formulation is causality (time integration in the response starts from 0), in the context of the Markov matrices formalism followed here comes from the algebraic expansion of $(1 - \mathcal{M})^{-1}$. The issues of convergence mentioned in the original paper by Ruelle can be translated in the rate of mixing of the system as determined by the properties of $\mathcal{M}$ discussed in Sect. 2.1.

4.2 Higher Order Terms

We can repeat the same construction to derive the higher order perturbation terms in the case of the continuous time dynamical systems. Inserting in Eqs. 15, 16 the expression 38 for $\mathcal{M}$ and expression 40 for m, we obtain for the second order the following expression for the perturbation to the invariant density:

$$v_2(d\mathbf{x}) = \epsilon^2 \int_0^\infty dt_1 \int_0^\infty dt_2 \nabla \cdot \left(\mathbf{X}(\mathbf{x}(t_1)\nabla_{\mathbf{x}(t_1)} \cdot (\mathbf{X}(\mathbf{x}(t_1 + t_2)))\right) v_0(d\mathbf{x}), \tag{50}$$

while the expression for the nth order correction reads like

$$v_n(d\mathbf{x}) = (-1)^n \epsilon^n \int_0^\infty dt_1 \ldots \int_0^\infty dt_n \nabla \cdot \left(\mathbf{X}(\mathbf{x}(t_1)) \ldots \nabla_{\mathbf{x}(t_1 + \ldots t_{n-1})}\right.$$
$$\left. \cdot (\mathbf{X}(\mathbf{x}(t_1 + \ldots t_n)))\right) v_0(d\mathbf{x}), \tag{51}$$

Considering the adjoint problem and computing the higher order corrections to the expectation value of the observable O, we derive the general response formula proposed by Ruelle

$$\delta[O]_n = \epsilon^n \int v_0(d\mathbf{x}) \int_0^\infty dt_1 \ldots \int_0^\infty dt_n \Lambda S^{t_1} \ldots S^{t_{n-1}} \Lambda S^{t_n} O(\mathbf{x}), \tag{52}$$

as reported in Eq. 2.

5 A Very Basic Numerical Experiment

In order to make a (very) preliminary assessment of the potential of some of the ideas presented in this paper, we have focused on investigating some properties of the celebrated Lorenz '63 system [44]:

$$
\begin{aligned}
\dot{x} &= \quad \sigma(y - x) \\
\dot{y} &= x(\rho - z) - y \\
\dot{z} &= \quad xy - \beta z
\end{aligned}
\tag{53}
$$

where we have chosen the standard value for the parameters $\sigma = 10$, $\rho = 28$, and $\beta = 8/3$. We remark that such a system is not an Axiom A, but instead a singular hyperbolic system [52], which possesses a chaotic attractor and an invariant SRB measure [53]. In a previous publication [34], we have performed an analysis of the linear and nonlinear response of the Lorenz '63 to perturbations, extending a previous investigation by Reick [54], which makes us confident that response theory can be safely applied at all practical purposes also in this case. We consider the special case of time-indepedent perturbations to the dynamics resulting from substituting $\rho \rightarrow \rho + \epsilon$ in Eq. 53, so that the perturbation flow can be written as $\epsilon \mathbf{X}(\mathbf{x}) = [0 \quad \epsilon x \quad 0]^{\top}$.

We have then identified a 3-dimensional box $\mathcal{B}$ containing the attractor, defined as $\mathcal{B} = \{(x, y, z) \in \mathcal{R}^3 | x \in [-20, 20], \quad y \in [-30, 30], \quad z \in [-0, 50]\}$, and subdivided it, á la Ulam, in smaller boxes of identical size using a regularly spaced cartesian grid. We have considered partitions obtained using small boxes with linear dimension given by $dx = 2 \times j$, $dy = 3 \times j$, and $dz = 2.5 \times j$, along the three directions, with $j = 1, 2, 4$, see Fig. 1. This amounts to partitioning $\mathcal{B}$ into $8000/j^3$ smaller boxes. Note that our construction delivers a much lower resolution with respect to what used in, e.g., [55].

We run the model with standard values of the parameters choosing as initial condition $[1 \quad 1 \quad 1]^{\top}$ (in fact, given the global attractivity and ergodicity of the Lorenz attractor, any initial condition can be chosen), and, after discarding a transient of 1000 time units, which brings us safely into the asymptotic regime, we run the model for 50,000 time units with a simple Runge–Kutta 4th order adaptive scheme and obtain the output with time step of 0.001 time units. This takes less than 10 minutes in a today's commercial laptop with stan-
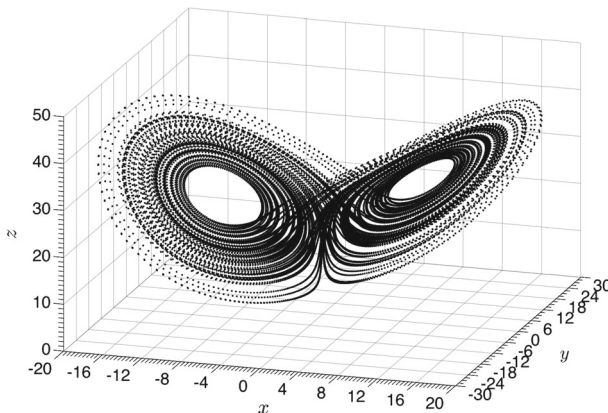

Fig. 1 Attractor of the Lorenz '63 system with indication of the cartesian grids used for constructing the partitions of its phase space. See *text*

Table 1 Expectation value of the observables x^2, y^2, z^2, and z and their linear response with respect to the perturbation $\rho \rightarrow \rho + \epsilon$

	$\langle x^2 \rangle$	$\langle y^2 \rangle$	$\langle z^2 \rangle$	$\langle z \rangle$	$\delta[x^2]_1$	$\delta[y^2]_1$	$\delta[z^2]_1$	$\delta[z]_1$
Lorenz '63 Model	62.9	81.2	630.0	25.6	2.8	3.7	50.3	1.01
MC, $j = 1$, $N_B^j = 770$	63.2	82.0	630.5	23.6	2.9	3.8	50.3	1.01
MC, $j = 2$, $N_B^j = 205$	64.3	84.2	632.2	23.6	3.0	3.5	49.7	1.02
MC, $j = 4$, $N_B^j = 56$	71.3	84.8	637.5	23.5	2.9	3.9	50.1	1.02

The first row refers to the integration of the Lorenz model given in Eq. 53. The other rows refer to the empirical discrete Markov chain constructed using boxes of different sizes. N_B^j refers to the number of states
The linear response of the observables defined in Eq. 32 has been obtained using Eq. 48. The derivative with respect to ϵ is estimated using finite differences with $\epsilon = 0.1$. See text

dard specifics using MATLAB®. We present results at such a low level of sophistication in order to clarify that the appracch proposed here is rather robust and of relatively simple implementation.

As the box-counting dimension or capacity of the attractor of the model given in Eq. 53 is $d_0 \sim 2.05$, we expect that the number of boxes B_k^j, $k = 1, \ldots, N_B^j$ needed to cover the attractor decreases $N_B^j \propto 1/j^{d_0}$. We obtain a slightly lower exponent ~ 1.9, which is perfectly acceptable as we are far from the asymptotic regime where the scaling given by d_0 is realized.

For each value of j, the boxes B_k^j define the discrete states ϕ_k^j, $k = 1, \ldots, N_B^j$. By counting the number of times the trajectory is included in each state ϕ_k^j and normalizing we derive experimentally the asymptotic normalized occupancies $\bar{u}_k^j$. Instead, by tracking the transitions between the various discrete states, we construct the estimate of the stochastic transition matrix $\mathcal{M}_{p,q}^j$ describing the probability that the state ϕ_q^j makes a transition to the state ϕ_p^j in one time step. By finding the eigenvector corresponding to the unique unitary eigenvalue of $\mathcal{M}_{p,q}^j$, we find the invariant measure, which agrees up to very high precision with the empirical occupancy rate $\bar{u}_k$ computed from the trajectory. As a first step, we evaluate the expectation values of four meaningful observables given by x^2, y^2, z^2, and z, as obtained from the time integration of the Lorenz model and from its discrete representation in terms of Markov chain. Table 1 shows that the agreement is rather good even when extremely coarse resolution is used.

We then show how to compute the response of the system to the perturbation due to the introduction of the vector field $\epsilon \mathbf{X}(\mathbf{x})$. We keep in mind that when continuous time dynamics is considered, there is a very simple linear relation between the perturbation flow and the corresponding perturbation to the Perron–Frobenius operator, see Eqs. 38–40.

Therefore, we repeat the the steps described above for the $\epsilon-$perturbed flow (we choose $\epsilon = 0.1$ in order to be on the safe side in terms of convergence), compute the new stochastic transition matrices $\mathcal{M}_{p,q}^{j,\epsilon}$, and derive the perturbation matrices $\epsilon m_{p,q}^j = \mathcal{M}_{p,q}^{j,\epsilon} - \mathcal{M}_{p,q}^j$. Once $m_{p,q}$ and $\mathcal{M}_{p,q}^j$ are known, we can use them to compute the response of the systems at all orders of nonlinearity using Eqs. 22 and 36.

One needs to note that because of the non-infinite integration time considered, of the non-infinitesimal perturbation applied, and of the somewhat arbitrary choice of the boxes, it can happen that the original and perturbed flow may be characterized by a different number of discrete states. We have observed such a difference only in the case $j = 1$, involving one

single extra state for the perturbed flow, with normalized relative occupancy ($\leq 10^{-6}$). This problem can be easily sorted out by imposing a cutoff and removing from the the discrete description all states with very low.

As discussed above, one needs to test accurately the well-posedness and convergence of the expansion in order to be sure to obtain meaningful results. This is not our goal at this stage given such a preliminary numerical test of our results. Therefore, we limit ourselves to the less ambitious yet interesting goal of computing the linear response defined in Eq. 32 for the observables indicated above, using Eq. 48. The results are reported in Table 1 and seem very encouraging. We have that the estimates of the response are very stable with respect to changes in the resolution of the boxes, and agree to a high degree of precision with the results one obtains by empirically evaluating the sensitivity of the observables with respect to the introduction of the perturbation flow using two integrations, as well, in the case of the z observable, with what reported in [34]. We note that the results are virtually unchanged if one uses instead of the high resolution time series with time step of 0.001 time units sparser observations corresponding to, e.g. a time step of 0.01 time units. Obviously, using a time resolution lower by a factor of s with respect to what considered here, one derives by tracking the transitions a stochastic transition matrix corresponding to the sth power of the one obtained at higher resolution. This does not affect the results as long as the sampling is much higher than the characteristic time scale of the system, which can be approximated in $\sim 1/\lambda_1 \sim 1.1$ time units, where λ_1 is the positive Lyapunov exponent of the system. On much longer time scales, instead, the stochastic matrix is quasi-degenerate, with all columns almost equal to the invariant measure

6 Conclusions

Taking the point of view of finite state Markov systems, we have been able to construct a perturbation theory for studying the impact of small perturbations to the background dynamics. While previous approaches focus on the constructing a theory able to account for the effect of adding small perturbations to the baseline flow, we focus on computing the change in the invariant measure and for the change in the expectation values of general observables (one problem being the adjoint of the other) occurring when the Markov transition matrix $\mathcal{M} \to \mathcal{M} + \epsilon m$.

The perturbation term ϵm has to be such that all the columns of the new stochastic matrix sum up to 1 and all entries are positive. All of our findings are obtained with rather simple linear algebra manipulations and using basic properties of the stochastic matrices. We can express the response as a perturbation series or, after suitable resummation, using compact exact formulas. We are also able to assess the convergence properties of the response theory by defining a value ϵ_{max}^* such that if $|\epsilon| \leq \epsilon_{max}^*$ the perturbative expansion converges. We have that the stronger is the mixing of the unperturbed system, the larger is the value of ϵ_{max}. These findings match well with previous results providing upper bounds to the sensitivity of stochastic matrices to perturbations.

Our results provide a direct algorithmic method for studying the response to perturbations for finite state Markov processes and have the advantage of allowing for an immediate and practical change of point of view between response theory seen in terms of changes of the invariant measure or in terms of changes in the expectation values of observables, by simply computing the transpose of the resulting finite dimensional linear operators. Our findings give

closed formulas for the linear and nonlinear response theory at all orders of perturbations through explicit matrix expressions that can be directly implemented in any coding language.

We can use our formulas to study the response to perturbations of finite state Markov processes constructed in order to have a simplified and treatable picture of a complex system. Given two different state spaces constructed using different finite partitions covering the attractor of the system, we cannot expect to obtain the same results for the change in the expectation value of a given observables. The results might indeed be model dependent, but this is the obvious price one has to pay because of the subjective choice of the reduced state space. An assessment of the robustness of the obtained results is key to applying our methods in the context of reduced models. Nonetheless, the extremely unsophisticated numerical study reported here on the Lorenz'63 model is quite encouraging at this regard, even if test should be made on much higher dimensional models.

If the underlying dynamics is Axiom A (or Axiom A equivalent, as in the cases where the chaotic hypothesis applies), one can impose conditions such that the response operators constructed using finer and finer partitions converge to to the actual corresponding response operators constructed on the SRB measure. Having in mind the Ulam method, the conditions are stricter than what needed in order to have convergence of the unperturbed measure, the basic reason being that Ruelle response operators correspond to nontrivial observables. One expects better convergence if the self-refining Markov partitions of the system are considered when constructing the finite state approximations.

Our results can be thought as intermediate steps at finite precision leading to the correct response formulas in the limit. One needs to add as a caveat that going from finite state to functional spaces is far from trivial and requires a high degree of mathematical precision, which is beyond the scopes of this paper. Nonetheless, the finite construction proposed here seems to somehow point at why some important mathematical issues emerge when the Perron–Frobenius operator formalism is considered in a continuum setting. In particular, the need for selecting suitable norms for vectors and linear operators in finite dimension points to the complex requirements in terms of functional spaces described in e.g. [51].

Interestingly, we can use the formulas obtained for finite state Markov processes to study the impact of perturbations to continuous time dynamical systems, after making a suitable identification between the considered transition matrices and the evolution operators for measures and observables. This operation is straightforward because there is a simple linear exact relation between the perturbation in the vector flow of the dynamical system and the perturbation in the Perron–Frobenius operator when infinitesimal time intervals are considered. As a result, we are able to derive in a very simple way previous formulas obtained studying the perturbations to the transfer operator as well as the original expressions proposed by Ruelle for the linear and higher order perturbations in the expectation values of observables. Using the results obtained in the finite state case, we propose a formula for the radius of expansion of the perturbative theory.

One can envision that in the case the underlying dynamics is discrete, there is not such a one-to-one correspondence between perturbations to the vector field and perturbations to the Markov transition matrix. This can be easily checked when constructing the perturbed Perron–Frobenius operator resulting from adding a ϵ correction to the vector field, which results into changes in the Perron–Frobenius operator at all orders in ϵ. Therefore, the perturbative expansion is different in the two cases. Agreement is instead found in the limit $\epsilon \to 0$, or, more practically, when we retain only the linear terms in ϵ perturbative expansion, i.e. when aiming only at the linear response function.

Future investigations will try, on the one side, to have a sharper mathematical look at the problem of going from finite to infinitely small partitions of the phase space, and, on the

other side, to delve in the numerical study of the effectiveness and efficiency of the proposed tools. Apart from testing the results on specific finite state Markov systems, we will test how robust the proposed methods are when studying finite state Markov processes that have been empirically constructed from time series of observations or of numerical simulations of high-dimensional complex systems. One may be led to hoping that it could be possible to have an accurate representation of the response of a high dimensional system to perturbations by constructing a smart finite state model well suited to studying specific observables of interest. Of course, in order to deal with the curse of dimensionality, one would like to be able to go beyond the Ulam method and deal with finite partition of reduced phase spaces where projection is applied on many or even most dimensions.

Our formulas may address the now long-standing problem of constructing suitable algorithms for studying the response of chaotic systems to perturbations. It is extremely hard to construct an algorithm for computing the (linear) response theory directly on the flow, because serious problems emerge when considering the contributions coming from the unstable directions in the tangent space. This might have great relevance for studying problems, like climate dynamics, where a direct construction of the response operator is especially challenging and slightly indirect methods have to be used [35] and a lot of effort has been devoted to defining the so-called atmospheric regimes and predicting their response to forcings [56].

References

1. Kubo, R.: Statistical-mechanical theory of irreversible processes. I. General theory and simple applications to magnetic and conduction problems. J. Phys. Soc. Jpn. **12**(6), 570–586 (1957)
2. Lucarini, V., Colangeli, M.: Beyond the linear fluctuation-dissipation theorem: the role of causality. J. Stat. Mech. **2012**(05), P05013 (2012)
3. Ruelle, D.: Differentiation of SRB states. Commun. Math. Phys. **187**(1), 227–241 (1997)
4. Ruelle, D.: Nonequilibrium statistical mechanics near equilibrium: computing higher-order terms. Nonlinearity **11**(1), 5–18 (1998)
5. Ruelle, D.: A review of linear response theory for general differentiable dynamical systems. Nonlinearity **22**(4), 855–870 (2009)
6. Ruelle, D.: Chaotic Evolution and Strange Attractors. Cambridge University Press, Cambridge (1989)
7. Lucarini, V., Sarno, S.: A statistical mechanical approach for the computation of the climatic response to general forcings. Nonlinear Process. Geophys. **18**, 7–28 (2011)
8. Colangeli, M., Lucarini, V.: Elements of a unified framework for response formulae. J. Stat. Mech. Theory E. **2014**, P01002 (2014)
9. Gallavotti, G., Cohen, E.G.D.: Dynamical ensembles in stationary states. J. Stat. Phys. **80**(5–6), 931–970 (1995)
10. Gallavotti, G.: Chaotic hypothesis: Onsager reciprocity and fluctuation-dissipation theorem. J. Stat. Phys. **84**(5–6), 899–925 (1996)
11. Vannitsem, S., Lucarini, V.: Statistical and dynamical properties of covariant Lyapunov vectors in a coupled atmosphere-ocean model—multiscale effects, geometric degeneracy, and error dynamics. *ArXiv e-prints*, October 2015

12. Gaspard, P.: Time-reversed dynamical entropy and irreversibility in markovian random processes. J. Stat. Phys. **117**, 599–615 (2004)

13. Gallavotti, G.: Stationary nonequilibrium statistical mechanics. In: Francoise, J.P., Naber, G.L., Tsun, T.S. (eds.) Encyclopedia of Mathematical Physics, vol. 3, pp. 530–539. Elsevier, Amsterdam (2006)

14. Gallavotti, G.: Nonequilibrium and irreversibility. Springer, New York (2014)

15. Baladi, V.: Positive Transfer Operators and Decay of Correlations. World Scientific, Singapore (2000)

16. Butterley, O., Liverani, C.: Smooth Anosov flows: correlation spectra and stability. J. Mod. Dyn. **1**(2), 301–322 (2007)

17. Liverani, C., Gouëzel, S.: Compact locally maximal hyperbolic sets for smooth maps: fine statistical properties. J. Differ. Geom. **79**, 433–477 (2008)

18. Baladi, Viviane: Linear response despite critical points. Nonlinearity **21**(6), T81 (2008)

19. Baladi, V.: Linear response, or else. *ArXiv e-prints*, August 2014

20. Engel, K.-J., Nagel, R.: One-parameter semigroups for linear evolution equations. Springer, New York (2001)

21. Chekroun, M.D., Neelin, D.J., Kondrashov, D., McWilliams, J.C., Ghil, M.: Rough parameter dependence in climate models and the role of ruelle-pollicott resonances. Proc. Natl. Acad. Sci. **111**(5), 1684–1690 (2014)

22. Tantet, A., Lucarini, V., Lunkeit, F., Dijkstra, H.A.: Crisis of the Chaotic Attractor of a Climate Model: A Transfer Operator Approach. *ArXiv e-prints*, July 2015

23. Hoffman, P.F., Kaufman, A.J., Halverson, G.P., Schrag, D.P.: On the initiation of a snowball earth. Science **281**, 1342 (2002)

24. Pierrehumbert, R.T., Abbot, D., Voigt, A., Koll, D.: Climate of the neoproterozoic. Annu. Rev. Earth Planet. Sci. **39**, 417 (2011)

25. Lucarini, V., Fraedrich, K., Lunkeit, F.: Thermodynamic analysis of snowball earth hysteresis experiment: efficiency, entropy production, and irreversibility. Q. J. R. Meteorol. Soc. **136**, 2–11 (2010)

26. Lucarini, V., Pascale, S., Boschi, V., Kirk, E., Iro, N.: Habitability and multistability in earth-like planets. Astronomische Nachrichten **334**(6), 576–588 (2013)

27. Intergovernmental Panel on Climate Change [Eds.: T. Stocker et al.]. Climate Change: The Physical Science Basis IPCC Working Group I Contribution to AR5. Cambridge University Press, Cambridge (2013). 2014

28. Abramov, R.V., Majda, A.J.: Blended response algorithms for linear fluctuation-dissipation for complex nonlinear dynamical systems. Nonlinearity **20**(12), 2793–2821 (2007)

29. Abramov, R.V., Majda, A.J.: New approximations and tests of linear fluctuation-response for chaotic nonlinear forced-dissipative dynamical systems. J. Nonlinear Sci. **18**, 303–341 (2008). doi:10.1007/s00332-007-9011-9

30. Eckmann, J.P., Ruelle, D.: Ergodic theory of chaos and strange attractors. Rev. Mod. Phys. **57**, 617–656 (1985)

31. Ginelli, F., Poggi, P., Turchi, A., Chaté, H., Livi, R., Politi, A.: Characterizing dynamics with covariant lyapunov vectors. Phys. Rev. Lett. **99**, 130601 (2007)

32. Eyink, G.L., Haine, T.W.N., Lea, D.J.: Ruelle's linear response formula, ensemble adjoint schemes and lvy flights. Nonlinearity **17**(5), 1867 (2004)

33. Wang, Qiqi: Forward and adjoint sensitivity computation of chaotic dynamical systems. J. Comput. Phys. **235**, 1–13 (2013)

34. Lucarini, V.: Evidence of dispersion relations for the nonlinear response of the Lorenz 63 system. J. Stat. Phys. **134**, 381–400 (2009). doi:10.1007/s10955-008-9675-z

35. Lucarini, V., Blender, R., Herbert, C., Ragone, F., Pascale, S., Wouters, J.: Mathematical and physical ideas for climate science. Rev. Geophys. **52**(4), 809–859 (2014)

36. Ragone, F., Lucarini, V., Lunkeit, F.: A new framework for climate sensitivity and prediction: a modelling perspective. Clim. Dyn. 1–13 (2015). doi:10.1007/s00382-015-2657-3

37. Ding, J., Li, T.Y., Zhou, A.: Finite approximations of markov operators. J. Comput. Appl. Math. **147**(1), 137–152 (2002)

38. Tantet, A., van der Burgt, F.R., Dijkstra, H.A.: An early warning indicator for atmospheric blocking events using transfer operators. Chaos **25**(3), 036406 (2015)

39. Froyland, G.: Approximating physical invariant measures of mixing dynamical systems in higher dimensions. Nonlinear Anal. **32**(7), 831–860 (1998)

40. Dellnitz, M., Junge, O.: On the approximation of complicated dynamical behavior. SIAM J. Numer. Anal. **36**(2), 491–515 (1999)

41. Blank, M., Keller, G., Liverani, C.: Ruelle–Perron–Frobenius spectrum for Anosov maps. Nonlinearity **15**(6), 1905 (2002)

42. Froyland, G.: On Ulam approximation of the isolated spectrum and eigenfunctions of hyperbolic maps. Discr. Contin. Dyn. Syst. **17**(3), 671–689 (2007)
43. Froyland, G.: Computer-assisted bounds for the rate of decay of correlations. Commun. Math. Phys. **189**(1), 237–257 (1997)
44. Lorenz, E.N.: Deterministic nonperiodic flow. J. Atmos. Sci. **20**, 130–141 (1963)
45. Schweitzer, P.J.: Perturbation theory and finite Markov chains. J. Appl. Probab. **5**(2), 401–413 (1968)
46. Mitrophanov, A.Yu.: Sensitivity and convergence of uniformly ergodic Markov chains. J. Appl. Probab. **42**, 1003–1014 (2005)
47. Seneta, A.: Explicit forms for ergodicity coefficients and spectrum localization. Linear Algebra Appl. **60**, 187–197 (1984)
48. Ipsen, I.C.F., Selee, T.M.: Ergodicity coefficients defined by vector norms. SIAM J. Matrix. Anal. Appl. **32**(1), 153–200 (2011)
49. Seneta, E.: Sensitivity of finite Markov chains under perturbation. Stat. Probab. Lett. **17**, 163–168 (1993)
50. Bódai, T.: Predictability of threshold exceedances in dynamical systems. *ArXiv e-prints*, August 2014
51. Liverani, C., Gouëzel, S.: Banach spaces adapted to Anosov systems. Ergodic Theory Dyn. Syst. **26**, 189–217 (2006)
52. Bonatti, C., Diaz, L.J., Viana, M.: Dynamics Beyond Uniform Hyperbolicity: A Global Geometric and Probabilistic Perspective. Springer, New York (2005)
53. Tucker, W.: The Lorenz attractor exists. C. R. Acad. Sci. Paris Sér. I Math. **328**(12), 1197–1202 (1999)
54. Reick, C.H.: Linear response of the Lorenz system. Phys. Rev. E **66**, 036103 (2002)
55. Froyland, G., Padberg, K.: Almost-invariant sets and invariant manifolds—connecting probabilistic and geometric descriptions of coherent structures in flows. Phys. D **238**, 1507–1523 (2009)
56. Corti, S., Molteni, F., Palmer, T.N.: Signature of recent climate change in frequencies of natural atmospheric circulation regimes. Nature **398**(6730), 799–802 (1999)