Stochastic Climate Dynamics: Random Attractors and Time-dependent Invariant Measures

Mickaël D. Chekroun[a,b,*], Eric Simonnet[d], Michael Ghil[a,b,c]

[a]*Department of Atmospheric Sciences and Institute of Geophysics and Planetary Physics, University of California, Los Angeles, CA 90095-1565, USA*
[b]*Environmental Research and Teaching Institute, École Normale Supérieure, 75231 Paris Cedex 05, France*
[c]*Geosciences Department and Laboratoire de Météorologie Dynamique (CNRS and IPSL), École Normale Supérieure, 75231 Paris Cedex 05, France*
[d]*Institut Non Linéaire de Nice (INLN)-UNSA, UMR 6618 CNRS, 1361, route des Lucioles 06560 Valbonne - France*

Abstract

This article attempts a unification of the two approaches that have dominated theoretical climate dynamics since its inception in the 1960s: the nonlinear deterministic and the linear stochastic one. This unification, via the theory of random dynamical systems (RDS), allows one to consider the detailed geometric structure of the random attractors associated with nonlinear, stochastically perturbed systems. A high-resolution numerical study of two highly idealized models of fundamental interest for climate dynamics allows one to obtain a good approximation of their global random attractors, as well as of the time-dependent invariant measures supported by these attractors; the latter are shown to be random Sinai-Ruelle-Bowen (SRB) measures. The first of the two models is a stochastically forced version of the classical Lorenz model. The second one is a low-dimensional, nonlinear stochastic model of the El Niño–Southern Oscillation (ENSO).

Keywords: Climate Dynamics, Dynamical Systems, El Niño, Random Dynamical Systems, Stochastic Forcing

The geometric [1] and ergodic [2] theory of dynamical systems represents a significant achievement of the last century. In the meantime, the foundations of the stochastic calculus also led to the birth of a rigorous theory of time-dependent random phenomena. Historically, theoretical developments in climate dynamics have been largely motivated by these two complementary approaches, based on the work of E. N. Lorenz [3] and that of K. Hasselmann [4], respectively.

It now seems clear that these two approaches complement, rather than exclude each other. Incomplete knowledge of small-, subgrid-scale processes, as well as computational limitations will always require one to account for these processes in a stochastic way. As a result of sensitive dependence on initial data and on parameters, numerical weather forecasts [5] as well as climate projections [6] are both expressed these days in probabilistic terms. In addition to the intrinsic challenge of addressing the nonlinearity along with the stochasticity of climatic processes, it is thus more convenient — and becoming more and more necessary — to rely on a model's (or set of models') probability density function (PDF) rather than on its individual, pointwise simulations or predictions.

We show in this paper that finer, highly relevant and still computable statistics exist for stochastic nonlinear systems, which provide meaningful physical information not described by the PDF alone. These statistics are supported by a random attractor that extends the concept of a strange attractor [3, 7] and of its invariant measures [2] from deterministic to stochastic dynamics.

The attractor of a deterministic dynamical system provides crucial geometric information about its asymptotic regime as $t \to \infty$, while the Sinaï-Ruelle-Bowen (SRB) measure provides, when it exists, the

---

statistics of the flow over this attractor [2, 8]. These concepts have been applied to climate dynamics — across a full hierarchy of models, from conceptual "toy" models via so-called intermediate models and all the way to high-resolution general circulation models (GCMs) — as well as to the related uncertainties [9, 10, 11]. Recent applications of ergodic theory to the problem of climate sensitivity, in the context of deterministic models of small and intermediate complexity, include [12, 13].

On the stochastic side, the crucial field of modeling subgrid-scale phenomena has been increasingly moving towards stochastic "parameterizations" [14, 15]. Such parameterizations have been studied in terms of their impact on the successful simulation of certain physical processes in GCMs, but not in terms of their global impact on model behavior. At a more fundamental level, the climate system is an open system and subject to variable forcing in time. The long-term effects of time-dependent forcing, whether deterministic or stochastic, have only started to be studied; examples include Quaternary glaciations and their relationship to orbital forcing [16, 17] or the interaction between the seasonal forcing and intrinsic variability in the Tropical Pacific [18, 19].

During the past two decades, the mathematical theory of random dynamical systems (RDS) [20] and of nonautonomous dynamical systems [21] has made substantial progress in describing the asymptotic behavior of open systems, subject to time-dependent forcing. The pertinent mathematical literature, however, is fairly technical and opaque. Its concepts and methods have, therefore, not become widely understood and applied to the physical sciences in general and to climate dynamics in particular; see [22] and references therein.

The main objective of this paper is twofold: (i) to introduce the key concepts and tools of RDS theory — from the point of view of ergodic theory [2, 8] — to a wider audience in the geosciences and macroscopic physics; and (ii) to present novel results for two highly idealized models of fundamental interest for climate dynamics. The first is a stochastically forced version of the Lorenz [3] model: We provide detailed geometric structure and novel statistical information by using a highly accurate numerical approximation of its global random attractor and of the invariant measures supported thereon; furthermore, these measures are shown to be random SRB measures [23]. The second one is a low-dimensional, nonlinear stochastic El Niño–Southern Oscillation (ENSO) model [24]; here we show how the information conveyed by its random attractor and invariant measures allow one to better understand the qualitative behavior of this model and to refine its physical interpretation.

In order to keep the presentation accessible to the intended audience, we refer for technical details to L. Arnold [20] and H. Crauel [25]. For the sake of brevity, statements about the rigorous existence of the mathematical objects being described are typically omitted.

1. Noise effect on model statistics: A change of paradigm

What is the the effect of random perturbations on a nonlinear deterministic system's phase portrait? To address this issue, especially in the case of a deterministically chaotic system, we introduce herewith the appropriate framework.

1.1. The classical forward approach

To analyze the effect of the noise on the invariant measure supported by the deterministic system's attractor, the traditional stochastic approach is to seek the fixed point of the associated Markov semigroup, *i.e.* to find stationary solutions of the Fokker-Planck equation. These solutions correspond precisely to the system's stationary measures. Numerically, it is most often easier to integrate the system forward in time, perform ensemble or time averages and call the resulting object the "PDF."

When a deterministic system is perturbed by noise, it is often observed that the support of such a numerically obtained PDF corresponds to a (small) phase-space neighborhood of the deterministic attractor; in particular, the topological structure of the deterministic attractor becomes fuzzy. Such an approach provides, therefore, purely statistical information, without a close link with the attractor's geometry. Even so, the effect of the noise can result in surprising changes, especially when the deterministic system is neither hyperbolic [26] nor stochastically stable [8].

The RDS approach is based on a drastically different view. Its fundamental objects are the random invariant measures of the dynamics rather than the stationary ones of the Markov semigroup. These invariant

2

measures are supported by a well-defined attractor, as will be explained below. In this approach, instead of integrating forward in time, the system is run from a distant point s in the past until the present time t, where it is "frozen." We refer to this as the *pullback approach*. Remarkably, by looking at the system in this way, the topological structures related to the stochastic dynamics emerge naturally and, even more surprisingly, there is no fuzziness in them. RDS theory thus reconciles the ergodic and geometric approaches in the stochastic context. We explain next the pullback approach, what an RDS and a random attractor are, and discuss the invariant measures such an attractor supports.

1.2. The pullback approach

This approach adopts a pathwise analysis, rather than the previous one, based on an ensemble of realizations. At first glance, this angle of attack may appear more laborious and less direct in providing statistical information. In fact, it yields much more detailed insights, along with the PDF, as will be seen below.

To understand this relatively novel approach, we first explain heuristically the concept of pullback attractor in the context of a deterministic, but nonautonomous dynamical system. For simplicity, we consider a finite-dimensional system, written in the form,

$$\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}), \tag{1}$$

where the law $\mathbf{f}$ governing the evolution of the state $\mathbf{x}$ depends explicitly on time t.

A simple example from climate dynamics is given by the oceans' wind-driven circulation [27]. The effect of the atmosphere on the mid-latitude oceans at zero order would be modeled by a time-independent forcing that yields an autonomous system [22]. At the next order, however, taking into account the seasonal cycle in the winds, the forcing would become time periodic and the system thus nonautonomous [28]. As the degree of realism increases — unless one were to switch to a fully coupled atmospheric-ocean model — the time-dependent aspects would become more and more elaborate and involve not only the forcing but also various coefficients, which eventually will include stochastic effects at some point. Another example that will be illustrated in the numerical section of this paper is an ENSO model, in which wind bursts are modeled stochastically; see *e.g.* [29, 30, 31].

Stochastic models, in particular, are nonautonomous, rough rather than smooth, and are indexed by the realizations of the random processes involved.

For such models, we ask the following question:

Q: For a fixed realization ω, and at a fixed time t — the time at which the system is observed — how does the "stochastic flow" transform the Lebesgue measure on the phase space, assuming we have started the system in a remote past $s << t$?

It is this question that motivates and guides our exposition, and our subsequent results.

To study **Q**, we need to recall several concepts. First, let us denote by $\varphi(s,t)\mathbf{x}$ the solution of (1) at time t, where $\mathbf{x}$ is the initial state at time $s \leq t$, *i.e.* $\varphi(s,s)\mathbf{x} = \mathbf{x}$. In general, the operator $\varphi(s,t)$ generates a two-parameter semi-group that provides a two-time description of the system's evolution, while in the autonomous case a one-parameter semi-group suffices to entirely determine this evolution. In the latter case, the system's evolution is invariant with respect to translation in time, *i.e.* $\varphi(s,t)\mathbf{x} = \varphi(t-s)\mathbf{x}$, while in the former, the time at which initial data are prescribed is of paramount importance. Thus, in the nonautonomous case, the limiting behavior when $s \to -\infty$ and t is fixed may differ from the one obtained in the forward situation, with $t \to \infty$ and s fixed, whereas in the autonomous case the two limits represent the same asymptotic behavior, due to the translation invariance of $\varphi(s,t) = \varphi(t-s)$.

To illustrate the fundamental character of this distinction, consider the simple scalar version of (1): $\dot{x} = -\alpha x + \sigma t$, with $\alpha > 0$, and $\sigma \geq 0$. We denote again by $\varphi(s,t)x_0$ the solution at time t, assuming that $x(s) = x_0$ at $s \leq t$. The forward approach yields blow-up as $t \to +\infty$ for any x_0, while an easy computation shows that $|\varphi(s,t)x_0 - \mathcal{A}(t)| \to 0$ as $s \to -\infty$, for all t and x_0, with $\mathcal{A}(t) := \sigma(t - 1/\alpha)/\alpha$.

It can be shown further that $\mathcal{A}(t)$ is invariant under the dynamics, *i.e.* $\varphi(s,t)\mathcal{A}(s) = \mathcal{A}(t)$, for every $s \leq t$. We have therefore exhibited a family of limiting objects $\mathcal{A}(t)$, which exist in actual time t rather than asymptotically in the future, and which convey the effect of the dissipation due to the term $-\alpha x$. In this example, $\mathcal{A}(t)$ is simply a time-dependent point that attracts all the initial data.

More generally, in the forced-dissipative case, one obtains for all t, by letting $s \to -\infty$, a collection $\bigcup_{t \in \mathbb{R}} \mathcal{A}(t)$ of objects $\mathcal{A}(t)$ that depend on time t; this collection is called a *pullback attractor*. Each $\mathcal{A}(t)$ may be more complicated than a point, and attract some subsets of initial data taken in the asymptotic past. In rigorous terms, a family of objects $\bigcup_{t \in \mathbb{R}} \mathcal{A}(t)$ in a finite-dimensional, complete metric phase space X is a pullback attractor if it satisfies the two following conditions:

(I) For all t, $\mathcal{A}(t)$ is a compact subset of X and is invariant with respect to the dynamics, namely,

$$\varphi(s,t)\mathcal{A}(s) = \mathcal{A}(t), \text{ for every } s \le t; \text{ and}$$

(II) for all t, pullback attraction occurs:

$$\lim_{s \to -\infty} d_X(\varphi(s,t)B, \mathcal{A}(t)) = 0, \text{ for all } B \in \mathbb{B}. \tag{2}$$

In (2), d_X denotes the Hausdorff semi-distance between two subsets in X, and B belongs to a collection $\mathbb{B}$ of sets in X that may itself exhibit some time dependence [20, 21].

A fundamental property of a system's pullback attractor is that it may support physically interesting invariant measures. In the present paper, this aspect is discussed in greater detail for stochastically perturbed systems. We provide here a simple deterministic, but nonautonomous illustration.

Going back to $\dot{x} = -\alpha x + \sigma t$, one can show that every x−interval in $\mathbb{R}$, taken at a time $s < t$, shrinks onto $\mathcal{A}(t)$ as $s \to -\infty$. In terms of measure, one can say that the Dirac measure $\delta_{\mathcal{A}(t)}$, supported by $\mathcal{A}(t)$, "pullback attracts" at time t the Lebesgue measure on $\mathbb{R}$. By invariance of $\mathcal{A}(t)$, $\delta_{\mathcal{A}(t)}$ is thus a globally stable, time-dependent, invariant measure of our scalar nonautonomous system, just as δ_0 is for the autonomous system $\dot{x} = -\alpha x$, when $\sigma = 0$.

In general, the simplest and most fundamental measures that are invariant under the dynamics are precisely these time-dependent invariant Dirac measures. For a nonautonomous system, they replace the role played by fixed points for autonomous ones: time dependence usually prevents the system from being at rest and traditional fixed points become the exception, rather than the rule.

It follows that, if a nonautonomous dynamical system involves dissipation, we may wish to consider its asymptotic behavior in a pullback sense. Indeed, dissipative properties, coupled with time-dependent forcing, lead to the existence of a dynamical object $\bigcup_{t \in \mathbb{R}} \mathcal{A}(t)$, rather than a static one; this pullback attractor describes the asymptotic regime at time t, by considering the system initialized in the asymptotic past. Furthermore, this object supports invariant measures that are time-dependent by nature. At this stage, we have traveled half the road that leads to answering question **Q**. We need now to consider the random case, in order to travel the other half.

1.3. The RDS approach

When the time-dependent forcing is random, the pullback attractor becomes a random pullback attractor or *random attractor* for short. This concept, however, is subtler than its "deterministic cousin" just discussed, and needs further clarification. In the 1980's, Kunita [32], among others, took an important step toward a geometrical description of "stochastic flows" by providing a pathwise two-parameter framework for describing the stochastic flows generated by fairly general stochastic differential equations (SDEs).

Roughly speaking, this framework allows one to show that, for almost all realizations ω living in some probability space Ω, the evolution in the phase space X of a stochastic system from time $s < t$ to time t is described by a two-parameter family of transformations $\varphi(s,t;\omega)$. It is tempting, therefore, to adopt the pullback approach just described above in an ω-parameterized version, in order to introduce the analog of a pullback attractor into the stochastic context. The problem with such a naive generalization is that the resulting object $\bigcup_{t \in \mathbb{R}} \mathcal{A}(t;\omega)$ does not exhibit any *a priori* relation between distinct realizations ω. As a matter of fact, this is one of the reasons why traditional approaches consider only expectations, *i.e.* ensemble means, rather than the stochastic flows.

The remedy to this problem comes from physical considerations. For an experiment to be repeatable, one has to have a reasonable description of its random aspects. These aspects may change in time, and thus the noise has to be modeled as a time-dependent stochastic process with certain known properties.

Representing mathematically such a stochastic process starts with a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathcal{F}$ is a σ-algebra of measurable subsets of Ω, called "events," and $\mathbb{P}$ is the probability measure [20]. Parameterizing noise by time, or equivalently, parameterizing the probability space by time, means that we should be able to connect the state ω of the random environment at time $t = 0$, say, with its state after a time t has elapsed; we call this connection $\theta_t\omega$ and set, of course, $\theta_0\omega = \omega$. This set-up establishes a map $\theta_t : \Omega \mapsto \Omega$ for all times t.

In practice, one requires $(t, \omega) \mapsto \theta_t\omega$ to be measurable, and to satisfy the one-parameter group property $\theta_{s+t} = \theta_s \circ \theta_t$ for any s and t; along with $\theta_0 = \mathrm{id}_\Omega$, these requirements lead to a time-dependent family $\{\theta_t\}$ of invertible transformations of Ω that keeps track of the noise. Furthermore, one requires a stationarity condition, namely that the statistics of the external noise are invariant under θ_t; mathematically, this means that the probability measure $\mathbb{P}$ is preserved by θ_t, *i.e.* $\theta_t\mathbb{P} = \mathbb{P}$.

No other properties are needed in general: even ergodicity of θ_t appears to be secondary [20]. A simple example is given by the class of two-sided Wiener processes that are so pervasive in SDEs, where $\Omega = \{W \in C_0(\mathbb{R}), W(0) = 0\}$, while θ_t is the shift operator that acts according to $W_s(\theta_t\omega) = W_{s+t}(\omega) - W_t(\omega)$, and thus preserves the Wiener measure. In this particular case, ergodicity holds [20].

At this stage, we realize that the evolution of a stochastic system $\varphi(s, \tau; \omega)$ for $\tau > s$ can be derived from a simpler description $\Phi(t, \omega)$; this operator is parameterized by a single time variable t and it ascribes the state of the system at any time $t > 0$, starting at time $s = 0$ and given a realization ω that characterizes the state of the noise at time 0. Indeed, given the one-parameter mapping $\Phi(t, \omega)$, it is easy to obtain the evolution of the system from any time s to any time τ via the two-parameter mapping $\varphi(s, \tau; \omega) = \Phi(\tau - s, \theta_s\omega)$.

This simplification allows one to understand why a pathwise approach to modeling the noise may be reasonable: a system influenced by an external stochastic process for a single realization ω can be interpreted via the *driving system* $\{\theta_t\}$ as wandering along a path $\theta_t\omega$ in Ω and thus may provide additional statistical information to the modeler. For this to hold true, however, one must show that the system's behavior can be described in a self-consistent manner along such a path $\theta_t\omega$; this is the case if the stochastic dynamics $\Phi(t, \omega)$ satisfies the so-called *cocycle property*, namely $\Phi(t + s, \omega) = \Phi(t, \theta_s\omega) \circ \Phi(s, \omega)$, [20, 22]. Remarkably, RDS theory shows that the cocycle property comes for free for a very large class of stochastic systems that includes standard SDEs; this result is mainly due to the group property of $\{\theta_t\}$ [20]. In short, fairly general SDEs generate RDSs.

The cocycle concept lies at the core of RDS theory, thus extending Kunita's results [32]. It appears that the system $\Theta(t) : (\omega, x) \mapsto (\theta_t\omega, \Phi(t, \omega)x)$, also referred to as *a skew product*, is indeed a dynamical system on the extended phase space $\Omega \times X$, *i.e.* (probability space)×(phase space). We can thus deal with a genuine stochastic flow on this extended space, where $\Theta(t + s) = \Theta(t) \circ \Theta(s)$ holds. The notion of stochastic flow mentioned in question **Q** is therewith defined rigorously. One thus recovers much of the classical theory for autonomous flows, although there is an additional difficulty. This difficulty arises because, in the extended phase space $\Omega \times X$, one can only use measurability concepts on Ω, without any topological tools. The pair (Φ, θ) is called an RDS [20].

1.4. *Random attractor*

With these concepts and tools in hand, we are now in a position to extend the notion of pullback attractor to the stochastic context. To do so, consider first the classical Langevin equation,

$$\mathrm{d}x = -\alpha x \mathrm{d}t + \sigma \mathrm{d}W_t, \text{with } \alpha > 0 \text{ and } \sigma \neq 0. \tag{3}$$

Using stochastic calculus, the properties of the Wiener process and the definition of θ_t as a shift introduced above, it follows that $|\Phi(t, \theta_{-t}\omega)x - a(\omega)| \to 0$ as $t \to +\infty$; here $a(\omega) := \sigma \int_{-\infty}^0 \exp(\alpha\tau)\mathrm{d}W_\tau(\omega)$ and $\Phi(t, \omega)$ as defined above. The random variable a is even invariant under the dynamics, *i.e.* $\Phi(t, \omega)a(\omega) = a(\theta_t\omega)$, for all $t \geq 0$ and all $\omega \in \Omega$.

We have thus exhibited a family of random invariant objects — each of which is a single point $a(\omega)$ — that describe the possible states that can be observed in the present, at $t = 0$, whatever the state it occupied in the infinitely distant past, at $t = -\infty$. This contraction of the Lebesgue measure under the stochastic

flow results, once more, from the effect of the dissipation that is "experienced" by the system in a pullback sense.

More generally, if a random subset $\mathcal{A}(\omega)$ of X satisfies certain measurability conditions [20], and furthermore

(I') $\mathcal{A}(\omega)$ is a compact subset of X and Φ-invariant, *i.e.*, for each $t \geq 0$, $\Phi(t, \omega)\mathcal{A}(\omega) = \mathcal{A}(\theta_t\omega)$; and

(II') $\mathcal{A}(\omega)$ is attracting in the pullback sense, *i.e.*,

$$\lim_{t \to +\infty} d_X(\varphi(t, \theta_{-t}\omega)B(\theta_{-t}\omega), \mathcal{A}(\omega)) = 0, \text{ for all } B \in \mathbb{B},$$

then $\bigcup_{\omega \in \Omega} \mathcal{A}(\omega)$ provides the complete picture of the only present states of the system that are likely to be observed. The second condition holds almost surely, with respect to the measure $\mathbb{P}$, while the $B(\omega)$ are now random subsets of X; see [20] for further details.

The resulting random compact set $\bigcup_{\omega \in \Omega} \mathcal{A}(\omega)$ is called a *random attractor*; it is also called a *strong attractor* since the convergence of remote initial data to the attractor holds almost surely for the Hausdorff semi-metric of the phase space X [20, 33, 38, 39]. Moreover, when $\{\theta_t\}$ is ergodic, then knowing $\bigcup_{t \in \mathbb{R}} \mathcal{A}(\theta_t\omega)$ yields $\bigcup_{\omega \in \Omega} \mathcal{A}(\omega)$, and vice-versa. In other words, knowing the random attractor along one path ω yields all the possible states of the attractor at a given t.

RDS theory also provides a natural link between the forward and pullback approach. Pullback attraction involves convergence almost surely (see above) and it implies weak convergence forward, that is, convergence in probability only. A *weak attractor*, as opposed to the strong one above, is defined similarly, except that now one requires only that

$$\lim_{t \to \infty} d_X(\Phi(t, \omega)B(\omega), \mathcal{A}(\theta_t\omega)) = 0 \text{ in probability [33]}.$$

A strong attractor is always a weak attractor, but the converse is obviously wrong [20, 33]. The argument runs as follows: $\{\theta_t\}$ is measure-preserving and thus one has $\mathbb{P}\{\omega : \Phi(t, \omega)x \in D\} = \mathbb{P}\{\omega : \Phi(t, \theta_{-t}\omega)x \in D\}$, for any $x \in X$ and D any measurable subset of X. If the limit of the right-hand side exists as $t \to +\infty$, then the left-hand side converges as well; that is, $\Phi(t, \omega)x$ converges, but only in probability. There is a duality between fuzziness from the present into the future versus "determinism" from the past up to the present. This duality arises from the asymmetric way in which time is addressed in the RDS approach, through the concepts of *past and future of an RDS*; see Sec. 1.7 of [20].

1.5. Random invariant measures

So far, we have focused on the more appealing and intuitive aspects of RDS theory, with random attractors playing the key role. The fundamental objects in RDS theory, though, are in fact the random invariant measures; they are intimately linked to random attractors in forced-dissipative systems.

When a global random attractor $\mathcal{A}$ exists, it supports all the invariant measures μ, and hence $\mu(\mathcal{A}) = 1$; this result is indeed similar to the deterministic situation. The invariant measures μ here are defined on the product space $\Omega \times X$ and invariance is with respect to the skew product Θ defined above, so that $\Theta(t)\mu = \mu$ for all t [20, 25].

The invariant measures μ "lift" the probability measure $\mathbb{P}$, defined on Ω, into the extended phase space $\Omega \times X$, so that the projection of μ on Ω equals $\mathbb{P}$. It is much more convenient, though, to work in the phase space X, rather than in the product space. Invariance of μ in X corresponds to the use of random measures $\omega \mapsto \mu_\omega$ on X called *sample measures* [23]; they are also known as *factorized or disintegrated measures* [20, 25].

Under very general conditions, one can show there is a one-to-one correspondence between any μ_ω on X and any μ on the product space whose Ω–projection equals $\mathbb{P}$; symbolically $\mu(d\omega, dx) = \mu_\omega(dx)\mathbb{P}(d\omega)$ [20]. In particular, to say that $\mu(\mathcal{A}) = 1$ is equivalent to $\mu_\omega(\mathcal{A}(\omega)) = 1$; *i.e.*, each sample of $\mathcal{A}$ supports the sample measure μ_ω. The invariance of μ can now be written, since we take θ_t here to be invertible, as

$$\Phi(t, \omega)\mu_\omega = \mu_{\theta_t\omega}, \text{ almost surely with respect to } \mathbb{P}.$$

6

Note that the time-dependent solutions of the Fokker-Planck equation must not be confused with the time-dependent sample measures $\mu_{\theta_t\omega}$ discussed herein. The main difference is apparent by returning now to question **Q**.

Indeed, the answer to our key question **Q** is now simply that the regions where the stochastic flow ends up at time t — for a realization ω of the system's random aspects — are determined by $\mathcal{A}(\omega)$ and distributed according to probabilities given by the invariant sample measure μ_ω. A condition for this to hold is that the measure be *physical*, *i.e.* that it satisfy almost surely (with respect to $\mathbb{P}$) the key identity:

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t G\circ\Phi(s,\theta_{-s}\omega)x\,\mathrm{d}s = \int_{\mathcal{A}(\omega)} G(x)\mu_\omega(\mathrm{d}x),\tag{4}$$

for almost every $x\in X$ (in the Lebesgue sense), and for every continuous observable $G:X\to\mathbb{R}$. Equation (4) is a direct generalization to the RDS framework of the concept of physical measure from the autonomous deterministic setting [2, 8]. In the next section, we discuss a particular class of physical measures of interest, namely *random SRB measures*. These measures are a sample version of the classical SRB measures found in the autonomous context; see [23] and the next section here. We focus now on the simplest invariant measures that are fundamental in RDS theory.

Going back to the Langevin equation (3), since the random point $a(\omega) = \sigma\int_{-\infty}^0 \exp(\alpha\tau)\mathrm{d}W_\tau(\omega)$ is the global random attractor of that system, we get that the random Dirac $\delta_{a(\omega)}$ supported by each point $a(\omega)$ is a globally stable invariant measure, which pullback attracts the Lebesgue measure of the real line. This result can be of course generalized to a higher-dimensional Langevin equation $\mathrm{d}\mathbf{x} = L\mathbf{x} + \mathrm{d}\mathbf{W}_t$, where L has good dissipative properties and where the random attractor $\mathcal{A}$ becomes the union of random vectors $\mathbf{a}(\omega)$ such that $\Phi(t,\omega)\mathbf{a}(\omega) = \mathbf{a}(\theta_t\omega)$ for all $t\geq 0$ and all $\omega\in\Omega$. For a more general RDS, a (measurable) random variable $\mathbf{a}:\Omega\to X$ is called a *random fixed point* of the system when this last equality holds. These random fixed points are the analogs of the steady states found in the autonomous setting.

When all the Lyapunov exponents are negative, then all the initial states are attracted to a single random fixed point. This well-known phenomenon occurs for general RDSs and is not restricted to linear dynamics with additive noise [35, 36]; it is sometimes called *synchronization* [20, 22]. When the random attractor is more complex than a random point, one may observe *intermittency*: for each realization ω, two arbitrary trajectories are either synchronized or not during variable time intervals. This on-off synchronization occurs exponentially fast; see Fig. 6 of [42]. It appears that the ENSO model studied below exhibits such intermittent behavior.

2. Numerical results and their RDS analysis

We consider here the Lorenz model [3] and the ENSO model of Timmerman and Jin [31]. The two models have three degrees of freedom each and we perturb them by linear multiplicative white noise [20] for the Lorenz model or by including white noise in the nonlinearity of the ENSO model.

2.1. A stochastically perturbed Lorenz model

In the deterministic context, geometric models were proposed in the 1970s [47] to interpret the dynamics observed numerically by E. N. Lorenz in [3]. These geometric models attracted considerable attention and it was shown that they possess a unique SRB measure [34, 46], *i.e.*, a time-independent measure that is invariant under the flow and has conditional measures on unstable manifolds that are absolutely continuous with respect to Lebesgue measure [2]. This result has been extended recently to the Lorenz flow [3] itself, in which the SRB measure is supported by a strange attractor of vanishing volume [48, 49].

Even though this result was only proven recently, the existence of such an SRB measure was suspected for a long time and has motivated numerous numerical methods to compute a PDF associated with the Lorenz model [3], by filtering out the stable manifolds; *e.g.* [44, 45] and references therein. The Lorenz attractor is then approximated by a 2-dimensional manifold, called the *branched manifold* [47], which supports this PDF. Based on such a strategy, the authors in [44] showed that the stationary solution of the Fokker-Planck equation for the Lorenz model [3] perturbed by white noise possesses a density with two components: the

7

PDF of the deterministic system supported by the branched manifold and a narrow Gaussian distribution transversal to that manifold.
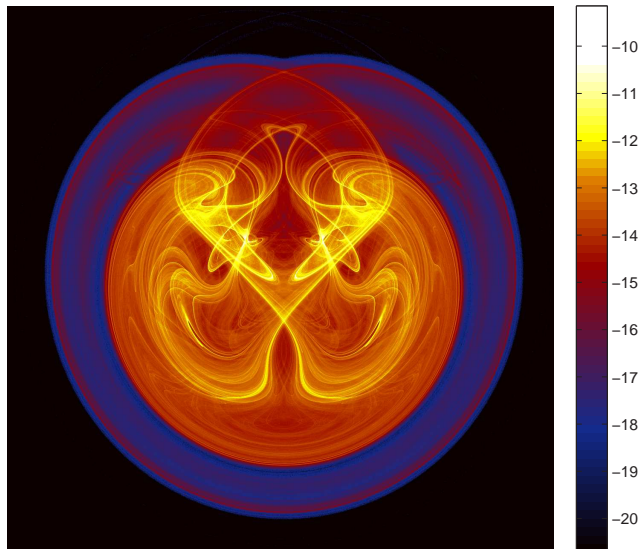

Figure 1: Snapshot of the Lorenz [3] model's random attractor $\mathcal{A}(\omega)$ and of the corresponding sample measure μ_ω, for a given, fixed realization ω. Each of the 3 equations of model [3] is perturbed by multiplicative noise, $dx = s(y - x)dt + \sigma x dW_t$, $dy = (rx - y - xz)dt + \sigma y dW_t$, $dz = (-bz + xy)dt + \sigma z dW_t$, where W_t is a Wiener process and $\sigma > 0$ the noise intensity; we call the resulting model [**SLM**]. The figure corresponds to projection onto the (y, z) plane, $\int \mu_\omega(x, y, z)dx$. One billion initial points have been used and the pullback attractor is computed for $t = 40$. The parameter values are the classical ones — $r = 28$, $s = 10$, and $b = 8/3$ — while $\sigma = 0.3$ and $\delta t = 5 \cdot 10^{-3}$. The color bar to the right is on a log-scale and quantifies the probability to end up in a particular region of phase space. Notice the interlaced filament structures between highly (yellow) and moderately (red) populated regions.

It follows that, in the presence of noise, the resulting PDF looks very much like that of the unperturbed system, only slightly fuzzier: the noise smoothes the small-scale structures of the attractor. More precisely, this fuzziness appears only in the forward approach, which compresses a lot of information; this information, to the contrary, comes into sharp focus in the pullback approach.

A quick look at Figs. 1, 2 and 3 is already enlightening. Figures 1 and 2 show two snapshots of the invariant measure μ_ω supported by the random attractor of our stochastic Lorenz model [**SLM**], for the same realization ω but for two different noise intensities. The exact formulation of model [**SLM**] appears in the caption of Fig. 1. Figure 3 provides four successive snapshots of the sample measure $\mu_{\theta_t \omega}$, for the same noise intensity as in Fig. 2, but with $t = t_0 + k\delta t$ and $k = 0, 1, 2, 3$ for some t_0.

The sample measures in these three figures, and in the associated short video given in the SM, exhibit amazing complexity, with fine, very intense filamentation; note logarithmic scale on color bars in the three figures. There is no fuzziness whatsoever in the topological structure of this filamentation, which evokes the Cantor-set foliation of the deterministic attractor [47]. Such a fine structure strongly suggests that these measures are supported by an object of vanishing volume.

Much more can be said, in fact, about these objects. RDS theory offers a rigorous way to define random versions of stable and unstable manifolds, via the Lyapunov spectrum, the Oseledec multiplicative theorem, and a random version of the Hartman-Grobman theorem [20]. These random invariant manifolds can support measures, like in the deterministic context. When the sample measures μ_ω of an RDS have absolutely continuous conditional measures on the random unstable manifolds, then μ_ω is called a *random SRB measure*.

We can prove rigorously, by relying on Theorem B of [23], that the sample measures of the discretized stochastic system obtained from the [**SLM**] model share the SRB property. Indeed, it can be shown for
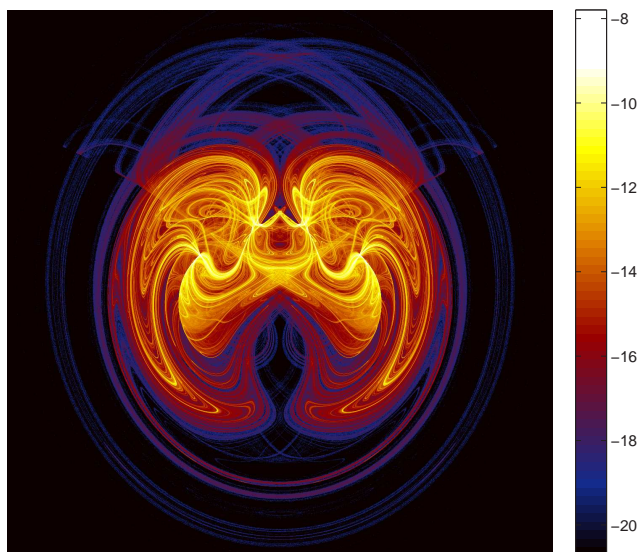
8

Figure 2: Same as Fig. 1, for the same realization ω but with noise intensity $\sigma = 0.5$. Interlaced filament structures between highly and moderately populated regions are now much more complex. Weakly populated regions cover an important part of the random attractor and are, in turn, entangled with "zero-probability" regions (black).
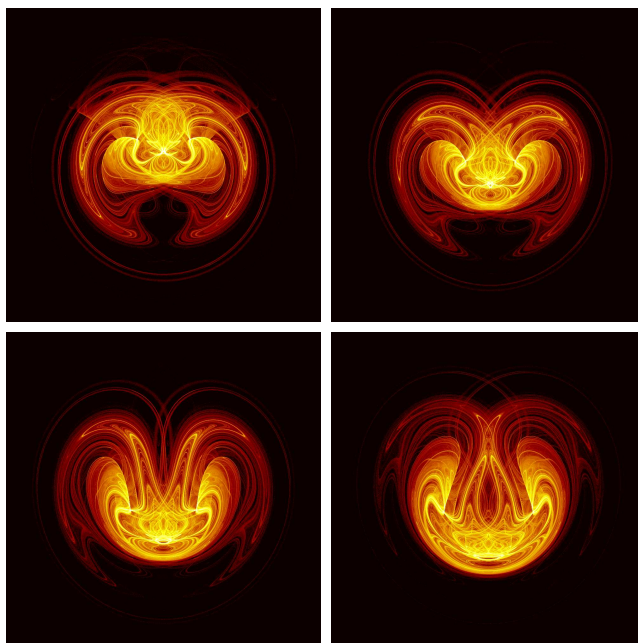

Figure 3: Four snapshots of the random attractor and sample measure supported on it, for the same parameter values as in Fig. 2. The time interval δt between two successive snapshots — moving from left to right and top to bottom — is $\delta t = 0.0875$. Note that the support of the sample measure may change quite abruptly, from time to time; see the related short video in the SM for details.

9

our discretized [**SLM**] model that a *Hörmander hypoellipticity condition* is satisfied, thus ensuring that the process has a smooth density $p(t, y)$ [41]. Standard arguments [43] can be used to prove that the stationary solution ρ of our system's Fokker-Planck equation is in fact absolutely continuous with respect to Lebesgue measure.

Since our simulations exhibit exactly one positive Lyapunov exponent, the absolute continuity of ρ implies that the sample measures seen in Figs. 1–3 are, actually, good numerical approximations of a genuine random SRB measure for our discretized [**SLM**], whenever δt is sufficiently small; see also the next section. In fact, Theorem B of [23] is a powerful result, which clearly shows that — in noisy systems, and subject to fairly general conditions — chaos can lead to invariant sample measures that are SRB.

Note that since the sample measures associated with the discrete [**SLM**] system are SRB here, they are physical measures and can thus be computed at any time t by simply flowing a large set of initial data from the remote past $s \ll t$ up till t, for a fixed realization ω; this is exactly how Figs. 1–3 were obtained. Given the SRB property, the nonzero density supported on the model's unstable manifolds delineates numerically these manifolds; Figs. 1–3 provide therefore an approximation of the *global* random attractor of our stochastic Lorenz system. Finally, these random measures are Markovian, in the sense that they are measurable with respect to the past σ-algebra of the noise [20]. The last statement results directly from the fact that these random measures are physical, cf. (4), and thus satisfy the required measurability conditions as a pullback-limit object.

The evolution of the sample measures $\mu_{\theta_t \omega}$ (see SM video) is quite complex, and two types of motion are present. First, a pervasive "jiggling" of the overall structure can be traced back to the roughness of the Wiener process and to the multiplicative way it enters into the [**SLM**] model. Second, there is a smooth and quite regular low-frequency motion present in the evolution of the sample measures, which seems to be driven by the deterministic system's unstable limit cycles and is thus related to the well-known lobe dynamics. The latter motion is clearly illustrated in Fig. 3.

Besides this low-frequency motion, abrupt changes in the global structure occur from time to time, with the support of the sample measure either shrinking or expanding suddenly. These abrupt changes recur frequently in the video associated with Fig. 3, which reproduces a relatively short sequence out of a very long stochastic model integration; see SM.

As the noise intensity σ tends to zero, the sample-measure evolution slows down, and one recovers numerically the measure of the deterministic Lorenz system (not shown). This convergence as $\sigma \to 0$ may be related to the concept of *stochastic stability* [8, 34]. Such a continuity property of the sample measures in the zero-noise limit does not, however, hold in general; it depends on properties of the noise, as well as of the unperturbed attractor [35, 36, 37].

As stated in the theoretical section, the forward approach is recovered by taking the expectation, $\mathbb{E}[\mu_\bullet] := \int_\Omega \mu_\omega \mathbb{P}(d\omega)$, of these invariant sample measures. In practice, $\mathbb{E}[\mu_\bullet]$ is closely related to ensemble or time averages that typically yield the previously mentioned PDFs. In addition, when the random invariant measures are Markovian and the Fokker-Planck equation possesses stationary solutions, $\mathbb{E}[\mu_\bullet] = \rho$, where ρ is such a solution. Subject to these conditions, there is even a one-to-one correspondence between Markovian invariant measures and stationary measures of the Markov semigroup [20, 25]. The inverse operation of $\mu \mapsto \rho = \mathbb{E}[\mu_\bullet]$ is then given by $\rho \mapsto \mu_\omega = \lim_{t \to \infty} \Phi(-t, \omega)^{-1}\rho$; the latter is in fact the pullback limit of ρ due to the cocycle property [25].

It follows readily from this result that RDS theory "sees" many more invariant measures than those seen by the Markov semi-group approach: non-Markovian measures appear to play an important role in stochastic bifurcation theory [20].

To summarize, one might say that the classical forward approach considers only expectations and PDFs, whereas the RDS approach "slices" the statistics very finely: the former takes a hammer to the problem, while the latter takes a scalpel. Clearly, distinct physical processes may lead to the same observed PDF: the RDS approach and, in particular, the pullback limit are able to discriminate between these processes and thus provide further insight into them.

2.2. Numerical stability of the sample measures

In this subsection, we perform simple numerical tests on the stability of the sample measures μ_ω computed in the previous subsection. We keep the same parameter values as in Fig. 2 and perturb slightly the noise intensity σ from its value $\sigma_0 = 0.5$.

Let C be a fixed cube in $\mathbb{R}^3$ such that the support of the measures lie always in C. We discretize C over a regular mesh with N^3 nodes and obtain a brute-force numerical approximation $\mu_\omega^{\sigma_0,N}$ of the measures $\mu_\omega^{\sigma_0}$. We consider the L_1-error $\delta_{N,\epsilon}(t) := \int_C |p_{\theta_t\omega}^{\sigma_0,N} - p_{\theta_t\omega}^{\sigma_0-\epsilon,N}| \, \mathrm{d}\mathbf{x}$, where $p_{\theta_t\omega}^{\sigma,N}(\mathbf{x})$ is the "probability density" of the discrete sample measure $\mu_{\theta_t\omega}^{\sigma_0,N}$ and the integral is evaluated on the N^3-mesh.
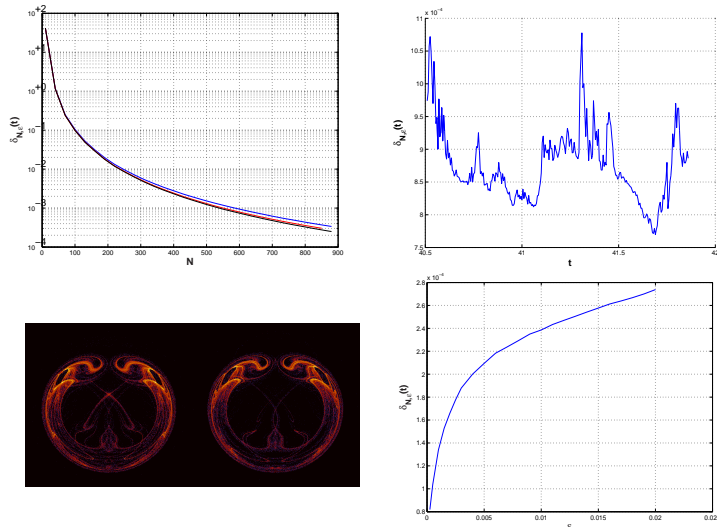

Figure 4: Numerical stability of the invariant sample measures of Fig. 2. The upper-left panel shows the L_1-error $\delta_{N,\epsilon}(t)$ as a function of N for $\epsilon = 10^{-2}$ and three different sets of initial data; the number of points n in the latter is $n = 50^3, 100^3$ and 300^3 for the blue, red and black curves, respectively. The upper-right panel displays $\delta_{N,\epsilon}(t)$ for t varying over 1.5 time units, $40.5 < t < 42.0$, while $N = 600$ and $\epsilon = 10^{-2}$. The lower-left panel plots a snapshot of the two sample measures that correspond to noise intensity σ_0 and $\sigma_0 - \epsilon$ at the end of the time series of $\delta_{N,\epsilon}(t)$ in the upper-right panel. The lower-right panel shows $\delta_{N,\epsilon}(t)$ as a function of ϵ for $N = 900$ and $n = 200^3$; this error clearly converges to zero as $\epsilon \to 0$.

The upper-left panel of Fig. 4 shows the dependency of $\delta_{N,\epsilon}(t)$ with respect to N for fixed t and $\epsilon = 10^{-2}$, and for three different sets of initial data, with an increasing number n of points (see caption). There is very little difference between the three curves, indicating that the number n is already large enough to guard against sampling error. An error of less than 1% is achieved for meshes of size $N \geq 200$.

The upper-right panel of Fig. 4 shows the evolution of $\delta_{N,\epsilon}(t)$ for the same ϵ and $N = 600$. It shows that, as the sample measures evolves with θ_t, they remain close to each other for all time; here $7.5 \cdot 10^{-4} \leq \delta_{N,\epsilon}(t) \leq 11 \cdot 10^{-4}$.

The lower-left panel shows an actual snapshot of the two measures at a fixed time t, at the end of the time series plotted in the upper-right panel (same N and ϵ). Tiny differences become visible when zooming in on the electronic file of the figure (see SM). As a matter of fact, for $\epsilon < 10^{-3}$, the two measures are no longer distinguishable by eye for the N used.

The lower-right panel exhibits $\delta_{N,\epsilon}(t)$ as a function of ϵ. *Statistical stability* is observed here, since the difference between the two measures tends to zero as $\epsilon \to 0$, in the sense of weak convergence. We conclude that the numerical results shown in this paper are very robust. Note that we cannot precisely estimate the level of accuracy with which the sample measures are computed, although the results of this subsection indicate they are quite good, due to the very large ensembles of initial states we used. Unfortunately, statistical methods for improving PDF estimates, like kernel density estimation, require densities that are at least twice differentiable, while the numerical evidence here is that the sample measures of our stochastic

system are strongly suspect of not being even absolutely continuous with respect to Lebesgue measure; in particular, they are not differentiable even once.

2.3. A stochastic ENSO model and its RDS analysis

In this subsection, we compute random invariant measures to obtain more detailed information on the stochastic Timmermann and Jin [31] model (hereafter [**STJ**]). These measures enable us to understand at a deeper level the interaction between noise and nonlinearity in this slightly more realistic climate model

Our theoretical laboratory is climate variability in the Tropical Pacific, which is characterized by the interannual ENSO oscillation. A variety of modeling studies and observations strongly suggest that the irregular, 2–7-year time scale of ENSO is produced by nonlinear ocean-atmosphere interactions in this region [18, 19]. In addition, this variability is bracketed by high-frequency, intraseasonal noise due to so-called "westerly wind bursts" in the surface winds, and by interdecadal changes in the global ocean circulation [29].

Aside from global effects that act on interdecadal time scales, these time scales may also arise from the interaction between noise and purely tropical effects [40]. Timmermann and Jin [31] argued, based on a dynamical analysis of the Jin model [24], that long-term changes in ENSO activity may result from perturbations of a homoclinic orbit specifically associated with the nonlinear advection terms in the model's sea surface temperature equation. Using the classical forward approach, they noted that such conclusions are robust against the introduction of wind-generated noise in the model.

Our aim here is to show that a pullback approach is better suited to study and rigorously quantify stochastic effects on the low-frequency dynamics than using merely a forward, PDF-type approach. We believe that this statement is likely to be true for any problem involving noise and nonlinearity.

Recall first that, if the deterministic model's variability is damped, adding even small-amplitude stochastic forcing can easily result in significant nonlinear effects; see [29, 30] and references therein. Such a noise-induced excitation of supercritical behavior at deterministically subcritical parameter values does indeed occur in our numerical study of the [**STJ**] model (not shown). We were able to confirm the presence of a Shil'nikov-type bifurcation to homoclinic orbits in the deterministic model, and show that the noise helps trigger horseshoe-like behavior in phase space for parameter ranges in which the deterministic model has only a stable limit cycle. All these noise-induced phenomena possess geometric features captured by the random attractors.

For instance — in the damped regime, with the right amount of noise — the global random attractor $\mathcal{A}(\theta_t \omega)$ of our [**STJ**] model is a closed curve, whose length and location in phase space vary with time, *i.e.*, a random periodic orbit [50] that pullback attracts the phase space's Lebesgue measure. This *random limit cycle* is associated with a broad spectral peak (not shown).

We proceed now to study the [**STJ**] model in a chaotic regime, *i.e.* when zonal advection is sufficiently strong. In the absence of noise (not shown), we effectively observe that the model exhibits interdecadal variability; this variability is due to the nearby presence in parameter space of a homoclinic orbit, whose characteristic amplitude modulation can be seen in the time series of T_2 (see Fig. 5). When including the noise that models wind-stress bursts, the sample measures shown in the figure still possess a complex structure.

The six sample measures $\mu_{\theta_t \omega}$ shown in the bottom panels of Fig. 5, at interannual intervals of $\Delta t = 1.6$ years, are even more obviously singular than those in Figs. 1–4: at every time t, the regions that are most populated by the stochastic flow are confined mainly to filaments near the sharp peak (white + sign) that is located in the upper-left corner of the $(h - T_2)$ plane. For the decadal time scale of $6 \times 1.6 \simeq 10$ years, the change in probability of occurrence of El Niño episodes, with warm T_2, is clearly visible. The underlying random attractor may, however, exhibit the same spectral signature as the random limit cycle does.

The distinction in behavior stems from the intermittency [42] that our [**STJ**] model exhibits in the chaotic regime; this is clearly visible in the T_2 time series in the upper-left panel of Fig. 5. The model's intermittency is encoded by the structure of its random attractor, which differs from that of the [**SLM**] model shown in Figs. 1–3. In Fig. 5 we have a situation that lies in between a random point — as observed in other ENSO models that are governed by linear dynamics with additive noise, cf. [51], for instance — and a noisy chaotic model with strong mixing, like [**SLM**].
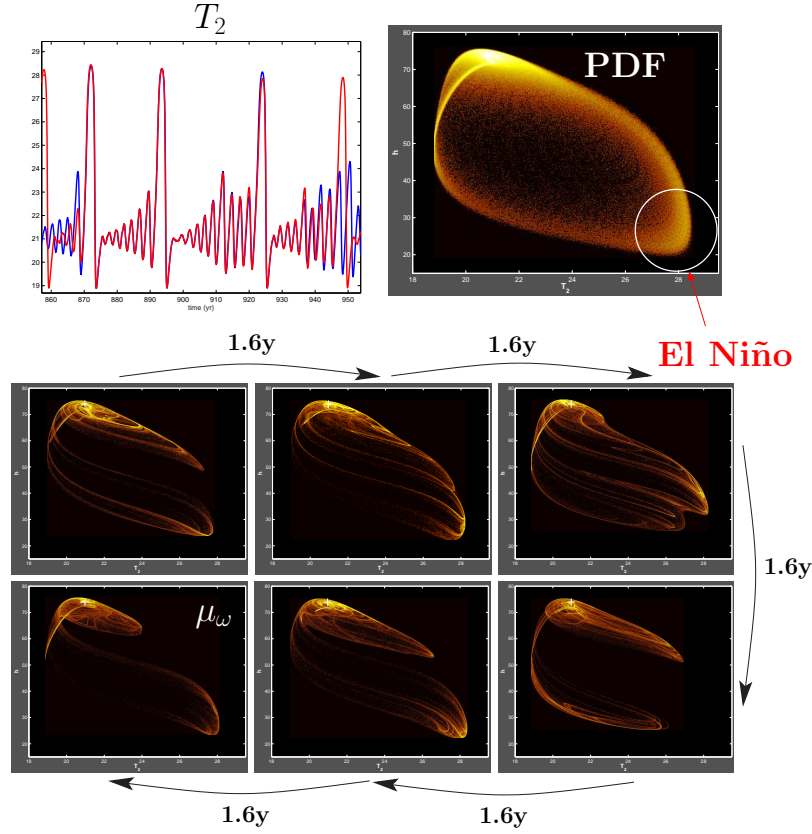
Figure 5: The **[STJ]** model [31] is governed by the following three SDEs: $\dot{T}_1 = -\alpha(T_1 - T_r) - (2\epsilon u/L)(T_2 - T_1)$, $\dot{T}_2 = -\alpha(T_2 - T_r) - (w/H_m)(T_2 - T_{sub}(h))$, and $\dot{h} = r(-h - bL\tau/2)$; here T_1 and T_2 are the sea surface temperatures in the western and eastern Tropical Pacific, and h is the thermocline depth, while $T_{sub}(h) = T_r - [(T_r - T_{r0})/2][1 - \tanh(H + bL\tau + h - z_0)/h^*]$ and $\tau = (a/\beta)(T_1 - T_2)(\xi_t - 1)$. Wind stress anomalies are given by τ, the equatorial upwelling by $w/H_m = -\beta\tau$, the zonal advection by $u = \beta L\tau/2$, and T_{sub} is the subsurface temperature; please see [31] for other subsidiary variables and parameters. Wind stress bursts are modeled as white noise ξ_t of variance σ, while ϵ measures the strength of the zonal advection and serves as a bifurcation parameter in [31]. Intermittency is illustrated in the upper-left panel, for two different initial states at $t = 0$ (blue and red curves) and the same realization ω; only the red curve appears where the two are visually indistinguishable. Six snapshots of the attractor and the sample measures μ_ω they support are shown at regular, 1.6-year intervals in the bottom panels; they are projected onto the $(h - T_2)$ plane, with T_2 on the abscissa, and their timing corresponds to interannual variability. The forward PDF is shown in the upper-right panel: it averages the sample measures μ_ω.

To summarize, the [**STJ**] model possesses two main types of random attractor: (i) a random limit cycle in the deterministically damped regime; and (ii) a random attractor associated with intermittency in chaotic regimes. The first one may be of interest in understanding certain features that are displayed by fairly realistic models of the tropical ocean driven by surface winds [30], while the second one needs further investigation, theoretically as well as practically. Both offer new perspectives in the understanding of ENSO variability.

3. Concluding remarks

We have briefly motivated and outlined the main concepts and tools of RDS theory, in particular how to rigorously define stochastic flows and random attractors, as well as the corresponding invariant random measures. It appears from this outline that a stochastically perturbed system's pullback, strong attractor [33] provides much more detailed information on the system's dynamics and statistics than its PDF.

Detailed computations of the invariant sample measures for the stochastic Lorenz model [**SLM**] reveal the amazing complexity that underlies its PDF; see Figs. 1–3. The numerical results were shown to be quite robust (Fig. 4) and suggest that the actual measures are Markovian random SRB measures [8] associated with one positive Lyapunov exponent.

We saw, moreover, that other noisy systems with a positive Lyapunov exponent may exhibit random attractors having a less striking geometry, which still support nontrivial sample measures and are associated with intermittent synchronization. This was illustrated on the nonlinear stochastic ENSO model [**STJ**] of [31]. We showed in Fig. 5 that the information conveyed by the sample measure's evolution in time can improve the physical interpretation of the dynamics.

Moving on to intermediate models, consider the numerical results of [52] obtained with a so-called hybrid coupled model that couples an empirical atmospheric component to an oceanic GCM; see also [9, 27]. The authors showed that noise can shift and broaden the model's spectral peaks (see Figs. 6 and 8 there). Here again, the RDS approach could provide deeper insights into this phenomenon.

On a longer, multidecadal time scale, the RDS approach could be combined with linear response theory [53]. For hyperbolic deterministic systems, precise estimates exist for the response of their SRB measures to perturbations [53, 54]. It appears feasible to extend such estimates to SRB sample measures [54]. In the climate setting, the response of forward PDFs to deterministic perturbations has been studied in [12, 13]; climate response to stochastic perturbations should be next and provide, maybe, better climate projections.

A key question arises here, as for many other new mathematical concepts and tools, when first applied in the climatic or, more generally, physical context. The question is how to extend these relatively novel ideas to more detailed and realistic models and even to observational data sets [9, 55]. This question is under investigation for certain intermediate ENSO models and results will be reported elsewhere.

References

[1] V.I. Arnol'd, Geometrical Methods in the Theory of Ordinary Differential Equations, Springer-Verlag, New York, 1983.
[2] J.-P. Eckmann, D. Ruelle, Ergodic theory of chaos and strange attractors, Rev. Modern. Phys. 57 (1985) 617–656.
[3] E.N. Lorenz, Deterministic nonperiodic flow, J. Atmos. Sci. 20 (1963) 130–141.
[4] K. Hasselmann, Stochastic Climate models, Part 1: Theory, Tellus 28 (1976) 473–485.
[5] E. Kalnay, Atmospheric Modeling, Data Assimilation and Predictability, Cambridge University Press, 2003.
[6] S. Solomon, *et al.*, Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the IPCC, Cambridge University Press, 2007.
[7] D. Ruelle, F. Takens, On the nature of turbulence, Commun. Math. Phys. 20 (1971) 167–192.
[8] L.-S. Young, What are SRB measures, and which dynamical systemes have them? J. Stat. Phys. 108 (2002) 733–754.
[9] M. Ghil, A.W. Robertson, Solving problems with GCMs: General circulation models and their role in the climate modeling hierarchy, in General Circulation Model Development: Past, Present and Future, D. Randall (Ed.), Academic Press, San Diego, (2000) 285–325.

[10] I.M. Held, The gap between simulation and understanding in climate modeling, Bull. Amer. Meteorol. Soc. 86 (2005) 1609–1614.

[11] J.C. McWilliams, Irreducible imprecision in atmospheric and oceanic simulations, Proc. Natl. Acad. Sci. USA. 104 (2007) 8709–8713.

[12] G.L. Eyink, T.W.N. Haine, D.J. Lea, Ruelle's linear response formula, ensemble adjoints schemes, and Lévy flights, Nonlinearity, 17 (2004) 1867-1889.

[13] R. Abramov, A. Majda, A new algorithm for low frequency climate response, J. Atmos. Sci. 66 (2009) 286–309.

[14] R. Buizza, M. Miller, T.N. Palmer, Stochastic representation of model uncertainties in the ECMWF ensemble prediction system, Quaterly J. Roy. Meteo. Soc. 125 (1999) 2887–2908.

[15] J.D. Neelin, O. Peters, J.W.-B. Lin, K. Hales, C.E. Holloway CE Rethinking convective quasi-equilibrium: observational constraints for stochastic convective schemes in climate models, Phil. Trans. Roy. Soc. London A 366 (2008) 2581–2604.

[16] M. Ghil, Cryothermodynamics: The chaotic dynamics of paleoclimate, Physica D 77 (1994) 130–159.

[17] B. Saltzman, Dynamical Paleoclimatology, Academic Press, 2002.

[18] F.-F. Jin, J.D. Neelin, M. Ghil, El Niño on the Devil's Staircase: Annual subharmonic steps to chaos, Science 264 (1994) 70–72.

[19] E. Tziperman, L. Stone, M. Cane, J. Jarosh, El Niño chaos: Overlapping of resonances between the seasonal cycle and the Pacific ocean-atmosphere oscillator, Science 264 (1994) 72–74.

[20] L. Arnold, Random Dynamical Systems, Springer-Verlag, NY, 1998.

[21] D. Cheban, Global Attractors of Nonautonomous Dissipative Dynamical Systems, Interdisciplinary Mathematical Sciences, World Scientific, 2004.

[22] M. Ghil, M.D. Chekroun, E. Simonnet, Climate dynamics and fluid mechanics: Natural variability and related uncertainties, Physica D 237 (2008) 2111–2126.

[23] F. Ledrappier, L.-S. Young, Entropy formula for random transformations, Prob. Theory Related Fields, 80 (1988) 217–240.

[24] F.-F. Jin, An equatorial ocean recharge paradigm for ENSO. Part I, J. Atmos. Sci. 54 (1997) 811–829.

[25] H. Crauel, Random Probability Measures on Polish Spaces, Vol. 11 in The Stochastic Monographs: Theory and Applications of Stochastic Process Series, Taylor & Francis, 2002.

[26] V.S. Anishchenko, T.E. Vadivasova, A.S. Kopeikin, G.I. Strelkova, J. Kurths, Influence of noise on statistical properties of nonhyperbolic attractors, Phys. Rev. E, 62 (2000) 7886–7893.

[27] H.A. Dijkstra, Nonlinear Physical Oceanography: A Dynamical Systems Approach to the Large Scale Ocean Circulation and El Niño (2nd ed.), Springer-Verlag, NY, 2005.

[28] L. Sushama, M. Ghil, K. Ide, Spatio-Temporal variability in a mid-latitude ocean basin subject to periodic wind forcing, Atmosphere-Ocean 45 (2007) 227–250.

[29] J.D. Neelin, *et al.*, ENSO theory, J. Geophys. Res. 103 (1998) 14261–14290.

[30] A.V. Fedorov, S.L. Harper, S.G. Philander, B. Winter, A. Wittenberg, How predictable is El Niño? Bull. Am. Meteorol. Soc. 84 (2003) 911–919.

[31] A. Timmermann, F.-F. Jin, A nonlinear mechanism for decadal El Niño amplitude changes, Geophys. Res. Lett. 29 (2002) 1003. doi: 10.10129/2001GL13369.

[32] H. Kunita, Stochastic Flows and Stochastic Differential Equations, Cambridge Univ. Press, 1990.

[33] H. Crauel, G. Dimitroff, M. Scheutzow, Criteria for strong and weak random attractors, J. Dyn. Diff. Eqn 21 (2009) 233–247.

[34] Y. Kifer, Ergodic Theory of Random Perturbations, Birkhauser, 1988.

[35] H. Crauel, White noise eliminates instablity, Arch. Math. 75 (2000) 472–480.

[36] H. Crauel, F. Flandoli, Additive noise destroys a pitchfork bifurcation, J. Dyn. Diff. Eqn. 10 (1998) 259–274.

[37] A.N. Carvalho, J.A. Langa, J.C. Robinson, Lower semicontinuity of attractors for non-autonomous dynamical systems, Ergod. Th.. Dyn. Syst. 29 (2009) 1765–1780.

[38] H. Crauel, Random point attractors versus random set attractors, J. London Math. Soc. 63 (2001) 413–427.

[39] H. Crauel, F. Flandoli, Attractors for random dynamical systems, Prob. Th. Rel. Fields 100 (1994) 365–393.

[40] B.P. Kirtman, P.S. Schopf, Decadal variability in ENSO predictability and prediction, J. Climate 11 (1998) 2804–2822.

[41] J.J. Kohn, Pseudo differential operators and hypoellipticity, Proc. Amer. Math. Soc. Symp. Pure Math. 23 (1973) 61–69.

[42] P. Ashwin, Attractors of a randomly forced electronic oscillator, Physica D 125 (1999) 302–310.

[43] C. Soize, The Fokker-Planck Equation for Stochastic Dynamical Systems and Its Explicit Steady States Solutions, World Scientic Publishing Co., 1994.

[44] M. Dorfle, R. Graham, Probability density of the Lorenz model, Phys. Rev. A 27 (1983) 1096–1105.

[45] A.K. Mittal, S. Dwivedi, R.S. Yadav, Probability distribution for the number of cycles between successive regime transitions for the Lorenz model, Physica D 233 (2007) 14–20.

[46] P. Collet, C. Tresser, Ergodic theory and continuity of the Bowen-Ruelle measure for geometrical flows, Fyzika 20 (1988) 33-48.

[47] J. Guckenheimer, P. Holmes, Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields, Springer-Verlag, NY, 1983.

[48] W. Tucker, Lorenz attractor exists, C. R. Acad. Sci. Paris 328 (12) (1999) 1197–1202.

[49] V. Araujo, M. Pacifico, R. Pujal, M. Viana, Singular-hyperbolic attractors are chaotic, Trans. Amer. Math. Soc. 361 (2009) 2431–2485.

[50] H. Zhao, Z.-H. Zheng, Random periodic solutions of random dynamical systems, J. Diff. Eqn. 246 (2009) 2020–2038.

[51] C. Penland, A stochastic model of IndoPacific sea surface temperatures using Linear Inverse Modeling, Physica D 98 (1996) 534–558.

[52] B. Blanke, J.D. Neelin, D. Gutzler, Estimating the effect of stochastic wind stress forcing on ENSO irregularity, J. Climate 10 (1997) 1473–1486.

[53] D. Ruelle, A review of linear response theory for general differentiable dynamical systems, Nonlinearity 22 (2009) 855-870.

[54] D. Ruelle, Differentiation of SRB states for hyperbolic flows, Ergod. Th. Dyn. Syst. 28 (2008) 613–631.

[55] M. Ghil, A.W. Robertson,"Waves" vs "particles" in the athmosphere's phase space: A pathway to long-range forecasting?, Proc. Natl. Acad. Sci. 99 (2002) 2493–2500.