



Regular Article

# Physical invariant measures and tipping probabilities for chaotic attractors of asymptotically autonomous systems

Peter Ashwin<sup>a</sup> and Julian Newman<sup>b</sup>

Department of Mathematics, Centre for Systems, Dynamics and Control, University of Exeter, Exeter, EX4 4QF, UK

Received 7 December 2020 / Accepted 14 April 2021  
© The Author(s) 2021

**Abstract** Physical measures are invariant measures that characterise “typical” behaviour of trajectories started in the basin of chaotic attractors for autonomous dynamical systems. In this paper, we make some steps towards extending this notion to more general nonautonomous (time-dependent) dynamical systems. There are barriers to doing this in general in a physically meaningful way, but for systems that have autonomous limits, one can define a physical measure in relation to the physical measure in the past limit. We use this to understand cases where rate-dependent tipping between chaotic attractors can be quantified in terms of “tipping probabilities”. We demonstrate this for two examples of perturbed systems with multiple attractors undergoing a parameter shift. The first is a double-scroll system of Chua et al., and the second is a Stommel model forced by Lorenz chaos.

## Contents

1	Tipping points and chaotic multistability . . . . .
2	Physical measures and pullback attractors . . . . .
3	Tipping between past and future attractors . . . . .
4	Example 1: double scroll with parameter shift . . . . .
5	Example 2: Lorenz–Stommel with parameter shift . . . . .
6	Discussion . . . . .
A	Proof of Theorem 6 . . . . .
	References . . . . .

## 1 Tipping points and chaotic multistability

For a deterministic autonomous dynamical system, an invariant measure is “physically relevant” [12] (also “physical” or “natural”) if it describes the statistics of a typical trajectory that started an arbitrarily long time ago in the past, i.e., without transients. There is an extensive literature that describes existence, hyperbolicity, dimension, entropy, regularity, and other properties of such measures, starting with the work of Sinai, Bowen, and Ruelle in the 1970s, reviewed in [27]. If there are a finite number of attractors  $\{A_i\}_{i=1}^n$ , there may be physical measures on each attractor and a splitting up of the phase space into the basins of attraction of the  $A_i$ .

As an example, models of the Atlantic Meridional Overturning Circulation (AMOC) [10] show the presence of two attractors corresponding to different tur-

bulent states of transport within the north Atlantic. There is evidence, across the model hierarchy, of bistability between an “on” attractor that transports large amounts of heat to high latitudes and an “off” attractor that does not. If there is a shift in parameters (e.g., caused by changes in forcing due to anthropogenic CO<sub>2</sub> emissions), it is important to know whether the shift in forcing may result in a “tipping”, i.e., a change to the qualitative state of the climate, such as the AMOC tipping from its current “on” state to the “off” state. Although there is an extensive literature on bifurcation-induced tipping based around what happens to dynamically simple (equilibrium) attractors losing stability by a fold bifurcation [20], much less is known about how more complex (chaotic) attractors may respond to changes in parameters. In addition to this point, the bifurcation-induced tipping scenario considered in most of the literature assumes a highly idealised scenario of an “infinitely slow” parameter drift through an essentially deterministic parameter-dependent autonomous dynamical system with simple attractors. By comparison, not so much is known regarding rate-induced tipping effects [5], where the realistic non-infinitesimal rate of real-time parameter drift may affect the realisation and/or timing of a tipping event.

The framework of nonautonomous systems [19] provides a promising approach to investigating tipping between chaotic attractors for dynamical models with a real-time parameter drift. The utility of the nonautonomous system framework has been noted in various applications in the life sciences [18] and machine learning [22], and indeed, in climate science, there have been numerical approximations of physical measures on pullback attractors at least in the case of stationary

<sup>a</sup> e-mail: [p.ashwin@exeter.ac.uk](mailto:p.ashwin@exeter.ac.uk) (corresponding author)

<sup>b</sup> e-mail: [j.m.i.newman@exeter.ac.uk](mailto:j.m.i.newman@exeter.ac.uk)

random forcing [8, 13]. In the case of a deterministic real-time parameter drift taking place within a compact time-interval (before which and after which the parameter is constant), [17] considers the time-evolution of physical measures of the pre-drift autonomous system, and uses this to describe the probability of tipping to different attractors of the post-shift autonomous system. More precisely, the “probability of tipping” quantifies the proportion (based on the physical measure of the pre-shift autonomous system) of the “snapshot attractor” that tips to a given attractor of the post-shift autonomous system. This proportion may be interpreted as a “probability” by virtue of the uncertainty of the exact state  $x(t)$  of the climate system being modelled as a solution of the deterministic nonautonomous equation. The case that this probability lies strictly between 0 and 1 corresponds to “splitting of the snapshot attractor”, a phenomenon that has been called *partial tipping* [1]. To understand tipping in such chaotic multistable systems for more general forcing, one needs to extend the concept of physical measure to nonautonomous systems. Such measures are explored practically for climate models in [11, 17] for parameter shifts within a compact time-interval, while theoretical generalizations of SRB measures (which are an important special case of physical measures) to periodically or randomly forced systems are reviewed in [28]; however, the general question of what should be a physical measure for a general nonautonomous system does not seem to have a simple answer.

In this paper, we propose a notion of physical measures appropriate for pullback attractors of nonautonomous systems, in the case that there are autonomous past and future limits with chaotic attractors. This corresponds to a parameter shift between two asymptotic values, one in the distant past and one in the distant future. Such systems, and their dependence on the rate of the parameter shift, have previously been investigated for shifts between equilibrium attractors [4, 5] and periodic or more general attractors [1, 2, 16]. Rigorous results on the existence of such nonautonomous physical measures are presented in a companion paper [24].

We then use our proposed notion of physical measures to define rigorously the probability of tipping to the different attractors of the future limit system. This involves formulating suitable technical conditions that will guarantee that the tipping probabilities are indeed well defined. Similarly to [17], these probabilities can be understood as the probability that the climate state  $x(t)$  will tip to a given attractor of the future limit system conditional on both the specific time-dependent dynamical system governing the evolution of  $x(t)$  and on knowing which of the attractors of the past-limit system the climate state was close to in the distant past, but without the knowledge of the *exact* state  $x(t)$  at any time  $t$ . We use these tipping probabilities to investigate tipping points arising in systems with a real-time continuous parameter shift. In particular, we will look at how the probability of tipping depends on the rate of parameter shift, and especially how the rate determines whether the pullback attractor undergoes “no tipping”

(i.e., the probability of tipping is zero), “full tipping” (i.e., the probability of tipping is one), or “partial tipping”, analogous to the phenomena in [1].

We consider systems modelled by a nonautonomous ODE of the form

$$\dot{x} = f(x, \Lambda(rt)) \quad (1)$$

for  $x \in \mathbb{R}^d$ , with  $f(x, \lambda)$  smooth in  $x$  and  $\lambda \in \mathbb{R}^p$ , and with a parameter  $r > 0$  controlling the rate of time-dependence of the system, which we call the *rate parameter*. As in [4], for two values  $\lambda^-$  and  $\lambda^+$ , we say there is a *rate-dependent parameter shift between  $\lambda^-$  and  $\lambda^+$*  if  $\Lambda(t)$  is a smooth function of  $t$ , such that

$$\lim_{t \rightarrow \pm\infty} \Lambda(t) = \lambda^\pm. \quad (2)$$

If  $\|\Lambda(t) - \lambda^\pm\|$  decays sufficiently fast as  $t \rightarrow \pm\infty$ , then the system (1) can be smoothly compactified [26] to create an autonomous system in an extended phase space. We refer to

$$\dot{x} = f(x, \lambda) \quad (3)$$

for  $\lambda \in \mathbb{R}^p$  as the *frozen systems* of (1,2), while the *past/future limit systems* of (1,2) correspond to the cases  $-/+$ , respectively, for

$$\dot{x} = f(x, \lambda^\pm). \quad (4)$$

We suppose that the past-limit system has a chaotic attractor  $A^-$  that carries a “physical measure”  $\mu^-$ . The structure of the paper will be as follows. Section 2 recalls some background on physical measures and proposes a way to extend a measure  $\mu^-$  for the past-limit system (4) to a time-dependent physical measure  $\mu_t$  for (1). Rigorous conditions for the existence of such time-dependent physical measures are presented in [24]. These time-dependent measures are used in Sect. 3 to define tipping probabilities, which depend on the rate of parameter shift (precisely on the basis of how the time-dependent measure  $\mu_t$  itself depends on the rate of parameter shift). We then illustrate physical measures and tipping probabilities through application to two example models. Section 4 considers a parameter shift applied to a prototypical dynamical system with two attractors, namely the double scroll of Chua et al. [9] with bistability between a chaotic attractor and a periodic attractor. Section 5 considers a parameter shift applied to a conceptual climate model consisting of a Lorenz system (representing internal variability) [21] forcing a Stommel model (representing the AMOC) [25]; for each time-frozen parameter value, this skew product system exhibits bistability between two chaotic attractors. Section 6 highlights some further questions and implications for the prediction of tipping points.

Some notations:  $\mathcal{B}(\mathbb{R}^d)$  denotes the Borel  $\sigma$ -algebra of  $\mathbb{R}^d$ ;  $B_\varepsilon(x) = \{y : |x - y| < \delta\}$  denotes the open ball of radius  $\varepsilon$  about  $x$ ; and for a probability measure

$\mu$ ,  $\text{supp}(\mu)$  denotes the support of  $\mu$ , i.e., the smallest closed set  $G$  for which  $\mu(G) = 1$ .

## 2 Physical measures and pullback attractors

For the deterministic nonautonomous model (1), a “physical measure” aims to describe, at each time  $t$ , the probability distribution for the location of a solution  $x(t)$  when the exact location is not known, but it is known that in the distant past, the solution was close to a particular attractor  $A^-$  of the past-limit system. The support of this time-dependent probability distribution corresponds to a “pullback attractor” of (1).

Consider the solution  $x(t)$  of (1) as a process [19]

$$x(t) = \Phi^{(r)}(t, s, x(s))$$

for any  $t > s$ . For the associated autonomous (frozen) system (3), we can write the solution

$$x(t) = \varphi_\lambda(t - s, x(s))$$

for any  $t > s$  and fixed  $\lambda$ , and we likewise write  $\varphi_{\lambda^\pm}$  in the case of the future- and past-limit system.

We first review physical measures in the autonomous setting, which we apply to the past-limit system, and then, we consider how to “extend” this from the past-limit system to the nonautonomous system. In the autonomous case

$$\dot{x} = f(x), \tag{5}$$

the solution flow  $\varphi$  is given by  $x(t) = \varphi(t - s, x(s))$ . The definitions that we now introduce for the autonomous case will later be applied to the past-limit system

$$\dot{x} = f(x, \lambda^-), \tag{6}$$

and from there, the analogous definitions for the nonautonomous system will be presented.

A (local) attractor of the system (5) is a closed bounded (i.e., compact) set  $A \subset \mathbb{R}^d$  admitting a neighbourhood  $U$ , such that

$$d_H(\overline{\varphi(t, U)}, A) \rightarrow 0 \quad \text{as } t \rightarrow \infty, \tag{7}$$

where  $d_H$  is the Hausdorff distance. The basin of attraction  $B$  of the attractor  $A$  is the union of all open neighbourhoods  $U$  of  $A$  fulfilling (7). It is not hard to show that for any bounded neighbourhood  $U$  of  $A$ ; if  $\bar{U} \subset B$ , then  $U$  satisfies (7). It is also not hard to show that if a set of attractors of (5) is pairwise disjoint, then their basins of attraction are likewise pairwise disjoint.

Now, for each  $x_0 \in \mathbb{R}^d$  and  $T > 0$ , we define the corresponding empirical measure  $\mu_{T, x_0}$  as follows: for each Borel set  $S \in \mathcal{B}(\mathbb{R}^d)$ ,  $\mu_{T, x_0}(S)$  is the proportion

of time  $t$  within the interval  $[0, T]$  for which  $x(t) \in S$ , where  $x(\cdot)$  is the solution of (5) with  $x(0) = x_0$ . In other words, if  $\ell$  denotes Lebesgue measure, then

$$\mu_{T, x_0}(S) = \frac{\ell(\{t \in [0, T] : \varphi(t, x_0) \in S\})}{T}.$$

Various possible definitions of “natural” or “physical” measures of autonomous systems exist. In general, these are not specifically defined with reference to attractors, but here, we will define a “physical measure on a given attractor  $A$ ”. Such measures are an important case of the general definition of “physical measures” given in [27].

**Definition 1** Given an attractor  $A$  of (5) with basin of attraction  $B$ , a physical measure on  $A$  is a probability measure  $\mu$  with  $\text{supp}(\mu) = A$ , such that for Lebesgue-almost every  $x_0 \in B$ , as  $T \rightarrow \infty$ , the empirical measure  $\mu_{T, x_0}$  converges weakly to  $\mu$ .

Under this definition, a physical measure is invariant measure but not necessarily ergodic<sup>1</sup> [23]. Furthermore, an attractor supports at most one physical measure. As an example, if (5) has a stable periodic orbit, then this periodic orbit supports a physical measure: namely, if  $A$  is a stable  $T$ -periodic orbit for some  $T > 0$ , then the physical measure  $\mu$  is precisely the time- $T$  empirical measure  $\mu_{T, p}$  starting at any point  $p \in A$ .

Now, in the example of a stable hyperbolic periodic orbit, if we observe at some time  $t$  the approximate state of some solution  $x(t)$  within the basin of attraction of the periodic orbit, then this observation determines forever afterwards the approximate asymptotic phase of this solution  $x(\cdot)$ ; thus, *some memory is retained forever*. We now introduce a type of physical measure where there is “no everlasting memory of previous observations”; this excludes examples such as stable hyperbolic periodic orbits, but still includes many examples of chaotic attractors. The significance from a computational point of view will also be discussed at the end of this section.

**Definition 2** Given an attractor  $A$  of (5) with basin of attraction  $B$ , a PF-attracting physical measure (or just “attracting measure” for short) is a physical measure  $\mu$  supported on  $A$ , such that for every probability measure  $\nu_0$  absolutely continuous w.r.t. Lebesgue where the density  $h \in L^1(\mathbb{R}^d)$  is supported within  $B$ , the pushforward measure

$$\nu_T(S) := \int_S (\mathcal{L}_{\varphi(T)} h)(y) dy = \int_{\mathbb{R}^d} \mathbb{1}_S(\varphi(T, x)) h(x) dx$$

converges weakly to  $\mu$  as  $T \rightarrow \infty$ . Here,  $\mathcal{L}_{\varphi(T)} : L^1(\mathbb{R}^d) \rightarrow L^1(\mathbb{R}^d)$  is the Perron–Frobenius (transfer) operator associated with the map  $\varphi(T, \cdot)$ .

<sup>1</sup> A probability measure  $\mu$  is ergodic if and only if for  $\mu$ -almost every  $x_0 \in \mathbb{R}^d$  (as opposed to “Lebesgue-almost every  $x_0 \in B$ ” in Definition 1), the empirical measure  $\mu_{T, x_0}$  converges weakly to  $\mu$ .

Under some mild assumptions (e.g., [6, Theorem 5.3]), Axiom A attractors for flows will support an attracting measure.

We now suggest an extension of these notions to the nonautonomous case (1) with an asymptotically autonomous past limit. In particular, the concept of an attractor will now become a “pullback attractor”. Pullback attractors have been defined for the general setting of nonautonomous dynamical systems [19, Definition 3.48(ii)], but here, we specifically wish to define a concept of pullback attractors for the setting of past-asymptotically autonomous systems, where we specify an attractor of the past-limit system as the “starting point” of the pullback attractor (see Definition 3).

Given a set  $\mathcal{A}$  of solutions  $x: \mathbb{R} \rightarrow \mathbb{R}^d$  of the nonautonomous system (1) defined on the entire timeline  $\mathbb{R}$ , we write  $\mathcal{A}(t) \subset \mathbb{R}^d$  for the set of locations of these trajectories at time  $t$ , that is

$$\mathcal{A}(t) = \{x(t) : x \in \mathcal{A}\}.$$

Given a probability measure  $\mu$  defined on such a set  $\mathcal{A}$  of solutions  $x: \mathbb{R} \rightarrow \mathbb{R}^d$  of the nonautonomous system (1),<sup>2</sup> we define the probability measure  $\mu_t$  on  $\mathbb{R}^d$  to be the corresponding distribution of locations of these trajectories at time  $t$ , that is

$$\mu_t(S) = \mu(\{x \in \mathcal{A} : x(t) \in S\}).$$

We define pullback attractors, physical measures, and pullback-PF-attracting physical measures relative to a given local attractor  $A^-$  of the past-limit system (6) with basin of attraction  $B^-$ .

**Definition 3** For the nonautonomous system (1), a *pullback attractor starting at  $A^-$*  is a set  $\mathcal{A}$  of two-sided solutions of (6), such that the following statements hold:

- (1)  $\bigcup_{t \in \mathbb{R}} \mathcal{A}(t)$  is bounded, and for each  $t \in \mathbb{R}$ ,  $\mathcal{A}(t)$  is closed;
- (2) for any bounded neighbourhood  $U$  of  $A^-$  with  $\bar{U} \subset B^-$ , for each  $t \in \mathbb{R}$  and  $\varepsilon > 0$ , taking sufficiently large-magnitude  $s < 0$  gives

$$\mathcal{A}(t) \subset \Phi^{(r)}(t, s, U) \subset B_\varepsilon(\mathcal{A}(t)).$$

For each  $x_0 \in \mathbb{R}^d$ ,  $\tau \in \mathbb{R}$ , and  $T > 0$ , define the *empirical measure*  $\mu_{\tau-T, \tau, x_0}$  as follows: for each Borel set  $S \in \mathcal{B}(\mathbb{R}^d)$ ,  $\mu_{\tau-T, \tau, x_0}(S)$  is the proportion of time  $t$  within the interval  $[\tau - T, \tau]$  for which  $x(t) \in S$ , where  $x(\cdot)$  is a solution of (1) with  $x(\tau - T) = x_0$ . In other words

$$\mu_{\tau-T, \tau, x_0}(S) = \frac{\ell(\{t \in [\tau - T, \tau] : \Phi^{(r)}(t, \tau - T, x_0) \in S\})}{T}.$$

**Definition 4** Given an attractor  $A^-$  for the past-limit system (6) and a pullback attractor  $\mathcal{A}$  of (1) starting at  $A^-$ , a *physical measure on  $\mathcal{A}$*  is a probability measure  $\mu$  on  $\mathcal{A}$ , such that  $\text{supp}(\mu_t) = \mathcal{A}(t)$  at each  $t \in \mathbb{R}$  and for Lebesgue-almost every  $x_0 \in B$ , the following holds: for each  $t \in \mathbb{R}$ , as  $T \rightarrow \infty$ , the empirical measure  $\mu_{t-T, t, x_0}$  converges weakly to  $\mu_t$ .

Let us emphasise that changing the value of the rate parameter  $r$  would change the pullback attractor  $\mathcal{A}$  and the physical measure  $\mu$  (or potentially cause them no longer to exist).

**Definition 5** Given an attractor  $A^-$  for the past-limit system (6) and a pullback attractor  $\mathcal{A}$  of (1) starting at  $A^-$ , a *pullback-PF-attracting physical measure* (or just “*pullback-attracting measure*” for short) on  $\mathcal{A}$  is a probability measure  $\mu$  on  $\mathcal{A}$ , such that  $\text{supp}(\mu_t) = \mathcal{A}(t)$  at each  $t \in \mathbb{R}$  and for every probability measure  $\nu_0$  of smooth density  $h$  supported within  $B$ , for each  $t \in \mathbb{R}$ , as  $T \rightarrow \infty$ , the measure  $\nu_T^t := \Phi^{(r)}(t, t - T, \nu_0)$  on  $\mathbb{R}^d$  computed as

$$\begin{aligned} \nu_T^t(S) &= \int_S (\mathcal{L}_{\Phi^{(r)}(t, t-T)} h)(y) dy \\ &= \int_{\mathbb{R}^d} \mathbb{1}_S(\Phi^{(r)}(t, t - T, x)) h(x) dx \end{aligned}$$

converges weakly to  $\mu_t$ . Here,  $\mathcal{L}_{\Phi^{(r)}(t, t-T)}: L^1(\mathbb{R}^d) \rightarrow L^1(\mathbb{R}^d)$  is the Perron–Frobenius operator for the map  $\Phi^{(r)}(t, t - T, \cdot)$ .

The significance of physical measures and pullback-attracting measures becomes clearer when considering them from a computational perspective. To simulate a pullback-attracting measure on a pullback attractor  $\mathcal{A}$  starting at  $A^-$ , we select a large ensemble of initial conditions from anywhere within a compact subset of the basin of attraction of  $A^-$ , choose a “starting time”  $t_0$  far back in the past, and then simulate the simultaneous evolution of the trajectories of these initial conditions forward in time from  $t_0$  according to (1). How far back in the past one needs to start depends on the chosen set of initial conditions themselves. This is clear, since, for any finite-time simulation, re-distributing the initial conditions will accordingly re-distribute the subsequent trajectories.

However, to simulate a physical measure that is not pullback-attracting, one cannot simply simulate the trajectories of initial conditions  $x_0$  starting at just one time sufficiently far back in the past. Rather, for an initial state  $x_0$ , one needs to take an ensemble of initial times  $t_0$  from which to evolve the trajectory  $x(t) = \Phi^{(r)}(t_0, t, x_0)$ . Specifically, this ensemble of initial times needs to be uniformly spread with a sufficiently high density across a sufficiently wide time-interval sufficiently far back in the past. However, for

<sup>2</sup> To be more precise, provided that the sets  $\mathcal{A}(t)$  are Borel subsets of  $\mathbb{R}^d$ , we can naturally equip  $\mathcal{A}$  with the  $\sigma$ -algebra  $\{\{x \in \mathcal{A} : x(t) \in S\} : S \in \mathcal{B}(\mathbb{R}^d)\}$  generated by the natural identification of  $\mathcal{A}$  with  $\mathcal{A}(t)$  by  $x \mapsto x(t)$ ; this  $\sigma$ -algebra does not depend on the time  $t$ . A probability measure  $\mu$  can then be defined on this  $\sigma$ -algebra.



this procedure for simulating physical measures, only one initial state  $x_0$  (rather than an ensemble of initial states) is needed. The distribution of the resulting ensemble of trajectories can be regarded as approximating the physical measure.

In [24], the question of when an (attracting) physical measure of the past-limit system can be extended to a (pullback-attracting) physical measure of the nonautonomous system is explored in detail.

### 3 Tipping between past and future attractors

Consider an asymptotically autonomous system (1) with some given rate parameter  $r$ , together with a given attractor  $A^-$  for the past-limit system (6), and suppose that the future limit system has several attractors. We will define the “probability of tipping” from the specified attractor of the past limit system to each of the attractors of the future limit system. More precisely:

- A1** Suppose that the past-limit system has an attractor  $A^-$  with basin of attraction  $B^-$ .
- A2** Suppose that the nonautonomous system (1) admits a pullback attractor  $\mathcal{A}$  starting at  $A^-$ , and that there is a physical measure  $\mu$  on this pullback attractor.
- A3** Suppose also that the future limit system has disjoint attractors  $A_1^+, \dots, A_{n_+}^+$  with basins of attraction  $B_1^+, \dots, B_{n_+}^+$ , respectively.

Since the parameter  $\lambda$  is shifting in real time and never necessarily reaches  $\lambda^+$  in finite time, we will need to assume that the attractors of the future limit system have some level of “robustness”:

- A4** For each  $j \in \{1, \dots, n_+\}$ , for every neighbourhood  $U$  of  $A_j^+$  with  $U \subset B_j^+$ , there exists a subneighbourhood  $V \subset U$  of  $A_j^+$  and  $T \in \mathbb{R}$ , such that for all  $s \geq T$ , we have  $\bigcup_{t \geq s} \Phi^{(r)}(t, s, V) \subset U$ .

With all this, we wish to define for each  $j \in \{1, \dots, n_+\}$  the probability of tipping from  $A^-$  to  $A_j^+$ . However, in general, it may be possible to “infinitely fine-tune” the parameters, and the shape of parameter shift in such a way that as  $t \rightarrow +\infty$ , a non-trivial proportion (defined by the physical measure  $\mu$ ) of the pullback attractor  $\mathcal{A}$  settles towards the boundaries of the basins of attraction for the future limit system; this would create difficulties in defining tipping probabilities. Therefore, we assume a non-degeneracy condition:

- A5** For all  $\varepsilon > 0$ , there is a neighbourhood  $O$  of  $\mathbb{R}^d \setminus \bigcup_{j=1}^{n_+} B_j^+$ , such that

$$\limsup_{t \rightarrow \infty} \mu_t(O) < \varepsilon.$$

The following theorem shows that it is possible to define transition probabilities from one attractor to another in an intuitive manner, at least under these assumptions.

**Theorem 6** *Suppose that assumptions A1–A5 above apply. Then*

- (A) *For each  $j \in \{1, \dots, n_+\}$ , the limit*

$$p_j := \lim_{t \rightarrow \infty} \mu_t(B_j^+)$$

*exists, and for every neighbourhood  $U$  of  $A_j^+$  with  $U \subset B_j^+$ , we have*

$$p_j = \lim_{t \rightarrow \infty} \mu_t(U).$$

- (B) *We have  $\sum_{j=1}^{n_+} p_j = 1$ .*

The proof is given in (Appendix A). We call the value  $p_j$  the *probability of tipping from  $A^-$  to  $A_j^+$*  for each  $j \in \{1, \dots, n_+\}$ . The next two sections illustrate that these tipping probabilities can be numerically estimated using appropriately chosen ensembles of initial conditions. In particular, since changing the rate  $r$  affects the pullback attractor  $\mathcal{A}$  and its physical measure  $\mu$ , we will explore how changing the rate  $r$  ultimately affects the probability of tipping in the examples that we will consider.

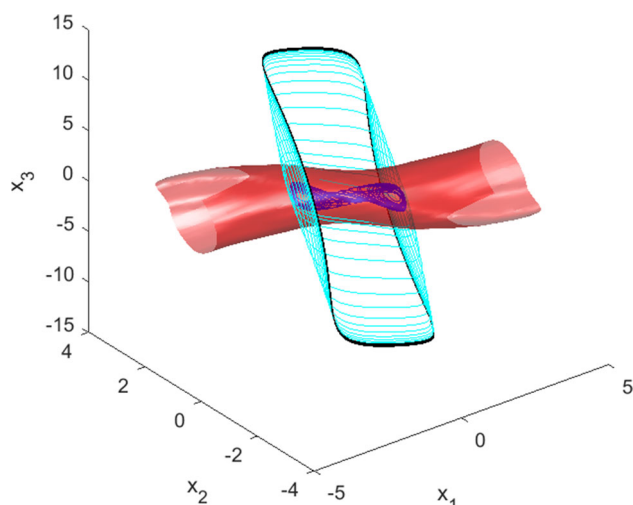
### 4 Example 1: double scroll with parameter shift

The double-scroll circuit, introduced by Chua *et al.* in [9], is a nonlinear analogue electronic circuit that exhibits bistability between two attractors for certain parameter values. The equations for the system are stated in [15, Sec. 16.5] as

$$\begin{aligned} \dot{x}_1 &= F_1(x_1, x_2, x_3) := a(x_2 - \phi(x_1)) \\ \dot{x}_2 &= F_2(x_1, x_2, x_3) := x_1 - x_2 + x_3 \\ \dot{x}_3 &= F_3(x_1, x_2, x_3) := -bx_2 \end{aligned} \tag{8}$$

for  $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ , with a cubic nonlinear function  $\phi(x_1) = x_1^3/16 - x_1/6$ . We fix values of the parameters  $a = 9, b = 14$  in [15] where this system is known to show bistability between two attractors: a “double scroll” chaotic attractor and a large-amplitude limit cycle.

Figure 1 illustrates the double-scroll chaotic attractor  $A_1$  in purple, and the large-amplitude periodic orbit  $A_2$  in black (with transient shown in cyan). The basin boundary between the two attractors is shown as a red tube—this is the stable manifold of a saddle periodic orbit (not shown). The attractors shown are obtained



**Fig. 1** Phase space of the double-scroll system (8) showing a chaotic attractor  $A_1$  (purple) enclosed within a tube-like basin of attraction with boundary shown in red. An initial condition outside the basin approaches (cyan) a large-amplitude limit cycle  $A_2$  (black)

by integrating initial conditions, while the basin boundary<sup>3</sup> is found by examining the fates of  $31^3$  initial conditions on a rectangular grid for the axes shown in Fig. 1.

We consider (8) undergoing a parameter shift that simply translates phase space in the direction  $(1, 1, 0)$ . More precisely, for  $x \in \mathbb{R}^3$ , we consider the system

$$\dot{x} = F(x - \Lambda(rt)(1, 1, 0)), \quad (9)$$

where

$$\Lambda(rt) = \frac{\sigma}{2} (1 + \tanh(rt)) \quad (10)$$

for real parameters  $r, \sigma$ . Note that (9) is a parameter shift [5] where  $\Lambda \rightarrow 0$  as  $t \rightarrow -\infty$  and  $\Lambda \rightarrow \sigma$  as  $t \rightarrow \infty$ . The effect of this is to shift the origin from  $(0, 0, 0)$  to  $(\sigma, \sigma, 0)$ . In all cases, the frozen systems are translated versions of (8) and so also have two attractors that are translations of  $A_1$  and  $A_2$ : we call these  $A_1(rt)$  and  $A_2(rt)$  and the limiting cases for  $t \rightarrow \pm\infty$  we write as  $A_{1,2}^\pm$ . Note that all of these (except  $A_{1,2}^-$ ) depend on  $\sigma$ .

In the slowly varying limit  $0 < r \ll 1$ , we conjecture that [1, Theorem III.1] can be applied to show there will be no tipping on the pullback attractor starting at  $A_1^-$ : the trajectories will track the branch  $A_1(rt)$ .

<sup>3</sup> We consider  $L(x) = 3x_2^2 + x_3^2$  and note that for  $L_0 = 8$  the surface  $L(x) = L_0$  strictly separates  $A_1$  and  $A_2$ . Hence we can compute the red tube by examining trajectories of  $31^3$  initial conditions  $x(0)$ : we calculate  $L(x(20))$  for (8) on this grid. The interpolated surface  $L(x(20)) = L_0$  from this grid gives an approximation of the boundary between the basins of attractors  $A_1$  and  $A_2$ .

In the limit  $r \gg 1$ , (9) has discontinuous right-hand side:

$$F(x - \Lambda(rt)(1, 1, 0)) = \begin{cases} f^-(x) & \text{for } t < 0 \\ f^+(x) & \text{for } t > 0, \end{cases} \quad (11)$$

where

$$f^-(x) = F(x), \quad f^+(x) = F(x - (\sigma, \sigma, 0)).$$

Hence, in this limit, the probability of tipping from  $A_1^-$  to  $A_2^+$  corresponds to the proportion of natural measure on  $A_1^-$  in the basin of attraction of  $A_2^+$ .

Figure 2 shows time-evolution of an ensemble of typical simulations of (9) for 200 initial conditions chosen at time  $T = -40$  uniformly and independently distributed in a cube within the basin of attraction of  $A_1^-$ . The parameters used are

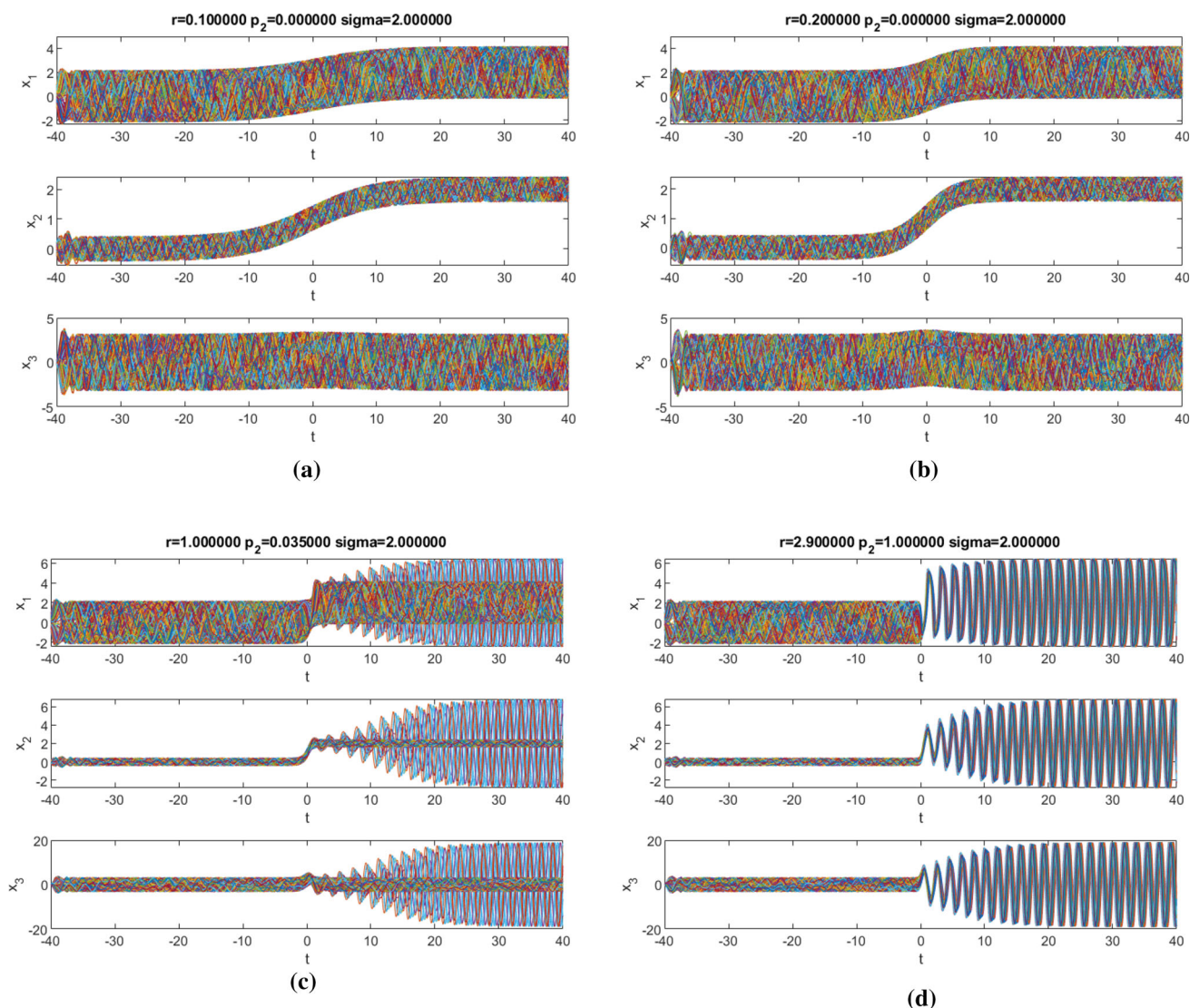
$$a = 9, \quad b = 14, \quad \sigma = 2, \quad (12)$$

and four rates  $r$  are used. During the parameter shift, some proportion of the trajectories follow the chaotic attractor and are asymptotic to the chaotic  $A_1^+$ , while others are asymptotic to the periodic  $A_2^+$ . The proportion that switch corresponds to the tipping probability  $p_2$  discussed in Sect. 3 (with  $A^-$  taken as the chaotic  $A_1^-$ ); theoretically, this is based on the assumption that there is a pullback-attracting measure on the pullback attractor starting at  $A_1^-$ . We clearly see that this proportion depends on the rate  $r$ . In Fig. 2a,b, all trajectories follow the branch of chaotic attractors and limit to  $A_1^+$ ; in (c), there is a mixture of tipping and not tipping (*partial tipping* in the terminology of [1]), and in case (d), all trajectories tip to the periodic attractor.

Each panel of Fig. 3 uses 400 simulations, with initial conditions chosen in the same manner as for Fig. 2, to estimate the tipping probability  $p_2$  for this system. We vary  $r$  but leave other parameters as in (12). From Fig. 3a, we infer the presence of two critical rates  $r_1 \approx 0.8$  and  $r_2 \approx 2.3$ , such that:

- For  $0 < r < r_1$ , the pullback attractor with past limit  $A_1^-$  has as its future limit the chaotic attractor  $A_1^+$ . This implies that  $p_1 = 1$  and  $p_2 = 0$ .
- For  $r_1 < r < r_2$ , there is partial tipping: some of the natural measure for the pullback attractor with past limit  $A_1^-$  has future limit supported on the chaotic attractor  $A_1^+$ , while the remainder has future limit supported on  $A_2^+$ . This implies  $0 < p_1 < 1$  and  $0 < p_2 < 0$ .
- For  $r > r_2$ , there is total tipping in the sense of [1]: the pullback attractor with past limit  $A_1^-$  has as its future limit the periodic attractor  $A_2^+$ . This implies that  $p_1 = 0$  and  $p_2 = 1$ .

This third case suggests in particular that for  $\sigma = 2$ ,  $A_1^-$  is contained within the basin of attraction of  $A_2^+$ . Fig. 3b shows analogous results to Fig. 3a, but for the



**Fig. 2** A sample of 400 initial conditions  $x(-T)$  for  $T = 40$  is chosen uniformly distributed within  $[-1, 1] \times [-0.25, 0.25]^2$ , and is then evolved up to  $T$  according to (9) with  $a = 9$ ,  $b = 14$ , and shift by  $\sigma = 2$  at rates **a**  $r = 0.1$ , **b**  $r = 0.82$ , **c**  $r = 1$ , and **d**  $r = 2.98$ . Each simulation is shown in a different colour. Note that all initial conditions

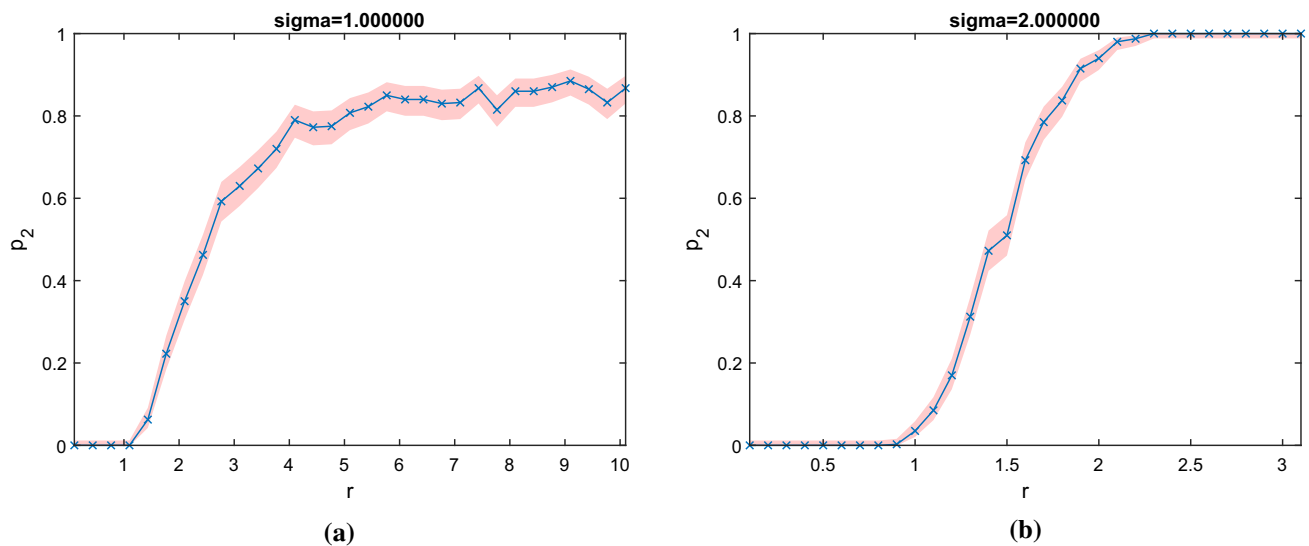
rapidly approach the double-scroll attractor, but at  $t \sim 0$ , the system changes to a translated version of the same system. In cases **(a,b)**, all states continue to follow the chaotic attractor. For **(c)**, only a proportion follows this attractor, while the remainder are asymptotic to the stable limit cycle. For **(d)**, all states are asymptotic to the stable limit cycle

case that  $\sigma = 1$  rather than  $\sigma = 2$ . In this case, we see the presence of a critical rate  $r_1 \approx 1.4$ , such that  $0 < r < r_1$  implies no tipping and  $r > r_1$  implies partial tipping; as  $r$  increases sufficiently, the probability of tipping seems to plateau around a value strictly less than 1; this suggests that for  $\sigma = 1$ , the past chaotic attractor  $A_1^-$  intersects the boundary between the basin of attraction of  $A_2^-$  and the basin of attraction of  $A_2^+$ . The simulations in Fig. 3 depend only weakly on choice of initial  $T$  or initial condition distribution, as long as  $T$  is chosen far enough in the past; this is demonstrated in Table 1, and is consistent with the notion that the physical measure on the pullback attractor starting at  $A_1^-$  is pullback-attracting. More specifically, we require

$T$  to be sufficiently negative that the initial distribution is well mixed by the time the parameter shift is underway.

Figure 4 illustrates, for different times  $T$ , a two-dimensional section ( $x_1 = 0$ ) of the set of points  $x \in \mathbb{R}^3$  for which  $\Phi^{(r)}(t, T, x)$  tends towards the future chaotic attractor  $A_1^+$  as  $t \rightarrow \infty$ ; these are all for parameter values  $\sigma = 2$  and  $r = 1.5$ , where we have partial tipping. Observe that for  $T = -10$ , there is very high sensitivity on initial state.

Figure 5 shows a numerical approximation (red) of the section  $x_1 = 0$  of the location at time  $T = -10$  (before the parameter shift is significantly underway) of the pullback attractor starting at  $A_1^-$ . Again, the



**Fig. 3** Estimate of the tipping probability from chaotic  $A_1^-$  to periodic  $A_2^+$  attractor for the double-scroll system with parameter shift (9,12) using an ensemble of 400 initial conditions  $x(-T)$  as in Fig. 2 started at time  $T = 40$ . Case (a) is for  $\sigma = 1$ : in this case, the asymptotically fast shift still gives partial tipping, corresponding to the attractor  $A_1^-$

partially intersecting the basin of  $A_1^+$ . Case (b) is for  $\sigma = 2$ : observe that for small enough  $r$ , there is tracking of the chaotic attractor, for a range  $r_1 < r < r_2$ , there is partial tipping, and for  $r > r_1$ , the tipping is total. A 95% confidence interval is given using the Agresti–Coulter estimate

**Table 1** Comparison of  $p_2$  computed for 800 initial conditions, otherwise as in Fig. 3 for  $\sigma = 2$  and  $r = 1.5$ . Initial conditions  $x(-T)$  are chosen uniformly in  $[-1, 1] \times [-z, z]$  for selection of  $T$  and  $z$ . For large enough  $T$ , the values of  $p_2$  obtained (95% confidence interval  $\pm 0.034$ ) are consistent with being independent of initial distribution parametrized by  $z$  by virtue of the initial conditions having become well mixed on the attractor

	T=10	T=40	T=80	T=120
$z=0.0025$	0.7351	0.5386	0.5871	0.5858
$z=0.025$	0.7104	0.5534	0.5560	0.6007
$z=0.25$	0.6381	0.5597	0.5896	0.5759.

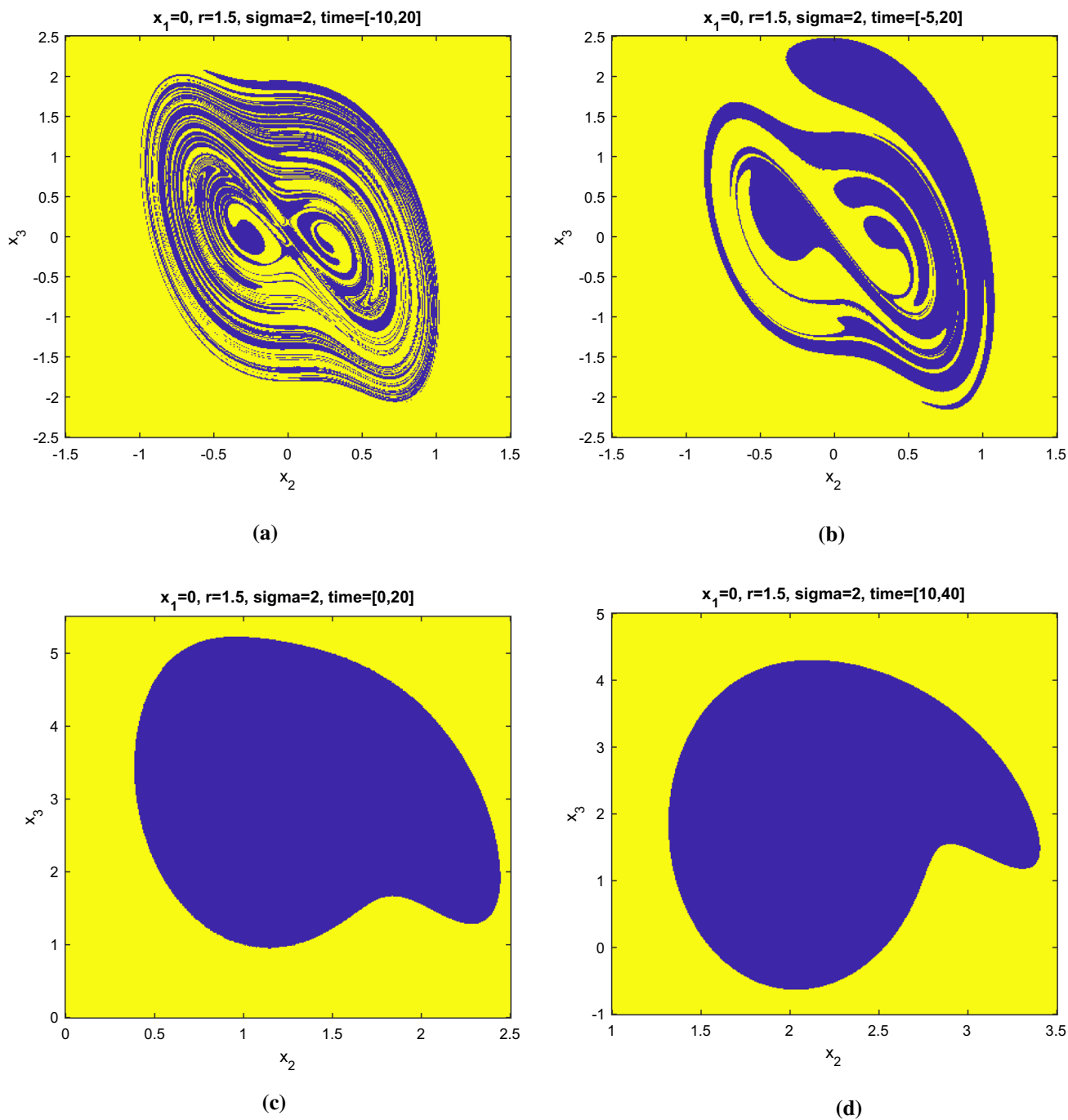
parameters are as in (12), with three increasing values of  $r$ . Panel (a) indicates that there is an open neighbourhood of the pullback attractor that is asymptotic to  $A_1^+$ . Panels (b) and (c) support the partial tipping seen in Fig. 3, namely that there is a non-trivial partition of the pullback attractor into parts that escape and parts that do not. For larger negative  $T$ , these become ever more intertwined.

The numerical results presented, and the fact that they depend little on choice of  $T$  or distribution size, support the hypothesis that the ensemble gives a reasonable approximation of the natural measure on the pullback attractor. It seems that the conclusions of Theorem 6 hold, even though the complexity of the system means that validity of assumptions **A1–A5** cannot be easily checked in this case.

## 5 Example 2: Lorenz–Stommel with parameter shift

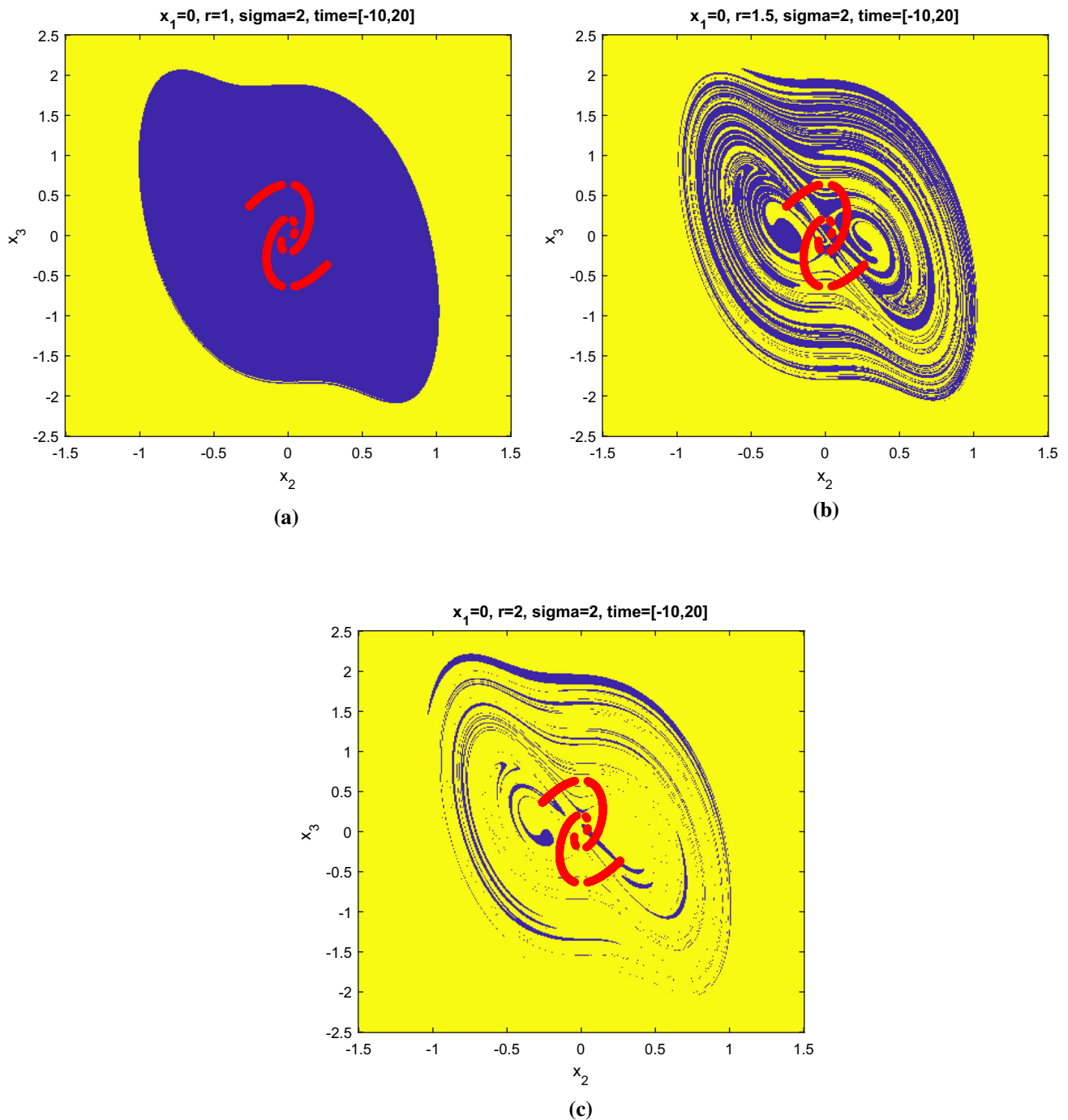
As a second example, we introduce and examine a conceptual model of AMOC tipping, based on the 2-box model of Stommel [25] with added “subgrid” chaotic forcing (a Lorenz model [21] that conceptually represents natural variability due to large-scale ocean transport variability) and an additional enforced change (a parameter shift modelling anthropogenic change imposed on the forcing). Previous work (reviewed for example in [10]) has highlighted the Stommel system as one that is susceptible to tipping under changes in the forcing. Previous work has also shown that these thresholds may depend not just on forcing levels but also rate and details of time-dependent forcing [3, 7].





**Fig. 4** Numerically approximated section ( $x_1 = 0$ ) through the set of initial conditions at a time  $T$  that are asymptotic to the future chaotic attractor  $A_1^+$  (blue) and the future periodic attractor (yellow), for the double-scroll system with parameter shift (9,12) for  $\sigma = 2$  and  $r = 1.5$ . In panels (a–d) are  $T = -10$ ,  $T = -5$ ,  $T = 0$ , and  $T = 10$ ,

respectively. We scan a  $400^2$  grid of initial values of  $x_2$  and  $x_3$ , and note that final state is assumed to be achieved for  $T \geq 20$ . Going backwards in time [from (d) back to (a)], the solutions of the nonautonomous system (9) that limit to each of the two attractors become increasingly intertwined



**Fig. 5** Section ( $x_1 = 0$ ) through the set of initial conditions at a time  $T = -10$  that are asymptotic to the future chaotic attractor  $A_1^+$  (blue) and the future periodic attractor (yellow), for the double-scroll system with parameter shift, calculated as in Fig. 4, for different rates **a**  $r = 1$ , **b**  $r = 1.5$ , **c**  $r = 2$ . In each case, the red region is the  $x_1 = 0$  section

intersection of the pullback attractor at  $T = -10$ . In **(a)**, all points in the pullback attractor are asymptotic to the future chaotic attractor. In **(b, c)**, there is partial tipping that manifests as a non-trivial partition of points near the pullback attractor that go to each of the future limit states

Without the shift, the model we consider is

$$\begin{aligned} \dot{x}_1 &= G_1(x) := \mu(x_2 - x_1) \\ \dot{x}_2 &= G_2(x) := x_1(\rho - x_3) - x_2 \\ \dot{x}_3 &= G_3(x) := x_1x_2 - \beta x_3 \\ \dot{x}_4 &= G_4(x) := \xi(x_1) - x_4(1 + |x_4 - x_5|) \\ \dot{x}_5 &= G_5(x) := \eta(x_1) - x_5(\zeta + |x_4 - x_5|). \end{aligned} \tag{13}$$

The first variables  $(x_1, x_2, x_3)$  are a standard Lorenz system [21]; note that we are not using this as a model of atmospheric motion, but simply as a model of chaotic variability. We assume the Lorenz system parametrically forces (via one of the variables—we take  $x_1$ ) a reduced non-dimensionalized Stommel model [25] where  $x_4$  corresponds to temperature gradient and  $x_5$  to salinity gradient between north and tropical Atlantic regions. We set

$$\xi(x) = \xi_1 + ax, \quad \eta(x) = \eta_1 + ax \tag{14}$$

and choose parameters

$$\xi_1 = 3.0, \quad \eta_1 = 1.0, \quad \zeta = 0.3, \quad \rho = 28, \quad \mu = 10, \quad \beta = 8/3. \tag{15}$$

The parameter  $a$  parametrises the amplitude of forcing of the Stommel variables by the Lorenz variables. For  $a = 0$ , there is bistability of the  $(x_4, x_5)$  system (see, e.g., [10]) between “AMOC on” and “AMOC off” states. For  $a$  small enough, the bistability is preserved, with two chaotic attractors close to these states and a chaotic saddle that is close to a saddle on the basin boundary for the  $a = 0$  system.

We now include a parameter shift by considering the ODE on  $x \in \mathbb{R}^5$  given by

$$\dot{x} = G(x - \Lambda(rt)v), \tag{16}$$

where  $G$  is defined as (13) and

$$\Lambda(rt) = \frac{\sigma}{2} (1 + \tanh(rt)), \tag{17}$$

and there are parameters  $r$  and

$$\sigma = 10, \quad v = (1, 0, 0, 0, 1). \tag{18}$$

(As in the previous section, this form of parameter shift is simply chosen for illustrative purposes as a mathematically straightforward example of a shift between two values, not for any particular physical significance.) Writing (16) in translated variables  $y = x - \Lambda(rt)v$ , note that

$$\dot{y} = G(y) - r\Lambda'(rt)v. \tag{19}$$

Note that the  $r = 0$  case of (19) coincides with (13). We will use an observable [10], such as

$$\Psi = y_4 - y_5 = x_4 - x_5 + \Lambda(rt)$$

to distinguish between attractors: if  $a$  is not too large, then the frozen systems of (16) have a chaotic attractor  $A_1(rt)$  corresponding to “AMOC on” state, where  $\Psi \approx 0.7$ , and also a chaotic attractor  $A_2(rt)$  corresponding to “AMOC off” state, where  $\Psi \approx 0$ . As in Sect. 4, we refer to the limiting cases as  $A_{1,2}^\pm$  for  $t \rightarrow \pm\infty$ .

In the case that  $r$  is small, just as in the example considered in Sect. 4, there is no tipping from whichever attractor the system starts on in the past limit. For larger  $r$ , the time-localised perturbation (namely  $-r\Lambda'(rt)v$ ) to (16) near  $t = 0$  may result in tipping.

We perform numerical simulations for this system both for individual and ensembles of initial conditions. These show rate dependence introduced by the parameter shift as well as initial condition dependence from the chaos that can lead to non-trivial tipping probabilities much as for the example in Sect. 4.

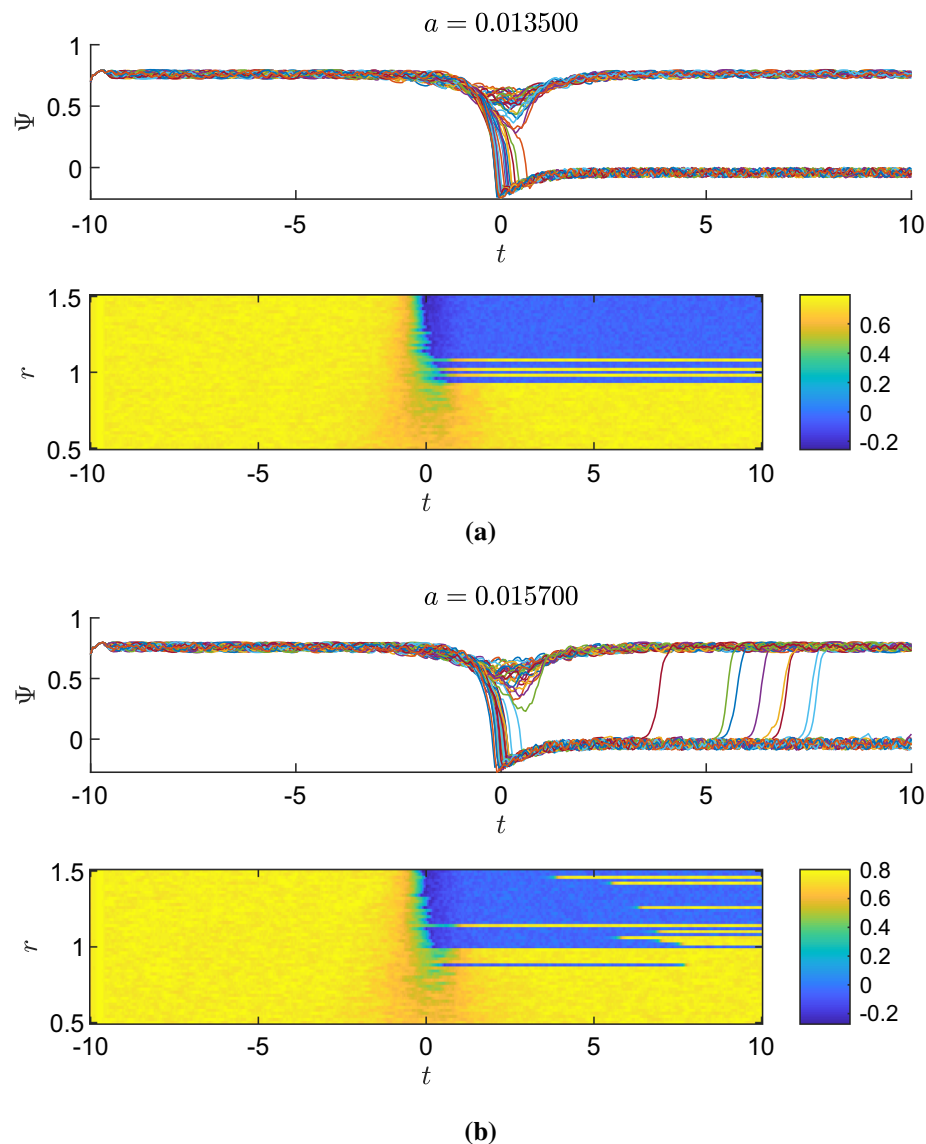
Figure 6a illustrates runs of (16) with shift as in (17) for fixed  $a = 0.135$  and a range of values of  $r$ , starting at the same one initial condition close to the ‘on’ state at time  $T = -10$ . The top panel shows the time series of the AMOC index  $\Psi$  for the range of values of  $r$ . The colour map (bottom panel) shows regions in yellow where the system is close to  $A_1$  and in blue where it is close to  $A_2$ . Both plots appear to indicate an interval of  $r$ -values for which it is “random” as to whether the trajectory will tip or not. This suggests partial tipping in which the question of whether a given trajectory  $x(t)$  will tip or not depends very sensitively on its initial condition in the distant past.

Figure 6b shows similar runs for a higher value of the amplitude of forcing  $a = 0.0157$ . In this case, there can be an escape from the “on” state, but this is apparently only transient at this amplitude, because the “off” attractor has undergone a crisis [14] leaving only a long chaotic transient near the off state.

For the same case as shown in Fig. 6a ( $a = 0.0135$ ) for an ensemble of initial conditions at  $T = -10$ , Fig. 7 now shows how one can quantify  $p_2$ , the probability of tipping from  $A_1^-$  to  $A_2^+$ , as a function of  $r$ . Similar to the example in Sect. 4, the probability of escape is apparently a continuous function of the rate  $r$  that increases from 0 to 1 over a range of values of  $r$  where there is partial tipping. This manifests itself as the random “striping” over the range of  $r$  with partial tipping in the colour map shown in Fig. 6a.

## 6 Discussion

This paper contributes to our understanding of tipping points in systems that are subject to nonautonomous variations of parameters and where chaotic attractors may be present. It has been previously rec-



**Fig. 6** Rate dependence for the Stommel–Lorenz model (16) with shift as in (17) for cases of (a) irreversible rate-dependent tipping  $a = 0.0135$  and (b) transient rate-dependent tipping  $a = 0.0157$ : for other parameters and details, see text. The top panels show time series of the AMOC index  $\Psi = x_4 - x_5 + \Lambda(rt)$  vs. time starting with the same initial condition  $x(-T)$  in the “on” state ( $\Psi \approx 0.7$ ) for  $T = 20$  and a range of values of  $r$ . Note chaotic oscillations before and after the parameter shift near  $t = 0$ . The bottom panels shows the value of  $\Psi$  vs.  $r$

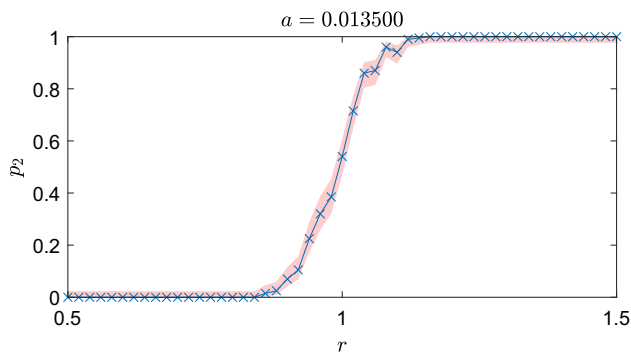
and  $t$  for the simulations in the top panel. For case (a) as a result of the parameter shift, some trajectories tip to the “off” state ( $\Psi \approx 0$ ) depending on the rate  $r$  and the state of the subgrid chaos at the time of tipping. For case (b), the amplitude  $a$  of subgrid chaotic forcing seems to be just above the critical amplitude for the destruction of the ‘off’ attractor, and so, the lower state is merely a transient ‘off’ state near where the ‘off’ attractor would have been if  $a$  were slightly less; trajectories eventually escape from this transient ‘off’ state and return to the ‘on’ attractor

ognized [8, 11, 17] that there will be physically relevant measures on pullback/snapshot attractors, but except in the case of ergodic forcing (Random Dynamical Systems), it is not obvious how to define a nonautonomous physical measure in a natural way. By restricting to nonautonomous systems that have autonomous limits in past and future times (so-called parameter shifts), we are able to give a deeper understanding of local pullback attractors and the statistics of trajectories which are governed by physical measures of the pullback attrac-

tors. In particular, Sect. 2 suggests an appropriate definition for these, and [24] gives criteria for existence of such nonautonomous physical measures.

The nonautonomous physical measures that start from particular past attractors are used in Sect. 3 to quantify the probability of tipping from one attractor to another. Although a number of technical assumptions are necessary to prove the main theoretical result (Theorem 6), the examples of the Double Scroll and chaotic Stommel models in Sects. 4 and 5 suggest that in prac-





**Fig. 7** Probability  $p_2$  of tipping from  $A_1$  to  $A_2$  as a function of rate  $r$  for the Stommel–Lorenz model in the case  $a = 0.0135$ , as shown in Fig. 6, computed using an ensemble of 200 randomly generated initial conditions close to a point in the basin of the “AMOC on” state for each value of  $r$ . Note that there is a range of rates for the parameter shift where the probability of tipping moves roughly monotonically from 0 to 1

tice, the result will apply even when verification of the technical assumptions is very difficult. In most studies, time-variable subgrid forcing is modelled by inclusion of a stochastic term; by contrast, the chaotic Stommel model discussed in Sect. 5 suggests a novel way to approach this. In particular, regarding the influence of subgrid forcing upon climate tipping events and their probabilities, Sect. 5 provides observations and insights that would not readily be gained through the traditional approach of modelling subgrid forcing as uncorrelated Gaussian noise.

Apart from verifying that the technical assumptions apply in the examples, there are a number of problems that remain to be tackled. First, it is not clear how to define natural measures in a useful, natural way for more general nonautonomous systems that do not have an autonomous limit in the past. Second, although we believe that the assumptions **A1–A5** are generic in some sense, there are special cases where these will be violated. For example, in some cases of tipping, the pullback attractor starting at some  $A^-$  may limit to a future saddle—in such a case, the physical measure on the pullback attractor may not be able to say much about typical trajectories.

Finally, it is a natural question to try and compute the critical rates (or other possible parameters of the shape of  $\lambda(t)$ ) for the onset of phenomena such as partial tipping or total tipping. For example, examining the dynamics of connecting orbits in an extended system has successfully been used to understand rate-dependent tipping between equilibria [4] and periodic orbits [1]. In the more general case of chaotic attractors, these critical rates will be associated with connections between attractors and basin boundaries.

**Acknowledgements** This is TiPES contribution #70. This project received funding from the European Union’s Horizon 2020 research and innovation programme under

grant agreement No 820970 (TiPES). Matlab code for a selection of the simulations and an animation of Fig. 1 are available from: <https://github.com/peterashwin/ashwin-newman-2021>.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

**Author contribution statement**

PA suggested formulating physical measures and tipping probabilities for asymptotically autonomous dynamical systems. PA and JN worked together on the rigorous definitions in Sect. 2, and on the setup in Sect. 3 as specified in **A1–A5**. JN formulated and proved Theorem 6. PA carried out the numerics in Sects. 4 and 5, and wrote up a first draft of these sections; then, PA and JN worked together on refining the discussion of the numerics.

**A Proof of Theorem 6**

(A) Suppose the statement is false. Therefore, there is a  $j \in \{1, \dots, n_+\}$ , a neighbourhood  $U$  of  $A_j^+$  with  $U \subset B_j^+$ , a sequence of times  $t_n \nearrow \infty$ , and distinct numbers  $q_1 < q_2$ , such that

$$\mu_{t_{2n}}(U) \rightarrow q_1 \quad \text{and} \quad \mu_{t_{2n+1}}(B_j^+) \rightarrow q_2 \quad (20)$$

as  $n \rightarrow \infty$ . On the basis of the “robustness” assumption **A4**, take an open neighbourhood  $V \subset U$  of  $A_j^+$  and  $T_1 \in \mathbb{R}$ , such that for all  $s \geq T_1$ , we have  $\bigcup_{t \geq s} \Phi^{(r)}(t, s, V) \subset U$ . On the basis of the non-degeneracy assumption **A5**, take an open neighbourhood  $O$  of  $\mathbb{R}^d \setminus \bigcup_{j=1}^{n_+} B_j^+$  and  $T_2 \in \mathbb{R}$ , such that  $\sup_{t \geq T_2} \mu_t(O) \leq \frac{q_2 - q_1}{2}$ . Let  $C$  be a closed bounded set, such that  $\bigcup_{t \in \mathbb{R}} \mathcal{A}(t) \subset C$ , and let  $K = C \cap B_j^+ \setminus O$ . Since  $O$  is an open neighbourhood of the boundary of  $B_j^+$ , the set  $K$  is itself closed. Let  $\Delta > 0$  be such that  $\varphi_{\lambda^+}(\Delta, K) \subset V$ . Since  $f(x, \lambda)$  is continuous in  $(x, \lambda)$  and  $A(rt) \rightarrow \lambda^+$  as  $t \rightarrow \infty$ , we can find  $T_3 \in \mathbb{R}$ , such that for all  $s \geq T_3$ ,  $\Phi^{(r)}(s, s + \Delta, K) \subset V$ . So then, for all  $s \geq \max(T_2, T_3)$ , for all  $t \geq \max(T_1, s + \Delta)$ , we have

$$\mu_t(U) \geq \mu_s(B_j^+) + \frac{q_2 - q_1}{2},$$

but this contradicts (20). (B) The non-degeneracy assumption **A5** implies that

$$\mu_t \left( \mathbb{R}^d \setminus \bigcup_{j=1}^{n_+} B_j^+ \right) \rightarrow 0$$

as  $t \rightarrow \infty$ , and hence, it follows that  $\sum_{j=1}^{n_+} p_j = 1$ .

## References

- Hassan Alkhayuon, Peter Ashwin, Rate-induced tipping from periodic attractors: Partial tipping and connecting orbits. *Chaos Interdiscip. J. Nonlinear Sci.* **28**(3), 033608 (2018)
- Hassan Alkhayuon, Peter Ashwin, Weak tracking in nonautonomous chaotic systems. *Phys. Rev. E* **102**(5), 052210 (2020)
- Hassan Alkhayuon, Peter Ashwin, Laura C. Jackson, Courtney Quinn, Richard A. Wood, Basin bifurcations, oscillatory instability and rate-induced thresholds for Atlantic meridional overturning circulation in a global oceanic box model. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **475**(2225), 20190051 (2019)
- Peter Ashwin, Clare Perryman, Sebastian Wieczorek, Parameter shifts for nonautonomous systems in low dimension: bifurcation- and rate-induced tipping. *Nonlinearity* **30**(6), 2185 (2017)
- Peter Ashwin, Sebastian Wieczorek, Renato Vitolo, Peter Cox, Tipping points in open systems: bifurcation, noise-induced and rate-dependent examples in the climate system. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **370**(1962), 1166–1184 (2012)
- Rufus Bowen, David Ruelle, The ergodic theory of Axiom A flows. *Invent. Math.* **29**, 181–202 (1975)
- D. Castellana, Sven Baars, Fred W. Wubs, H.A. Dijkstra, Transition probabilities of noise-induced transitions of the Atlantic Ocean circulation. *Sci. Rep.* **9**, 20284 (2019)
- Mickaël D. Chekroun, Eric Simonnet, Michael Ghil, Stochastic climate dynamics: Random attractors and time-dependent invariant measures. *Phys. D Nonlinear Phenom.* **240**(21), 1685–1700 (2011)
- L. Chua, M. Komuro, T. Matsumoto, The double scroll family. *IEEE Trans. Circ. Syst.* **33**(11), 1072–1118 (1986)
- Henk A. Dijkstra, *Nonlinear Climate Dynamics* (Cambridge University Press, Cambridge, 2013)
- Gábor Drótos, Tamás Bódai, Tamás Tél, On the importance of the convergence to climate attractors. *Eur. Phys. J. Spec. Top.* **226**(9), 2031–2038 (2017)
- J.P. Eckmann, D. Ruelle, Ergodic theory of chaos and strange attractors. *Rev. Mod. Phys.* **57**, 617–656 (1985)
- Michael Ghil, Eric Simonnet, *Geophysical Fluid Dynamics, Nonautonomous Dynamical Systems, and the Climate Sciences* (Springer, Cham, 2020), pp. 3–81
- Celso Grebogi, Edward Ott, James A. Yorke, Crises, sudden changes in chaotic attractors, and transient chaos. *Phys. D Nonlinear Phenom.* **7**(1–3), 181–200 (1983)
- Morris W. Hirsch, Stephen Smale, Robert L. Devaney, *Differential Equations, Dynamical Systems, and an Introduction to Chaos* (Elsevier Inc., Amsterdam, 2013)
- Bálint Kaszás, Ulrike Feudel, Tamás Tél, Tipping phenomena in typical dynamical systems subjected to parameter drift. *Sci. Rep.* **9**(1), 1–12 (2019)
- Bálint Kaszás, Tímea Haszpra, Mátyás Herein, The snowball earth transition in a climate model with drifting parameters: Splitting of the snapshot attractor. *Chaos Interdiscip. J. Nonlinear Sci.* **29**(11), 113102 (2019)
- P.E. Kloeden, C. Pötzsche, *Nonautonomous Dynamical Systems in the Life Sciences*, Lecture Notes in Mathematics (Springer, Berlin, 2014)
- Peter E. Kloeden, Martin Rasmussen, *Nonautonomous Dynamical Systems* (American Mathematical Soc., USA, 2011)
- Timothy M. Lenton, Hermann Held, Elmar Kriegler, Jim W. Hall, Wolfgang Lucht, Stefan Rahmstorf, Hans Joachim Schellnhuber, Tipping elements in the Earth's climate system. *Proc. Natl. Acad. Sci.* **105**(6), 1786–1793 (2008)
- Edward N. Lorenz, Deterministic nonperiodic flow. *J. Atmos. Sci.* **20**(2), 130–141 (1963)
- G. Manjunath, P. Tino, H. Jaeger, Theory of input driven dynamical systems. European symposium on artificial neural networks, computational intelligence and machine learning, Bruges, p 1–12, (2012)
- Esteban Muñoz-Young, Andrés Navas, Enrique Pujals, Carlos H. Vásquez, A continuous Bowen-Mane type phenomenon. *Discret. Continuous Dyn. Syst. A* **20**(3), 713–724 (2007)
- Newman J, Ashwin P, Theory of nonautonomous natural measures. (in Preparation), (2020)
- Henry Stommel, Thermohaline convection with two stable regimes of flow. *Tellus* **13**(2), 224–230 (1961)
- Wieczorek S, Xie C, Chris KRT. Jones, compactification for asymptotically autonomous dynamical systems: theory, applications and invariant manifolds (2020). [arXiv:2001.08733v2](https://arxiv.org/abs/2001.08733v2)
- Lai-Sang Young, What are SRB measures, and which dynamical systems have them? *J. Stat. Phys.* **108**(5–6), 733–754 (2002)
- Lai-Sang Young, Generalizations of SRB measures to nonautonomous, random and infinite dimensional systems. *J. Stat. Phys.* **166**, 4924–515 (2016)