

# 1 Introduction

Complex dynamical systems have found widespread use in the modelling of real-world phenomena given their capability of describing the collective, macroscopic properties of those systems made of a multitude of interactive components. The analysis of complex systems has thus become particularly useful for scientific areas such as (s.a.) climate science, ecosystems, biology, sociology and economics. The non-linear nature of the models makes the rigorous characterisation of these systems notoriously challenging as it enables the emergence of specific properties (self-organising pattern formation and onset of chaos to cite a few) that are shared among very different scientific fields. Recently however, some specific communities (notably ecology and climate science) have focused their attention on one of those property in particular, that is the existence of abrupt dynamical regimes shifts often referred to as critical transitions or tipping points. Loosely speaking a complex system, whose time-evolution may be deterministic or stochastic, can undergo a tipping point whenever its state variables abruptly shift away from a stable equilibrium and transition to a new regime that may have catastrophic implications for the properties of the system itself. It is therefore unsurprising that the reliable and timely prediction of such critical thresholds has attracted much attention from ecologists and climate scientists in the past two decades with disrupting phenomena noticed experimentally in a range of catastrophic events in the past and that are likely to happen in the near future as well. This report concerns the investigation of indicators of critical transitions in complex dynamical systems, i.e. the identification of robust, measurable and detectable quantities that can provide early-warning signals (EWS) of incoming catastrophic tipping events.

## Structure

This document is organised in 4 Chapters detailing and motivating the research of EWS. The content is structured as follows: in the present (first) Chapter we will outline the incentives for the study and analysis of these precursors (Section 1.1), as provided by empirical evidence from the natural world, followed by a detailed review of the history of EWS proposed in the last 25 years (Section 1.2), their mathematical foundation and the scientific disciplines in which they have been successfully detected; in the second Chapter we will review the basic, fundamental theories upon which the research of catastrophic events is based on, namely differential equations (Sections ?? and ??) and stochastic calculus (Section ??); in the third Chapter we focus on the construction of a mathematical framework for critical transitions and their EWS in low-dimensional (Section 3.1) and high-dimensional (Section 3.2) systems; this will motivate the research itself reported in the following Chapters, in particular in the fourth Chapter we address limitations of current methods in finding precursors in terms of lack or universality given their *top-down* nature; Chapter 5 concerns the investigation of spatially-extended systems and the rich dynamics that allows the formation of patterned instabilities.

## 1.1 Motivation

As briefly introduced above, critical transitions are often associated to catastrophic and abrupt regime shifts observed in a variety of natural and human complex systems. As the system evolves by (slowly) tracking a stable equilibrium a sudden and often unforeseen change in the dynamics brings the state observables to be repelled away from the attractor in a rapid fashion. The state becomes unstable and it evolves much more quickly with respect to (w.r.t.) the timescale that characterised its dynamics prior to the regime shift. Despite the rather severe tone, these type of events appear to be ubiquitous in complex systems. One classical example is given by the global tipping points of the Earth's climate [43] driven by anthropogenic climate change, however paleoclimatic extreme events [25] s.a. the end of the Greenhouse

Earth at around 34 million years ago (Figure 1(a)), the end of the last ice age that occurred about 10,000 years ago (Figure 1(b)) and the most recent end of the African humid period at approximately 3000 BCE (Figure 1(c)) all happened without human input. Similarly to climate critical thresholds, catastrophic regime shifts also appear quite frequently as population collapses of living, interacting organisms observed in natural ecosystems [11] as well as in-vitro (Figure 1(d)) and in-silico [20] lab experiments. In addition, seemingly unrelated natural sciences s.a. physics, biology and medicine also feature similar tipping points. Evidence of abrupt transitions were found recently in magnetic quantum phase transitions (Figure 1(e)). In vital human biological systems some examples are the exacerbation of ventilation defects prior to the onset of asthma [18], pre-ictal dynamic changes before epileptic seizures [10, 13], mood instabilities symptomatic of major depressive disorder (MDD) [57] (albeit lacking substantial empirical proof [54]) and pathological sleep-quality disruptions [71]. Finally human sciences have also started to recently characterise disruptive phenomena as critical transitions with particular emphasis being pursued in real [78] and simulated [58, 66] social networks and opinion dynamics [19] as well as in the study of the collapse of financial markets (Figure 1(f)) in complex macroeconomic systems [63, 75]. Given this wide and impactful range of different phenomena, the topic of reliably detecting indicators of critical transitions has attracted ever growing attention and in the past two decades several efforts have tried to associate signs of incoming criticalities with measurable quantities as we describe in the next Section.

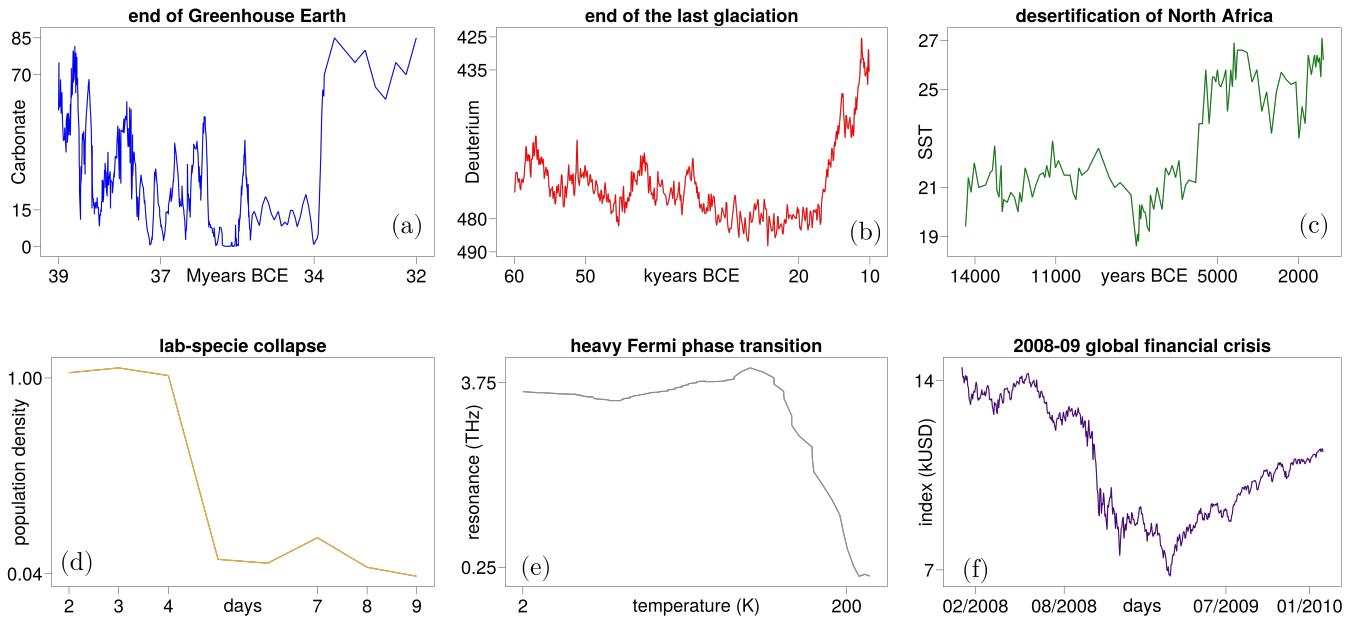


Figure 1: Examples of real-world timeseries exhibiting a sudden regime shift (tipping event): (a) transition from Greenhouse to Icehouse Earth measured by calcium carbonate ( $\text{CaCO}_3$ ) concentration in sediments of the tropical Pacific; (b) end of the last glacial period (LGP), colloquially referred to as the last ice age, detected as a transition of Deuterium ( ${}^2\text{H}$ ) concentrations in the ice core of Vostok, Antarctica; (c) abrupt end of the African humid period and onset of desertification from mean sea surface temperature (SST) measured in oceanic drilled holes off the cost of North-East Africa; (d) population collapse of a lab-grown ecosystem of budding yeast; (e) quantum phase transition from antiferromagnetic to heavy Fermi liquid of  $\text{YbRh}_2\text{Si}_2$  measured as fermionic breakdown causing an abrupt decay of the resonant (spectroscopy) frequency; (f) global financial crisis caused by the collapse of subprime mortgages reflected by the Standard and Poor's 500 (S&P 500) stock market index. Information about the sources of these study and the data points for the timeseries is provided in the Supplementary material.

## 1.2 Research context

When considering EWS of critical transitions a distinction has to be made between the observed precursor, that is a phenomoenological pattern that the system exhibits as it approaches the threshold, and the indicator itself, which is a specific measurable signal that characterises the registered phenomenon. This discernment is fundamental in the characterisation of EWS given by the fact that the same precursor can be measured by different indicators (see Table 1). As an example consider *critical slowing down* (CSD), arguably the most studied and well understood precursor in climate and ecological systems. It specifically refers to the observed behaviour of a dynamical system to become less resilient to perturbations of its stable equilibrium as a critical transition is approached. In other terms if the state of the system tracking an attractor comes close to a bifurcation then CSD is the symptomatic slower recovery towards the attractor w.r.t. small (noisy) perturbations. This phenomenon can be measured in a variety of ways and thus a number of different indicators (both analytical and statistical) have been proposed throughout the years. Other forms of precursors are *flickering* and *pattern-formation* and they also have been associated to the detection of multiple signals. Some questions naturally arise following this separation, namely is there a better indicator than others for the clear and early detection of a critical transition? Is such indicator robust i.e. can it provide consistently a signal for the given precursor observed from different systems? Does it exist a universal measure that characterises critical transitions a-priori i.e. without prior knowledge of neither the presence of a bifurcation nor the functional model of reference? Answers to these queries have been the subject of several efforts in the last 10 years and, to some extent, generic and robust indicators of abrupt regimes shifts have been achieved for a subset of simplified models. For the rest of the present section we will outline a condensed but representative genealogy of EWS of critical transitions.

### 1.2.1 Critical slowing down as loss of resilience

Detection of early signs of incoming catastrophic events can take several forms. As mentioned above, the first occurrence that was proposed as a precursor of such events is a phenomenon known as CSD. The origin of the adoption of this term in dynamical systems can be traced back to 1994 Strogatz's seminal book [62] (which further mentions the name to be coined in statistical mechanics) and it referred to the lethargic (algebraic) decay of solutions to the stable equilibrium in the proximity of a supercritical pitchfork bifurcation (as opposed to the fast, exponential decay away from the bifurcation).

#### 1980s and 1990s

This was proved mathematically one decade earlier in [3] where it was shown how the return time of the system to its attractor depends on the inverse of the leading eigenvalue of the linearisation of the dynamics around the equilibrium. As such attractor approaches a critical threshold it would thus exhibit infinitely slower return rate from a small perturbation. As far as we know, the characteristic return time is the first indicator foretelling a critical transition. The derivation of this “universal law” was motivated by the study of thresholds of instabilities in multi-stable ecological systems [1] which in retrospective appears to be somehow prophetic of the communities that will extensively apply CSD forewarning catastrophic population collapses 2 decades later. The influence of noise in such perturbations was characterised shortly thereafter in [4] which linked the frequency range of the disturbance with the simplest codim-1 bifurcations showing that small-signal amplification was indeed a feature of these autonomous dynamical systems. In 1995 this concept was (arguably independently) explored in [5] for stochastic dynamical systems. In there the *resilience* of ecological systems was quantified statistically to be the ratio of the variability of population densities to variability in population growth rates. The study mentions the return time to equilibrium as the deterministic equivalent of stochastic resilience.

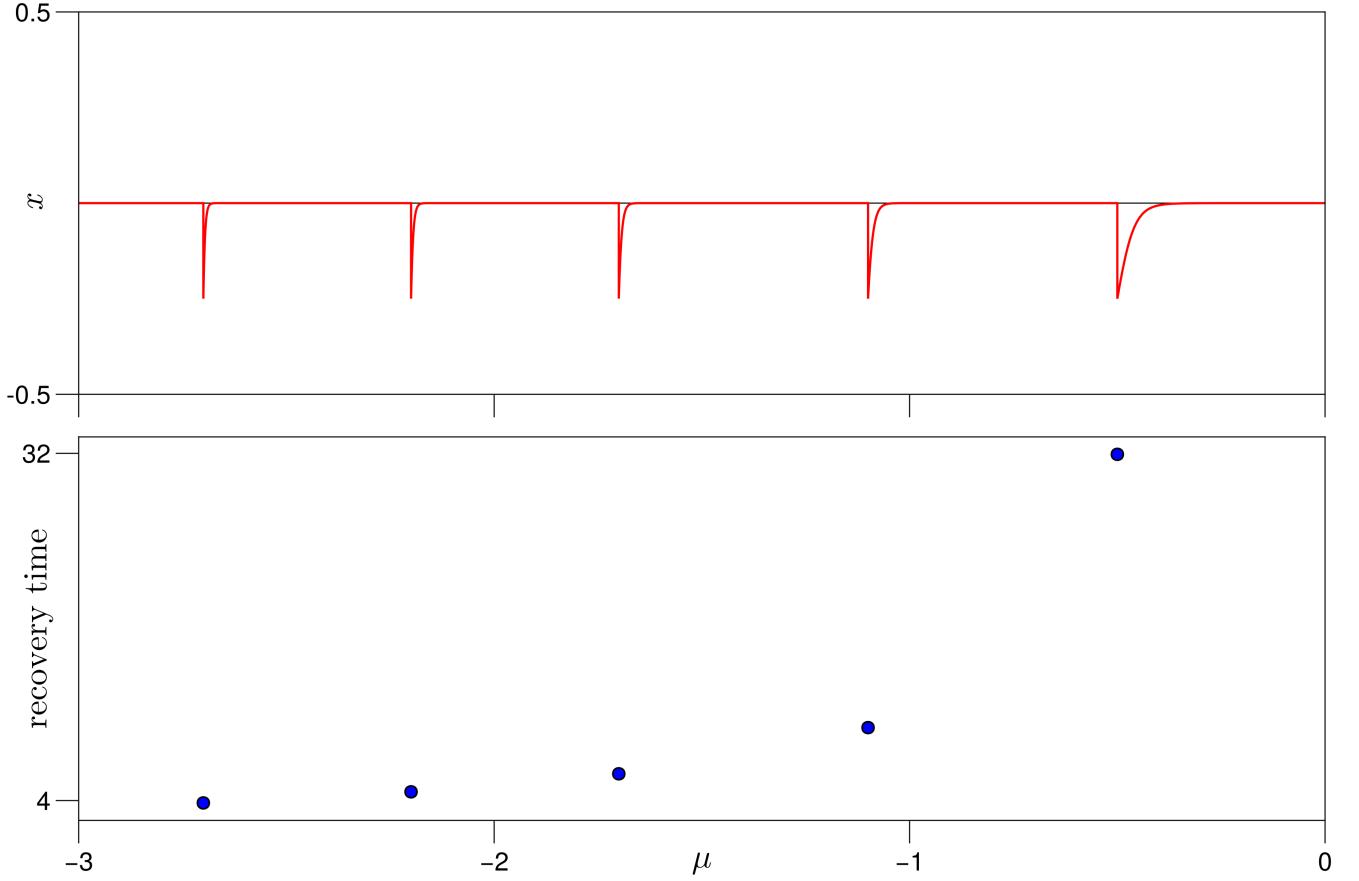


Figure 2: The phenomenon of CSD illustrated for a transcritical normal form. As the parameter approaches the bifurcation value  $\mu = 0$  small perturbations of the state (red line) tracking the attractor  $x = 0$  result in increasingly longer recovery. The return time is estimated numerically (blue dots) for each perturbation.

It then took almost one decade to the ecologists to realise that, as catastrophic switches between stable population's states occurred unannounced when one only monitors the state variables (as for a saddle-node bifurcation [11]) the derivation of EWS was necessary to predict these disruptive shifts and act upon the external forcing to prevent them.

## Mid 2000s

In 2005 it was observed [16] that the spatial dynamics of a logistic model organised on a square lattice exhibits fractal-like subdivisions of the competing species and suggested that the scaling law of the spatial variance near the extinction can be used as a EWS. It was one year later when Carpenter and Brock [20] formalised this idea and conjectured how increases in the variance of the timeseries of a stochastic differential equation (SDE) can be a clue of impending regime shifts. The rise in variance of SDEs as an indicator of CSD will not be rigorously proved mathematically for another 5 years when Kuehn [42] formalised it in the framework of fast-slow dynamics (discussed later in Section ??). Nevertheless there is a pleasant intuition for this measure of CSD as discussed in [31]. As CSD is associated to a slower recovery from forced perturbations, due to the leading eigenvalue of the respective Jacobian approaching the immaginary axis, the intrinsic rate of change of the realizations of the SDE diminishes. This in turn is reflected in the timeseries to exhibit memory of its past i.e. the state at the next iteration will be somewhat related to some order of its past states. This intuition is already enough to suggest the

autocorrelation of lagging order  $p$  ( $\text{AC}(p)$ ) to be an indicator of CSD. However if one is willing to force some form of equivalence between a more sluggish timeseries w.r.t. perturbations and a random walk (which has a monotonic trend in variance depending linearly on time) then rise in the variance can be reasonably assumed as an EWS of an upcoming bifurcation. Perhaps the major significance of Carpenter's and Brock's 2006 paper for the ecology community was the suggested implication that, by relying merely on timeseries data, a catastrophic population collapse can be predicted early enough even without knowing the underlying model (or assuming high levels of uncertainties in the estimation of parameters and noise levels). Prior to their proposal however, earlier works in the climate community proposed different types of indicators for estimating the distance to bifurcations of the North Atlantic thermoaline circulation (THC) under (stochastic) freshwater forcing driven by climate change. In particular in 2003 a simplified box model of the THC was deployed to estimate (analytically) the dependence of shifts in the frequency spectrum of the idealised Ornstein-Uhlenbeck process (OUP) of the deviations of the salinity gradient from equilibrium to the (bi-stable) saddle-node bifurcation in the model [12] (which became known as spectral reddening). One year later the same first author also co-authored an investigation for larger, more realistic models of the THC in a numerical setting [14] and concluded that an increased variability is a better diagnostic tool at anticipating the overturning circulation. These works spurred a vast uptake by climate scientist and ecologist in the role of CSD and in general in the potential of statistical measures that are prognostics of catastrophic regime shifts in their respective systems.

### 1.2.2 Criticism of EWS from climate and ecosystem timeseries

A substantial wealth of works was published in the following decade marking the beginning of intense research and categorisation of EWS from a dynamic perspective. We can think of the years spanning from 2007 to 2015 as the *gold rush* era of EWS in climate and ecology.

#### Late 2000s

Regarding the former, in 2007 a method based on detrended fluctuation analysis (DFA) was introduced by Livina and Lenton [22] to improve upon the estimation of the decay rate, and hence the proximity to a climate bifurcation, from paleoclimate records. They realised that the non-stationarity of (climate) timeseries approaching a bifurcation somehow pollute the statistical scaling properties used to estimate the non-linear shifts. The importance of detrending dynamical timeseries will be made clear in the third Chapter of the present work. In the same year van Nes and Scheffer [23] took note of this and explored the estimation of the recovery rate from perturbations by fitting an exponential process to the timeseries generated by 6 simple, 1-dimensional models of ecosystems. They also curcially mentioned the potential detection of *false positives* and *false negatives*. This crossover between the two communities eventually led to the 2008 pivotal work by Dakos, Scheffer, van Nes et al. [25] in which the first coefficient of a fitted autoregressive AR(1) model (which measures the autocorrelation at lag-1) was shown to consistently predict the onset of critical transitions in 8 major paleoclimate events (3 of which are reported in Figures 1(a)-1(c) above), i.e. using real-world data rather than model-based simulations. In the following 3 years several forms of indicators prognostic of CSD were proposed and applied to different population models. We mention, among others: kurtosis and skewness [27] used in a fisheries food web model (which found weaker signals compared to AR(1) and reddening of the spectral density ratio); the analytical recovery rate of Hopf and transcritical bifurcations of the 2-dimensional predatory-prey and Lotka-Volterra models [28]; variance, autocorrelation and skewness in the transcritical bifurcation of an experimental population in a deteriorating environment [36]; variance, return rate, skewness and spectral ratio in a monitored lake ecosystem [40]; the total variance of a fitted drift-diffusion-jump (DDJ) model [39] for non-parametric, data-generating processes. Soon after reviews of EWS of critical transitions became available, most notably

Scheffer, Carpenter, Dakos et al. [31] and Lenton [43]. In 2009 Scheffer also published a book [30] on critical transitions compiled until then in natural and human sciences.

## Early 2010s

With this new discipline gaining traction it also, naturally, attracted some criticism. In 1977 the implications of catastrophic events in biological and social sciences was put in question [2] by “*incorrect reasoning*” and “*far-fetched assumptions*” regarding cusp bifurcations. With regards to the robustness of the statistical indicators a more recent work in 2010 [35] argued that not all paleoclimate records show critical transitions due to bifurcation and are rather caused by stochastic fluctuations (known today as N-tippings, see Section ??) which have very limited predictability. The authors noted that by assessing the statistical significance of the variance and autocorrelation of records of the Dansgaard-Oeschger events (analysed by the ice core data in the North Greenland ice core project, or NGRIP) one showed a monotonic trend while the other showed no signal. This, the authors argued, was evidence that no bifurcation was approached by the system which instead tipped over to an alternative stable regime driven by noise alone, which is contrary to what was claimed in previous works [25, 31]. In addition, the same year, another work showed that abrupt population shifts in particular ecosystems (specifically when the magnitude of the stochastic fluctuations increases, i.e. higher noise levels) occurred without forewarning [37] of CSD (i.e. false negatives or *missed alarms*). The mathematical basis of the authors was that all the previous works showed EWS for (simplified) systems that have a smooth dependence of the potential landscape on the underlying parameter and did not consider real, complex systems exhibiting, for example, chaotic attractors. Ensemble simulations of the Ricker map did in fact lead to a failure of the aforementioned EWS in detecting the population collapse. Increasing evidence supporting the lack of consistency of the proposed indicators was further provided in 2013 by Dakos, Scheffer et al. [52] and others [50]. In particular the latter followed the examples laid out in the previous papers, i.e. providing evidence of *silent catastrophes* while the former went in the opposite direction and reported that variance and autocorrelation give a trend even for non-catastrophic events (i.e. false positives). These 4 works pointed out the urgency of:

1. a robust characterisation (in terms of statistical significance and some form of universality across different complex systems) of EWS;
2. a deeper (and desirably rigorous) understanding of the mathematical structure of the analysis motivating the existence of EWS.

The second of those two needs began to be investigated by Kuehn in 2011 [42] and we discuss it in the next Chapter (see Sections ?? and ??). The questions regarding the robustness and universality of the leading indicators of CSD also started to be addressed shortly thereafter and we will explore them further ahead (see Sections ?? and 3.1).

### 1.2.3 Flickering as an alternative sign of critical thresholds

Contemporary to the rise of indicators for CSD and their critics, another phenomenon was being observed to be a precursor of regime shifts. We mentioned above some indicators s.a. increase in variance and skewness to be prognostic of CSD albeit showing weaker signals w.r.t. autocorrelation [27, 36, 40]. Changing skewness in the distribution of the realizations of the timeseries was first proposed in 2008 [26] as an indicator of increasing asymmetry in a number of simulated ecosystems. In such paper the origin of a trend in skewness as the models approached a critical transition was linked to the asymmetric change of the potential landscape of the SDE. This prescribes that the state of the system has particular directions,

towards which it may tip to, that are more favourable than others. This observation is considerably different from the flattening of the potential landscape that is found at the core of CSD and as such suggested that a novel precursor was being measured by the proposed EWS. This new phenomenon (it was shown) did not rely on the linearisation of the dynamics around the stable equilibrium and instead considered the effects of high noise levels to be dominant close to the critical threshold where alternative attractors became “available” to the system under large fluctuations. It would be thus reasonable to expect that systems that do not feature such asymmetry in the presence of multi-stable states will not exhibit stronger signals than those that are instead purely associated to the flattening of the potential landscape as prescribed by CSD.

## Mid 2010s

In 2010 this intuition was further validated by approximating the shape of the potential function (and thus estimating the number of states available to the system at given parameter values) by polynomial fitting of the empirical distribution of timeseries data [38]. The method, termed potential analysis, successfully detected an increased number of alternative stable states for ice-core proxy records of paleotemperature transitions at NGRIP, thus providing yet another indicator of asymmetric changes in the potential landscape of a model. Similarly to the increase in skewness, also the increase in variance, previously linked to CSD, was discovered to occur in unsteady excursions of the state towards alternate basins of attraction. Brock and Carpenter [32] thus distinguished the two sources of increased variance in the already known CSD and the new phenomenon which they named flickering.

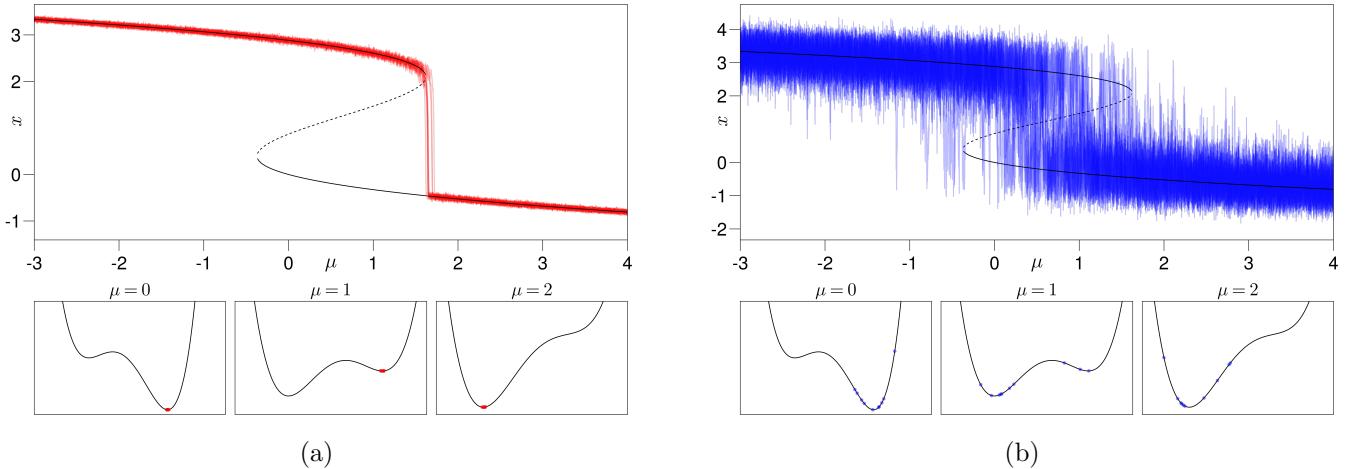


Figure 3: Ensemble sample paths of a bistable saddle-node bifurcation in low-noise (a) and highly stochastic (b) regimes. The potential function is displayed in the bottom panels at parameter values  $\mu = 0, 1$  in the bistability region and  $\mu = 2$  past the bifurcation; distribution of the ensemble states are plotted as dots on the potential function. We can clearly see how the high noise levels in (b) causes flickering anticipating substantially the critical threshold and this results in the ensemble states being considerably less clustered in the potential minima w.r.t. the low-noise case in (a). Details on the system and the potential function are provided in the Supplementary material.

The newly introduced precursor, by its very nature, complicated the analysis of the dynamics far from the bifurcation since high level of noise may cause the system to switch back and forth between alternative stable states and, as previously established, the prediction of these noise-induced tipping is considerably harder than those that are associated to bifurcations. In 2012 the influence of *strong noise* in a multi-decadal lake eutrophication experiment showed that skewness and autocorrelation display opposite trends

w.r.t. the increase in variance [49] which led to the conclusion that the latter could have only be caused by flickering rather than CSD. The meaningful consequence of this is that, in “*highly stochastic dynamics*”, the switch to bistable regimes sets in quite earlier than CSD (see Figure 3), i.e. when the alternative attractor becomes strong enough such that the noise-induced fluctuations will cause excursions across the entire basin of attraction. Further evidence of this was given one year later by Dakos, van Nes and Scheffer [51] which systematically registered increase in variance and decrease of autocorrelation (symptomatic of flickering) in highly noisy but low-resolution timeseries well ahead of the CSD associated to a saddle-node bifurcation. While all these results firmly placed skewness as an indicator of flickering, the increase in variance was now linked to both CSD and flickering depending on the noise-level embedded in the realizations of the stochastic process and the size shrinking of the basin of attraction in a multistable regime. The underlying implication is that it would not always be possible to rely on timeseries data alone to distinguish between the two different drivers of critical transitions given the low dimensionality of the model, by which we mean that the mere observation of the temporal realizations of the stochastic processes will forcibly discard potentially useful information that could otherwise help to determine robustly the distance from upcoming thresholds. As such, other works were starting to address the possibility of including spatial dependence in their models in the hope of enriching the forewarning capability of EWS.

#### 1.2.4 Investigation of spatially extended systems

The first evidence of accounting spatial information in the context of critical transitions is from the 1998 series of works by Gandhi and colleagues [6] which realised that, whenever the dynamics of competing species are dominated by spatial clusters (as opposed to the usual assumption of well-mixing of the population densities) proxies of CSD marking the distance to extinction can be determined s.a. the correlation length (average cluster size). These works also noted that spatial aggregation through mean-field approximation (MFA) averages out the fluctuations that become dominant as the phase transition is approached thereby providing some insights on the loss of information when space is not accounted in dynamical models.

#### 2000s

Some years later in 2004, the MFA of well-mixed ecosystems with bistable regimes was again called into question when considering natural catastrophic shifts that feature the emergence of self-organising patches of consumers and resources [15], s.a. those found in arid and savanna vegetation patterns. This called for the extension to spatial domains to recognise self-organised patchiness as a precursor of regime shifts. One year later, self-organised patchiness (equivalently known as pattern-formation or Turing instability) was again suggested as a forerunner of catastrophic shifts for asthma attacks [18]. It compared experimental evidence via positron emission tomography (PET) of cluster ventilation defects (CVDs) in the lung with the effect of spatial heterogeneity breaking the uniform smooth muscle activation in a computational model of a symmetric bronchial tree. It was observed specifically how the local bistability of the airway tree induces the patchy self-organisation of CVDs. Further evidence of spatial patterns preluding ecosystem collapses was presented in 2007 when it was observed that the size distribution of the vegetation patches in Mediterranean arid environments (found in Southern Spain, Greece and North-Western Morocco) follows a power law with the number of such patches prior to desertification [21]. As it became increasingly clear that pattern-formation was yet another form of precursor capturing the insofar neglected spatial information, a consistent statistical measure of such a phenomenon was still lacking. In 2009 Guttal et al. [29] put a remedy to this deficiency by introducing spatial variance and spatial skewness as a first form of leading indicators. In particular they show how an increase in the former and changes in the latter provide clear indications of pattern-forming transitions to bistability regimes in a spatially extended population model. Furthermore it was demonstrated how these two indicators can overcome the limitations imposed by low

temporal resolution of the timeseries of the MFA of the model.

## Mid 2010s

The 2010s thus marked the beginning of the analysis of spatio-temporal EWS, particularly by the ecologists. Carpenter and Brock [33] used the discrete Fourier transform (DFT) to extend the previously conceived concept of spectral reddening of a timeseries experiencing CSD [12] to shifts to lower spatial frequencies for period-doubling (Ricker map) and bistable (predator-prey and harvester-prey) transitions in 4 ecological models. Concurrently Dakos, van Nes, Scheffer et al. [34] proposed spatial correlation as the leading indicator of CSD prior to the bifurcations in overharvesting, eutrophication and vegetation turbidity reaction-diffusion models. Specifically they compared the signal provided by Moran's two-points correlation coefficient (used to calculate the spatial correlation of neighboring cells in a 2-dimensional lattice) with that of the temporal autocorrelation at lag-1 of the spatially aggregated model and found that the former consistently outperforms the latter. In 2011 the aforementioned authors and others [41] applied the same ideas to spatially-patterned vegetation changes at the onset of desertification and concluded that CSD was not detected by rising temporal variance and autocorrelation as opposed to their spatial counterparts which were able to show positive trends leading to both CSD and Turing instabilities. Experimental evidence from a whole-ecosystem food-web monitoring of predators' distribution in a lake [55] also indicated that spatial variance and shifts to low spatial frequencies using the DFT provided EWS for the emergence of prey fish patchiness one year in advance of the bistable transition in the populations.

## Late 2010s and 2020s

Nevertheless, pattern-formation is only one of the many dynamical phenomena that affect the stability of spatio-temporal systems (as we will explore in the third Chapter of the present work). In fact Kuehn et al. in 2013 [53] and Leemput, van Nes and Scheffer in 2015 [59] independently considered travelling waves of invasion fronts connecting two well-separated initial populations distributions. They both indicated the slow-down (or *pinning*) of the wavespeed as a precursor to the critical point at which the resilience of two alternative stable states of high and low biomass is equal (called Maxwell point). Investigation of spatio-temporal EWS also extended to network-based analysis in the later part of the decade with the maximum element of the covariance matrix [58], the noise-dependent characteristic return time of a dimensionality reduction of a coupled network [66], the spatial correlation of a multiplex disease network [67] and the slower recovery from perturbation of mutualistic communities [73] all providing candidate spatial indicators. The latest efforts of the current decade also proposed the adoption of techniques from statistical mechanics [68, 72, 79] for spatio-temporal systems as well as deep-learning [76, 80] for the automatic detection of combined signals albeit with questionable insights gained from the interpretation of their meaning.

## 2 Background

### 3 Critical transitions

In the first Chapter we explored how several different EWS have been derived for three different forms of precursors of critical transitions in natural phenomena. In most of the cited literature the data from which these signals arose came from idealised mathematical models. This is particularly the case for ecological and climate systems where meaningful real-life data gathering is complex and requires a substantial undertaking in terms of resources. As a result of this idealisation the quality and robustness of a proposed EWS heavily relies on the accuracy with which such simplified models capture the essence of a phenomenon of interest. To that end in this Chapter we will devote our attention on how such models are built in an increasingly complex setting. We will start with the most simple case derived with restrictive assumptions regarding the time-evolution of the systems under investigation and what impact those assumptions have on the prognosis of tipping events. We will thus evaluate how, as we subsequently drop some of these assumptions, the models become more realistic at the cost of increased complexity in the analysis. This construction will aid the intuition on how to mathematically investigate catastrophic events in dynamical systems with the aim of understanding how global properties of the model can be exploited in order to predict them.

#### 3.1 Low-dimensional dynamical systems

The naive interpretation of a system exhibiting a critical transition is that the temporal evolution  $\dot{x} = f(x)$  of a set of observables  $x \in \mathbb{R}^n$  undergoes a sudden, qualitative change at particular values of an intrinsic (potentially hidden) parameter  $\mu \in \mathbb{R}^m$ . Here the dimensionality of the state space  $\mathbb{R}^n$  is important to characterise as  $n \sim \mathcal{O}(1) \ll \infty$  as it fundamentally enables us to apply classical dynamical systems theory for the discovery of equilibria, (hyperbolic) invariant sets and their bifurcation. In those low-dimensional systems there are multiple causes that may drive the dynamics to tip towards different regimes (two of which we have already informally introduced in the previous Chapter):

- ▶ bifurcation-induced (B-tipping): also known as dynamic bifurcations, they are the result of dynamical systems passing through a bifurcation of their parameters causing one or multiple attractors to lose their stability;
- ▶ rate-induced (R-tipping): when the evolution of a system fails to track the time-changing attractor caused by the rate of change of the parameters rather than their values (as opposed to what happens at bifurcation points);
- ▶ noise-induced (N-tipping): affecting those systems perturbed by sufficiently high noise which causes the state to depart far enough from a neighbour of a stable equilibrium so that it eventually escapes the basin of attraction.

It must be emphasized that given any dynamical system, critical transitions may occur because of one or multiple concurrent causes described above and in general it is not possible to discern the original source of such tipping.

##### 3.1.1 Deterministic transitions: bifurcation-induced (B-)tipping

In classical dynamical systems, the location (in state space) and the stability of equilibria and their manifolds is studied at fixed values of the parameter  $\mu$ . In the context of bifurcation theory in fact the loss of hyperbolicity of these objects may cause the system to undergo a local or global bifurcation which inherently changes the stability or even the existence of the attractors. This causes the model's state to shift

the dynamical regime from slow stable time-evolution to a fast unstable one. The theory upon which these systems are investigated has been known since the 1930s<sup>1</sup> and ever since expanded upon to include complex dynamics such as chaos theory and delayed differential equations (DDEs). However in real-world phenomena it is reasonable to assume that while the system's dynamics dictates the evolution of the observables in phase space at specific parameter values, the system properties also indirectly affect the parameter itself. Therefore the first assumption to be dropped in our quest of deriving these high-fidelity models is the stationarity of the parameter  $\mu$  and the formulation of an augmented (albeit still low-dimensional) version of the original system. In the simplest form the time variation of the parameter is monotonic and linear with the ramping characterised by a timescale  $\tau := \varepsilon t$ ,  $0 < \varepsilon \ll 1$  much slower than that of the observables' dynamics. The resulting system is therefore called fast-slow

$$\begin{cases} \dot{x} := \frac{d}{dt}x = f(x; \mu), \\ \dot{\mu} := \frac{d}{dt}\mu = \varepsilon, \end{cases} \longleftrightarrow \begin{cases} x' := \frac{d}{d\tau}x = \varepsilon^{-1}f(x; \mu), \\ \mu' := \frac{d}{d\tau}\mu = 1, \end{cases} \quad (1)$$

and it has been the cornerstone object for the development of geometric singular perturbation theory (GSPT) [74], extensively applied in the analysis of biological mechanisms s.a. calcium cell signaling [70]. In [42] such formulation is the starting point in the discussion of a general and robust framework for the characterisation of critical transitions in low-dimensional systems. Notably the fast-slow dynamics is approximated by resolving the layer equation

$$\begin{cases} \dot{x} = f(x; \mu), \\ \dot{\mu} = 0, \end{cases} \quad (2)$$

which is obtained by taking the singular limit  $\varepsilon \rightarrow 0$  of the system in the fast timescale (i.e. the leftmost system in (1)), and the reduced equation

$$\begin{cases} 0 = f(x; \mu), \\ \mu' = 1, \end{cases} \quad (3)$$

which is derived via the singular limit of the system in slow timescale (which is the rightmost system in (1)). The centerpiece of the whole analysis is the derivation of the critical manifold  $C := \{(x, \mu) \in \mathbb{R}^{n+m} : f(x; \mu) = 0\}$  as it constraints the solutions of the slow subsystem (3) to a unique codim- $m$  object as shown in the example depicted in Figure 4 below. An important result from GSPT allows us to characterise a family of (infinitely many) slow manifolds  $C_\varepsilon := \{(x, \mu) \in \mathbb{R}^{N+M} : x = h_\varepsilon(\mu) = h(\mu) + \mathcal{O}(\varepsilon)\}$  where  $h : \mu \rightarrow x$  is given by the Implicit Function Theorem.

### Theorem 3.1: Fenichel's (1<sup>st</sup>) Theorem

Let  $C$  be a compact, normally hyperbolic manifold then for  $\varepsilon > 0$  sufficiently small there exists  $C_\varepsilon$  locally invariant and diffeomorphic to  $C$  that lies at a Hausdorff distance  $O(\varepsilon)$  from  $C$ .

These singularly perturbed slow manifolds will become paramount later on in the derivation of statistical measure as precursors of critical transitions which we hereby define.

**Definition 3.1** (Critical transition). Let  $\gamma(t)$  be a homeomorphic image of the real subset  $(a, b)$  s.t. for each partitioning  $a = t_0 < t_1, \dots, t_{N-1} < t_N = b$  the image  $\gamma(t_{j-1}, t_j)$  is an orientation preserving trajectory of either the layer (2) or the reduced equation (3) and let  $p \in C$  be a point where the critical manifold loses its hyperbolicity, then  $p$  is a critical transition if there exist a candidate trajectory  $\gamma$  s.t.  $p = \gamma(t_j)$  is a transition between fast and slow regimes and  $\gamma(t_{j-1}, t_j)$  is hyperbolic and attracting.

<sup>1</sup>[http://www.scholarpedia.org/article/History\\_of\\_dynamical\\_systems](http://www.scholarpedia.org/article/History_of_dynamical_systems)

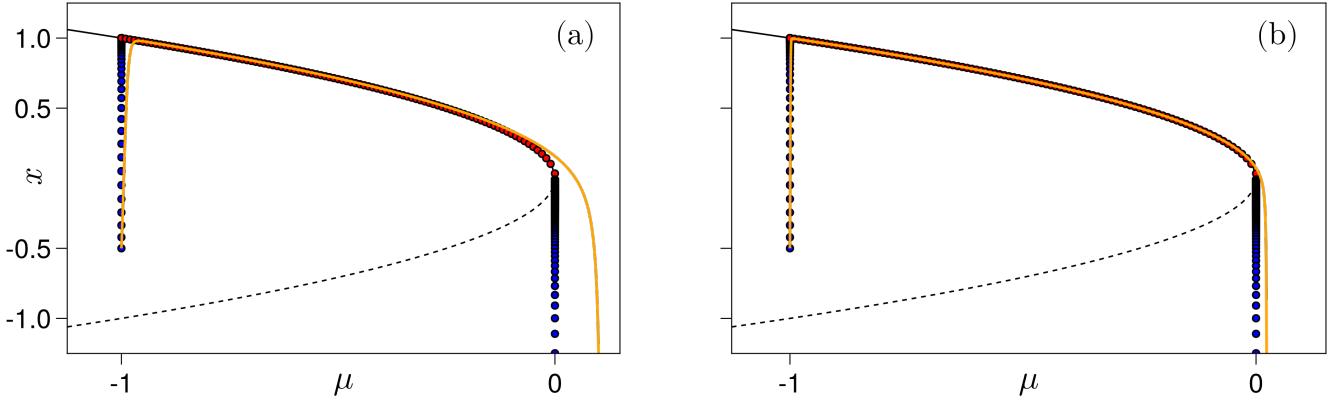


Figure 4: Comparison of the (numerical) fast-slow trajectory (yellow curve) of the codim-1 saddle-node normal form with different linear rampings ((a):  $\varepsilon = 10^{-2}$ ; (b):  $\varepsilon = 10^{-3}$ ) and its approximation given by different regimes in the singular limit. Solutions of the layer problem (2) in blue dots, with initial conditions  $(-0.5, -1)$  and  $(0, 0)$ , connected by a solution of the reduced problem (3) in red dots with initial condition  $(1, -1)$ . Notice how the dynamics on the critical manifold prescribes the solution of the slow subsystem and the delay of the critical transition in the fast regime is determined by the value of the timescale separation constant  $\varepsilon$ .

### 3.1.2 Different timescales: rate-induced (R-)tipping

The modelling of critical transitions in the fast-slow formulation also introduces an additional mechanism driving the tipping in which the ramping of the parameter is itself the cause of sudden regime shifts in the dynamics. In [45] the rate of change of  $\mu$  is shown to cause, for some specific critical values, sudden regime shifts in the dynamics of a fast-slow system that does not feature a deterministic bifurcation, otherwise known as R-tipping.

**Definition 3.2.** Let  $x \in \mathbb{R}^n$  be a state variable,  $\mu \in \mathbb{R}$  a time-changing parameter,  $R > 0$  a constant and  $M$  a fixed, stable and invertible linear operator then the linearised system

$$x'(t) = M(x - x^*(\mu)),$$

is said to be (adiabatically) tracking the stable equilibria  $x_{\text{eq}}(\mu)$  if  $\forall t |x(t) - x_{\text{eq}}(\mu(t))| < R$  and viceversa it tips beyond the equilibria if  $\exists t_0$  s.t.  $|x(t_0) - x_{\text{eq}}(\mu(t_0))| = R$ .

The so-called tipping radius  $R$  is identified by constructing the condition upon which a fast-slow system fails to track the (slow) drift of the attractor To study the dependence of the tipping event from the rate of change of the time-varying parameter we formally introduce the drift of the stable equilibria

$$r(t) := \frac{d}{dt}(x_{\text{eq}}(\mu(t))) = \frac{dx_{\text{eq}}}{d\mu} \frac{d\mu}{dt} = \varepsilon \frac{dx_{\text{eq}}}{dt}.$$

With these quantities defined we are ready to define a criterion for R-tipping for systems with steady drift.

#### Theorem 3.2

Let (1) be a system with tipping radius  $R > 0$  and  $r(t) = r > 0$  is constant (i.e. steady drift of the parameter), then

$$\|M\|^{-1}|r| > R$$

is a sufficient condition for R-tipping to occur.

*Proof.* Given an initial condition  $x(0) = x_0$  the general solution of (1) is

$$x(t) = e^{Mt}x_0 + \int_{s=0}^t e^{M(t-s)}M x_{\text{eq}}(\mu(s)) ds = - \int_{u=0}^{\infty} e^{Mu} M x_{\text{eq}}(\mu(t-u)) du,$$

where the dependence on the initial condition is dropped from the second integral by assuming an arbitrarily long past while setting  $u := t - s$ . Integrating by parts the general solution gives us

$$x(t) = -[e^{Mu} x_{\text{eq}}(\mu(t-u))]_0^{\infty} - \int_0^{\infty} e^{Mu} \frac{d}{dt}(x_{\text{eq}}(\mu(t-u))) du \Rightarrow x(t) - x_{\text{eq}}(\mu(t)) = - \int_0^{\infty} e^{Mu} r(\mu(t-u)) du,$$

where the (exponential) stability of the linear operator ( $|e^{Mt} \rightarrow 0|$  as  $t \rightarrow \infty$ ) has been used. We integrate by parts again and obtain

$$x(t) - x_{\text{eq}}(\mu(t)) = M^{-1}r(\mu(t-u)) - \int_0^{\infty} e^{Mu} M^{-1} r'(\mu(t-u)) du.$$

Assuming now steady drift the above simplifies to

$$x(t) - x_{\text{eq}}(\mu(t)) = M^{-1}r \Rightarrow |x(t) - x_{\text{eq}}(\mu(t))| = |M^{-1}r|.$$

By choosing the matrix norm for the linear operator  $\|M\| = \sup_{v \neq 0} \frac{|Mv|}{|v|}$  then it trivially follows

$$\|M\|^{-1}|r| \leq |M^{-1}r| \leq \|M^{-1}\||r|.$$

We now recall from Definition 3.2 the condition for R-tipping which gives us the sufficient condition  $\|M\|^{-1}|r| > R$  for the solution to slip off the stable equilibria thereby causing the rate-induced critical transition.  $\square$

This result allows us to generalise the analysis to systems that have no known deterministic bifurcations therefore dropping the additional constraint of restricting ourselves to problems that can be recast in a normal form.

### Example 3.1

As an example we consider the following 2-dimensional model of compost-bomb instability [44] in the slow timescale

$$\begin{cases} \varepsilon x' = y + \mu + x(x - 1), \\ y' = -x - x^2 - x^3 - x^4 - x^5, \end{cases}$$

with steady (linear) ramping of the parameter  $\mu \in \mathbb{R}$ . The system has one stable equilibrium at  $(0, -\mu)$  and in the singular limit we are able to derive the critical manifold

$$C = \{(x, y) \in \mathbb{R}^2 : y = -\mu - x(x - 1)\},$$

with attracting  $C^{(a)} = \{C \cap \{x < \frac{1}{2}\}\}$  and repelling  $C^{(r)} = \{C \cap \{x > \frac{1}{2}\}\}$  submanifolds separated by a saddle line which is tangent to  $C$  along  $x = \frac{1}{2}$ . Note that the distance along  $x$  between the stable equilibrium and the fold of the critical manifold is exactly  $\frac{1}{2} = R$  which will thus provide us the tipping radius for our system. We now differentiate the critical manifold w.r.t. the slow timescale to obtain

$$\begin{aligned} 0 &= \dot{x} + \dot{\mu} + \frac{d}{d\tau}(x^2 - x) = f_2(x, y) + r + \dot{x}(2x - 1) \Rightarrow \\ \Rightarrow \dot{x} &= -\frac{f_2(x, y) + r}{2x - 1}, \end{aligned} \tag{4}$$

where we have  $f_2(x, y) := y = -x - x^2 - x^3 - x^4 - x^5 = -\sum_{n=1}^5 x^n$  for the slow dynamics as well as the definition for the drift  $\dot{\mu} =: r$ . We immediately realise that (4) is singular along the fold of the critical manifold therefore, by recalling that  $\tau = \epsilon t \Rightarrow \frac{dt}{d\tau} = \epsilon^{-1}$ , we shift the temporal framework onto the fast timescale  $t$  obtaining

$$\begin{cases} x' = r - \sum_{n=1}^5 x^n, \\ \mu' = -r(2x - 1), \end{cases} \tag{5}$$

The equilibria for (5) will finally give us the (parametrised) invariant set of  $x$  points to which trajectories starting within the attracting submanifolds will converge to

$$M(r) = \{(x, y) \in \mathbb{R}^2 : \sum_{n=1}^5 x^n = r\}$$

To find the critical value  $r_c$  for the drift of the parameter that causes the state to tip and slip off the critical manifold we only need to substitute the value for the tipping radius previously derived in  $M(r)$  which thus gives us  $r_c \approx 0.9687$ . We conclude that drift values  $r > r_c$  will move the invariant set over and across the fold of the critical manifold meaning that even those trajectories that start within  $C^{(a)}$  will slip off the stable equilibria and tip in a rate-induced critical transition.

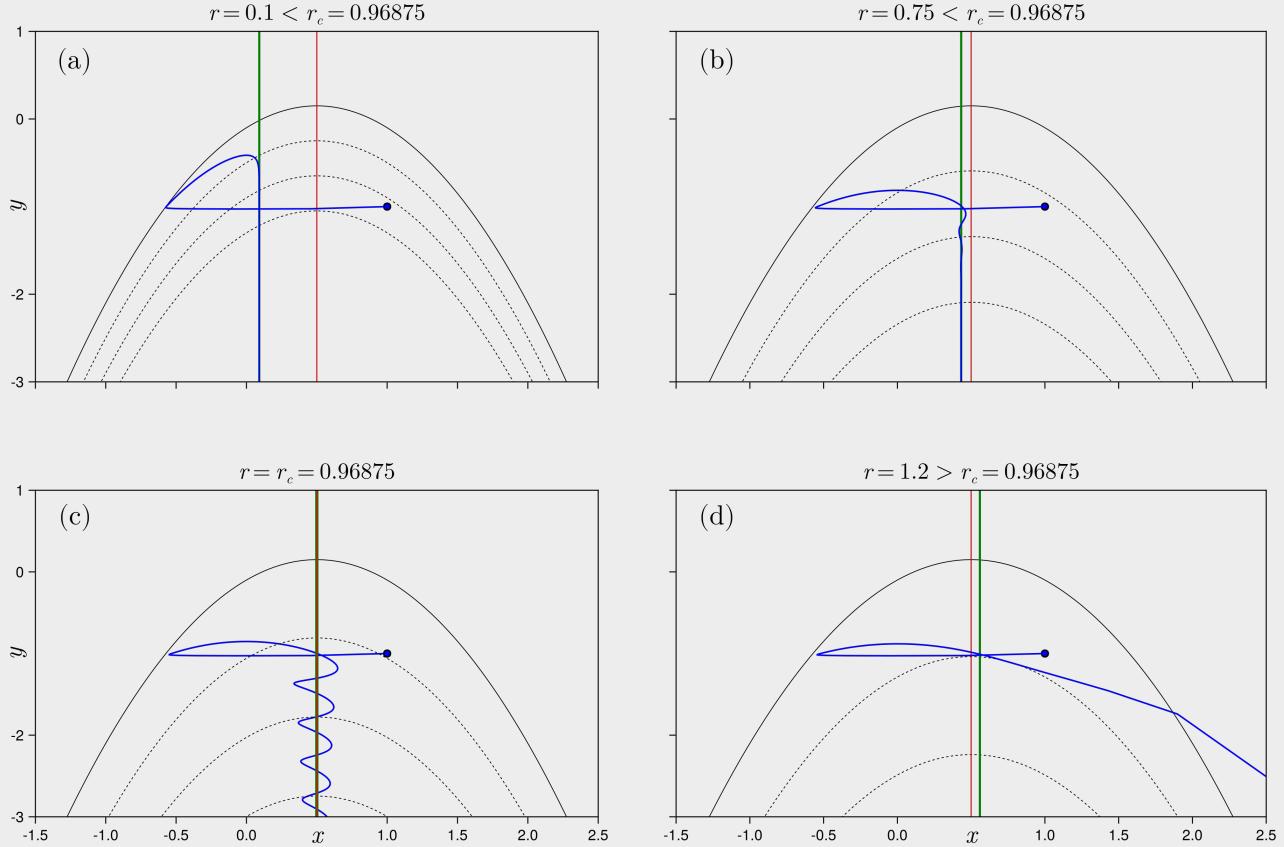


Figure 5: Forward-time trajectories (blue curves) of the compost-bomb fast-slow system propagated from the IC  $(1, -1)$  (blue dots) located in the inner part of the attracting submanifold  $C^{(a)}$  with timescale separation value  $\varepsilon = 2 \cdot 10^{-2}$  and parameter  $\mu$  varied linearly in  $[0, 2]$  at different rates  $r(t)$ . Notice how as we keep increasing the drift of the parameter while keeping it below the critical value (a)~(b) the invariant set (solid red line) moves closer to the fold of the critical manifold (here depicted as the solid black curve for the initial parameter value of  $\mu = 0.1$  and as dashed black curves at increasing parameter values) but never crosses it, allowing the trajectories to eventually settle on the invariant set (solid green lines) meaning that they will remain close (within tipping radius distance) to the stable equilibrium. For a drift corresponding to the critical value  $r_c$  (c) we notice large oscillations while the trajectories try to settle along the invariant set which has now moved over and overlapping with the saddle. Finally as we keep increasing the drift over the critical value (d) we clearly see that now the invariant line has crossed over the fold and thus the trajectories are now unable to track the stable equilibrium thereby slipping off the critical manifold.

As mentioned in the description of Figure 4 the non-stationarity of the parameter implies a delay in the tipping event. As such a natural question arises in how this delay reflects in the detection of EWS of R-tipping events. In 2016 Ritchie and colleagues [64] observed that an increase in variance and autocorrelation also feature the same delay w.r.t. the critical parameter drift value and analytically characterised it for a saddle-node normal form. 6 years later in 2023 Ritchie et al. [82] revisited this concept for more complex settings of predator-prey ecological systems and the possible collapse of the Atlantic meridional overturning circulation (AMOC), both with non-linear ramps of the forcing parameter. The same climate model was used 2 years prior [77] to conclude how R-tipping could in principle lead to a cascading effect of tipping

points predicted by an increase in variance and autocorrelation.

### 3.1.3 Stochastic perturbations: noise-induced (N-)tipping

One substantially strong assumption that we need to discard in order to build a more realistic dynamical system for the prediction of critical transition is the deterministic character of its dynamics. Whether it is due to the intrinsic stochastic nature of the phenomenon under investigation or to the noise naturally embedded in the sensory data collected in scientific monitoring, the stochastic character of real-world timeseries data is inevitable. We thus need to move away from the classical theory of flows in ordinary differential equations (ODEs) and into the realm of stochastic processes. In [42] the fast-slow problem (1) in the slow timescale  $\tau = \varepsilon t$  is reformulated as an Itô differential form

$$\begin{cases} dx = \epsilon^{-1} f(x; \mu) d\tau + \frac{\sigma_f}{\sqrt{\epsilon}} dW, \\ d\mu = g(x; \mu) d\tau, \end{cases} \quad (6)$$

whith  $f$  and  $g$  modelling the deterministic drift of an ensemble's average of trajectories,  $\sigma_f$  modelling the stochastic diffusion on such ensemble and  $W$  representing Brownian motion. We can now exploit the slow manifold  $C_\varepsilon$  introduced in Theorem 3.1 to derive a criterion to discern an ensemble's trajectory close to the critical transition from one that is in a stable regime. According to Definition 3.1 in fact, a sample path of an ensemble of trajectories of (6) that stays close (in probabilistic terms) to  $C_\varepsilon$  will be in a regime of slow, stable evolution. It is thus intuitive to define a new stochastic process  $\xi(\tau) = x(\tau) - h_\varepsilon(\mu)$  which quantifies the statistical departure of the sample path  $x(\tau)$  from the critical manifold. We then apply Ito's formula to this new process to derive the associated SDE; after that we define once again a new stochastic process  $X(\tau) := \sigma_f^{-2} \text{Var}(\xi)$  which satisfies the following deterministic fast-slow system

$$\begin{cases} \dot{X} = 2\epsilon^{-1} A_\varepsilon(\mu) X + 1, \\ \dot{\mu} = g(h_\varepsilon(\mu), \mu), \end{cases} \quad (7)$$

with  $A_\varepsilon(\mu) = (D_x f)(h_\varepsilon(\mu), \mu) - \epsilon(D_\mu h_\varepsilon)(\mu)(D_x g)(h_\varepsilon(\mu), \mu)$ . Applying Theorem 3.1 again for (7) we can construct a new slow manifold for the sample paths of the new stochastic fast-slow system

$$C_\varepsilon^{(X)} := \left\{ (X, \mu) \in \mathbb{R}^2 : X = H_\varepsilon(\mu) := -\frac{1}{2A_\varepsilon(\mu) + O(\epsilon)} \right\}, \quad (8)$$

and from that we finally can define a neighbour of the slow manifold which scales with the variance of sample paths

$$N_\rho(C_\varepsilon) := \left\{ (x, \mu) \in \mathbb{R}^2 : \frac{(x - h_\varepsilon(\mu))^2}{H_\varepsilon(\mu)} < \rho^2 \right\}. \quad (9)$$

We formalise the above derivation in the Theorem below.

#### Theorem 3.3

Let (6) be a scalar stochastic fast-slow system (i.e.  $n = m = 1$ ) with additive noise (i.e.  $\sigma_f(x) = \sigma \in \mathbb{R}$ ) and

$$N_\rho(C_\varepsilon) := \left\{ (x, \mu) \in \mathbb{R}^2 : -\frac{(x - h_\varepsilon(\mu))^2}{2A_\varepsilon(\mu) + O(\epsilon)} < \rho^2 \right\}, \quad A_\varepsilon(\mu) = (D_x f)(h_\varepsilon, \mu) - \epsilon(D_\mu h_\varepsilon)(D_x g)(h_\varepsilon, \mu),$$

be a neighbour of the singularly perturbed slow manifold, then if a sample path has an initial condition on the slow manifold  $C_\varepsilon$  it will stay bounded within  $N_\rho(C_\varepsilon)$  with high probability.

Arguably of greater importance for the context of EWS detection is the corollary that follows.

**Corollary 3.1.** *Since the measure of  $N_\rho(C_\varepsilon)$  scales with the variance of  $\xi(\tau)$  (which coincides with the variance of  $x(\tau)$  from trivial properties of the variance of a sum of stochastic processes) then an increase in the variance implies the neighbour to grow bigger.*

Whenever a stochastic process tracking a (deterministic) slow stable manifold is *kicked-off* due to the perturbations of the noise and escapes the basin of attraction we are presented with a case of the previously discussed N-tipping. Despite the notorious difficulty of anticipating N-tipping events, the corollary above provides us with a robust mathematical characterisation of either CSD or flickering through the increase in variance of a stochastic process approaching B-tipping. The intuition behind CSD for B-tipping is that as the bifurcation is approached the neighbor  $N_\rho(C_\varepsilon)$ , in which the realisations of the stochastic process are bounded, increases in size allowing for more probability to a N-tipping event to occur.

### Example 3.2

In [42] an example is provided to validate this result using an OUP

$$\begin{cases} dx = -\varepsilon^{-1}\alpha x d\tau + \frac{\sigma}{\sqrt{\varepsilon}} dW, \\ d\mu = d\tau, \end{cases} \quad (10)$$

as benchmark. Let us consider a sample path realised by the OUP in (10) with drift constant  $\alpha = 1$ , timescale constant  $\epsilon = 2 \cdot 10^{-2}$  and noise level  $\sigma = 0.1$ . We want to construct the neighborhood  $N_\rho(C_\varepsilon)$  and confirm that the numerical sample paths stay in fact bounded within such neighborhood with high probability. To do so we must first construct the (deterministic) slow manifold  $C_\varepsilon$  and for that we need in turn to derive the map  $h_\varepsilon(\mu) = h(\mu) + \mathcal{O}(\varepsilon)$ . With the critical manifold coinciding with the  $x = 0$  axis

$$C = \{(x, \mu) \in \mathbb{R}^2 : f(x, \mu) = -\alpha x = 0 \Rightarrow x = 0 \forall \alpha \neq 0, \forall \mu \in \mathbb{R}\},$$

we can easily check that the existance of  $h : \mu \mapsto x$  is guranteed by Theorem ?? given that  $C$  is normally hyperbolic at any point. It follows similarly that the only map  $x = h(\mu) = 0, \forall \mu \in \mathbb{R}$  is the null function yielding

$$h_\varepsilon(\mu) = h(\mu) + \mathcal{O}(\varepsilon) = \mathcal{O}(\varepsilon) \implies C_\varepsilon = \{(x, \mu) \in \mathbb{R}^2 : x = O(\varepsilon)\}.$$

This tells us that the slow manifold is a narrow strip wrapped around the critical manifold whose width is twice the Hausdorff distance from  $C$ . The next step in our construction is to derive a new slow manifold  $C_\varepsilon^{(X)}$  for the process

$$X = \sigma^{-2}\text{Var}(\xi) = \sigma^{-2}\text{Var}(x - h_\varepsilon(\mu)) = \sigma^{-2}\text{Var}(x),$$

where we recall the invariance of the variance under additive constant perturbations. First we compute the quantity  $A_\varepsilon(\mu) = (D_x f)(h_\varepsilon(\mu), \mu) - \varepsilon(D_\mu h_\varepsilon)(\mu)(D_x g)(h_\varepsilon(\mu), \mu) = -\alpha$  and then by using the definition

$$C_\varepsilon^{(X)} := \left\{ (X, \mu) \in \mathbb{R}^2 : X = H_\varepsilon(\mu) := -\frac{1}{2A_\varepsilon(\mu) + \mathcal{O}(\varepsilon)} \right\},$$

we get

$$C_\varepsilon^{(X)} = \{(X, \mu) \in \mathbb{R}^2 : X = \frac{1}{2\alpha} + O_1(\varepsilon)\}.$$

Furthermore from the definition

$$N(\rho; C_\varepsilon) := \left\{ (x, \mu) \in \mathbb{R}^2 : \frac{(x - h_\varepsilon(\mu))^2}{H_\varepsilon(\mu)} < \rho^2 \right\},$$

we get

$$N(\rho; C_\varepsilon) = \left\{ \frac{\sigma^{-2} \text{Var}(x) + O_2(\varepsilon)}{\frac{1}{2\alpha} + O_1(\varepsilon)} < \rho^2 \right\} = \left\{ \frac{\sigma^{-2} \text{Var}(x) + M_2 \varepsilon}{\frac{1}{2\alpha} + M_1 \varepsilon} < \rho^2 \right\}, \quad (11)$$

for which  $M_1, M_2$  are two distinct, positive constants that need to be quantified in order to compute  $\rho$  which will give us the boundaries  $\partial N(\rho; C_\varepsilon) = \sqrt{\rho^2}$ . To do so let us observe that  $X = H_\varepsilon(\mu) = \frac{1}{2A_\varepsilon(\mu)} + O_1(\varepsilon) \Rightarrow |X| = |\sigma^{-2} \text{Var}(x)| < M_1 \varepsilon$ . We recall that  $x(\tau)$  is in fact an OUP for which we have an analytical expression for its variance readily available

$$\text{Var}(x(\tau)) = \left( x_0 - \frac{\sigma^2}{2\alpha} \right) e^{-2\frac{\alpha\tau}{\varepsilon}}. \quad (12)$$

We can plug in the values for  $\alpha, \sigma$  and  $\varepsilon$  in (12) to get  $M_1 = 225$ . Similarly for the other constant we get  $x = h_\varepsilon(\mu) = O_2 \Rightarrow |x| < M_2 \varepsilon$ ; here we assume that  $\max_{\tau \in \mathbb{R}^+} |x(\tau)| = 0.25$  which lead us to the value  $M_2 = 12.5$ . Putting all together into (11) finally gives us  $\rho = \pm 0.35$ . We are now able to compute the closure of the neighborhood  $N(\rho; C_\varepsilon)$  of the critical manifold  $C_\varepsilon$  within which our sample paths  $x(\tau)$  for the OUP (which are normally distributed) will stay bounded with high probability.

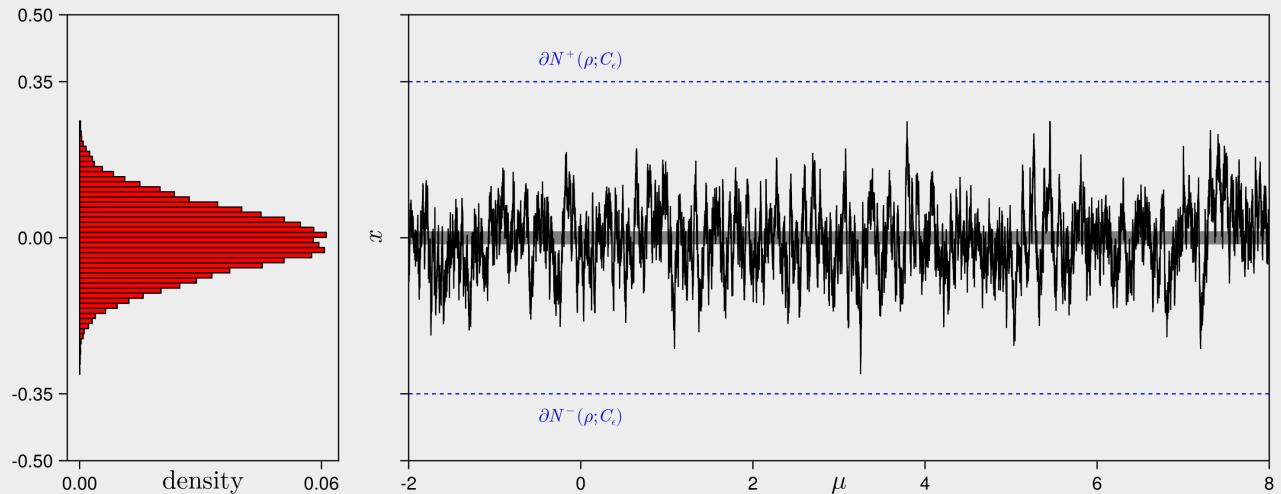


Figure 6: States distribution (left) of a sample path (right) of the OUP (10), with drift  $\alpha = 1$ , timescale separation  $\varepsilon = 2 \cdot 10^{-2}$  and noise level  $\sigma = 0.1$ , starting on the slow manifold  $C_\varepsilon$  (gray strip). Note how the states stay bounded within the neighborhood  $N_\rho(C_\varepsilon)$  (dashed blue lines) with probability 1.

To substantiate the result given by the corollary we can generalise the approach in detecting CSD through the increase in variance of a sample path of the fast-slow problem by computing the moments of the probability distribution solving the Fokker-Planck equation (FPE) associated to the SDE. Considering (9) again for the case of additive noise, the FPE reads

$$\partial_t p(x, t) = -\partial_x J(x, t), \quad J(x, t) := f(x; \mu)p(x, t) - \frac{\sigma^2}{2}\partial_x p(x, t), \quad (13)$$

where  $J(x, t)$  is the probability density current. The fundamental assumption of the analysis carried in [42] is that there exist an asymptotic steady-state  $x_s$  whose distribution  $p_s(x) := p(x_s, t)$  satisfies the stationary limit of (13) i.e.  $J'(x) = 0 \Rightarrow J(x) = \text{constant}(x)$ . We immediately realise that equation (13) is defined over the entirety of the real (spatial) domain, meaning that in order to restrict it to a bounded and limited subinterval  $[a, b] \subset \mathbb{R}$  we must enforce the entirety of the probability current to never leave such domain, resulting in the specification of reflecting boundary conditions. The reason for this immediately follows by the constraint of the stationary Fokker-Planck density to remain normalised in any subdomain

$[a, b]$  we restrict ourselves to. Following the imposition of reflecting boundary conditions we get that the only possible constant that satisfies (13) is  $J(x) = J = 0$ . From the definition of  $J(x, t)$  this leads to a first-order, homogeneous, non-linear ODE whose general integral can be explicitly calculated

$$p_s(x) = \frac{1}{N} e^{2 \int_a^x \sigma^{-2} f(y; \mu) dy}, \quad N = \int_a^b \int_a^z \sigma^{-2} f(y; \mu) dy dz. \quad (14)$$

We can now analytically derive the variance from (14) for simple forms of the drift  $f(x; \mu)$ . In [42] this is done for the saddle-node, transcritical and subcritical pitchfork normal forms and we compare those with the ensemble variance of 1000 sample paths as reported in Figure 7. One clear observation that we can gather from Figure 7 is that in all three cases the (analytic) variance is non-monotonic and it reaches a local maxima in the proximity of the bifurcation at  $\mu = 0$  before decreasing. This apparent anomaly is a result of imposing reflecting boundary conditions for the probability current. As we previously established in fact, the approach of a critical transition implies that the stationary density  $p_s(x)$  becomes less confined around the stable equilibrium leading to N-tipping to occur with higher and higher probability.

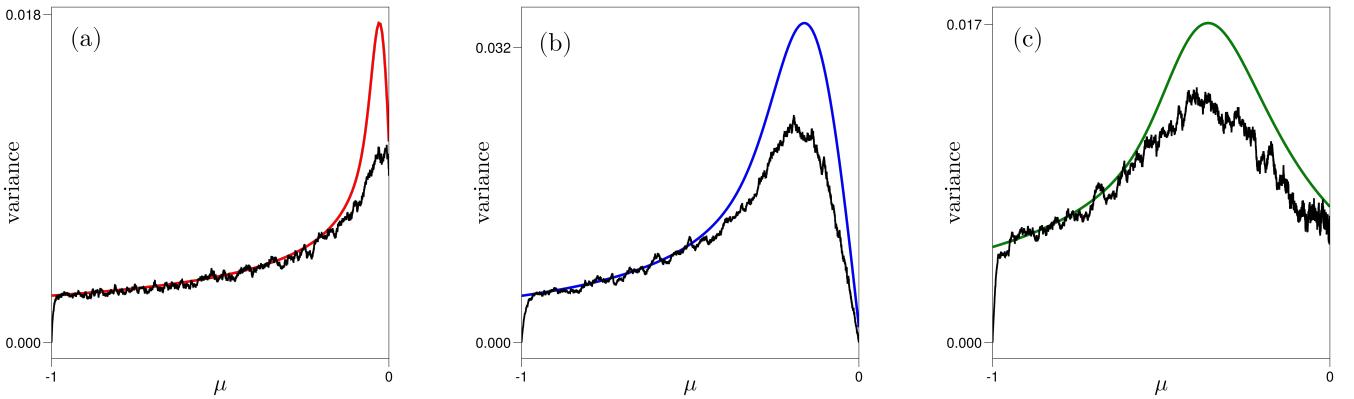


Figure 7: Comparison between the analytical variance of the stationary Fokker-Planck distribution of the saddle-node (a), pitchfork (b) and transcritical (c) normal forms (colored lines) against the numerical variance of an ensemble of 1000 sample paths (black lines) for the same cases (noise level  $\sigma = 0.1$ ).

The second most evident observation concerns the discrepancies between the analytical and numerical variance near the local maxima. In all three cases in fact we observe that the two curves agree when the parameter's values are far away from the deterministic critical transition however they largely differ as the parameter value gets closer to the local maxima for the analytical variance. We can ascribe such discrepancy to (a) the choice that we made in counting the ensemble's sample paths as escaped (i.e. in counting N-tipping events) the exact moment they cross the unstable branch of the critical manifold and (b) the fact that while the analytical variance of the FPE is derived in the singular limit  $\varepsilon = 0$  (due to the stationarity assumption) the ensemble variance is computed within the fast-slow framework. The first condition (a) actually imposes absorbing boundary conditions to our FPE which quite substantially differ from the reflecting boundary conditions used in the analytical derivation especially when boundary and noise effects become dominant (i.e. close to the deterministic critical transition as the basin of attraction shrinks further and further in size). The second condition (b) emphasizes as the solution of the reduced equation in the singular limit is only an approximation of the full fast-slow dynamics. This argument is validated further by observing how the discrepancy between the analytical and numerical variance for the saddle-node bifurcation is significantly less than those for the remaining two normal forms; based on this observation we expect fewer trajectories to escape the stable equilibrium for the saddle-node normal form and this is exactly what we observe in the ensemble simulations as depicted below in Figure 8.

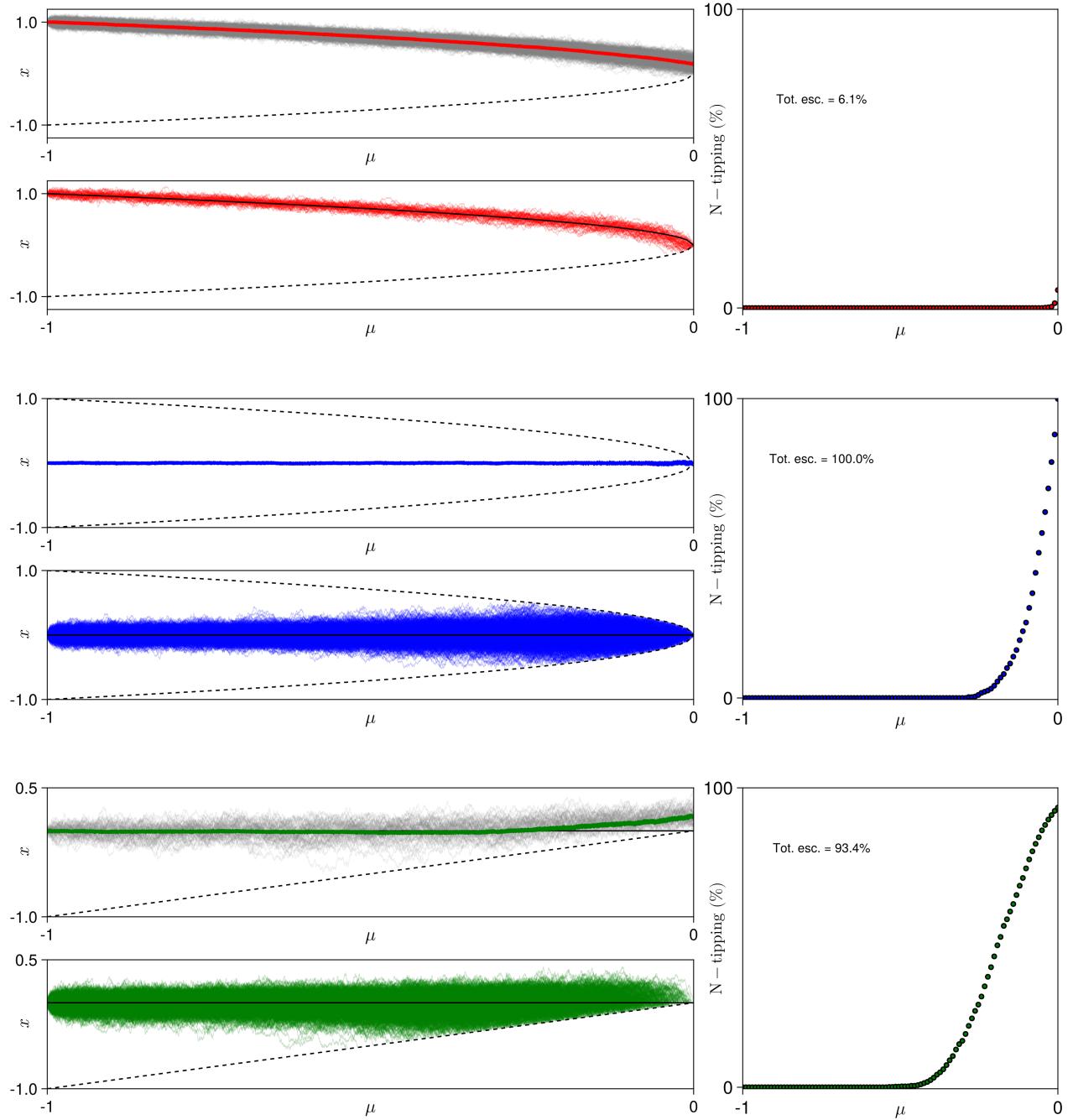


Figure 8: Ensemble sample paths of saddle-node (top), pitchfork (middle) and transcritical (bottom) normal forms showing the trajectories that stayed bounded on the stable branch of the critical manifold (top panels: gray lines indicate the ensembles trajectories and the colored lines show the ensemble mean) and those that escaped due to N-tipping (bottom panels). On the left of each plot a percentage of the number of escaped trajectories as  $\mu$  varies is shown.

### Example 3.3

Consider the case of a transcritical normal form with multiplicative noise

$$dx = (\mu x - x^2 + \frac{1}{2}\sigma^2 x) dt + \sigma x dW,$$

as presented in [42]. We write the associated FPE which, in the singular limit, it reads

$$0 = -\frac{d}{dx}((yx - x^2 + \frac{1}{2}\sigma^2 x)p_s(x; y)) + \frac{d^2}{dx^2}(\frac{1}{2}\sigma^2 x^2 p_s(x; y)).$$

One normalizable solution of the above stationary FPE is

$$p_s(x) = \frac{1}{N_\mu} x^{\frac{2\mu}{\sigma^2}-1} e^{-\frac{2x}{\sigma^2}}, \quad N_\mu = \int_0^{+\infty} x^{\frac{2\mu}{\sigma^2}-1} e^{-\frac{2x}{\sigma^2}} dx = \left(\frac{\sigma^2}{2}\right)^{\frac{2\mu}{\sigma^2}} \Gamma\left(\frac{2\mu}{\sigma^2}\right).$$

We recall the definition of the  $\Gamma$ -probability density function

$$f_{\alpha,\beta}(x) := \frac{1}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \beta^\alpha \quad (15)$$

where  $\Gamma(z) := \int_0^\infty y^{z-1} e^{-y} dy$  is the Euler's Gamma function for  $z \in \mathbb{C}$  while  $\alpha, \beta \in \mathbb{R}$  parametrise the distribution. The first central moments of (14) are well known, in particular

$$\mathbb{E}_f(x; \alpha, \beta) = \frac{\alpha}{\beta}, \quad \text{Var}_f(x; \alpha, \beta) = \frac{\alpha}{\beta^2}. \quad (16)$$

With a simple change of variables we immediately verify that the normalised solution of the FPE of the transcritical stochastic process with linear, multiplicative noise, is equivalent to (14)

$$\alpha = \frac{2\mu}{\sigma^2}, \quad \beta = -\frac{2}{\sigma^2}.$$

Putting the above into (16) yields

$$\mathbb{E}_{p_s}(x; \mu) = \mu, \quad \text{Var}_{p_s}(x; \mu) = \frac{\sigma^2}{2} \mu.$$

What is left to verify is that under the above change of variables the normalisation constant  $N_\mu$  coincides with that of the  $\Gamma$ -distribution. We write

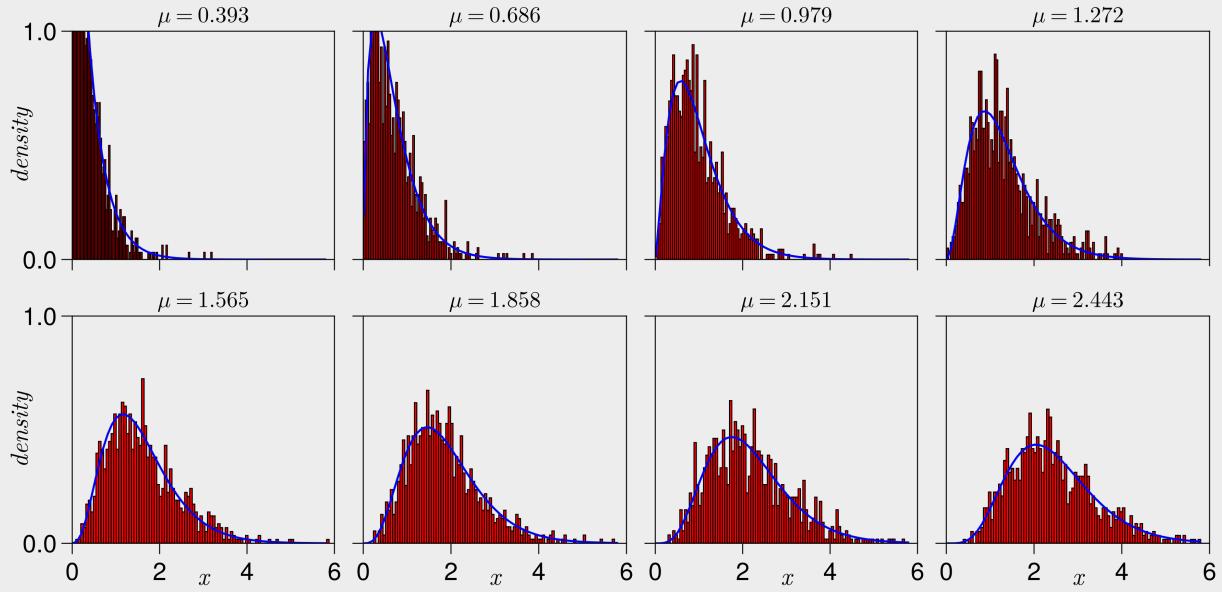
$$N_\mu = \int_0^\infty x^{\frac{2\mu}{\sigma^2}-1} e^{-\frac{2x}{\sigma^2}} dx,$$

and perform a  $u$ -substitution  $x = \frac{\sigma^2}{2}u \rightarrow dx = \frac{\sigma^2}{2}du$  to get

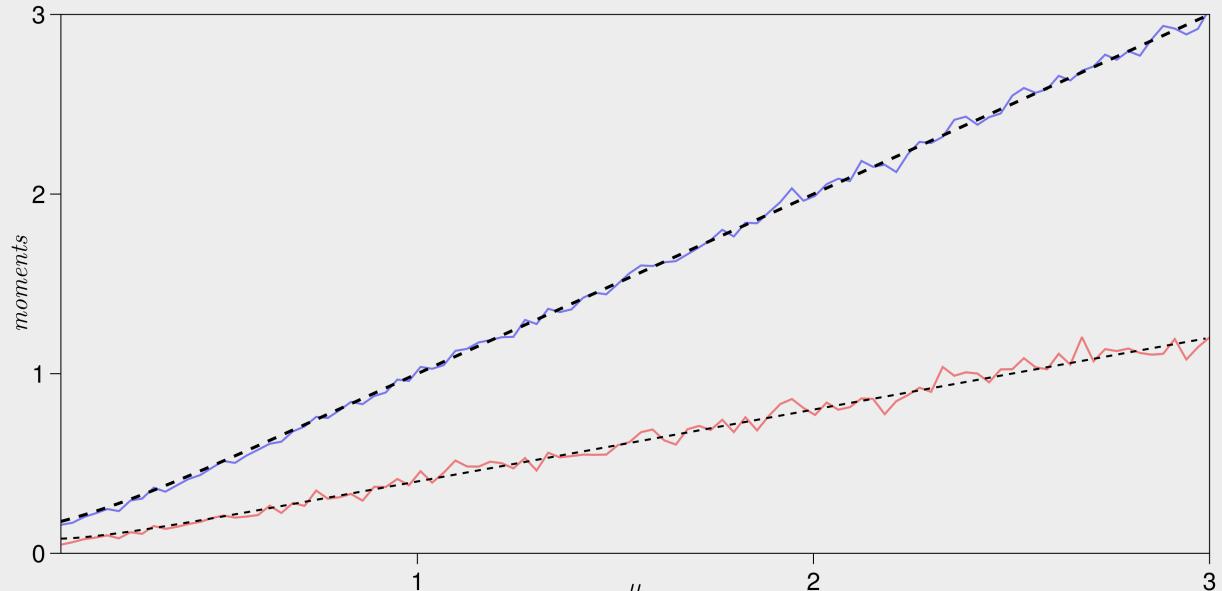
$$N_\mu = \int_0^\infty \frac{\sigma^2}{2} \left(\frac{u\sigma^2}{2}\right)^{\frac{2\mu}{\sigma^2}-1} e^{-u} du,$$

which can be integrated by parts yielding

$$N_\mu = \left(\frac{\sigma^2}{2}\right)^{\frac{2\mu}{\sigma^2}-1} \left(\frac{\sigma^2}{2}\right) \int_0^\infty u^{\frac{2\mu}{\sigma^2}-1} e^{-u} du = \left(\frac{\sigma^2}{2}\right)^\alpha \int_0^\infty u^\alpha e^{-u} du = \left(\frac{\sigma^2}{2}\right)^\alpha \Gamma(\alpha).$$



(a)



(b)

Figure 9: (a) Snapshots of the stationary Fokker-Plank distribution  $p_s(x)$  (solid blue lines) for the transcritical normal form with multiplicative noise scaled by  $\sigma = \sqrt{0.8}$  at different values of the parameter overlaid with histograms (red bars) associated to an ensemble simulation of 1000 sample paths. (b) Comparison between the analytical parameter variation of the mean and variance of the stationary FPE (dashed black lines) with the sample estimates from the ensemble trajectories (ensemble average in blue while variance is in red).

## 3.2 High-dimensional systems

The construction of a dynamical model that exhibits tipping points driven by different phenomena has so far relied on a number of assumptions that have been subsequently relaxed in order to achieve a more flexible and realistic mathematical description of observed critical events. Throughout the previous analysis we have however restricted ourselves in considering a finite number of observables whose time-evolution, if taken to be a model of real phenomena defined on a spatio-temporal domain, represent the dynamics of spatially-aggregate measures. In other words we have so far only consider observables that are intrinsically defined over a spatial domain whose change with time happens uniformly throughout space. This simplification is powerful, as it allowed the derivation of robust mathematical frameworks in e.g. [42] and [45] for the analysis of tipping events and their precursors however it also implies the discard of any potential instantaneous change in the spatial distribution of those state variables. Indeed ignoring how observables change in space and only considering the time-evolution of their properties automatically restrict the analysis of either events that are not defined on a spatial domain or spatially heterogeneous models that we can only measure in an aggregate sense (i.e. by considering spatial means of the observables). In both instances, relying on such simplification heavily implies the loss of meaningful information and methods of mathematical analysis thus preventing the detection of precursors s.a. self-organising pattern-formation. As such this last strong assumption of spatial homogeneity will be dropped as we will now consider observables  $x \mapsto u(x)$ ,  $x \in \Omega \subseteq \mathbb{R}^d$  ( $d = 1, 2, 3$ ) as (scalar) functions varying in both space and time. These functions  $u(x, t) \in V \times \mathbb{R}^+$  can be thought as solutions of space-time partial differential equations (PDEs) on a functional space  $V$  which is infinite dimensional. We will thus refer to these spatio-temporal dynamical systems, whose phase space coincides with  $V$ , as high-dimensional as opposed to the low-dimensional case insofar treated for a set of finite observables  $x \in \mathbb{R}^n$  with  $n \ll \infty$ .

### 3.2.1 Intrinsic discrete phenomena

As an intermediate step between low and high-dimensional problems, lattice dynamical systems (LDSs) have been investigated in the context of EWS detection. These are (large) systems of coupled ODEs organised on a  $d$ -dimensional regular lattice; each of such ODEs models the dynamics of the same observable  $u$  at a given lattice site (node). These systems are therefore intrinsically discrete in space while still retaining spatial heterogeneity via the coupling of the system. As such one can use the temporal indicators derived above for low-dimensional systems while enriching the analysis by considering spatial statistics at each timestep of observation. One specific model of interest has captivated the efforts of recent works towards the construction of a robust theory of spatio-temporal EWS and that is the reaction-diffusion problem. Given a 2-dimensional rectangular lattice of  $n \times m$  nodes, a generic reaction-diffusion problem perturbed by additive noise can be written, in cartesian coordinates, as

$$du_{j,k} = (D_{j,k} \phi(u_{j,k}, u_{j\pm 1,k}, u_{j,k\pm 1}) + R_{j,k} \psi(u_{j,k})) dt + \sigma dW_{j,k}, \quad (17)$$

where  $\phi$  represents the nearest-neighbour coupling modelling the diffusion component while  $\psi$  is responsible for the reaction of  $u$  at lattice site  $(j, k)$ . A trivial transformation  $\ell = n(k-1) + (j+1)$  allows us to recast (16) in linear coordinates

$$du_\ell = (D_\ell \phi(u_\ell, u_{\ell\pm 1}, u_{\ell\pm n}) + R_\ell \psi(u_\ell)) dt + \sigma dW_\ell, \quad (18)$$

which is more convenient for our analysis. The inclusion of a slowly-varying parameter  $\mu$  in (18) completes our intermediate model to a more general case

$$\begin{cases} du_\ell = (D_\ell(\mu) \phi(u_\ell, u_{\ell\pm 1}, u_{\ell\pm n}; \mu) + R_\ell(\mu) \psi(u_\ell; \mu)) dt + \sigma dW_\ell, \\ d\mu = \varepsilon dt. \end{cases} \quad (19)$$

We end this section by noticing that  $D$  and  $R$  are (generally inhomogeneous) scalar fields assigning relative (local) magnitudes to the diffusion and reaction phenomena respectively and they are paramount in the analysis of the onset of instabilities in these problems as addressed in the following Section.

### 3.2.2 Instabilities in reaction-diffusion problems

Reaction-diffusion equations are of particular interest in mathematical modelling and analysis of differential equations because of how different instabilities of homogeneous stable states may cause very different asymptotic solutions: examples are travelling (stationary) waves, solitons, regular and irregular patterns. In the context of this report and at large as the primary object of investigation of this research proposal going forward, the rise of spatial patterns from homogeneous solutions (Turing instabilities) will be considered as the tipping event of interest connecting two, very different steady-states of the system. We choose to target these particular events for two reasons:

- ▶ to narrow the scope of an otherwise prohibitively broad type of instabilities that can occur in high-dimensional systems;
- ▶ since they are observed to be peculiar to natural systems that exhibit transition from a *healthy* state (e.g. homogeneous ventilation in the airways of a lung or locally homogeneous vegetation distribution in a florid ecosystem) to an *unhealthy* one (onset of CVDs in a lung undergoing asthma and the rise of patches of turbidity prior desertification).

To illustrate simple techniques in detecting pattern formations in 2–dimensional reaction-diffusion LDSs we consider first a model of vegetation turbidity analysed in [68]. From (19) we set

$$\begin{aligned}\phi(u_\ell, u_{\ell \pm 1}, u_{\ell \pm n}; \mu) &= u_{\ell+1} + u_{\ell+1} + u_{\ell-1} + u_{\ell+n} + u_{\ell-n} - 4u_\ell, \\ \psi(u_\ell; \mu) &= r_v u_\ell \left(1 - u_\ell \frac{r_\ell^4 + E_\ell^4(\mu)}{r_\ell^4}\right), \quad E_\ell(\mu) = \mu \frac{h_v}{h_v + u_\ell},\end{aligned}$$

with  $r_\ell$  being sampled uniformly at random in  $[0.6, 1.0]$  and independently for each lattice site while  $r_v = 0.5$  and  $h_v = 0.2$  are uniform throughout the domain. In the paper the authors aim at detecting the formation of irregular patterns by means of tracking the leading eigenvalue of the spatial covariance matrix  $\Sigma \in \mathbb{R}^{N \times N}$  where  $N := n \cdot m$  is the dimension of a single snapshot of the solution defined on a regular  $n \times m$  lattice at each timestep. The matrix is assembled considering  $W$  consecutive snapshots collected in  $\mathcal{X} \in \mathbb{R}^{N \times W}$  across a temporal sliding window and the eigendecomposition of  $\Sigma$  will provide the proposed EWS

$$\begin{aligned}\mathcal{X}(w) = \begin{pmatrix} \mathbf{x}(t_w) & \dots & \mathbf{x}(t_{w+W}) \end{pmatrix} \Rightarrow \{\Sigma(w)\}_{j,k} &= \text{cov}(\mathbf{x}_j(t_{w \rightarrow w+W}), \mathbf{x}_k(t_{w \rightarrow w+W})) \\ &= \mathbb{E}(\mathbf{x}_j(t_{w \rightarrow w+W}) \mathbf{x}_k(t_{w \rightarrow w+W})) - \\ &\quad - \mathbb{E}(\mathbf{x}_j(t_{w \rightarrow w+W})) \mathbb{E}(\mathbf{x}_k(t_{w \rightarrow w+W})),\end{aligned}\tag{20}$$

where  $\mathbf{x}_\ell(t_{w \rightarrow w+W}) \in \mathbb{R}^W$  is a vector of  $W$  observations (truncated time-series) of site  $x_\ell \in \mathcal{L}$  across the sliding window. The motivation behind this proposal comes from the combination of the conceptual CSD observed in low-dimensional systems discussed before and the intuition that, as the system moves away from a state of spatially homogeneous distribution to one of patched heterogeneity the spatial variance shall increase.

### Theorem 3.4

Let  $\{\sigma_1, \dots, \sigma_N\}$  be the eigenvalues of the covariance matrix defined in (20) and  $\{\lambda_1, \dots, \lambda_N\}$  be the eigenvalues (with negative real parts) of the Jacobian around an attractor and denote with  $\sigma_{\max}$  and  $\lambda_{\max}$  the leading eigenvalues of the covariance and Jacobian matrix respectively, then as  $\lambda_{\max} \rightarrow 0$  (i.e. the bifurcation is approached)  $\sigma_{\max} >> \sigma_{j=1,\dots,N} \neq \sigma_{\max}$ .

*Proof.* TBA... □

This proposed EWS is a representative example of a spatio-temporal indicator of pattern-formation as it merges together the spatial information of increased heterogeneity with the loss of resilience to stochastic perturbations of the stable attractor as the leading eigenmode of the linearisation becomes weaker approaching the Turing instability (B-tipping). Below in Figure 10 we report our experimental detection of such EWS for the vegetation turbidity model introduced above.

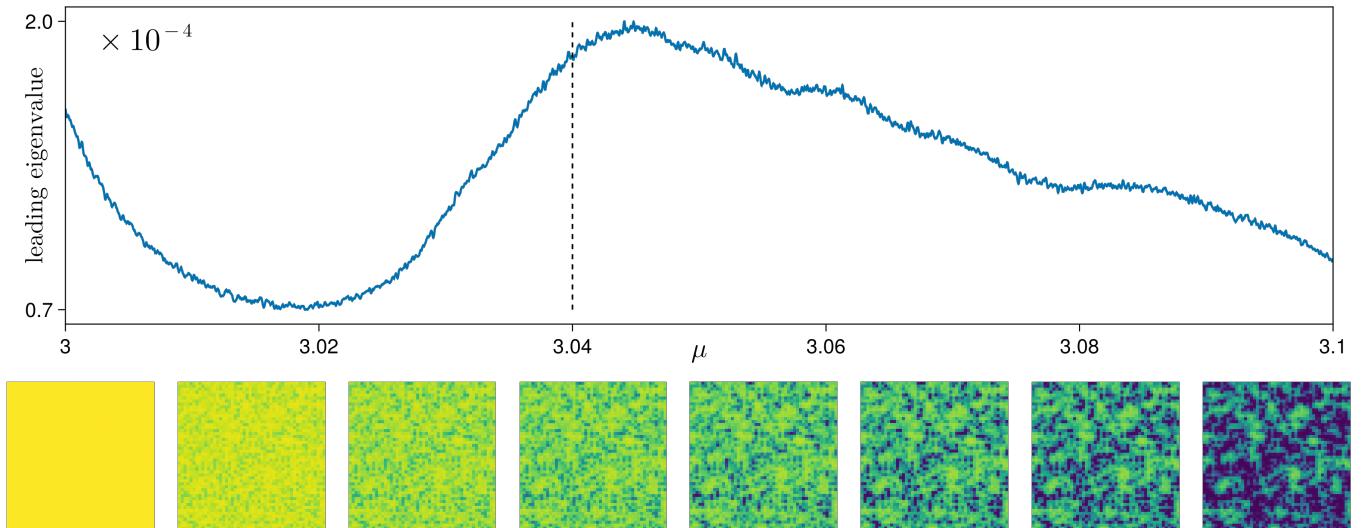


Figure 10: Time evolution of the vegetation turbidity LDS on a square  $50 \times 50$  lattice with slow ( $\varepsilon = 10^{-2}$ ), linear ramping of the bifurcation parameter, noise level  $\sigma = 10^{-3}$  and uniform reaction  $R = 1$  and diffusion  $D = 0.2$ . The top pannel shows the timeseries of the leading eigenvalue of the covariance matrix assembled over a window of  $W = 20$  snapshots. The dashed vertical line indicates the visual detection of irregular patterns forming in the solution, shown in the bottom panels at increasing timesteps (left to right).

The fundamental idea behind this analysis is to bring together the spatial and temporal variation of the observables in a (prebafarly unique) measure that is capable of capturing, with robustness and mathematical interpretability the onset of pattern-formation in spatially extended dynamical systems. This concept is further investigated in [79] where the dynamic mode decomposition (DMD) of the Koopman operator linearising the evolution of the reaction-diffusion LDS is employed. Suppose  $\mathcal{X}$  is collecting snapshots from timestep  $t_w$  to  $t_{w+W}$  while in  $\mathcal{X}'$  we store the snapshots one stride of the sliding window away i.e. capturing the snapshots from  $t_{w+S}$  to  $t_{(w+S)+W}$ . The DMD will provide us with a more direct and efficient way of capturing CSD and Turing instability by way of approximating the linearisation around the steady-state directly as per the following result.

### Theorem 3.5

Let  $A : \mathcal{X} \rightarrow \mathcal{X}'$  be a linear operator and  $\mathcal{X} \approx \mathbf{U}\Sigma\mathbf{V}^T$  be the singular valued decomposition (SVD) of  $\mathcal{X}$ , then the spectrum of  $\mathbf{S} := \mathbf{U}^T \mathcal{X}' \mathbf{V} \Sigma^{-1}$  is the best approximation of the spectrum of A ( $\sigma(A)$ ) and its eigenvectors are also optimally approximated by those of  $\mathbf{S}$  multiplied from the left by  $\mathbf{U}$ .

*Proof.* Here  $\mathcal{X}$  will represent a matrix of snapshots of the solution at different timesteps over the square  $N$ -dimensional lattice (i.e.  $N := n^2$ ) collected across a sliding window of fixed width  $W$ . Since  $\mathcal{X} \in \mathbb{R}^{N \times W}$  then the SVD will give us the left and right singular vectors in  $\mathbf{U}, \mathbf{V} \in \mathbb{R}^{N \times W}$  respectively while the singular values will be stored along the (main) diagonal of  $\Sigma \in \mathbb{R}^{W \times W}$  (which is not the covariance matrix as in the previous case). From straightforward algebraic manipulation we get

$$A = \mathcal{X}' \mathcal{X}^{-1} \approx \mathcal{X}' (\mathbf{U} \Sigma \mathbf{V}^T)^{-1} = \mathcal{X}' \mathbf{V}^{-T} \Sigma^{-1} \mathbf{U}^{-1} = \mathcal{X}' \mathbf{V} \Sigma^{-1} \mathbf{U}^T \Rightarrow \mathbf{U}^{-1} A \mathbf{U} \approx \mathbf{U}^T \mathcal{X}' \mathbf{V} \Sigma^{-1} =: \mathbf{S},$$

where we used the fact that both  $\mathbf{U}$  and  $\mathbf{V}$  are orthogonal. The eigendecomposition of  $\mathbf{S}$  will then give

$$\mathbf{U}^{-1} A \mathbf{U} \approx \mathbf{S} = \mathbf{Q} \Lambda \mathbf{Q}^{-1} \Rightarrow A \approx (\mathbf{U} \mathbf{Q}) \Lambda (\mathbf{U} \mathbf{Q})^{-1}.$$

□

As an illustrative example of how the leading eigenvalue of the DMD can provide a prognostic of pattern-formation events here we consider the LDS formulation of a lung ventilation model introduced in [60]. Again from (19) we write

$$\begin{aligned} \phi(u_\ell, u_{\ell \pm 1}, u_{\ell \pm n}; \mu) &= \sigma(-r_\ell), \\ \psi(u_\ell; \mu) &= -u_\ell, \end{aligned}$$

where  $\sigma(\theta) = \frac{1}{1+e^{-\theta}}$  is the sigmoid function and

$$r_\ell := -P_b(t) + \mu \frac{k}{u_\ell} + P_b A (u_\ell^4 + u_{\ell+1}^4 + u_{\ell-1}^4 + u_{\ell+n}^4 + u_{\ell-n}^4) (1 - 1 - u_\ell + \frac{3}{2}(1 - u_\ell)^2),$$

with  $P_b(t) := \frac{P_b(0)N}{\sum_{\ell=1}^N u_\ell^4}$ . Notice that this last term enforces global-coupling of the lattice sites alongside the traditional local, nearest-neighbour interactions of the discrete diffusion operator. Finally the remaining homogeneous constants are set to  $k = 14.1$ ,  $A = 0.63$ ,  $P_i = 0.96$ ,  $P_b(0) = 7.25$ . In Figure ?? below we report the detection of the proposed spatio-temporal EWS for this model.

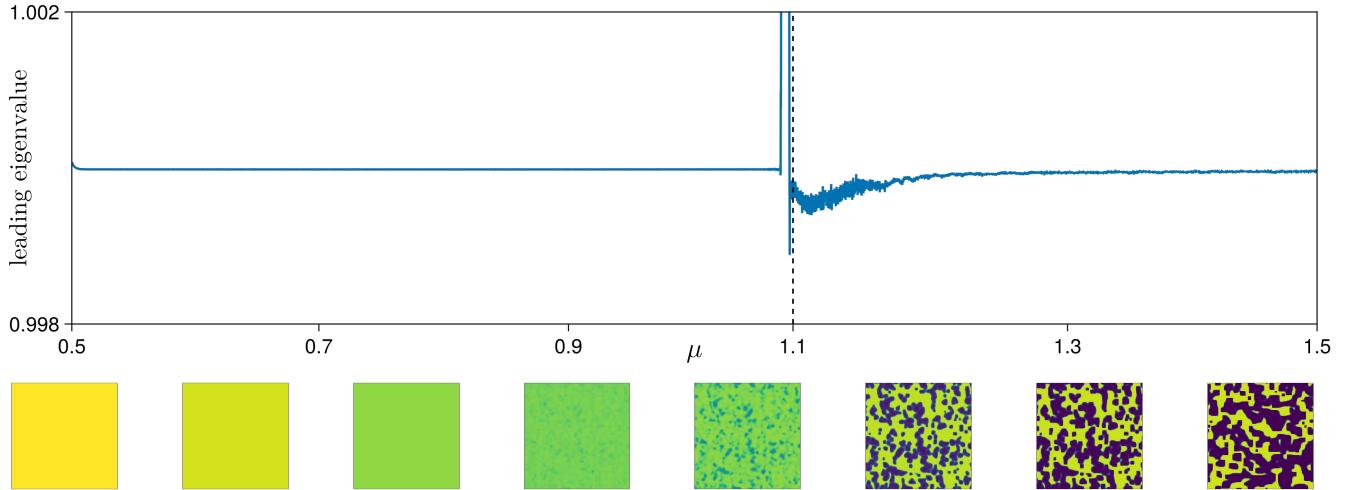


Figure 11: Time evolution of the lung ventilation LDS on a square  $300 \times 300$  lattice with slow ( $\varepsilon = 10^{-2}$ ), linear ramping of the parameter and noise level  $\sigma = 10^{-3}$ . In the top panel the leading mode of the DMD is reported, showing an abrupt upward burst shortly before the formation of patterns (dashed vertical line). In the bottom pannel several snapshots increasing timesteps are reported.

### 3.3 A comprehensive catalogue of spatio-temporal early-warning signals

The two indicators of pattern-forming tipping events discussed above are but among the latest entries of a much larger ecosystem of EWS that have been proposed throughout the past two decades. Their distinctive property w.r.t. the rest of the earlier measures is that they both combine spatial and temporal information rather than looking at either dimensions individually. As addressed in the earlier literature review, the discovery and characterisation of tipping events was preceeded by the observation of critical phenomena s.a. CSD and flickering whose measurement through statistical and dynamical indicators become later EWS. As discussed in the previous Chapter, the robustness of the signals of these precursors is of uttermost importance. Specifically an optimal EWS of tipping events should address:

- ▶ the existence (or lack thereof) of a sound mathematical framework linking the properties and structure of the models exhibiting a tipping event with the underlying theorised or observed phenomena (i.e. CSD, flickering etc...);
- ▶ a reasonably wide scope of applicability (or degree of universality) to different natural systems modelled by the same mechanism (e.g. reaction-diffusion equations);
- ▶ the sensitivity of the indicator to downsampling and other practical limitations in the monitoring of complex real-world systems (s.a. climate change, ecosystem collapse, onset of asthma etc...);
- ▶ the reliance of the methods detecting the signal on single and multivariate timeseries analysis, including detrending and other preprocessing techniques.

Addressing how robust a spatio-temporal EWS is w.r.t. the aforementioned metrics is arguably among the most important tasks ahead in the discovery and characterisation of catastrophic regime shifts. To conclude this Chapter, here we limit ourselves in categorizing the indicators of the precursors discussed in the literature review. A tabulated list of indicators is not new and as a matter of fact we found 5 works [43, 47, 48, 56, 61] published in the first half of the 2010s compiling a comprehensive collection of the EWS being proposed until then and their connections to the relevant precursors. Remarkably [61] was also

accompanied by the creation of a website<sup>2</sup> collecting and detailing all the progresses that have been done insofar in the detection of both ad-hoc and generic EWS. In Table 1 below we merge these lists by selecting only those indicators that have shown promises of robustness to some degree while enriching them with newer signals proposed after the publication of the aforementioned papers. We remark that the discovery of critical phenomena preceding a critical transition and characterising them using indicators has been an ongoing effort by multiple communities with a strong focus in the ecological and climate sciences. This means that the mathematical ground upon which many of those precursors are based varies greatly, from semi-rigorous and well-developed frameworks for idealised cases [42, 53] to more pragmatic, data-driven and intuition-based indicators of experimental monitorings for scientific applications [31, 43].

Indicator	Phenomenon	First proposed
recovery time	CSD	1995 ([5])
spectral reddening	CSD	2003 ([12])
autoregressive model fitting	CSD	2004 ([14])
lag-1 autocorrelation	CSD	2004 ([14])
<b>size distribution of patches</b>	Turing instability	2004 ([15])
<b>spatial variance</b>	CSD, flickering, Turing instability	2005 ([16])
variance	CSD, flickering	2006 ([20])
detrended fluctuation analysis	CSD	2007 ([22])
<b>leading eigenvalue of the covariance</b>	CSD, Turing instability	2008 ([24])
skewness	flickering	2008 ([26])
mean exit time	CSD	2008 ([26])
kurtosis	flickering	2009 ([27])
<b>spatial skewness</b>	flickering, Turing instability	2009 ([29])
<b>spatial spectral reddening</b>	CSD, Turing instability	2010 ([33])
<b>spatial correlation</b>	CSD, Turing instability	2010 ([36])
potential analysis	flickering	2010 ([38])
<b>traveling wavespeed</b>	flickering	2013 ([53])
<b>leading mode of the DMD</b>	CSD, Turing instability	2020 ([72])
<b>mutual information</b>	CSD	2024 ([84])

Table 1: Catalogue of robust EWS of critical transitions sorted from top to bottom by the year in which they were first proposed over the past 2 decades. In bold we indicate spatial and spatio-temporal indicators. In gray we highlight the most promising EWS in spatially extended dynamical systems which will also be the focus of the next Chapters.

<sup>2</sup><https://www.early-warning-signals.org/>

## 4 A probabilistic early-warning from statistical mechanics

### 4.1 Escape problem

#### 4.1.1 The many ways to derive the problem

In the following we review several, equivalent methods to approach the problem of establishing the exit time of a stochastically perturbed particle sitting in a local minimum of a potential function  $U(x)$  (stable equilibrium of a conservative vector field  $f(x) = -V'(x)$ ). The general framework starts with a SDE in Itô form with the only assumption of the noise being additive

$$dx = f(x)dt + \sigma dW, \quad (21)$$

and its associated FPE (or Kolmogoroff forward equation, KFE)

$$\partial_t p(x, t) = \mathcal{L}_{FP} p(x, t) = D\partial_{xx}^2 p(x, t) - \partial_x(f(x)p(x, t)) = D\partial_{xx}^2 p(x, t) + \partial_x(V'(x)p(x, t)), \quad (22)$$

where  $D = \frac{\sigma^2}{2}$  defines the diffusion coefficient. In the following derivations we will also use the adjoint of the FP operator hence we hereby compute it. By definition, given a bounded, linear operator  $\mathcal{L} : U \rightarrow U$  on an inner product space  $U$ , its adjoint  $\mathcal{L}^\dagger$  is constructed to satisfy the relation

$$\langle \mathcal{L}p, q \rangle = \langle p, \mathcal{L}^\dagger q \rangle$$

where the bra-ket notation identifies the inner product on  $V$ . A suitable choice of  $V$  for the FP operator would be

$$V = \left\{ f \in H_0^1(\mathbb{R}) : \int_{-\infty}^{+\infty} f dx = 1 \right\}$$

i.e. it would be the space of infinitely differentiable functions which vanish at  $\pm\infty$ , alongside their first derivative, and whose integral is 1 (in order to make them *pdfs*). The derivation of the adjoint follows

$$\begin{aligned} \langle \mathcal{L}_{FPP}, q \rangle &= \int_{-\infty}^{+\infty} (\mathcal{L}_{FPP}(x))q(x)dx = \int_{-\infty}^{+\infty} \left( Dp'' + (V'p)' \right) q dx = \int_{-\infty}^{+\infty} \left( Dp' + V'p \right)' q dx = \\ &= \lim_{x \rightarrow \pm\infty} (Dp' + V'p) \overbrace{q}^0 - \int_{-\infty}^{+\infty} \left( Dp' + V'p \right) q' dx = -D \int_{-\infty}^{+\infty} p' q' dx - \int_{-\infty}^{+\infty} V' p q' dx = \\ &= -D \left( \lim_{x \rightarrow \pm\infty} p q' \overbrace{-}^0 \int_{-\infty}^{+\infty} p q'' dx \right) - \int_{-\infty}^{+\infty} p V' q' dx = \int_{-\infty}^{+\infty} p \left( Dq'' - V' q' \right) dx = \\ &= \int_{-\infty}^{+\infty} p(x)(\mathcal{L}_{FP}^\dagger q(x))dx = \langle p, \mathcal{L}_{FP}^\dagger q \rangle. \end{aligned}$$

The adjoint FPE (or Kolmogoroff backward equation, KBE) thus reads

$$\partial_t p(x, t) = -\mathcal{L}_{FP}^\dagger p(x, t) = V'(x)\partial_x p(x, t) - D\partial_{xx}^2 p(x, t). \quad (23)$$

#### Derivation from first passage time

(WRONG: needs review!)

Suppose you know the solution  $p(x, t)$  of (22). Suppose as well you have two subsets  $\Omega_0, \Omega_T \subset \Omega$  of the sample space. The first passage time problem concerns the derivation of the probability distribution  $h(t)$

of the time  $t = T$  that a particle takes to reach  $\Omega_T$  starting at time  $t = 0$  in  $\Omega_0$  (henceforth referred to as *hitting time distribution*). Let us fix a specific time  $t = T > 0$ ; from the FP distribution we can compute

$$P_T(x \notin \Omega_T) = P_T(x \in \Omega_T^c) = \int_{\Omega_T^c} p(x, t = T) dx, \quad (24)$$

which is the probability that, at time  $t = T$  the particle has not reached  $\Omega_T$ . If instead of looking at a particular point in time we look at the entire history of the process (i.e. at any  $t > 0$ ) then Equation (24) will read

$$S(t) = \int_{\Omega_T^c} p(x, t) dx, \quad (25)$$

which gives us a marginal probability distribution of the time of survival of the particle (see Figure 12a).

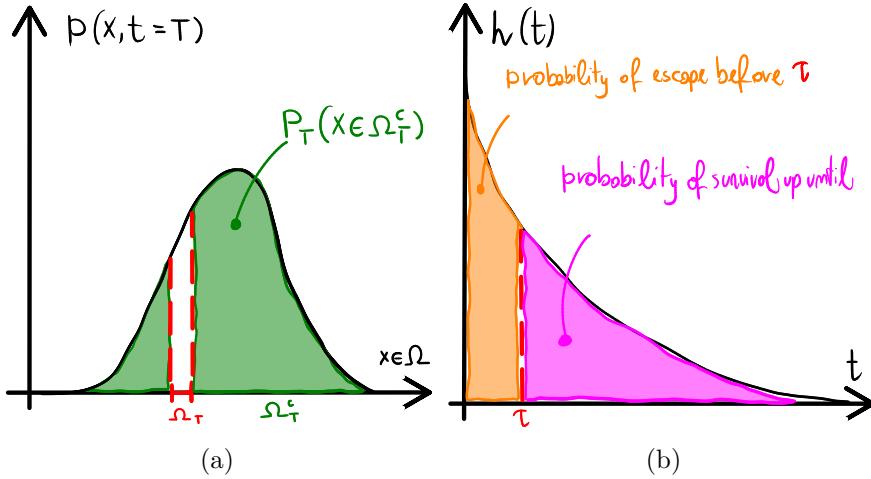


Figure 12

Now suppose instead that you already know the hitting time distribution  $h(t)$ . Again let us fix  $t = T > 0$ ; the following quantifies the probability of the particle reaching  $\Omega_T$  before  $T$

$$P_T(x \in \Omega_T) = \int_0^T h(t) dt. \quad (26)$$

From Equation (26) we can easily recover the probability of survival up to time  $T$

$$\begin{aligned} P_T(x \in \Omega_T^c) &= 1 - P_T(x \in \Omega_T) = \int_0^{+\infty} h(t) dt - \int_0^T h(t) dt \\ &= \left( \int_0^T h(t) dt + \int_T^{+\infty} h(t) dt \right) - \int_0^T h(t) dt = \int_T^{+\infty} h(t) dt. \end{aligned} \quad (27)$$

As we did in Equation (25), we now let  $t$  to vary and therefore we recover the same distribution of survival time for our particle

$$S(t) = \int_t^{+\infty} h(\tau) d\tau. \quad (28)$$

If we differentiate Equation (28) w.r.t. time we get

$$\partial_t S(t) = \int_t^{+\infty} \frac{\partial}{\partial \tau} h(\tau) d\tau = \left( \lim_{\tau \rightarrow +\infty} h(\tau) \right)^0 - h(t) = -h(t), \quad (29)$$

and if we differentiate Equation (25) w.r.t. time we get

$$\partial_t S(t) = \int_{\Omega_T^c} \partial_t p(x, t) dx \stackrel{\text{Eq.(22)}}{=} \int_{\Omega_T^c} \mathcal{L}_{FP} p(x, t) dx. \quad (30)$$

So from the equality of Equation (29) and (30) we get

$$h(t) = - \int_{\Omega_T^c} \mathcal{L}_{FP} p(x, t) dx. \quad (31)$$

Equation (31) says that the marginal w.r.t. the sample space (the complement of the target set to be precise) of the FP operator gives us the distribution of first passage time from  $\Omega_0$  to  $\Omega_T$ . However, to assemble  $h(t)$  in practice we will need to first know  $p(x, t)$ , then compute its FP functional  $\mathcal{L}_{FP} p(x, t)$  and finally compute the marginal of that. In most cases we will not be able to compute  $p(x, t)$  explicitly. From the definition of the  $n$ -th central moment of  $h(t)$

$$E(t^n) = \int_0^{+\infty} t^n h(t) dt \stackrel{\text{Eq.(31)}}{=} - \int_0^{+\infty} t^n \left( \int_{\Omega_T^c} \mathcal{L}_{FP} p(x, t) dx \right) dt = - \int_0^{+\infty} t^n \partial_t \left( \int_{\Omega_T^c} p(x, t) dx \right) dt,$$

we set  $n = 1$  and write the mean of  $h(t)$ , which we label the *mean first passage time* or *MFPT*

$$E(t) = \langle t \rangle =: T = - \int_0^{+\infty} \underbrace{t^n}_{u(t)} \underbrace{\partial_t \left( \int_{\Omega_T^c} p(x, t) dx \right)}_{v'(t)} dt. \quad (32)$$

We can integrate Equation (32) by parts to get

$$\begin{aligned} T &= \left[ u(t)v(t) \right]_0^{+\infty} - \int_0^{+\infty} u'(t)v(t) dt = \int_0^{+\infty} \left( \int_{\Omega_T^c} p(x, t) dx \right) dt = \int_{\Omega_T^c} \left( \int_0^{+\infty} p(x, t) dt \right) dx = \\ &= \int_{\Omega_T^c} G(x) dx, \end{aligned} \quad (33)$$

where in the first step we assumed that  $\lim_{t \rightarrow \infty} v(t) = \lim_{t \rightarrow \infty} \int_{\Omega_T^c} p(x, t) dx = 0$ , i.e. the probability of the particle to stay bounded vanishes after long enough time. Equation (33) tells us that, in order to find the MFPT, we need to derive the marginal distribution of the states  $G(x)$  across the entire history. This is in general difficult to do since the FP distribution is unknown; instead we trade-off the explicit definition of the MFPT in order to simplify the RHS of Equation (33). By applying the FP operator to both sides we get

$$\begin{aligned} \mathcal{L}_{FP} T(x) &= \int_{\Omega_T^c} \mathcal{L}_{FP} G(x) dx \stackrel{\text{Eq.(22)}}{=} \int_{\Omega_T^c} \partial_t G(x) dx = \int_{\Omega_T^c} \left( \int_0^{+\infty} \partial_t p(x, t) dt \right) dx = \\ &= \int_{\Omega_T^c} \left( \lim_{t \rightarrow +\infty} p(x, t) - p(x, 0) \right) dx = \underbrace{\lim_{t \rightarrow +\infty} \int_{\Omega_T^c} p(x, t) dx}_0 - \int_{\Omega_T^c} p(x, 0) dx = \\ &= - \int_{\Omega_T^c} p(x, 0) dx = - \int_{\Omega_T^c} \delta(x) dx = -1, \end{aligned} \quad (34)$$

where the IC  $p(x, 0) = \delta(x)$  is chosen to be deterministic. Solutions of the second-order ODE (34) will give the state distribution of the MFPT conditioned on the starting state  $x \in \Omega_0$ . Written explicitly the ODE reads

$$DT''(x) + (V'(x)T(x))' = -1, \quad (35)$$

where for the first time we introduce the potential function  $V(x)$  prescribing the dynamics of the particle through its (deterministic) drift term  $f(x) = -V'(x)$ . In order to solve (35) we can consider a 1-dimensional domain  $\Omega \equiv \mathbb{R}$  while we set  $\Omega_0 = \{-\infty < x < b\}$  with  $b$  being the (local) maximum of  $V(x)$  (i.e. the unstable equilibrium of  $f(x)$ ) and  $\Omega_T = \{b \leq x < +\infty\}$ . This choice entails that the local minimum  $x = a$  of the potential is found somewhere in  $\Omega_0$  which, being bounded from above by  $b$ , is the basin of attraction of  $a$ . As such we can specify the following BCs for our desired mean escape time

$$\begin{aligned} T(x = b) &= 0, \quad (\text{i.e. at the unstable equilibrium the escape occurs immediately}), \\ T'(x = b) &= 0, \quad (\text{the mean escape time at } x = b \text{ is a minima of the time distribution}). \end{aligned}$$

With the above we can integrate (35) between an arbitrary  $x \in \Omega_0$  and the unstable equilibrium  $b$  bounding the basin of attraction

$$\begin{aligned} \int_x^b (DT''(y) + (V'(y)T(y))') dy &= D(T(b) \overset{0}{\cancel{T}}(b) - T(x)) + (V'(b)T(b) \overset{0}{\cancel{T}}(b) - V'(x)T(x)) = -DT'(x) - V'(x)T(x) = \\ &= - \int_x^b 1 dx = -(b - x) \quad \Rightarrow \quad T'(x) + \frac{V'(x)}{D} T(x) = \frac{x - b}{D}, \end{aligned} \quad (36)$$

The first-order ODE in (36) can be solved explicitly using Theorem ???. The integrating factor is readily computed

$$I(x) = e^{\int \frac{V'(x)}{D} dx} = e^{\frac{V(x)}{D}}.$$

Multiplying both sides of Equation (36) by  $I(x)$  and integrating gives

$$\int (T(x)I(x))' dx = T(x)I(x) - C = \int \frac{x - b}{D} I(x) dx \quad \Rightarrow \quad T(x) = \frac{\int \frac{x - b}{D} e^{\frac{V(x)}{D}} dx + C}{e^{\frac{V(x)}{D}}}. \quad (37)$$

To derive the constant of integration  $C$  in Equation (37) we impose the aforementioned BC at  $x = b$

$$T(x = b) = 0 = \frac{\int_b^b \cancel{e}^{\frac{V(x)}{D}} dx + C}{e^{\frac{V(b)}{D}}} \quad \Rightarrow \quad C = 0 \quad \Rightarrow \quad T(x) = e^{-\frac{V(x)}{D}} \int \frac{x - b}{D} e^{\frac{V(x)}{D}} dx, \quad (38)$$

The closed form above gives the *MFPT* implicitly. To derive an explicit expression we can expand the potential in a Taylor series centered around its local maximum  $x = b$

$$\begin{aligned} V(x) &= V(b) + V'(b)(x - b) + \frac{V''(b)}{2}(x - b)^2 + \mathcal{O}(x^3), \\ e^{\frac{V(x)}{D}} &\approx e^{\frac{V(b)}{D}} \left( V(b) + \frac{V''(b)}{2}(x - b)^2 \right), \end{aligned}$$

which plugged into Equation (38) yields

$$\begin{aligned} T(x) &\approx e^{-\frac{V(x)}{D}} \left( \int \frac{x - b}{D} e^{\frac{V(b)}{D}} \left( V(b) + \frac{V''(b)}{2}(x - b)^2 \right) dx \right) = e^{-\frac{V(x)}{D}} \left( e^{\frac{V(b)}{D}} \int \frac{x - b}{D} e^{\frac{V''(b)}{2D}(x - b)^2} dx \right) = \\ &= e^{-\frac{V(x)}{D}} \left( \frac{e^{\frac{V(b)}{D}}}{V''(b)} \int \frac{d}{dx} \left( e^{\frac{V''(b)}{2D}(x - b)^2} \right) dx \right) = e^{-\frac{V(x)}{D}} \left( \frac{e^{\frac{V(b)}{D}}}{V''(b)} \left( e^{\frac{V''(b)}{2D}(x - b)^2} + C \right) \right) \\ &= \frac{e^{\frac{\Delta V(x)}{D}}}{V''(b)} \left( e^{\frac{V''(b)}{2D}(x - b)^2} + C \right) \stackrel{T(b)=0}{=} \frac{e^{\frac{\Delta V(x)}{D}}}{V''(b)} \left( e^{\frac{V''(b)}{2D}(x - b)^2} - 1 \right), \quad \Delta V(x) = V(b) - V(x). \end{aligned} \quad (39)$$

## Derivation from the (stationary and homogeneous) probability current

We start by rewriting the FPE (22) in terms of the probability density current  $J(x, t)$

$$\partial_t p(x, t) = D \partial_{xx}^2 p(x, t) + \partial_x(V'(x)p(x, t)) = \partial_x(D \partial_x p(x, t) + V'(x)p(x, t)) = \partial_x J(x, t). \quad (40)$$

We now proceed by rewriting the probability current itself

$$\begin{aligned} J(x, t) &= D \partial_x p(x, t) + V'(x)p(x, t) = D e^{-\frac{V(x)}{D}} \left( e^{\frac{V(x)}{D}} \partial_x p(x, t) + \underbrace{\frac{V'(x)}{D} e^{\frac{V(x)}{D}} p(x, t)}_{\partial_x e^{\frac{V(x)}{D}}} \right) = \\ &= D e^{-\frac{V(x)}{D}} \left( e^{\frac{V(x)}{D}} \partial_x p(x, t) + p(x, t) \partial_x e^{\frac{V(x)}{D}} \right) = D e^{-\frac{V(x)}{D}} \partial_x \left( e^{\frac{V(x)}{D}} p(x, t) \right), \end{aligned}$$

from which we get

$$\partial_x \left( e^{\frac{V(x)}{D}} p(x, t) \right) = D^{-1} J(x, t) e^{\frac{V(x)}{D}}. \quad (41)$$

To further develop the calculations we assume that  $x = a$  is a local minima of  $V(x)$  (i.e. a stable equilibrium for the vector field  $f(x)$ ) so that, integrating (41) between  $a$  and an arbitrary  $x > a$  yields

$$e^{\frac{V(x)}{D}} p(x, t) - e^{\frac{V(a)}{D}} p(a, t) = \frac{1}{D} \int_a^x J(y, t) e^{\frac{V(y)}{D}} dy. \quad (42)$$

From Equation (42) we can first assume stationarity of the stochastic process ( $\partial_t p(x, t) = 0 = \partial_x J(x, t) \Rightarrow J(x, t) = J$ ) which will thus evolve in a bounded region around the stable equilibrium  $x = a$ . As such if we define  $x = b > a$  as the unstable equilibrium, then for any  $x > b$ , we can reasonably assume that  $e^{\frac{V(x)}{D}} \gg p(x, t) \approx 0$  (see Figure 13a), and thus we retrieve

$$-e^{\frac{V(a)}{D}} p(a) = \frac{J}{D} \int_a^x e^{\frac{V(y)}{D}} dy \Rightarrow J = \frac{-D e^{\frac{V(a)}{D}} p(a)}{\int_a^x e^{\frac{V(y)}{D}} dy}, \quad \forall x \geq b. \quad (43)$$

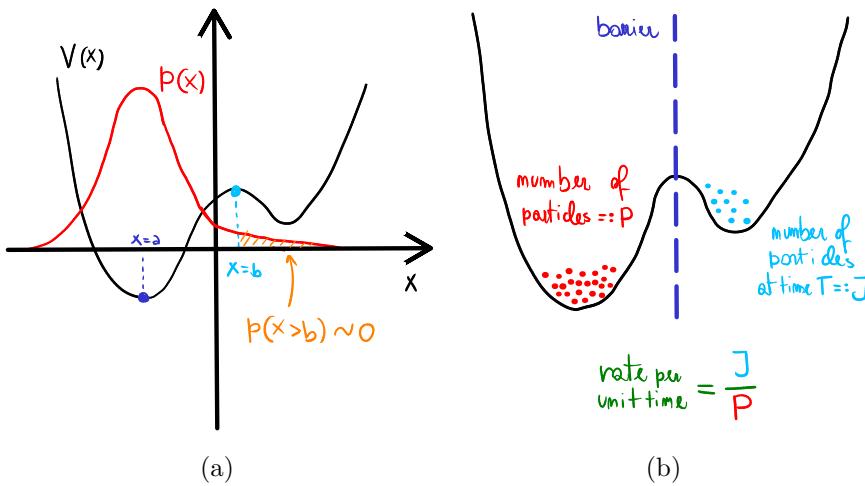


Figure 13

We can now switch our interpretation of the probability density current from a sample path perspective to an ensemble one (under the ergodic assumption). From this point the probability of the stochastic process'

realizations to be bounded in a region centered around the local minima  $x = a$  is the total number of an ensemble of the same family of processes that at time  $t$  are within such region. If we call such probability as  $P := \int_{a-\epsilon}^{a+\epsilon} p(x, t) dx$  and interpret  $J$  as the flux per unit time of trajectories in the ensemble that cross the potential barrier at  $x = b$  to escape (see Figure 13b), then we trivially derive the escape rate per unit  $R$  as

$$R = \frac{J}{P} = \frac{\frac{-De^{\frac{V(a)}{D}} p(a)}{\int_a^x e^{\frac{V(y)}{D}} dy}}{\int_{a-\epsilon}^{a+\epsilon} p(x) dx}. \quad (44)$$

Since we assumed stationarity for our process, we can rewrite the FP solution using Theorem ?? as  $p(x) = Ne^{-\frac{V(x)}{D}}$ . Furthermore in the *low-noise regime*  $D \ll \Delta V(x) = V(x) - V(a)$  we can obtain an approximation of the normalisation constant and thus rewrite the stationary distribution as  $p(x) = p(a)e^{-\frac{\Delta V(x)}{D}}$ . When using this assumption, the term  $P$  in Equation (44) becomes

$$P = \int_{a-\epsilon}^{a+\epsilon} p(x) dx = p(a)e^{\frac{V(a)}{D}} \int_{a-\epsilon}^{a+\epsilon} e^{-\frac{V(x)}{D}} dx. \quad (45)$$

In order to compute the integral in Equation (45) we expand the integrand in its Taylor series centered around  $x = a$

$$\begin{aligned} V(x) &= V(a) + \cancel{V'(a)(x-a)}^0 + \frac{V''(a)}{2}(x-a)^2 + \mathcal{O}(x^3), \\ e^{-\frac{V(x)}{D}} &\approx e^{-\frac{1}{D}(V(a) + \frac{V''(a)}{2}(x-a)^2)}, \end{aligned}$$

which gives us

$$\begin{aligned} P &\approx p(a)e^{\frac{V(a)}{D}} \int_{a-\epsilon}^{a+\epsilon} e^{-\frac{1}{D}(V(a) + \frac{V''(a)}{2D}(x-a)^2)} dx = p(a)e^{\cancel{\frac{V(a)}{D}}} \int_{a-\epsilon}^{a+\epsilon} e^{-\frac{V(a)}{D}} e^{\frac{V''(a)}{2D}(x-a)^2} dx = \\ &= p(a) \int_{a-\epsilon}^{a+\epsilon} e^{\frac{(x-a)^2}{\frac{2D}{V''(a)}}} dx = p(a) \sqrt{\frac{2\pi D}{V''(a)}}. \end{aligned} \quad (46)$$

Plugging Equation (46) in (44) results in

$$R = \frac{-De^{\frac{V(a)}{D}} p(a)}{p(a) \sqrt{\frac{2\pi D}{V''(a)}}} \frac{1}{\int_a^x e^{\frac{V(y)}{D}} dy} = \frac{-De^{\frac{V(a)}{D}} \sqrt{\frac{V''(a)}{2\pi D}}}{\int_a^x e^{\frac{V(y)}{D}} dy}, \quad (47)$$

and finally the integral at the denominator of Equation (47) can be approximated by repeating the same Taylor expansion of  $V(x)$  carried out above only this time centered around  $a < b < x$

$$R \approx \frac{-De^{\frac{V(a)}{D}} \sqrt{\frac{V''(a)}{2\pi D}}}{e^{\frac{V(b)}{D}} \sqrt{\frac{2\pi D}{V''(b)}}} = \frac{D}{2\pi} \sqrt{V''(a)V''(b)} e^{-\frac{\Delta V}{D}}. \quad (48)$$

where we defined  $\Delta V = V(b) - V(a)$  to be the height of the (potential) barrier.

## Derivation from the conditional solution of the FPE

This method is similar to the first one we derived at the beginning of this paragraph in the sense that it also treats the problem from the perspective of *first passage time*. The main difference in what follows is

the use of the conditional FPE and the imposition of absorbing BCs. As such we reformulate Equation (22) as an IBVP

$$\begin{cases} \partial_t p(x, t|x_0, t_0) = \mathcal{L}_{FPP}(x, t|x_0, t_0), & x \in \Omega \\ p(x, t_0|x_0, t_0) = \delta(x - x_0), & x \in \Omega \quad (\text{deterministic}) \text{ IC}, \\ p(x, t|x_0, t_0) = 0, & x \in \partial\Omega \quad (\text{absorbing}) \text{ BC}, \end{cases} \quad (49)$$

where, due to the absorbing BC we can deduce that

$$\lim_{t \rightarrow +\infty} p(x, t|x_0, t_0) = 0,$$

i.e. in the time asymptotic limit any trajectory starting inside  $\Omega$  would eventually hit its boundary  $\partial\Omega$  and get absorbed (escape). For a finite time  $t < \infty$  however, we can compute the probability of a stochastic process to stay bounded in  $\Omega$  by summing over all possible states  $x \in \Omega$

$$S(t|x_0, t_0) = \int_{\Omega} p(x, t|x_0, t_0) dx. \quad (50)$$

Notice that Equation (50) quantifies the same survival time distribution that we defined in Equation (25) with the only difference of being conditioned on the starting point of the process being  $x_0 \in \Omega$  at time  $t = t_0$ . Similarly to what we did in the first derivation, we can define such probability also in terms of the marginal of the distribution  $h(t|x_0, t_0)$  of the *hitting times*

$$S(t|x_0, t_0) = \int_t^{+\infty} h(\tau|x_0, t_0) d\tau, \quad (51)$$

which corresponds to Equation (28). The FTC for Equation (51) then gives

$$h(t|x_0, t_0) = -\frac{d}{dt} S(t|x_0, t_0), \quad (52)$$

from which, adopting the notation  $x_0 =: x$  and  $t_0 = 0$  for convenience, we compute the *MFPT* as done in Equation (32)

$$T(x) = - \int_0^{+\infty} t \frac{d}{dt} S(t|x, 0) dt = \int_0^{+\infty} S(t|x, 0) dt \stackrel{\text{Eq.(50)}}{=} \int_0^{+\infty} \left( \int_{\Omega} p(y, t|x, 0) dy \right) dt, \quad (53)$$

where in the integration by parts we implicitly used the property of the *pdf* vanishing in the time asymptotic limit due to the absorbing BCs we imposed in Equation (49). From here we depart from the first derivation where, in Equation (33), we switched the order of integration and defined a marginal distribution  $G(x)$  over the states across the history of the process. Instead we observe that, if the process is Markovian, then if we assume that there are infinitely many states  $x \in \Omega$  that the process can visit, in finite time  $t$ , from the initial state  $x(t_0) = x_0$  to the final state  $x(t_n) = x_n$ , we can write the probability of being at  $x_n$  conditioned that we start at  $x_0$  via the Chapman-Kolmogorov equation (see Figure 14)

$$p(x_n, t_n|x_0, t_0) = \int p(x_n, t_n|x, t) p(x, t|x_0, t_0) dx, \quad t_0 < t < t_n. \quad (54)$$

Differentiating Equation (54) w.r.t. time yields

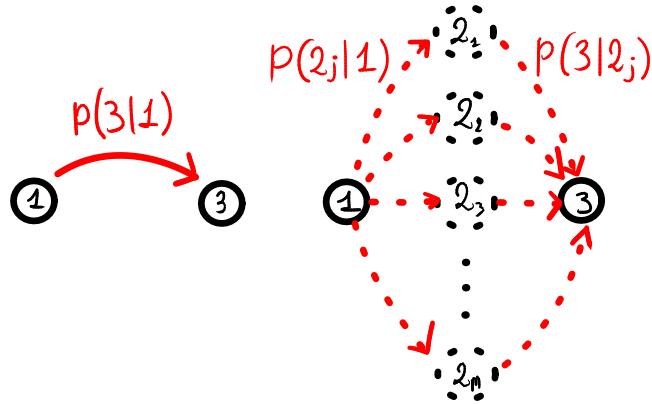
$$\begin{aligned} \partial_t p(x_n, t_n|x_0, t_0) &= 0 = \int \partial_t \left( p(x_n, t_n|x, t) p(x, t|x_0, t_0) \right) dx = \\ &= \int \left( p(x, t|x_0, t_0) \partial_t p(x_n, t_n|x, t) + p(x_n, t_n|x, t) \partial_t p(x, t|x_0, t_0) \right) dx \stackrel{\text{Eq.(49)}}{=} \\ &= \int \left( p(x, t|x_0, t_0) \partial_t p(x_n, t_n|x, t) + p(x_n, t_n|x, t) \mathcal{L}_{FPP}(x, t|x_0, t_0) \right) dx \stackrel{\text{Eq.(23)}}{=} \\ &= \int \left( p(x, t|x_0, t_0) \partial_t p(x_n, t_n|x, t) + p(x, t|x_0, t_0) \mathcal{L}_{FPP}^\dagger(x, t|x, t) \right) dx = \\ &= \int \left( \partial_t p(x_n, t_n|x, t) + \mathcal{L}_{FPP}^\dagger(x, t|x, t) \right) p(x, t|x_0, t_0) dx, \end{aligned}$$

from which we recover the adjoint FPE in Equation (23)

$$\partial_t p(x_n, t_n | x, t) + \mathcal{L}_{FP}^\dagger p(x_n, t_n | x, t) = 0.$$

Applying the adjoint FP operator to Equation (53) gives us an ODE for the *MFPT*

$$\begin{aligned} \mathcal{L}_{FP}^\dagger T(x) &= - \int_0^{+\infty} \left( \int_{\Omega} \partial_t p(y, t | x, 0) dx \right) dt = - \int_{\Omega} \left( \int_0^{+\infty} \partial_t p(y, t | x, 0) dt \right) dx = \\ &= \int_{\Omega} p(y, 0 | x, 0) dx \stackrel{\text{Eq.(49)}}{=} \int_{\Omega} \delta(y - x) dx = 1, \end{aligned} \quad (55)$$



$$p(3|1) = \sum_{j=1}^m p(2_j|1)p(3|2_j)$$

Figure 14

Written explicitly the ODE in Equation (55) reads

$$DT''(x) - V'(x)T'(x) = 1, \quad (56)$$

which we notice it is the adjoint of Equation (35). Similarly to what we did in the derivation from the stationary probability current, we multiply both sides of (56) by  $\frac{e^{-\frac{V(x)}{D}}}{D}$  and proceed by algebraic manipulation to get

$$\begin{aligned} \frac{e^{-\frac{V(x)}{D}}}{D} &= e^{-\frac{V(x)}{D}} \frac{d^2}{dx^2} T(x) - \frac{V'(x)}{D} e^{-\frac{V(x)}{D}} \frac{d}{dx} T(x) = e^{-\frac{V(x)}{D}} \frac{d^2}{dx^2} T(x) + \left( \frac{d}{dx} e^{-\frac{V(x)}{D}} \right) \left( \frac{d}{dx} T(x) \right) = \\ &= \frac{d}{dx} \left( e^{-\frac{V(x)}{D}} \frac{d}{dx} T(x) \right), \end{aligned} \quad (57)$$

We now assume that the potential landscape has a form as depicted in Figure 13a, i.e. it has a local minimum in  $x = a > -\infty$  and a local maximum at  $+\infty > x = b > a$ . As such, integrating (57) in  $[-\infty, x]$  yields

$$e^{-\frac{V(x)}{D}} \frac{d}{dx} T(x) - \lim_{x \rightarrow \infty} e^{-\frac{V(x)}{D}} \frac{d}{dx} T(x) = \frac{1}{D} \int_{-\infty}^x e^{-\frac{V(y)}{D}} dy,$$

and integrating again in  $[x, b]$  gives us an implicit expression for the *MFPT*

$$\mathcal{T}(b) - T(x) = \frac{1}{D} \int_x^b e^{\frac{V(y)}{D}} \left( \int_{-\infty}^y e^{-\frac{V(z)}{D}} dz \right) dy. \quad (58)$$

By recalling the Taylor expansion of the potential function around a stationary point  $x = c \in \{a, b\}$  we have

$$e^{\frac{V(y)}{D}} \approx e^{\frac{1}{D} \left( V(b) + \frac{V''(b)}{2} (x-b)^2 \right)}$$

$$e^{-\frac{V(z)}{D}} \approx e^{-\frac{1}{D} \left( V(a) + \frac{V''(a)}{2} (x-a)^2 \right)},$$

which, plugged in Equation (58), yields the closed form for the *MFPT*

$$\begin{aligned} -T(x) &\approx \frac{1}{D} \int_x^b e^{\frac{V(b)}{D}} e^{\frac{V''(b)}{2} (y-b)^2} \left( \int_{-\infty}^y e^{-\frac{V(a)}{D}} e^{-\frac{V''(a)}{2} (z-a)^2} dz \right) dy = \\ &= \frac{e^{\frac{V(b)-V(a)}{D}}}{D} \int_x^b e^{\frac{V''(b)}{2} (y-b)^2} \left( \int_{-\infty}^y e^{-\frac{V''(a)}{2} (z-a)^2} dz \right) dy \\ &= \frac{e^{\frac{\Delta V}{D}}}{D} \int_x^b e^{\frac{V''(b)}{2} (y-b)^2} \left( \int_{-\infty}^y e^{-\frac{V''(a)}{2} (z-a)^2} dz \right) dy. \end{aligned} \quad (59)$$

We can drop the dependence on the state  $x$  from Equation (59), which implicitly assumes that the escape time is compute for a stochastic process bounded in a small neighbour of the local minimum  $x = a$  of the potential function (i.e. the stable equilibrium of the vector field  $f(x) = -V'(x)$ ). In doing so we can substitute the bounds of integration in the outer integral  $x \mapsto -\infty$ ,  $b \mapsto +\infty$  and the upper bound in the inner integral  $y \mapsto +\infty$  thus getting

$$\begin{aligned} -T &= \frac{e^{\frac{\Delta V}{D}}}{D} \int_{-\infty}^{+\infty} e^{\frac{V''(b)}{2} (y-b)^2} \left( \int_{-\infty}^{+\infty} e^{-\frac{V''(a)}{2} (z-a)^2} dz \right) dy = \\ &= \frac{e^{\frac{\Delta V}{D}}}{D} \left( \int_{-\infty}^{+\infty} e^{-\left(-\frac{V''(b)}{2}\right)(y-b)^2} dy \right) \left( \int_{-\infty}^{+\infty} e^{-\frac{V''(a)}{2} (z-a)^2} dz \right) = \\ &= \frac{e^{\frac{\Delta V}{D}}}{D} \sqrt{\frac{2\pi}{-V''(b)}} \sqrt{\frac{2\pi}{V''(a)}} = \frac{2\pi}{D} \left( \sqrt{V''(a)|V''(b)|} \right)^{-1} e^{\frac{\Delta V}{D}}, \end{aligned} \quad (60)$$

where we can substitute  $-V''(b) \mapsto |V''(b)|$  given that  $V''(b) < 0$  ( $x = b$  is a local maximum for  $V(x)$ ). By taking the inverse of (60) we recover Kramers' escape rate given by Equation (48)

$$R = -\frac{D}{2\pi} \sqrt{V''(a)|V''(b)|} e^{-\frac{\Delta V}{D}}.$$

## 4.2 Learning the potential function from data

When an SDE evolves around a stable equilibrium of the vector field the stochastic diffusion dominates the dynamics over the drift. The corresponding solution of the FPE will be an equilibrium probability density function  $p(x, t)$  which is stationary i.e.  $p(x, t) = p(x)$ . We can exploit this to reconstruct a local approximation of the scalar potential function that generates the data of the solution of the unknown SDE.

#### 4.2.1 Inversion of the equilibrium distribution

Solutions of the stationary FPE [85, Theorem 3.1, p. 3] can be used to reconstruct the potential function. Assuming the appropriate BCs for the FPE, we can write the equilibrium distribution as

$$p(x) = N e^{-\frac{V(x)}{D}}, \quad (61)$$

where  $V(x)$  is the potential function of the drift term of the SDE and  $D = \frac{\sigma^2}{2}$  is the diffusion coefficient. A simple inversion gives us

$$V(x) = -D \ln \left( \frac{p(x)}{N} \right) = D \ln N - D \ln p(x), \quad (62)$$

which tells us that if we have access to the (approximate) equilibrium distribution  $p(x)$  and we know the level of noise in the system  $D$  then we can recover the shape of the potential function  $V(x)$  up to an arbitrary shift given by  $D \ln N$ , where  $N > 0$  is a normalisation constant for  $p$ .

#### Example 4.1

Let us consider a saddle-node normal form for the drift of a SDE with additive noise

$$dx = -(\mu + x^2)dt + \sigma dW.$$

The stationary FPE equation reads

$$\left( p' + \frac{1}{D} (\mu + x^2)p \right)' = 0.$$

We follow the procedure outlined in proof of [85, Theorem 3.2, pp. 4-6] to solve the second-order IVP and specify that  $p(x), p'(x) \rightarrow 0$  as  $x \rightarrow +\infty$ . Integrating the above in  $[x, +\infty)$  yields

$$p'(x) + \frac{\mu + x^2}{D} p(x) = 0.$$

The integrating factor is

$$h(x) = e^{\frac{1}{D} \int (\mu + x^2) dx} = e^{\frac{\mu x + \frac{x^3}{3}}{D}},$$

which, when multiplied to both sides of the first-order, homogeneous ODE, gives us

$$\left( e^{\frac{\mu x + \frac{x^3}{3}}{D}} p(x) \right)' = 0 \Rightarrow p(x) = N e^{-\frac{\mu x + \frac{x^3}{3}}{D}},$$

where  $N > 0$  is a normalisation constant. By definition

$$\int_{-\infty}^{+\infty} p(x) dx = 1 = N \int_{-\infty}^{+\infty} e^{-\frac{\mu x + \frac{x^3}{3}}{D}} dx \Rightarrow N = \frac{1}{\int_{-\infty}^{+\infty} e^{-\frac{\mu x + \frac{x^3}{3}}{D}} dx},$$

which gives us

$$p(x) = \left( \int_{-\infty}^{+\infty} e^{-\frac{\mu x + \frac{x^3}{3}}{D}} dx \right)^{-1} e^{-\frac{\mu x + \frac{x^3}{3}}{D}}.$$

Notice that since  $e^{-\frac{\mu x + \frac{x^3}{3}}{D}} \rightarrow +\infty$  as  $x \rightarrow -\infty$  then  $N = \infty$  and as such we conclude that the stationary solution of the FPE for the saddle-node normal form does not exist over the domain  $(-\infty, +\infty)$ . In order to get a normalisable solution that exists we need to truncate the domain to  $[a, +\infty)$ . When we do that we impose reflecting (i.e. zero-flux) BCs s.t.  $p'(a) = 0, p(a) \neq 0$ . We choose the lower boundary of the domain to be the unstable equilibrium of the saddle-node normal form, i.e.  $a = -\sqrt{-\mu} =: x_u$ , which obviously entails that  $V'(x_u) = 0$  and  $V''(x_u) < 0$ . We thus retrieve a stationary solution of the form

$$p(x) = \left( \int_{x_u}^{+\infty} e^{-\frac{\mu x + \frac{x^3}{3}}{D}} dx \right)^{-1} e^{-\frac{\mu x + \frac{x^3}{3}}{D}}.$$

## The problem with the normalisation constant

As per the example above, we incur in the issue of computing the normalisation constant with a higher-than-second-degree polynomial as argument of the exponential, which is rarely solvable in explicit terms. In the more general sense when we plug the explicit expression for the normalisation constant of a generic potential  $V(x)$  back into (62) we get a trascendental equation of the form

$$V(x) = -D \ln \left( p(x) \int_{-\infty}^{+\infty} e^{-\frac{V(x)}{D}} dx \right). \quad (63)$$

The question now thus becomes, under what further assumptions regarding  $V(x)$  we can compute the normalisation constant analytically so that (63) becomes algebraic?

**Linearised approach: topological equivalence with an OUP** When the solution  $x_t$  of a non-linear SDE  $dx = f(x, \mu)dt + \sigma dW$  with additive noise  $\sigma$  settles onto a stable equilibrium  $x_s$  of  $f(x, \mu)$  then the system is topologically equivalent to the stationary solution of the OUP equation

$$dx = -\theta(x - x_s)dt + \sigma dW.$$

This is easy to see by linearising the dynamics onto the equilibrium

$$f(x) = \sum_{n=0}^{\infty} \frac{(D^n f)(x_s)}{n!} (x - x_s)^n = \cancel{f(x_s)}^0 + f'(x_s)(x - x_s) + O(x^3) \approx f'(x_s)(x - x_s),$$

meaning that the parameter of the topologically equivalent OUP is  $\theta = -f'(x_s)$ . We can use this equivalence to compute the normalisation constant; we know in fact [85, Theorem 3.2, pp. 4-6] that if  $x_t \sim \mathcal{N}(x_s, \frac{\sigma^2}{2\theta})$ , in the time-asymptotic limit  $t \rightarrow \infty$ , then normalisation constant is  $\sqrt{\frac{\theta}{\sigma^2 \pi}}$ . This assumption however necessarily implies that  $V(x) \sim -\theta(\frac{x^2}{2} - x_s x)$  which, unsurprisingly, has a unique global minima at  $x_s$  if  $\theta > 0$ . Suppose now we are only presented with the solution  $x_t$  in the form of a (weakly) stationary timeseries, knowing anything about the SDE that generated the process. Under the linearised approach and with the OUP assumption we can “learn” the parameters of the process via linear least-squares regression. From the timeseries data we can construct a histogram that approximates the stationary solution  $p(x)$  of a FPE (see Figure 15).

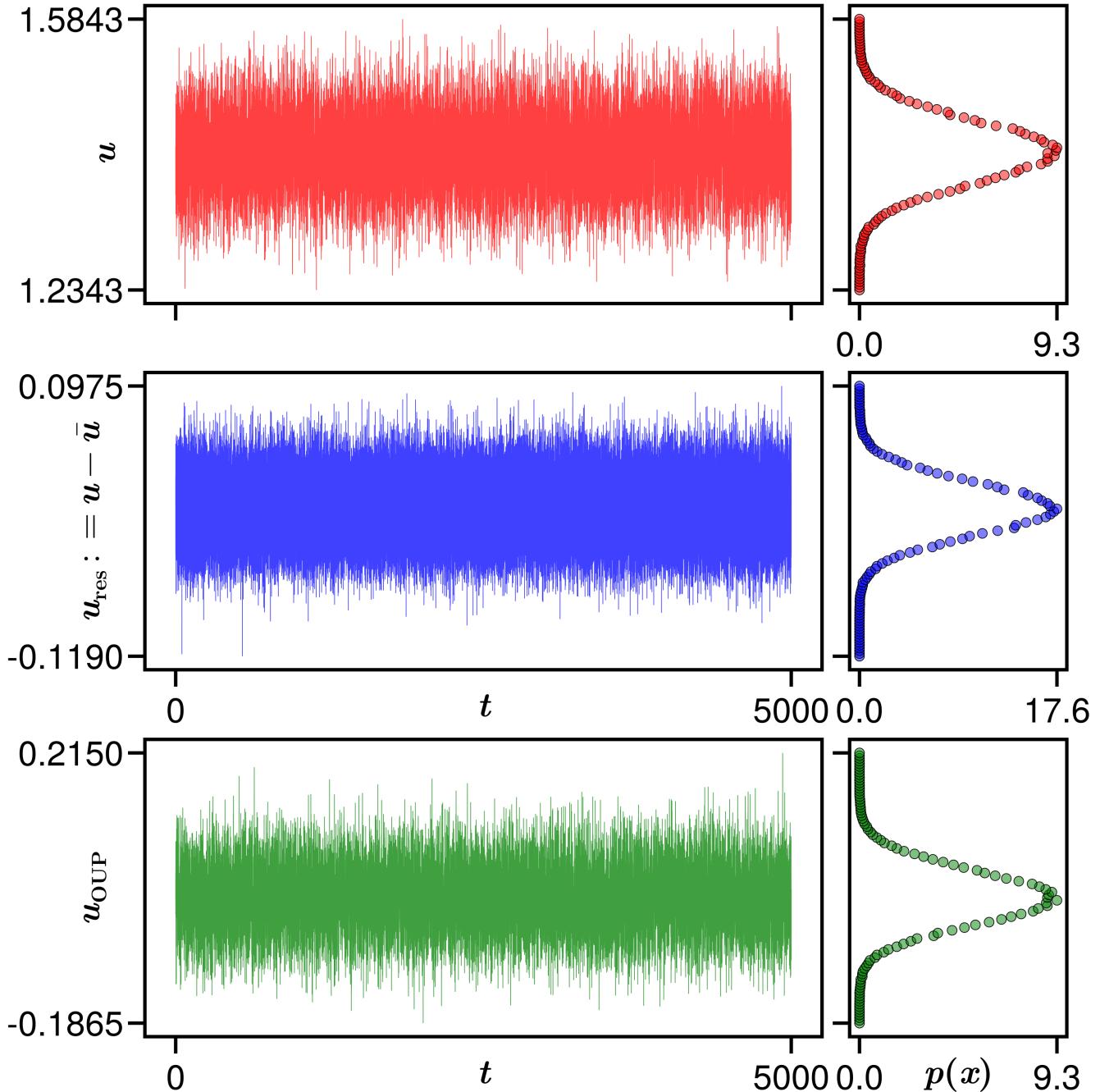


Figure 15: Timeseries and histograms of a saddle-node normal form (top, red) at fixed parameter value  $\mu = -2$ , its detrended residuals (middle, blue) and a topologically equivalent OUP (bottom, green) with  $\theta \approx 2.82$  and  $x_s \approx 1.414$ . All 3 systems have been simulated for  $n = 10^5$  realisations with (additive) noise level  $\sigma = 0.1$ .

The values  $y_n =: p(x_n)$  of this histogram can be put into (63), which is now algebraic under the OUP assumption for the normalisation constant

$$V(x) = -D \ln \left( \sqrt{\frac{\theta}{\sigma^2 \pi}} p(x) \right), \quad (64)$$

and the linear least-squares problem can thus be formulated

$$\text{find } \{c_0, c_1, c_2\} \text{ s.t. } \sqrt{\sum_{n=1}^{N_{\text{data}}} (U(x_n) - V_n)^2} \text{ is minimal,} \quad (65)$$

$$V(x_n) =: V_n = -D \ln \left( \sqrt{\frac{\theta}{\sigma^2 \pi}} y_n \right), \quad U(x) = \sum_{m=0}^2 c_m x^m.$$

**Generic cubic: topological equivalence with a saddle-node normal form** For the purposes of EWS of incoming tipping-points we need to bound the basin of attraction of  $x_s$  which necessarily entails the existence of at least 1 unstable equilibrium  $x_u$ . For that to be true we conclude that in order to derive the probabilistic EWS based on the escape problem from a potential well, the minimum working model is a (generic) cubic exponential

$$U(x) = \sum_{m=0}^3 c_m x^m = c_0 + c_1 x + c_2 x^2 + c_3 x^3,$$

$$U'(x) = \sum_{m=1}^3 m c_m x^{m-1} = c_1 + 2 c_2 x + 3 c_3 x^2,$$

$$U''(x) = \sum_{m=2}^3 m(m-1) c_m x^{m-2} = 2 c_2 + 6 c_3 x.$$

As established in Example 4.1 in order for the stationary FPE to exist in a cubic potential we need to truncate the domain at the unstable equilibrium  $x_u$  and thereby impose reflecting BCs. This already provides us with a constraint on the coefficients of the cubic potential since we want it to have a local minima (i.e. the stable equilibrium  $x_s$ ) and a local maxima (i.e. the unstable equilibrium  $x_u$ ). This translates in imposing that  $U'(x) = 0$ , which is a quadratic polynomial, has to have exactly 2 real and distinct roots  $x_{1,2}$  which provides

$$x_{1,2} = \frac{-2c_2 \pm \sqrt{\Delta}}{6c_3}, \quad \Delta = 4(c_2^2 - 3c_3c_1) > 0 \iff -\sqrt{3c_3c_1} > c_2 > \sqrt{3c_3c_1}. \quad (66)$$

Furthermore we can find closed-form expressions for both equilibria in terms of the unknown cubic coefficients since  $U''(x_s) > 0$  and  $U''(x_u) < 0$  leading us to

$$x_{s,u} = \pm \frac{1}{3c_3} \left( \sqrt{c_2^2 - 3c_3c_1} \mp c_2 \right). \quad (67)$$

Finally in order to choose the appropriate domain of integration we must be able to discern the case  $x_s > x_u$ , in which case the integration bounds are  $[x_u, +\infty)$ , from  $x_s < x_u$ , which yields the integration domain  $(-\infty, x_u]$  instead. By using the closed-form expressions (67) we obtain

$$x_s > x_u \iff \frac{2}{3c_3} \left( \sqrt{c_2^2 - 3c_3c_1} \right) > 0 \iff c_3 > 0, \quad (68)$$

as expected. We now use conditions (67)-(68) to write down the stationary solution of the FPE (62) explicitly

$$p(x) = \left( \int_{-\frac{1}{3c_3} \left( \sqrt{c_2^2 - 3c_3c_1} + c_2 \right)}^{+\infty} e^{-\frac{U(x)}{D}} dx \right)^{-1} e^{-\frac{U(x)}{D}}. \quad (69)$$

Notice that (69) is again transcendental given the fact that the improper integral does not have a closed-form expression for a generic cubic potential  $U(x) = \sum_{m=0}^3 c_m x^m$  and it has to be approximated either

numerically by quadrature or other methods. Now again we suppose that we have access to some timeseries data which we think of as the solution  $x_t$  of a scalar, non-linear SDE with additive noise. We want to quantify the probability of escape outside the basin of attraction of the stable equilibrium of the underlying SDE. In order to do that we assume a cubic potential as a minimum model for our system. Again we use the timeseries data to construct a histogram whose values  $y_n$  correspond to an approximation of the stationary solution  $p(x_n)$  of the FPE (63). To quantify the coefficients  $\{c_m\}_{m=0,\dots,3}$  we use (69), together with condition (66), to formulate the following constrained, non-linear optimisation problem

$$\text{find } \{c_0, c_1, c_2, c_3\} \text{ s.t. } \sqrt{\sum_{n=1}^{N_{\text{data}}} (p(x_n) - y_n)^2} \text{ is minimal,} \quad (70)$$

$$p(x_n) =: p_n = \left( \int_{-\frac{1}{3c_3}(\sqrt{c_2^2 - 3c_3c_1} + c_2)}^{+\infty} e^{-\frac{U(x)}{D}} dx \right)^{-1} e^{-\frac{U(x_n)}{D}}, \quad -\sqrt{3c_3c_1} > c_2 > \sqrt{3c_3c_1}.$$

**Laplace's method** This method provides an analytical approximation of definite integrals of bounded exponential functions the form

$$\int_a^b e^{Mf(x)} dx \approx \sqrt{\frac{2\pi}{M|f''(x_0)|}} e^{Mf''(x_0)}, \quad M \rightarrow +\infty, \quad (71)$$

where  $M > 0$  is large and  $f$  is at least twice-differentiable with a global maximum  $x_0$  in  $[a, b]$ . Under these assumption the major contribution to the definite integral comes from a small region centered on  $x_0$  as the exponential quickly decays to 0 at  $\pm\infty$  if  $M \rightarrow +\infty$ . This implies that one can expand  $f$  in its Taylor series around the global maximum  $x_0$  and truncate it to the second order. Extending the bounds of integration from  $[a, b]$  to  $(-\infty, +\infty)$  will thus have a negligible error in the computation of (71) and thus one can readily use the formula for a Gaussian integral to derive the approximate result. This method effectively fits a Gaussian function of the form  $e^{Mx^2}$  which, for our purposes, essentially boils down on fitting a stationary OUP to our histogram. It should not come as a surprise in fact that the normalisation constant in (71) coincides with the one we computed for the stationary OUP above with the scaling  $M = \frac{2\theta}{\sigma^2}$ . Notice how the approximate formula in (71) does not depend on either the bounds of integration but only on the global maximum of  $f$ . Applying such method to compute the normalisation constant in (70) thus yields

$$\text{find } \{c_0, c_1, c_2, c_3\} \text{ s.t. } \sqrt{\sum_{n=1}^{N_{\text{data}}} (p(x_n) - y_n)^2} \text{ is minimal,} \quad (72)$$

$$p(x_n) =: p_n = \left( \sqrt{\frac{2\pi D}{c_2^2 - 3c_3c_1}} e^{-\frac{U\left(\frac{1}{3c_3}(\sqrt{c_2^2 - 3c_3c_1} - c_2)\right)}{D}} \right) e^{-\frac{U(x_n)}{D}}, \quad -\sqrt{3c_3c_1} > c_2 > \sqrt{3c_3c_1}.$$

which does not depend on  $x_u$  but rather on  $x_s = \frac{1}{3c_3}(\sqrt{c_2^2 - 3c_3c_1} - c_2)$  which is taken as a local maximum in  $[x_u, +\infty)$ . Notice that for the method to be accurate  $M \rightarrow +\infty$  we must be in the low-noise regime  $\sigma \rightarrow 0$ .

## References

- [1] May, R. “Thresholds and breakpoints in ecosystems with a multiplicity of stable states”. *Nature*, 269 (1977), pp. 471–477. [DOI](#).
- [2] Zahler, R. and Sussmann, H. “Claims and accomplishments of applied catastrophe theory”. *Nature*, 269 (1977), pp. 759–763. [DOI](#).

- [3] Wissel, C. “A universal law of the characteristic return time near thresholds”. *Oecologia*, 65 (1984), pp. 101–107. [DOI](#).
- [4] Wiesenfeld, K. and McNamara, B. “Small-signal amplification in bifurcating dynamical systems”. *Physical Review A*, 33 (1 1986), pp. 629–642. [DOI](#).
- [5] Ives, A. R. “Measuring resilience in stochastic systems”. *Ecological monographs*, 65 (1995), pp. 217–233. [DOI](#).
- [6] Gandhi, A., Levin, S., and Orszag, S. “Critical slowing down in time-to-extinction: an example of critical phenomena in ecology”. *Journal of theoretical biology*, 192 (1998), pp. 363–376. [DOI](#).
- [7] Vandewalle, N., Boveroux, P., Minguet, A., and Ausloos, M. “The crash of October 1987 seen as a phase transition: amplitude and universality”. *Physica A: statistical mechanics and its applications*, 255 (1998), pp. 201–210. [DOI](#).
- [8] Petit, J. et al. “Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica”. *Nature*, 399 (1999), pp. 429–436. [DOI](#).
- [9] deMenocal, P., Ortiz, J., Guilderson, T., Adkins, J., Sarnthein, M., Baker, L., and Yarusinsky, M. “Abrupt onset and termination of the African Humid Period: rapid climate responses to gradual insolation forcing”. *Quaternary science reviews*, 19 (2000), pp. 347–361. [DOI](#).
- [10] Litt, B. et al. “Epileptic seizures may begin hours in advance of clinical onset: a report of five patients”. *Neuron*, 30 (2001), pp. 51–64. [DOI](#).
- [11] Scheffer, M., Carpenter, S., Foley, J., Folke, C., and Walker, B. “Catastrophic shifts in ecosystems”. *Nature*, 413 (2001), pp. 591–596. [DOI](#).
- [12] Kleinen, T., Held, H., and Petschel-Held, G. “The potential role of spectral properties in detecting thresholds in the Earth system: application to the thermohaline circulation”. *Ocean dynamics*, 53 (2003), pp. 53–63. [DOI](#).
- [13] McSharry, P., Smith, L., and Tarassenko, L. “Prediction of epileptic seizures: are nonlinear methods relevant?” *Nature medicine*, 9 (2003), pp. 241–242. [DOI](#).
- [14] Held, H. and Kleinen, T. “Detection of climate system bifurcations by degenerate fingerprinting”. *Geophysical research letters*, 31 (2004). [DOI](#).
- [15] Rietkerk, M., Dekker, S. C., Ruiter, P. C. de, and Koppell, J. van de. “Self-organized patchiness and catastrophic shifts in ecosystems”. *Science*, 305 (2004), pp. 1926–1929. [DOI](#).
- [16] Oborny, G. M. and György Szabó, B. “Dynamics of populations on the verge of extinction”. *Oikos*, 109 (2005), pp. 291–296. [DOI](#).
- [17] Tripati, A., Backman, J., Elderfield, H., and Ferretti, P. “Eocene bipolar glaciation associated with global carbon cycle changes”. *Nature*, 436 (2005), pp. 341–346. [DOI](#).
- [18] Venegas, J., Winkler, T., Musch, G., Vidal Melo, M., Layfield, D., Tgavalekos, N., Fischman, A., Callahan, R., Bellani, G., and Harris, S. “Self-organized patchiness in asthma as a prelude to catastrophic shifts”. *Nature*, 434 (2005), pp. 777–782. [DOI](#).
- [19] Brock, W. A. “Chapter 3: tipping points, abrupt opinion changes and punctuated policy change”. *Punctuated equilibrium and the dynamics of U.S. environmental policy*. Ed. by R. Repetto. New Haven: Yale University Press, 2006, pp. 47–77. [DOI](#).
- [20] Carpenter, S. R. and Brock, W. A. “Rising variance: a leading indicator of ecological transition”. *Ecology letters*, 9 (2006), pp. 311–318. [DOI](#).
- [21] Kéfi, S., Rietkerk, M., Alados, C., Pueyo, Y., Papanastasis, V., ElAich, A., and Ruiter, P. de. “Spatial vegetation patterns and imminent desertification in Mediterranean arid ecosystems”. *Nature*, 449 (2007), pp. 213–217. [DOI](#).

- [22] Livina, V. N. and Lenton, T. M. “A modified method for detecting incipient bifurcations in a dynamical system”. *Geophysical research letters*, 34 (2007). [DOI](#).
- [23] van Nes, E. H. and Scheffer, M. “Slow recovery from perturbations as a generic indicator of a nearby catastrophic shift.” *The American naturalist*, 169 (2007), pp. 738–747. [DOI](#).
- [24] Carpenter, S. R., Brock, W. A., Cole, J. J., Kitchell, J. F., and Pace, M. L. “Leading indicators of trophic cascades”. *Ecology letters*, 11 (2008), pp. 128–138. [DOI](#).
- [25] Dakos, V., Scheffer, M., van Nes, E. H., Brovkin, V., Petoukhov, V., and Held, H. “Slowing down as an early warning signal for abrupt climate change”. *Proceedings of the National Academy of Sciences*, 105 (2008), pp. 14308–14312. [DOI](#).
- [26] Guttal, V. and Jayaprakash, C. “Changing skewness: an early warning signal of regime shifts in ecosystems”. *Ecology letters*, 11 (2008), pp. 450–460. [DOI](#).
- [27] Biggs, R., Carpenter, S. R., and Brock, W. A. “Turning back from the brink: detecting an impending regime shift in time to avert it”. *Proceedings of the National Academy of Sciences*, 106 (2009), pp. 826–831. [DOI](#).
- [28] Chisholm, R. A. and Filotas, E. “Critical slowing down as an indicator of transitions in two-species models”. *Journal of theoretical biology*, 257 (2009), pp. 142–149. [DOI](#).
- [29] Guttal, V. and Jayaprakash, C. “Spatial variance and spatial skewness: leading indicators of regime shifts in spatial ecological systems”. *Theoretical ecology*, 2 (2009), pp. 3–12. [DOI](#).
- [30] Scheffer, M. *Critical transitions in nature and society*. Princeton University Press, 2009.
- [31] Scheffer, M., Bascompte, J., Brock, W., Brovkin, V., Carpenter, S., Dakos, V., Held, H., van Nes, E., Rietkerk, M., and Sugihara, G. “Early-warning signals for critical transitions”. *Nature*, 461 (2009). [DOI](#).
- [32] Brock, W. A. and Carpenter, S. R. “Interacting regime shifts in ecosystems: implication for early warnings”. *Ecological monographs*, 80 (2010), pp. 353–367. [DOI](#).
- [33] Carpenter, S. R. and Brock, W. A. “Early warnings of regime shifts in spatial dynamics using the discrete Fourier transform”. *Ecosphere*, 1 (2010). [DOI](#).
- [34] Dakos, V., van Nes, E., Donangelo, R., Fort, H., and Scheffer, M. “Spatial correlation as leading indicator of catastrophic shifts”. *Theoretical ecology*, 3 (2010), pp. 163–174. [DOI](#).
- [35] Ditlevsen, P. D. and Johnsen, S. J. “Tipping points: early warning and wishful thinking”. *Geophysical research letters*, 37 (2010). [DOI](#).
- [36] Drake, J. and Griffen, B. “Early warning signals of extinction in deteriorating environments”. *Nature*, 467 (2010), pp. 456–459. [DOI](#).
- [37] Hastings, A. and Wysham, D. B. “Regime shifts in ecological systems can occur with no warning”. *Ecology letters*, 13 (2010), pp. 464–472. [DOI](#).
- [38] Livina, V. N., Kwasniok, F., and Lenton, T. M. “Potential analysis reveals changing number of climate states during the last 60 kyr”. *Climate of the past*, 6 (2010), pp. 77–82. [DOI](#).
- [39] Carpenter, S. R. and Brock, W. A. “Early warnings of unknown nonlinear shifts: a nonparametric approach”. *Ecology*, 92 (2011), pp. 2196–2201. [DOI](#).
- [40] Carpenter, S. R. et al. “Early warnings of regime shifts: a whole-ecosystem experiment”. *Science*, 332 (2011), pp. 1079–1082. [DOI](#).
- [41] Dakos, V., Kéfi, S., Rietkerk, M., van Nes, E. H., and Scheffer, M. “Slowing down in spatially patterned ecosystems at the brink of collapse”. *The American naturalist*, 177 (2011), pp. 153–166. [DOI](#).

- [42] Kuehn, C. “A mathematical framework for critical transitions: bifurcations, fast–slow systems and stochastic dynamics”. *Physica D: nonlinear phenomena*, 240 (2011), pp. 1020–1035. [DOI](#).
- [43] Lenton, T. “Early warning of climate tipping points”. *Nature climate change*, 1 (2011), pp. 201–209. [DOI](#).
- [44] Wieczorek, S., Ashwin, P., Luke, C. M., and Cox, P. M. “Excitability in ramped systems: the compost-bomb instability”. *Proceedings of the Royal Society A*, 467 (2011), pp. 2733–2733. [DOI](#).
- [45] Ashwin, P., Wieczorek, S., Vitolo, R., and Cox, P. “Tipping points in open systems: bifurcation, noise-induced and rate-dependent examples in the climate system”. *Philosophical Transactions of the Royal Society A*, 370 (2012), pp. 1166–1184. [DOI](#).
- [46] Dai, L., Vorselen, D., Korolev, K. S., and Gore, J. “Generic indicators for loss of resilience before a tipping point leading to population collapse”. *Science*, 336 (2012), pp. 1175–1177. [DOI](#).
- [47] Dakos, V. et al. “Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data”. *PLOS ONE*, 7 (2012), pp. 1–20. [DOI](#).
- [48] Scheffer, M. et al. “Anticipating critical transitions”. *Science*, 338 (2012), pp. 344–348. [DOI](#).
- [49] Wang, R., Dearing, J., and Langdon, P. “Flickering gives early warning signals of a critical transition to a eutrophic lake state”. *Nature*, 492 (2012), pp. 419–422. [DOI](#).
- [50] Boerlijst, M. C., Oudman, T., and Roos, A. M. de. “Catastrophic collapse can occur without early warning: examples of silent catastrophes in structured ecological models”. *PLOS ONE*, 8 (2013), pp. 1–6. [DOI](#).
- [51] Dakos, V., van Nes, E., and Scheffer, M. “Flickering as an early warning signal”. *Theoretical ecology*, 6 (2013), pp. 309–317. [DOI](#).
- [52] Kéfi, S., Dakos, V., Scheffer, M., van Nes, E. H., and Rietkerk, M. “Early warning signals also precede non-catastrophic transitions”. *Oikos*, 122 (2013), pp. 641–648. [DOI](#).
- [53] Kuehn, C. “Warning signs for wave speed transitions of noisy Fisher–KPP invasion fronts”. *Theoretical ecology*, 6 (2013), pp. 295–308. [DOI](#).
- [54] Bos, E. H. and Jonge, P. D. ““Critical slowing down in depression” is a great idea that still needs empirical proof”. *Proceedings of the National Academy of Sciences*, 111 (2014), pp. 878–878. [DOI](#).
- [55] Cline, T. J., Seekell, D. A., Carpenter, S. R., Pace, M. L., Hodgson, J. R., Kitchell, J. F., and Weidel, B. C. “Early warnings of regime shifts: evaluation of spatial indicators from a whole-ecosystem experiment”. *Ecosphere*, 5 (2014). [DOI](#).
- [56] Kéfi, S., Guttal, V., Brock, W. A., Carpenter, S. R., Ellison, A. M., Livina, V. N., Seekell, D. A., Scheffer, M., van Nes, E. H., and Dakos, V. “Early warning signals of ecological transitions: methods for spatial patterns”. *PLOS ONE*, 9 (2014), pp. 1–13. [DOI](#).
- [57] Leemput, I. A. van de et al. “Critical slowing down as early warning for the onset and termination of depression”. *Proceedings of the National Academy of Sciences*, 111 (2014), pp. 87–92. [DOI](#).
- [58] Suweis, S. and D’Odorico, P. “Early warning signs in social-ecological networks”. *PLOS ONE*, 9 (2014), pp. 1–6. [DOI](#).
- [59] van de Leemput, I. A., van Nes, E. H., and Scheffer, M. “Resilience of alternative states in spatially extended ecosystems”. *PLOS ONE*, 10 (Feb. 2015), pp. 1–17. [DOI](#).
- [60] Donovan, G. and Kritter, T. “Spatial pattern formation in the lung”. *Journal of mathematical biology*, 70 (2015), pp. 1119–1149. [DOI](#).
- [61] Scheffer, M., Carpenter, S. R., Dakos, V., and van Nes, E. H. “Generic indicators of ecological resilience: inferring the chance of a critical transition”. *Annual review of ecology, evolution, and systematics*, 46 (2015), pp. 145–167. [DOI](#).

- [62] Strogatz, S. *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering* (2nd ed.) CRC Press, 2015. [DOI](#).
- [63] Kozłowska, M., Denys, M., Wiliński, M., Link, G., Gubiec, T., Werner, T., Kutner, R., and Struzik, Z. “Dynamic bifurcations on financial markets”. *Chaos, solitons & fractals*, 88 (2016), pp. 126–142. [DOI](#).
- [64] Ritchie, P. and Sieber, J. “Early-warning indicators for rate-induced tipping”. *Chaos: an interdisciplinary journal of nonlinear science*, 26 (2016). [DOI](#).
- [65] Dmitriev, A., Dmitriev, V., Sagaydak, O., and Tsukanova, O. “The application of stochastic bifurcation theory to the early detection of economic bubbles”. *Procedia computer science*, 122 (2017). 5th International Conference on Information Technology and Quantitative Management, ITQM 2017, pp. 354–361. [DOI](#).
- [66] Liang, J., Hu, Y., Chen, G., and Zhou, T. “A universal indicator of critical state transitions in noisy complex networked systems”. *Scientific reports*, 7 (2017). [DOI](#).
- [67] Jentsch, P. C., Anand, M., and Bauch, C. T. “Spatial correlation as an early warning signal of regime shifts in a multiplex disease-behaviour network”. *Journal of theoretical biology*, 448 (2018), pp. 17–25. [DOI](#).
- [68] Chen, S., O’Dea, E. B., Drake, J. M., and Epureanu, B. I. “Eigenvalues of the covariance matrix as early warning signals for critical transitions in ecological systems”. *Scientific reports*, 9 (2019). [DOI](#).
- [69] Diks, C., Hommes, C., and Wang, J. “Critical slowing down as an early warning signal for financial crises?” *Empirical economics*, 57 (2019), pp. 1201–1228. [DOI](#).
- [70] Pages, N., Vera-Sigüenza, E., Rugis, J., Kirk, V., Yule, D., and Sneyd, J. “A model of  $Ca^{2+}$  dynamics in an accurate reconstruction of parotid acinar cells”. *Bulletin of mathematical biology*, 81 (2019), pp. 1394–1426. [DOI](#).
- [71] de Mooij, S. M., Blanken, T. F., Grasman, R. P., Ramautar, J. R., Van Someren, E. J., and van der Maas, H. L. “Dynamics of sleep: exploring critical transitions and early warning signals”. *Computer methods and programs in biomedicine*, 193 (2020), p. 105448. [DOI](#).
- [72] Gottwald, G. and Gugole, F. “Detecting regime transitions in time series using dynamic mode decomposition”. *Journal of statistical physics*, 179 (2020), pp. 1028–1045. [DOI](#).
- [73] Lever, J. J., Leemput, I. A. van de, Weinans, E., Quax, R., Dakos, V., Nes, E. H. van, Bascompte, J., and Scheffer, M. “Foreseeing the future of mutualistic communities beyond collapse”. *Ecology letters*, 23 (2020), pp. 2–15. [DOI](#).
- [74] Wechselberger, M. *Geometric singular perturbation theory beyond the standard form*. Springer, 2020. [DOI](#).
- [75] Xing, K. and Yang, X. “Predicting default rates by capturing critical transitions in the macroeconomic system”. *Finance research letters*, 32 (2020), p. 101107. [DOI](#).
- [76] Bury, T. M., Sujith, R. I., Pavithran, I., Scheffer, M., Lenton, T. M., Anand, M., and Bauch, C. T. “Deep learning for early warning signals of tipping points”. *Proceedings of the National Academy of Sciences*, 118 (2021). [DOI](#).
- [77] Lohmann, J., Castellana, D., Ditlevsen, P. D., and Dijkstra, H. A. “Abrupt climate change as a rate-dependent cascading tipping point”. *Earth system dynamics*, 12 (2021), pp. 819–835. [DOI](#).
- [78] Scheffer, M., Nes, E. H. van, Bird, D., Bocinsky, R. K., and Kohler, T. A. “Loss of resilience preceded transformations of pre-hispanic pueblo societies”. *Proceedings of the National Academy of Sciences*, 118 (2021). [DOI](#).
- [79] Donovan, G. and Brand, C. “Spatial early warning signals for tipping points using dynamic mode decomposition”. *Physica A: statistical mechanics and its applications*, 596 (2022), p. 127152. [DOI](#).

- [80] Dylewsky, D., Lenton, T. M., Scheffer, M., Bury, T. M., Fletcher, C. G., Anand, M., and Bauch, C. T. “Universal early warning signals of phase transitions in climate systems”. *Journal of the Royal Society Interface*, 20 (2023), p. 20220562. [DOI](#).
- [81] Proverbio, D., Skupin, A., and Gonçalves, J. “Systematic analysis and optimization of early warning signals for critical transitions using distribution data”. *IScience*, 26 (2023), p. 107156. [DOI](#).
- [82] Ritchie, P. D. L., Alkhayuon, H., Cox, P. M., and Wieczorek, S. “Rate-induced tipping in natural and human systems”. *Earth system dynamics*, 14 (2023), pp. 669–683. [DOI](#).
- [83] Yang, C., Kliemt, K., and Krellner, C. a. “Critical slowing down near a magnetic quantum phase transition with fermionic breakdown”. *Nature physics*, 19 (2023), pp. 1605–1610. [DOI](#).
- [84] Deb, S. and Dutta, P. S. “Critical transitions in spatial systems induced by Ornstein–Uhlenbeck noise: spatial mutual information as a precursor”. *Proceedings of the Royal Society A*, 480 (2024). [DOI](#).
- [85] Papapicco, D. *Notes on stochastic processes*. [Link](#).

## References

- [1] May, R. “Thresholds and breakpoints in ecosystems with a multiplicity of stable states”. *Nature*, 269 (1977), pp. 471–477. [DOI](#).
- [2] Zahler, R. and Sussmann, H. “Claims and accomplishments of applied catastrophe theory”. *Nature*, 269 (1977), pp. 759–763. [DOI](#).
- [3] Wissel, C. “A universal law of the characteristic return time near thresholds”. *Oecologia*, 65 (1984), pp. 101–107. [DOI](#).
- [4] Wiesenfeld, K. and McNamara, B. “Small-signal amplification in bifurcating dynamical systems”. *Physical Review A*, 33 (1 1986), pp. 629–642. [DOI](#).
- [5] Ives, A. R. “Measuring resilience in stochastic systems”. *Ecological monographs*, 65 (1995), pp. 217–233. [DOI](#).
- [6] Gandhi, A., Levin, S., and Orszag, S. “Critical slowing down in time-to-extinction: an example of critical phenomena in ecology”. *Journal of theoretical biology*, 192 (1998), pp. 363–376. [DOI](#).
- [7] Vandewalle, N., Boveroux, P., Minguet, A., and Ausloos, M. “The crash of October 1987 seen as a phase transition: amplitude and universality”. *Physica A: statistical mechanics and its applications*, 255 (1998), pp. 201–210. [DOI](#).
- [8] Petit, J. et al. “Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica”. *Nature*, 399 (1999), pp. 429–436. [DOI](#).
- [9] deMenocal, P., Ortiz, J., Guilderson, T., Adkins, J., Sarnthein, M., Baker, L., and Yarusinsky, M. “Abrupt onset and termination of the African Humid Period: rapid climate responses to gradual insolation forcing”. *Quaternary science reviews*, 19 (2000), pp. 347–361. [DOI](#).
- [10] Litt, B. et al. “Epileptic seizures may begin hours in advance of clinical onset: a report of five patients”. *Neuron*, 30 (2001), pp. 51–64. [DOI](#).
- [11] Scheffer, M., Carpenter, S., Foley, J., Folke, C., and Walker, B. “Catastrophic shifts in ecosystems”. *Nature*, 413 (2001), pp. 591–596. [DOI](#).
- [12] Kleinen, T., Held, H., and Petschel-Held, G. “The potential role of spectral properties in detecting thresholds in the Earth system: application to the thermohaline circulation”. *Ocean dynamics*, 53 (2003), pp. 53–63. [DOI](#).

- [13] McSharry, P., Smith, L., and Tarassenko, L. “Prediction of epileptic seizures: are nonlinear methods relevant?” *Nature medicine*, 9 (2003), pp. 241–242. [DOI](#).
- [14] Held, H. and Kleinen, T. “Detection of climate system bifurcations by degenerate fingerprinting”. *Geophysical research letters*, 31 (2004). [DOI](#).
- [15] Rietkerk, M., Dekker, S. C., Ruiter, P. C. de, and Koppell, J. van de. “Self-organized patchiness and catastrophic shifts in ecosystems”. *Science*, 305 (2004), pp. 1926–1929. [DOI](#).
- [16] Oborny, G. M. and György Szabó, B. “Dynamics of populations on the verge of extinction”. *Oikos*, 109 (2005), pp. 291–296. [DOI](#).
- [17] Tripati, A., Backman, J., Elderfield, H., and Ferretti, P. “Eocene bipolar glaciation associated with global carbon cycle changes”. *Nature*, 436 (2005), pp. 341–346. [DOI](#).
- [18] Venegas, J., Winkler, T., Musch, G., Vidal Melo, M., Layfield, D., Tgavalekos, N., Fischman, A., Callahan, R., Bellani, G., and Harris, S. “Self-organized patchiness in asthma as a prelude to catastrophic shifts”. *Nature*, 434 (2005), pp. 777–782. [DOI](#).
- [19] Brock, W. A. “Chapter 3: tipping points, abrupt opinion changes and punctuated policy change”. *Punctuated equilibrium and the dynamics of U.S. environmental policy*. Ed. by R. Repetto. New Haven: Yale University Press, 2006, pp. 47–77. [DOI](#).
- [20] Carpenter, S. R. and Brock, W. A. “Rising variance: a leading indicator of ecological transition”. *Ecology letters*, 9 (2006), pp. 311–318. [DOI](#).
- [21] Kéfi, S., Rietkerk, M., Alados, C., Pueyo, Y., Papanastasis, V., ElAich, A., and Ruiter, P. de. “Spatial vegetation patterns and imminent desertification in Mediterranean arid ecosystems”. *Nature*, 449 (2007), pp. 213–217. [DOI](#).
- [22] Livina, V. N. and Lenton, T. M. “A modified method for detecting incipient bifurcations in a dynamical system”. *Geophysical research letters*, 34 (2007). [DOI](#).
- [23] van Nes, E. H. and Scheffer, M. “Slow recovery from perturbations as a generic indicator of a nearby catastrophic shift.” *The American naturalist*, 169 (2007), pp. 738–747. [DOI](#).
- [24] Carpenter, S. R., Brock, W. A., Cole, J. J., Kitchell, J. F., and Pace, M. L. “Leading indicators of trophic cascades”. *Ecology letters*, 11 (2008), pp. 128–138. [DOI](#).
- [25] Dakos, V., Scheffer, M., van Nes, E. H., Brovkin, V., Petoukhov, V., and Held, H. “Slowing down as an early warning signal for abrupt climate change”. *Proceedings of the National Academy of Sciences*, 105 (2008), pp. 14308–14312. [DOI](#).
- [26] Guttal, V. and Jayaprakash, C. “Changing skewness: an early warning signal of regime shifts in ecosystems”. *Ecology letters*, 11 (2008), pp. 450–460. [DOI](#).
- [27] Biggs, R., Carpenter, S. R., and Brock, W. A. “Turning back from the brink: detecting an impending regime shift in time to avert it”. *Proceedings of the National Academy of Sciences*, 106 (2009), pp. 826–831. [DOI](#).
- [28] Chisholm, R. A. and Filotas, E. “Critical slowing down as an indicator of transitions in two-species models”. *Journal of theoretical biology*, 257 (2009), pp. 142–149. [DOI](#).
- [29] Guttal, V. and Jayaprakash, C. “Spatial variance and spatial skewness: leading indicators of regime shifts in spatial ecological systems”. *Theoretical ecology*, 2 (2009), pp. 3–12. [DOI](#).
- [30] Scheffer, M. *Critical transitions in nature and society*. Princeton University Press, 2009.
- [31] Scheffer, M., Bascompte, J., Brock, W., Brovkin, V., Carpenter, S., Dakos, V., Held, H., van Nes, E., Rietkerk, M., and Sugihara, G. “Early-warning signals for critical transitions”. *Nature*, 461 (2009). [DOI](#).
- [32] Brock, W. A. and Carpenter, S. R. “Interacting regime shifts in ecosystems: implication for early warnings”. *Ecological monographs*, 80 (2010), pp. 353–367. [DOI](#).

- [33] Carpenter, S. R. and Brock, W. A. “Early warnings of regime shifts in spatial dynamics using the discrete Fourier transform”. *Ecosphere*, 1 (2010). [DOI](#).
- [34] Dakos, V., van Nes, E., Donangelo, R., Fort, H., and Scheffer, M. “Spatial correlation as leading indicator of catastrophic shifts”. *Theoretical ecology*, 3 (2010), pp. 163–174. [DOI](#).
- [35] Ditlevsen, P. D. and Johnsen, S. J. “Tipping points: early warning and wishful thinking”. *Geophysical research letters*, 37 (2010). [DOI](#).
- [36] Drake, J. and Griffen, B. “Early warning signals of extinction in deteriorating environments”. *Nature*, 467 (2010), pp. 456–459. [DOI](#).
- [37] Hastings, A. and Wysham, D. B. “Regime shifts in ecological systems can occur with no warning”. *Ecology letters*, 13 (2010), pp. 464–472. [DOI](#).
- [38] Livina, V. N., Kwasniok, F., and Lenton, T. M. “Potential analysis reveals changing number of climate states during the last 60 kyr”. *Climate of the past*, 6 (2010), pp. 77–82. [DOI](#).
- [39] Carpenter, S. R. and Brock, W. A. “Early warnings of unknown nonlinear shifts: a nonparametric approach”. *Ecology*, 92 (2011), pp. 2196–2201. [DOI](#).
- [40] Carpenter, S. R. et al. “Early warnings of regime shifts: a whole-ecosystem experiment”. *Science*, 332 (2011), pp. 1079–1082. [DOI](#).
- [41] Dakos, V., Kéfi, S., Rietkerk, M., van Nes, E. H., and Scheffer, M. “Slowing down in spatially patterned ecosystems at the brink of collapse”. *The American naturalist*, 177 (2011), pp. 153–166. [DOI](#).
- [42] Kuehn, C. “A mathematical framework for critical transitions: bifurcations, fast–slow systems and stochastic dynamics”. *Physica D: nonlinear phenomena*, 240 (2011), pp. 1020–1035. [DOI](#).
- [43] Lenton, T. “Early warning of climate tipping points”. *Nature climate change*, 1 (2011), pp. 201–209. [DOI](#).
- [44] Wieczorek, S., Ashwin, P., Luke, C. M., and Cox, P. M. “Excitability in ramped systems: the compost-bomb instability”. *Proceedings of the Royal Society A*, 467 (2011), pp. 2733–2733. [DOI](#).
- [45] Ashwin, P., Wieczorek, S., Vitolo, R., and Cox, P. “Tipping points in open systems: bifurcation, noise-induced and rate-dependent examples in the climate system”. *Philosophical Transactions of the Royal Society A*, 370 (2012), pp. 1166–1184. [DOI](#).
- [46] Dai, L., Vorselen, D., Korolev, K. S., and Gore, J. “Generic indicators for loss of resilience before a tipping point leading to population collapse”. *Science*, 336 (2012), pp. 1175–1177. [DOI](#).
- [47] Dakos, V. et al. “Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data”. *PLOS ONE*, 7 (2012), pp. 1–20. [DOI](#).
- [48] Scheffer, M. et al. “Anticipating critical transitions”. *Science*, 338 (2012), pp. 344–348. [DOI](#).
- [49] Wang, R., Dearing, J., and Langdon, P. “Flickering gives early warning signals of a critical transition to a eutrophic lake state”. *Nature*, 492 (2012), pp. 419–422. [DOI](#).
- [50] Boerlijst, M. C., Oudman, T., and Roos, A. M. de. “Catastrophic collapse can occur without early warning: examples of silent catastrophes in structured ecological models”. *PLOS ONE*, 8 (2013), pp. 1–6. [DOI](#).
- [51] Dakos, V., van Nes, E., and Scheffer, M. “Flickering as an early warning signal”. *Theoretical ecology*, 6 (2013), pp. 309–317. [DOI](#).
- [52] Kéfi, S., Dakos, V., Scheffer, M., van Nes, E. H., and Rietkerk, M. “Early warning signals also precede non-catastrophic transitions”. *Oikos*, 122 (2013), pp. 641–648. [DOI](#).
- [53] Kuehn, C. “Warning signs for wave speed transitions of noisy Fisher–KPP invasion fronts”. *Theoretical ecology*, 6 (2013), pp. 295–308. [DOI](#).

- [54] Bos, E. H. and Jonge, P. D. ““Critical slowing down in depression” is a great idea that still needs empirical proof”. *Proceedings of the National Academy of Sciences*, 111 (2014), pp. 878–878. [DOI](#).
- [55] Cline, T. J., Seekell, D. A., Carpenter, S. R., Pace, M. L., Hodgson, J. R., Kitchell, J. F., and Weidel, B. C. “Early warnings of regime shifts: evaluation of spatial indicators from a whole-ecosystem experiment”. *Ecosphere*, 5 (2014). [DOI](#).
- [56] Kéfi, S., Guttal, V., Brock, W. A., Carpenter, S. R., Ellison, A. M., Livina, V. N., Seekell, D. A., Scheffer, M., van Nes, E. H., and Dakos, V. “Early warning signals of ecological transitions: methods for spatial patterns”. *PLOS ONE*, 9 (2014), pp. 1–13. [DOI](#).
- [57] Leemput, I. A. van de et al. “Critical slowing down as early warning for the onset and termination of depression”. *Proceedings of the National Academy of Sciences*, 111 (2014), pp. 87–92. [DOI](#).
- [58] Suweis, S. and D’Odorico, P. “Early warning signs in social-ecological networks”. *PLOS ONE*, 9 (2014), pp. 1–6. [DOI](#).
- [59] van de Leemput, I. A., van Nes, E. H., and Scheffer, M. “Resilience of alternative states in spatially extended ecosystems”. *PLOS ONE*, 10 (Feb. 2015), pp. 1–17. [DOI](#).
- [60] Donovan, G. and Kritter, T. “Spatial pattern formation in the lung”. *Journal of mathematical biology*, 70 (2015), pp. 1119–1149. [DOI](#).
- [61] Scheffer, M., Carpenter, S. R., Dakos, V., and van Nes, E. H. “Generic indicators of ecological resilience: inferring the chance of a critical transition”. *Annual review of ecology, evolution, and systematics*, 46 (2015), pp. 145–167. [DOI](#).
- [62] Strogatz, S. *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering* (2nd ed.) CRC Press, 2015. [DOI](#).
- [63] Kozłowska, M., Denys, M., Wiliński, M., Link, G., Gubiec, T., Werner, T., Kutner, R., and Struzik, Z. “Dynamic bifurcations on financial markets”. *Chaos, solitons & fractals*, 88 (2016), pp. 126–142. [DOI](#).
- [64] Ritchie, P. and Sieber, J. “Early-warning indicators for rate-induced tipping”. *Chaos: an interdisciplinary journal of nonlinear science*, 26 (2016). [DOI](#).
- [65] Dmitriev, A., Dmitriev, V., Sagaydak, O., and Tsukanova, O. “The application of stochastic bifurcation theory to the early detection of economic bubbles”. *Procedia computer science*, 122 (2017). 5th International Conference on Information Technology and Quantitative Management, ITQM 2017, pp. 354–361. [DOI](#).
- [66] Liang, J., Hu, Y., Chen, G., and Zhou, T. “A universal indicator of critical state transitions in noisy complex networked systems”. *Scientific reports*, 7 (2017). [DOI](#).
- [67] Jentsch, P. C., Anand, M., and Bauch, C. T. “Spatial correlation as an early warning signal of regime shifts in a multiplex disease-behaviour network”. *Journal of theoretical biology*, 448 (2018), pp. 17–25. [DOI](#).
- [68] Chen, S., O’Dea, E. B., Drake, J. M., and Epureanu, B. I. “Eigenvalues of the covariance matrix as early warning signals for critical transitions in ecological systems”. *Scientific reports*, 9 (2019). [DOI](#).
- [69] Diks, C., Hommes, C., and Wang, J. “Critical slowing down as an early warning signal for financial crises?” *Empirical economics*, 57 (2019), pp. 1201–1228. [DOI](#).
- [70] Pages, N., Vera-Sigüenza, E., Rugis, J., Kirk, V., Yule, D., and Sneyd, J. “A model of  $Ca^{2+}$  dynamics in an accurate reconstruction of parotid acinar cells”. *Bulletin of mathematical biology*, 81 (2019), pp. 1394–1426. [DOI](#).
- [71] de Mooij, S. M., Blanken, T. F., Grasman, R. P., Ramautar, J. R., Van Someren, E. J., and van der Maas, H. L. “Dynamics of sleep: exploring critical transitions and early warning signals”. *Computer methods and programs in biomedicine*, 193 (2020), p. 105448. [DOI](#).

- [72] Gottwald, G. and Gugole, F. “Detecting regime transitions in time series using dynamic mode decomposition”. *Journal of statistical physics*, 179 (2020), pp. 1028–1045. [DOI](#).
- [73] Lever, J. J., Leemput, I. A. van de, Weinans, E., Quax, R., Dakos, V., Nes, E. H. van, Bascompte, J., and Scheffer, M. “Foreseeing the future of mutualistic communities beyond collapse”. *Ecology letters*, 23 (2020), pp. 2–15. [DOI](#).
- [74] Wechselberger, M. *Geometric singular perturbation theory beyond the standard form*. Springer, 2020. [DOI](#).
- [75] Xing, K. and Yang, X. “Predicting default rates by capturing critical transitions in the macroeconomic system”. *Finance research letters*, 32 (2020), p. 101107. [DOI](#).
- [76] Bury, T. M., Sujith, R. I., Pavithran, I., Scheffer, M., Lenton, T. M., Anand, M., and Bauch, C. T. “Deep learning for early warning signals of tipping points”. *Proceedings of the National Academy of Sciences*, 118 (2021). [DOI](#).
- [77] Lohmann, J., Castellana, D., Ditlevsen, P. D., and Dijkstra, H. A. “Abrupt climate change as a rate-dependent cascading tipping point”. *Earth system dynamics*, 12 (2021), pp. 819–835. [DOI](#).
- [78] Scheffer, M., Nes, E. H. van, Bird, D., Bocinsky, R. K., and Kohler, T. A. “Loss of resilience preceded transformations of pre-hispanic pueblo societies”. *Proceedings of the National Academy of Sciences*, 118 (2021). [DOI](#).
- [79] Donovan, G. and Brand, C. “Spatial early warning signals for tipping points using dynamic mode decomposition”. *Physica A: statistical mechanics and its applications*, 596 (2022), p. 127152. [DOI](#).
- [80] Dylewsky, D., Lenton, T. M., Scheffer, M., Bury, T. M., Fletcher, C. G., Anand, M., and Bauch, C. T. “Universal early warning signals of phase transitions in climate systems”. *Journal of the Royal Society Interface*, 20 (2023), p. 20220562. [DOI](#).
- [81] Proverbio, D., Skupin, A., and Gonçalves, J. “Systematic analysis and optimization of early warning signals for critical transitions using distribution data”. *IScience*, 26 (2023), p. 107156. [DOI](#).
- [82] Ritchie, P. D. L., Alkhayuon, H., Cox, P. M., and Wieczorek, S. “Rate-induced tipping in natural and human systems”. *Earth system dynamics*, 14 (2023), pp. 669–683. [DOI](#).
- [83] Yang, C., Kliemt, K., and Krellner, C. a. “Critical slowing down near a magnetic quantum phase transition with fermionic breakdown”. *Nature physics*, 19 (2023), pp. 1605–1610. [DOI](#).
- [84] Deb, S. and Dutta, P. S. “Critical transitions in spatial systems induced by Ornstein–Uhlenbeck noise: spatial mutual information as a precursor”. *Proceedings of the Royal Society A*, 480 (2024). [DOI](#).
- [85] Papapicco, D. *Notes on stochastic processes*. [Link](#).

## 5 Supplementary material

### 5.1 Data sources for real-world timeseries in Figure 1

We hereby list all the sources from which the historical data used for the timeseries of Figure 1. The **end of Greenhouse Earth** (a) timeseries is a collection of irregularly spaced samples from tropical Pacific sediment core records for the Eocene transition from Greenhouse to Icehouse Earth's climate [17]; the data itself was downloaded from the World Data Center for Paleoclimatology<sup>1</sup> hosted by the National Oceanic and Atmospheric Administration (NOAA). The **end of the last glaciation** (b) represents the concentration of Deuterium found in Vostok's Antarctic ice core [8] from which temperature reconstruction for the past 420,000 years was performed; the dataset was downloaded from the aforementioned Data Center<sup>1</sup>. The **desertification of North Africa** (c), coinciding with the abrupt end of the African humid period, is depicted as a measurement of the mean Sea Surface Temperature (SST) reconstructed from the Ocean Drilling Program (ODP) site 658C [9]; the data is sourced from the georeferenced database PANGAEA<sup>2</sup>. The **yeast population collapse** (d) data comes from a controlled experiment monitoring the density of budding yeast cells at steady dilution factors of sucrose [46]; these series of experiments were then used to form hypothesis on the role of CSD (and its indicators) preceding systemic collapse in ecosystems [81]. The **heavy Fermi compound phase transition** (e) experimental datapoints are related to the Fermionic breakdown shown as a sharp decrease in the (THz) spectroscopy resonance of YbRh<sub>2</sub>Si<sub>2</sub> at constant magnetic field (130 mT) and varying temperature; these results were added as supplementary data in [83]. The **2009 global financial crisis** (f) is reflected, among other indicators, as a sudden collapse of the historical S&P 500 stock market index and recent studies [65, 69] suggest a CSD for this and other market crashes [7]; the timeseries is collected from the Federal Reserve Bank of St. Louis<sup>3</sup>.

### 5.2 Potential of the bistable saddle-node bifurcation in Figure 3

We consider an additive noise Ito process in potential form

$$dx = -U'(x; \mu)dt + \sigma dW,$$

where we set  $U(x; \mu) = \mu x + x^2 - x^3 + \frac{1}{5}x^4$  to be the primitive dynamics. The system has two saddle-node bifurcations at  $\mu \approx -0.38$  and  $\mu \approx 1.62$ . The two ensembles simulated in (a) and (b) have a total of 100 independent trajectories each with slow, linear ramping of the parameter with time  $\varepsilon = 10^{-2}$  and noise levels  $\sigma = 0.1$  and  $\sigma = 1.2$  respectively.

---

<sup>1</sup><https://www.ncei.noaa.gov/products/paleoclimatology>

<sup>2</sup><https://www.pangaea.de/>

<sup>3</sup><https://fred.stlouisfed.org/series/SP500>